Universidade Federal Fluminense

Instituto de Matemática e Estatística

GET00143 – TEORIA DAS PROBABILIDADES II
Variáveis Aleatórias Unidmensionais

Ana Maria Lima de Farias
Jessica Quintanilha Kubrusly
Mariana Albi de Oliveira Souza

Departamento de Estatística

Conteúdo

Capítulo 1

Variáveis Aleatórias

Neste capítulo será introduzido o conceito de Variável Aleatória e de Função de Distribuição. Antes disso, na primeira seção do capítulo, será feita uma breve revisão a fim de destacar alguns resultados importantes para o conteúdo que segue.

1.1 Uma Breve Revisão

Um experimento que pode gerar diferentes resultados se realizado mais de uma vez sob as mesmas condições é chamado de ***experimento aleatório***. Como exemplo de experimentos aleatórios temos: o lançamento de um dado e a observação da face superior; a observação do tempo de duração de um certo dispositivo eletrônico; lançamentos consecutivos de uma moeda até que saia a face cara.

O conjunto formado por todos os possíveis resultados de um experimento aleatório é chamado de ***espaço amostral*** e é representado pela letra grega maiúscula Ω. Em geral usamos a letra grega minúscula ω para representar um resultado específico de um experimento aleatório. Nesse caso podemos escrever $\omega \in \Omega$.

Seja Ω um espaço amostral de algum experimento aleatório. Todo subconjunto $A \subseteq \Omega$ será chamado de ***evento***.

Exemplo 1.1
Para os experimentos aleatórios listados a seguir defina um espaço amostral e apresente alguns eventos.

(a) o lançamento de um dado e a observação da face superior;

(b) a observação do tempo de duração de um certo dispositivo eletrônico;

(c) lançamentos consecutivos de uma moeda até que saia a face cara.

Solução:

(a) $\Omega = \{1, 2, 3, 4, 5, 6\}$.
Eventos:
$E_1 = \{$ sair uma face par $\} = \{2, 4, 6\}$;
$E_2 = \{$ sair um número maior que 1 $\} = \{2, 3, 4, 5, 6\}$;
$E_3 = \{$ sair o número 3 $\} = \{3\}$;
$E_4 = \emptyset$, o evento vazio.

(b) $\Omega = \{t \in \mathbb{R} \mid t \geq 0\}$, onde t é o tempo de duração em minutos.
Eventos:
$E_1 = \{$ o dispositivo durar entre 10 e 30 minutos $\} = (10, 30)$;
$E_2 = \{$ o dispositivo durar menos de 15 minutos $\} = [0, 15)$;
$E_3 = \{$ o dispositivo durar exatos 5 minutos e 30 segundos $\} = \{5.5\}$;
$E_4 = \{$ o dispositivo pelo menos 1 hora $\} = [60, \infty)$.

(c) $\Omega = \{K, CK, CCK, CCCK, CCCCK, \ldots\}$, onde K representa a face cara e C a face coroa.
Eventos:
$E_1 = \{$ sair 2 coroas $\} = \{CCK\}$;
$E_2 = \{$ ocorrer mais de 2 lançamentos da moeda $\} = \{CCK, CCCK, \ldots\}$;
$E_3 = \{$ ocorrer menos de 5 lançamentos $\} = \{K, CK, CCK, CCCK, CCCCK\}$;
$E_4 = \{$ ocorrer qualquer saída $\} = \Omega$, o próprio espaço amostral.

♦♦

Definição 1.2 σ**-álgebra**
*Seja Ω um espaço amostral de algum experimento aleatório. Uma classe de subconjuntos de Ω, representada por $\mathcal{F}$, é denominada uma **σ-álgebra** se satisfaz as seguintes propriedades:*

A_1. $\Omega \in \mathcal{F}$;

A_2. *Se* $A \in \mathcal{F} \Rightarrow A^c \in \mathcal{F}$;

A_3. *Se* $A_i \in \mathcal{F} \ \forall \ i \Rightarrow \cup_{i=1}^{\infty} A_i \in \mathcal{F}$.

Veja que uma σ-álgebra é um conjunto de eventos, ou seja, um conjunto de subconjuntos de Ω. Além disso, dada uma σ-álgebra $\mathcal{F}$ qualquer sempre temos $\emptyset \in \mathcal{F}$, uma vez que $\Omega \in \mathcal{F}$ e o complementar de qualquer elemento também é um elemento da σ-álgebra (A_2).

O conjunto de todos os subconjuntos de Ω é chamado de ***conjunto das partes*** de Ω.

Exemplo 1.3
Considere o experimento de lançar de um dado e observar a sua face superior. Para esse experimento defina $\Omega = \{1, 2, 3, 4, 5, 6\}$. Apresente duas possíveis σ-álgebras para esse espaço amostral Ω.

Solução:
A primeira σ-álgebra $\mathcal{F}_1$ será definida pelo conjunto das partes de Ω, que tem $2^6 = 64$ elementos. Entre eles está o único subconjunto com zero elementos, $\emptyset$, todos os 6 subconjuntos com 1 elemento, $\{1\}, \{2\}, \{3\}, \{4\}, \{5\}$ e $\{6\}$, todos os 15 subconjuntos com 2 elementos, $\{1, 2\}, \{1, 3\}, \{1, 4\}, \{1, 5\}, \{1, 6\}, \{2, 3\}, \{2, 4\}, \{2, 5\}, \{2, 6\}, \ldots, \{5, 4\}$ e $\{5, 6\}$, assim por diante até o único subconjunto com 6 elementos, que é o próprio $\Omega = \{1, 2, 3, 4, 5, 6\}$.

$$\begin{aligned}\mathcal{F}_1 = \{ \quad & \emptyset, \{1\}, \{2\}, \ldots, \{6\}, \{1, 2\}, \{1, 3\}, \ldots \{5, 6\}, \{1, 2, 3\}, \ldots \{4, 5, 6\}, \\ & \{1, 2, 3, 4\} \ldots \{3, 4, 5, 6\}, \{1, 2, 3, 4, 5\} \ldots \{2, 3, 4, 5, 6\}, \{1, 2, 3, 4, 5, 6\} \quad \}.\end{aligned}$$

Um outro exemplo de σ-álgebra é $\mathcal{F}_2 = \{\emptyset, \Omega\}$.

Verifique que tanto $\mathcal{F}_1$ quanto $\mathcal{F}_2$ satisfazem as propriedades A_1, A_2 e A_3 listadas na Definição 1.2.

♦♦

Definição 1.4 Probabilidade
Seja Ω um espaço amostral de um experimento aleatório e $\mathcal{F}$ uma σ-álgebra de subconjuntos de Ω. Uma função $\mathrm{P} : \mathcal{F} \mapsto [0,1]$ é uma ***probabilidade*** *se satisfaz os seguintes axiomas, chamados de Axiomas de Kolmogorov:*

Ax_1. $\mathrm{P}(\Omega) = 1$;

Ax_2. *Para todo evento* $A \in \mathcal{F}$, $\mathrm{P}(A) \geq 0$;

Ax_3. *Para toda sequência de eventos* $A_1, A_2, \ldots \in \mathcal{F}$ *mutuamente exclusivos, isto é, tais que* $A_i \cap A_j = \emptyset \ \forall \ i \neq j$, *temos*

$$\mathrm{P}\left(U_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mathrm{P}(A_i).$$

A trinca $(\Omega, \mathcal{F}, \mathrm{P})$ é denominada *Espaço de Probabilidade*.

Em alguns livros, por exemplo James (2004), os autores fazem distinção entre evento e evento aleatório. Enquanto evento é qualquer subconjunto do espaço amostral o termo *Evento Aleatório* é definido como qualquer evento ao qual é atribuído uma probabilidade, ou seja, os eventos da σ-álgebra. Neste texto, por se tratar de um primeiro curso de probabilidade, não vamos fazer essa distinção. Sempre que qualquer um dos dois termos for mencionado ele se refere a um evento ao qual é atribuído uma probabilidade.

Proposição 1.5 Propriedades da Probabilidade
Seja $(\Omega, \mathcal{F}, \mathrm{P})$ *um espaço de probabilidade. Sejam também A, B e* A_i *eventos nesse espaço de probabilidade. Então valem as seguintes propriedades:*

P_1. $\mathrm{P}(A) = 1 - \mathrm{P}(A^c)$;

P_2. $\mathrm{P}(B) = \mathrm{P}(B \cap A) + \mathrm{P}(B \cap A^c)$;

P_3. *Se* $A \subset B \Rightarrow \mathrm{P}(A) \leq \mathrm{P}(B)$;

P_4. $\mathrm{P}(A \cup B) = \mathrm{P}(A) + \mathrm{P}(B) - \mathrm{P}(A \cap B)$ *(ou* $\mathrm{P}(A \cap B) = \mathrm{P}(A) + \mathrm{P}(B) - \mathrm{P}(A \cup B)$*)*;

P_5. $\mathrm{P}\left(U_{i=1}^{\infty} A_i\right) \leq \sum_{i=1}^{\infty} \mathrm{P}(A_i)$.

Demonstração:
Para a demonstração da Proposição 1.5 veja Proposição 1.1 de Magalhães (2011).

□

Exemplo 1.6 Problema do Aniversário
Em um grupo com r pessoas qual a probabilidade de pelo menos duas fazerem aniversário no mesmo dia? Qual o menor número de pessoas tal que a probabilidade de pelo menos duas fazerem aniversário no mesmo dia seja maior que 99%?

Solução:
Vamos primeiro resolver para $r = 5$.

Qual é o experimento aleatório que está sendo realizado? O experimento pode ser definido como sortear 5 datas de aniversários (aniversário dos indivíduos), entre as 365 possíveis. Então o espaço amostral para esse experimento pode ser representado por:

$$\Omega = \{(1,1,1,1,1), (1,1,1,1,2), , (1,1,1,2,1), \ldots, (2,4,123,201,250), \ldots, (365,365,365)\}.$$

O elemento $(1,1,1,1,1)$ representa a saída em que todos os aniversários sorteados foram 01/jan, o elemento $(365,365,365)$ representa a saída em que todos os aniversários sorteados foram 31/dez e o elemento $((2,4,123,201,250)$ representa a saída em que um aniversário foi o segundo dia do ano (02/jan), o segundo aniversário foi o quarto dia do ano (04/jan), e assim por diante.

Veja que $\#\Omega = 365^5$.

Como queremos calcular uma probabilidade devemos definir o evento para o qual queremos encontrar a probabilidade. Que evento é esse?

$E = \{$pelo menos duas pessoas no grupo de 5 fazendo aniverário no mesmo dia$\}$.

Veja que $(1,2,3,4,5) \notin E$, mas $(1,1,1,1,1) \in E$ e $(1,10,20,53,10) \in E$. Supondo que todos os dias do ano são igualmente prováveis de serem escolhidos a probabilidade associada a cada elemento de Ω será $1/365^5$, pois nesse caso o espaço amostral será equiprovável. Então para saber $\mathrm{P}(E)$ basta saber quantos elementos estão contidos em E.

Mais fácil do que contar o número de elementos de E será contar o número de elementos de E^c. Como podemos usar $\mathrm{P}(E^c)$ para encontrar $\mathrm{P}(E)$, veja Proposição 1.5, vamos resolver por esse caminho. Para isso precisamos definir E^c:

$$E^c = \{\text{nenhum par de pessoas fazendo aniverário no mesmo dia}\}.$$

Veja que $\#E = 365 \times 364 \times 363 \times 362 \times 361$. Então

$$\mathrm{P}(E^c) = \frac{365 \times 364 \times 363 \times 362 \times 361}{365^5} = 0,9729 \;\Rightarrow\; \mathrm{P}(E) = 0,0271.$$

Repetindo as mesmas contas para diferentes valores de r podemos construir a tabela abaixo.

r	5	10	20	30	40	50	60
P(E)	0,0271	0,1169	0,4114	0,7063	0,8912	0,9704	0,9941

♦♦

1.2 Variáveis Aleatórias

Considere o experimento aleatório definido pelo sorteio de uma amostra de 20 funcionários de uma empresa que tem 500 funcionários. O espaço amostral deste experimento é formado por todas as amostras possíveis com 20 funcionários da empresa. Como a ordem que os funcionários são sorteados não importa e não deve haver repetição de funcionários em um sorteio, o número total de tais amostras é $\binom{500}{20}$. Cada elemento desse espaço amostral é formado pela relação dos 20 funcionários sorteados.

Em situações como essa, em geral, o interesse não está no funcionário em si, mas, sim, em alguma característica deste funcionário, por exemplo, sua altura, se tem curso superior ou não, número de dependentes. Dessa forma, poderíamos calcular a altura média dos funcionários da amostra, o número médio de dependentes, a proporção de funcionários com curso superior, etc. Então, a cada amostra possível, ou seja, a cada ponto do espaço amostral associamos um número. Essa é a definição de *variável aleatória*.

Definição 1.7 Variável aleatória
*Seja $(\Omega, \mathcal{F}, \mathcal{P})$ um espaço de probabilidade. Uma **variável aleatória** X é qualquer função $X : \Omega \longrightarrow \mathbb{R}$ tal que*

$$X^{-1}(I) = \{\omega \in \Omega : X(\omega) \in I\} \in \mathcal{F}$$

para todo intervalo $I \subset \mathbb{R}$, ou seja, X é tal que sua imagem inversa de qualquer $I \subset \mathbb{R}$ pertence à $\sigma-$álgebra $\mathcal{F}$.

Dito de outra forma, uma *variável aleatória* é uma função real (isto é, que assume valores em $\mathbb{R}$) definida no espaço amostral Ω de um experimento aleatório. Isso significa que uma variável aleatória é uma função que associa a cada elemento de Ω um número real.

Por questões de simplicidade, muitas vezes abreviaremos a expressão variável aleatória por "v.a.". A convenção usual para representar uma v.a. consiste em usar letras maiúsculas como X, Y, etc. Um valor específico, mas genérico, desta variável será representado pela letra minúscula correspondente: x, y, etc.

Como uma variável aleatória X é uma função muitas vezes estaremos interessados na sua imagem, $Im(X)$, que é definida pelo conjunto de valores que a variável aleatória pode assumir. Veja os exemplos a seguir.

Exemplo 1.8
Considere o experimento de lançar três vezes uma moeda. Seja X a variável aleatória definida pelo número de caras nos três lançamentos.

(a) Defina o espaço amostral Ω deste experimento.

(b) Determine $X(\omega)$ para cada $\omega \in \Omega$.

(c) Determine $Im(X)$.

Solução:
Para definir o espaço amostral vamos representar por K a cara e por C a coroa.

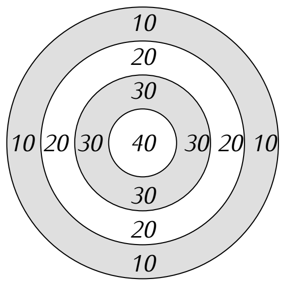
(a) $\Omega = \{KKK, KKC, KCK, KCC, CKK, CKC, CCK, CCC\}$.

(b) $X(KKK) = 3,\ X(KKC) = 2,\ X(KCK) = 2,\ X(KCC) = 1,\ X(CKK) = 2,$
$X(CKC) = 1,\ X(CCK) = 1,\ X(CCC) = 0.$

(c) $Im(X) = \{0, 1, 2, 3\}$.

♦♦

Exemplo 1.9 Jogo de Dardo
Considere o experimento de lançar um dardo em um alvo formado por quatro círculos concêntricos de raios 10, 20, 30 e 40, como o da figura ao lado. Para esse experimento defina duas variáveis aleatórias X e Y: seja X a distância entre o dardo e o centro do alvo e Y a pontuação ganha no lançamento do dardo, indicada também na figura. Defina um espaço amostral do experimento e determine $Im(X)$ e $Im(Y)$.



Solução:
Vamos supor que se o dardo for lançado fora do alvo o experimento é repetido. Nesse caso podemos representar cada saída do experimento pelo ponto no plano representado pelo local em que o dardo foi lançado dentro do alvo. Assim o espaço amostral será:

$$\Omega = \left\{ (i,j) \mid \sqrt{i^2 + j^2} \leq 40 \right\}.$$

Lembre-se que $\sqrt{i^2 + j^2}$ é a distância do ponto (i, j) à origem, isto é, ao centro do alvo. Como a v.a. X é definida justamente pela distância entre o dardo e o centro alvo podemos concluir que, $\forall\ (i,j) \in \Omega$, $X(i,j) = \sqrt{i^2 + j^2}$. Além disso, a menor distância possível é 0 e a maior é 40, logo $0 \leq X(i,j) \leq 40$ e $Im(X) = [0, 40]$.

Já a v.a. Y, que é definida pelos pontos ganhos, só pode assumir os valores 40, 30, 20 ou 10, de acordo com a pontuação indicada na figura. Então $Im(Y) = \{10, 20, 30, 40\}$.

♦♦

Exemplo 1.10
Considere o experimento de observar, durante um dia, o número de nascimentos e o sexo dos bebês em uma certa maternidade. Para esse experimento considere o evento A = nasceram mais homens que mulheres no dia de observação.

(a) Defina um espaço amostral para o experimento.

(b) Defina o evento A a partir dos elementos de Ω.

(c) Verifique que $\mathcal{F} = \{\Omega, \emptyset, A, A^c\}$ é uma σ-álgebra.

(d) Verifique que a função I_A definida a seguir é uma variável aleatória.

$$I_A(\omega) = \begin{cases} 1 & , \ se\ \omega \in A \\ 0 & , \ se\ \omega \notin A \end{cases}$$

(e) Determine $Im(I_A)$.

Solução:

(a) Vamos representar cada elemento de Ω por um par ordenado (x, y) de forma que x indique o número de nascimentos de bebês do sexo feminino e y o número de nascimentos de bebês do sexo masculino. Como tanto x quanto y indicam uma contagem, ambos têm que ser números naturais. Logo, $\Omega = \{(x, y) \in \mathbb{N}^2\}$.

(b) O evento A é definido pelo subconjunto de Ω tal que cada elemento representa mais nascimentos de meninos do que de meninas. Por exemplo, $\{(0, 1), (2, 4), (1, 3)\} \subset A$ e $\{(5, 2), (0, 0), (1, 0)\} \nsubseteq A$. Podemos então definir $A = \{(x, y) \in \mathbb{N}^2 \mid x < y\}$.

(c) Para verificar que $\mathcal{F}$ é σ-álgebra vamos verificar as propriedades A_1, A_2 e A_3 da Definição 1.2. A_1 é imediata, pois $\Omega \in \mathcal{F}$. Como $\mathcal{F}$ só tem 4 elementos, podemos, para cada um deles, ver que o seu complementar também pertence à $\mathcal{F}$ e assim verificar A_2. O mesmo vale para A_3, podemos verificar que qualquer união também será elemento de $\mathcal{F}$.

(d) Para verificar que I_A é v.a. é preciso verificar que a imagem inversa de qualquer intervalo $I \in \mathbb{R}$ pertence à σ-álgebra $\mathcal{F}$. Seja $I \in \mathbb{R}$ um intervalo qualquer. Se $0 \in I$ e $1 \in I$, então $I_A^{-1}(I) = \Omega \in \mathcal{F}$. Se $0 \notin I$ e $1 \notin I$, então $I_A^{-1}(I) = \emptyset \in \mathcal{F}$. Se $0 \notin I$ e $1 \in I$, então $I_A^{-1}(I) = A \in \mathcal{F}$. Se $0 \in I$ e $1 \notin I$, então $I_A^{-1}(I) = A^c \in \mathcal{F}$. Logo, I_A é variável aleatória.

(e) $Im(I_A) = \{0, 1\}$.

♦♦

A função I_A definida no Exemplo 1.10 é chamada de *função indicadora* do conjunto A. Sempre que A for um evento da σ-álgebra, I_A será variável aleatória. Nesse caso I_A indica a ocorrência ou não do evento A. Veja a definição formal a seguir.

Definição 1.11 Variável Aleatória Indicadora
Seja A um evento qualquer no espaço de probabilidade $(\Omega, \mathcal{F}, P)$. A função I_A definida por

$$I_A(\omega) = \begin{cases} 1 & , \text{ se } \omega \in A \\ 0 & , \text{ se } \omega \notin A \end{cases}$$

é uma variável aleatória e é chamada de ***variável aleatória indicadora*** *(ou v.a. dummy, ou v.a. binária).*

1.3 Função de Distribuição Acumulada

Seja X uma variável aleatória definida no espaço de probabilidade $(\Omega, \mathcal{F}, \mathrm{P})$. Veja que, pela própria definição de variável aleatória, qualquer que seja o intervalo $I \subset \mathbb{R}$

$$E = \{\omega \in \Omega \mid X(\omega) \in I\}$$

é um evento e logo $\mathrm{P}(E)$ é bem definido. Ou seja, podemos calcular $\mathrm{P}(E)$. Por abuso de notação costuma-se escrever essa probabilidade da seguinte maneira,

$$\mathrm{P}(E) = \mathrm{P}(\{\omega \in \Omega \mid X(\omega) \in I\}) = \mathrm{P}(X \in I),$$

que é a probabilidade da v.a. X assumir valores no intervalo I.

Em particular podemos escolher o intervalo $I = (-\infty, x]$, qualquer que seja $x \in \mathbb{R}$. Nesse caso $\mathrm{P}(E) = \mathrm{P}(\omega \in \Omega \mid X(\omega) \in (-\infty, x]) = \mathrm{P}(X \leq x)$. Se conhecermos $\mathrm{P}(X \leq x)$ para todo $x \in \mathbb{R}$ podemos calcular a probabilidade de qualquer evento. Por esse motivo é tão importante a definição a seguir.

Definição 1.12 Função de Distribuição Acumulada
Dada uma variável aleatória X, a ***função de distribuição acumulada*** *de X, ou simplesmente* ***função de distribuição***, *é definida por*

$$F_X(x) = \mathrm{P}(X \leq x) \qquad \forall\, x \in \mathbb{R}.$$

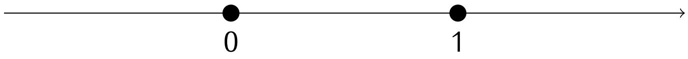
É interessante notar que a função F_X está definida para todo número real x, independente dos valores que a variável aleatória X possa assumir. Além disso também vale notar que $F_X : \mathbb{R} \to [0, 1]$.

Exemplo 1.13
Considere o experimento de lançar uma moeda. Para esse experimento seja X a v.a. indicadora da ocorrência da face cara, isto é,

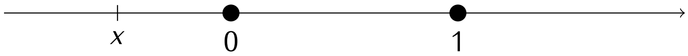
$$X = \begin{cases} 1 & , \text{ se sair cara;} \\ 0 & , \text{ se sair coroa.} \end{cases}$$

Encontre a função de distribuição de X e esboce seu gráfico.

Solução:
Queremos encontrar $F_X(x) = \mathrm{P}(X \leq x)\ \forall\ x \in \mathbb{R}$. Veja que $Im(X) = \{0, 1\}$,



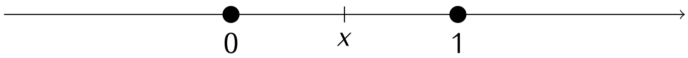
vamos pensar em $F_X(x)$ para valores de x separadamente nos conjuntos $(-\infty, 0)$, $\{0\}$, $(0, 1)$, $\{1\}$ e $(1, \infty)$. Primeiro veja que se $x \in (-\infty, 0)$.



X não pode assumir valores menores que x, veja que todos os elementos de $Im(X)$ estão à direta de x na reta acima. Então $F_X(x) = 0$ sempre que $x < 0$.

E se $x = 0$? $F_X(0) = \mathrm{P}(X \leq 0) = \mathrm{P}(X = 0) = \mathrm{P}(\text{sair coroa}) = 1/2$.

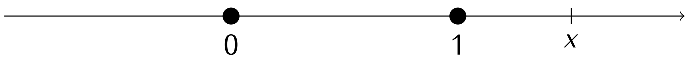
Vejamos agora a situação em que $0 < x < 1$, $x \in (0, 1)$.



A única possibilidade de X assumir valores menores que x é quando $X = 0$, veja a reta acima. Então $F_X(x) = \mathrm{P}(X \leq x) = \mathrm{P}(X = 0) = 1/2$ sempre que $0 < x < 1$.

Se $x = 1$, $F_X(1) = \mathrm{P}(X \leq 1) = \mathrm{P}(X = 0 \text{ ou } X = 1) = \mathrm{P}(\text{sair cara ou coroa}) = \mathrm{P}(\Omega) = 1$.

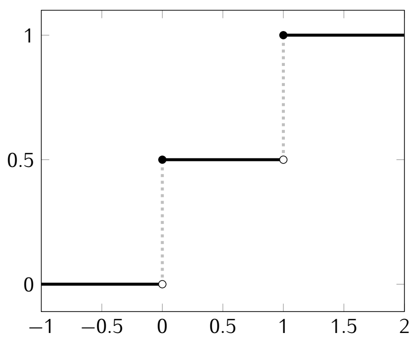
Para terminar, considere $x > 1$, isto é, $x \in (1, \infty)$.



Nesse caso $F_X(x) = \mathrm{P}(X \leq x) = \mathrm{P}(X = 0 \text{ ou } X = 1) = \mathrm{P}(\text{sair cara ou coroa}) = \mathrm{P}(\Omega) = 1$.

Concluindo, a função F_X e seu gráfico estão definidos a seguir.

$$F_X(x) = \begin{cases} 0 & , \text{ se } x < 0; \\ 1/2 & , \text{ se } 0 \leq x < 1; \\ 1 & , \text{ se } x \geq 1. \end{cases}$$



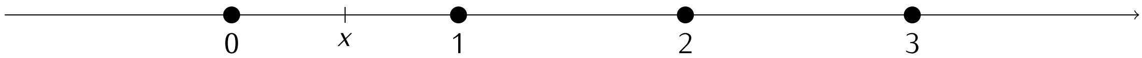
♦♦

Exemplo 1.14
Considere o experimento e a v.a. descritos no Exemplo 1.8. Encontre e esboce o gráfico da função de distribuição de X, definida pelo número de caras em 3 lançamentos de uma moeda.

Solução:
Já vimos que $Im(X) = \{0, 1, 2, 3\}$. Vamos primeiro determinar $F_X(x)$ para $x \in (-\infty, 0)$.

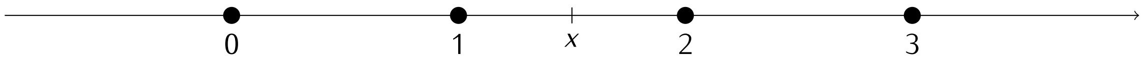
Veja que se $x < 0$ não tem como a v.a. X assumir valores menores ou iguais a x, pois todos os elementos da $Im(X)$ são maiores que x. Por isso, se $x < 0$ temos $F_X(x) = \mathrm{P}(X \leq x) = 0$. Já se $x = 0$, temos $F_X(0) = \mathrm{P}(X \leq 0) = \mathrm{P}(X = 0) = \mathrm{P}(\text{sair } KKK) = 1/8$.

Agora vajamos o caso em que $x \in (0, 1)$.



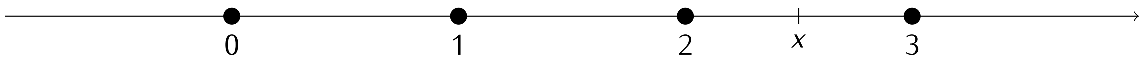
Nesse caso temos $F_X(x) = \mathrm{P}(X \leq x) = \mathrm{P}(X = 0) = \mathrm{P}(KKK) = 1/8$. Já se $x = 1$ temos $F_X(1) = \mathrm{P}(X \leq 1) = \mathrm{P}(X = 0 \ \cup \ X = 1) = \mathrm{P}(\{KKK, KKC, KCK, CKK\}) = 4/8 = 1/2$.

Suponha agora $x \in (1, 2)$.



Nesse caso temos $F_X(x) = \mathrm{P}(X \leq x) = \mathrm{P}(X = 0 \ \cup \ X = 1) = \mathrm{P}(\{KKK, KKC, KCK, CKK\}) = 4/8 = 1/2$. Já se $x = 2$ temos $F_X(2) = \mathrm{P}(X \leq 2) = \mathrm{P}(X = 0 \ \cup \ X = 1 \ \cup \ X = 2) = \mathrm{P}(\{KKK, KKC, KCK, CKK, KCC, CKC, CCK\}) = 7/8$.

Ainda temos o caso em que $x \in (2, 3)$.



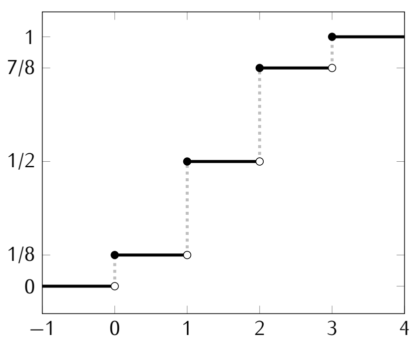
Nesse caso temos $F_X(x) = \mathrm{P}(X \leq x) = \mathrm{P}(X = 0 \ \cup \ X = 1 \ \cup \ X = 2) = 7/8$. Já se $x = 3$ temos $F_X(3) = \mathrm{P}(X \leq 3) = \mathrm{P}(X = 0 \ \cup \ X = 1 \ \cup \ X = 2 \ \cup X = 3) = \mathrm{P}(\Omega) = 1$.

Para terminar, seja $x \in (3, \infty)$.

Nesse caso $F_X(x) = \mathrm{P}(X \leq x) = \mathrm{P}(X = 0 \ \cup \ X = 1 \ \cup \ X = 2 \ \cup X = 3) = \mathrm{P}(\Omega) = 1$.

Concluindo, a função F_X e seu gráfico estão definidos a seguir.

$$F_X(x) = \begin{cases} 0 & \text{, se } x < 0; \\ 1/8 & \text{, se } 0 \leq x < 1; \\ 1/2 & \text{, se } 1 \leq x < 2; \\ 7/8 & \text{, se } 2 \leq x < 3; \\ 1 & \text{, se } x \geq 3. \end{cases}$$



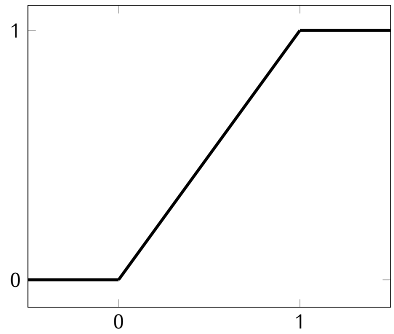
♦♦

Exemplo 1.15
Suponha o experimento de sortear de forma aleatória um ponto no intervalo $(0, 1)$. Para esse experimento defina X como sendo o ponto sorteado. Encontre a função de distribuição de X.

Solução:
Veja que $Im(X) = (0, 1)$. Dado $x \in (0, 1)$ qualquer, $P(X \leq x) = x$. Se $x \leq 0$, $P(X \leq x) = 0$, pois não há números possíveis de serem sorteados menores que 0. Já se $x \geq 1$, $P(X \leq x) = 1$, pois todos os números possíveis de serem sorteados são menores ou iguais a 1.

Então a função de distribuição F_X pode ser definida por:

$$F_X(x) = \begin{cases} 0 & , \text{ se } x \leq 0; \\ x & , \text{ se } 0 < x < 0; \\ 1 & , \text{ se } x \geq 1. \end{cases}$$



♦♦

Proposição 1.16 Propriedades da Função de Distribuição
Seja F_X a função de distribuição de alguma variável aleatória X. Então F_X satisfaz as seguintes propriedades.

F_1. $\lim_{x\to\infty} F_X(x) = 1$ *e* $\lim_{x\to-\infty} F_X(x) = 0$.

F_2. *F_X é função não decrescente, isto é, se $a < b \Rightarrow F_X(a) \leq F_X(b)$.*

F_3. *F_X é função contínua à direita, isto é, $F_X(b^+) \triangleq \lim_{h\to 0} F_X(b + h) = F_X(b)$.*

Demonstração:
As propriedades a serem demonstradas são decorrentes dos axiomas e das propriedades da probabilidade, veja Definição 1.4 e Proposição 1.5.

F_1. $\lim_{x\to\infty} F_X(x) = \lim_{x\to\infty} P(X \leq x) = P(X \leq \infty) = P(\Omega) = 1.$

$\lim_{x\to-\infty} F_X(x) = \lim_{x\to-\infty} P(X \leq x) = P(X \leq -\infty) = P(\emptyset) = 0.$

F_2. Se $a < b$, então o evento $\{X \leq a\} \subset \{X \leq b\}$ e, portanto, $P(\{X \leq a\}) \leq P(\{X \leq b\})$, ou seja, $F_X(a) \leq F_X(b)$.

F_3. $F_X(b^+) \triangleq \lim_{h\to 0} F_X(b + h) = \lim_{h\to 0} P(X \leq b + h) =$
$\lim_{h\to 0} P(X < b \ \cup \ b \leq X \leq b + h) = \lim_{h\to 0} \{P(X < b) + P(b \leq X \leq b + h)\} =$
$P(X < b) + \lim_{h\to 0} P(b \leq X \leq b + h) = P(X < b) + P(X = b) = P(X \leq b) = F_X(b).$

□

Exemplo 1.17 *Verifique as quatro propriedades da função de distribuição, listadas na Proposição 1.16, para a função F_X encontrada no Exemplo 1.8, onde X é o número de caras em 3 lançamentos de uma moeda.*

Conhecendo a função de distribuição de uma variável aleatória qualquer é possível calcular a probabilidade dessa variável aleatória pertencer a qualquer intervalo da reta. Veja no Exemplo 1.18 alguns exemplos em que o cálculo de probabilidades são convertidos para a função de distribuição.

Exemplo 1.18 Cálculo de probabilidades a partir da função de distribuição
Seja X uma v.a. e F_X a sua função de distribuição. Sejam a e b números reais tais que $a < b$. Reescreva as probabilidades a seguir em termos de F_X.

(a) $\mathrm{P}(X \leq a)$	*(b)* $\mathrm{P}(X > a)$	*(c)* $\mathrm{P}(a < X \leq b)$	*(d)* $\mathrm{P}(X = a)$
(e) $\mathrm{P}(X < a)$	*(f)* $\mathrm{P}(X \geq a)$	*(g)* $\mathrm{P}(a < X < b)$	*(h)* $\mathrm{P}(a \leq X \leq b)$

Solução:
A ideia é transformar o evento para o qual queremos calcular a probabilidade em um evento do tipo $\{X \leq c\}$, pois nesse caso sabemos que $\mathrm{P}(X \leq c) = F_X(c)$. Para isso vamos usar as propriedades da probabilidade (Proposição 1.5).

(a) $\mathrm{P}(X \leq a) = F_X(a)$.

(b) $\mathrm{P}(X > a) = 1 - \mathrm{P}(X \leq a) = 1 - F_X(a)$.

(c) $\mathrm{P}(a < X \leq b) = \mathrm{P}(\{X \leq b\} \cap \{X > a\}) = \mathrm{P}(X \leq b) - \mathrm{P}(\{X \leq a\}) = F_X(b) - F_X(a)$.

(d) $\mathrm{P}(X = a) = \mathrm{P}(\{X \leq a\} \cap \{X \geq a\}) = \mathrm{P}(X \leq a) - \mathrm{P}(X < a) = F_X(a) - \lim_{h \to 0^+} \mathrm{P}(X \leq a - h) = F_X(a) - F_X(a^-)$, que é o tamanho do salto da função F_X no ponto a.

(e) $\mathrm{P}(X < a) = \mathrm{P}(X \leq a) - \mathrm{P}(X = a) = F_X(a) - \mathrm{P}(X = a)$, lembrando que $\mathrm{P}(X = a)$ é o tamanho do salto da função F_X no ponto a.

(f) $\mathrm{P}(X \geq a) = 1 - P(X < a) = 1 - (F_X(a) - \mathrm{P}(X = a)) = 1 - F_X(a) + \mathrm{P}(X = a)$.

(g) $\mathrm{P}(a < X < b) = \mathrm{P}(X < b) - \mathrm{P}(X \leq a) = F_X(b) - \mathrm{P}(X = b) - F_X(a)$.

(h) $\mathrm{P}(a \leq X \leq b) = \mathrm{P}(X \leq b) - \mathrm{P}(X < a) = F_X(b) - (F_X(a) - \mathrm{P}(X = a)) = F_X(b) - F_X(a) + \mathrm{P}(X = a)$.

♦♦

O Exemplo 1.18 acima nos mostra que conhecendo a função de distribuição de uma variável aleatória podemos calcular qualquer probabilidade envolvendo essa variável. Ou seja, o comportamento de uma variável aleatória fica perfeitamente determinado através da sua função de distribuição.

Exemplo 1.19
Seja X a v.a. que representa o tempo de vida (em dias) de um dispositivo elétrico. Suponha que

$$F_X(x) = \begin{cases} 0 & , x < 0; \\ 1 - e^{-x/20} & , x \geq 0. \end{cases}$$

(a) *Esboce o gráfico da função F_X.*

(b) *Verifique as propriedades da função de distribuição para F_X (Proposição 1.16).*

(c) *Calcule a probabilidade do dispositivo durar mais de 30 dias.*

(d) *Calcule a probabilidade do dispositivo durar entre 10 e 30 dias.*

(e) *Suponha que o dispositivo está em funcionamento há 10 dias. Qual a probabilidade do seu tempo de vida ultrapassar os 30 dias?*

Solução:

(c) P(dispositivo durar mais de 30 dias) $= \mathrm{P}(X > 30) = 1 - \mathrm{P}(X \leq 30) = 1 - F_X(30) = 1 - (1 - e^{-30/20}) = e^{-3/2} = 0,2231$.

(d) P(dispositivo durar entre 10 e 30 dias) $= \mathrm{P}(10 \leq X \leq 30) = \mathrm{P}(X \leq 30) - \mathrm{P}(X < 10) = \mathrm{P}(X \leq 30) - (\mathrm{P}(X \leq 10) - \mathrm{P}(X = 10)) = F_X(30) - (F_X(10) - 0) = F_X(30) - F_X(10) = 1 - e^{-30/20} - 1 + e^{-10/20} = e^{-1/2} - e^{-3/2} = 0,6065 - 0,2231 = 0,3834$.

(e) P(tempo de vida ultrapassar os 30 dias | dispositivo está em funcionamento há 10 dias) $= \mathrm{P}(X > 30 \mid X > 10) = \mathrm{P}(X > 30 \cap X > 10)/\,\mathrm{P}(X > 10) = \mathrm{P}(X > 30)/\,\mathrm{P}(X > 10) = (1 - \mathrm{P}(X \leq 30))/(1 - \mathrm{P}(X \leq 10)) = (1 - F_X(30))/(1 - F_X(10)) = e^{-30/20}/e^{-10/20} = e^{-1} = 0,3679$.

♦♦

Exemplo 1.20
Seja X uma v.a. cuja função de distribuição é definida por:

$$F_X(x) = \begin{cases} 0 & ,x < -1; \\ 1/5 & ,-1 \leq x < 1; \\ 2/5 & ,1 \leq x < 2; \\ 4/5 & ,2 \leq x < 4; \\ 1 & ,x \geq 4. \end{cases}$$

(a) Esboce o gráfico da função F_X.

(b) Verifique as quatro propriedades da função de distribuição para F_X.

(c) Calcule: $\mathrm{P}(X > 0)$, $\mathrm{P}(X = 0)$, $\mathrm{P}(X = 2)$, $\mathrm{P}(X < 2)$, $\mathrm{P}(0 < X < 4)$ *e* $\mathrm{P}(X > 0 \mid X \leq 2)$.

Solução:

(c) $\mathrm{P}(X > 0) = 1 - \mathrm{P}(X \leq 0) = 1 - F_X(0) = 1 - 1/5 = 4/5$.

$\mathrm{P}(X = 0) = F_X(0) - F_X(0^-) = 0$, pois não há salto no gráfico de F_X em $x = 0$.

$\mathrm{P}(X = 2) = F_X(2) - F_X(2^-) = 4/5 - 2/5 = 2/5$, que é o tamanho do salto no gráfico de F_X em $x = 2$.

$\mathrm{P}(X < 2) = \mathrm{P}(X \leq 2) - \mathrm{P}(X = 2) = F_X(2) - \mathrm{P}(X = 2) = 4/5 - 2/5 = 2/5$.

$\mathrm{P}(0 < X < 4) = \mathrm{P}(X < 4) - \mathrm{P}(X \leq 0) = \mathrm{P}(X \leq 4) - \mathrm{P}(X = 4) - \mathrm{P}(X \leq 0) = F_X(4) - (F_X(4) - F_X(4^-)) - F_X(0) = 1 - (1 - 4/5) - 1/5 = 4/5 - 1/5 = 3/5$.

$\mathrm{P}(X > 0 \mid X \leq 2) = \mathrm{P}(X > 0 \cap X \leq 2)/\,\mathrm{P}(X \leq 2) = \mathrm{P}(0 < X \leq 2)/\,\mathrm{P}(X \leq 2) = (\mathrm{P}(X \leq 2) - \mathrm{P}(X \leq 0))\,/\,\mathrm{P}(X \leq 2) = (F_X(2) - F_X(0))\,/F_X(2) = (4/5 - 1/5)\,/(4/5) = (3/5)/(4/5) = 3/4$.

♦♦

1.4 Classificação de Variáveis Aleatórias

Cada variável aleatória pode ser classificada como: discreta, contínua, singular ou mista. Neste curso, por se tratar de um curso introdutório da Teoria das Probabilidades, estudaremos apenas as variáveis discretas e contínuas.

Definição 1.21 Variável Aleatória Discreta
*Uma variável aleatória X é classificada como **discreta** se a sua imagem é um conjunto finito ou enumerável.*

Nos Exemplos 1.8 e 1.10 as variáveis X e I_A são discretas. Em ambos os casos a imagem delas é finita. Também no Exemplo 1.9 (problema do dardo) a variável Y, definida pela pontuação ganha, é uma variável aleatória discreta com imagem finita, mas a variável X, definida pela distância entre o dardo e o centro, não é variável aleatória discreta.

Podemos citar um exemplo de variável aleatória discreta com imagem infinita e enumerável. Considere o experimento de lançar um dado até sair a face 6. Seja X a variável definida pelo número de lançamentos desse dado. Veja que $Im(X) = \{1, 2, 3, \ldots\}$, logo $Im(X)$ é um conjunto infinito porém enumerável e por isso a variável aleatória X é discreta.

O gráfico da função de distribuição das variáveis aleatórias discretas sempre será do tipo escada. Cada degrau dessa escada está localizado no elementos da imagem da variável aleatória. Alem disso, o tamanho do salto no degrau localizado em $X = x$ é exatamente $\mathrm{P}(X = x)$. Reveja o Exemplo 1.14 e verifique essa característica no gráfico de F_X.

Exemplo 1.22 Nota média de dois alunos
Um curso tem cinco alunos com coeficiente de rendimento (CR) superior a 8,5. Os CRs desses alunos são: 8,8; 9,2; 8,9; 9,5; 9,0. Dentre esses cinco alunos, dois serão sorteados para receber uma bolsa de estudos.

(a) Designando por A, B, C, D, E os alunos, defina um espaço amostral para o experimento definido pelo sorteio dos dois alunos que receberão a bolsa de estudos.

Seja X = CR médio dos alunos sorteados.

(b) Defina $Im(X)$.

(c) Classifique X como variável aleatória discreta ou não. Justifique sua resposta.

(d) Liste o evento $\{X \geq 9,0\}$, isto é, apresente os elementos de Ω que pertencem a esse evento.

(e) Apresente a função de distribuição de X.

(f) Calcule $\mathrm{P}(X \geq 9)$.

Solução:

(a) Note que não importa a ordem que os alunos são sorteados; logo, $n(\Omega) = \binom{5}{2} = 10$. Mais especificamente,

$$\Omega = \left\{ \begin{array}{c} (A, B), (A, C), (A, D), (A, E), (B, C), \\ (B, D), (B, E), (C, D), (C, E), (D, E) \end{array} \right\}$$

(b) Usando uma tabela de duas entradas, podemos representar os valores de X da seguinte forma:

	$A(8,8)$	$B(9,2)$	$C(8,9)$	$D(9,5)$	$E(9,0)$
$A(8,8)$		$\frac{8,8+9,2}{2} = \mathbf{9,0}$	$8,85$	$\mathbf{9,15}$	$8,90$
$B(9,2)$			$\mathbf{9,05}$	$\mathbf{9,35}$	$\mathbf{9,10}$
$C(8,9)$				$\mathbf{9,20}$	$8,95$
$D(9,5)$					$\mathbf{9,25}$
$E(9,0)$					

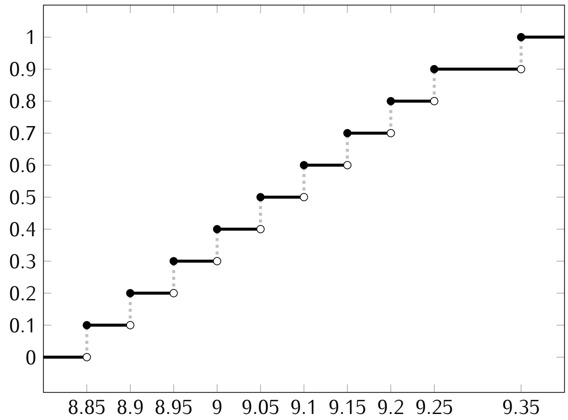
Logo, $Im(X) = \{8,85\ ,\ 8,90\ ,\ 8,95\ ,\ 9,00\ ,\ 9,05\ ,\ 9,10\ ,\ 9,15\ ,\ 9,20\ ,\ 9,25\ ,\ 9,35\}$.

(c) Como $Im(X)$ é um conjunto finito, em particular tem 10 elementos, podemos dizer que X é variável aleatória discreta.

(d) Como todos os pontos do espaço amostral são equiprováveis (o sorteio é aleatório), podemos afirmar que $P(\omega) = 1/10$ para qualquer $\omega \in Im(X)$. Usando as propriedades da função de distribuição de variáveis aleatórias discretas (já sabemos que é uma função escada) podemos concluir que:

(e) $\{X \geq 9\} = \{(A,B)\,,(A,D)\,,(B,C)\,,(B,D)\,,(B,E)\,,(C,D)\,,(D,E)\}$.

$$F_X(x) = \begin{cases} 0 & ,\ \text{se } x < 8,85; \\ 1/10 & ,\ \text{se } 8,85 \leq x < 8,90; \\ 2/10 & ,\ \text{se } 8,90 \leq x < 8,95; \\ 3/10 & ,\ \text{se } 8,95 \leq x < 9,00; \\ 4/10 & ,\ \text{se } 9,00 \leq x < 9,05; \\ 5/10 & ,\ \text{se } 9,05 \leq x < 9,10; \\ 6/10 & ,\ \text{se } 9,10 \leq x < 9,15; \\ 7/10 & ,\ \text{se } 9,15 \leq x < 9,20; \\ 8/10 & ,\ \text{se } 9,20 \leq x < 9,25; \\ 9/10 & ,\ \text{se } 9,25 \leq x < 9,35; \\ 1 & ,\ \text{se } x \geq 9,35. \end{cases}$$



(f) Podemos calcular $P(X \geq 9)$ pela função de distribuição ou pela listagem do evento $\{X \geq 9\}$. Vamos fazer dos dois jeitos. Primeiro pela função de distribuição.

$$\begin{aligned} P(X \geq 9) &= 1 - P(X < 9) = 1 - (P(X \leq 9) - P(X = 9)) = \\ &= 1 - F_X(9) + P(X = 9) = 1 - \frac{4}{10} + (\frac{4}{10} - \frac{3}{10}) = \frac{7}{10}. \end{aligned}$$

Agora pela listagem do evento $\{X \geq 9\}$.

$$P(X \geq 9) = P((A,B)\,,(A,D)\,,(B,C)\,,(B,D)\,,(B,E)\,,(C,D)\,,(D,E)) = 7 \times \frac{1}{10} = \frac{7}{10}.$$

♦♦

Então se uma variável aleatória tem gráfico do tipo escada, já sabemos de que se trata de uma variável aleatória discreta.

Veremos a definição mais usual de variável aleatória contínua no Capítulo 4, mas a princípio temos como verificar se uma variável aleatória é contínua a partir da sua função de distribuição.

Definição 1.23 Variável Aleatória Contínua
Uma variável aleatória X é classificada como ***contínua*** *quando a sua função de distribuição é uma função absolutamente contínua.*

Precisamos ter atenção com o termo **função absolutamente contínua**, que não é a mesma coisa que **função contínua**. Mas para facilitar o entendimento de variáveis aleatórias contínuas no contexto de um primeiro curso de Probabilidade, podemos usar o resultado apresentado na Proposição 1.24.

Proposição 1.24
Seja F é uma função que satisfaz:

(i) F é contínua em $\mathbb{R}$;

(ii) F é diferenciável em $\mathbb{R}$, exceto em um conjunto finito de pontos.

Então F é absolutamente contínua.

Demonstração:
A demostração desta proposição será omitida.

□

Então, combinando os resultados da Definição 1.23 e da Proposição 1.24, toda vez que uma variável aleatória X apresentar função de distribuição F satisfazendo as condições (i) e (ii) da Proposição 1.24, podemos afirmar que X é variável aleatória contínua. Veja que a volta não é verdadeira, ou seja, podem existir variáveis aleatórias contínuas tais que sua função de distribuição F não satisfaça (i) ou (ii). Mas esses casos não serão tratados nesse curso.

Nos Exemplos 1.15 e 1.19 as variáveis aleatórias definidas são contínuas. Veja que em ambos os exemplos a função de distribuição F_X é contínua. Além disso F_X só não é diferenciável em um conjunto finito de pontos: no Exemplo 1.15 F_X só não é diferenciável em dois pontos, $x = 0$ e $x = 1$, e no Exemplo 1.19 a função F_x só não é diferenciável no ponto $x = 0$.

Já vimos que o comportamento de uma variável aleatória fica perfeitamente determinado através da sua função de distribuição. Além de calcular as probabilidades envolvendo a variável aleatória em questão podemos também, através do conhecimento da função de distribuição, encontrar a imagem da variável aleatória. No caso de X ser uma variável aleatória discreta a imagem de X é formado pelos pontos do domínio de F_X onde ocorrem os saltos no seu gráfico. Já se X é variável aleatória contínua, a sua imagem é formada pelos pontos do domínio de F_X onde esta função é crescente, ou seja, não constante.

Exemplo 1.25
Dada a função

$$F(x) = \begin{cases} 0 & , x < 1 \\ 1/2\,, & 1 \leq x < 2 \\ k & , 2 \leq x < 3 \\ 3/4\,, & 3 \leq x < 4 \\ 1 & , x \geq 4 \end{cases}$$

em que k é uma constante, determine os possíveis valores de k para que F seja a função de distribuição acumulada de uma variável aleatória discreta X. Em seguida, determine a imagem de X.

Solução:
Como a função de distribuição de qualquer v.a. X tem que ser uma função não decrescente, concluímos que k tem que ser maior ou igual a $\frac{1}{2}$. Pela mesma razão, k tem que ser menor ou igual a $\frac{3}{4}$. Dessa forma, os possíveis valores de k pertencem ao intervalo $\left[\frac{1}{2}, \frac{3}{4}\right]$. Os valores possíveis da v.a. X correspondem aos pontos de descontinuidade da função $F(x)$. Logo, $Im(X) = \{1, 2, 3, 4\}$.

♦♦

Exemplo 1.26
Dada a função

$$F(x) = \begin{cases} 0 & , x < 0 \\ x^2/2 & , 0 \leq x < 1 \\ kx & , 1 \leq x < 2 \\ 1 & , x \geq 2 \end{cases}$$

em que k é uma constante, determine os possíveis valores de k para que F seja a função de distribuição acumulada de uma variável aleatória contínua X. Em seguida, determine a imagem de X.

Solução:
Para que F seja função de distribuição de uma variável aleatória contínua é necessário que não haja "saltos" no gráfico de F. Para que isso ocorra o valor de k tem que ser tal que $\lim_{x\to 1^-} F(x) = \lim_{x\to 1^+} F(x)$ e $\lim_{x\to 2^-} F(x) = \lim_{x\to 2^+} F(x)$. Isto é, $1/2 = k$ e $2k = 1$, ou seja, $k = 1/2$. Para terminar, a imagem de F é formada pelos valores do domínio onde a função é crescente, ou seja, $Im(X) = [0, 2]$.

♦♦

Para terminar esse capítulo, vejamos um exemplo de uma variável aleatória que não é nem discreta nem contínua.

Exemplo 1.27
Seja X variável aleatória com função de distribuição definida por:

$$F_X(x) = \begin{cases} 0 & , x < 0 \\ x & , 0 \leq x < 1/2 \\ 1 & , x \geq x < 1/2 \end{cases}$$

(a) Faça o gráfico de F_X.

(b) Verifique que F_X é função de distribuição.

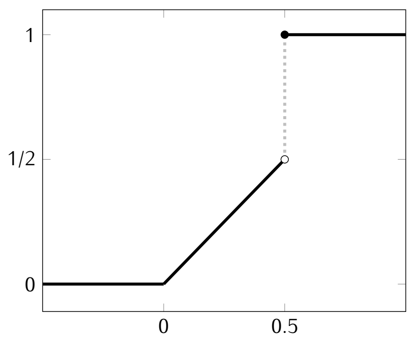
(c) Calcule $\mathrm{P}(X > 1/3)$, $\mathrm{P}(X = 1/4)$ *e* $\mathrm{P}(X = 1/2)$.

(d) Verifique que X não é variável aleatória discreta.

(e) Verifique que X não é variável aleatória contínua.

Solução:

(a) Segue o gráfico de F_X.



(b) Para verificar o que se pede precisamos verificar as propriedades da Proposição 1.16.

- Primeiro veja pelo gráfico que $\lim_{x\to\infty} F_X(x) = 1$ e $\lim_{x\to-\infty} F_X(x) = 0$.
- Veja que F_X é função crescente no intervalo $(0, 1/2)$ e constante no restante da reta. Então F_X é função não decrescente.
- Veja que o único ponto de descontinuidade de F_X é o $(1/2, 1/2)$. Precisamos então mostras que nesse ponto F_X é contínua à direita.
  Seja $\{x_i\}$ uma sequencia tal que $x_i \to 1/2$ e $x_i \leq 1/2$, isto é, $\{x_i\}$ é uma sequência que converge para 1/2 pela direita, então $F_X(x_i) = 1 \ \forall\ i$. Logo a sequência $\{F_X(x_i)\}$ é constante e igual a 1 e por isso $F_X(x_i) \to 1 = F_X(0.5)$. Assim mostramos que F_X é contínua a direita em 0.5.

(c) $P(X > 1/3) = 1 - F_X(1/3) = 1 - 1/3 = 2/3$.
$P(X = 1/4) = F_X(1/4) - F_X(1/4^-) = 0$. Veja que não existe salto no gráfico de F_X em $x = 1/4$.
$P(X = 1/2) = F_X(1/2) - F_X(1/2^-) = 1/2$. Veja que o tamanho do salto no gráfico de F_X em $x = 1/2$ é $1/2$.

(d) Veja que $Im(X) = [0, 1/2]$. Como esse conjunto não é enumerável, X não é variável aleatória discreta.

(e) Veja que o gráfico de F_X não é contínuo. Logo X não é variável aleatória contínua.

♦♦

Capítulo 2

Variáveis Aleatórias Discretas

Nesse capítulo, vamos estudar em mais detalhes as variáveis aleatórias discretas.

2.1 Função de Probabilidade

Já vimos no Capítulo 1 que uma variável aleatória é chamada de discreta se a sua imagem é um conjunto finito ou enumerável. Nesse caso o gráfico da sua função de distribuição é do tipo escada e o tamanho do degrau localizado em $X = x$ é exatamente $P(X = x)$.

Além disso, o comportamento de uma variável aleatória qualquer, discreta ou não, é perfeitamente determinado pela sua função de distribuição. No caso das variáveis aleatórias discretas temos outra opção para determinar o comportamento da variável aleatória, a *função de probabilidade*.

Definição 2.1 Função de Probabilidade
Seja X uma variável aleatória discreta. A ***função de probabilidade*** *de X é a função p_X definida por:*

$$p_X(x) = P(X = x)$$

Para encontrar a função de probabilidade de uma variável aleatória temos que primeiro encontrar a sua imagem e em seguida calcular a probabilidade de ocorrer cada elemento da imagem. Todos os valores reais que não pertencem a imagem tem probabilidade nula de ocorrer, logo para esses valores a função de probabilidade também é nula.

Exemplo 2.2 Máximo entre dois dados
Considere o lançamento de dois dados equilibrados.

(a) Apresente um espaço amostral para este experimento, isto é, apresente Ω.

Suponha que nosso interesse esteja no máximo das faces dos dois dados. Neste caso, defina a variável aleatória X = "máximo das 2 faces".

(b) Encontre $X(\omega)$ para todo $\omega \in \Omega$.

(c) Defina $Im(X)$.

(d) Verifique que X é v.a. discreta.

(e) Encontre a função de probabilidade de X e esboce seu gráfico.

(f) Encontre a função de distribuição de X e esboce o seu gráfico.

Solução:

(a)
$$\begin{aligned}\Omega &= \{(i,j) \mid i,j = 1,2,3,4,5,6\} \\ &= \{(1,1),(1,2),\ldots,(1,6),(2,1),(2,2),\ldots,(2,6),\ldots,(6,6)\}.\end{aligned}$$

(b) A Tabela 2.1 abaixo apresenta $X(\omega)$ para todo $\omega \in \Omega$.

Tabela 2.1 – Variável aleatória X = "máximo das faces de 2 dados"

Pontos do espaço amostral (ω)	Valor de $X(\omega)$
(1,1)	1
(1,2)(2,2),(2,1)	2
(1,3),(2,3),(3,3),(3,2),(3,1)	3
(1,4),(2,4),(3,4),(4,4),(4,3),(4,2),(4,1)	4
(1,5),(2,5),(3,5),(4,5),(5,5),(5,4),(5,3),(5,2),(5,1)	5
(1,6),(2,6),(3,6),(4,6),(5,6),(6,6),(6,5),(6,4),(6,3),(6,2),(6,1)	6

(c) De acordo com a Tabela 2.1, $Im(X) = \{1,2,3,4,5,6\}$.

(d) Veja que o conjunto $Im(X)$ é finito, logo X é discreta.

(e) Queremos calcular $p_X(x) = \mathrm{P}(X = x)$ para todo $x \in Im(X)$.

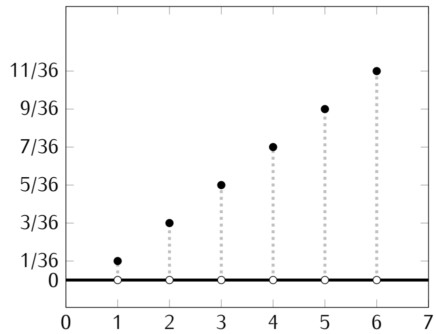
Vamos começar com $x = 1$: $p_X(1) = \mathrm{P}(X = 1) = \mathrm{P}(\{(1,1)\}) = 1/36$, considerando que cada um dos 36 elementos de Ω tem mesma probabilidade de ocorrer.

Vamos calcular agora p_X nos demais elementos da $Im(X)$:

$$\begin{aligned}
p_X(2) &= \mathrm{P}(X=2) = && \mathrm{P}(\{(1,2),(2,1),(2,2)\}) = 3/36.\\
p_X(3) &= \mathrm{P}(X=3) = && \mathrm{P}(\{(1,3),(3,1),(2,3),(3,2),(3,3)\}) = 5/36.\\
p_X(4) &= \mathrm{P}(X=4) = && \mathrm{P}(\{(1,4),(4,1),(2,4),(4,2),(3,4),(4,3),(4,4)\}) = 7/36.\\
p_X(5) &= \mathrm{P}(X=5) = && \mathrm{P}(\{(1,5),(5,1),(2,5),(5,2),(3,5),(5,3),(4,5),(5,4),(5,5)\})\\
& && = 9/36.\\
p_X(6) &= \mathrm{P}(X=6) = && \mathrm{P}(\{(1,6),(6,1),(2,6),(6,2),(3,6),(6,3),(4,6),(6,4),(5,6),(6,5),\\
& && (6,6)\}) = 11/36.
\end{aligned}$$

Então,

$$p_X(x) = \begin{cases} 1/36 & , \text{se } x = 1 \\ 3/36 & , \text{se } x = 2 \\ 5/36 & , \text{se } x = 3 \\ 7/36 & , \text{se } x = 4 \\ 9/36 & , \text{se } x = 5 \\ 11/36 & , \text{se } x = 6 \\ 0 & , \text{caso contrário.} \end{cases}$$

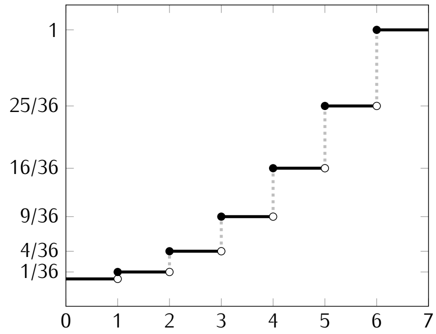


Também podemos definir a função de probabilidade p_X em forma de uma tabela:

x	1	2	3	4	5	6
$p_X(x)$	$\frac{1}{36}$	$\frac{3}{36}$	$\frac{5}{36}$	$\frac{7}{36}$	$\frac{9}{36}$	$\frac{11}{36}$

(f) Como já conhecemos $P(X = x)$ para todo $x \in Im(X)$ basta construir uma função escada, com os degraus ocorrendo em cada ponto de $Im(X)$ e o tamanho do degrau equivalente à $P(X = x)$. Uma vez o gráfico pronto podemos definir a expressão para F_X.

$$F_X(x) = \begin{cases} 0 & \text{, se } x < 1; \\ 1/36 & \text{, se } 1 \leq x < 2; \\ 4/36 & \text{, se } 2 \leq x < 3; \\ 9/36 & \text{, se } 3 \leq x < 4; \\ 16/36 & \text{, se } 4 \leq x < 5; \\ 25/36 & \text{, se } 5 \leq x < 6; \\ 1 & \text{, se } x \geq 6. \end{cases}$$



♦♦

Proposição 2.3 Propriedades da Função de Probabilidade
Seja p_X a função de probabilidade de uma variável aleatória discreta qualquer. Então valem as seguintes propriedades:

(i) $p_X(x) \geq 0$ *para todo* $x \in \mathbb{R}$*;*

(ii) $\sum_{\forall x \in Im(X)} p_X(x) = 1$.

Demonstração:
As propriedades a serem demonstradas são decorrentes dos axiomas da probabilidade, veja Definição 1.4.

(i) Como $P(A) \geq 0$ para todo $A \in \mathcal{F}$, também é verdade que $P(X = x) \geq 0$ para todo $x \in \mathbb{R}$. Logo, $p_X(x) = P(X = x) \geq 0$ para todo $x \in \mathbb{R}$.

(ii) Veja que $\sum_{\forall x \in Im(X)} p_X(x) = \sum_{\forall x \in Im(X)} P(X = x) = P\left(\cup_{\forall x \in Im(X)} \{X = x\}\right) = P(\Omega) = 1$.

□

Exemplo 2.4 *Verifique as duas propriedades listadas na Proposição 2.3 para a função de probabilidade da variável aleatória do Exemplo 2.2.*

Exemplo 2.5 Demanda por produto
A demanda por um certo produto pode ser vista como uma variável aleatória X cuja função de probabilidade p_X é estimada por

Número de unidades demandadas x	1	2	3	4
$p_X(x) = P(X = x)$	0,25	0,45	0,15	0,15

(a) Verifique que p_X realmente define uma função de probabilidade.

(b) Obtenha a função de distribuição acumulada de X.

(c) Usando a função de distribuição calculada no item anterior, calcule $P(X \leq 3,5)$.

Solução:

(a) $0,25 + 0,45 + 0,15 + 0,15 = 1$ e todos os valores são não negativos. Logo, f_X é uma função de probabilidade.

(b)
$$F_X(x) = \begin{cases} 0 & \text{se } x < 1 \\ 0{,}25 & \text{se } 1 \leq x < 2 \\ 0{,}70 & \text{se } 2 \leq x < 3 \\ 0{,}85 & \text{se } 3 \leq x < 4 \\ 1{,}00 & \text{se } x \geq 4 \end{cases}$$

(c) Temos que
$$\mathrm{P}(X \leq 3,5) = F_X(3,5) = 0,85$$

♦♦

Exemplo 2.6
Uma variável aleatória discreta X tem a seguinte função de probabilidade

$$p_X(x) = \begin{cases} \frac{k}{(x+2)!} & , \text{ se } x = 0 \text{ ou } x = 1 \\ \\ 0 & , \text{ se } x \neq 0 \text{ e } x \neq 1 \end{cases}$$

onde k é uma constante real.

(a) Determine o valor de k para que p_X seja realmente função de probabilidade.

(b) Encontre a função de distribuição F_X e esboce seu gráfico.

Solução:

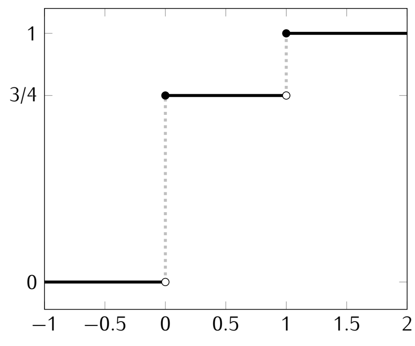
(a) Os valores possíveis da v.a. são 0 e 1, isto é, $Im(X) = \{0, 1\}$. Então, temos que ter

$$p_X(0) + p_X(1) = 1 \;\Rightarrow\; \frac{k}{2!} + \frac{k}{3!} = 1 \;\Rightarrow\; \frac{k}{2} + \frac{k}{6} = 1 \;\Rightarrow\; \frac{3k+k}{6} = 1 \;\Rightarrow\; k = \frac{6}{4} = \frac{3}{2}.$$

Logo, $p_X(0) = \frac{3/2}{2} = \frac{3}{4}$ e $p_X(1) = \frac{3/2}{6} = \frac{1}{4}$.

(b) A função de distribuição de X é

$$F_X(x) = \begin{cases} 0 & , \text{ se } x < 0; \\ 3/4 & , \text{ se } 0 \leq x < 1; \\ 1 & , \text{ se } x \geq 1. \end{cases}$$



♦♦

Exemplo 2.7 Soma das faces de dois dados
Considere novamente o experimento de lançar dois dados equilibrados. Agora o nosso interesse está na soma das faces dos dois dados. Neste caso, defina a variável aleatória X = "soma das faces".

(a) Defina $Im(X)$.

(b) Verifique que X é v.a. discreta.

(c) *Encontre a função de probabilidade de X, verifique as propriedades da função de probabilidade e esboce seu gráfico.*

(d) *Encontre a função de distribuição de X e esboce o seu gráfico.*

(e) *Calcule a probabilidade da soma ser menor que 6.*

Solução:

(a) $Im(X) = \{2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}$.

(b) Como $Im(X)$ é um conjunto finito, X é variável aleatória discreta.

(c) Para encontrar p_X precisamos encontrar $P(X = x)$ para todo $x \in Im(X)$. O espaço amostral desse experimento é o mesmo do Exemplo 2.2, ou seja, ele tem 36 elementos e todos têm mesma probabilidade de ocorre.

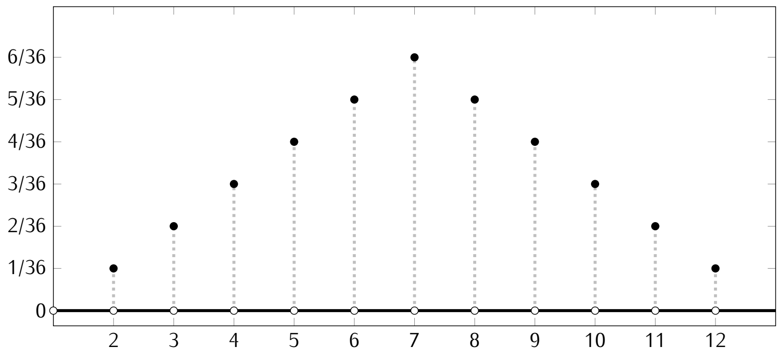
$$\begin{array}{lll}
p_X(2) & = P(X=2) = & P(\{(1,1)\}) = 1/36. \\
p_X(3) & = P(X=3) = & P(\{(1,2),(2,1)\}) = 2/36. \\
p_X(4) & = P(X=4) = & P(\{(1,3),(3,1),(2,2)\}) = 3/36. \\
p_X(5) & = P(X=5) = & P(\{(1,4),(4,1),(2,3),(3,2)\}) = 4/36. \\
p_X(6) & = P(X=6) = & P(\{(1,5),(5,1),(2,4),(4,2),(3,3)\}) = 5/36. \\
p_X(7) & = P(X=7) = & P(\{(1,6),(6,1),(2,5),(5,2),(3,4),(4,3)\}) = 6/36. \\
p_X(8) & = P(X=8) = & P(\{(2,6),(6,2),(3,5),(5,3),(4,4)\}) = 5/36. \\
p_X(9) & = P(X=9) = & P(\{(3,6),(6,3),(4,5),(5,4)\}) = 4/36. \\
p_X(10) & = P(X=10) = & P(\{(4,6),(6,4),(5,5)\}) = 3/36. \\
p_X(11) & = P(X=11) = & P(\{(5,6),(6,5)\}) = 2/36. \\
p_X(12) & = P(X=12) = & P(\{(6,6)\}) = 1/36.
\end{array}$$

Para qualquer outro valor $x \in \mathbb{R}$ temos $p_X(x) = 0$. Em forma de tabela podemos representar p_X por:

x	2	3	4	5	6	7	8	9	10	11	12
$p_X(x)$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

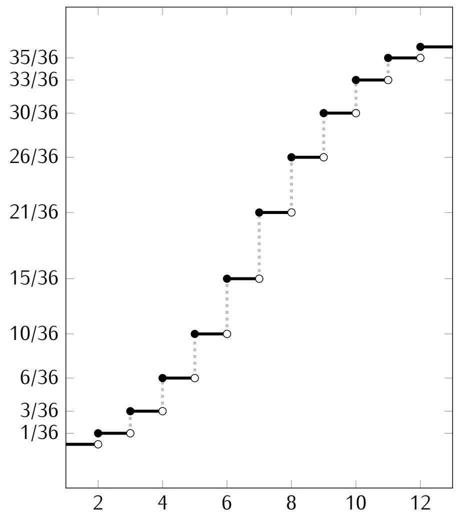
Veja que $p_X(x) \geq 0$ para todo x e que $\sum_{x \in Im(X)} p_X(x) = \frac{1}{36} + \frac{2}{36} + \frac{3}{36} + \frac{4}{36} + \frac{5}{36} + \frac{6}{36} + \frac{5}{36} + \frac{4}{36} + \frac{3}{36} + \frac{2}{36} + \frac{1}{36} = 1$. Logo as propriedades foram verificadas.

Para terminar esse item só falta o gráfico de p_X.



(d) Como já conhecemos $P(X = x)$ para todo $x \in Im(X)$ basta construir uma função escada, com os degraus ocorrendo em cada ponto de $Im(X)$ e o tamanho do degrau equivalente à $P(X = x)$. Uma vez o gráfico pronto podemos definir a expressão para F_X.

$$F_X(x) = \begin{cases} 0 & , \text{ se } x < 2; \\ 1/36 & , \text{ se } 2 \leq x < 3; \\ 3/36 & , \text{ se } 3 \leq x < 4; \\ 6/36 & , \text{ se } 4 \leq x < 5; \\ 10/36 & , \text{ se } 5 \leq x < 6; \\ 15/36 & , \text{ se } 6 \leq x < 7; \\ 21/36 & , \text{ se } 7 \leq x < 8; \\ 26/36 & , \text{ se } 8 \leq x < 9; \\ 30/36 & , \text{ se } 9 \leq x < 10; \\ 33/36 & , \text{ se } 10 \leq x < 11; \\ 35/36 & , \text{ se } 11 \leq x < 12; \\ 1 & , \text{ se } x \geq 12. \end{cases}$$



(e) Podemos fazer o cálculo usando a função de probabilidade ou a função de distribuição. Veja primeiro usando a função de probabilidade.

P($X > 6$) = P($X = 7$ ou $X = 8$ ou $X = 9$ ou $X = 10$ ou $X = 11$ ou $X = 12$) = P($X = 7$) + P($X = 8$) + P($X = 9$) + P($X = 10$) + P($X = 11$) + P($X = 12$) = $p_X(7) + p_X(8) + p_X(9) + p_X(10) + p_X(11) + p_X(12) = \frac{6}{36} + \frac{5}{36} + \frac{4}{36} + \frac{3}{36} + \frac{2}{36} + \frac{1}{36} = \frac{21}{36}$.

Agora usando a função de distribuição.

P($X > 6$) = 1 − P($X \leq 6$) = 1 − $F_X(6)$ = 1 − 15/36 = 21/36.

♦♦

Exemplo 2.8
Um homem possui quatro chaves em seu bolso. Como está escuro, ele não consegue ver qual a chave correta para abrir a porta de sua casa, que se encontra trancada. Ele testa uma chave de cada vez até encontrar a correta.

(a) Defina um espaço amostral para esse experimento.

Defina a variável aleatória X = número de chaves experimentadas até conseguir abrir a porta (inclusive a chave correta).

(b) Encontre $Im(X)$.

(c) Encontre a função de probabilidade de X e esboce o seu gráfico.

(d) Encontre a função de distribuição de X e esboce o seu gráfico.

(e) Qual a probabilidade do homem experimentar no máximo duas chaves para abrir a porta?

Solução:

(a) Vamos representar pela letra C o sorteio da chave correta, aquela que abre a porta, e por E o sorteio de uma chave errada, uma entre as 3 chaves que não abrem a porta. Então podemos representar o espaço amostral por:

$$\Omega = \{C, EC, EEC, EEEC\}$$

onde $\{C\}$ representa situação em que o homem acertou a chave de primeira, $\{EC\}$ ele errou na primeira tentativa e acertou na segunda, $\{EEC\}$ ele errou as duas primeiras tentativas e acertou na terceira e $\{EEEC\}$ representa a situação em que ele testou todas as chaves até achar aquela que abre a porta.

(b) Veja que o homem pode experimentar de 1 a 4 chaves antes de abrir a porta. Logo $Im(X) = \{1, 2, 3, 4\}$.

(c) $P(X = 1) = P(C) = 1/4$.

$P(X = 2) = P(EC) = 3/4 \times 1/3 = 1/4$.

$P(X = 3) = P(EEC) = 3/4 \times 2/3 \times 1/2 = 1/4$.

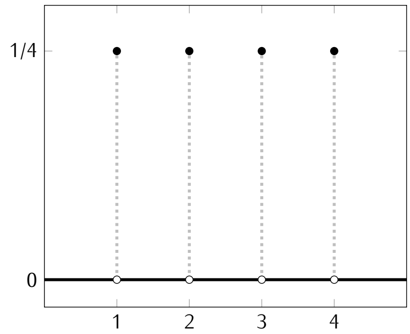
$P(X = 4) = P(EEEC) = 3/4 \times 2/3 \times 1/2 \times 1 = 1/4$.

Logo, a função de probabilidade de X e o seu gráfico são definidos por:

$$p_X(x) = \begin{cases} 1/4 & \text{, se } x = 1, 2, 3, 4 \\ 0 & \text{, caso contrário.} \end{cases}$$

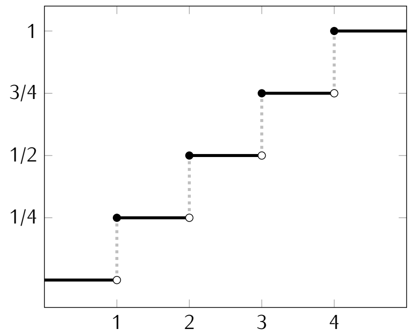
ou

x	1	2	3	4
$P(X = x)$	1/4	1/4	1/4	1/4



(d)

$$F_X(x) = \begin{cases} 0 & \text{, se } x < 1; \\ 1/4 & \text{, se } 1 \leq x < 2; \\ 1/2 & \text{, se } 2 \leq x < 3; \\ 3/4 & \text{, se } 3 \leq x < 4; \\ 1 & \text{, se } x \geq 4. \end{cases}$$



(e) $P(X \leq 2) = P(X = 1) + P(X = 2) = 1/4 + 1/4 = 1/2$.

♦♦

Exemplo 2.9

Suponha agora que o homem com as 4 chaves não tenha controle das chaves que já foram experimentadas e por isso a cada nova tentativa ele escolhe, de forma aleatória, uma entre as quatro chaves.

(a) Defina um espaço amostral para esse experimento.

Defina a variável aleatória X = número de chaves experimentadas até conseguir abrir a porta (inclusive a chave correta).

(b) Encontre $Im(X)$.

(c) Encontre a função de probabilidade de X e esboce o seu gráfico.

(d) Qual a probabilidade do homem experimentar no máximo duas chaves para abrir a porta?

Solução:

(a) Representando por C a seleção da chave correta e E a seleção de uma chave errada,

$$\Omega = \{C, EC, EEC, EEEC, EEEEC, EEEEEC, \ldots\}$$

(b) Veja que, diferente do Exemplo 2.8, o homem pode experimentar mais de 4 chaves antes de encontrar a correta. Para esse exemplo temos $Im(X) = \{1, 2, 3, 4, 5, 6, \ldots\} = \mathbb{N}$. Veja que esse é um exemplo de variável aleatória discreta com imagem infinita. Podemos dizer que X é discreta uma vez que $Im(X)$, apesar de infinito, é um conjunto enumerável.

(c) Como a $Im(X)$ é um conjunto infinito não vamos encontrar o valor de $p_X(x)$ para cada $x \in Im(X)$. Vamos tentar encontrar uma expressão geral que represente p_X. Para isso vamos começar calculando p_X em alguns pontos.

Veja que não importa quantas chaves já foram testadas em cada nova tentativa a probabilidade do homem pegar a chave certa é 1/4 e de pegar uma chave errada é de 3/4.

$\mathrm{P}(X = 1) = \mathrm{P}(C) = 1/4$.

$\mathrm{P}(X = 2) = \mathrm{P}(EC) = 3/4 \times 1/4 = 3/4^2$.

$\mathrm{P}(X = 3) = \mathrm{P}(EEC) = 3/4 \times 3/4 \times 1/4 = 3^2/4^3$.

$\mathrm{P}(X = 4) = \mathrm{P}(EEEC) = 3/4 \times 3/4 \times 3/4 \times 1/4 = 3^3/4^4$.

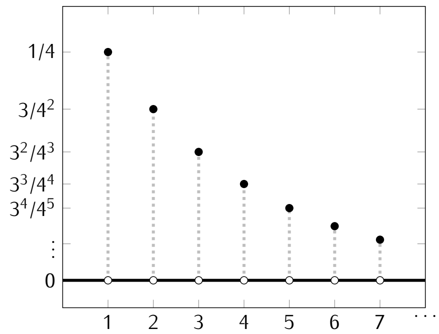
$\mathrm{P}(X = 5) = \mathrm{P}(EEEEC) = 3/4 \times 3/4 \times 3/4 \times 3/4 \times 1/4 = 3^4/4^5$.

De forma geral,

$\mathrm{P}(X = x) = 3/4 \times 3/4 \times 3/4 \ldots \times 3/4 \times 1/4 = 3^{x-1}/4^x$.

Logo, a função de probabilidade de X e o seu gráfico são definidos por:

$$p_X(x) = \begin{cases} \frac{3^{x-1}}{4^x} & \text{, se } x = 1, 2, 3, 4, 5, \ldots \\ 0 & \text{, caso contrário.} \end{cases}$$



No caso de $Im(X)$ ser um conjunto infinito não temos como representar p_X através de uma tabela.

(d) $\mathrm{P}(X \leq 2) = \mathrm{P}(X = 1) + \mathrm{P}(X = 2) = 1/4 + 3/4^2 = 7/16$.

♦♦

Exemplo 2.10
Verifique as propriedades de p_X para a variável aleatória X definida no Exemplo 2.9.

Solução:
No Exemplo 2.9 encontramos

$$p_X(x) = \begin{cases} \frac{3^{x-1}}{4^x} & \text{, se } x = 1, 2, 3, 4, 5, \ldots \\ 0 & \text{, caso contrário.} \end{cases}$$

Queremos mostrar que $p_X(x) \geq 0$ para todo $x \in \mathbb{R}$ e que $\sum_{x \in Im(X)} p_X(x) = 1$.

Primeiro veja que como $3^{x-1}/4^x \geq 0$ sempre que $x \in \mathbb{N}$, logo a primeira propriedade foi verificada: $p_X(x) \geq 0$ para todo $x \in \mathbb{R}$.

Para mostrar que $\sum_{x \in Im(X)} p_X(x) = 1$ vamos precisar de alguns resultados de séries.

$$\begin{aligned}\sum_{x \in Im(X)} p_X(x) &= \sum_{x=1}^{\infty} \frac{3^{x-1}}{4^x} = \sum_{x=1}^{\infty} \frac{1}{3} \left(\frac{3}{4}\right)^x = \frac{1}{3} \sum_{x=1}^{\infty} \left(\frac{3}{4}\right)^x = \frac{1}{3} \left(\sum_{x=0}^{\infty} \left(\frac{3}{4}\right)^x - 1 \right) = \\ &= \frac{1}{3} \left(\frac{1}{1 - 3/4} - 1 \right) = \frac{1}{3}(4-1) = 1.\end{aligned}$$

Veja na Equação A.4 do Apêndice A que $\sum_{i=0}^{\infty} r^i = 1/(1-r)$ sempre que $|r| < 1$.

♦♦

Exemplo 2.11

Suponha que um dado seja lançado quatro vezes. Seja X = o número de lançamentos cujo resultado foi 6.

(a) Defina um espaço amostral para o experimento e diga quantos elementos ele tem.

(b) Apresente os valores de X em alguns elementos de Ω e depois encontre $Im(X)$.

(c) Encontre a função de probabilidade de X.

(d) Verifique as propriedades de p_X.

Solução:

(a) Nesse caso será conveniente considerar que a ordem com que os dados foram lançados importa. Podemos definir o espaço amostral por:

$$\Omega = \left\{ \begin{array}{c} (1,1,1,1), (1,1,1,2), \ldots (1,1,1,6), (1,1,2,1), (1,1,2,2), \ldots (1,1,2,6), \\ \vdots \\ (5,6,3,1), (5,6,3,2), \ldots (5,6,3,6), (5,6,4,1), (5,6,4,2), \ldots (5,6,4,6), \\ \vdots \\ (6,6,5,1), (6,6,5,2), \ldots (6,6,5,6).(6,6,6,1), (6,6,6,2), \ldots (6,6,6,6). \end{array} \right\}$$

Veja que o número de elementos de Ω é $6 \times 6 \times 6 \times 6 = 1296$.

(b) Veja que $X((1,1,1,1) = 0$, $X(1,2,6,4) = 1$, $X(5,6,1,6) = 2$, $X(6,6,4,6) = 3$ e $X(6,6,6,6) = 4$.
$Im(X) = \{0,1,2,3,4\}$.

(c) Queremos encontrar $P(X = x)$ para todo $x \in Im(X)$. Primeiro veja que o espaço amostral é equiprovável, isto é, cada elemento $\omega \in \Omega$ tem mesma probabilidade de ocorrer uma vez que estamos supondo o dado justo. Isto é, $P(1,1,1,1) = P(6,6,6,6) = P(1,2,3,4) = 1/1296$. Ou melhor, $P(\omega) = 1/1296$ para qualquer $\omega \in \Omega$.

Vamos pensar em cada valor de x separadamente.

Primeiro veja que $P(X = 0) =$ P(nenhuma face 6 nos 4 lançamentos). Sabemos que em cada lançamento do dado a probabilidade de sair o 6 é 1/6 e a probabilidade de sair outra face qualquer é 5/6. Então, $P(X = 0) = \frac{5}{6} \times \frac{5}{6} \times \frac{5}{6} \times \frac{5}{6} = \frac{625}{1296}$. Outra maneira de pensar seria definir o evento {nenhuma face 6 nos 4 lançamentos} e verificar que neste evento existem $5 \times 5 \times 5 \times 5 = 652$ elementos.

Se quisermos calcular $P(X = 4) =$ P(todas as faces 6 nos 4 lançamentos) $= \frac{1}{6} \times \frac{1}{6} \times \frac{1}{6} \times \frac{1}{6} = \frac{1}{1296}$. Pensando em eventos no espaço amostral, só existem um elemento no evento {todas as faces 6 nos 4 lançamentos}, que é $(6,6,6,6)$.

Já $P(X = 1) = P(1$ face 6 e 3 faces diferentes de 6 nos 4 lançamentos). A probabilidade do primeiro lançamento ser 6 e os outros não é $\frac{1}{6}\times\frac{5}{6}\times\frac{5}{6}\times\frac{5}{6} = \frac{125}{1296}$. Mas também podemos ter o segundo lançamento igual a 6 e os outros não, a probabilidade disso ocorrer é $\frac{5}{6}\times\frac{1}{6}\times\frac{5}{6}\times\frac{5}{6} = \frac{125}{1296}$. Na verdade, não importa em qual dos quatro lançamentos ocorreu a face 6, o importante é que ela tenha ocorrido uma única vez. Então $P(X = 1) = 4\times\frac{1}{6}\times\frac{5}{6}\times\frac{5}{6}\times\frac{5}{6} = \frac{500}{1296}$, pois temos 4 possibilidades para o lançamento em que ocorreu a face 6.

Agora vejamos $P(X = 2) = P(2$ faces 6 e 2 faces diferentes de 6 nos 4 lançamentos). Nesse caso estamos pensando no evento em que dois lançamentos resultaram em 6 e os outros dois em faces diferentes de 6. A probabilidade dos dois primeiros lançamentos ser 6 e os outros dois serem diferentes de 6 é $\frac{1}{6}\times\frac{1}{6}\times\frac{5}{6}\times\frac{5}{6} = \frac{25}{1296}$. Mas podemos escolher de $\binom{4}{2} = 6$ maneiras diferentes quais os dois lançamentos que resultaram em 6 e quais os dois que resultaram em uma face diferente de 6. Logo, $P(X = 2) = 6\times\frac{1}{6}\times\frac{1}{6}\times\frac{5}{6}\times\frac{5}{6} = \frac{150}{1296}$.

Seguindo a mesma linha de raciocínio concluímos que $P(X = 3) = P(3$ faces 6 e 1 faces diferentes de 6 nos 4 lançamentos) $= 4\times\frac{1}{6}\times\frac{1}{6}\times\frac{1}{6}\times\frac{5}{6} = \frac{20}{1296}$.

Então,

x	0	1	2	3	4
$p_X(x)$	625/1296	500/1296	150/1296	20/1296	1/1296

(d) Veja que $p_X(x) \geq 0$ para todo $x \in \mathbb{R}$ e que

$$\sum_{x=0}^{4} p_X(x) = \frac{625}{1296} + \frac{500}{1296} + \frac{150}{1296} + \frac{20}{1296} + \frac{1}{1296} = 1.$$

♦♦

2.2 Funções de Variáveis Aleatórias Discretas

Dada uma variável aleatória discreta X, podemos obter outras variáveis aleatórias discretas através de funções de X e, da mesma forma que calculamos a função de probabilidade de X, podemos calcular a função de probabilidade dessas novas variáveis.

O importante é ficar claro que se X é variável aleatória então qualquer função real de X também será variável aleatória. Isto é, se X é variável aleatória então $Y = g(X)$ também é variável aleatória, com g função real. Veja que Y continua sendo uma função do espaço amostral na reta, definida pela composição da função g com a função X:

$$\begin{array}{rccc} X: & \Omega & \longrightarrow & \mathbb{R} \\ & \omega & \mapsto & X(\omega) \end{array} \quad \text{e} \quad \begin{array}{rccccc} Y = g(X(\cdot)): & \Omega & \longrightarrow & \mathbb{R} & \longrightarrow & \mathbb{R} \\ & \omega & \mapsto & X(\omega) & \mapsto & g(X(\omega)). \end{array}$$

Exemplo 2.12

Considere a variável aleatória X cuja função de probabilidade é dada na tabela abaixo:

x	-2	-1	0	1	2	3
$p_X(x)$	0,1	0,2	0,2	0,3	0,1	0,1

Defina $Y = g(X) = X^2$.

(a) *Encontre $Im(Y)$.*

(b) Encontre a função de probabilidade de Y, isto é, p_Y.

Solução:

(a) Como $Y = X^2$ podemos aproveitar os valores de $Im(X)$ para definir o conjunto $Im(Y)$. Temos $Im(X) = \{-2, -1, 0, 1, 2, 3\}$, sempre que X assumir o valor $x \in Im(X)$ a v.a. Y irá assumir o valor x^2. Assim, $Im(Y) = \{4, 1, 0, 9\}$.

(b) Para calcular as probabilidades desses valores, temos que identificar os valores de X que originaram cada um deles.

$$\begin{aligned} \mathrm{P}(Y=0) &= \mathrm{P}(X=0) = 0,2 \\ \mathrm{P}(Y=1) &= \mathrm{P}(X=-1) + \mathrm{P}(X=1) = 0,5 \\ \mathrm{P}(Y=4) &= \mathrm{P}(X=-2) + \mathrm{P}(X=2) = 0,2 \\ \mathrm{P}(Y=9) &= \mathrm{P}(X=3) = 0,1 \end{aligned}$$

Podemos resumir essa função de probabilidade como

y	0	1	4	9
$p_Y(y)$	0,2	0,5	0,2	0,1

♦♦

Proposição 2.13 Função de variável aleatória discreta
Seja X uma variável aleatória discreta com função de probabilidade p_X. Seja $Y = g(X)$, onde g é uma função real qualquer. Então,

(i) Y é variável aleatória discreta.

(ii) A função de probabilidade de Y é calculada como

$$p_Y(y) = \mathrm{P}(Y = y) = \sum_{\{x \mid g(x)=y\}} p_X(x)$$

Demonstração:

(i) Para mostrar que Y é variável aleatória discreta basta mostrar que $Im(Y)$ é um conjunto finito ou enumerável. Como X é v.a. discreta sabemos que $Im(X)$ é um conjunto finito ou enumerável. Como $Im(Y) = g(Im(X)) = \{y \in \mathbb{R} \mid \exists\ x \in Im(X) \text{ com } y = g(x)\}$ podemos afirmar que a cardinalidade de $Im(X)$ é sempre maior ou igual à cardinalidade de $Im(Y)$, pois cada elemento de $Im(Y)$ é associado a um ou mais elementos em $Im(X)$. Então, como $Im(X)$ é um conjunto finito ou enumerável, $Im(Y)$ só pode ser um conjunto finito ou enumerável. Com isso concluímos que Y é v.a. discreta.

(ii) $p_Y(y) = \mathrm{P}(Y = y) = \mathrm{P}(g(X) = y) = \sum_{\{x \mid g(x)=y\}} p_X(x)$.

□

Exemplo 2.14
No Exemplo 2.11 definimos X = número de faces 6 em 4 lançamentos de um dado. Defina agora Y = número de faces diferentes de 6 em 4 lançamentos de um dado.

(a) Escreva Y como função de X.

(b) Defina $Im(Y)$ a partir de $Im(X)$.

(c) Encontre a função de probabilidade de Y.

Solução:

(a) Veja que podemos definir $Y = 4-X$, pois em quatro lançamentos o número de resultados diferentes de 6 é o total de lançamentos, 4, menos o número de resultados iguais a 6.

(b) Vimos no Exemplo 2.11 que $Im(X) = \{0, 1, 2, 3, 4\}$. Veja que quando $X = x$ temos $Y = 4 - x$, logo, $Im(Y) = \{4, 3, 2, 1, 0\}$.

(c) Também no Exemplo 2.11 encontramos a função de probabilidade de X, definida na tabela abaixo.

x	0	1	2	3	4
$p_X(x)$	625/1296	500/1296	150/1296	20/1296	1/1296

Veja que $P(Y = y) = P(4 - X = y) = P(X = 4 - y)$. Então, $P(Y = 0) = P(X = 4)$, $P(Y = 1) = P(X = 3)$, $P(Y = 2) = P(X = 2)$, $P(Y = 3) = P(X = 1)$ e $P(Y = 4) = P(X = 0)$. Logo,

y	0	1	2	3	4
$p_Y(y)$	1/1296	20/1296	150/1296	500/1296	625/1296

♦♦

Exemplo 2.15
No Exemplo 2.9 definimos X = número de chaves testadas até abrir a porta, considerando que a cada nova tentativa qualquer uma das quatro chaves podem ser selecionadas com mesma probabilidade. Considerando as mesmas condições do experimento, defina agora Y = número de chaves erradas até abrir a porta.

(a) Escreva Y como função de X.

(b) Defina $Im(Y)$ a partir de $Im(X)$.

(c) Encontre a função de probabilidade de Y.

(d) Verifique as propriedades da função de probabilidade em p_Y.

Solução:

(a) Veja que $Y = X - 1$, pois sempre a última chave testada será a chave correta e todas as anteriores serão erradas.

(b) No Exemplo 2.9 encontramos $Im(X) = \{1, 2, 3, 4, \ldots\}$. Então, $Im(Y) = \{0, 1, 2, 3, \ldots\}$.

(c) No Exemplo 2.9 encontramos

$$p_X(x) = \begin{cases} \frac{3^{x-1}}{4^x} & \text{, se } x = 1, 2, 3, 4, \ldots \\ 0 & \text{, caso contrário.} \end{cases}$$

Veja que $p_Y(y) = P(Y = y) = P(X - 1 = y) = P(X = y + 1) = p_X(y + 1)$. Então podemos definir:

$$p_Y(y) = p_X(y+1) = \begin{cases} \frac{3^{y+1-1}}{4^{y+1}} & \text{, se } y + 1 = 1, 2, 3, 4, \ldots \\ 0 & \text{, caso contrário.} \end{cases} = \begin{cases} \frac{3^{y}}{4^{y+1}} & \text{, se } y = 0, 1, 2, 3, \ldots \\ 0 & \text{, caso contrário.} \end{cases}$$

(d) Claramente $p_Y(y) \geq 0$ para todo $y \in \mathbb{R}$. Basta então verificar que $\sum_{y \in Im(Y)} p_Y(y) = 1$.

$$\sum_{y \in Im(Y)} p_Y(y) = \sum_{y=0}^{\infty} \frac{3^y}{4^{y+1}} = \sum_{y=0}^{\infty} \frac{1}{4}\left(\frac{3}{4}\right)^y = \frac{1}{4}\sum_{y=0}^{\infty}\left(\frac{3}{4}\right)^y = \frac{1}{4}\left(\frac{1}{1-3/4}\right) = \frac{1}{4}4 = 1.$$

Novamente usamos o resultado da Equação A.4 do Apêndice A, que nos mostra que $\sum_{i=0}^{\infty} r^i = 1/(1-r)$ sempre que $|r| < 1$.

♦♦

2.3 Esperança de Variáveis Aleatórias Discretas

No estudo de variáveis aleatórias e suas distribuições de probabilidades, associamos números aos pontos do espaço amostral, ou seja, o resultado é sempre uma variável quantitativa (note que os resultados cara e coroa não definem uma variável aleatória; para tal, temos que associar números, 0 e 1, por exemplo, a esses resultados). Sendo assim, faz sentido perguntar "qual é o valor médio da variável aleatória X?"

A ideia de valor médio diz respeito à média dos valores da variável aleatória se o experimento fosse realizado infinitas vezes. Suponha que X seja variável aleatória discreta tal que $Im(X) = \{x_1, x_2, x_3, \ldots\}$. Suponha também que ao realizar o experimento n vezes a variável X assuma os seguintes valores na seguinte ordem: x_2, x_4, x_3, x_4, x_2, x_1, x_2, ... Então o valor médio de X nas n realizações do experimento é:

$$\frac{x_2 + x_4 + x_3 + x_4 + x_2 + x_1 + x_2 \ldots}{n} = \frac{n_1 x_1 + n_2 x_2 + n_3 x_3 + n_4 x_4 \ldots}{n},$$

onde n_i representa o número de vezes que a variável X assumiu o valor x_i.

Chamamos de esperança o limite do valor médio de X quando $n \to \infty$. Veja que quando $n \to \infty$ a frequência de vezes que X assume o valor x_i, definida por $\frac{n_i}{n}$, converge para $\mathrm{P}(X = x_i)$. Então podemos escrever:

$$\lim_{n \to \infty} \frac{n_1 x_1 + n_2 x_2 + n_3 x_3 + n_4 x_4 \ldots}{n} = x_1 \mathrm{P}(X = x_1) + x_2 \mathrm{P}(X = x_2) + x_3 \mathrm{P}(X = x_3) + \ldots$$

Definição 2.16 Esperança de Variável Aleatória Discreta
*Seja X uma variável aleatória discreta. A **esperança** ou **média** da variável aleatória X é definida como*

$$\mathrm{E}(X) = \sum_{x \in Im(X)} x p_X(x) = \sum_{x \in Im(X)} x \mathrm{P}(X = x),$$

desde que o somatório convirja.

Podemos ver, então, que a esperança de X é uma média dos seus valores, ponderada pelas respectivas probabilidades.

Veja que quando $Im(X)$ é um conjunto finito, $\mathrm{E}(X)$ é definida por um somatório com finitas parcelas e, nesse caso, o somatório sempre converge. Logo $\mathrm{E}(X)$ existe e $\mathrm{E}(X) \in \mathbb{R}$. Já quando $Im(X)$ é um conjunto infinito e enumerável o somatório que define $\mathrm{E}(X)$ tem infinitas parcelas, ou seja, trata-se de uma série. Esta série pode convergir para um número real ou não. Sempre que essa série convergir para um número real dizemos que $\mathrm{E}(X)$ existe e $\mathrm{E}(X) \in \mathbb{R}$. Caso contrário, dizemos que $\mathrm{E}(X)$ não existe. Veja então que podem existir variáveis aleatórias que não tem esperança.

Um resultado simples e imediato é a esperança de uma constante, ou seja, a esperança de uma variável aleatória discreta que pode assumir um único valor com probabilidade 1, também chamada de **variável aleatória degenerada**.

Proposição 2.17
Seja X uma variável aleatória tal que $\mathrm{P}(X = c) = 1$. Então $\mathrm{E}(X) = c$, ou seja, $\mathrm{E}(c) = c$.

Demonstração:
Como $\mathrm{P}(X = c) = 1$, temos $Im(X) = \{c\}$. Então, $\mathrm{E}(X) = c\,\mathrm{P}(X = c) = c$.

□

Exemplo 2.18 Vendas e comissões
Em determinado setor de uma loja de departamentos, o número de produtos vendidos em um dia pelos funcionários é uma variável aleatória P com a seguinte distribuição de probabilidades (esses números foram obtidos dos resultados de vários anos de estudo):

Número de produtos	*0*	*1*	*2*	*3*	*4*	*5*	*6*
Probabilidade de venda	*0,1*	*0,4*	*0,2*	*0,1*	*0,1*	*0,05*	*0,05*

Cada vendedor recebe comissões de venda, distribuídas da seguinte forma: se ele vende até dois produtos em um dia, ele ganha uma comissão de R$10,00 por produto vendido; a partir da terceira venda, a comissão passa para R$50,00 por produto.

Qual é o número médio de produtos vendidos por cada vendedor em um dia de trabalho e qual a comissão diária média de cada um deles?

Solução:
O número médio de produtos vendidos em um dia por um funcionário é $\mathrm{E}(P)$. Veja que $Im(P) = \{0, 1, 2, 3, 4, 5, 6\}$, então

$$\mathrm{E}(P) = 0 \times 0,1 + 1 \times 0,4 + 2 \times 0,2 + 3 \times 0,1 + 4 \times 0,1 + 5 \times 0,05 + 6 \times 0,05 = 2,05$$

ou seja, o número médio de produtos vendidos em um dia por cada funcionário é 2,05 unidades.

Para encontrar a comissão média precisamos primeiro definir a variável aleatória C = comissão diária do funcionário e depois calcular $\mathrm{E}(C)$. Veja que C pode ser definido como função de P. A relação entre C e P, assim como as probabilidades associadas a cada valor dessas duas variáveis aleatórias, estão apresentadas na tabela a seguir.

Número de produtos P	0	1	2	3	4	5	6
Comissão C	0	10	20	70	120	170	220
Probabilidade de venda	0,1	0,4	0,2	0,1	0,1	0,05	0,05

Dessa forma vimos que $Im(C) = \{0, 10, 20, 70, 120, 170, 220\}$ e também conhecemos a probabilidade de C assumir cada um desses valores. Podemos então calcular $\mathrm{E}(C)$:

$$\mathrm{E}(C) = 0 \times 0,1 + 10 \times 0,4 + 20 \times 0,2 + 70 \times 0,1 + 120 \times 0,1 + 170 \times 0,05 + 220 \times 0,05 = 46,5$$

ou seja, a comissão média diária de cada vendedor é de R$ 46,50.

♦♦

Note que a esperança de X tem a mesma unidade de medida dos valores de X. Para o Exemplo 2.18 acima a unidade de medida da variável aleatória P é unidades de produtos, logo $\mathrm{E}(P)$ também terá essa mesma unidade de medida. O mesmo vale para a variável aleatória C do mesmo exemplo, sua unidade de medida é reais, assim como a unidade de medida de $\mathrm{E}(C)$.

Exemplo 2.19
Em um jogo de dados Cláudio paga R\$20,00 para Lúcio e lança 3 dados. Se sair a face 1 em um dos dados, Cláudio ganha R\$20,00 e sai do jogo sem prejuízo ou lucro. Se sair a face 1 em dois dos dados, Cláudio ganha R\$ 50,00. Se sair a face 1 nos três dados, Cláudio ganha R\$80,00. Se não sair a face 1, Cláudio não ganha nada. Calcule o ganho médio de Cláudio no jogo. Você acha que vale a pena participar desse jogo?

Solução:
Para calcularmos o ganho médio de Cláudio vamos definir a variável aleatória X = ganho do Cláudio e depois encontrar $\mathrm{E}(X)$. Veja que $Im(X) = \{-20, 0, 30, 60\}$. Precisamos agora encontrar p_X, isto é, $\mathrm{P}(X = x)$ para todo $x \in Im(X)$.

$$\begin{aligned}
p_X(-20) &= \mathrm{P}(X = -20) = \mathrm{P}(\text{não sair a face 1}) = 5/6 \times 5/6 \times 5/6 = 125/216. \\
p_X(0) &= \mathrm{P}(X = 0) = \mathrm{P}(\text{sair a face 1 em um dado}) = 3 \times 1/6 \times 5/6 \times 5/6 = 75/216. \\
p_X(30) &= \mathrm{P}(X = 30) = \mathrm{P}(\text{sair a face 1 em dois dados}) = 3 \times 1/6 \times 1/6 \times 5/6 = 15/216. \\
p_X(60) &= \mathrm{P}(X = 60) = \mathrm{P}(\text{sair a face 1 nos três dados}) = 1/6 \times 1/6 \times 1/6 = 1/216.
\end{aligned}$$

Veja que $\sum_{\forall x \in Im(X)} p_X(x) = 1$.

Agora já podemos calcular $\mathrm{E}(X)$,

$$\mathrm{E}(X) = -20 \times \frac{125}{216} + 0 \times \frac{75}{216} + 30 \times \frac{15}{216} + 60 \times \frac{1}{216} = -\frac{1990}{216} \approx -9,12.$$

Ou seja, em média Cláudio perde R\$ 9,12 com o jogo. Isso significa que se Cláudio tivesse bastante dinheiro para jogar muitas partidas seguidas desse jogo, ao final ele teria um prejuízo em torno de R\$9,12. Não vale a pena participar do jogo.

♦♦

Exemplo 2.20
Voltando ao Exemplo 2.11, considere X = o número de lançamentos cujo resultado foi 6 em quatro lançamentos de um dado. Calcule a esperança de X, ou seja, o número médio de faces iguais a 6 em quatro lançamentos de um dado.

Solução:
No Exemplo 2.11 chegamos no resultado:

x	0	1	2	3	4
$p_X(x)$	625/1296	500/1296	150/1296	20/1296	1/1296

Então temos

$$\mathrm{E}(X) = 0 \times \frac{625}{1296} + 1 \times \frac{500}{1296} + 2 \times \frac{150}{1296} + 3 \times \frac{20}{1296} + 4 \times \frac{1}{1296} = \frac{861}{1296} \approx 0,6636.$$

Ou seja, o número médio de faces iguais a 6 em quatro lançamentos de um dado é 0,6636 face, menor que 1 face.

♦♦

Exemplo 2.21
Calcule $\mathrm{E}(X)$ *para* X = *número de chaves testadas até conseguir abrir a porta, definida nos Exemplos 2.8 e 2.9.*

Solução:
Vamos começar com X definida no Exemplo 2.8, onde X = número de chaves testadas até abrir a porta, considerando que a cada nova tentativa as chaves já testadas são descartadas. Para esse caso encontramos $Im(X) = \{1, 2, 3, 4\}$ e $p_X(x) = 1/4$ para todo $x \in Im(X)$. Então,

$$\mathrm{E}(X) = 1 \times \frac{1}{4} + 2 \times \frac{1}{4} + 3 \times \frac{1}{4} + 4 \times \frac{1}{4} = \frac{10}{4} = 2,5.$$

Ou seja, se as chaves já testadas forem descartadas nas próximas tentativas, em média o homem precisa de 2,5 tentativas para abrir a porta.

Vejamos agora o caso do Exemplo 2.9, onde X = número de chaves testadas até abrir a porta, considerando que a cada nova tentativa qualquer uma das quatro chaves podem ser selecionadas com mesma probabilidade. Para esse caso encontramos $Im(X) = \{1, 2, 3, 4, 5, 6, \ldots\}$ e $p_X(x) = \frac{3^{x-1}}{4^x}$ para todo $x \in Im(X)$. Então,

$$\mathrm{E}(X) = \sum_{\forall x \in Im(X)} x p_X(x) = \sum_{x=1}^{\infty} x \frac{3^{x-1}}{4^x} = \sum_{x=1}^{\infty} (x + 1 - 1) \frac{3^{x-1}}{4^x} = \underbrace{\sum_{x=1}^{\infty} \frac{3^{x-1}}{4^x}}_{S_1} + \underbrace{\sum_{x=1}^{\infty} (x-1) \frac{3^{x-1}}{4^x}}_{S_2}.$$

Primeiro veja que $S_1 = 1$, pois $S_1 = \sum_{\forall x \in Im(X)} p_X(x)$ e pelas propriedades da função de probabilidade sabemos que essa soma tem que ser 1.

Vamos agora calcular S_2. Para isso faremos a seguinte troca de variável: $y = x - 1$.

$$\sum_{x=1}^{\infty} (x-1) \frac{3^{x-1}}{4^x} = \sum_{y=0}^{\infty} y \frac{3^y}{4^{y+1}} = \left(0 \frac{3^0}{4^1}\right) + \sum_{y=1}^{\infty} y \frac{3^y}{4^{y+1}} = \sum_{y=1}^{\infty} y \frac{3^y}{4^{y+1}} = \frac{3}{4} \sum_{y=1}^{\infty} y \frac{3^{y-1}}{4^y} = \frac{3}{4} \mathrm{E}(X).$$

Assim chegamos ao resultado:

$$\mathrm{E}(X) = S_1 + S_2 = 1 + \frac{3}{4} \mathrm{E}(X) \Rightarrow \mathrm{E}(X) - \frac{3}{4} \mathrm{E}(X) = 1 \Rightarrow \frac{1}{4} \mathrm{E}(X) = 1 \Rightarrow \mathrm{E}(X) = 4.$$

Ou seja, se as chaves já testadas não forem descartadas nas próximas tentativas, em média o homem precisa de 4 tentativas para abrir a porta.

♦♦

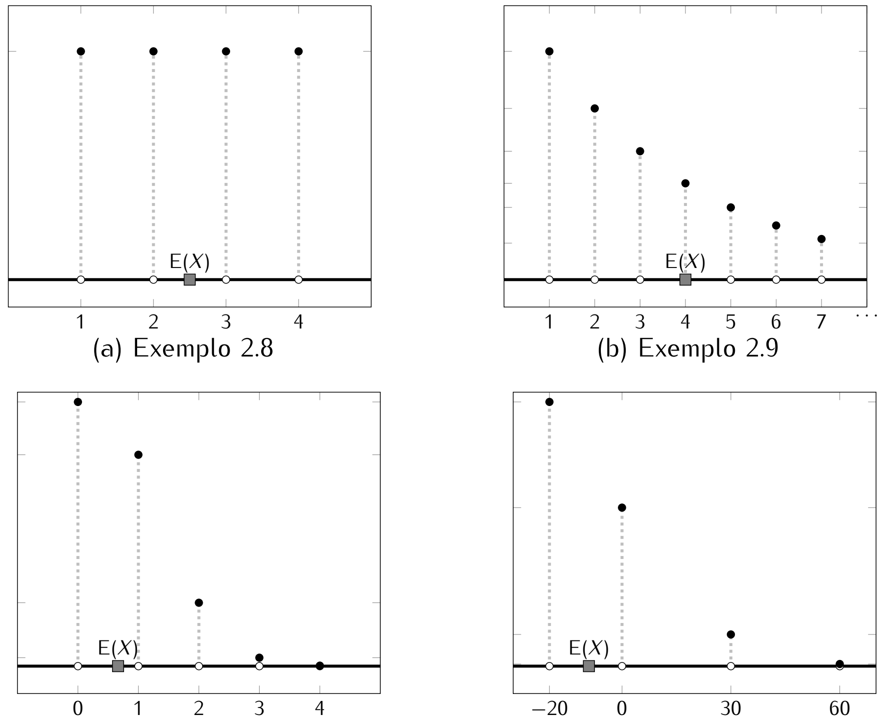
Para as variáveis aleatórias definidas nos Exemplos 2.8, 2.9, 2.11 e 2.19 veja na Figura 2.1 os gráficos das funções de probabilidade junto com a esperança, já calculada para cada variável aleatória.

Observando os gráficos da Figura 2.1 podemos ver que a esperança de uma variável aleatória X é o centro de gravidade da distribuição de probabilidades. Sendo assim, a esperança é uma medida de *posição*. Com isso temos o resultado intuitivo de que $\mathrm{E}(X)$ está sempre entre o menor e o maior valor do conjunto $Im(X)$, resultado apresentado na Proposição 2.22 a seguir.

Proposição 2.22
Seja X variável aleatória discreta, $x_{min} = \min\{Im(X)\}$ e $x_{max} = \max\{Im(X)\}$. Se $\mathrm{E}(X) \in \mathbb{R}$ então, $x_{min} \leq \mathrm{E}(X) \leq x_{max}$.

Demonstração:
Primeiro vamos mostrar que $x_{min} \leq \mathrm{E}(X)$.

$$\mathrm{E}(X) = \sum_{\forall x \in Im(X)} x p_X(x) \geq \sum_{\forall x \in Im(X)} x_{min} p_X(x) = x_{min} \underbrace{\sum_{\forall x \in Im(X)} p_X(x)}_{1} = x_{min}.$$



(a) Exemplo 2.8 (b) Exemplo 2.9 (c) Exemplo 2.11 (d) Exemplo 2.19

Figura 2.1 – Gráfico de algumas funções de probabilidade junto com a Esperança.

Para mostrar que $\mathrm{E}(X) \leq x_{max}$ faremos um desenvolvimento análogo.

$$\mathrm{E}(X) = \sum_{\forall x \in Im(X)} x p_X(x) \leq \sum_{\forall x \in Im(X)} x_{max} p_X(x) = x_{max} \underbrace{\sum_{\forall x \in Im(X)} p_X(x)}_{1} = x_{max}.$$

□

Já vimos que é possível obter uma nova variável aleatória Y a partir de função de uma variável X, $Y = g(X)$. Também vimos que através da função de probabilidade de X podemos obter a função de probabilidade de Y. Sendo assim, podemos calcular a esperança de Y uma vez que conhecemos a sua função de probabilidade.

O interessante que será apresentado na Proposição 2.23 a seguir é que podemos encontrar a esperança de Y sem precisar antes encontrar a sua função de probabilidade. O cálculo da esperança de Y pode ser feito somente com o conhecimento da função g que relaciona X e Y e da função de probabilidade de X, p_X.

Proposição 2.23 Esperança de funções de variável aleatória discreta
Seja X uma variável aleatória discreta com função de probabilidade p_X e $Y = g(X)$. Então,

$$\mathrm{E}(Y) = \mathrm{E}(g(X)) = \sum_{\forall x \in Im(X)} g(x) p_X(x),$$

desde que o somatório convirja.

Demonstração:

$$\begin{aligned}
\mathrm{E}(Y) &= \sum_{\forall y \in Im(Y)} y p_Y(y) = \sum_{\forall y \in Im(Y)} y \, \mathrm{P}(Y = y) = \sum_{\forall y \in Im(Y)} y \, \mathrm{P}(g(X) = y) = \\
&= \sum_{\forall y \in Im(Y)} y \sum_{\{x|g(x)=y\}} \mathrm{P}(X = x) = \sum_{\forall y \in Im(Y)} \sum_{\{x|g(x)=y\}} y \, \mathrm{P}(X = x) = \\
&= \sum_{\forall y \in Im(Y)} \sum_{\{x|g(x)=y\}} g(x) p_X(x) = \sum_{\forall x \in Im(X)} g(x) p_X(x).
\end{aligned}$$

□

Exemplo 2.24
Considere as variáveis aleatórias X e $Y = X^2$ definidas no Exemplo 2.12. Encontre $\mathrm{E}(Y)$ a partir de p_Y e a partir de p_X, usando o resultado da Proposição 2.23.

Solução:
No Exemplo 2.12 foi dado p_X e encontramos p_Y, as duas funções de probabilidade estão apresentadas a seguir.

x	-2	-1	0	1	2	3
$p_X(x)$	0,1	0,2	0,2	0,3	0,1	0,1

y	0	1	4	9
$p_Y(y)$	0,2	0,5	0,2	0,1

Primeiro vamos encontrar $\mathrm{E}(Y)$ a partir de p_Y:

$$\mathrm{E}(Y) = \sum_{\forall y \in Im(Y)} y p_Y(y) = 0 \times 0,2 + 1 \times 0,5 + 4 \times 0,2 + 9 \times 0,1 = 2,2.$$

Agora se quisermos usar o resultado da Proposição 2.23 e calcular $\mathrm{E}(Y)$ usando p_X devemos seguir assim:

$$\begin{aligned}
\mathrm{E}(Y) = \mathrm{E}(X^2) &= \sum_{\forall x \in Im(X)} x^2 p_X(x) = \\
&= (-2)^2 \times 0,1 + (-1)^2 \times 0,2 + 0^2 \times 0,2 + 1^2 \times 0,3 + 2^2 \times 0,1 + 3^2 \times 0,1 \\
&= 2,2.
\end{aligned}$$

♦♦

Exemplo 2.25
Seja X variável aleatória discreta com função de distribuição definida por:

$$F_X(x) = \begin{cases} 0 & , x < -2 \\ 1/8 & , -2 \leq x < 1 \\ 5/8 & , 1 \leq x < 2 \\ 7/8 & , 2 \leq x < 4 \\ 1 & , x \geq 4. \end{cases}$$

Calcule $\mathrm{E}(X)$, $\mathrm{E}(X^2)$, $\mathrm{E}(5 - 2X)$ *e* $\mathrm{E}(|X|)$.

Solução:
Primeiro vamos encontrar $Im(X)$ e p_X. Veja que $Im(X) = \{-2, 1, 2, 4\}$ e que $p_X(x)$ em cada $x \in Im(X)$ está definido na tabela a seguir:

x	-2	1	2	4
$p_X(x)$	1/8	1/2	1/4	1/8

Agora podemos calcular o que se pede.

$$\begin{aligned}
\mathrm{E}(X) &= \sum_{\forall x \in Im(X)} x p_X(x) = (-2) \times \frac{1}{8} + 1 \times \frac{1}{2} + 2 \times \frac{1}{4} + 4 \times \frac{1}{8} = \frac{10}{8} = \frac{5}{4}. \\
\mathrm{E}(X^2) &= \sum_{\forall x \in Im(X)} x^2 p_X(x) = (-2)^2 \times \frac{1}{8} + 1^2 \times \frac{1}{2} + 2^2 \times \frac{1}{4} + 4^2 \times \frac{1}{8} = \frac{32}{8} = 4. \\
\mathrm{E}(5-2X) &= \sum_{\forall x \in Im(X)} (5-2x) p_X(x) \\
&= (5 - 2 \times (-2)) \times \frac{1}{8} + (5 - 2 \times 1) \times \frac{1}{2} + (5 - 2 \times 2) \times \frac{1}{4} + (5 - 2 \times 4) \times \frac{1}{8} \\
&= \frac{20}{8} = \frac{5}{2}. \\
\mathrm{E}(|X|) &= \sum_{\forall x \in Im(X)} |x| p_X(x) = |-2| \times \frac{1}{8} + |1| \times \frac{1}{2} + |2| \times \frac{1}{4} + |4| \times \frac{1}{8} = \frac{14}{8} = \frac{7}{4}.
\end{aligned}$$

♦♦

Proposição 2.26 Propriedade de Linearidade da Esperança
Seja X variável aleatória discreta tal que $\mathrm{E}(X) \in \mathbb{R}$. Sejam $a, b \in \mathbb{R}$. Então,

$$\mathrm{E}(aX + b) = a\,\mathrm{E}(X) + b.$$

Demonstração:

$$\begin{aligned}
\mathrm{E}(aX+b) &= \sum_{x \in Im(X)} (ax+b) p_X(x) = \\
&= \sum_{x \in Im(X)} (ax p_X(x) + b p_X(x)) = \\
&= \sum_{x \in Im(X)} ax p_X(x) + \sum_{x \in Im(X)} b p_X(x) = \\
&= a \underbrace{\sum_{x \in Im(X)} x p_X(x)}_{\mathrm{E}(X)} + b \underbrace{\sum_{x \in Im(X)} p_X(x)}_{1} = a\,\mathrm{E}(X) + b.
\end{aligned}$$

□

Exemplo 2.27
No Exemplo 2.9 definimos X = número de chaves testadas até abrir a porta, considerando que a cada nova tentativa qualquer uma das quatro chaves podem ser selecionadas com mesma probabilidade. No Exemplo 2.21 encontramos $\mathrm{E}(X) = 4$. Considere Y = número de chaves erradas até abrir a porta definida no Exemplo 2.15. Calcule Y de duas maneiras diferentes, primeiro usando p_Y encontrada no Exemplo 2.15 e depois usando o resultado da Proposição 2.26.

Solução:

No Exemplo 2.15 encontramos $p_Y(y) = \begin{cases} \frac{3^y}{4^{y+1}} & \text{, se } y = 0, 1, 2, 3, \ldots \\ 0 & \text{, caso contrário.} \end{cases}$

Segue o cálculo de $\mathrm{E}(Y)$ a partir de p_Y.

$$\begin{aligned}
\mathrm{E}(Y) &= \sum_{\forall y \in Im(Y)} y p_Y(y) = \sum_{y=0}^{\infty} y \frac{3^y}{4^{y+1}} = \sum_{y=0}^{\infty} (y+1-1) \frac{3^y}{4^{y+1}} = \\
&= \underbrace{\sum_{y=0}^{\infty} (y+1) \frac{3^y}{4^{y+1}}}_{S_1} - \underbrace{\sum_{y=0}^{\infty} \frac{3^y}{4^{y+1}}}_{S_2}.
\end{aligned}$$

Veja primeiro que $S_2 = \sum_{y=0}^{\infty} \frac{3^y}{4^{y+1}} = \sum_{\forall y \in Im(Y)} p_Y(y) = 1$. Vamos agora encontrar S_1, para isso faremos a troca de variável $z = y + 1$.

$$S_1 = \sum_{y=0}^{\infty}(y+1)\frac{3^y}{4^{y+1}} = \sum_{z=1}^{\infty} z\frac{3^{z-1}}{4^z} = \frac{4}{3}\sum_{z=1}^{\infty} z\frac{3^z}{4^{z+1}} = \frac{4}{3}\underbrace{\sum_{z=0}^{\infty} z\frac{3^z}{4^{z+1}}}_{\mathrm{E}(Y)} = \frac{4}{3}\,\mathrm{E}(Y).$$

Assim temos,

$$\mathrm{E}(Y) = \frac{4}{3}\,\mathrm{E}(Y) - 1 \Rightarrow 1 = \frac{4}{3}\,\mathrm{E}(Y) - \mathrm{E}(Y) \Rightarrow 1 = \frac{1}{3}\,\mathrm{E}(Y) \Rightarrow \mathrm{E}(Y) = 3.$$

Chegamos na mesma resposta, de forma bem mais simples, usando o resultado da Proposição 2.26. Já encontramos $\mathrm{E}(X) = 4$ no Exemplo 2.21 e temos $Y = X - 1$, veja Exemplo 2.15. Logo $\mathrm{E}(Y) = \mathrm{E}(X-1) = \mathrm{E}(X) - 1 = 4 - 1 = 3$.

♦♦

2.4 Variância e Desvio-Padrão de Variável Aleatória

Já vimos que a esperança de uma variável aleatória X é o centro de gravidade da distribuição de probabilidades, logo uma medida de posição. No entanto, é possível que duas variáveis bem diferentes tenham a mesma esperança, o que nos mostra que a esperança não é uma medida que descreve completamente a variável aleatória.

Exemplo 2.28
Considere X_1 e X_2 variáveis aleatórias discretas com funções de probabilidade p_{X_1} e p_{X_2} definidas por:

$$p_{X_1}(x) = \begin{cases} 0,05 & , x = 1,2,8,9 \\ 0,1 & , x = 3,7 \\ 0,15 & , x = 4,6 \\ 0,3 & , x = 5 \\ 0 & , \text{caso contrário;} \end{cases} \qquad p_{X_2}(x) = \begin{cases} 0,1 & , x = 3,7 \\ 0,3 & , x = 4,6 \\ 0,2 & , x = 5 \\ 0 & , \text{caso contrário.} \end{cases}$$

(a) Mostre que $\mathrm{E}(X_1) = \mathrm{E}(X_2)$.

(b) Esboce e compare os gráficos de p_{X_1} e p_{X_2}. O que você pode perceber de diferente entre essas duas variáveis?

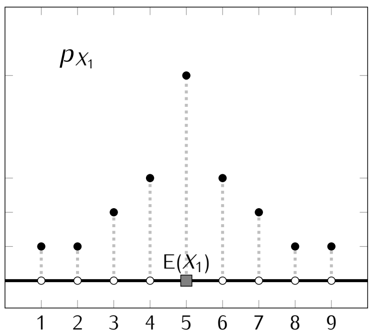
Solução:

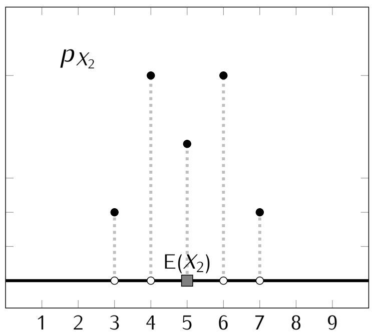
(a) Ambas as esperanças são iguais à 5, basta fazer as contas:

$\mathrm{E}(X_1) = 1 \times 0,05 + 2 \times 0,05 + 3 \times 0,1 + 4 \times 0,15 + 5 \times 0,3 + 6 \times 0,15 + 7 \times 0,1 + 8 \times 0,05 + 9 \times 0,05 = 5$.

$\mathrm{E}(X_2) = 3 \times 0,1 + 4 \times 0,3 + 5 \times 0,2 + 6 \times 0,3 + 7 \times 0,1 = 5$.

(b) A seguir estão os gráficos das funções de probabilidade de X_1 e X_2.





Uma diferença que pode ser notada de imediato é a a dispersão dos valores para cada variável aleatória. Veja que X_1 tem valores mais dispersos em torno da média que X_2, isto é, mais espalhados. Ou seja, os valores de $Im(X_2)$ são mais concentrados em torno da média do que são os valores de $Im(X_1)$. Exitem várias maneiras de medir a dispersão de uma variável aleatória, vamos definir inicialmente uma dessas medidas de dispersão, chamada de variância.

♦♦

Definição 2.29 Variância de Variável Aleatória Discreta
Seja X variável aleatória discreta tal que $\mathrm{E}(X) \in \mathbb{R}$. *A* ***variância*** *de X é definida como*

$$\mathrm{Var}\,(X) = \mathrm{E}\left[(X - \mathrm{E}\,(X))^2\right] = \sum_{\forall x \in Im(X)} (x - \mathrm{E}\,(X))^2\, p_X(x),$$

desde que o somatório convirja.

Veja que $\mathrm{Var}(X)$ é definida como $\mathrm{E}(g(X))$, com $g(X) = (X - \mathrm{E}(X))^2$. O termo $X - \mathrm{E}(X)$ é chamado de *desvio em torno da média*. Sendo assim, a variância é a média dos desvios quadráticos em torno da média $\mathrm{E}(X)$.

Veja também que para existir $\mathrm{Var}(X)$ é necessário que exista $\mathrm{E}(X)$, isto é, que $\mathrm{E}(X) \in \mathbb{R}$. Além disso é necessário que o somatório $\sum_{\forall x \in Im(X)} (x - \mathrm{E}\,(X))^2\, p_X(x)$ convirja.

A Proposição 2.30 a seguir apresenta uma expressão alternativa para o cálculo da variância, que em geral é mais usada do que a própria expressão apresentada na Definição 2.29.

Proposição 2.30
Seja X variável aleatória discreta tal que $\mathrm{E}(X) \in \mathbb{R}$ *e* $\mathrm{E}(X^2) \in \mathbb{R}$. *Então,* $\mathrm{Var}(X) \in \mathbb{R}$ *e*

$$\mathrm{Var}(X) = \mathrm{E}(X^2) - (\mathrm{E}(X))^2$$

Demonstração:

$$
\begin{aligned}
\operatorname{Var}(X) &= \sum_{\forall x \in Im(X)} (x - \mathrm{E}(X))^2 p_X(x) \\
&= \sum_{\forall x \in Im(X)} \left(x^2 - 2x\,\mathrm{E}(X) + (\mathrm{E}(X))^2 \right) p_X(x) = \\
&= \sum_{\forall x \in Im(X)} \left(x^2 p_X(x) - 2x\,\mathrm{E}(X) p_X(x) + (\mathrm{E}(X))^2 p_X(x) \right) = \\
&= \sum_{\forall x \in Im(X)} x^2 p_X(x) - \sum_{\forall x \in Im(X)} 2x\,\mathrm{E}(X) p_X(x) + \sum_{\forall x \in Im(X)} (\mathrm{E}(X))^2 p_X(x) = \\
&= \underbrace{\sum_{\forall x \in Im(X)} x^2 p_X(x)}_{\mathrm{E}(X^2)} - 2\,\mathrm{E}(X) \underbrace{\sum_{\forall x \in Im(X)} x p_X(x)}_{\mathrm{E}(X)} + (\mathrm{E}(X))^2 \underbrace{\sum_{\forall x \in Im(X)} p_X(x)}_{1} = \\
&= \mathrm{E}(X^2) - 2\,\mathrm{E}(X)\,\mathrm{E}(X) + (\mathrm{E}(X))^2 = \\
&= \mathrm{E}(X^2) - (\mathrm{E}(X))^2 .
\end{aligned}
$$

□

O resultado da Proposição 2.30 pode ser lido, de maneira informal, como *a variância é a esperança do quadrado menos o quadrado da esperança*.

Da definição de variância, resulta que sua unidade de medida é o quadrado da unidade de medida da variável em estudo, sendo assim, uma unidade sem significado físico. Para se ter uma medida de dispersão na mesma unidade dos dados, define-se o *desvio-padrão* como a raiz quadrada da variância.

Definição 2.31 Desvio-padrão
O ***desvio-padrão*** *de uma variável aleatória X é definido como a raiz quadrada de sua variância:*

$$\mathrm{DP}(X) = \sqrt{\operatorname{Var}(X)}$$

Exemplo 2.32
Calcule $\operatorname{Var}(X_1)$, $\mathrm{DP}(X_1)$, $\operatorname{Var}(X_2)$ *e* $\mathrm{DP}(X_2)$ *para as variáveis definidas no Exemplo 2.28 e comente os resultados.*

Solução:
Primeiro as contas para a varável X_1:

$$
\begin{aligned}
\mathrm{E}(X_1^2) &= 1^2 \times 0,05 + 2^2 \times 0,05 + 3^2 \times 0,1 + 4^2 \times 0,15 + 5^2 \times 0,3 + \\
& \quad 6^2 \times 0,15 + 7^2 \times 0,1 + 8^2 \times 0,05 + 9^2 \times 0,05 \\
&= 28,6. \\
\operatorname{Var}(X_1) &= \mathrm{E}(X_1^2) - \mathrm{E}(X_1)^2 = 28,6 - 5^2 = 3,6. \\
\mathrm{DP}(X_1) &= \sqrt{\operatorname{Var}(X_1)} = \sqrt{3,6} \approx 1,9.
\end{aligned}
$$

Agora as contas para a varável X_2:

$$
\begin{aligned}
\mathrm{E}(X_2^2) &= 3^2 \times 0,1 + 4^2 \times 0,3 + 5^2 \times 0,2 + 6^2 \times 0,3 + 7^2 \times 0,1 = 26,4. \\
\operatorname{Var}(X_2) &= \mathrm{E}(X_2^2) - \mathrm{E}(X_2)^2 = 26,4 - 5^2 = 1,4. \\
\mathrm{DP}(X_2) &= \sqrt{\operatorname{Var}(X_2)} = \sqrt{1,4} \approx 1,2.
\end{aligned}
$$

Como já comentado no Exemplo 2.28, a dispersão de X_1 é maior que a dispersão de X_2, uma vez que $\operatorname{Var}(X_1) > \operatorname{Var}(X_2)$, ou $\mathrm{DP}(X_1) > \mathrm{DP}(X_2)$.

♦♦

Já vimos que se X é variável aleatória degenerada, isto é, X é tal que $\mathrm{P}(X = c) = 1$, então $\mathrm{E}(X) = c$ (Proposição 2.17). Vejamos agora como é a variância de uma variável aleatória degenerada.

Proposição 2.33
Seja X uma variável aleatória tal que $\mathrm{P}(X = c) = 1$. *Então* $\mathrm{Var}(X) = 0$*, ou seja,* $\mathrm{Var}(c) = 0$.

Demonstração:
Como $\mathrm{P}(X = c) = 1$, temos $\mathrm{E}(X) = c$. Veja que $Y = X^2$ também é variável aleatória degenerada e $\mathrm{P}(Y = c^2) = 1$. Logo, $\mathrm{E}(Y) = \mathrm{E}(X^2) = c^2$. Então, $\mathrm{Var}(X) = \mathrm{E}(X^2) - \mathrm{E}(X)^2 c^2 - c^2 = 0$.

□

Vejamos mais algumas propriedades da variância, agora para qualquer que seja a variável aleatória.

Proposição 2.34 Propriedades da Variância e do Desvio-padrão
Seja X variável aleatória tal que $\mathrm{Var}(X) \in \mathbb{R}$. *Então valem as seguintes propriedades:*

(i) $\mathrm{Var}(X) \geq 0$

(ii) $\mathrm{DP}(X) \geq 0$

(iii) $\mathrm{Var}(aX + b) = a^2 \,\mathrm{Var}(X)$

(iv) $\mathrm{DP}(aX + b) = |a| \,\mathrm{DP}(X)$

Demonstração:

(i) Veja que $\mathrm{Var}(X) = \mathrm{E}(g(X))$, com $g(X) = (X - \mathrm{E}(X))^2$. Veja também que $g(X)$ é variável aleatória não negativa, isto é, $x_{min} = \min\{Im(g(X))\} \geq 0$. Logo, pela propriedade apresentada na Proposição 2.22, podemos afirmar que $\mathrm{E}(g(X)) \geq x_{min} \geq 0$. Então, $\mathrm{Var}(X) \geq 0$.

(ii) Como $\mathrm{Var}(X) \geq 0$, $\mathrm{DP}(X) = \sqrt{\mathrm{Var}(X)} \geq 0$.

(iii)

$$
\begin{aligned}
\mathrm{Var}(aX + b) &= \mathrm{E}\left((aX + b - \mathrm{E}(aX + b))^2\right) \\
&= \mathrm{E}\left(\left(aX + \cancel{b} - a\,\mathrm{E}(X) - \cancel{b}\right)^2\right) \\
&= \mathrm{E}\left((aX - a\,\mathrm{E}(X))^2\right) \\
&= \mathrm{E}\left(a^2 (X - \mathrm{E}(X))^2\right) \\
&= a^2\,\mathrm{E}\left((X - \mathrm{E}(X))^2\right) \\
&= a^2\,\mathrm{Var}(X).
\end{aligned}
$$

(iv) $\mathrm{DP}(aX + b) = \sqrt{\mathrm{Var}(aX + b)} = \sqrt{a^2\,\mathrm{Var}(X)} = |a|\sqrt{\mathrm{Var}(X)} = |a|\,\mathrm{DP}(X)$.

□

Exemplo 2.35
Considere a v.a. Y com função de probabilidade dada por

y	-3	-1	0	2	5	8	9
$p_Y(y)$	$0,25$	$0,30$	$0,20$	$0,10$	$0,07$	$0,05$	$0,03$

e seja $Z = 2Y - 3$. Vamos calcular a esperança e a variância de Y e Z.

Solução:
$E(Y) = -3 \times 0,25 - 1 \times 0,30 + 0 \times 0,20 + 2 \times 0,10 + 5 \times 0,07 + 8 \times 0,05 + 9 \times 0,03 = 0,17.$

$E(Z) = 2 \times E(Y) - 3 = 2 \times 0,17 - 3 = -2,66.$

Vamos calcular agora $E(Y^2)$:

$E(Y^2) = 9\times 0,25+1\times 0,30+0\times 0,20+4\times 0,10+25\times 0,07+64\times 0,05+81\times 0,03 = 10,33.$

Logo, $\text{Var}(Y) = 10,33 - 0,17^2 = 10,3011.$

Usando as propriedades da variância, temos que $\text{Var}(Z) = 2^2 \times \text{Var}(Y) = 41,2044.$

♦♦

Exemplo 2.36
Um lojista mantém extensos registros das vendas diárias de certo aparelho. O quadro a seguir dá a distribuição de probabilidades do número de aparelhos vendidos em uma semana. Se o lucro por unidade vendida é de R$500,00, qual o lucro esperado em uma semana? Qual é o desvio-padrão do lucro?

x= número de aparelhos	*0*	*1*	*2*	*3*	*4*	*5*
$p_X(x)$	*0,1*	*0,1*	*0,2*	*0,3*	*0,2*	*0,1*

Solução:
Seja X o número de aparelhos vendidos. Vamos calcular a média e o desvio padrão de X, isto é, do número de aparelhos vendidos.

$$\begin{aligned} E(X) &= 0 \times 0,1 + 1 \times 0,1 + 2 \times 0,2 + 3 \times 0,3 + 4 \times 0,2 + 5 \times 0,1 \\ &= 2,7 \text{ aparelhos} \end{aligned}$$

$$\begin{aligned} E\left(X^2\right) &= 0^2 \times 0,1 + 1^2 \times 0,1 + 2^2 \times 0,2 + 3^2 \times 0,3 + 4^2 \times 0,2 + 5^2 \times 0,1 \\ &= 10,2 \text{ aparelhos}^2 \end{aligned}$$

$$\text{Var}(X) = 10,2 - (2,7)^2 = 2,91 \text{ aparelhos}^2$$

$$\text{DP}(X) = 1,706 \text{ aparelhos}$$

Seja L o lucro semanal. Veja que $L = 500X$. Como conhecemos $E(X)$ e $\text{DP}(X)$ podemos usar as propriedades da esperança e do desvio padrão para encontrar o que se pede: $E(L)$ e $\text{DP}(L)$.

$$E(L) = 500\, E(X) = R\$1350,00$$

$$\text{DP}(L) = 500\, \text{DP}(X) = R\$852,94$$

♦♦

Exemplo 2.37
Seja X a variável aleatória definida no Exemplo 2.25, isto é, X é tal que

$$F_X(x) = \begin{cases} 0 & , x < -2 \\ 1/8 & , -2 \le x < 1 \\ 5/8 & , 1 \le x < 2 \\ 7/8 & , 2 \le x < 4 \\ 1 & , x \ge 4. \end{cases}$$

No Exemplo 2.25 calculamos $E(X)$, $E(X^2)$, $E(5-2X)$ *e* $E(|X|)$. *Agora calcule* $\text{Var}(X)$, $\text{Var}(X^2)$, $\text{Var}(5-2X)$ *e* $\text{Var}(|X|)$.

Solução:
Os resultados encontrados no Exemplo 2.25 foram: $\mathrm{E}(X) = \frac{5}{4}$, $\mathrm{E}(X^2) = 4$, $\mathrm{E}(5-2X) = \frac{5}{2}$ e $\mathrm{E}(|X|) = \frac{7}{4}$.

Vamos calcular o que se pede.

$$\begin{aligned}
\mathrm{Var}(X) &= \mathrm{E}(X^2) - \mathrm{E}(X)^2 = 4 - \left(\frac{5}{4}\right)^2 = \frac{64-25}{16} = \frac{39}{16}.\\
\mathrm{Var}(X^2) &= \mathrm{E}(X^4) - \mathrm{E}(X^2)^2 = \left(\sum_{\forall x \in Im(X)} x^4 p_X(x)\right) - 4^2\\
&= \left((-2)^4 \times \frac{1}{8} + 1^4 \times \frac{1}{2} + 2^4 \times \frac{1}{4} + 4^4 \times \frac{1}{8}\right) - 16\\
&= \left(\frac{312}{8}\right) - 16 = 39 - 16 = 23.\\
\mathrm{Var}(5-2X) &= 4\,\mathrm{Var}(X) = 4 \times \frac{39}{16} = \frac{39}{4}.\\
\mathrm{Var}(|X|) &= \mathrm{E}(|X|^2) - \mathrm{E}(|X|)^2 = \mathrm{E}(X^2) - \mathrm{E}(|X|)^2 = 4 - \left(\frac{7}{4}\right)^2 = \frac{64-49}{16} = \frac{15}{16}.
\end{aligned}$$

♦♦

Exemplo 2.38
Seja uma v.a. X com função de probabilidade dada na tabela a seguir:

x	1	2	3	4	5
$p_X(x)$	p^2	p^2	p	p	p^2

(a) *Encontre o valor de p para que $p_X(x)$ seja, de fato, uma função de probabilidade.*

(b) *Calcule* $\mathrm{P}(X \geq 4)$ *e* $\mathrm{P}(X < 3)$.

(c) *Calcule* $\mathrm{P}(|X-3| \geq 2)$.

(d) *Calcule* $\mathrm{E}(X)$ *e* $\mathrm{Var}(X)$.

Solução:

(a) Como $\sum_x p_X(x) = 1$, temos que ter:

$$3p^2 + 2p = 1 \Rightarrow 3p^2 + 2p - 1 = 0 \Rightarrow$$
$$p = \frac{-2 \pm \sqrt{4+12}}{6} = \frac{-2 \pm 4}{6} \Rightarrow \begin{cases} p = -1 \\ \text{ou} \\ p = \frac{1}{3} \end{cases}$$

Como p é uma probabilidade, temos que ter $p \geq 0$. Logo, o valor correto é $p = \frac{1}{3}$.

(b) $\mathrm{P}(X \geq 4) = \mathrm{P}(X=4) + \mathrm{P}(X=5) = p + p^2 = \frac{1}{3} + \frac{1}{9} = \frac{4}{9}$.
$Pr(X < 3) = \mathrm{P}(X=1) + \mathrm{P}(X=2) = 2p^2 = \frac{2}{9}$.

(c) Aqui temos que notar o seguinte fato sobre a função módulo, ilustrado na **Figura 2.2**. Valores $y = |x|$ no eixo vertical menores que k (abaixo da linha horizontal sólida) correspondem a valores de x no intervalo $(-k, k)$ e valores y no eixo vertical maiores que k correspondem ou a $x > k$ ou a $x < -k$. Mais precisamente,

$$\begin{array}{lcl} |\,x\,| \geq k & \Leftrightarrow & x \geq k \quad \text{ou} \quad x \leq -k \\ |\,x\,| \leq k & \Leftrightarrow & -k \leq x \leq k \end{array}$$

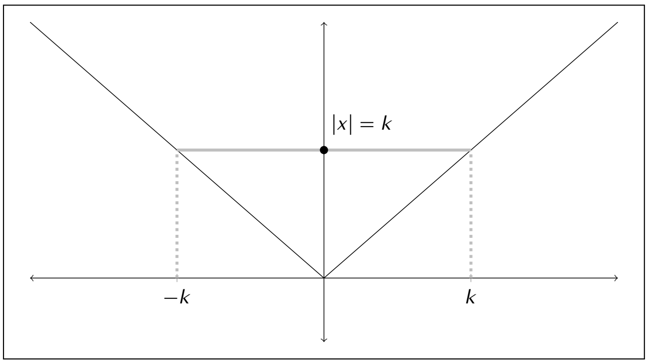


Figura 2.2 – Função módulo

Usando esses fatos, temos que

$$\begin{aligned} \mathrm{P}\left(|X-3| \geq 2\right) &= \mathrm{P}\left(\{X-3 \leq -2\} \;\cup\; \{X-3 \geq 2\}\right) = \\ &= \mathrm{P}\left(X-3 \leq -2\right) + \mathrm{P}\left(X-3 \geq 2\right) = \\ &= \mathrm{P}\left(X \leq 1\right) + \mathrm{P}\left(X \geq 5\right) = \\ &= \mathrm{P}\left(X = 1\right) + \mathrm{P}(X = 5) = \\ &= 2p^2 = \frac{2}{9} \end{aligned}$$

(d) Temos que

$$\begin{aligned} \mathrm{E}(X) &= 1 \times p^2 + 2 \times p^2 + 3 \times p + 4 \times p + 5 \times p^2 \\ &= \frac{1}{9} + \frac{2}{9} + 1 + \frac{4}{3} + \frac{5}{9} \\ &= \frac{29}{9} = 3,2222 \end{aligned}$$

$$\begin{aligned} \mathrm{E}(X^2) &= 1^2 \times p^2 + 2^2 \times p^2 + 3^2 \times p + 4^2 \times p + 5^2 \times p^2 \\ &= \frac{1}{9} + \frac{4}{9} + 3 + \frac{16}{3} + \frac{25}{9} \\ &= \frac{105}{9} = \frac{35}{3} \\ \mathrm{Var}(X) &= \frac{35}{3} - \left(\frac{29}{9}\right)^2 = \frac{14}{81} \end{aligned}$$

♦♦

Exemplo 2.39 Jogo de dados

Um jogador A paga R$5,00 a B e lança um dado. Se sair face 3, ganha R$20,00. Se sair face 4, 5, ou 6, perde. Se sair face 1 ou 2, tem o direito de jogar novamente. Desta vez, lança dois dados. Se saírem duas faces 6, ganha R$50,00. Se sair uma face 6, recebe o dinheiro de volta. Nos demais casos, perde. Seja L o lucro líquido do jogador A nesse jogo. Calcule a função de probabilidade de L e o lucro esperado do jogador A.

Solução:
Sabemos que o dado é honesto e que os lançamentos são independentes. O diagrama de para o espaço amostral desse experimento está apresentado na Figura 2.3.

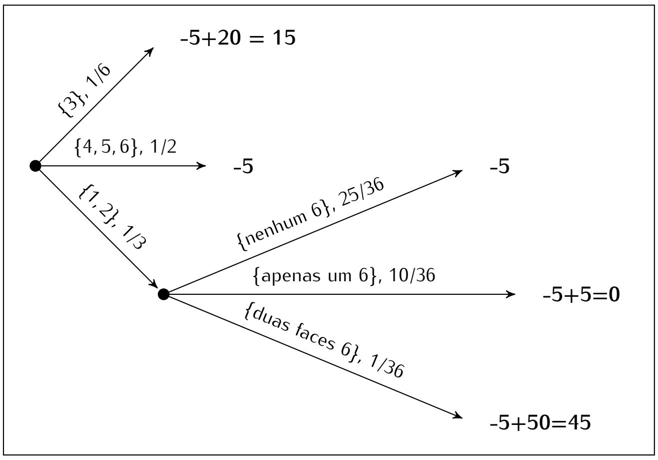


Figura 2.3 – Diagrama para o espaço amostral do Exemplo 2.39

Para calcular a probabilidade dos eventos associados aos lançamentos dos dois dados (parte inferior da árvore), usamos o fato de que a probabilidade da interseção de eventos independentes é o produto das probabilidades.

No cálculo da probabilidade de uma face 6 temos: $2 \times (1/6) \times (5/6)$. Multiplicamos por 2, porque a face 6 pode estar em qualquer um dos dois dados.

Vemos que os valores do lucro L são: -5; 0; 15; 45 e a função de probabilidade de L é

Lucro ℓ	-5	0	15	45
$\mathrm{P}(L=\ell)$	$\frac{1}{2}+\frac{2}{6}\times\frac{5}{6}\times\frac{5}{6}=\frac{158}{216}$	$\frac{2}{6}\times 2\times\frac{1}{6}\times\frac{5}{6}=\frac{20}{216}$	$\frac{1}{6}=\frac{36}{216}$	$\frac{2}{6}\times\frac{1}{6}\times\frac{1}{6}=\frac{2}{216}$

Assim,

$$\mathrm{E}(L) = -5 \times \frac{158}{216} + 15 \times \frac{36}{216} + 45 \times \frac{2}{216} = -\frac{160}{216} = -0,74.$$

♦♦

Capítulo 3

Algumas Distribuições Discretas

3.1 Introdução

Considere as seguintes situações:

1. (a) Lança-se uma moeda viciada e observa-se o resultado obtido e (b) pergunta-se a um eleitor se ele vai votar no candidato A ou B.
2. (a) Lança-se uma moeda n vezes e observa-se o número de caras obtidas e (b) de uma grande população, extrai-se uma amostra de n eleitores e pergunta-se a cada um deles em qual dos candidatos A ou B eles votarão e conta-se o número de votos do candidato A.
3. (a) De uma urna com P bolas vermelhas e Q bolas brancas, extraem-se n bolas sem reposição e conta-se o número de bolas brancas e (b) de uma população com P pessoas a favor do candidato A e Q pessoas a favor do candidato B, extrai-se uma amostra de tamanho n sem reposição e conta-se o número de pessoas a favor do candidato A na amostra.

Em cada uma das situações anteriores, os experimentos citados têm algo em comum: em certo sentido, temos a "mesma situação ", mas em contextos diferentes. Por exemplo, na situação 1, cada um dos experimentos tem dois resultados possíveis e observamos o resultado obtido. Na situação 3, temos uma população dividida em duas categorias e dela extraímos uma amostra sem reposição; o interesse está no número de elementos de uma determinada categoria.

Na prática, existem muitas outras situações que podem se "encaixar" nos modelos acima e mesmo em outros modelos. O que veremos nesse capítulo são alguns *modelos* de variáveis aleatórias discretas que podem descrever situações como as listadas anteriormente. Nesse contexto, um modelo será definido por uma variável aleatória e sua função de probabilidade, explicitando-se claramente as hipóteses de validade. De posse desses elementos, poderemos analisar diferentes situações práticas para tentar "encaixá-las" em algum dos modelos dados.

Neste capítulo, serão descritas as distribuições de probabilidade discretas mais usuais.A introdução de cada uma delas será feita através de um exemplo clássico (moeda, urna, baralho etc.) e, em seguida, serão explicitadas as características do experimento. Tais características são a ferramenta necessária para sabermos qual modelo se aplica a uma determinada situação prática. Definida a distribuição, calculam-se a média e a variância.

3.2 Distribuição Uniforme Discreta

Suponha que seu professor de Estatística decida dar de presente a um dos alunos um livro de sua autoria. Não querendo favorecer qualquer aluno em especial, ele decide sortear aleatoriamente o ganhador, dentre os 45 alunos da turma. Para isso, ele numera os nomes dos alunos que constam do diário de classe de 1 a 45, escreve esses números em pedaços iguais de papel, dobrando-os ao meio para que o número não fique visível, e sorteia um desses papéis depois de bem misturados. Qual é a probabilidade de que você ganhe o livro? Qual é a probabilidade de que o aluno que tirou a nota mais baixa na primeira prova ganhe o livro? E o que tirou a nota mais alta?

O importante a notar nesse exemplo é o seguinte: o professor tomou todos os cuidados necessários para não favorecer qualquer aluno em especial. Isso significa que todos os alunos têm a mesma chance de ganhar o livro. Temos, assim, um exemplo da distribuição uniforme discreta.

Definição 3.1 Distribuição Uniforme Discreta
Uma variável aleatória X tem distribuição de ***Uniforme Discreta*** *com parâmetro n se $Im(X)$ é um conjunto finito com n elementos e a probabilidade de X assumir qualquer um do n elementos é a mesma, independente do elemento.*

Note que, em uma distribuição discreta uniforme, todos os valores são igualmente prováveis. Veja que o parâmetro n é o número de valores que a variável aleatória pode assumir e por isso n pode ser qualquer valor no conjunto $\mathbb{N}$. Chamamos de **espaço paramétrico** o conjunto de valores que o parâmetro de uma distribuição pode assumir. Nesse caso, o espaço paramétrico para o parâmetro n é o conjunto dos números naturais, isto é, $\mathbb{N}$.

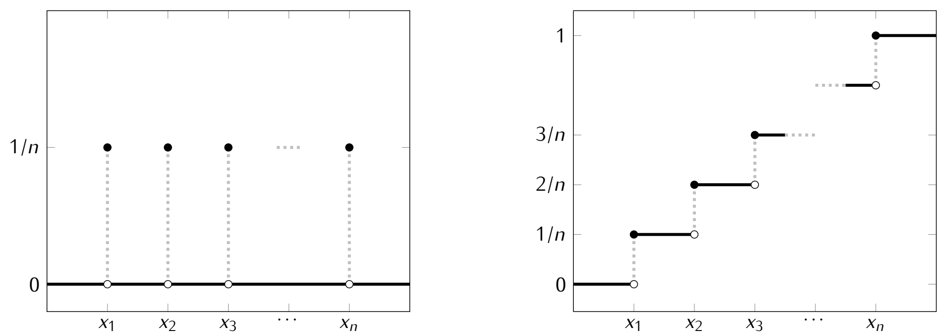
Vamos denotar a distribuição Uniforme Discreta com parâmetro n por $Unif(n)$. Nesse caso, se quisermos indicar que uma variável aleatória X segue a distribuição Uniforme Discreta com parâmetro n podemos simplesmente escrever: $X \sim Unif(n)$ (lê-se: a variável aleatória X tem distribuição Uniforme Discreta com parâmetro n).

Função de Probabilidade e Função de Distribuição

Seja $X \sim Unif(n)$ e suponha $Im(X) = \{x_1, x_2, \ldots, x_n\}$. Logo a sua função de probabilidade é definida por

$$p_X(x_i) = \mathrm{P}(X = x_i) = \frac{1}{n} \qquad \forall i = 1, 2, \ldots, n. \tag{3.1}$$

Na Figura 3.1 a seguir estão os gráficos da função de probabilidade e função de distribuição de uma variável aleatória discreta. Veja que como a probabilidade associada a cada elemento x_i de $Im(X)$ é o mesmo $\forall$ i, os degraus no gráfico da Função de Distribuição tem mesmo tamanho, Figura 3.1(b).



(a) Função de Probabilidade (b) Função de Distribuição

Figura 3.1 – Distribuição Uniforme Discreta.

Esperança e Variância

Seja X uma v.a. discreta uniforme que assume valores $x_1, x_2, \ldots, x_n$. Seguindo a Definição 2.16 podemos calcular $\mathrm{E}(X)$,

$$\mathrm{E}(X) = \frac{1}{n}x_1 + \frac{1}{n}x_2 + \cdots + \frac{1}{n}x_n = \overline{x}, \tag{3.2}$$

ou seja, $\mathrm{E}(X)$ é a média aritmética dos valores possíveis de X.

Com relação à variância, temos a Definição 2.29, de onde tiramos que

$$\mathrm{Var}(X) = \mathrm{E}\left[X - \mathrm{E}(X)\right]^2 = \frac{1}{n}(x_1 - \overline{x})^2 + \frac{1}{n}(x_2 - \overline{x})^2 + \cdots + \frac{1}{n}(x_n - \overline{x})^2 = \sigma_X^2 \tag{3.3}$$

Exemplo 3.2 Lançamento de uma moeda
Considere o lançamento de uma moeda. Vamos definir a seguinte variável aleatória X associada a esse experimento:

$$X = \begin{cases} 0 & \text{, se ocorre cara} \\ 1 & \text{, se ocorre coroa} \end{cases}$$

Verifique se X é variável aleatória uniforme discreta e calcule sua média e variância.

Solução:
Para que essa v.a. tenha distribuição uniforme, é necessário supor que a moeda seja honesta e, nesse caso,

$$\begin{aligned} p_X(0) &= p_X(1) = \frac{1}{2} \\ \mathrm{E}(X) &= \frac{0+1}{2} = \frac{1}{2} \\ \mathrm{Var}(X) &= \frac{1}{2} \times \left(0 - \frac{1}{2}\right)^2 + \frac{1}{2} \times \left(1 - \frac{1}{2}\right)^2 \\ &= \frac{1}{2} \times \frac{1}{4} + \frac{1}{2} \times \frac{1}{4} = \frac{1}{4} \end{aligned}$$

♦♦

Exemplo 3.3 Conserto de máquina
Uma máquina pode apresentar 5 tipos diferentes de defeitos, que ocorrem aproximadamente na mesma frequência. Dependendo do tipo de defeito, o técnico leva 1, 2, 3, 4 ou 5 horas para consertar a máquina.

(a) Descreva o modelo probabilístico apropriado para representar a duração do tempo de reparo da máquina.

(b) Qual é o tempo médio de reparo desta máquina? E o desvio-padrão deste tempo de reparo?

(c) São 15 horas e acaba de ser entregue uma máquina para reparo. A jornada normal de trabalho do técnico termina às 17 horas. Qual é a probabilidade de que o técnico não precise fazer hora extra para terminar o conserto desta máquina?

Solução:
Seja T = "tempo de reparo, em horas".

(a) Como os defeitos ocorrem na mesma frequência, o modelo probabilístico apropriado é uma distribuição uniforme:

t	1	2	3	4	5
$p_T(t) = \mathrm{P}(T = t)$	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{5}$

(b) $\mathrm{E}(T) = \dfrac{1+2+3+4+5}{5} = 3$ horas

$\mathrm{Var}(T) = \dfrac{1^2+2^2+3^2+4^2+5^2}{5} - 9 = 2 \Longrightarrow DP(T) = 1,41$ horas

(c) Seja E o evento "técnico vai ter que fazer hora extra". Então

$$\mathrm{P}(E) = \mathrm{P}(T > 2) = \frac{3}{5} = 0,6$$

Logo, a probabilidade de que ele não tenha que fazer hora extra é 0,4.

♦♦

3.3 Distribuição de Bernoulli

Considere o lançamento de uma moeda. A característica de tal experimento aleatório é que ele possui apenas dois resultados possíveis. Uma situação análoga surge quando da extração da carta de um baralho, em que o interesse está apenas na cor (preta ou vermelha) da carta sorteada.

Definição 3.4 Ensaio de Bernoulli
*Um **ensaio de Bernoulli**, ou experimento de Bernoulli, é um experimento aleatório com apenas dois resultados possíveis; por convenção, um deles é chamado "sucesso" e o outro, "fracasso".*

Suponha que seja realizado um ensaio de Bernoulli e, baseado nesse experimento, seja definida a variável aleatória X:

$$X = \begin{cases} 1 & \text{, se ocorre sucesso} \\ 0 & \text{, se ocorre fracasso} \end{cases}$$

Veja que a variável aleatória assume o valor 1 no caso de sucesso e 0 no caso de fracasso. X é chamada de v.a. de Bernoulli, como mostra a Definição 3.5.

Definição 3.5 Distribuição de Bernoulli
Uma variável aleatória X tem ***distribuição de Bernoulli*** *com parâmetro p se ela é uma variável indicadora de algum evento, denominado "sucesso", com probabilidade p de ocorrência.*

Primeiro veja que para qualquer ensaio de Bernoulli é sempre possível definir uma v.a. de Bernoulli, como já comentado acima. Veja também que $Im(X) = \{0, 1\}$, uma vez que X é v.a. indicadora (Definição 1.11), $P(X = 1) = p$, probabilidade de sucesso, e $P(X = 0) = 1 - p$, probabilidade de fracasso. Como p é uma probabilidade, o espaço paramétrico é o intervalo $[0, 1]$. É comum denotar a probabilidade de fracasso por q, isto é, $q = 1 - p$ e $0 \leq q \leq 1$.

Vamos denotar a distribuição de Bernoulli com parâmetro p por $Bern(p)$. Nesse caso, se quisermos indicar que uma variável aleatória X segue a distribuição de Bernoulli com parâmetro p podemos simplesmente escrever: $X \sim Bern(p)$ (lê-se: a variável aleatória X tem distribuição de Bernoulli com parâmetro p).

Função de Probabilidade e Função de Distribuição

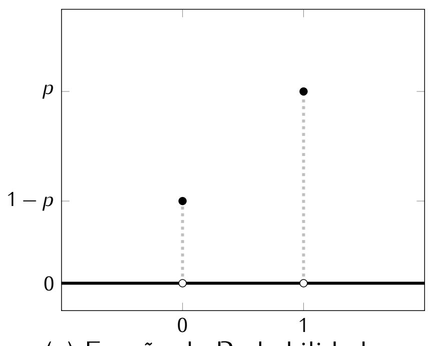
A função de probabilidade de $X \sim Bern(p)$ pode também ser escrita da seguinte forma:

$$p_X(x) = P(X = x) = p^x (1 - p)^{1-x} \qquad x = 0, 1. \tag{3.4}$$

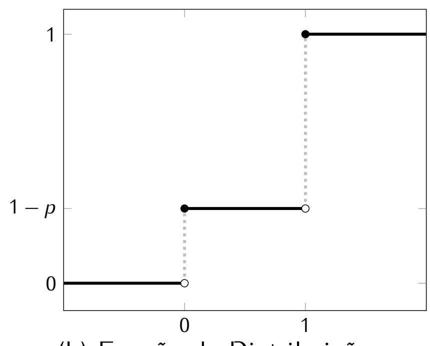
Verifique que $P(X = 1) = p$ e $P(X = 0) = 1 - p$. Já a sua função de distribuição acumulada é dada por:

$$F_X(x) = \begin{cases} 0 & \text{se } x < 0 \\ 1 - p & \text{se } 0 \leq x < 1 \\ 1 & \text{se } x \geq 1 \end{cases} \tag{3.5}$$

Na Figura 3.2, temos os gráficos da função de probabilidade e da função de distribuição acumulada de uma variável de Bernoulli. Como $Im(X)$ é um conjunto com apenas dois elementos, $Im(X) = \{0, 1\}$, a função de distribuição de X só tem dois pontos de descontinuidade, em 0 e em 1.



(a) Função de Probabilidade



(b) Função de Distribuição

Figura 3.2 – Distribuição de Bernoulli.

Esperança e Variância

Seja $X \sim Bern(p)$. Então,

$$\begin{aligned}
\mathrm{E}(X) &= 0 \times (1-p) + 1 \times p = p \\
\mathrm{E}(X^2) &= 0^2 \times (1-p) + 1^2 \times p = p \\
\mathrm{Var}(X) &= \mathrm{E}(X^2) - [\mathrm{E}(X)]^2 = p - p^2 = p(1-p)
\end{aligned}$$

Em resumo, se $X \sim Bern(p)$ temos

$$\mathrm{E}(X) = p \tag{3.6}$$
$$\mathrm{Var}(X) = p(1-p) \tag{3.7}$$

Exemplo 3.6 Lançamento de uma moeda
Considere, assim como no Exemplo 3.2, o lançamento de uma moeda e a seguinte variável aleatória X associada a esse experimento:

$$X = \begin{cases} 0 & \text{, se ocorre cara} \\ 1 & \text{, se ocorre coroa.} \end{cases}$$

Seja p a probabilidade de cara, $0 < p < 1$. Já vimos que se $p = 1/2$ então X é uniforme discreta. Encontre a distribuição de X qualquer que seja o valor de p.

Solução:
Como $Im(X) = \{0, 1\}$, X tem distribuição de Bernoulli com parâmetro p, qualquer que seja p. Nesse caso o "sucesso" é definido como a saída cara, e ocorre com probabilidade p, e o "fracasso" a saída coroa.

Note que se $p = 1/2$ X pode ser considerada uma v.a. de Bernoulli ou uniforme discreta, para os outros valores de p X só pode ser considerada v.a. de Bernoulli. Nesse caso, a Bernoulli com parâmetro $p = 1/2$ é equivalente à distribuição uniforme.

♦♦

Exemplo 3.7 Auditoria da Receita Federal
Um auditor da Receita Federal examina declarações de Imposto de Renda de pessoas físicas, cuja variação patrimonial ficou acima do limite considerado aceitável. De dados históricos, sabe-se que 10% dessas declarações são fraudulentas. Considere o experimento correspondente ao sorteio aleatório de uma dessas declarações e defina a v.a. indicadora do evento A = foi sorteada uma declaração fraudulenta (Definição 1.11). Encontre o modelos probabilístico adequado para essa v.a.

Solução:
Primeiro veja que o experimento correspondente ao sorteio aleatório de uma dessas declarações em que é observado se a declaração sorteada ou não é fraudulenta é um experimento de Bernoulli.

A v.a. indicadora do evento em questão pode ser definida por

$$I_A = \begin{cases} 1 & \text{, se a declaração sorteada é fraudulenta} \\ 0 & \text{, caso contrário.} \end{cases}$$

Veja que $Im(I_A) = \{0, 1\}$ e por isso podemos dizer que esta v.a. tem distribuição de Bernoulli. O parâmetro p é a probabilidade de "sucesso", que por definição é o evento associado ao valor 1 da variável aleatória. Da forma com que I_A foi construída o "sucesso" para esse ensaio de Bernoulli é encontrar uma declaração fraudulenta. Então $p = 0,1$.

Esse exemplo ilustra o fato de que "sucesso", nesse contexto, nem sempre significa uma situação feliz na vida real. Aqui, sucesso é definido de acordo com o interesse estatístico no problema. Em uma situação mais dramática, "sucesso" pode indicar a morte de um paciente, por exemplo.

♦♦

3.4 Distribuição Binomial

Vamos introduzir a distribuição binomial, uma das mais importantes distribuições discretas, através de dois um exemplo. Em seguida, discutiremos as hipóteses feitas e apresentaremos os resultados formais sobre tal distribuição e novos exemplos.

Exemplo 3.8 Lançamentos de uma moeda
Considere o seguinte experimento: uma moeda é lançada 4 *vezes. Seja* $p = \mathrm{P}(\textit{sair a face cara})$*. Vamos definir a seguinte variável aleatória associada a este experimento:*

$$X = \textit{ número de faces caras considerando os quatro lançamentos da moeda.}$$

Encontre a função de probabilidade da v.a. X*.*

Solução:
Como visto antes, cada lançamento da moeda representa um experimento de Bernoulli e como o interesse está no número de caras, vamos definir sucesso = cara.

Para encontrar a função de probabilidade de X, o primeiro fato a notar é que os valores possíveis de X são: 0, que equivale à ocorrência de nenhuma cara e, portanto, de 4 coroas; 1, que equivale à ocorrência de apenas 1 cara e, portanto, 3 coroas; 2, que equivale à ocorrência de 2 caras e, portanto, 2 coroas; 3, que equivale à ocorrência de 3 caras e 1 coroa e, finalmente, 4, que equivale à ocorrência de 4 caras e nenhuma coroa. Assim, os possíveis valores de X são

$$X = 0, 1, 2, 3, 4$$

Vamos, agora, calcular a probabilidade de cada um desses valores, de modo a completar a especificação da função de probabilidade de X. Para isso, vamos representar por K_i o evento "cara no i-ésimo lançamento" e por C_i o evento "coroa no i-ésimo lançamento".

- $X = 0$

  Temos a seguinte equivalência de eventos:

  $$\{X = 0\} \equiv C_1 \cap C_2 \cap C_3 \cap C_4$$

  É razoável supor que os lançamentos da moeda sejam eventos independentes, ou seja, o resultado de um lançamento não interfere no resultado de qualquer outro lançamento.

  Dessa forma, os eventos C_i e K_j são independentes para $i \neq j$. (Note que os eventos C_i e K_i são mutuamente exclusivos e, portanto, não são independentes – se sair cara em um lançamento específico, não é possível sair coroa nesse mesmo lançamento e vice-versa).

  Analogamente, os eventos C_i e C_j são independentes para $i \neq j$, bem como os eventos K_i e K_j, $i \neq j$. Pela regra da probabilidade da interseção de eventos independentes, resulta que

$$\begin{aligned} \mathrm{P}\left(C_1 \cap C_2 \cap C_3 \cap C_4\right) &= \mathrm{P}(C_1) \times \mathrm{P}(C_2) \times \mathrm{P}(C_3) \times \mathrm{P}(C_4) \\ &= (1-p) \times (1-p) \times (1-p) \times (1-p) \\ &= (1-p)^4 \end{aligned}$$

- $X = 1$

  O evento $X = 1$ corresponde à ocorrência de 1 cara e, consequentemente, de 3 coroas. Uma sequência possível de lançamentos é

  $$K_1 \cap C_2 \cap C_3 \cap C_4.$$

  Vamos calcular a probabilidade desse resultado. Como antes, os lançamentos são eventos independentes e, portanto,

  $$\begin{aligned} \mathrm{P}(K_1 \cap C_2 \cap C_3 \cap C_4) &= \mathrm{P}(K_1) \times \mathrm{P}(C_2) \times \mathrm{P}(C_3) \times \mathrm{P}(C_4) \\ &= p \times (1-p) \times (1-p) \times (1-p) \\ &= p(1-p)^3 \end{aligned}$$

  Mas qualquer sequência com 1 cara resulta em $X = 1$, ou seja, a face cara pode estar em qualquer uma das quatro posições e todas essas sequências resultam em $X = 1$. Além disso, definida a posição da face cara, as posições das faces coroas já estão determinadas – são as posições restantes. Então, temos a seguinte equivalência:

  $$\begin{aligned} \{X = 1\} \equiv\ & \{K_1 \cap C_2 \cap C_3 \cap C_4\} \cup \{C_1 \cap K_2 \cap C_3 \cap C_4\} \cup \\ & \{C_1 \cap C_2 \cap K_3 \cap C_4\} \cup \{C_1 \cap C_2 \cap C_3 \cap K_4\} \end{aligned}$$

  Mas os eventos que aparecem no lado direito da expressão anterior são eventos mutuamente exclusivos. Logo,

  $$\begin{aligned} \mathrm{P}(X = 1) &= \mathrm{P}(K_1 \cap C_2 \cap C_3 \cap C_4) \\ &\quad + \mathrm{P}(C_1 \cap K_2 \cap C_3 \cap C_4) \\ &\quad + \mathrm{P}(C_1 \cap C_2 \cap K_3 \cap C_4) \\ &\quad + \mathrm{P}(C_1 \cap C_2 \cap C_3 \cap K_4) \\ &= p \times (1-p) \times (1-p) \times (1-p) \\ &\quad +(1-p) \times p \times (1-p) \times (1-p) \\ &\quad +(1-p) \times (1-p) \times p \times (1-p) \\ &\quad +(1-p) \times (1-p) \times (1-p) \times p \\ &= 4p(1-p)^3 \end{aligned}$$

- $X = 2$

  O evento $X = 2$ corresponde à ocorrência de 2 caras e, consequentemente, de 2 coroas. Qualquer uma dessas sequêcias tem probabilidade $p^2(1-p)^2$.

  As sequências de lançamentos com 2 caras e 2 coroas são as seguintes:

  $$\begin{array}{c} K_1K_2C_3C_4 \\ K_1C_2K_3C_4 \\ K_1C_2C_3K_4 \\ C_1C_2K_3K_4 \\ C_1K_2C_3K_4 \\ C_1K_2K_3C_4 \end{array}$$

Todas essas 6 sequências têm a mesma probabilidade e correspondem a eventos mutuamente exclusivos. Temos a seguinte equivalência:

$$\begin{aligned}\{X=2\} \quad \equiv \quad & (K_1 \cap K_2 \cap C_3 \cap C_4) \cup (K_1 \cap C_2 \cap K_3 \cap C_4) \cup \\ & (K_1 \cap C_2 \cap C_3 \cap K_4) \cup (C_1 \cap C_2 \cap K_3 \cap K_4) \cup \\ & (C_1 \cap K_2 \cap C_3 \cap K_4) \cup (C_1 \cap K_2 \cap K_3 \cap C_4)\end{aligned}$$

e, portanto,

$$\begin{aligned}\mathrm{P}(X=2) &= \mathrm{P}(K_1 \cap K_2 \cap C_3 \cap C_4) + \mathrm{P}(K_1 \cap C_2 \cap K_3 \cap C_4) + \\ & \quad \mathrm{P}(K_1 \cap C_2 \cap C_3 \cap K_4) + \mathrm{P}(C_1 \cap C_2 \cap K_3 \cap K_4) + \\ & \quad \mathrm{P}(C_1 \cap K_2 \cap C_3 \cap K_4) + \mathrm{P}(C_1 \cap K_2 \cap K_3 \cap C_4) \\ &= 6p^2(1-p)^2\end{aligned}$$

- $X = 3$ e $X = 4$

  Os casos $X = 3$ e $X = 4$ são análogos aos casos $X = 1$ e $X = 0$, respectivamente; basta trocar caras por coroas e vice-versa. Assim,

$$\mathrm{P}(X=3) = 4p^3(1-p)$$
$$\mathrm{P}(X=4) = p^4$$

Dessa forma chegamos à resposta:

$$p_X(x) = \begin{cases} (1-p)^4 & \text{, se x} = 0 \\ 4p(1-p)^3 & \text{, se x} = 1 \\ 6p^2(1-p)^2 & \text{, se x} = 2 \\ 4p^3(1-p) & \text{, se x} = 3 \\ p^4 & \text{, se x} = 4 \\ 0 & \text{, caso contrário.} \end{cases}$$

É importante notar que a hipótese de independência dos lançamentos da moeda foi absolutamente fundamental na solução do exemplo; foi ela que nos permitiu multiplicar as probabilidades dos resultados de cada lançamento para obter a probabilidade da sequência completa de n lançamentos. Embora essa hipótese seja muito razoável nesse exemplo, ainda assim é uma hipótese "subjetiva".

Outra propriedade utilizada foi a da probabilidade da união de eventos mutuamente exclusivos. Mas aqui essa propriedade é óbvia, ou seja, não há qualquer subjetividade: os eventos $C_1 \cap K_2$ e $K_1 \cap C_2$ são mutuamente exclusivos, pois no primeiro lançamento ou sai cara ou sai coroa; não pode sair cara e coroa no primeiro lançamento, ou seja, cada lançamento é um experimento de Bernoulli.

♦♦

Exemplo 3.9 Bolas em uma urna

Uma urna contém quatro bolas brancas e seis bolas verdes. Três bolas são retiradas dessa urna, com reposição, isto é, depois de tirada a primeira bola, ela é recolocada na urna e sorteia-se a segunda, que também é recolocada na urna para, finalmente, ser sorteada a terceira bola. Vamos definir a seguinte variável aleatória associada a esse experimento:

$$X = \text{ número de bolas brancas sorteadas.}$$

Encontre a função de probabilidade da v.a. X.

Solução:
O importante a notar aqui é o seguinte: como cada bola sorteada é recolocada na urna antes da próxima extração, a composição da urna é sempre a mesma e o resultado de uma extração não afeta o resultado de outra extração qualquer. Dessa forma, em todas as extrações a probabilidade de bola branca (e também bola verde) é a mesma e podemos considerar as extrações como independentes. Assim, temos uma situação análoga à do exemplo anterior: temos três repetições de um experimento (sorteio de uma bola), essas repetições são independentes, em cada uma delas há dois resultados possíveis – bola branca (sucesso) ou bola verde (fracasso) – e as probabilidades de sucesso e fracasso são as mesmas. Assim, cada extração equivale a um experimento de Bernoulli e como o interesse está nas bolas brancas, vamos considerar sucesso = bola branca e da observação anterior resulta que

$$P(sucesso) = \frac{4}{10}$$

Os valores possíveis de X são $0, 1, 2, 3$, uma vez que são feitas três extrações. Vamos calcular a probabilidade de cada um dos valores de X. Como antes, vamos denotar por V_i o evento "bola verde na i-ésima extração" e por B_i o evento "bola branca na i-ésima extração". Da discussão anterior, resulta que, para $i \neq j$, os eventos V_i e B_j são independentes, assim como os eventos B_i e B_j e os eventos V_i e V_j.

- $X = 0$

  Esse resultado equivale à extração de bolas verdes em todas as três extrações.

  $$\{X = 0\} \equiv \{V_1 \cap V_2 \cap V_3\}$$

  Logo,

  $$\begin{aligned} P(X = 0) &= P(V_1 \cap V_2 \cap V_3) \\ &= P(V_1) \times P(V_2) \times P(V_3) \\ &= \frac{6}{10} \times \frac{6}{10} \times \frac{6}{10} = \left(\frac{6}{10}\right)^3 \end{aligned}$$

- $X = 1$

  Esse resultado equivale à extração de uma bola branca e, por consequência, duas bolas verdes. A bola branca pode sair em qualquer uma das três extrações e, definida a posição da bola branca, as posições das bolas verdes ficam totalmente estabelecidas. Logo,

  $$P(X = 1) = 3\left(\frac{4}{10}\right)\left(\frac{6}{10}\right)^2$$

- $X = 2$ e $X = 3$

  Os casos $X = 2$ e $X = 3$ são análogos aos casos $X = 1$ e $X = 0$, respectivamente; basta trocar bola branca por bola verde e vice-versa. Assim,

  $$P(X = 2) = 3\left(\frac{4}{10}\right)^2\left(\frac{6}{10}\right) \qquad \text{e} \qquad P(X = 3) = \left(\frac{4}{10}\right)^3$$

♦♦

Esses dois exemplos ilustram a *distribuição binomial*, que depende de dois parâmetros: o número de repetições n e a probabilidade de sucesso p de cada ensaio de Bernoulli. No Exemplo 3.8, $n = 4$ e temos uma probabilidade de sucesso qualquer p. No Exemplo 3.9, $n = 3$ e $p = \frac{4}{10}$.

Nos dois exemplos anteriores, tínhamos repetições de um ensaio de Bernoulli que podiam ser consideradas independentes e a probabilidade de sucesso p se mantinha constante ao longo de todas as repetições. Essas são as condições definidoras de um *experimento binomial*.

Definição 3.10 Experimento Binomial
Um ***experimento binomial*** *consiste em repetições independentes de ensaios de Bernoulli com probabilidade p de sucesso, probabilidade essa que permanece constante em todas as repetições.*

Definição 3.11 Distribuição Binomial
Uma variável aleatória X tem ***distribuição Binomial*** *com parâmetros n e p quando esta representa o número de sucessos em n ensaios de Bernoulli, independentes, cada um com probabilidade p de sucesso.*

Veja que se X tem distribuição Binomial, então os valores possíveis de X são $0, 1, 2, \ldots, n$, isto é, $Im(X) = \{0, 1, 2, \ldots, n\}$. O espaço paramétrico para o parâmetro n é o conjunto $\mathbb{N}$ e para o parâmetro p o intervalo $[0, 1]$.

Vamos denotar a distribuição Binomail com parâmetros n e p por $Binom(n, p)$. Nesse caso, se quisermos indicar que uma variável aleatória X segue a distribuição Binomial com parâmetros n e p podemos simplesmente escrever: $X \sim B(n, p)$ ou $X \sim bin(n, p)$ (lê-se: a variável aleatória X tem distribuição Binomial com parâmetros n e p).

Função de Probabilidade

Nessa seção ser apresentada a função de probabilidade de $X \sim B(n, p)$. Para isso precisamos definir $\mathrm{P}(X = x)$ para $x = 0, 1, 2, \ldots, n$. Veja que o evento $\{X = x\}$ corresponde a todas as sequências de resultados com x sucessos e $n - x$ fracassos. Como as repetições são independentes, cada uma dessas sequências tem probabilidade $p^x(1 - p)^{n-x}$. O número total de tais sequências é dado pelo *coeficiente binomial* $\binom{n}{x}$, apresentado na Definição 3.12.

Definição 3.12 Coeficiente Binomial
Para n e k inteiros define-se o ***coeficiente binomial*** $\binom{n}{k}$ *como o número de maneiras com que k elementos podem ser escolhidos dentro de um total de n elementos. Para $k \leq n$ esse número pode ser calculado por:*

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}.$$

Por convenção definimos $\binom{n}{k} = 0$ *quando $k > n$, pois não há como escolher mais elementos do que o total de elementos disponíveis.*

Veja que $\binom{n}{k}$ pode ser interpretado como o número de subconjuntos com k elementos dentro de um conjunto com n elementos. Lembre-se que $n!$ representa o fatorial de n, definido como $n! = n \times (n-1) \times (n-2) \times \cdots \times 2 \times 1$. Por definição, $0! = 1$.

Dessa forma chegamos na seguinte função de probabilidade para $X \sim B(n,p)$.

$$p_X(x) = \mathrm{P}(X = x) = \binom{n}{x} p^x (1-p)^{n-x}, \qquad x = 0, 1, 2, \ldots, n. \tag{3.8}$$

Vamos verificar que a expressão apresentada na Equação 3.8 realmente é uma função de probabilidade. Para isso é preciso mostrar que $p_X(x) \geq 0$ para qualquer $x \in \mathbb{R}$ e $\sum_{\forall x \in Im(X)} p_X(x) = 1$. É imediato ver, da Equação 3.8, que $p_X(x) \geq 0$. Para mostrar que a soma das probabilidades é 1, usaremos o Teorema 3.13 do binômio de Newton.

Teorema 3.13 Binômio de Newton
Dados dois números reais quaisquer x e a e um inteiro qualquer n, então

$$(x+a)^n = \sum_{k=0}^{n} \binom{n}{k} a^k x^{n-k}$$

Vamos agora ao resultado que falta.

$$\sum_{\forall x \in Im(X)} p_X(x) = \sum_{x=0}^{n} \binom{n}{x} p^x (1-p)^{n-x} \overset{\star}{=} (1 + (1-p))^n = 1$$

Onde a passagem $\overset{\star}{=}$ é uma aplicação direta do Binômio de Newton, Teorema 3.13. Assim, a Equação 3.8 realmente define uma função de probabilidade.

Esperança e Variância

Vamos calcular primeiro $\mathrm{E}(X)$, para $X \sim B(n,p)$.

$$\mathrm{E}(X) = \sum_{k=0}^{n} k \, \mathrm{P}(X = k) = \sum_{k=0}^{n} k \binom{n}{k} p^k (1-p)^{n-k} = \sum_{k=0}^{n} k \frac{n!}{k!\,(n-k)!} p^k (1-p)^{n-k}$$

Quando $k = 0$, a parcela correspondente no somatório é nula. Logo, podemos escrever (note o índice do somatório!):

$$\mathrm{E}(X) = \sum_{k=1}^{n} k \frac{n!}{k!\,(n-k)!} p^k (1-p)^{n-k} = \sum_{k=1}^{n} k \frac{n!}{k\,(k-1)!\,(n-k)!} p^k (1-p)^{n-k}$$

e como $k \neq 0$, podemos dividir o numerador e o denominadr por k, o que resulta na simplificação

$$\begin{aligned}
\mathrm{E}(X) &= \sum_{k=1}^{n} \frac{n!}{(k-1)!\,(n-k)!} p^k (1-p)^{n-k} = \sum_{k=1}^{n} \frac{n\,(n-1)!}{(k-1)!\,(n-k)!} \left(p \times p^{k-1}\right) (1-p)^{n-k} = \\
&= np \sum_{k=1}^{n} \frac{(n-1)!}{(k-1)!\,(n-k)!} p^{k-1} (1-p)^{n-k} = np \sum_{k=1}^{n} \binom{n-1}{k-1} p^{k-1} (1-p)^{n-k}
\end{aligned}$$

Fazendo $j = k - 1$, temos que

$$\begin{aligned} k &= j + 1 \\ k &= 1 \Rightarrow j = 0 \\ k &= n \Rightarrow j = n - 1 \end{aligned}$$

Logo,

$$\mathrm{E}(X) = np \underbrace{\sum_{j=0}^{n-1} \binom{n-1}{j} p^j (1-p)^{n-1-j}}_{1}$$

Veja que a expressão marcada com chaves é igual a 1 pois é o somatório da função de probabilidade de uma distribuição binomial com parâmetros $(n-1)$ e p para todos os valores possíveis dessa variável aleatória. Portanto,

$$X \sim B(n,p) \qquad \Rightarrow \qquad \mathrm{E}(X) = np \tag{3.9}$$

Vamos, agora, calcular $\mathrm{E}\left(X^2\right)$. Usando raciocínio análogo ao usado no cálculo da esperança, temos que:

$$\begin{aligned} \mathrm{E}\left(X^2\right) &= \sum_{k=0}^{n} k^2 \binom{n}{k} p^k (1-p)^{n-k} = \sum_{k=1}^{n} k^2 \frac{n!}{k!\,(n-k)!} p^k (1-p)^{n-k} = \\ &= \sum_{k=1}^{n} k^2 \frac{n!}{k\,(k-1)!\,(n-k)!} p^k (1-p)^{n-k} = \\ &= \sum_{k=1}^{n} k \frac{n!}{(k-1)!\,(n-k)!} p^k (1-p)^{n-k} = \\ &= \sum_{k=1}^{n} k \frac{n\,(n-1)!}{(k-1)!\,(n-k)!} \left(p \times p^{k-1}\right) (1-p)^{n-k} = \\ &= np \sum_{k=1}^{n} k \frac{(n-1)!}{(k-1)!\,(n-k)!} p^{k-1} (1-p)^{n-k} \end{aligned}$$

Faça a seguinte troca de variável: $j = k - 1$. Continuando,

$$\begin{aligned} \mathrm{E}(X) &= np \sum_{j=0}^{n-1} (j+1) \frac{(n-1)!}{j!\,(n-j-1)!} p^j (1-p)^{n-j-1} = \\ &= np \sum_{j=0}^{n-1} j \frac{(n-1)!}{j!\,(n-j-1)!} p^j (1-p)^{n-j-1} + np \sum_{j=0}^{n-1} \frac{(n-1)!}{j!\,(n-1-j)!} p^j (1-p)^{n-j-1} = \\ &= \underbrace{np \sum_{j=0}^{n-1} j \binom{n-1}{j} p^j (1-p)^{n-1-j}}_{(n-1)p} + np \underbrace{\sum_{j=0}^{n-1} \binom{n-1}{j} p^j (1-p)^{n-1-j}}_{1} \end{aligned}$$

O primeiro somatório é a esperança de uma binomial com parâmetros $(n-1)$ e p; portanto, pelo resultado apresentado na Equação 3.9, é igual a $(n-1)\,p$. Já o segundo somatório é a soma das probabilidades dos valores de uma binomial com esses mesmos parâmetros, logo, é igual a 1. Segue, então, que

$$\mathrm{E}\left(X^2\right) = np\,(n-1)\,p + np \times 1 = n^2p^2 - np^2 + np$$

e, portanto,

$$\operatorname{Var}(X) = \mathrm{E}(X^2) - [\mathrm{E}(X)]^2 = n^2p^2 - np^2 + np - (np)^2 = np - np^2$$

ou seja,

$$X \sim B(n,p) \qquad \Rightarrow \qquad \operatorname{Var}(X) = np\,(1-p) \tag{3.10}$$

Note que a esperança e a variância da binomial são iguais à esperança e à variância da distribuição de Bernoulli, multiplicadas por n, o número de repetições. Pode-se pensar na distribuição de Bernoulli como uma distribuição binomial com parâmetros $n = 1$ e p.

Exemplo 3.14 Tiro ao alvo
Um atirador acerta, na mosca do alvo, 20% dos tiros. Se ele dá 10 tiros, qual a probabilidade de ele acertar na mosca no máximo uma vez?

Solução:
Podemos pensar os tiros como experimentos de Bernoulli independentes, em que o sucesso é acertar no alvo e a probabilidade de sucesso é 0,20. Então, o problema pede $\mathrm{P}(X \leq 1)$, em que $X =$ número de acertos em 10 tiros. Logo, $X \sim bin(10; 0,20)$ e

$$\begin{aligned}
\mathrm{P}(X \leq 1) &= \mathrm{P}(X=0) + \mathrm{P}(X=1) \\
&= \binom{10}{0}(0,20)^0\,(0,80)^{10} + \binom{10}{1}(0,20)^1\,(0,80)^9 \\
&= 0,37581
\end{aligned}$$

♦♦

Exemplo 3.15 Partidas de um jogo
Dois adversários A e B disputam uma série de oito partidas de um determinado jogo. A probabilidade de A ganhar uma partida é 0,6 e não há empate. Qual é a probabilidade de A ganhar a série?

Solução:
Note que só podem ocorrer vitórias ou derrotas, o que significa que temos repetições de um experimento de Bernoulli com probabilidade 0,6 de sucesso (vitória do jogador A). Assumindo a independência das provas, se definimos $X =$ número de vitórias de A, então $X \sim bin(8; 0,6)$ e o problema pede $\mathrm{P}(X \geq 5)$, isto é, probabilidade de A ganhar mais partidas que B.

$$\begin{aligned}
\mathrm{P}(X \geq 5) &= \mathrm{P}(X=5) + \mathrm{P}(X=6) + \mathrm{P}(X=7) + \mathrm{P}(X=8) \\
&= \binom{8}{5}(0,6)^5\,(0,4)^3 + \binom{8}{6}(0,6)^6\,(0,4)^2 + \\
&\quad + \binom{8}{7}(0,6)^7\,(0,4)^1 + \binom{8}{8}(0,6)^8\,(0,4)^0 \\
&= 0,5940864
\end{aligned}$$

♦♦

Exemplo 3.16 Parâmetros da binomial
Em uma distribuição binomial, sabe-se que a média é 4,5 e a variância é 3,15. Encontre os valores dos parâmetros da distribuição.

Solução:
Temos que

$$\begin{aligned}
np &= 4,5 \\
np(1-p) &= 3,15
\end{aligned}$$

Substituindo a primeira equação na segunda, resulta

$$\begin{aligned} 4,5(1-p) &= 3,15 \Rightarrow \\ 1-p &= 0,7 \Rightarrow \\ p &= 0,3 \end{aligned}$$

Substituindo na primeira equação, obtemos que

$$n = 4,5/0,3 = 15.$$

♦♦

Forma da Distribuição Binomial

Se $X \sim bin(n;p)$, então temos

$$\begin{aligned} \frac{\mathrm{P}(X=k+1)}{\mathrm{P}(X=k)} &= \frac{\dfrac{n!}{(k+1)!(n-k-1)!}p^{k+1}(1-p)^{n-k-1}}{\dfrac{n!}{k!(n-k)!}p^k(1-p)^{n-k}} \\ &= \frac{n-k}{k+1}\frac{p}{1-p} \qquad k=0,1,\ldots,n \Rightarrow \\ \mathrm{P}(X=k+1) &= \frac{n-k}{k+1}\frac{p}{1-p}\mathrm{P}(X=k) \end{aligned} \tag{3.11}$$

Temos, asim, uma forma recursiva de calcular probabilidades binomiais.

Suponhamos, agora, que a probabilidade máxima ocorra em k_0. Usando o resultado acima, temos que ter

$$\begin{aligned} \frac{\mathrm{P}(X=k_0+1)}{\mathrm{P}(X=k_0)} &\le 1 \Rightarrow \frac{n-k_0}{k_0+1}\frac{p}{1-p} \le 1 \Rightarrow np-k_0p \le k_0-k_0p+1-p \Rightarrow \\ np &\le k_0+1-p \Rightarrow k_0 \ge p(n+1)-1 \end{aligned}$$

Analogamente, temos que ter

$$\begin{aligned} \frac{\mathrm{P}(X=k_0)}{\mathrm{P}(X=k_0-1)} &\ge 1 \Rightarrow \frac{(n-(k_0-1))}{k_0}\frac{p}{1-p} \ge 1 \Rightarrow np-k_0p+p \le k_0-k_0p \Rightarrow \\ k_0 &\le p(n+1) \end{aligned}$$

Resulta, então, que se k_0 é ponto de máximo, então

$$p(n+1)-1 \le k_0 \le p(n+1)$$

e como k_0 tem que ser inteiro, temos que ter

$$k_0 = [p(n+1)] \tag{3.12}$$

em que $[x]$ representa o maior inteiro menor ou igual a x. Por exemplo, se $n=7$ e $p=0,2$, temos que ter $0,2\times 8-1 \le k_0 \le 0,2\times 8$, ou seja, $0,6 \le k_0 \le 1,6$ e o ponto de máximo ocorre em $k_0 = 1$.

Note que, se $p(n+1)$ for inteiro, então a distribuição é bimodal, com moda em $k_{01} = p(n+1)$ e $k_{02} = p(n+1)-1$ pois

$$
\begin{aligned}
\frac{\mathrm{P}\left(X = p(n+1)\right)}{\mathrm{P}\left(X = p(n+1) - 1\right)} &= \frac{\mathrm{P}(X = np + p)}{\mathrm{P}(X + np + p - 1} = \frac{n - (np + p - 1)}{np + p} \cdot \frac{p}{1 - p} \\
&= \frac{n - np - p + 1}{p(n+1)} \cdot \frac{p}{1-p} = \frac{(n+1) - p(n+1)}{(n+1)(1-p)} \\
&= \frac{(n+1)(1-p}{(n+1)(1-p)} = 1
\end{aligned}
$$

Para o caso em que $p = 0,5$, esses resultados implicam que a distribuição será unimodal quando $n+1$ for ímpar e bimodal quando $n+1$ for par, ou seja, se $p = 0,5$, a distribuição é unimodal para n par e bimodal para n ímpar. Além disso, se $p = 0,5$, a distribuição será simétrica, pois

$$
\mathrm{P}(X = k) = \binom{n}{k} 0,5^k (1 - 0,5)^{n-k} = \binom{n}{n-k} 0,5^{n-k} (1 - 0,5)^{n-(n-k)} = \mathrm{P}(X = n - k)
$$

Esses resultados estão ilustrados na Figura 3.3, onde temos gráficos da distribuição binomial para diferentes valores dos parâmetros n e p. Note que a distribuição é assimétrica quando $p \neq 0,5$ e a assimetria é positiva quando $p < 0,5$ e negativa quando $p > 0,5$.

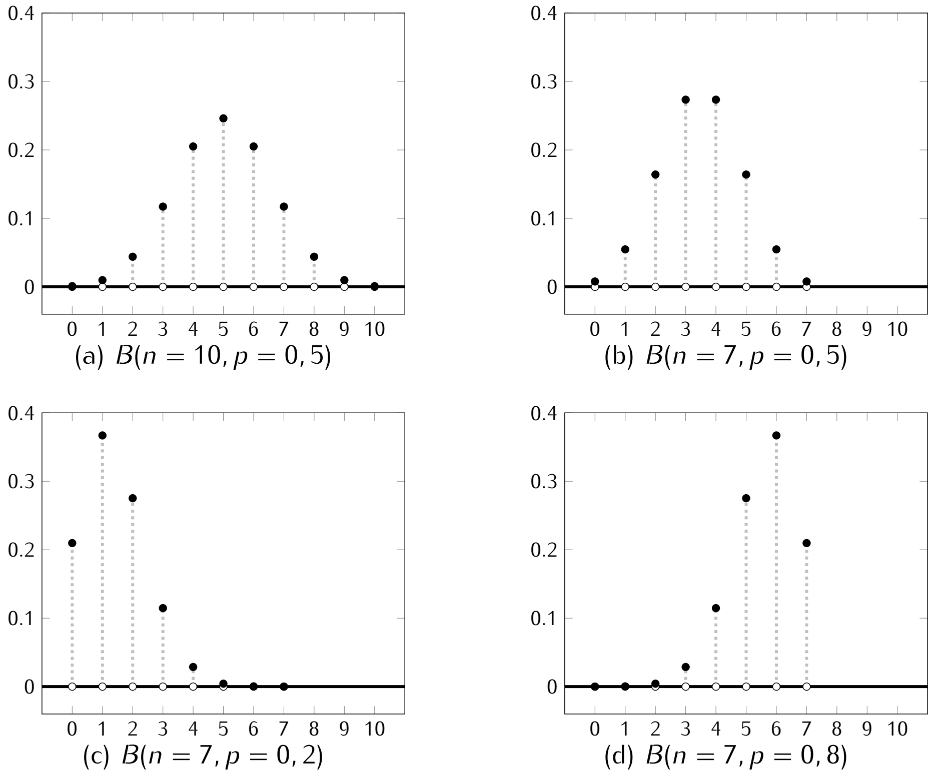

(a) $B(n = 10, p = 0,5)$ (b) $B(n = 7, p = 0,5)$ (c) $B(n = 7, p = 0,2)$ (d) $B(n = 7, p = 0,8)$

Figura 3.3 – Função de Probabilidade da Distribuição Binomial.

3.5 Distribuição Hipergeométrica

Assim como na seção anterior, nessa começar esta com um exemplo antes da formalização da nova distribuição.

Exemplo 3.17 Bolas em uma urna
Considere a situação do Exemplo 3.9, em que 3 bolas eram retiradas de uma urna composta por 4 bolas brancas e 6 bolas verdes. Naquele exemplo, as extrações eram feitas com reposição. Vamos supor, agora, que as extrações sejam feitas sem reposição. Vamos definir novamente a variável aleatória:

$$X = \text{ número de bolas brancas sorteadas.}$$

Encontre a função de probabilidade da v.a. X.

Solução:
No exemplo 3.9 as bolas eram retiradas com reposição, agora estamos supondo que as extrações sejam feitas sem reposição. O que muda? A primeira observação é a de que as probabilidades em cada extração dependem das extrações anteriores, ou seja, não temos mais independência ou probabilidades constantes. A segunda observação é que cada subconjunto de 3 bolas é igualmente provável, já que as extrações são aleatórias. O número total de subconjuntos de 3 bolas retiradas das 10 bolas da urna é $\binom{10}{3}$ e, portanto, cada subconjunto tem probabilidade $\frac{1}{\binom{10}{3}}$.

Vamos considerar, novamente, a variável aleatória X que representa o número de bolas brancas extraídas. Como há 6 bolas verdes na urna, é possível que todas as três bolas extraídas sejam verdes, isto é, $X = 0$ é o valor mínimo. No outro extremo, como há 4 brancas na urna, é possível que todas as bolas extraídas sejam brancas, ou seja, $X = 3$ é o valor máximo. É possível obter, também, 1 ou 2 bolas brancas. Logo, os valores possíveis de X são 0, 1, 2, 3. Vamos calcular a probabilidade de cada um desses valores.

- $X = 0$

  Ter 0 bola branca na amostra, significa que todas as 3 bolas são verdes. Como há 6 bolas vedes, o número de maneiras que podemos retirar 3 bolas verdes é dado por $\binom{6}{3}$. Logo,

$$P(X = 0) = \frac{\binom{6}{3}}{\binom{10}{3}}$$

- $X = 1$

  Ter 1 bola branca na amostra, significa que as outras 2 são verdes. Como há 4 bolas brancas, o número de maneiras que podemos retirar 1 bola branca é dado por $\binom{4}{1}$. Analogamente, como há 6 bolas verdes, o número de maneiras que podemos retirar 2 bolas verdes é $\binom{6}{2}$. Pelo Princípio Fundamental da Multiplicação, o número de maneiras que podemos retirar 1 bola branca *e* 2 bolas verdes é $\binom{4}{1} \times \binom{6}{2}$. Logo,

$$P(X = 1) = \frac{\binom{4}{1} \times \binom{6}{2}}{\binom{10}{3}}$$

- $X = 2$ e $X = 3$

  Analogamente,

$$P(X = 2) = \frac{\binom{4}{2} \times \binom{6}{1}}{\binom{10}{3}} \qquad \text{e} \qquad P(X = 3) = \frac{\binom{4}{3} \times \binom{6}{0}}{\binom{10}{3}}$$

Suponhamos que nossa amostra seja, agora, de 5 bolas. Como só há 4 bolas brancas, não é possivel obter uma amostra formada apenas por bolas brancas. Mas, vamos pensar, por um momento, que pudéssemos ter $X = 5$. Seguindo o raciocínio anterior, teríamos que

$$P(X = 5) = \frac{\binom{4}{5} \times \binom{6}{0}}{\binom{10}{5}} = 0$$

uma vez que $\binom{4}{5} = 0$, já que $4 < 5$, veja a Definição 3.12 de coeficiente binomial. Isso resulta numa probabilidade nula para o evento impossível $X = 5$.

♦♦

A generalização do contexto do exemplo anterior é a seguinte: temos uma população de tamanho N (a urna com as bolas) dividida em 2 classes (duas cores). A classe que estamos interessados em observar será considerada "sucessos" e a outra "fracassos". Seja r o número de indivíduos da classe sucesso entre os N indivíduos da população, logo a classe fracasso possui $N - r$ indivíduos. Dessa população vamos extrair uma amostra de tamanho n sem reposição. A variável aleatória de interesse é

$$X = \text{número de sucessos na amostra.}$$

Essa variável aleatória tem distribuição Hipergeométrica, como mostra a Definição 3.18 abaixo.

Definição 3.18 Distribuição Hipergeométrica
*Uma variável aleatória X tem **distribuição Hipergeométrica** com parâmetros N, r e n quando esta representa o número de sucessos dentro de uma amostra de tamanho n extraída, sem reposição, de uma população de tamanho N formada por r sucessos e $N - r$ fracassos.*

Vamos primeiro analisar o espaço paramétrico para os parâmetros N, r e n. Como N é o tamanho da população, então $N \in \mathbb{N}$. Como n é o tamanho da amostra retirada da população de tamanho N, então $0 \leq n \leq N$. Por último, o parâmetro r indica o número de sucessos dentro da população de tamanho N, então $0 \leq r \leq N$.

Para determinar os valores possíveis de X, isto é, $Im(X)$, temos que considerar diferentes possibilidades para a composição da população em termos dos números de sucessos e fracassos relativos ao tamanho da amostra.

- Se for possível ter uma amostra só com sucessos ou só com fracassos, isto é, se $r \geq n$ e $N - r \geq n$, então os possíveis valores de X variam de 0 (amostra formada só por fracassos) a n (amostra formada só por sucessos).

- Se for possível ter uma amostra só com sucessos, mas não só com fracassos, isto é, se $r \geq n$ e $N - r < n$, então o número máximo de fracassos na amostra é $N - r$ e, portanto, o número mínimo de sucessos é $n - (N - r)$. Logo, os valores de X variam de $n - (N - r)$ a n.

- Se for possível ter uma amostra só com fracassos, mas não só com sucessos, isto é, se $N - r > n$ e $r < n$, então o número mínimo de sucessos na amostra é 0 e o número máximo é r.

Vejamos algumas situações que ilustram os itens destacados acima.

- Se $N = 6$, $n = 3$, $r = 3 \Rightarrow n \geq r$ e $N - r = 6 - 3 = 3 \geq n = 3$. Então podemos ter uma amostra só com sucesso ou só com fracasso. Nesse caso $Im(X) = \{0, 1, 2, 3\}$. De forma geral, $Im(X) = \{0, \ldots, n\}$.

- Se $N = 6$, $n = 3$, $r = 4 \Rightarrow n \geq r$ e $N - r = 6 - 4 = 2 < n$. Então o número máximo de fracassos na amostra é $N - r = 2$, logo no mínimo temos que ter $n - (N - r) = 1$ secesso, ou seja, $Im(X) = \{1, 2, 3\}$. De forma geral, $Im(X) = \{n - (N - r), \ldots, n\}$.

- Se $N = 6$, $n = 3$, $r = 2 \Rightarrow r < n$ e $N - r = 6 - 2 = 4 \geq n$. Então o número máximo de sucessos na amostra é $r = 2$, ou seja, $Im(X) = \{0, 1, 2\}$. De forma geral, $Im(X) = \{0, \ldots, r\}$.

Vamos denotar a distribuição Hipergeométrica com parâmetros N, r e n por $X \sim hiper(N, r, n)$ ou $X \sim Hgeo(N, r, n)$. Nesse caso, se quisermos indicar que uma variável aleatória X segue a distribuição Hipergeométrica com parâmetros N, r e n podemos simplesmente escrever: $X \sim hiper(N, r, n)$ ou $X \sim Hgeo(N, r, n)$ (lê-se: a variável aleatória X tem distribuição Hipergeométrica com parâmetros N, r e n). Atenção: alguns livros usam letras diferentes para representar cada parâmetro.

Função de Probabilidade

O número total de amostras de tamanho n que podem ser extraídas de uma população de tamanho N, sem reposição, é $\binom{N}{n}$. Note que isso equivale ao número de subconjuntos de tamanho n do conjunto universo de tamanho N.

Para calcular a probabilidade de k sucessos na amostra, $\mathrm{P}(X = k)$, vamos considerar as 3 situações anteriores:

- Sucessos e fracassos suficientes: $r \geq n$ e $N - r \geq n$.

  Há $\binom{r}{k}$ maneiras de tirarmos k sucessos e $\binom{N-r}{n-k}$ maneiras de tirar os fracassos que completam a amostra. Logo,

$$\mathrm{P}(X = k) = \frac{\dbinom{r}{k}\dbinom{N-r}{n-k}}{\dbinom{N}{n}} \tag{3.13}$$

  e nesse caso, os valores de k vão de 0 a n.

- Sucessos suficientes, mas não fracassos: $r \geq n$ e $N - r < n$.

  Vimos que os valores de X variam de $n-(N-r)$ a n. Para qualquer valor $k < n-(N-r)$, o coeficiente binomial $\binom{N-r}{n-k}$ será 0, pela Definição 3.12. Por exemplo, se $k = n-(N-r)-1$, resulta

$$\binom{N-r}{n-k} = \binom{N-r}{n-[n-(N-r)-1]} = \binom{N-r}{N-r+1} = 0$$

  Então, ainda podemos usar a Equação 3.13 para calcular as probabilidades, considerando k variando de 0 a n.

- Fracassos suficientes, mas não sucessos: $N - r > n$ e $r < n$.

  Vimos que os valores X variam de 0 a r. Para qualquer valor $k > r$, o coeficiente binomial $\binom{r}{k}$ será 0, pela Definição 3.12. Então, ainda podemos usar a Equação 3.13 para calcular as probabilidades, considerando k variando de 0 a n.

Resumindo, se $X \sim hiper(N, r, n)$ então a sua função de probabilidade é definida por:

$$p_X(x) = \mathrm{P}(X = x) = \frac{\binom{r}{x}\binom{N-r}{n-x}}{\binom{N}{n}} \qquad x = 0, \ldots, n \tag{3.14}$$

Para provar que a Equação 3.14 realmente define uma função de probabilidade, note inicialmente que $\mathrm{P}(X = x) \geq 0 \ \forall \ x \in \mathbb{R}$. Além disso temos que provar que a probabilidade do espaço amostral é 1, isto é, que $\sum_{x|x \in Im(X)} \mathrm{P}(X = x) = 1$, ou seja

$$\sum_{x=0}^{n} \frac{\binom{r}{x}\binom{N-r}{n-x}}{\binom{N}{n}} = 1 \Longrightarrow \frac{\sum_{x=0}^{n}\binom{r}{x}\binom{N-r}{n-x}}{\binom{N}{n}} = 1 \Longrightarrow \sum_{k=0}^{n}\binom{r}{x}\binom{N-r}{n-x} = \binom{N}{n}$$

Então basta provar

$$\sum_{x=0}^{n}\binom{r}{k}\binom{N-r}{n-x} = \binom{N}{n}. \tag{3.15}$$

Para tal, vamos precisar do resultado apresentado na Proposição 3.19 a seguir.

Proposição 3.19 Fórmula de Euler
Sejam m, n e s inteiros tais que $s < n + m$. Então é verdade que:

$$\sum_{k=0}^{s}\binom{m}{k}\binom{n}{s-k} = \binom{m+n}{s}$$

Demonstração:
Veja que se $m = 0$ e $n = 0$ a igualdade é trivialmente satisfeita. Vamos analisar então o caso em que $m > 0$ ou $n > 0$.

Seja C um conjunto não vazio com $m + n$ elementos. Veja que C pode ser partido em dois conjuntos disjuntos com m e n elementos, respectivamente. Vamos denotar tais conjuntos por C_n e C_m, onde C_n tem n elementos, C_m tem m elementos, $C_n \cap C_m = \emptyset$ e $C_n \cup C_m = C$.

Fixando $0 \leq k \leq s$, veja que $\binom{m}{k}$ representa o número de subconjuntos de C_m com k elementos e $\binom{n}{s-k}$ representa o número de subconjuntos de C_n com $s - k$ elementos. Logo, $\binom{m}{k}\binom{n}{s-k}$ representa o número de subconjuntos de C com s elementos tais que k elementos pertencem à C_m e os outros $s - k$ elementos pertencem à C_m.

Então, $\sum_{k=0}^{s}\binom{m}{k}\binom{n}{s-k}$ representa o número de subconjuntos de C com s elementos, isto é, $\binom{m+n}{s}$.

□

Podemos ver que a Equação 3.15 nada mais é que a fórmula de Euler com $k = x$, $m = r$, $n = N - r$ e $s = n$, o que resulta que

$$\sum_{k=0}^{n}\binom{r}{k}\binom{N-r}{n-k} = \binom{r+N-r}{n} = \binom{N}{n}$$

e isso completa prova de que a Equação 3.14 realmente define uma função de probabilidade.

Esperança e Variância

Nessa seção iremos calcular a esperança e variância de $X \sim hiper(N, r, n)$. Vamos começar pelo cálculo de $\mathrm{E}(X)$.

$$\mathrm{E}(X) = \sum_{x=0}^{n} x \frac{\binom{r}{x}\binom{N-r}{n-x}}{\binom{N}{n}} \stackrel{\star}{=} \sum_{x=1}^{n} x \frac{\binom{r}{x}\binom{N-r}{n-x}}{\binom{N}{n}}$$

Na passagem $\stackrel{\star}{=}$ acontece a mudança do índice pois quando $x = 0$ o somatório se anula. Continuando,

$$\begin{aligned} \mathrm{E}(X) &= \sum_{x=1}^{n} x \frac{\binom{r}{x}\binom{N-r}{n-x}}{\binom{N}{n}} = \frac{1}{\binom{N}{n}} \sum_{x=1}^{n} x \binom{r}{x}\binom{N-r}{n-x} = \\ &= \frac{1}{\binom{N}{n}} \sum_{x=1}^{n} x \frac{r!}{x!\,(r-x)!}\binom{N-r}{n-x} = \frac{1}{\binom{N}{n}} \sum_{x=1}^{n} \cancel{x} \frac{r(r-1)!}{\cancel{x}(x-1)!\,(r-x)!}\binom{N-r}{n-x} \end{aligned}$$

Como $x \neq 0$ podemos simplificar a expressão acima e seguir com as contas abaixo.

$$\mathrm{E}(X) = \frac{1}{\binom{N}{n}} \sum_{x=1}^{n} \frac{r(r-1)!}{(x-1)!\,(r-x)!}\binom{N-r}{n-x} = \frac{r}{\binom{N}{n}} \sum_{x=1}^{n} \frac{(r-1)!}{(x-1)!\,(r-x)!}\binom{N-r}{n-x}$$

Faremos agora a seguinte troca de variável no somatório: $y = x - 1$.

$$\mathrm{E}(X) = \frac{r}{\binom{N}{n}} \sum_{y=0}^{n-1} \frac{(r-1)!}{y!\,(r-1-y)!}\binom{N-r}{n-1-y} = \frac{r}{\binom{N}{n}} \underbrace{\sum_{y=0}^{n-1} \binom{r-1}{y}\binom{N-r}{n-1-y}}_{\star}$$

Veja que na expressão destacada por $\star$ podemos aplicar a Fórmula de Euler, Proposição 3.19, com $n = N - r$, $m = r - 1$, $s = n - 1$. Assim seguimos com

$$\mathrm{E}(X) = \frac{r}{\binom{N}{n}}\binom{N-1}{n-1} = r\frac{n!\cancel{(N-n)!}}{N!}\frac{(N-1)!}{(n-1)!\cancel{(N-n)!}} = \frac{rn\cancel{(n-1)!}}{N\cancel{(N-1)!}}\frac{\cancel{(N-1)!}}{\cancel{(n-1)!}} = \frac{rn}{N}$$

Logo,

$$X \sim hiper(N, r, n) \quad \Rightarrow \quad \mathrm{E}(X) = n\frac{r}{N} \tag{3.16}$$

Vamos agora calcular $\mathrm{E}(X^2)$.

$$\mathrm{E}\left(X^2\right) = \sum_{x=0}^{n} x^2 \frac{\binom{r}{x}\binom{N-r}{n-x}}{\binom{N}{n}} \stackrel{\star}{=} \sum_{x=1}^{n} x^2 \frac{\binom{r}{x}\binom{N-r}{n-x}}{\binom{N}{n}}$$

Novamente, na passagem $\stackrel{\star}{=}$ acontece a mudança do índice pois quando $x = 0$ o somatório se anula. Continuando,

$$\mathrm{E}\left(X^2\right) = \sum_{x=1}^{n} x^2 \frac{\binom{r}{x}\binom{N-r}{n-x}}{\binom{N}{n}} = \frac{1}{\binom{N}{n}} \sum_{x=1}^{n} x^2 \binom{r}{x}\binom{N-r}{n-x} = \frac{1}{\binom{N}{n}} \sum_{x=1}^{n} x^{\cancel{2}} \frac{r(r-1)!}{\cancel{x}(x-1)!\,(r-x)!}\binom{N-r}{n-x}$$

Como $x \neq 0$ podemos simplificar a expressão acima. As contas abaixo segue já com a troca de variável $y = x - 1$ e com a retirada da constante r de dentro do somatório.

$$\begin{aligned}
\mathrm{E}(X) &= \frac{r}{\binom{N}{n}} \sum_{y=0}^{n-1} (y+1) \frac{(r-1)!}{y!\,(r-1-y)!}\binom{N-r}{n-1-y} = \frac{r}{\binom{N}{n}} \sum_{y=0}^{n-1} (y+1)\binom{r-1}{y}\binom{N-r}{n-1-y} \\
&= \frac{r}{\binom{N}{n}} \left[\sum_{y=0}^{n-1} y \binom{r-1}{y}\binom{N-r}{n-1-y} + \sum_{y=0}^{n-1} \binom{r-1}{y}\binom{N-r}{n-1-y} \right] \\
&= \frac{r\binom{N-1}{n-1}}{\binom{N}{n}} \left[\underbrace{\sum_{j=0}^{n-1} j \frac{\binom{r-1}{j}\binom{N-1-(r-1)}{n-1-j}}{\binom{N-1}{n-1}}}_{S_1} + \underbrace{\sum_{j=0}^{n-1} \frac{\binom{r-1}{j}\binom{N-1-(r-1)}{n-1-j}}{\binom{N-1}{n-1}}}_{S_2} \right]
\end{aligned}$$

Mas o primeiro somatório é a esperança de uma hipergeométrica com parâmetros $N - 1, n - 1$ e $r - 1$, ou seja, $S_1 = (n-1)\frac{r-1}{N-1}$. Já o segundo somatório é a soma das probabilidades no espaço amostral de uma hipergeométrica com os mesmos parâmetros, ou seja, $S_2 = 1$. Segue, então, que

$$\mathrm{E}\left(X^2\right) = \frac{r\binom{N-1}{n-1}}{\binom{N}{n}} \left[(n-1)\,\frac{r-1}{N-1} + 1\right] = \frac{rn}{N} \times \frac{(n-1)\,(r-1) + N - 1}{N-1}$$

e, portanto,

$$\begin{aligned}
\mathrm{Var}\,(X) &= \mathrm{E}(X^2) - \mathrm{E}(X)^2 = \frac{rn}{N} \times \frac{(n-1)\,(r-1) + N - 1}{N-1} - \frac{n^2r^2}{N^2} \\
&= \frac{rn}{N} \times \left[\frac{(n-1)\,(r-1) + N - 1}{N-1} - \frac{nr}{N}\right] = \frac{rn}{N} \times \left[\frac{nr - n - r + 1 + N - 1}{N-1} - \frac{nr}{N}\right] \\
&= \frac{rn}{N} \times \left[\frac{nr - n - r + N}{N-1} - \frac{nr}{N}\right] = \frac{rn}{N} \times \left[\frac{Nnr - nN - Nr + N^2 - Nnr + nr}{N\,(N-1)}\right] \\
&= \frac{rn}{N} \times \left[\frac{-nN - Nr + N^2 + nr}{N\,(N-1)}\right] = n\frac{r}{N}\frac{N\,(N-n) - r\,(N-n)}{N\,(N-1)} = n\frac{r}{N}\frac{(N-n)\,(N-r)}{N\,(N-1)}
\end{aligned}$$

Ou seja,

$$X \sim hiper(N, r, n) \qquad \Rightarrow \qquad \mathrm{Var}\,(X) = n\frac{r}{N}\frac{N-r}{N}\frac{N-n}{N-1} \tag{3.17}$$

Exemplo 3.20 Equipe de programadores

Entre os 16 programadores de uma empresa, 12 são do sexo masculino. A empresa decide sortear 5 programadores para fazer um curso avançado de programação. Qual é a probabilidade dos 5 sorteados serem do sexo masculino? Quantos

Solução:
Sucesso = sexo masculino. Se X = número de homens sorteados, então $X \sim hiper(16; 12; 5)$ e o problema pede

$$\Pr(X = 5) = \frac{\binom{12}{5}}{\binom{16}{5}} = \frac{12 \times 11 \times 10 \times 9 \times 8}{16 \times 15 \times 14 \times 13 \times 12} = \frac{33}{14 \times 13} = 0,181319$$

♦♦

Distribuição Binomial *versus* Distribuição Hipergeométrica

Vamos fazer agora algumas comparações entre as distribuições binomial e hipergeométrica, considerando que elas descrevem a extração de amostra de tamanho n. No contexto da binomial, a amostra é retirada *com* reposição, enquanto na hipergeométrica as extrações são feitas *sem* reposição.

A esperança da binomial é igual ao produto do tamanho da amostra pela probabilidade de sucesso; Na hipergeométrica, a esperança também é o produto do tamanho da amostra pela probabilidade de sucesso, probabilidade essa tomada apenas na primeira extração.

A variância da binomial é igual ao produto do tamanho da amostra pelas probabilidades de sucesso e fracasso. Na hipergeométrica, considerando apenas a primeira extração, a variância é igual a esse produto, mas corrigido pelo fator $\dfrac{N-n}{N-1}$.

Em pesquisas estatísticas por amostragem, normalmente lidamos com amostragem sem reposição. No entanto, os resultados teóricos sobre amostragem com reposição são bem mais simples, pois envolvem variáveis independentes); assim, costuma-se usar uma aproximação, sempre que possível. Ou seja, quando a população (tamanho N) é suficientemente grande (de modo que podemos encará-la como uma população infinita) e o tamanho da amostra é relativamente pequeno, podemos "ignorar" o fato de as extrações serem feitas sem reposição. Lembre-se que a probabilidade em extrações sucessivas são $\frac{1}{N}, \frac{1}{N-1}, \ldots, \frac{1}{N-n}$. Então, se N é "grande" e n é pequeno, temos que $N \approx N-1 \approx \cdots \approx N-n$. Nessas condições, extrações com e sem reposição podem ser consideradas como equivalentes. O termo que aparece na variância da hipergeométrica, $\frac{N-n}{N-1}$, é chamado correção para populações finitas, exatamente porque, se a população é pequena, não podemos ignorar o fato de as extrações estarem sendo feitas sem reposição.

3.6 Distribuição Geométrica

Considere as seguintes situações: (i) uma moeda com probabilidade p de cara é lançada até que apareça cara pela primeira vez; (ii) em uma população muito grande (pense na população mundial), p% das pessoas sofrem de uma rara doença desconhecida e portadores precisam ser encontrados para estudos clínicos. Quais são as semelhanças entre essas daus situações? No primeiro caso, matematicamente falando, poderíamos ter que fazer "infinitos" lançamentos. No segundo caso, "muito grande" pode ser uma aproximação para "infinitos". Em ambos os casos, temos repetições de um experimento de Bernoulli. No primeiro caso, as repetições certamente podem ser consideradas independentes. No segundo caso, também podemos assumir independência, desde que esqueçamos os fatores genéticos por um momento. Então, temos repetições independentes de um experimento de Bernoulli e estamos interessados no *número de repetições até a ocorrência do primeiro sucesso* (cara no caso da moeda, pessoa portadora no estudo epidemiológico).

Vamos, agora, formalizar a definição de tal variável e obter a sua função de probabilidade. Consideremos, então, repetições independentes de um experimento de Bernoulli com probabilidade p de sucesso. Vamos definir a variável aleatória que representa o número de repetições até a ocorrência do primeiro sucesso.

Definição 3.21 Distribuição Geométrica
*Uma variável aleatória X tem **distribuição Geométrica** com parâmetro p quando esta representa o número tentativas até a ocorrência do primeiro sucesso em consecutivos ensaios de Bernoulli independentes, cada um com probabilidade p de sucesso.*

Veja que os possíveis valores dessa variável aleatória são: 1 (primeiro sucesso na primeira repetição), 2 (primeiro sucesso na segunda repetição e, portanto, fracasso na primeira), 3 (primeiro sucesso na terceira repetição e, portanto, fracasso nas duas primeiras), etc. Logo, $Im(X) = \mathbb{N}$. Esse é um exemplo de v.a. discreta em que o espaço amostral, enumerável, é infinito.

O espaço paramétrico para o parâmetro p é o intervalo $[0, 1]$, uma vez que p é a probabilidade de sucesso em cada ensaio de Bernoulli realizado.

As características definidoras desse modelo são: (i) repetições de um mesmo experimento de Bernoulli, o que significa que em todas elas a probabilidade de sucesso (e, portanto, de fracasso) é a mesma e (ii) as repetições são independentes. No caso do lançamento de uma moeda essas hipóteses são bastante plausíveis mas no caso da doença a hipótese de independência pode não ser satisfeita; por exemplo, pode haver um componente de hereditariedade.

Vamos denotar a distribuição Geométrica com parâmetro p por $Geom(p)$. Nesse caso, se quisermos indicar que uma variável aleatória X segue a distribuição Geométrica com parâmetro p podemos simplesmente escrever: $X \sim Geom(p)$ (lê-se: a variável aleatória X tem distribuição Geométrica com parâmetro p).

Função de Probabilidade

Para calcular a probabilidade do evento $X = x$, $x = 1, 2, 3, \ldots$, devemos notar que tal evento corresponde à ocorrência de fracassos nas $x - 1$ primeiras repetições e sucesso na x-ésima repetição. Denotando por F_i e S_i a ocorrência de fracasso e sucesso na i-ésima repetição respectivamente, temos a seguinte equivalência de eventos:

$$\{X = x\} = \{F_1 \cap F_2 \cap \cdots \cap F_{x-1} \cap S_x\}$$

Como as repetições são independentes, segue que

$$\begin{aligned} \mathrm{P}(X = x) &= \mathrm{P}(F_1 \cap F_2 \cap \cdots \cap F_{x-1} \cap S_x) = \mathrm{P}(F_1) \times \mathrm{P}(F_2) \times \cdots \times \mathrm{P}(F_{x-1}) \times \mathrm{P}(S_x) = \\ &= (1-p) \times (1-p) \times \cdots \times (1-p) \times p \end{aligned}$$

Ou seja, se $X \sim Geom(p)$ a sua função de probabilidade é definida por:

$$p_X(x) = \mathrm{P}(X = x) = (1-p)^{x-1} p \qquad x = 1, 2, 3, \ldots \tag{3.18}$$

Para mostrar que a Equação 3.18 realmente define uma função de probabilidade, temos que mostrar que a soma das probabilidades, isto é, a probabilidade do espaço amostral é 1 (obviamente, $P(X = x) \geq 0$). Para isso faremos uso do resultado sobre séries geométricas apresentado a seguir.

Proposição 3.22 Série Geométrica
A série geométrica $\sum_{i=0}^{\infty} r^i$ *diverge se* $|r| \geq 1$ *e converge se* $|r| < 1$. *Temos também que*

$$\sum_{i=0}^{\infty} r^i = \frac{1}{1-r} \qquad se\ |r| < 1.$$

Demonstração:
A demonstração encontra-se na Seção A.1 do Apêndice A.

□

Vamos às contas. Queremos mostrar que $\sum_{x=1}^{\infty} P(X = x) = 1$.

$$\sum_{x=1}^{\infty} P(X = x) = \sum_{x=1}^{\infty} (1-p)^{x-1} p = p \sum_{x=1}^{\infty} (1-p)^{x-1}.$$

Fazendo $y = x - 1$, temos que quando $x = 1 \Rightarrow y = 0$ e quando $x = \infty \Rightarrow y = \infty$. Portanto,

$$\sum_{x=1}^{\infty} P(X = x) = p \sum_{y=0}^{\infty} (1-p)^y$$

Usando o resultado da Proposição 3.22 com $r = 1 - p$, obtém-se que:

$$\sum_{x=1}^{\infty} P(X = x) = p \frac{1}{1-(1-p)} = 1$$

como queríamos mostrar.

Esperança e Variância

Nessa seção iremos calcular a esperança e variância de $X \sim Geom(p)$. Vamos começar pelo cálculo de $E(X)$.

$$E(X) = \sum_{x=1}^{\infty} xp(1-p)^{x-1} = \sum_{x=1}^{\infty} (x-1+1)p(1-p)^{x-1} = \sum_{x=1}^{\infty} (x-1)p(1-p)^{x-1} + \underbrace{\sum_{x=1}^{\infty} p(1-p)^{x-1}}_{S_1}$$

Veja que $S_1 = 1$, pois trata-se da soma da função de probabilidade para todos os valores em $Im(X)$ para $X \sim Geom(p)$. Seguimos então fazendo a seguinte troca de variável: $y = x - 1$.

$$E(X) = \sum_{y=0}^{\infty} yp(1-p)^y + 1 \overset{\star}{=} \sum_{y=1}^{\infty} yp(1-p)^y + 1 = \sum_{y=1}^{\infty} yp(1-p)(1-p)^{y-1} + 1 = (1-p) \underbrace{\sum_{y=1}^{\infty} yp(1-p)^{y-1}}_{S_2} + 1$$

A passagem $\stackrel{\star}{=}$ é justificada pois quando $y = 0$ o somatório se anula. Veja que S_2 é justamente $\mathrm{E}(X)$ para $X \sim Geom(p)$, isto é, aquilo que estamos calculando. Então podemos escrever.

$$\mathrm{E}(X) = (1-p)\,\mathrm{E}(X) + 1 \quad \Rightarrow \quad (1-1+p)\,\mathrm{E}(X) = 1 \quad \Rightarrow \quad \mathrm{E}(X) = \frac{1}{p}$$

Logo,

$$X \sim Geom(p) \qquad \Rightarrow \qquad \mathrm{E}(X) = \frac{1}{p} \tag{3.19}$$

Vamos calcular agora $\mathrm{E}(X^2)$.

$$\begin{aligned}
\mathrm{E}(X^2) &= \sum_{x=1}^{\infty} x^2 p(1-p)^{x-1} = \sum_{x=1}^{\infty} (x-1+1)^2 p(1-p)^{x-1} \\
&= \sum_{x=1}^{\infty} ((x-1)^2 + 2(x-1) + 1) p(1-p)^{x-1} \\
&= \underbrace{\sum_{x=1}^{\infty} (x-1)^2 p(1-p)^{x-1}}_{S_1} + \underbrace{\sum_{x=1}^{\infty} 2(x-1)p(1-p)^{x-1}}_{S_2} + \underbrace{\sum_{x=1}^{\infty} p(1-p)^{x-1}}_{S_3}
\end{aligned}$$

Veja primeiro que $S_3 = 1$. Vamos calcular S_1 e S_2 separadamente.

$$\begin{aligned}
S_2 &= \sum_{x=1}^{\infty} 2(x-1)p(1-p)^{x-1} = 2\left(\sum_{x=1}^{\infty} xp(1-p)^{x-1} - \sum_{x=1}^{\infty} p(1-p)^{x-1}\right) \\
&= 2\,(\mathrm{E}(X) - 1) = 2\left(\frac{1}{p} - 1\right) = 2\frac{(1-p)}{p}
\end{aligned}$$

Para seguir com as contas de S_1 faremos a troca de variável $y = x - 1$.

$$\begin{aligned}
S_1 &= \sum_{x=1}^{\infty} (x-1)^2 p(1-p)^{x-1} = \sum_{y=0}^{\infty} y^2 p(1-p)^y = \sum_{y=1}^{\infty} y^2 p(1-p)^y \\
&= \sum_{y=1}^{\infty} y^2 p(1-p)(1-p)^{y-1} = (1-p)\sum_{y=1}^{\infty} y^2 p(1-p)^{y-1} = (1-p)\,\mathrm{E}(X^2)
\end{aligned}$$

Assim seguimos com,

$$\begin{aligned}
\mathrm{E}(X^2) &= (1-p)\,\mathrm{E}(X^2) + 2\tfrac{(1-p)}{p} + 1 \\
(1-1+p)\,\mathrm{E}(X^2) &= \tfrac{2-2p+p}{p} \\
p\,\mathrm{E}(X^2) &= \tfrac{2-p}{p} \\
\mathrm{E}(X^2) &= \tfrac{2-p}{p^2}.
\end{aligned}$$

Do resultado acima segue que

$$\mathrm{Var}(X) = \mathrm{E}(X^2) - [\mathrm{E}(X)]^2 = \frac{2-p}{p^2} - \frac{1}{p^2} = \frac{1-p}{p^2}.$$

Logo,

$$X \sim Geom(p) \qquad \Rightarrow \qquad \mathrm{Var}(X) = \frac{1-p}{p^2} \tag{3.20}$$

Exemplo 3.23 Tiro ao alvo
Um atirador acerta na mosca do alvo, 20% dos tiros. Qual a probabilidade de ele acertar na mosca pela primeira vez no 10º tiro? E quantos tiros serão dados em média até que o atirador acerte na mosca?

Solução:
Podemos pensar os tiros como experimentos independentes de Bernoulli (acerta ou não acerta). A probabilidade de sucesso (acertar no alvo) é $p = 0,20$. Estamos querendo o número de tiros até o primeiro acerto e calcular a probabilidade de que esse número seja 10. Seja X = número de tiros até primeiro acerto. Então, $X \sim Geom(0,20)$ e $\mathrm{P}(X = 10) = 0,8^9 \times 0,2 = 0,02684$.

Como já identificamos $X \sim Geom(0,20)$, para calcular o número médio de tiros que o atirador fará basta usar o resultado da Equação 3.19 e calcular $\mathrm{E}(X)$.

$$\mathrm{E}(X) = \frac{1}{p} = \frac{1}{0,20} = 5.$$

♦♦

Exemplo 3.24 Lançamento de dado
Joga-se um dado equilibrado. Qual é a probabilidade de serem necessários 10 lançamentos até a primeira ocorrência de um seis?

Solução:
Nesse caso, sucesso é a ocorrência de face seis. Logo, P(sucesso) $= p = \frac{1}{6}$ e P(fracasso) $= 1 - p = \frac{5}{6}$. Seja X = número de lançamentos até o primeiro seis. Então, $X \sim geom\left(\frac{1}{6}\right)$ e o problema pede $\mathrm{P}(X = 10) = \left(\frac{5}{6}\right)^9 \left(\frac{1}{6}\right) = 0,03230$.

♦♦

Forma Alternativa da Distribuição Geométrica

Em vez de definir a variável geométrica como o "número de repetições até primeiro sucesso", podemos usar a seguinte definição alternativa:

$$Y = \text{"número de fracassos antes do primeiro sucesso"}$$

Neste caso, a distribuição de Y é

$$\mathrm{P}(Y = k) = (1 - p)^k p \qquad k = 0, 1, 2, \ldots \tag{3.21}$$

Em termos da variável geométrica X apresentada anteriormente na Definição 3.21, podemos ver que

$$Y = X - 1$$

e essa distribuição de probabilidade é, às vezes, chamada distribuição geométrica deslocada (em inglês, *shifted geometric distribution*). Usando essa relação entre X e Y e conhecendo as propriedades de esperança e variância apresentadas nas Proposições 2.26 e 2.34, podemos rapidamente calcular $\mathrm{E}(Y)$ e $\mathrm{Var}(Y)$:

$$\begin{aligned}
\mathrm{E}(Y) &= \mathrm{E}(X) - 1 = \frac{1}{p} - 1 = \frac{1-p}{p} \\
\mathrm{Var}(Y) &= \mathrm{Var}(X) = \frac{1-p}{p^2}
\end{aligned}$$

Propriedade de Falta de Memória

Nessa seção vamos apresentar o conceito de falta de memória de uma variável aleatória discreta. Em seguida vamos mostrar que a distribuição geométrica segue tal propriedade.

Definição 3.25 Propriedade de Falta de Memória - v.a. Discreta
Seja X uma variável aleatória discreta e m e n números reais quaisquer. Se,

$$P(X = m + n \mid X > m) = P(X = n)$$

*então dizemos que X tem a propriedade de **falta de memória**.*

Queremos mostrar nessa seção que uma variável aleatória com distribuição geométrica tem a propriedade de falta de memória. Seja $X \sim geo(p)$ e $x \in \grave{I}m(X)$, isto é, x é um inteiro positivo. Então,

$$\begin{aligned} P(X > x) &= 1 - P(X \leq x) = 1 - \sum_{i=1}^{x} P(X = i) = 1 - \sum_{i=1}^{x} p(1-p)^{i-1} \\ &= 1 - p\sum_{i=1}^{x}(1-p)^{i-1} \stackrel{\star}{=} 1 - p\underbrace{\sum_{j=0}^{x-1}(1-p)^{j}}_{S_1} \end{aligned}$$

Na passagem $\stackrel{\star}{=}$ foi feita a seguinte troca de variável: $j = i - 1$. O somatório S_1 é a soma até o termo $x - 1$ de uma progressão geométrica, e, de acordo com a Equação A.2 do Apêndice A, concluímos que $S_1 = \dfrac{1 - (1-p)^{x-1+1}}{p} = \dfrac{1-(1-p)^x}{p}$. Continuando,

$$P(X > x) = 1 - p\left(\frac{1-(1-p)^x}{1-(1-p)}\right) = 1 - (1 - (1-p)^x) = (1-p)^x.$$

Agora já é possível mostrar que $X \sim geo(p)$ tem a propriedade da falta de memória. Para simplificar as contas, suponha a e δ inteiros positivos, então

$$P(X = m + n \mid X > m) = \frac{P(X = m+n)}{P(X > m)} = \frac{p(1-p)^{m+n-1}}{(1-p)^m} = p(1-p)^{n-1} = P(X = n).$$

Na prática isso significa que se já foram realizados m ensaios de Bernoulli sem nenhum sucesso, a probabilidade do primeiro sucesso ocorrer no $m + n$-ésimo ensaio é exatamente a probabilidade do primeiro sucesso ocorrer no n-ésimo ensaio, contado desde o início do experimento. Ou seja, se jogamos uma moeda 5 vezes a não saiu nenhuma cara, não significa que tem mais chance de sair cara no próximo lançamento do que no 1º lançamento do experimento.

3.7 Distribuição Binomial Negativa

Consideremos novamente repetições independentes de um experimento de Bernoulli com probabilidade p de sucesso. Vamos considerar agora uma generalização da v.a. geométrica, no seguinte sentido: em vez de fixarmos o primeiro sucesso, vamos fixar um número qualquer r de sucessos, ou seja, definimos

$$X = \text{número de repetições necessárias até a obtenção do } r\text{-ésimo sucesso, } r \geq 1.$$

Note que $r = 1$ corresponde à distribuição geométrica.

Definição 3.26 Distribuição Binomial Negativa
Uma variável aleatória X tem ***distribuição Binomial Negativa*** *com parâmetros r e p quando esta representa o número tentativas até a ocorrência do r-ésimo sucesso em consecutivos ensaios de Bernoulli independentes, cada um com probabilidade p de sucesso.*
OBS: Essa distribuição é também conhecida como distribuição de Pascal.

Para definir os possíveis valores de X, isto é, $Im(X)$, devemos notar que para ter r sucessos, são necessários no mínimo r repetições. Logo, $Im(X) = r, r+1, r+2, \ldots$. Esse é mais um exemplo de v.a. discreta cuja imagem é um conjunto infinito.

Veja que o parâmetro r pode assumir qualquer valor inteiro positivo, logo seu espaço paramétrico é o conjunto $\mathbb{N}^*$. Já para o parâmetro p, o espaço paramétrico é o intervalo $[0, 1]$, uma vez que p é a probabilidade de sucesso em cada ensaio de Bernoulli realizado

Vamos denotar a distribuição Geométrica com parâmetros r e p por $BinNeg(r, p)$. Nesse caso, se quisermos indicar que uma variável aleatória X segue a distribuição Binomial Negatova com parâmetros r e p podemos simplesmente escrever: $X \sim BinNeg(r, p)$ (lê-se: a variável aleatória X tem distribuição Binomial Negativa com parâmetros r e p).

Função de Probabilidade

O evento $\{X = k\}$ indica que foram necessárias k repetições para obter r sucessos e, portanto, $k - r$ fracassos. Pela definição da variável, a última repetição resultou em sucesso e os outros $r - 1$ sucessos podem estar em quaisquer das $k - 1$ posições restantes (veja Figura 3.4).

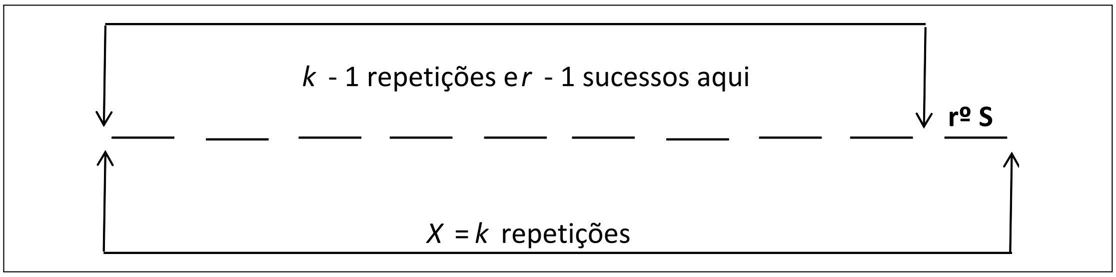


Figura 3.4 – Ilustração dos resultados da binomial negativa

Uma sequência possível de resultados é ter os $r - 1$ sucessos nas primeiras posições, os $k - r$ fracassos nas posições seguintes e o último sucesso na última posição:

$$S_1, \ldots, S_{r-1}, F_r, \ldots, F_{k-1}, S_k$$

A probabilidade de tal sequência é dada pelo produto das probabilidades, já que as repetições são independentes, isto é:

$$
\begin{gathered}
\mathrm{P}\left(S_1 \cap \ldots \cap S_{r-1} \cap F_r \cap \ldots F_{k-1} \cap S_k\right) = \\
\mathrm{P}\left(S_1\right) \times \cdots \times \mathrm{P}\left(S_{r-1}\right) \times \mathrm{P}\left(F_r\right) \times \cdots \times \mathrm{P}\left(F_{k-1}\right) \times \mathrm{P}\left(S_k\right) = \\
p \times \cdots \times p \times (1-p) \times \cdots \times (1-p) \times p = p^r\,(1-p)^{k-r}
\end{gathered}
$$

Mas existem $\binom{k-1}{r-1}$ maneiras de arrumar $r-1$ sucessos em $k-1$ posições e as sequências resultantes têm todas a probabilidade acima. Como elas constituem eventos mutuamente exclusivos, a probabilidade da união delas, que é $\mathrm{P}\left(X=k\right)$, é a soma das probabilidades, ou seja:

$$
\mathrm{P}\left(X=k\right) = p^r\,(1-p)^{k-r} + p^r\,(1-p)^{k-r} + \cdots + p^r\,(1-p)^{k-r}
$$

em que o número de parcelas é $\binom{k-1}{r-1}$. Logo

$$
\mathrm{P}\left(X=k\right) = \binom{k-1}{r-1} p^r (1-p)^{k-r} \qquad k \geq r
$$

Ou seja, se $X \sim BinNeg(p)$ a sua função de probabilidade é definida por:

$$
p_X(x) = \mathrm{P}(X=x) = \binom{x-1}{r-1} p^r (1-p)^{x-r} \qquad x \geq r \tag{3.22}
$$

Como nas seções anteriores, queremos agora mostrar que a Equação 3.22 realmente define uma função de probabilidade. Como $\mathrm{P}\left(X=x\right) \geq 0$, para todo $x \in \mathbb{R}$, fica faltando apenas mostrar que

$$
\sum_{x=r}^{\infty} \mathrm{P}\left(X=x\right) = \sum_{x=r}^{\infty} \binom{x-1}{r-1} p^r (1-p)^{x-r} = 1.
$$

Vamos às contas. Fazendo $x-r=j$, temos que $x=r+j$ e $x=r \Rightarrow j=0$. Logo

$$
\sum_{x=r}^{\infty} \mathrm{P}\left(X=x\right) = \sum_{j=0}^{\infty} \binom{r+j-1}{r-1} p^r (1-p)^j = p^r \sum_{j=0}^{\infty} \binom{r-1+j}{r-1} (1-p)^j
$$

Usando o resultado da Equação A.11) do Apêndice A.1, com $k=r-1$ e $r=1-p$, obtemos que:

$$
\sum_{x=r}^{\infty} \mathrm{P}\left(X=x\right) = p^r \times \frac{1}{\left[1-(1-p)\right]^{r-1+1}} = 1
$$

Na distribuição binomial negativa, o número de repetições do experimento de Bernoulli é variável e o número de sucessos é um número fixo (pré-determinado); note o contraste com a distribuição binomial, em que o número de repetições é fixo e a variável de interesse é o número de sucessos.

Parâmetro	Distribuição	
	Binomial	Binomial negativa
Número de repetições	fixo	variável aleatória
Número de sucessos	variável aleatória	fixo

Exemplo 3.27 Fabricação de peças

Deseja-se produzir 5 peças boas, em uma máquina que dá 20% de peças defeituosas. Qual é a probabilidade de ser necessário fabricar 8 peças para se conseguir as 5 peças boas?

Solução:
Seja X = número de peças fabricadas até a obtenção de 5 boas (sucesso). Temos que P(peça boa) = $0,80$ e P(peça defeituosa) = $0,20$. Logo, $X \sim BinNeg\,(5; 0,80)$. O problema pede $P(X = 8) = \binom{7}{4}(0,80)^5(0,20)^3 = 0,0917504$.

♦♦

Esperança e Variância

Nessa seção iremos calcular a esperança e variância de $X \sim BinNeg(r,p)$. Antes disse veja na Proposição 3.28 a seguir um resultado que será usado tanto na conta de $E(X)$ quanto na conta de $E(X^2)$.

Proposição 3.28
Sejam k e r número inteiros positivos. Então,

$$r\binom{k}{r} = k\binom{k-1}{r-1}.$$

Demonstração:
Veja que se $k < r$ ambos os lados da igualdade são iguais a 0, logo a afirmação é verdadeira. Vamos então tratar o caso em que $k \geq x$.

$$r\binom{k}{r} = r\frac{k!}{r!(k-r)!} = \cancel{r}\frac{k(k-1)!}{\cancel{r}(r-1)!(k-r)!} = k\frac{(k-1)!}{(r-1)!(k-r)!} = k\binom{k-1}{r-1}.$$

□

Vamos começar pelo cálculo de $E(X)$.

$$E(X) = \sum_{x=r}^{\infty} x\binom{x-1}{r-1}p^r(1-p)^{x-r} \overset{\star}{=} \sum_{x=r}^{\infty} r\binom{x}{r}p^r(1-p)^{x-r} = \frac{r}{p}\sum_{x=r}^{\infty}\binom{x}{r}p^{r+1}(1-p)^{x-r}$$

Na passagem $\overset{\star}{=}$ foi aplicado o resultado da Proposição 3.28. Faça agora $y = x+1$.

$$E(X) = \frac{r}{p}\underbrace{\sum_{y=r+1}^{\infty}\binom{y-1}{r}p^{r+1}(1-p)^{y-1-r}}_{S_1} = \frac{r}{p}$$

Veja que $S_1 = 1$ pois S_1 é a soma da função de probabilidade de $Y \sim BinNeg(r+1,p)$ para todos $y \in Im(Y)$. Chegamos então no seguinte resultado,

$$X \sim BinNeg(r,p) \qquad \Rightarrow \qquad E(X) = \frac{r}{p} \tag{3.23}$$

Veja que se $r = 1$ temos a esperança de uma geométrica, veja Equação 3.19.

Vamos agora calcular $E(X^2)$ para depois obtermos $Var(X)$.

$$E(X^2) = \sum_{x=r}^{\infty} x^2\binom{x-1}{r-1}p^r(1-p)^{x-r} \overset{\star}{=} \sum_{x=r}^{\infty} xr\binom{x}{r}p^r(1-p)^{x-r} = \frac{r}{p}\sum_{x=r}^{\infty}x\binom{x}{r}p^{r+1}(1-p)^{x-r}$$

Na passagem $\overset{\star}{=}$ foi aplicado o resultado da Proposição 3.28. Faça agora $y = x+1$.

$$E(X^2) = \frac{r}{p}\underbrace{\sum_{y=r+1}^{\infty}(y-1)\binom{y-1}{r}p^{r+1}(1-p)^{y-1-r}}_{S_2}.$$

Veja que $S_2 = \mathrm{E}(Y-1)$, com $Y \sim BinNeg(r+1,p)$. Usando o resultado da Equação 3.23 e as propriedades de linearidade do valor esperado, apresentadas na Proposição 2.26, concluímos que

$$S_1 = \mathrm{E}(Y-1) = \mathrm{E}(Y) - 1 = \frac{r+1}{p} - 1 \Rightarrow \mathrm{E}(X^2) = \frac{r}{p}\left(\frac{r+1}{p} - 1\right) = \frac{r^2+r-rp}{p^2}.$$

Segue então que,

$$\mathrm{Var}(X) = \mathrm{E}(X^2) - \mathrm{E}(X)^2 = \frac{r^2+r-rp}{p^2} - \frac{r^2}{p^2} = \frac{r-rp}{p^2} = r\frac{1-p}{p^2}$$

Com isso chegamos à expressão da variância de uma v.a. Binomial Negativa.

$$X \sim BinNeg(r,p) \qquad \Rightarrow \qquad Var(X) = r\frac{1-p}{p^2} \tag{3.24}$$

Novamente, veja que se $r=1$ temos a variância de uma geométrica, veja Equação 3.20.

Forma Alternativa da Distribuição Binomial Negativa

Assim como no caso da distribuição geométrica, podemos definir a variável binomial negativa como

$$Y = \text{“número de fracassos antes do } r\text{-ésimo sucesso”} \tag{3.25}$$

Isso significa que a última repetição corresponde ao $r-$ésimo sucesso e antes dele temos k fracassos e $r-1$ sucessos. Logo, antes do último sucesso, temos um total de $k+r-1$ repetições. Veja a Figura 3.5.

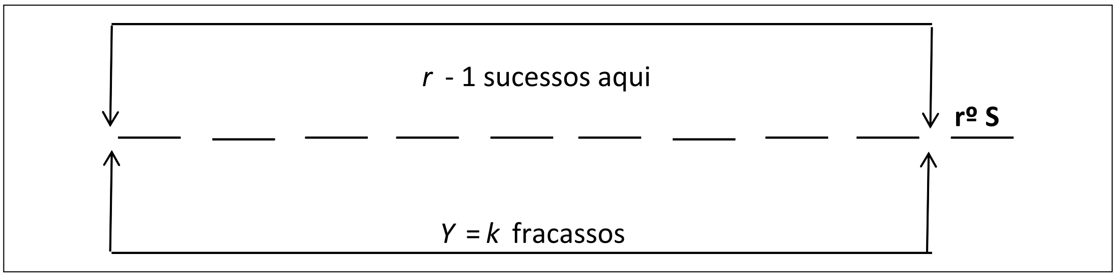


Figura 3.5 – Ilustração dos resultados da binomial negativa deslocada

Resulta, então, que

$$\mathrm{P}(Y=k) = \binom{k+r-1}{r-1} p^r (1-p)^k \qquad k = 0, 1, 2, \ldots \tag{3.26}$$

Note a seguinte relação: se X tem distribuição dada pela Equação 3.22, então

$$\begin{aligned} &Y = k \Rightarrow X = k + r - 1 + 1 = k + r \therefore X = Y + r \Rightarrow \\ &Y = X - r \end{aligned} \tag{3.27}$$

Com a forma alternativa, temos

$$\begin{aligned} \mathrm{E}(Y) &= \mathrm{E}(X-r) = \frac{r}{p} - r = \frac{r(1-p)}{p} \\ \mathrm{Var}(Y) &= \mathrm{Var}(X-r) = \mathrm{Var}(X) = \frac{r(1-p)}{p^2} \end{aligned}$$

Por que binomial negativa?

Quando definimos o número binomial $\binom{n}{k}$, supusemos que $n > 0$ e $0 < k < n$ e vimos que

$$\binom{n}{k} = \frac{n!}{k!(n-k)!} = \frac{n(n-1)\cdots(n-k+1)(n-k)!}{k!(n-k)!} = \frac{n(n-1)\cdots(n-k+1)}{k!}$$

De forma análoga, definimos

$$\binom{-n}{k} = \frac{(-n)(-n-1)\cdots(-n-k+1)}{k!}$$

que pode ser escrito como

$$\begin{aligned}\binom{-n}{k} &= \frac{[(-1)(n)][(-1)(n+1)]\cdots[(-1)(n+k-1)]}{k!} \\ &= \frac{(-1)^k n(n+1)\cdots(n+k-1)}{k!}\end{aligned}$$

Vamos, agora, analisar a expressão (3.26):

$$\begin{aligned}\mathrm{P}(Y=k) &= \binom{k+r-1}{r-1}p^r(1-p)^k = \binom{k+r-1}{k}p^r(1-p)^k = \frac{(k+r-1)!}{k!(r-1)!}p^r(1-p)^k \\ &= \frac{(k+r-1)(k+r-2)\cdots(k+r-k+1)(k+r-k)(k+r-k-1)!}{k!(r-1)!}p^r(1-p)^k \\ &= \frac{(k+r-1)(k+r-2)\cdots(k+r-k+1)(k+r-k)}{k!}p^r(1-p)^k \\ &= \frac{r(r+1)\cdots(r+k-2)(r+k-1)}{k!}p^r(1-p)^k \\ &= \frac{[(-1)(-r)][(-1)(-r-1)]\cdots[(-1)(-r-k+1]}{k!}p^r(1-p)^k \\ &= (-1)^k\binom{-r}{k}p^r(1-p)^k\end{aligned}$$

Logo, uma outra forma de escrever a distribuição binomial negativa é

$$\mathrm{P}(Y=k) = (-1)^k\binom{-r}{k}p^r(1-p)^k \qquad k = 0, 1, 2, \ldots$$

daí o nome binomial negativa. Note que essa distribuição corresponde à variável $Y =$"número de fracassos antes do r-ésimo sucesso".

3.8 Distribuição de Poisson

Suponhamos que estejamos observando um determinado fenômeno de interesse por um certo período de tempo de comprimento $T = 1$ com o interesse de contar o número de vezes X que um determinado evento ocorre.

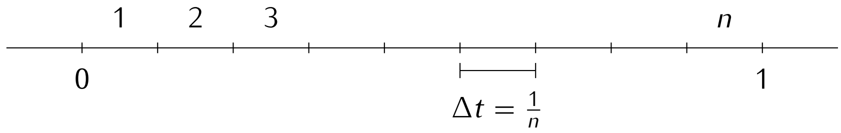
Vamos fazer as seguintes hipóteses sobre a forma de ocorrência desse evento:

H_1) Em um intervalo de tempo suficientemente curto, apenas 0 ou 1 evento ocorre, ou seja, 2 ou mais ocorrências não podem acontecer simultaneamente. Então, em cada um desses intervalos temos um *experimento de Bernoulli*.

H_2) A probabilidade de exatamente 1 ocorrência nesse pequeno intervalo de tempo, de comprimento Δt, é proporcional a esse comprimento, ou seja, é $\lambda\Delta t$. Logo, a probabilidade de nenhuma ocorrência é $1 - \lambda\Delta t$.

H_3) As ocorrências em intervalos pequenos e disjuntos são experimentos de Bernoulli *independentes*.

Estamos interessados na v.a. X = número de ocorrências do evento no intervalo $(0, 1]$. Particionando esse intervalo em n pequenos subintervalos de comprimento Δt, temos que o número total de ocorrências será a soma do número de ocorrências em cada subintervalo.



Mas em cada subintervalo podemos aplicar as hipóteses H_1, H_2 e H_3. Logo, X é uma variável binomial com parâmetros $n = \dfrac{1}{\Delta t}$ (note que $\Delta t = 1/n$) e probabilidade de sucesso $p = \lambda\Delta t$ pela hipótese H_2. Veja que $np = \lambda$. Então, para $x = 0, 1, 2, \dots, n$ temos que:

$$
\begin{aligned}
\mathrm{P}(X = x) &= \binom{n}{x} (\lambda\Delta t)^x (1 - \lambda\Delta t)^{n-x} = \binom{n}{x} \left(\frac{\lambda}{n}\right)^x \left(1 - \frac{\lambda}{n}\right)^{n-x} \\
&= \frac{n!}{x!(n-x)!} \times \frac{1}{n^x} \times \lambda^x \times \left(1 - \frac{\lambda}{n}\right)^n \times \left(1 - \frac{\lambda}{n}\right)^{-x} \\
&= \frac{n(n-1)\cdots(n-x+1)(n-x)!}{(n-x)!} \times \frac{1}{n^x} \times \frac{\lambda^x}{x!} \times \left(1 - \frac{\lambda}{n}\right)^n \times \left(1 - \frac{\lambda}{n}\right)^{-x} \\
&= \frac{n(n-1)\cdots(n-x+1)}{n^x} \times \frac{\lambda^x}{x!} \times \left(1 - \frac{\lambda}{n}\right)^n \times \left(1 - \frac{\lambda}{n}\right)^{-x} = \\
&= \frac{n}{n} \times \frac{n-1}{n} \times \cdots \times \frac{n-x+1}{n} \times \frac{\lambda^x}{x!} \times \left(1 - \frac{\lambda}{n}\right)^n \times \left(1 - \frac{\lambda}{n}\right)^{-x}
\end{aligned}
$$

Consideremos, agora, a situação em que $\Delta t \to 0$, ou equivalentemente, $n \to \infty$. Nesse caso $p \to 0$, uma vez que $p = \lambda\Delta t$. Vamos supor que $\lambda = np$ não diverge e nem converge para 0, ou seja, permanece uma constante não nula quando $n \to \infty$. Nesse caso, a variável aleatória X pode assumir qualquer valor inteiro não negativo e

$$
\begin{aligned}
\lim_{n\to\infty} \mathrm{P}(X = x) &= \lim_{n\to\infty} \left[\frac{n}{n} \times \frac{n-1}{n} \times \cdots \times \frac{n-x+1}{n} \times \frac{\lambda^x}{x!} \times \left(1 - \frac{\lambda}{n}\right)^n \times \left(1 - \frac{\lambda}{n}\right)^{-x}\right] \\
&= 1 \times 1 \times \cdots \times 1 \times \frac{\lambda^x}{x!} \times \lim_{n\to\infty} \left(1 - \frac{\lambda}{n}\right)^n \times 1 \stackrel{\star}{=} \frac{\lambda^x}{x!} e^{-\lambda}
\end{aligned}
$$

Na passagem $\stackrel{\star}{=}$ fêz-se uso da Equação A.24 do Apêndice A.6. Temos, assim, uma aproximação para a função de probabilidade Binomial. Ou seja, se $X \sim B(n, p)$ com n muito grande, $p \approx 0$ e $\lambda = np$ limitado, isto é, um número real, então

$$\mathrm{P}(X = x) \approx \frac{\lambda^x}{x!} e^{-\lambda}.$$

A expressão acima também é uma função de probabilidade, como veremos mais a frente. A variável aleatória com essa função de probabilidade leva o nome de Poisson, como mostra a Definição 3.29 a seguir.

Definição 3.29 Distribuição de Poisson
Diz-se que uma variável aleatória X tem ***distribuição de Poisson*** *com parâmetro λ se sua função de probabilidade é dada por*

$$P(X = x) = p_X(x) = \frac{\lambda^x}{x!}e^{-\lambda} \qquad x = 0, 1, 2, \ldots$$

O único parâmetro de uma variável aleatória de Poisson é o λ, que pode assumir qualquer valor real positivo. Veja que, de acordo com a hipótese H_2, a probabilidade de ocorrer uma vez o evento em questão dentro de um intervalo de comprimento Δt é $\lambda\Delta t$. Então, o espaço paramétrico para λ é $\mathbb{R}^+$.

Veja que X representa o número de ocorrências de um certo evento no intervalo $(0, 1]$. Então X pode assumir qualquer valore natural, pois podemos ter 0, 1, 2, ... ocorrências do evento. Assim $Im(X) = \mathbb{N}$. Esse é mais um exemplo de v.a. discreta cuja imagem é um conjunto infinito.

Vamos denotar a distribuição de Poisson com parâmetro λ por $Poi(\lambda)$. Nesse caso, se quisermos indicar que uma variável aleatória X segue a distribuição de Poisson com parâmetros λ podemos simplesmente escrever: $X \sim Poi(\lambda)$ (lê-se: a variável aleatória X tem distribuição de Poisson com parâmetro λ).

Relação entre Binomial e Poisson

A função de probabilidade de uma variável aleatória Binomial converge para a função de probabilidade de uma Poisson sempre que $n \to \infty$, $p \to 0$ e $np \to \lambda$, com λ constante não nula, isto é, np não diverge e nem converge para 0. Na prática isso significa que se $X \sim B(n, p)$ com n muito grande, maior que 100 em geral, e $p \approx 0$ então podemos usar a seguinte aproximação

$$P(X = x) \approx \frac{\lambda^x}{x!}e^{-\lambda} \qquad \text{, com } \lambda = np. \tag{3.28}$$

Se tais condições não forem satisfeitas, isto é, se n não for grande ou se p não for tão pequeno, a aproximação não será tão boa.

Exemplo 3.30 Aproximação da Binomial pela Poisson
Sabe-se que 5% das lampadas pisca-pisca para árvore de natal são produzidas com defeito. O setor de qualidade de um fabricante seleciona uma amostra com 150 lampadas pisca-pisca, vindas direto da linha de produção. Qual a probabilidade da amostra recolhida pelo setor de qualidade conter mais de 10 lampadas com defeito?

Solução:
Defina X = como o número de lampadas pisca-pisca com defeito dentro da amostra de 150 lampadas. Veja que $X \sim B(n = 150, p = 0,05)$. Queremos encontrar $P(X \geq 10)$. Para isso é mais fácil encontrar $P(X < 10)$ e depois fazer $P(X \geq 10) = 1 - P(X < 10)$.

Vamos começar fazendo as contas exatas.

$$\begin{aligned} P(X < 10) &= \sum_{x=0}^{9} \binom{150}{x} 0,05^x 0,95^{52-x} \\ &= \binom{150}{0} 0,05^0 0,95^{52-0} + \ldots + \binom{150}{8} 0,05^8 0,95^{52-8} + \binom{150}{9} 0,05^9 0,95^{52-9} \end{aligned}$$

Veja que as combinações da conta acima são bem custosas de serem encontradas. Por exemplo, para chegar à resposta precisamos encontrar

$$\binom{150}{9} = \frac{150!}{141!9!} = \frac{150 \times 149 \times \ldots \times 142}{9 \times 8 \times \ldots \times 1}$$

Com a ajuda de um computador chegamos à resposta final $P(X < 10) = 0,7808835$.

Vejamos agora como usar a aproximação da Binomial pela Poisson para encontrar a resposta do problema. Para isso vamos usar a expressão apresentada na Equação 3.28 considerando $\lambda = np = 150 \times 0,05 = 7,5$.

$$P(X < 10) \approx \sum_{x=0}^{9} \frac{7,5^x}{x!} e^{-7,5} = \frac{7,5^0}{0!} e^{-7,5} + \ldots + \frac{7,5^8}{8!} e^{-7,5} + \frac{7,5^9}{9!} e^{-7,5} = 0,7764076.$$

A partir das contas exatas chegamos em $P(X \geq 10) = 1 - 0,7808835 = 0.2191165$. Já usando as contas aproximadas chegamos em $P(X \geq 10) = 1 - 0,7764076 = 0,2235924$. Veja como próximos são os dois resultados.

♦♦

Função de Probabilidade

A própria definição de Distribuição de Poisson já nos passa que se $X \sim Poi(\lambda)$ a sua função de probabilidade é definida por:

$$p_X(x) = P(X = x) = \frac{\lambda^x}{x!} e^{-\lambda} \qquad x = 0, 1, 2, \ldots \tag{3.29}$$

Para mostrar que a Equação 3.29 realmente define uma função de probabilidade, temos que provar que $\sum_{x=0}^{\infty} P(X = x) = 1$. Para isso vamos rever o resultado da série de Taylor em torno de $x_0 = 0$ para $f(x) = e^x$, Equação A.26 do Apêndice A.6, de onde concluímos que

$$e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!}.$$

Assim, temos que:

$$\sum_{x=0}^{\infty} P(X = x) = \sum_{x=0}^{\infty} \frac{\lambda^x}{x!} e^{-\lambda} = e^{-\lambda} \underbrace{\sum_{x=0}^{\infty} \frac{\lambda^x}{x!}}_{e^{\lambda}} = e^{-\lambda} e^{\lambda} = 1.$$

Esperança e Variância

Vamos agora calcular a esperança e a variância de tal distribuição.

$$\begin{aligned} E(X) &= \sum_{x=0}^{\infty} x \frac{\lambda^x}{x!} e^{-\lambda} \stackrel{\star}{=} \sum_{x=1}^{\infty} x \frac{\lambda^x}{x!} e^{-\lambda} = \sum_{x=1}^{\infty} x \frac{\lambda \lambda^{x-1}}{x(x-1)!} e^{-\lambda} \\ &= \lambda \sum_{x=1}^{\infty} \frac{\lambda^{x-1}}{(x-1)!} e^{-\lambda} \stackrel{\star\star}{=} \lambda \underbrace{\sum_{y=0}^{\infty} \frac{\lambda^y}{y!} e^{-\lambda}}_{1} = \lambda \end{aligned}$$

A passagem $\stackrel{\star}{=}$ se justifica pois o somatório se anula quando $x = 0$, por isso podemos mudar o índice do somatório para começar em 1. Na passagem $\stackrel{\star\star}{=}$ foi feita a seguinte troca de variável: $y = x - 1$. O somatório destacado pela chave vale 1 pois ele é a soma da função de probabilidade de uma variável aleatória de Poisson para todos os valores na imagem.

Portanto,

$$X \sim Poi(\lambda) \qquad \Rightarrow \qquad \mathrm{E}(X) = \lambda \tag{3.30}$$

Veja que o parâmetro λ da distribuição de Poisson é também a média dessa variável aleatória. Isso significa que durante um intervalo de tempo de tamanho 1 acontece em média λ ocorrências do evento em questão.

A unidade de tempo não foi definida porque ela pode ser qualquer uma. O importante é que o valor de λ esteja de acordo com a unidade de tempo adotada. Veja alguns exemplos nos itens a seguir.

- Se X = número de ocorrência de um certo evento em 1 dia, então a unidade de tempo é 1 dia e λ será o número médio de ocorrências em 1 dia;
- Se X = número de ocorrência de um certo evento em 1 semana, então a unidade de tempo é 1 semana e λ será o número médio de ocorrências em 1 semana;
- Se X = número de ocorrência de um certo evento em 2 horas, então a unidade de tempo é 2 horas e λ será o número médio de ocorrências em 2 horas.

Exemplo 3.31 Central telefônica

Uma central telefônica recebe uma média de 5 chamadas por minuto. Supondo que as chamadas cheguem segundo uma distribuição de Poisson. Qual é a probabilidade da central telefônica

(a) não receber nenhuma chamada em um minuto?

(b) receber no máximo 2 chamadas em 2 minutos?

Solução:

(a) Nesse item estamos interessados em saber o número de chamadas no intervalo de 1 minuto. Então a unidade de tempo será 1 minuto e vamos definir X = número de chamadas por minuto. De acordo com o enunciado, $X \sim Poi(\lambda)$, onde λ é o número médio de chamadas durante 1 minuto, ou seja, $\lambda = 5$. Queremos encontrar $\mathrm{P}(X = 0)$.

$$\mathrm{P}(X = 0) = \frac{5^0}{0!}e^{-5} = e^{-5} \approx 0,006737947.$$

(b) Agora o nosso interesse é no número de ligações no intervalo de 2 minutos. Então a unidade de tempo adotada será 2 minutos. Vamos definir Y = número de chamadas em 2 minutos. De acordo com o enunciado, $Y \sim Poi(\lambda)$, onde λ é o número médio de chamadas em 2 minutos, ou seja, $\lambda = 10$. Queremos encontrar $\mathrm{P}(Y \leq 2)$.

$$\begin{aligned} \mathrm{P}(Y \leq 2) &= \mathrm{P}(Y = 0) + \mathrm{P}(Y = 1) + \mathrm{P}(Y = 2) \\ &= \frac{10^0}{0!}e^{-10} + \frac{10^1}{1!}e^{-10} + \frac{10^2}{2!}e^{-10} = 0.002769396 \end{aligned}$$

♦♦

Vamos agora encontrar Var(X). Para isso vamos primeiro calcular E(X^2).

$$\begin{aligned}
E(X^2) &= \sum_{x=0}^{\infty} x^2 \frac{\lambda^x}{x!} e^{-\lambda} \overset{\star}{=} \sum_{x=1}^{\infty} x^2 \frac{\lambda^x}{x!} e^{-\lambda} = \sum_{x=1}^{\infty} x^2 \frac{\lambda\lambda^{x-1}}{x(x-1)!} e^{-\lambda} \\
&= \lambda \sum_{x=1}^{\infty} x \frac{\lambda^{x-1}}{(x-1)!} e^{-\lambda} \overset{\star\star}{=} \lambda \sum_{y=0}^{\infty} (y+1) \frac{\lambda^y}{y!} e^{-\lambda} = \lambda \left(\underbrace{\sum_{y=0}^{\infty} y \frac{\lambda^y}{y!} e^{-\lambda}}_{S_1} + \underbrace{\sum_{y=0}^{\infty} \frac{\lambda^y}{y!} e^{-\lambda}}_{S_2} \right)
\end{aligned}$$

A passagem $\overset{\star}{=}$ se justifica pois o somatório se anula quando $x = 0$, por isso podemos mudar o índice do somatório para começar em 1. Na passagem $\overset{\star\star}{=}$ foi feita a seguinte troca de variável: $y = x - 1$. Veja que $S_1 = E(Y)$, com $Y \sim Poi(\lambda)$, logo, $S_1 = \lambda$. Já $S_2 = 1$ pois ele é a soma da função de probabilidade de uma variável aleatória de Poisson para todos os valores na imagem. Assim chegamos em,

$$E(X^2) = \lambda(\lambda + 1) = \lambda^2 + \lambda.$$

Logo,

$$\mathrm{Var}(X) = E(X^2) - E(X)^2 = \lambda^2 + \lambda - \lambda^2 = \lambda$$

Com isso chegamos à expressão da variância de uma v.a. de Poisson.

$$X \sim Poi(\lambda) \qquad \Rightarrow \qquad Var(X) = \lambda \tag{3.31}$$

A variável aleatória de Poisson tem média e variância iguais à λ.

Apesar da motivação da distribuição de Poisson ser para a contagem do número de ocorrência ao longo do tempo, como citado anteriormente, tal distribuição também pode ser usada para modelar situações que não necessariamente envolvam tempo, como mostra o Exemplo 3.32 a seguir. Mas o importante é saber que as características dessa distribuição são adequadas para modelar variáveis aleatórias de contagem.

Exemplo 3.32 Fabricação de fita magnética

Em um certo tipo de fabricação de fita magnética, ocorrem cortes a uma taxa de um corte por 2000 pés. Qual é a probabilidade de que um rolo com comprimento de 4000 pés apresente

(a) no máximo dois cortes?

(b) pelo menos dois cortes?

Solução:

Seja Y = número de cortes num rolo de 4.000 pés. Então a unidade usada é 4.000 pés. Como ocorre em média 1 corte a cada 2.000 pés, o número médio de cortes em 4.000 pés são 2. Logo, $Y \sim Poi(2)$.

(a) Queremos encontrar $P(Y \leq 2)$.

$$\begin{aligned}
P(Y \leq 2) &= P(Y = 0) + P(Y = 1) + P(Y = 2) \\
&= \frac{2^0}{0!}\exp(-2) + \frac{2^1}{1!}\exp(-2) + \frac{2^2}{2!}\exp(-2) = 0,676676
\end{aligned}$$

(b) Nesse item queremos $P(Y \geq 2)$.

$$\begin{aligned} P(Y \geq 2) &= 1 - P(Y < 2) = 1 - (P(Y = 0) + P(Y = 1)) \\ &= 1 - \left(\frac{2^0}{0!} \exp(-2) + \frac{2^1}{1!} \exp(-2) \right) = 0,593994 \end{aligned}$$

♦♦

Formas da Distribuição de Poisson

Se $X \sim Poi(\lambda)$, então temos

$$\frac{P(X = k+1)}{P(X = k)} = \frac{\frac{\lambda^{k+1}}{(k+1)!} e^{-\lambda}}{\frac{\lambda^k}{k!} e^{-\lambda}} = \frac{\lambda}{k+1} \qquad k = 0, 1, 2, \ldots$$

Temos, assim, uma forma recursiva de calcular probabilidades de Poisson.

Suponhamos, agora, que a probabilidade máxima ocorra em k_0. Usando o resultado acima, temos que ter

$$\frac{P(X = k_0 + 1)}{P(X = k_0)} \leq 1 \Rightarrow \frac{\lambda}{k_0 + 1} \leq 1 \Rightarrow k_0 \geq \lambda - 1$$

e também

$$\frac{P(X = k_0)}{P(X = k_0 - 1)} \geq 1 \Rightarrow \frac{\lambda}{k_0} \geq 1 \Rightarrow k_0 \leq \lambda$$

Resulta, então, que se k_0 é ponto de máximo, então

$$\lambda - 1 \leq k_0 \leq \lambda$$

e como k_0 tem que ser inteiro, temos que ter $k_0 = [\lambda]$, ou seja, a moda ocorre em $k_0 = [\lambda]$, o maior inteiro menor ou igual a λ.

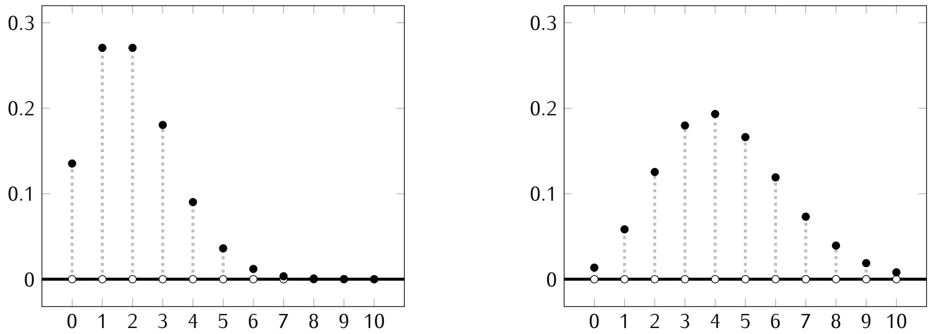
Se λ for inteiro, então a distribuição é bimodal, com moda em $k_{01} = \lambda$ e $k_{02} = \lambda - 1$ pois, nesse caso,

$$P(X = \lambda) = \frac{\lambda^\lambda}{\lambda!} e^{-\lambda} = \frac{\lambda \cdot \lambda^{\lambda-1}}{\lambda \cdot (\lambda - 1)!} e^{-\lambda} = P(X = \lambda - 1)$$

Nas Figuras 3.6(a) e 3.6(b) ilustra-se a distribuição de Poisson para dois valores do parâmetro: $\lambda = 2$ (distribuição bimodal, com modas 1 e 2) e $\lambda = 4,3$ (distribuição unimodal, com moda 4)

3.9 Mais Exemplos

Exemplo 3.33
Um grupo de 30 políticos, sendo 12 do partido de esquerda, resolve formar, de forma aleatória, uma comissão com 5 políticos. Qual a probabilidade dos políticos de esquerda serem maioria na comissão?



(a) $Poi(\lambda = 2)$ (b) $Poi(\lambda = 4,3)$

Figura 3.6 – Função de Probabilidade da Distribuição de Poisson.

Solução:
Defina X = número de políticos de esquerda na comissão. Veja que $X \sim hiper(N = 30, r = 12, n = 5)$. Queremos $P(X \geq 3) = P(X = 3) + P(X = 4) + P(X = 5)$.

$$
\begin{aligned}
P(X = x) &= \frac{\binom{12}{x}\binom{30-12}{5-x}}{\binom{30}{5}} = \frac{\binom{12}{x}\binom{18}{5-x}}{\binom{30}{5}} \\
P(X = 3) &= \frac{\binom{12}{3}\binom{18}{5-3}}{\binom{30}{5}} = \frac{\frac{12!}{3!9!}\frac{18!}{2!16!}}{\frac{30!}{5!25!}} = \ldots = 0,2362 \\
P(X = 4) &= \frac{\binom{12}{4}\binom{18}{5-4}}{\binom{30}{5}} = \frac{\frac{12!}{4!8!}\frac{18!}{1!17!}}{\frac{30!}{5!25!}} = \ldots = 0,0625 \\
P(X = 5) &= \frac{\binom{12}{5}\binom{18}{5-5}}{\binom{30}{5}} = \frac{\frac{12!}{5!7!}\frac{18!}{0!18!}}{\frac{30!}{5!25!}} = \ldots = 0,0055
\end{aligned}
$$

Logo, $P(X \geq 3) = 0,2362 + 0,0625 + 0,0055 = 0,3042$.

♦♦

Exemplo 3.34
Sabe-se que a eficácia de uma vacina para uma certa doença é de 92%. Isso significa que a probabilidade de um indivíduo que tomou a vacina ser imunizado é de 92%.

(a) Qual a probabilidade de entre 10 pacientes vacinados no máximo 1 não ter sido imunizado?

(b) Quantos indivíduos, em média, devem ser vacinados até que seja encontrado um que não tenha sido imunizado?

Solução:

(a) Nesse item queremos a probabilidade de 0 ou 1 paciente entre 10 vacinados não ter sido imunizado. Defina X = número de pacientes não imunizados entre os 10 que tomaram a vacina. Veja que $X \sim B(n = 10, p = 0,08)$. Queremos $\mathrm{P}(X \leq 1)$.

$$\begin{aligned}
\mathrm{P}(X \leq 1) &= \mathrm{P}(X = 0) + \mathrm{P}(X = 1) = \binom{10}{0} 0,08^0 0,92^{10} + \binom{10}{1} 0,08^1 0,92^9 \\
&= 1 \times 0,4343885 + 10 \times 0,1 \times 0,4721614 = 0,9065499
\end{aligned}$$

(b) Já nesse item queremos o número médio de pacientes vacinados até encontrar o 1º que não foi imunizado. Defina Y = número de pacientes vacinados até encontrarmos 1 não imunizados. Veja que $Y \sim geo(p = 0,08)$. Queremos $\mathrm{E}(Y) = 1/p = 12,5$. Ou seja, em média são vacinados 12,5 pacientes até que se encontre o 1º que não foi imunizado.

♦♦

Exemplo 3.35

Uma rodovia registra, em média, 2 acidentes por dia. Considere a variável aleatória X = número de acidentes na rodovia.

(a) Indique uma distribuição adequada para a variável aleatória X. Justifique sua resposta.

(b) Qual a probabilidade de nas próximas 6 horas acontecer pelo menos 1 acidente.

Solução:

(a) A distribuição de Poisson é adequada para dados de contagem. No caso estamos contando o número de acidentes, então é razoável considerar $X \sim Poi(\lambda)$.

(b) Estamos interessados no número de acidentes no período de 6 horas. Vamos definir então X = número de acidentes no período de 6 horas e assim temos $X \sim Poi(\lambda)$, onde λ é o número médio de acidentes em 6 horas, no caso, $\lambda = 0,5$. Queremos $\mathrm{P}(X \geq 1)$.

$$\mathrm{P}(X \geq 1) = 1 - \mathrm{P}(X = 0) = 1 - \frac{0,5^0}{0!} e^{-0,5} = 1 - 0.6065307 = 0.3934693.$$

♦♦

Exemplo 3.36

Suponha que em uma cidade 15% dos cidadãos votaram no vereador João. A fim de avaliar a satisfação de seus eleitores, João decide fazer uma pesquisa direcionada. Para isso João aborda, de forma aleatória e com reposição, cidadãos na rua a fim de encontrar seus eleitores e realizar a pesquisa.

(a) Qual a probabilidade de nas 10 primeiras abordagens João não encontrar eleitor algum seu?

(b) Se João planeja entrevistar 20 dos seus eleitores, em média, quantos cidadãos serão abordados?

Solução:

(a) Esse item podemos fazer de duas maneiras diferentes. Vamos apresentar aqui as duas alternativas de solução.

Primeira solução. Seja Y = número de eleitores de João entre os 10 primeiros cidadãos abordados. Veja que $Y \sim B(n = 10, p = 0,15)$. Queremos $\mathrm{P}(Y = 0)$.

$$\mathrm{P}(Y = 0) = \binom{10}{0} 0,15^0 0,85^{10-0} = 0,85^{10} = 0,1968744.$$

Para esse exemplo as contas da solução anterior são mais fácies, mas vale à pena ver que o mesmo problema pode ser resolvido de formas totalmente diferentes.

Segunda solução. Considere X = número de cidadãos abordados até encontrar o primeiro eleitor de João. Veja que $X \sim geo(p = 0,150)$. Queremos $P(X > 10) = 1 - P(X \leq 10)$.

$$\begin{aligned} P(X > 10) &= 1 - P(X \leq 10) = 1 - (P(X = 1) + P(X = 2) + \ldots + P(X = 10)) \\ &= 1 - \sum_{x=1}^{10} P(X = x) = 1 - \sum_{x=1}^{10} 0,15(0,85)^x = 1 - 0,15 \sum_{x=1}^{10} (0,85)^x \\ &= 1 - 0,15 \frac{1 - 0,85^{10}}{0,15} = 0,85^{10} = 0,1968744 \end{aligned}$$

(b) Precisamos definir quantas abordagens em média são realizadas até que João encontre 20 de seus eleitores. Defina W = número de abordagens até João encontrar 20 de seus eleitores. Veja que $W \sim BinNeg(r = 20, p = 0,15)$. Queremos $E(W) = r/p = 20/0,15 = 133,33$. Ou seja, em média, João vai precisar abordar em torno de 134 indivíduos para conseguir encontrar 20 de seus eleitores.

♦♦

Exemplo 3.37 Problema dos pontos - Exemplo 2.16 de Magalhães (2011)
Ensaios do tipo sucesso-fracasso são realizados de forma independente, sendo p a probabilidade de sucesso. Qual é a probabilidade de ocorrerem n sucessos antes de m fracassos?

Solução:
Veja a probabilidade de ocorrerem n sucessos antes de m fracassos é exatamente a probabilidade de ocorrerem n sucessos antes da $n + m$-ésima tentativa. Então podemos encontrar a resposta do problema calculando a probabilidade do n-ésimo sucesso acontecer antes da $n + m$-ésima tentativa. Para isso defina X = como o número de ensaios antes do n-ésimo sucesso. Veja que $X \sim BinNeg(n, p)$. Queremos encontrar $P(X < n + m)$,

$$P(X < n + m) = \sum_{x=n}^{n+m-1} \binom{x-1}{n-1} p^n (1-p)^{x-n}.$$

Poderíamos também resolver pela forma alternativa da Binomial Negativa. Seja Y = número de fracassos antes do n-ésimo sucesso. Veja que a probabilidade de interesse é $P(Y < m)$, ou seja, a probabilidade do n-ésimo sucesso ocorrer antes do m-ésimo fracasso. Nesse caso,

$$P(Y < m) = \sum_{y=0}^{m-1} \binom{y+n-1}{n-1} p^n (1-p)^y.$$

Vejamos um caso numérico. Supondo que os ensaios sejam o lançamento de uma moeda justa, ou seja, $p = 1/2$. Vamos definir a ocorrência de cara como sucesso. Suponha que queremos saber a probabilidade de sair 5 caras antes da 3 corroa, ou seja, $n = 5$ e $m = 3$.

$$P(X < 8) = \sum_{x=5}^{7} \binom{x-1}{n-1} \frac{1}{2}^5 \frac{1}{2}^{x-5} = \binom{4}{4} \frac{1}{2}^5 \frac{1}{2}^0 + \binom{5}{4} \frac{1}{2}^5 \frac{1}{2}^1 + \binom{6}{4} \frac{1}{2}^5 \frac{1}{2}^2 \approx 0.2265625.$$

$$P(Y < 3) = \sum_{y=0}^{2} \binom{y+4}{4} \frac{1}{2}^5 \frac{1}{2}^y = \binom{4}{4} \frac{1}{2}^5 \frac{1}{2}^0 + \binom{5}{4} \frac{1}{2}^5 \frac{1}{2}^1 + \binom{6}{4} \frac{1}{2}^5 \frac{1}{2}^2 \approx 0.2265625.$$

♦♦

Capítulo 4

Variáveis Aleatórias Contínuas

Já vimos, no Capítulo 1, conceitos iniciais sobre as variáveis aleatórias contínua. Nesse capítulo vamos estudar mais sobre esse tipo de variável aleatória.

4.1 Função Densidade de Probabilidade

A definição de variável aleatória contínua normalmente apresentada nos livros de probabilidade não é citada no Capítulo 1, que define X contínua quando a sua função de distribuição acumulada é uma função absolutamente contínua. A forma mais comum de definir variável aleatória contínua está na Definição 4.1 a seguir, onde junto aparece o conceito de função densidade.

Definição 4.1 Variável Aleatória Contínua e Função Densidade
Uma variável aleatória X com função de distribuição F é classificada como ***contínua*** *quando existe uma função não negativa f tal que*

$$F(x) = \int_{-\infty}^{x} f(t)dt \quad , \quad \text{para todo } x \in \mathbb{R}.$$

Nesse caso, a função f é denominada ***função densidade de probabilidade****, ou simplesmente* ***função densidade****, da variável aleatória X.*

Proposição 4.2 Propriedades da Função Densidade
Seja f_X a função densidade de alguma variável aleatória X. Então f_X satisfaz as seguintes propriedades.

(i) $f_X(x) \geq 0 \ \forall \ x \in \mathbb{R}$.

(ii) $\int_{-\infty}^{\infty} f(t)dt = 1$

Demonstração:
A função densidade é não negativa por definição. Para mostrar que $\int_{-\infty}^{\infty} f(t)dt = 1$ veja que

$$\int_{-\infty}^{\infty} f(t)dt = \lim_{x\to\infty} \int_{-\infty}^{x} f(t)dt = \lim_{x\to\infty} F(x) = 1,$$

pelas propriedades de função distribuição. $\square$

Já vimos como calcular $P(a \leq X \leq b)$ se conhecermos a função de distribuição de X (Exemplo 1.18). O que veremos agora é que esse cálculo também pode ser feito a partir da função densidade.

Proposição 4.3
Seja X variável aleatória contínua com função densidade f_X e $a, b \in \mathbb{R}$ tais que $a \leq b$. Então

$$P(a \leq X \leq b) = \int_a^b f_X(t)dt.$$

Demonstração:
Seja X variável aleatória contínua e F_X a sua função de distribuição. Já vimos que:

$$P(a \leq X \leq b) = F_X(b) - F_X(a). \tag{4.1}$$

Pela Definição 4.1 temos que $F(x) = \int_{-\infty}^x f(t)dt$, $\forall\ x \in \mathbb{R}$. Ou seja,

$$F(a) = \int_{-\infty}^a f(t)dt \qquad \text{e} \qquad F(b) = \int_{-\infty}^b f(t)dt. \tag{4.2}$$

Juntando as Equações 4.1 e 4.2 e aplicando algumas propriedades da integral temos:

$$P(a \leq X \leq b) = F_X(b) - F_X(a) = \int_{-\infty}^b f(t)dt - \int_{-\infty}^a f(t)dt = \int_a^b f(t)dt.$$

$\square$

Veja que a Proposição 4.3 nos mostra que a probabilidade de $X \in [a, b]$ é a área embaixo da curva da sua função densidade. Quando $a = b$ a proposição ainda se aplica e temos uma demonstração para o seguinte resultado já visto no Capítulo 1:

$$P(X = a) = P(a \leq X \leq a) = \int_a^a f_X(t)dt = 0.$$

O resultado da Proposição 4.3 também se estende para $a = -\infty$ ou $b = \infty$.

$$\begin{aligned} P(-\infty \leq X \leq b) &= P(X \leq b) = F_X(b) = \int_{-\infty}^b f_X(t)dt \\ P(a \leq X \leq \infty) &= P(X \geq a) = 1 - F_X(a) = 1 - \int_{-\infty}^a f_X(t)dt = \int_a^\infty f_X(t)dt, \end{aligned}$$

uma vez que $\int_{-\infty}^a f_X(t)dt + \int_a^\infty f_X(t)dt = \int_{-\infty}^\infty f_X(t)dt = 1$.

Assim sabemos fazer contas de probabilidades envolvendo uma variável aleatória contínua a partir da sua função densidade. Dessa forma, conhecendo a função densidade de uma variável aleatória contínua temos toda a informação sobre a sua distribuição. Isso significa que tanto faz conhecer F_X ou f_X, em ambos os casos temos toda a informação sobre a variável aleatória.

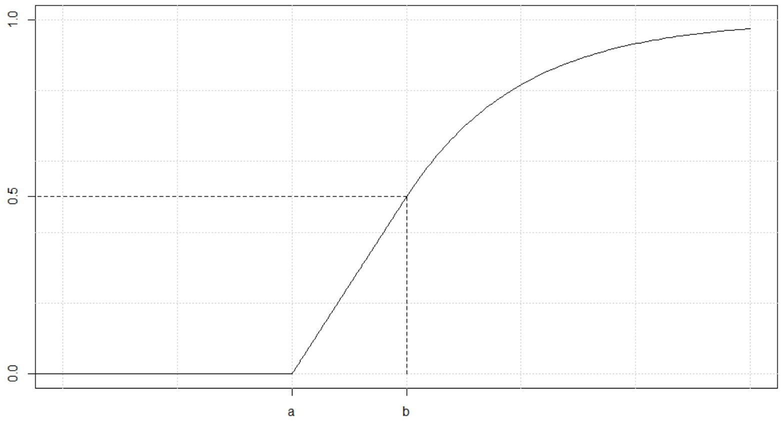
Mais que isso, conhecendo uma das funções F_X ou f_X temos como encontrar a outra, veja como. A Definição 4.1 nos mostra como encontrar F_X se conhecermos f_X:

$$F_X(x) = \int_{-\infty}^x f_X(t)dt.$$

Aplicando o Teorema Fundamental do Cálculo concluímos que:

$$f_X(x) = \frac{d}{dx}F_X(x).$$

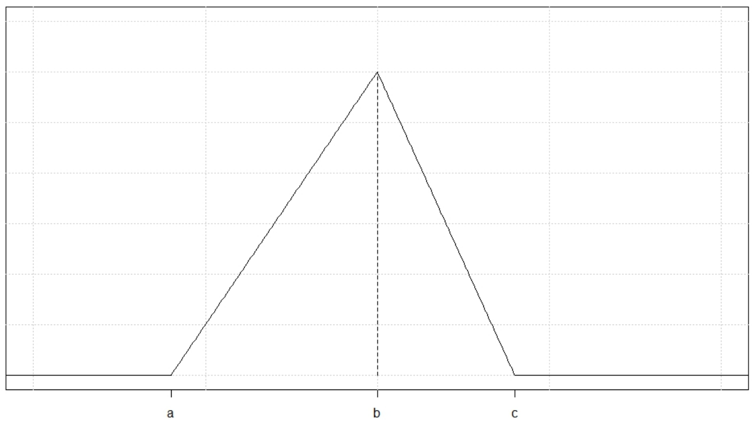
Exemplo 4.4
Seja X uma variável aleatória cuja função de distribuição tem seu gráfico esboçado a seguir. Apresente um esboço para o gráfico da função densidade dessa variável aleatória.



Solução:

♦♦

Exemplo 4.5
Seja X uma variável aleatória cuja função densidade tem seu gráfico esboçado a seguir. Apresente um esboço para o gráfico da função de distribuição dessa variável aleatória.



Solução:

♦♦

Exemplo 4.6
Seja X variável aleatória contínua cuja função densidade é definida por:

$$f_X(x) = \begin{cases} a(1+x) & , \textit{se } -1 \leq x < 0 \\ a(1-x) & , \textit{se } 0 \leq x \leq 1 \\ 0 & , \textit{caso contrário.} \end{cases}$$

(a) *Faça um esboço do gráfico de f_X.*

(b) *Encontre o valor de a para que f_X seja realmente função densidade.*

(c) *Quais valores a variável aleatória X pode assumir? Ou seja, qual a imagem de X?*

(d) *Calcule* $P(X > 0)$, $P(-1/2 \leq X \leq 1/2)$ *e* $P(X = 0)$.

(e) *Encontre a função de distribuição de X, F_X, e esboce seu gráfico.*

(f) *Verifique as propriedades da função de distribuição.*

(g) *Repita os itens ?? e ?? resolvendo agora a partir da função F_X.*

Solução:

♦♦

Exemplo 4.7
Seja X variável aleatória cuja função de distribuição é definida por:

$$F_X(x) = \begin{cases} 0 & , \text{ se } x < 2 \\ (x-2)/4 & , \text{ se } 2 \leq x < 4 \\ 1/2 & , \text{ se } 4 \leq x < 6 \\ (x-4)/4 & , \text{ se } 6 \leq x < 8 \\ 1 & , \text{ se } x \geq 8 \end{cases}$$

(a) *Faça um esboço do gráfico de F_X.*

(b) *Quais valores a variável aleatória X pode assumir? Ou seja, qual a imagem de X?*

(c) *Classifique a variável aleatória X.*

(d) *Calcule* $P(X \geq 3)$ *e* $P(3 < X \leq 5)$.

(e) *Encontre a função densidade de X, f_X, e esboce seu gráfico.*

(f) *Verifique as propriedades da função densidade.*

(g) *Repita os itens ?? e ?? resolvendo agora a partir da função f_X.*

Solução:

♦♦

Exemplo 4.8
Seja X variável aleatória contínua tal que sua função densidade é definida por:

$$f(x) = \begin{cases} c(4x - 2x^2) & , \text{ se } 0 < x < 2 \\ 0 & , \text{ caso contrário.} \end{cases}$$

(a) *Encontre o valor de c para que f seja realmente função densidade.*

(b) *Esboce o gráfico de f.*

(c) *Calcule* $P(X > 1)$.

(d) *Encontre F_X.*

Solução:

♦♦

Exemplo 4.9
O tempo de vida em horas de um certo dispositivo eletrônico pode ser considerado uma variável aleatória com função densidade definida por

$$f_X(x) = \frac{1}{100} e^{-x/100} \qquad , \qquad x \geq 0.$$

(a) *Qual a probabilidade de um desses dispositivos eletrônicos durar entre 50 e 150 horas?*

(b) *Qual a probabilidade de um desses dispositivos eletrônicos durar menos de 100 horas?*

Solução:

♦♦

Exemplo 4.10
O tempo de vida em horas de um certo tipo de tubo de rádio é uma variável aleatória com função densidade

$$f(x) = \begin{cases} \frac{100}{x^2} & \text{, se } x > 100 \\ 0 & \text{, caso contrário.} \end{cases}$$

(a) *Qual a probabilidade de um tubo desse tipo durar mais que 150h?*

(b) *Qual a probabilidade de exatos 2 entre 5 desses tubos terem que ser trocados antes de 150h de operação?*

Solução:

♦♦

4.2 Esperança de Variáveis Aleatórias Contínuas

A ideia de valor médio apresentada no Capítulo 3, que diz respeito à média dos valores da variável aleatória se o experimento fosse realizado infinitas vezes, também pode ser introduzida para o caso de variáveis aleatórias contínuas.

Porém no caso contínuo não é possível tirar a média dos valores da variável aleatória ponderados pela probabilidade de cada um ocorrer, uma vez que não há uma quantidade enumerável de valores e a probabilidade de cada um ocorrer e sempre nula. Por isso, no caso contínuo a definição será diferente do caso discreto.

Vamos tentar generalizar. Sem perdas, suponha X variável aleatória contínua com função densidade f_X e $Im(X) = [0, 1]$. Vamos dividir esse intervalo em subintervalos $[x, x + dx]$, com $dx = 1/n$ e $x = 0, 1/n, 2/n, \ldots, n - 1$. Veja que se n for grande, então

$$f_X(x)dx \approx \mathrm{P}(x \leq X \leq x + dx)$$

Seguindo a definição de esperança para o caso discreto, podemos supor para o caso contínuo que, se n for grande,

$$\mathrm{E}(X) \approx \sum_x x\,\mathrm{P}(x \leq X \leq x + dx) \approx \sum_x x f_X(x)dx$$

Pensando no limite quando $n \to \infty$, ou $dx \to 0$, temos

$$\mathrm{E}(X) = \lim_{n\to\infty} \sum_x x\,\mathrm{P}(x \leq X \leq x + dx) = \lim_{n\to\infty} \sum_x x f_X(x)dx = \int_x x f_X(x)dx.$$

E assim chegamos na Definição 4.11.

Definição 4.11 Esperança de Variável Aleatória Contínua
*Seja X uma variável aleatória contínua com função densidade de probabilidade f_X. A **esperança** (ou **média** ou **valor esperado**) de X é definida como*

$$\mathrm{E}(X) = \int_{-\infty}^{+\infty} x f_X(x)dx$$

desde que a integral esteja bem definida.

Exemplo 4.12
Calcule $\mathrm{E}(X)$ para os exemplos 4.8, 4.9 e 4.10.

Solução:

♦♦

Assim como no caso discreto, se X é variável aleatória e $Y = g(X)$, Y também será variável aleatória e o cálculo da sua esperança pode ser feito somente com o conhecimento da função g que relaciona X e Y e da função densidade de X, f_X.

Proposição 4.13 Esperança de funções de variável aleatória contínua
Seja X uma variável aleatória contínua com função densidade f_X e seja $Y = g(X)$. Então,

$$\mathrm{E}\left(Y\right) = \mathrm{E}\left(g\left(X\right)\right) = \int_{-\infty}^{\infty} g(x) f_X(x),$$

desde que a integral esteja bem definida.

Demonstração:
A demonstração será omitida.

□

Exemplo 4.14
Seja X variável aleatória com função densidade definida por

$$f(x) = \begin{cases} (x+1)^2 & , \text{se } -1 < x < 0 \\ 1 - x^2 & , \text{se } 0 \leq x < 1 \end{cases}$$

(a) Verifique as propriedades da função densidade e esboce o gráfico de f.

(b) Calcule $\mathrm{E}(X)$.

(c) Calcule $\mathrm{E}(1/(X+1))$.

Solução:

♦♦

Assim como no caso discreto, continuam valendo para o caso contínuo as seguintes propriedades da esperança.

Proposição 4.15
Seja X variável aleatória contínua, $x_{min} = \min\{Im(X)\}$ e $x_{max} = \max\{Im(X)\}$. Se $\mathrm{E}(X) \in \mathbb{R}$ então, $x_{min} \leq \mathrm{E}(X) \leq x_{max}$.

Demonstração:
Primeiro vamos mostrar que $x_{min} \leq \mathrm{E}(X)$.

$$\mathrm{E}(X) = \int_{-\infty}^{\infty} x f_X(x) \geq \int_{-\infty}^{\infty} x_{min} f_X(x) = x_{min} \underbrace{\int_{-\infty}^{\infty} f_X(x)}_{1} = x_{min}.$$

Para mostrar que $\mathrm{E}(X) \leq x_{max}$ faremos um desenvolvimento análogo.

$$\mathrm{E}(X) = \int_{-\infty}^{\infty} x f_X(x) \leq \int_{-\infty}^{\infty} x_{max} f_X(x) = x_{max} \underbrace{\int_{-\infty}^{\infty} f_X(x)}_{1} = x_{max}.$$

$\square$

Proposição 4.16 Propriedade de Linearidade da Esperança
Seja X variável aleatória contínua tal que $\mathrm{E}(X) \in \mathbb{R}$. Sejam $a, b \in \mathbb{R}$. Então,

$$\mathrm{E}(aX + b) = a\,\mathrm{E}(X) + b.$$

Demonstração:

$$\begin{aligned}
\mathrm{E}(aX + b) &= \int_{-\infty}^{\infty} (ax + b) f_X(x) dx = \\
&= \int_{-\infty}^{\infty} \left(ax f_X(x) + b f_X(x)\right) dx = \\
&= \int_{-\infty}^{\infty} ax f_X(x) dx + \int_{-\infty}^{\infty} b f_X(x) dx = \\
&= a \underbrace{\int_{-\infty}^{\infty} x f_X(x) dx}_{\mathrm{E}(X)} + b \underbrace{\int_{-\infty}^{\infty} f_X(x) dx}_{1} = a\,\mathrm{E}(X) + b.
\end{aligned}$$

$\square$

4.3 Variância

A variância de uma variável aleatória, discreta ou contínua, é uma medida de dispersão em torno da média. O que muda do caso discreto para o contínuo é apenas como as contas são feitas.

Definição 4.17 Variância de Variável Aleatória Contínua
*Seja X variável aleatória contínua tal que $\mathrm{E}(X) \in \mathbb{R}$. A **variância** de X é definida como*

$$\mathrm{Var}(X) = \mathrm{E}\left[(X - \mathrm{E}(X))^2\right] = \int_{-\infty}^{\infty} (x - \mathrm{E}(X))^2 f_X(x) dx,$$

desde que a integral esteja bem definida.

O **desvio-padrão** é definido como a raiz quadrada de sua variância, independente do tipo de variável.

$$\mathrm{DP}(X) = \sqrt{\mathrm{Var}(X)}$$

Usando as propriedades do cálculo integral e representando por μ a esperança de X (note que μ é uma constante, um número real), temos que:

$$\begin{aligned}
\mathrm{Var}(X) &= \int_{-\infty}^{+\infty} [x-\mu]^2 f_X(x)dx \\
&= \int_{-\infty}^{+\infty} \left(x^2 - 2\mu x + \mu^2\right) f_X(x)dx \\
&= \int_{-\infty}^{+\infty} x^2 f_X(x)dx - 2\mu \int_{-\infty}^{+\infty} x f_X(x)dx + \mu^2 \int_{-\infty}^{+\infty} f_X(x)dx
\end{aligned}$$

Se definimos $h(x) = x^2$, a primeira integral nada mais é que $\mathrm{E}(X^2)$, pela Proposição 4.13. A segunda integral é $\mathrm{E}(X) = \mu$ e a terceira integral é igual a 1, pela definição de função densidade. Logo,

$$\mathrm{Var}(X) = \mathrm{E}(X^2) - 2\mu^2 + \mu^2 = \mathrm{E}(X^2) - \mu^2$$

o que nos leva ao resultado já visto para variáveis discretas:

$$\mathrm{Var}(X) = \mathrm{E}(X^2) - [\mathrm{E}(X)]^2 \tag{4.3}$$

De forma resumida: a variância é a esperança do quadrado de X menos o quadrado da esperança de X.

As propriedades apresentadas na Proposição 2.34 valem também para o caso contínuo. Veja que a sua demonstração em momento algum usou o fato da variável ser discreta, apesar da proposição ter sido enunciada no capítulo sobre variáveis aleatórias discretas. Vejamos novamente tais propriedades.

Se X variável aleatória tal que $\mathrm{Var}(X) \in \mathbb{R}$, então valem as seguintes propriedades:

(i) $\mathrm{Var}(X) \geq 0$

(ii) $\mathrm{DP}(X) \geq 0$

(iii) $\mathrm{Var}(aX + b) = a^2 \,\mathrm{Var}(X)$

(iv) $\mathrm{DP}(aX + b) = |a| \,\mathrm{DP}(X)$

Exemplo 4.18
Para a variável aleatória apresentada no Exemplo 4.14 calcule: $\mathrm{Var}(X)$, $\mathrm{Var}(3X+1)$ *e* $\mathrm{Var}(X^2)$.

Solução:

♦♦

Exemplo 4.19 Função linear
Seja X variável aleatória cuja função densidade f_X está apresentada na Figura 4.1.

(a) Encontre o valor de k para que f_X seja função densidade de probabilidade de uma variável aleatória contínua e determine a equação que define f_X.

(b) Calcule $\mathrm{P}(2 \leq X \leq 3)$.

(c) Encontre a função de distribuição acumulada de X.

(d) Determine o valor de q tal que $\mathrm{P}(X \leq q) = 0,6$.

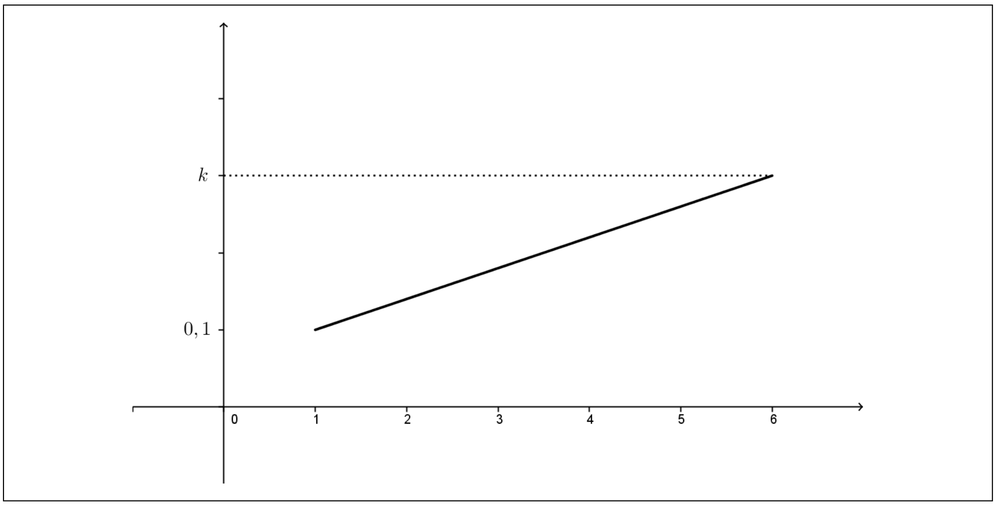


Figura 4.1 – Função densidade de probabilidade para o Exemplo 4.19

(e) *Calcule a esperança e a variância de X.*

Solução:

(a) O valor de k pode ser encontrado de forma que a área embaixo da curva seja 1. Para isso vamos somar a área do retângulo com a do triângulo para obter a área do trapézio.

$$\text{área } = 5 \times 0,1 + 5 \times (k - 0,1) \times \frac{1}{2} = 1 \Rightarrow k = 0,3.$$

Para encontrar a expressão que define f_X basta encontrar a equação da reta. Veja que essa reta passa pelos pontos $(1; 0,1)$ e $(6; 0,3)$, o que fornece o seguinte sistema de equações:

$$\begin{cases} 0,1 = a + b \\ 0,3 = a + 6b \end{cases}$$

Subtraindo a primeira equação da segunda, obtemos $b = 0,04$ e $a = 0,06$. Logo,

$$f_X(x) = \begin{cases} 0,06 + 0,04x & \text{se } 1 \leq x \leq 6 \\ 0 & casocontrário \end{cases}$$

(b) Veja a Figura 4.3, em que a área sombreada corresponde à probabilidade pedida:

$$P(2 \leq X \leq 3) = \int_2^3 (0,06 + 0,04x)dx = \left(0,06x + 0,04\frac{x^2}{2}\right)\Big|_2^3 = 0,06 + 0,04\left(\frac{9}{2} - \frac{4}{2}\right) = 0,16$$

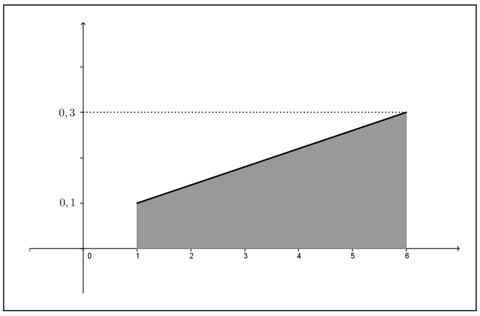


Figura 4.2 – Área sob a função densidade f_X do Exemplo 4.19

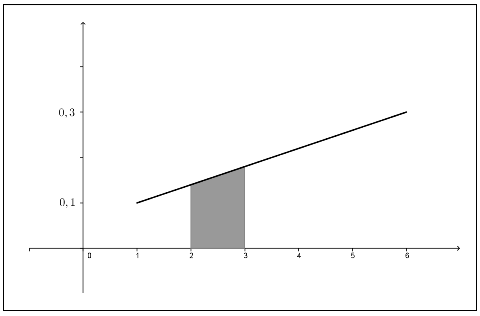


Figura 4.3 – $P(2 \leq X \leq 3)$ para o Exemplo 4.19

(c) Veja a Figura 4.4; aí podemos ver que, para $x \in [1, 6]$, $F_X(x)$ é a área de um trapézio de altura $x - 1$; base maior igual a $f_X(x)$ e base menor igual a $f_X(1) = 0,1$. Logo,

$$\begin{aligned} F_X(x) &= \frac{(0,06 + 0,04x) + 0,1}{2} \times (x - 1) \\ &= (0,08 + 0,02x)(x - 1) \end{aligned}$$

ou seja,

$$F_X(x) = \begin{cases} 0 & \text{se } x < 1 \\ 0,02x^2 + 0,06x - 0,08 & \text{se } 1 \leq x \leq 6 \\ 1 & \text{se } x > 6 \end{cases}$$

Na Figura 4.5 é dado o gráfico da função de distribuição.

Usando integral, temos que

$$\begin{aligned} F(x) &= \int_1^x (0,06 + 0,04t)dt = \left(0,06t + \frac{0,04t^2}{2}\right)\bigg|_1^x \\ &= \left(0,06x + 0,02x^2\right) - (0,06 + 0,02) \\ &= 0,02x^2 + 0,06x - 0,08 \qquad 1 \leq x \leq 6 \end{aligned}$$

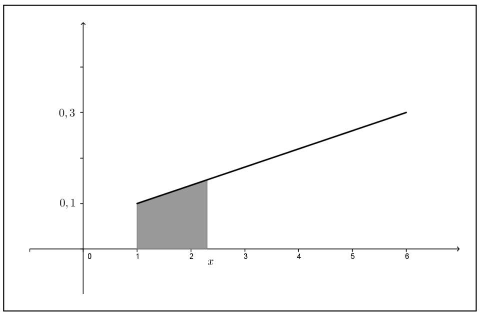


Figura 4.4 – F_X como área – Exemplo 4.19

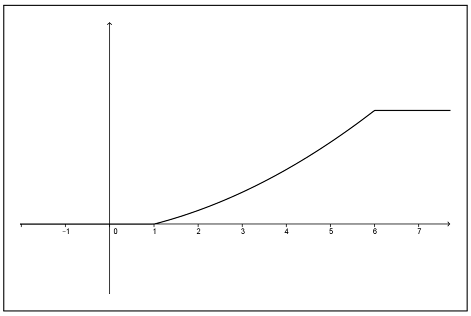


Figura 4.5 – F_X – Exemplo 4.19

(d) Queremos determinar q tal que $F_X(q) = 0,6$. Logo,

$$\begin{aligned} & 0,6 = 0,02q^2 + 0,06q - 0,08 \Rightarrow \\ & 0,02q^2 + 0,06q - 0,68 = 0 \Rightarrow \\ & q^2 + 3q - 34 = 0 \Rightarrow \\ q &= \frac{-3 \pm \sqrt{9 + 4 \times 34}}{2} \end{aligned}$$

A raiz que fornece resultado dentro do domínio de variação de X é

$$q = \frac{-3 + \sqrt{9 + 4 \times 34}}{2} \approx 4,5208$$

(e) Temos que

$$\begin{aligned} \mathrm{E}(X) &= \int_1^6 x\,(0,06 + 0,04x)\,dx = \left(0,06\frac{x^2}{2} + 0,04\frac{x^3}{3}\right)\bigg|_1^6 \\ &= \left(0,03 \cdot 36 + 0,04 \cdot \frac{6^3}{3}\right) - \left(0,03 \cdot 1 + \frac{0,04}{3}\right) \\ &= 1,08 + 2,88 - 0,03 - \frac{0,04}{3} = \frac{11,75}{3} = 3,9167 \end{aligned}$$

$$
\begin{aligned}
\mathrm{E}(X^2) &= \int_1^6 x^2\,(0,06+0,04x)\,dx = \left(0,06\frac{x^3}{3}+0,04\frac{x^4}{4}\right)\Big|_1^6 \\
&= (0,02\cdot 216 + 0,01\cdot 1296) - (0,02+0,01) \\
&= 4,32 + 12,96 - 0,03 = 17,25
\end{aligned}
$$

$$
\mathrm{Var}(X) = 17,25 - \left(\frac{11,75}{3}\right)^2 = \frac{155,25-138,0625}{9} = 1,9097
$$

♦♦

4.4 Densidade Simétrica

Se interpretamos a função densidade de probabilidade de X como uma distribuição de massa na reta real, então $\mathrm{E}(X)$ é o centro de massa desta distribuição. Essa interpretação nos permite concluir, por exemplo, que se f_X é simétrica, então $\mathrm{E}(X)$ é o valor central que define o eixo de simetria.

Exemplo 4.20 Densidade triangular
*Considere a função f_X apresentada na **Figura 4.6**:*

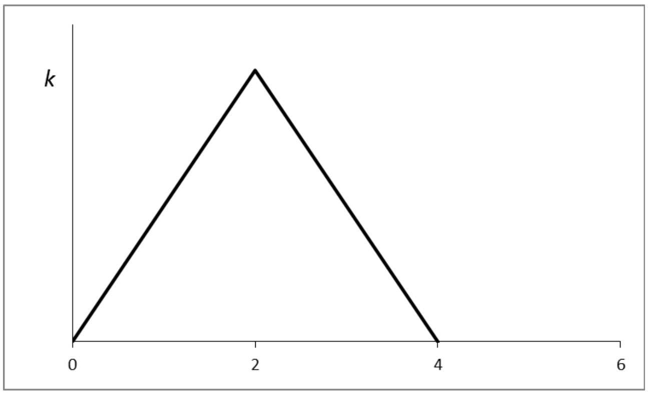


Figura 4.6 – Função de densdidade de probabilidade

(a) Encontre o valor de k para que f_X seja uma função densidade de probabilidade de uma v.a. X (note que o triângulo é isósceles).

(b) Determine a equação que define f_X.

(c) Calcule $\mathrm{P}(1 \leq X \leq 3)$.

(d) Encontre $\mathrm{E}(X)$.

(e) Determine o valor de Q tal que $\mathrm{P}(X \leq Q) = 0,6$.

(f) Encontre a função de distribuição acumulada de X.

Solução:

(a) Como a área tem que ser 1, resulta

$$
1 = \frac{1}{2} \times (4-0) \times k \Rightarrow k = \frac{1}{2}
$$

(b) A função f_X é dada por 2 equações de reta. A primeira é uma reta de inclinação positiva que passa pelos pontos $(0,0)$ e $\left(2, \frac{1}{2}\right)$. A segunda é uma reta de inclinação negativa, que passa pelos pontos $\left(2, \frac{1}{2}\right)$ e $(4,0)$.

Para achar a equação de cada uma das retas, basta substituir as coordenadas dos dois pontos na equação geral de uma reta $f(x) = a + bx$ e resolver o sistema.

Para a primeira reta, temos o seguinte sistema:

$$\begin{aligned} 0 &= a + b \times 0 \\ \frac{1}{2} &= a + b \times 2 \end{aligned}$$

Da primeira equação resulta que $a = 0$ (é o ponto onde a reta cruza o eixo y) e substituindo esse valor de a na segunda equação, temos $b = \frac{1}{4}$.

Para a segunda reta, temos o seguinte sistema:

$$\begin{aligned} 0 &= a + b \times 4 \\ \frac{1}{2} &= a + b \times 2 \end{aligned}$$

Subtraindo a segunda equação da primeira, resulta:

$$0 - \frac{1}{2} = (a - a) + (4b - 2b) \Rightarrow b = -\frac{1}{4}$$

Substituindo na primeira equação, encontramos que $a = 1$.

Combinando essas duas equações, obtemos a seguinte expressão para f_X:

$$f_X(x) = \begin{cases} \dfrac{x}{4} & \text{se } 0 \leq x < 2 \\[2ex] 1 - \dfrac{x}{4} & \text{se } 2 \leq x \leq 4 \\[2ex] 0 & \text{se } x < 0 \text{ ou } x > 4 \end{cases}$$

(c) A probabilidade pedida é a área sombreada em cinza-escuro na Figura 4.7.

Figura 4.7 – Cálculo de $P(1 \leq X \leq 3)$

Os dois triângulos sombreados de cinza-claro têm a mesma área, por causa da simetria. Assim, podemos calcular a probabilidade usando a regra do complementar, uma vez que a área total é 1.

A altura dos dois triângulos é $\frac{1}{4} = f_X(1) = f_X(3)$. Logo, a área de cada um dos triângulos é $\frac{1}{2} \times 1 \times \frac{1}{4} = \frac{1}{8}$ e, portanto,

$$P(1 \leq X \leq 3) = 1 - 2 \times \frac{1}{8} = \frac{6}{8} = \frac{3}{4}$$

(d) Como a função é simétrica, resulta que $E(X) = 2$.

(e) O primeiro ponto a observar é o seguinte: o ponto $x = 2$ divide a área ao meio, ou seja, $x = 2$ é a mediana da distribuição. Como temos que $P(X \leq Q) = 0,6$, resulta que Q tem que ser maior que 2. Veja a Figura 4.8:

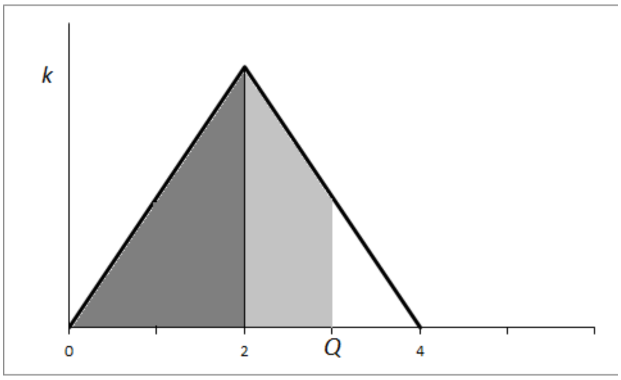


Figura 4.8 – Cálculo de Q tal que $P(X \leq Q) = 0,6$

Novamente, vamos usar a regra do complementar: como a área (probabilidade) abaixo de Q tem de ser 0,6, a área (probabilidade) acima de Q, dada pela área do triângulo superior, tem de ser 0,4. A altura desse triângulo é obtida substituindo-se o valor $x = Q$ na equação da segunda reta, o que nos dá $h = 1 - \frac{Q}{4}$. Substituindo na fórmula que dá a área de um triângulo, resulta:

$$0,4 = \frac{1}{2} \times (4 - k) \times \left(1 - \frac{k}{4}\right) \Rightarrow$$

$$\begin{aligned}
0,4 &= \frac{1}{2}\left(4 - k - k + \frac{k^2}{4}\right) \Rightarrow \\
0,8 &= \frac{16 - 8k + k^2}{4} \Rightarrow \\
3,2 &= k^2 - 8k + 16 \Rightarrow \\
& \quad k^2 - 8k + 13,2 = 0 \Rightarrow \\
k &= \frac{8 \pm \sqrt{64 - 4 \times 13,2}}{2} = \frac{8 \pm \sqrt{11,2}}{2}
\end{aligned}$$

A raiz $\frac{8+\sqrt{11,2}}{2}$ está fora do domínio de definição da função; logo, essa solução não serve. A solução para o problema, então, é:

$$Q = \frac{8 - \sqrt{11,2}}{2} = 2,3267$$

(f) Assim como a função densidade, a função de distribuição também será definida por 2 equações: uma para os valores de x no intervalo $[0,2)$ e outra para valores de x no intervalo $[2,4]$. Para $x \in [0,2)$, $F_X(x)$ é a área do triângulo sombreado na Figura 4.9 e para $x \in [2,4]$, $F_x(x)$ é a área sombreada na Figura 4.10.

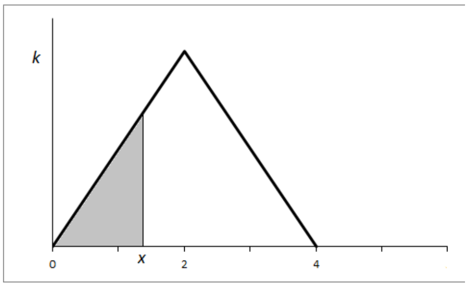


Figura 4.9 – Cálculo de $F_X(x)$ para $0 \leq x < 2$

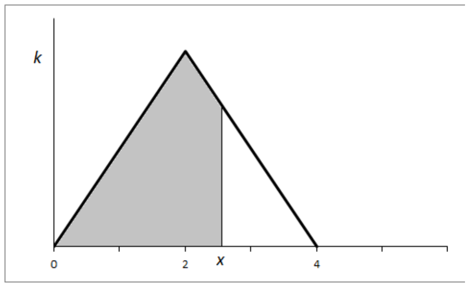


Figura 4.10 – Cálculo de $F_X(x)$ para $2 \leq x < 4$

Logo,

$$F_X(x) = \frac{1}{2}(x-0) \times \frac{x}{4} \qquad x \in [0,2)$$

$$F_X(x) = 1 - \frac{1}{2}(4-x)\left(1 - \frac{x}{4}\right).$$

Combinando os resultados obtidos, resulta a seguinte expressão para F_X :

$$F_X(x) = \begin{cases} 0 & \text{se } x < 0 \\ \frac{1}{8}x^2 & \text{se } 0 \leq x < 2 \\ 1 - \frac{1}{8}(4-x)^2 & \text{se } 2 \leq x \leq 4 \\ 1 & \text{se } x > 4 \end{cases}$$

Veja a Figura 4.11; para $0 \leq x < 2$, o gráfico de F_X é uma parábola côncava para cima; para $2 \leq x \leq 4$, o gráfico de F_X é uma parábola côncava para baixo.

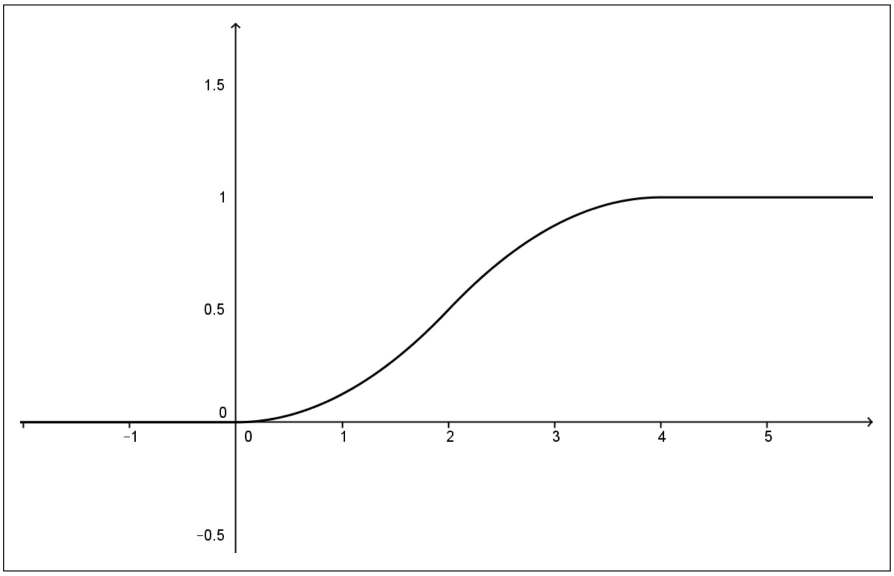


Figura 4.11 – Função de distribuição acumulada para a densidade triangular

♦♦

O fato da esperança da variável aleatória no exemplo anterior ter sido o ponto de simetria não foi uma coincidência, esse é um resultado geral sempre que a variável aleatória tiver função densidade simétrica. Veja Proposição 4.21 a seguir.

Proposição 4.21
Seja X uma variável aleatória com função densidade $f_X(x)$ simétrica em torno de μ, isto é,

$$f_X(\mu - x) = f_X(\mu + x).$$

Então, $E(X) = \mu$.

Demonstração:
Seja X variável aleatória com função densidade $f_X(x)$ simétrica em torno de μ. Vamos calcular a sua esperança de X.

- $\mu = 0$

  Nesse caso, $f_X(x)$ é uma função par, isto é, $f_X(x) = f_X(-x)$. Por definição,

  $$E(X) = \int_{-\infty}^{+\infty} x f_X(x) dx$$

  Vamos denotar o integrando por $h(x)$, isto é, $h(x) = x f_X(x)$. Resulta da simetria de f_X que

  $$h(-x) = -x f_X(-x) = -x f_X(x) = -h(x)$$

  ou seja, $h(x)$ é uma função ímpar e pelo Teorema A.10,

  $$\int_{-\infty}^{+\infty} h(x) dx = 0$$

  Provamos, assim, que se X é uma variável aletaória com densidade simétrica em torno de $x = 0$ então $E(X) = 0$.

- μ qualquer

  Vamos definir uma nova variável aleatória $Y = X - \mu$. Então,

  $$\begin{aligned} F_Y(y) &= P(Y \leq y) = P(X - \mu \leq y) = P(X \leq \mu + y) = F_X(\mu + y) \Rightarrow \\ F_Y'(y) &= f_Y(y) = F_X'(\mu + y) \cdot 1 = f_X(\mu + y) \end{aligned}$$

  Como f_X é simétrica em torno de μ, temos que

  $$f_Y(y) = f_X(\mu + y) = f_X(\mu - y) = f_Y(-y)$$

  ou seja, a densidade de Y é simétrica em torno de 0 e, portanto,

  $$E(Y) = 0 \Rightarrow E(X - \mu) = 0 \Rightarrow E(X) = \mu$$

□

Capítulo 5

Algumas Distribuições Contínuas

Assim como no caso discreto, algumas distribuições de variáveis aleatórias contínuas são bastante usadas e por isso merecem um estudo mais detalhado. Esse é o objetivo deste capítulo, estudar algumas distribuições contínuas.

5.1 Distribuição Uniforme

A motivação principal da distribuição Uniforme é definir a função densidade de uma variável aleatória X de forma que $\mathrm{P}(X \in I)$ dependa apenas do tamanho do intervalo I. Considere a função densidade f_X apresentada na Figura 5.1, em que a e b são números conhecidos.

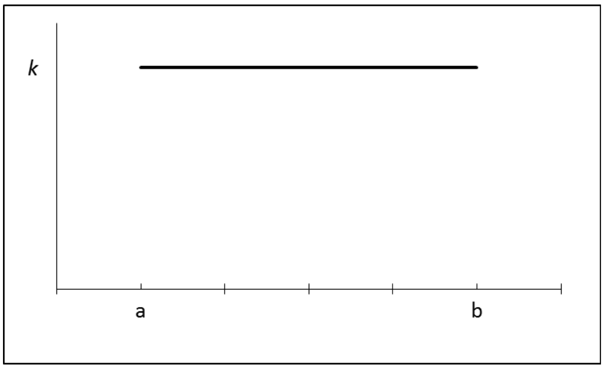


Figura 5.1 – Função densidade uniforme

Note que, para dois subintervalos de mesmo comprimento contidos em $[a, b]$, a área será igual, uma vez que temos áreas de retângulos com mesma altura. Assim, intervalos de mesmo comprimento têm a mesma probabilidade. Esse fato leva à denominação de tal densidade como *densidade uniforme*.

Qual deve ser o valor de k para que f_X seja função densidade?

A primeira condição é a função densidade é uma função não-negativa, logo isso implica em $k \geq 0$. A segunda condição é que a área e embaixo da curva de f_X tem que ser 1, o resulta em

$$1 = (b - a) \times k \Rightarrow k = \frac{1}{b - a}.$$

Definição 5.1 Distribuição Uniforme
Uma variável aleatória contínua X tem distribuição uniforme no intervalo $[a, b]$ se sua função densidade é dada por

$$f(x) = \begin{cases} \dfrac{1}{b-a} & \text{, se } x \in [a, b] \\ 0 & \text{, caso contrário.} \end{cases}$$

Os valores a e b são chamados parâmetros da distribuição uniforme.

Note que ambos os parâmetros da densidade uniforme têm de ser finitos para que a integral seja igual a 1. Além disso, $a \leq b$. Dessa forma, o espaço paramétrico para o vetor de parâmetros (a, b) é $\{a \in \mathbb{R}, b \in \mathbb{R}, a \leq b\}$.

Vamos denotar a distribuição Uniforme com parâmetros a e b por $U(a, b)$. Nesse caso, se quisermos indicar que uma variável aleatória contínua X segue a distribuição Uniforme no intervalo $[a, b]$ podemos simplesmente escrever $X \sim U(a, b)$ (lê-se: a variável aleatória X tem distribuição Uniforme no intervalo $[a, b]$).

Quando $a = 0$ e $b = 1$ temos a distribuição Uniforme Padrão, denotada por $U(0, 1)$.

Função Densidade e Função de Distribuição

Já vimos se $X \sim U(a, b)$ a sua função densidade é definida por:

$$f(x) = \begin{cases} \dfrac{1}{b-a} & \text{, se } x \in [a, b] \\ 0 & \text{, caso contrário.} \end{cases}$$

Por definição, a função de distribuição acumulada é

$$F_X(x) = \mathrm{P}(X \leq x)$$

e essa probabilidade é dada pela área sob a curva densidade à esquerda de x, conforme ilustrado na Figura 5.2.

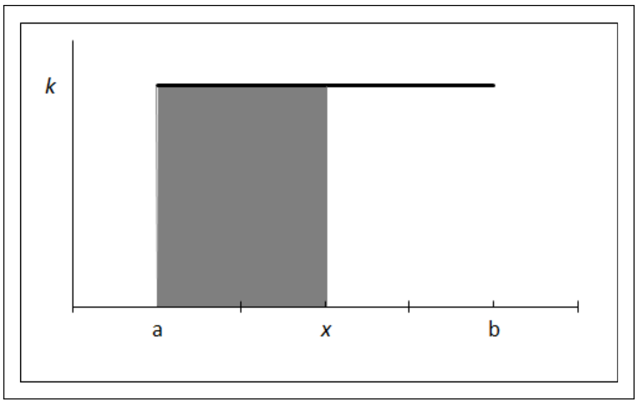


Figura 5.2 – Função de distribuição acumulada da densidade uniforme como área

Essa é a área de um retângulo com base $(x - a)$ e altura $\dfrac{1}{b-a}$. Logo,

$$F_X(x) = \begin{cases} 0 & , \text{se } x < a \\ \dfrac{x-a}{b-a} & , \text{se } a \le x \le b \\ 1 & , \text{se } x > b \end{cases} \tag{5.1}$$

O gráfico dessa função de distribuição acumulada é apresentado na Figura 5.3.

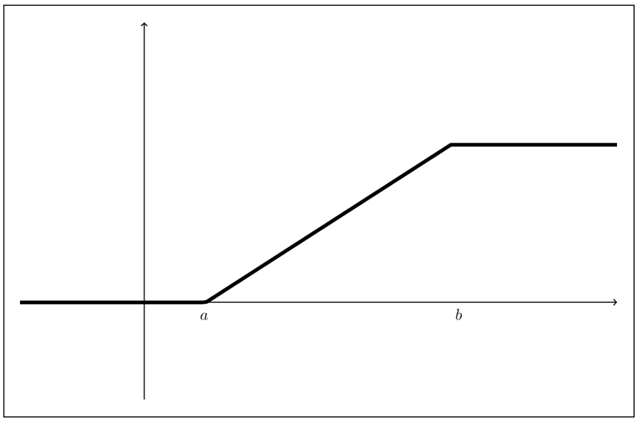


Figura 5.3 – Função de distribuição acumulada da densidade uniforme

Esperança e Variância

Das propriedades da esperança e das características da densidade uniforme (função simétrica), sabemos que $E(X)$ é o ponto médio do intervalo $[a, b]$. Logo

$$X \sim U(a, b) \qquad \Rightarrow \qquad E(X) = \frac{a+b}{2} \tag{5.2}$$

Vamos, agora, calcular $E(X^2)$.

$$E(X^2) = \int_a^b x^2 f(x)dx = \int_a^b x^2 \frac{1}{b-a} dx = \frac{b^3 - a^3}{3(b-a)} = \frac{a^2 + b^2 + ab}{3}.$$

Logo,

$$\begin{aligned} \operatorname{Var}(X) &= \frac{a^2 + b^2 + ab}{3} - \left(\frac{a+b}{2}\right)^2 = \frac{4a^2 + 4b^2 + 4ab - 3a^2 - 6ab - 3b^2}{12} \\ &= \frac{b^2 - 2ab + a^2}{12} = \frac{(b-a)^2}{12}. \end{aligned}$$

Assim chegamos na expressão,

$$X \sim U(a, b) \qquad \Rightarrow \qquad \operatorname{Var}(X) = \frac{(b-a)^2}{12} \tag{5.3}$$

Exemplo 5.2 Latas de coca-cola
Latas de coca-cola são enchidas num processo automático segundo uma distribuição uniforme no intervalo (em ml) [345,355].

(a) Qual é a probabilidade de uma lata conter mais de 353 ml?

(b) Qual é a probabilidade de uma lata conter menos de 346 ml?

(c) Qualquer lata com volume 4 ml abaixo da média pode gerar reclamação do consumidor e com volume 4 ml acima da média pode transbordar no momento de abertura, devido à pressão interna. Qual é a proporção de latas problemáticas?

Solução:
Seja X = "conteúdo da lata de coca-cola". Então, $X \sim \mathcal{U}[345, 355]$

(a) Pede-se
$$P(X > 353) = \frac{355 - 353}{355 - 345} = 0,2$$

(b) Pede-se
$$P(X < 346) = \frac{346 - 345}{355 - 345} = 0,1$$

(c) Pede-se
$$P(X < 350 - 4) + P(X > 350 + 4) = \frac{346 - 345}{355 - 345} + \frac{355 - 354}{355 - 345} = 0,2$$

Logo, a proporção de latas problemáticas é de 20%. Note que essa é uma proporção bastante alta!

♦♦

Exemplo 5.3 Determinação dos parâmetros
Uma variável aleatória X tem distribuição uniforme no intervalo $[a, b]$ com média 7,5 e variância 6,75. Determine os valores de a e b, sabendo que $b > a > 0$.

Solução:
$$\begin{cases} \dfrac{a+b}{2} = 7,5 \\ \\ \dfrac{(b-a)^2}{12} = 6,75 \end{cases}$$

Da segunda equação, resulta que $(b - a)^2 = 81$ e, portanto, $|b - a| = \sqrt{81}$. Como estamos supondo que $a < b$, resulta que $b - a = 9$, o que nos leva ao sistema

$$\begin{cases} a + b = 15 \\ b - a = 9 \end{cases}$$

cuja solução é $a = 3$ e $b = 12$, ou seja, $X \sim \mathcal{U}(3, 12)$.

♦♦

5.2 Distribuição Exponencial

Suponha que eventos ocorram ao longo do tempo e que o número de eventos ocorridos dentro de um intervalo de tempo unitário possa ser considerado uma variável aleatória com distribuição de Poisson de média λ. A motivação principal para a criação da distribuição exponencial vem da distribuição do tempo até a ocorrência do primeiro evento.

Seja T o instante em que ocorre o primeiro evento. Veja que T é variável aleatória. Quais os valores possíveis para T, ou seja, quem é $Im(T)$? $Im(T) = (0, \infty)$. Vamos encontrar a distribuição da variável aleatória T.

Primeiro vamos buscar a função de distribuição da variável aleatória T.

$$F_T(t) = \mathrm{P}(T \leq t) = 1 - \mathrm{P}(T > t) = 1 - \mathrm{P}(\text{não ocorrer evento algum em}(0, t)).$$

Seja X = número de ocorrências dentro do intervalo $(0, t)$. Então $X \sim Poi(\lambda t)$ e $p_X(x) = \dfrac{(\lambda t)^x}{x!} e^{-\lambda t}$. Assim,

$$\mathrm{P}(T > t) = \mathrm{P}(X = 0) = e^{-\lambda t} \qquad \Rightarrow \qquad F_T(t) = 1 - e^{-\lambda t}.$$

Para encontrar a função densidade basta derivar F_T em relação à t. Assim temos,

$$f_T(t) = \frac{d}{dt} F_T(t) = \lambda e^{-\lambda t}.$$

Uma variável aleatória com essa distribuição é chamada de Exponencial, como mostra a Definição 5.4 a seguir.

Definição 5.4 Distribuição Exponencial
Uma variável aleatória X contínua tem distribuição exponencial com parâmetro λ se sua função densidade de probabilidade é dada por

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & , x > 0 \\ 0 & , x \leq 0 \end{cases}$$

Como essa função depende apenas do valor de λ, esse é o parâmetro da densidade exponencial.

Veja que a distribuição Exponencial tem apenas um parâmetro, λ. O espaço paramétrico para ele é $\lambda > 0$, uma vez que λ indica o número médio de ocorrências em um intervalo de tempo unitário.

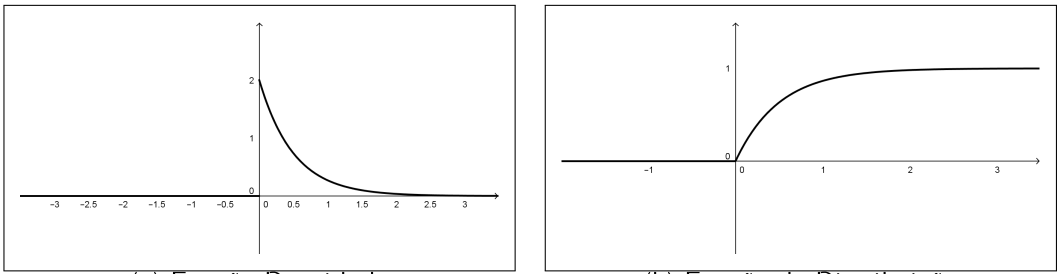
Usaremos a seguinte notação para indicar que uma variável aleatória tem distribuição exponencial com parâmetro λ: $X \sim \exp(\lambda)$.

Função Densidade e Função de Distribuição

No início dessa seção já encontramos a função densidade e a função de distribuição de uma variável aleatória $X \sim exp(\lambda)$. Só para relembrar, essas funções são

$$F_X(x) = 1 - e^{-\lambda x}, x > 0 \qquad \text{e} \qquad f_X(x) = \lambda e^{-\lambda x}, x > 0.$$

Veja nas Figuras 5.4(a) e 5.4(b) os gráficos da função densidade e de distribuição, respectivamente, para $X \sim exp(2)$.



(a) Função Densidade (b) Função de Distribuição

Figura 5.4 – Distribuição Exponencial com $\lambda = 2$.

Esperança e Variância

Vamos fazer as contas para encontrar $\mathrm{E}(X)$ e $\mathrm{Var}(X)$ para $X \sim exp(\lambda)$. Para o cálculo de $\mathrm{E}(X)$ faremos primeiro uma integração por partes e em seguida por substituição.

$$\mathrm{E}(X) = \int_0^\infty x\lambda e^{-\lambda x}dx = -xe^{-\lambda x}\Big|_0^\infty + \int_0^\infty e^{-\lambda x}dx = 0 - \lim_{x\to\infty} xe^{-\lambda x} - \frac{e^{-\lambda x}}{\lambda}\Big|_0^\infty = \frac{1}{\lambda}.$$

Portanto,

$$X \sim exp(\lambda) \qquad \Rightarrow \qquad \mathrm{E}\left(X\right) = \frac{1}{\lambda}. \tag{5.4}$$

Vamos agora calcular $\mathrm{E}(X^2)$ para depois encontrar $\mathrm{Var}(X)$. Novamente faremos uma integração por partes.

$$\begin{aligned} \mathrm{E}(X^2) &= \int_0^\infty x^2\lambda e^{-\lambda x}dx = -x^2e^{-\lambda x}\Big|_0^\infty + \int_0^\infty 2xe^{-\lambda x}dx \\ &= 0 - \lim_{x\to\infty} x^2e^{-\lambda x} + 2\int_0^\infty xe^{-\lambda x}dx = \frac{2}{\lambda}\int_0^\infty x\lambda e^{-\lambda x}dx = \frac{2}{\lambda}\mathrm{E}(X) = \frac{2}{\lambda^2}. \end{aligned}$$

Assim chegamos no resultado,

$$\mathrm{Var}(X) = \mathrm{E}(X^2) - \mathrm{E}(X)^2 = \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}.$$

Portanto,

$$X \sim exp(\lambda) \qquad \Rightarrow \qquad \mathrm{Var}\left(X\right) = \frac{1}{\lambda^2}. \tag{5.5}$$

Exemplo 5.5 Relação entre Poisson e Exponencial
Suponha que o número de acidentes em uma rodovia seja uma variável aleatória de Poisson com média de 1 acidente a cada 3 dias.

(a) Qual a probabilidade do primeiro acidente do mês acontecer antes do terceiro dia?

(b) Qual a probabilidade de não ser registrado acidente algum na primeira semana do mês?

Solução:
Cada um dos dois itens pode ser feito usando a distribuição Poisson ou Exponencial. Faremos dos dois jeitos.

(a) Primeiro faremos pensando na distribuição de X = número de acidentes nos primeiros 2 dias do mês. Queremos encontrar $P(X > 0)$. Veja que $X \sim Poi(\lambda = 2/3)$. Logo, $P(X > 0) = 1 - P(X = 0) = 1 - e^{-2/3}$.

Podemos resolver o mesmo exercício a partir da variável Y = intervalo de tempo (em dias) até o primeiro acidente. Queremos $P(Y < 2)$. Veja que $Y \sim exp(\lambda = 1/3)$. Logo, $P(Y < 2) = F_Y(2) = 1 - e^{-2/3}$.

(b) Vamos primeiro resolver a partir de X = número de acidentes na primeira semana. Queremos $P(X = 0)$. Veja que$X \sim Poi(\lambda = 7/3)$. Logo, $P(X = 0) = e^{-7/3}$.

Podemos resolver o mesmo exercício a partir da variável Y = intervalo de tempo (em dias) até o primeiro acidente. Queremos $P(Y > 7)$. Veja que $Y \sim exp(\lambda = 1/3)$. Logo, $P(Y > 7) = 1 - F_Y(7) = e^{-7/3}$.

♦♦

Exemplo 5.6
Seja X uma variável aleatória exponencial com média 4. Calcule

(a) $P(X > 1)$

(b) $P(1 \leq X \leq 2)$

Solução:

(a) A função densidade de X é $f_X(x) = \frac{1}{4}e^{-x/4}$ e a função de distribuição é $F(x) = 1 - e^{-x/4}$. Podemos resolver essa questão por integração da função densidade

$$P(X > 1) = \int_1^\infty f(x)dx = \int_1^\infty \frac{1}{4}e^{-\frac{x}{4}}dx = -e^{-\frac{x}{4}}\Big|_1^\infty = 0 - (-e^{0,25}) = e^{-0,25} = 0,7788$$

ou pelo uso direto da função de distribuição

$$P(X > 1) = 1 - P(X \leq 1) = 1 - F(1) = 1 - [1 - e^{-1/4}] = e^{-0,25} = 0,7788$$

(b)

$$\begin{aligned} P(1 \leq X \leq 2) &= P(X \leq 2) - P(X < 1) \\ &= P(X \leq 2) - P(X \leq 1) \\ &= F(2) - F(1) = [1 - e^{-2/4}] - [1 - e^{-1/4}] \\ &= e^{-0,25} - e^{-0,5} = 0,17227 \end{aligned}$$

♦♦

Parametrização Alternativa

Muitos livros apresentam a distribuição exponencial em termos do parâmetro $\beta = \frac{1}{\lambda} > 0$. Neste caso, a função densidade passa a ser

$$f_X(x) = \begin{cases} \frac{1}{\beta}e^{-x/\beta} & , x > 0 \\ 0 & x \leq 0. \end{cases}$$

Além disso, usando essa nova parametrização temos:

$$E(X) = \beta \;, \qquad E(X^2) = 2\beta^2 \quad \text{e} \quad Var(X) = \beta^2.$$

Exemplo 5.7
Seja X variável aleatória com distribuição exponencial. Calcule $\mathrm{P}(X > E(X))$.

Solução:
Vamos usar a notação alternativa, mas poderia ser resolvida pela notação tradicional sem problema algum.

$$\mathrm{P}(X > E(X)) = 1 - \mathrm{P}(X \leq E(X)) = 1 - F(\mathrm{E}(X)) = 1 - \left[1 - e^{-\beta/\beta}\right] = e^{-1}$$

Note que essa é a probabilidade de uma variável aleatória exponencial ser maior que o seu valor médio; o que mostramos é que essa probabilidade é constante, qualquer que seja o parâmetro.

♦♦

Propriedade de Falta de Memória

Nessa seção vamos apresentar o conceito de falta de memória de uma variável aleatória contínua. Em seguida vamos mostrar que a distribuição exponencial segue tal propriedade.

Definição 5.8 Propriedade de Falta de Memória - v.a. Contínua
Seja X uma variável aleatória contínua e a e δ números reais quaisquer. Se,

$$\mathrm{P}(X > a + \delta \mid X > a) = \mathrm{P}(X > \delta)$$

então dizemos que X tem a propriedade de ***falta de memória****.*

Queremos mostrar nessa seção que uma variável aleatória com distribuição exponencial tem a propriedade de falta de memória.

Seja $X \sim \exp(\lambda)$, então

$$\begin{aligned}
\mathrm{P}(X > a + \delta \mid X > a) &= \frac{\mathrm{P}(X > a + \delta \text{ e } X > a)}{\mathrm{P}(X > a)} = \frac{\mathrm{P}(X > a + \delta)}{\mathrm{P}(X > a)} \\
&= \frac{1 - \mathrm{P}(X \leq a + \delta)}{1 - \mathrm{P}(X \leq a)} = \frac{e^{-\lambda(a+\delta)}}{e^{-\lambda a}} \\
&= \frac{e^{-\lambda a} e^{-\lambda \delta)}}{e^{-\lambda a}} = e^{-\lambda \delta} \\
&= \mathrm{P}(X > \delta).
\end{aligned}$$

Exemplo 5.9
Ao comprar um carro usado você não perguntou para o antigo dono quando foi a última vez que ele trocou a bateria. O fabricante da bateria diz que o tempo de vida de cada bateria pode ser considerado uma variável aleatória exponencial com média de 4 anos. Sabendo que a bateria está funcionando quando o carro foi comprado, qual a probabilidade de você não precisar trocar essa bateria pelos próximos 2 anos?

Solução:
Considere X = tempo de vida (em anos) da bateria e a = tempo (em anos) desde que a bateria atual foi instalada até a compra do carro. O que queremos encontrar é

$$\mathrm{P}(X > a + 2 | X > a).$$

Sabemos que $X \sim exp(\lambda = 1/4)$, mas não conhecemos o valor de a. Porém essa informação não é necessária, uma vez que a distribuição exponencial tem a propriedade de falta de memória. Nesse caso temos, pela propriedade da falta de memória (Definição 5.8),

$$P(X > a + 2|X > a) = P(X > 2) = 1 - P(X \leq 2) = 1 - F_X(2) = e^{-\frac{1}{4}2} = e^{-\frac{1}{2}} = 0,6065.$$

♦♦

5.3 Distribuição Gama

A Função Gama

Para apresentar a distribuição Gama primeiro precisamos conhecer a função Gama e suas propriedades.

Definição 5.10 Função Gama
Chama-se função Gama a função de variável real, estritamente positiva, definida por:

$$\Gamma(\alpha) = \int_0^\infty x^{\alpha-1}e^{-x}dx \qquad , \alpha > 0$$

Veja que a função Γ é não negativa, uma vez que a função $x^{\alpha-1}e^{-x} \geq 0$, qualquer que seja $\alpha > 0$ e $x > 0$.

Além disso, o que não será demonstrado nessa apostila, $\Gamma(\alpha)$ é um número real sempre que $\alpha > 0$, isto é, a integral acima converge qualquer que seja $\alpha > 0$. E se $\alpha \leq 0$ a integral $\int_0^\infty x^{\alpha-1}e^{-x}dx$ não converge, isto é, $\int_0^\infty x^{\alpha-1}e^{-x}dx = \infty$.

A seguir veremos 3 propriedades importantes, e bastante úteis, da função Γ.

Proposição 5.11
Seja Γ a função apresentada na Definição 5.10. Então valem as seguintes propriedades.

(G1) Dados $\alpha > 0$ e $\lambda > 0$, $\dfrac{\Gamma(\alpha)}{\lambda^\alpha} = \int_0^\infty x^{\alpha-1}e^{-\lambda x}dx$.

(G2) Dado $\alpha > 1$, $\Gamma(\alpha) = (\alpha - 1)\Gamma(\alpha - 1)$.

(G3) Dado $n \in \{1, 2, 3, \ldots\}$, $\Gamma(n) = (n - 1)!$.

Demonstração:

(G1) Para mostrar essa propriedade basta fazer contas e mostrar que $\Gamma(\alpha) = \lambda^\alpha \int_0^\infty x^{\alpha-1}e^{-\lambda x}dx$, qualquer que seja $\lambda > 0$. Para isso, nas contas a seguir, logo de início faremos a substituição $y = x/\lambda$ (ou $x = \lambda y$), supondo $\lambda > 0$.

$$\Gamma(\alpha) = \int_0^\infty x^{\alpha-1}e^{-x}dx = \int_0^\infty (\lambda y)^{\alpha-1}e^{-\lambda y}\lambda dy = \lambda^\alpha \int_0^\infty y^{\alpha-1}e^{-\lambda y}dy$$

(G2) Vamos tentar resolver a integral que define a função Γ por partes. Logo de início faremos então: $u = x^{\alpha-1} \Rightarrow du = (\alpha - 1)x^{\alpha-2}dx$ e $dv = e^{-x}dx \Rightarrow v = -e^{-x}$.

$$\begin{aligned}
\Gamma(\alpha) &= \int_0^\infty x^{\alpha-1}e^{-x}dx = -x^{\alpha-1}e^{-x}\Big|_0^\infty + \int_0^\infty e^{-x}(\alpha-1)x^{\alpha-2}dx \\
&= \int_0^\infty e^{-x}(\alpha-1)x^{\alpha-2}dx = (\alpha-1)\int_0^\infty x^{\alpha-2}e^{-x}dx = (\alpha-1)\Gamma(\alpha-1).
\end{aligned}$$

(G3) Por último vamos considerar $\alpha = n$ inteiro positivo. Usando o resultado da Propriedade (G2) podemos demonstrar, por indução, que $\Gamma(n) = (n-1)!$.

Primeiro veja que a afirmação é verdadeira para $n = 1$, isto é, que $\Gamma(1) = 0! = 1$.

$$\Gamma(1) = \int_0^\infty x^{1-1} e^{-x} dx = \int_0^\infty e^{-x} dx = 1.$$

Agora, supondo que é verdade para n qualquer, isto, supondo que $\Gamma(n) = (n-1)!$, vejamos que isso implica em ser verdade para $n+1$. Vamos começar usando o resultado da Propriedade (G2).

$$\Gamma(n+1) = n\Gamma(n) = n(n-1)! = n!$$

□

Função Densidade

Definição 5.12 Distribuição Gama
Uma variável aleatória contínua X tem distribuição gama com parâmetros $\alpha > 0$ e $\lambda > 0$ se sua função densidade é dada por

$$f_X(x) = \begin{cases} \dfrac{1}{\Gamma(\alpha)} x^{\alpha-1} \lambda^\alpha e^{-\lambda x} & \text{, se } x > 0 \\ 0 & \text{, caso contrário.} \end{cases}$$

Note que, quando $\alpha = 1$, resulta a densidade exponencial com parâmetro λ, ou seja, a densidade exponencial é um caso particular da densidade gama.

Os parâmetros da distribuição gama são α e λ. O espaço paramétrico para tais parâmetros é definido por $\alpha > 0$ e $\lambda > 0$. Usaremos a notação $X \sim gama(\alpha; \beta)$ para indicar que a variável aleatória X tem distribuição gama com parâmetros α, β.

Para verificar que a função apresentada na Definição 5.12 realmente define uma função densidade, notamos inicialmente que $f(x) \geq 0$. Além disso, fazendo o uso da propriedade (G1), apresentada na Proposição 5.11, temos

$$\int_{-\infty}^\infty f_X(x) dx = \int_0^\infty \frac{1}{\Gamma(\alpha)} x^{\alpha-1} \lambda^\alpha e^{-\lambda x} = \frac{\lambda^\alpha}{\Gamma(\alpha)} \underbrace{\int_0^\infty x^{\alpha-1} e^{-\lambda x}}_{\frac{\Gamma(\alpha)}{\lambda^\alpha}} = 1.$$

Logo, as duas condições para uma função densidade são satisfeitas.

Esperança e Variância

Seja $X \sim gama(\alpha, \lambda)$. Queremos encontrar $\mathrm{E}(X)$ e $\mathrm{Var}(X)$. Vamos começar pela esperança. Nas contas a seguir faremos o uso das Propriedades ?? e (G2) apresentadas na Proposição 5.11.

$$\begin{aligned} \mathrm{E}(X) &= \int_{-\infty}^\infty x f_X(x) dx = \int_0^\infty x \frac{1}{\Gamma(\alpha)} x^{\alpha-1} \lambda^\alpha e^{-\lambda x} = \frac{\lambda^\alpha}{\Gamma(\alpha)} \int_0^\infty x^\alpha e^{-\lambda x} \\ &= \frac{\lambda^\alpha}{\Gamma(\alpha)} \frac{\Gamma(\alpha+1)}{\lambda^{\alpha+1}} = \frac{\lambda^\alpha}{\Gamma(\alpha)} \frac{\alpha\Gamma(\alpha)}{\lambda^{\alpha+1}} = \frac{\alpha}{\lambda} \end{aligned}$$

Portanto,

$$X \sim gama(\alpha, \lambda) \qquad \Rightarrow \qquad \mathrm{E}\,(X) = \frac{\alpha}{\lambda}. \tag{5.6}$$

De modo análogo, vamos calcular $\mathrm{E}(X^2)$.

$$\begin{aligned}
\mathrm{E}(X^2) &= \int_{-\infty}^{\infty} x^2 f_X(x)dx = \int_0^{\infty} x^2 \frac{1}{\Gamma(\alpha)} x^{\alpha-1}\lambda^{\alpha} e^{-\lambda x} = \frac{\lambda^{\alpha}}{\Gamma(\alpha)} \int_0^{\infty} x^{\alpha+1} e^{-\lambda x} \\
&= \frac{\lambda^{\alpha}}{\Gamma(\alpha)} \frac{\Gamma(\alpha+2)}{\lambda^{\alpha+2}} = \frac{\lambda^{\alpha}}{\Gamma(\alpha)} \frac{\alpha(\alpha+1)\Gamma(\alpha)}{\lambda^{\alpha+2}} = \frac{\alpha(\alpha+1)}{\lambda^2}
\end{aligned}$$

Logo,

$$\mathrm{Var}(X) = \mathrm{E}(X^2) - \mathrm{E}(X)^2 = \frac{\alpha(\alpha+1)}{\lambda^2} - \frac{\alpha^2}{\lambda^2} = \frac{\alpha(\alpha+1)-\alpha^2}{\lambda^2} = \frac{\alpha}{\lambda^2}$$

Portanto,

$$X \sim gama(\alpha, \lambda) \qquad \Rightarrow \qquad \mathrm{Var}\,(X) = \frac{\alpha}{\lambda^2}. \tag{5.7}$$

Parametrização Alternativa

Assim como no caso da distribuição exponencial, a distribuição gama também tem uma parametrização alternativa, onde em vez do parâmetro λ é usado o parâmetro $\beta = \frac{1}{\lambda} > 0$. Neste caso, a função densidade passa a ser

$$f_X(x) = \begin{cases} \dfrac{1}{\Gamma(\alpha)\beta^{\alpha}} x^{\alpha-1} e^{-x/\beta} & \text{, se } x > 0 \\ 0 & \text{, caso contrário.} \end{cases}$$

Além disso, usando essa nova parametrização temos:

$$\mathrm{E}(X) = \alpha\beta \ , \qquad \mathrm{E}(X^2) = \alpha(\alpha+1)\beta^2 \quad \text{e} \quad \mathrm{Var}(X) = \alpha\beta^2.$$

Os Gráficos da Função Densidade da Gama

Nessa seção foi considerada a parametrização alternativa para a distribuição gama.

Qualquer que seja o valor do parâmetro β, temos que

$$\begin{aligned}
\lim_{x\to\infty} f(x) &= 0 \\
\lim_{x\to 0} f(x) &= \begin{cases} 0 & \text{, se } \alpha \geq 1 \\ \infty & \text{, se } \alpha < 1 \end{cases}
\end{aligned}$$

Vamos, agora, calcular as derivadas primeira e segunda de f para estudar pontos de máximo, mínimo, regiões de crescimento ou decrescimento, concavidade.

Derivada primeira

$$
\begin{aligned}
f'(x) &= \frac{1}{\Gamma(\alpha)\beta^\alpha}\left[(\alpha-1)x^{\alpha-2}e^{-x/\beta} - \frac{1}{\beta}x^{\alpha-1}e^{-x/\beta}\right] \\
&= \frac{1}{\Gamma(\alpha)\beta^\alpha}\left[\left(x^{\alpha-2}e^{-x/\beta}\right)\left(\alpha-1-\frac{1}{\beta}x\right)\right] \qquad (5.8)
\end{aligned}
$$

O sinal da derivada primeira depende do sinal de

$$g(x) = \alpha - 1 - \frac{1}{\beta}x$$

Vamos analisar a derivada primeira estudando o comportamento de g.

- $\alpha \leq 1$

  Como $x > 0$, resulta que $f'(x) < 0$, ou seja, se $\alpha \leq 1$ a densidade gama é uma função estritamente decrescente.

- $\alpha > 1$

$$
\begin{aligned}
f'(x) &= 0 \Leftrightarrow x = \beta(\alpha-1) \\
f'(x) &> 0 \Leftrightarrow x < \beta(\alpha-1) \qquad f \text{ crescente} \\
f'(x) &< 0 \Leftrightarrow x > \beta(\alpha-1) \qquad f \text{ decrescente}
\end{aligned}
$$

  Logo,

$$x_0 = \beta(\alpha-1) \text{ é um ponto de máximo}$$

Derivada segunda

$$
\begin{aligned}
f''(x) &= \frac{1}{\Gamma(\alpha)\beta^\alpha}\left[\begin{array}{c}(\alpha-2)(\alpha-1)x^{\alpha-3}e^{-x/\beta} - \frac{1}{\beta}(\alpha-1)x^{\alpha-2}e^{-x/\beta} - \frac{1}{\beta}(\alpha-1)x^{\alpha-2}e^{-x/\beta} \\ +\frac{1}{\beta^2}x^{\alpha-1}e^{-x/\beta}\end{array}\right] \\
&= \frac{1}{\Gamma(\alpha)\beta^\alpha}\left\{\left(x^{\alpha-3}e^{-x/\beta}\right)\left[(\alpha-2)(\alpha-1) - \frac{2}{\beta}(\alpha-1)x + \frac{1}{\beta^2}x^2\right]\right\} \\
&= \frac{1}{\Gamma(\alpha)\beta^\alpha}\left\{\frac{\left(x^{\alpha-3}e^{-x/\beta}\right)}{\beta^2}\left[\beta^2(\alpha-2)(\alpha-1) - 2\beta(\alpha-1)x + x^2\right]\right\} \qquad (5.9)
\end{aligned}
$$

O sinal da derivada segunda depende do sinal da expressão entre colchetes, que é uma função do segundo grau. Vamos denotar essa expressão por $h(x)$, de modo que

$$h(x) = x^2 - 2\beta(\alpha-1)x + \beta^2(\alpha-2)(\alpha-1)$$

Vamos, agora, analisar o sinal da derivada segunda da densidade gama, estudando o sinal do discriminante da equação dada pela expressão de $h(x)$, para analisar a concavidade da densidade gama:

$$\Delta = 4\beta^2(\alpha-1)^2 - 4\beta^2(\alpha-2)(\alpha-1) = 4\beta^2(\alpha-1)(\alpha-1-\alpha+2) = 4\beta^2(\alpha-1)$$

- $\alpha < 1$

  Nesse caso, o discriminante é negativo e a equação não tem raízes reais e terá sempre o sinal do coeficiente do termo quadrático, que é 1. Assim, neste caso, a concavidade é para cima. Resumindo, se $\alpha < 1$ a densidade gama é decrescente com concavidade para cima.

- $\alpha = 1$

  Nesse caso, o discriminante é nulo e $h(x) = x^2 \geq 0$; logo, a concavidade é para cima, ou seja, se $\alpha = 1$ a densidade gama é decrescente com concavidade para cima.

- $\alpha > 1$

  Neste caso, o discriminante é sempre positivo, ou seja, temos duas raizes reais distintas, calculadas da seguinte forma:

$$
\begin{aligned}
(\alpha-2)(\alpha-1) - \frac{2}{\beta}(\alpha-1)x + \frac{1}{\beta^2}x^2 = 0 &\Longleftrightarrow \\
\beta^2(\alpha-2)(\alpha-1) - 2\beta(\alpha-1)x + x^2 = 0 &\Longleftrightarrow \\
x = \frac{2\beta(\alpha-1) \pm \sqrt{}}{2} &\Longleftrightarrow \\
x = \frac{2\beta(\alpha-1) \pm 2\beta\sqrt{(\alpha-1)(\alpha-1-\alpha+2)}}{2} &\Longleftrightarrow \\
x = \beta(\alpha-1) \pm \beta\sqrt{\alpha-1} &\Longleftrightarrow \\
x = \beta\sqrt{\alpha-1}\left(\sqrt{\alpha-1} \pm 1\right)
\end{aligned}
$$

o que fornece as raízes

$$
\begin{aligned}
r_1 &= \beta\sqrt{\alpha-1}\left(\sqrt{\alpha-1} - 1\right) \\
r_2 &= \beta\sqrt{\alpha-1}\left(\sqrt{\alpha-1} + 1\right)
\end{aligned}
$$

A raiz é r_2 é sempre positiva. Já a raiz r_1 só será positiva se $\sqrt{\alpha-1} - 1 > 0$, ou seja, se $\alpha > 2$.

Considerando a função de segundo grau $h(x)$ que define o sinal da derivada segunda, vemos que o coeficiente do termo quadrático é 1; assim, a função é negativa (sinal oposto ao de a) para valores de x entre as raízes, e positiva (mesmo sinal de a) fora das raízes. Veja a Figura 5.5; aí podemos ver que, se $\alpha > 2$, a derivada segunda muda de sinal em dois pontos dentro do domínio de definição da densidade gama. Isso não ocorre se $\alpha < 2$ (ou $\alpha = 2$), uma vez que, neste caso a menor raíz é negativa (nula).

Mais precisamente, temos a seguinte situação:

$\star$ $\alpha > 2$

$$
\begin{aligned}
f''(x) &< 0 \text{ se } r_1 < x < r_2 \\
f''(x) &> 0 \text{ se } x > r_2 \text{ ou } x < r_1
\end{aligned}
$$

ou seja, a função densidade é côncava para cima se se $x > \beta\sqrt{\alpha-1}\left(\sqrt{\alpha-1}+1\right)$ ou $x < \beta\sqrt{\alpha-1}\left(\sqrt{\alpha-1}-1\right)$ e é côncava para baixo se $\beta\sqrt{\alpha-1}\left(\sqrt{\alpha-1}-1\right) < x < \beta\sqrt{\alpha-1}\left(\sqrt{\alpha-1}+1\right)$, o que indica a ocorrência de dois pontos de inflexão: r_1 e r_2.

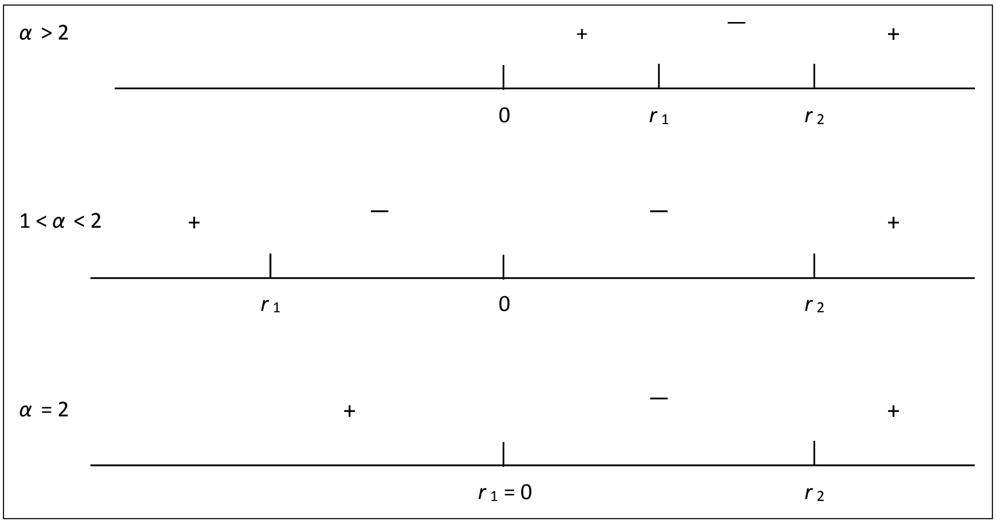


Figura 5.5 – Ilustração do sinal da derivada segunda da função densidade gama para $\alpha > 1$.

⋆ $1 < \alpha \leq 2$

$$f''(x) < 0 \text{ se } 0 < x < r_1$$
$$f''(x) > 0 \text{ se } x > r_2$$

ou seja, a função densidade gama é côncava para cima se $x > \beta\sqrt{\alpha-1}\left(\sqrt{\alpha-1}+1\right)$ e é côncava para baixo se $0 < x < \beta\sqrt{\alpha-1}\left(\sqrt{\alpha-1}-1\right)$, o que indica a ocorrência de apenas um ponto de inflexão em r_2.

Na tabela 5.1 temos um resumo do comportamento da densidade gama em função do parâmetro α.

Tabela 5.1 – Características do gráfico da densidade gama em função do parâmetro α

<table>
<tr><td>$0 < \alpha \leq 1$</td><td>Estritamente decrescente</td><td>Côncava para cima</td></tr>
<tr><td>$1 < \alpha \leq 2$</td><td rowspan="2">Ponto de máximo:
• $x_0 = \beta(\alpha-1)$

Crescente
• $x < \beta(\alpha-1)$

Decrescente
• $x > \beta(\alpha-1)$</td><td>Ponto de inflexão:
• $r_2 = \beta\sqrt{\alpha-1}\left(\sqrt{\alpha-1}+1\right)$
Concavidade:
• para cima se $x > r_2$
• para baixo se $x < r_2$</td></tr>
<tr><td>$\alpha > 2$</td><td>Pontos de inflexão:
• $r_1 = \beta\sqrt{\alpha-1}\left(\sqrt{\alpha-1}-1\right)$
• $r_2 = \beta\sqrt{\alpha-1}\left(\sqrt{\alpha-1}+1\right)$
Concavidade:
• para cima se $x < r_1$ ou $x > r_2$
• para baixo se $r_1 < x < r_2$</td></tr>
</table>

Em cada uma das Figuras 5.6 e 5.7, o parâmetro β está fixo e temos o gráfico da função densidade para $\alpha = 1, 2, 4$. Nas Figuras 5.8 e 5.9, fixamos o parâmetro α e variamos β. Analisando esses gráficos,podemos ver que α tem grande influência sobre a forma da distribuição, enquanto β afeta mais fortemente a dispersão. Por essas razões, α é chamado *parâmetro de forma* da densidade gama e β é o *parâmetro de escala*.

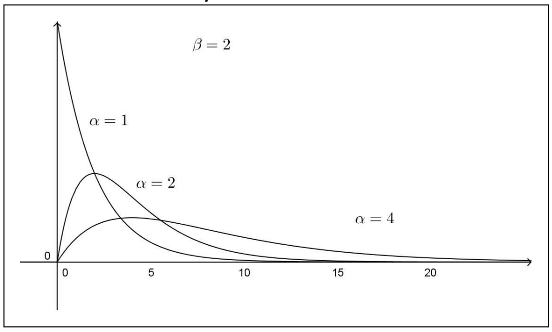


Figura 5.6 – Efeito do parâmetro α – $\beta = 2; \alpha = 1, 2, 4$

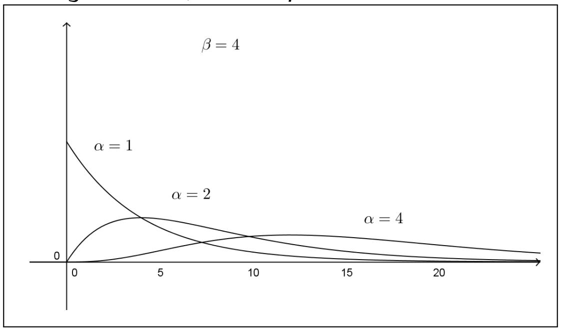


Figura 5.7 – Efeito do parâmetro α – $\beta = 4; \alpha = 1, 2, 4$

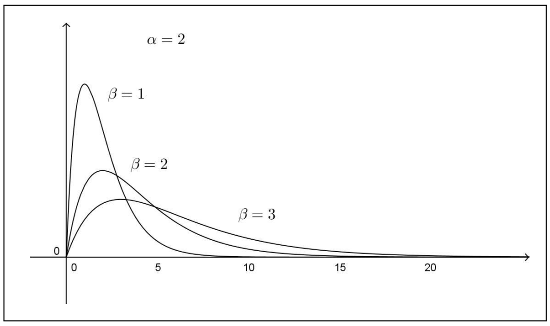


Figura 5.8 – Efeito do parâmetro β – $\alpha = 2; \beta = 1, 2, 3$

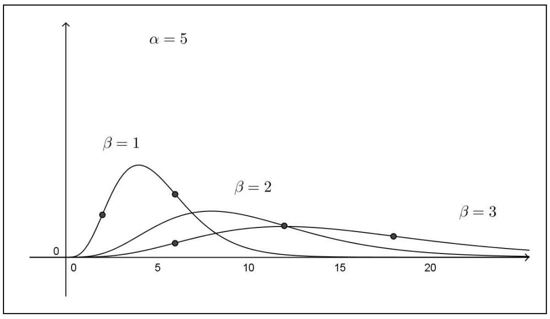


Figura 5.9 – Efeito do parâmetro β – $\alpha = 5; \beta = 1, 2, 3$

Na Figura 5.9 estão marcados os 2 pontos de inflexão de cada densidade.

Casos Particulares da Distribuição Gama

Nessa seção voltamos a trabalhar com a parametrização tradicional da gama, com os parâmetros α e λ. Veremos a seguir três casos particulares da distribuição Gama, entre eles a distribuição Exponencial, como já comentado ao longo do texto.

Distribuição Exponencial

Quando, na distribuição gama, o parâmetro $\alpha = 1$ temos a distribuição exponencial com parâmetro λ. Nesse caso, para $x > 0$, sua função densidade é definida por

$$f_X(x) = \frac{1}{\Gamma(1)} x^{1-1} \lambda^1 e^{-\lambda x} = \lambda e^{-\lambda x}.$$

Além disso,

$$\mathrm{E}(X) = \frac{1}{\lambda} \quad \text{e} \quad \mathrm{Var}(X) = \frac{1}{\lambda^2}.$$

Veja que os resultados acima batem com aqueles encontrados na seção 5.2.

Distribuição de Erlang

Quando o parâmetro α da distribuição gama é um inteiro positivo, $\alpha = k$, a distribuição gama é conhecida como **distribuição de Erlang** de ordem k e sua função densidade, para $x > 0$, é definida por:

$$f_X(x) = \frac{1}{\Gamma(k)} x^{k-1} \lambda^k e^{-\lambda x} = \frac{1}{(k-1)!} x^{k-1} \lambda^k e^{-\lambda x}$$

Ou seja, se X tem distribuição de Erlang de ordem k e parâmetro λ (notação: $X \sim Erl_k(\lambda)$), a sua função densidade é:

$$f_X(x) = \begin{cases} \frac{1}{(k-1)!} x^{k-1} \lambda^k e^{-\lambda x} & , \text{se } x > 0 \\ 0 & , \text{caso contrário.} \end{cases} \tag{5.10}$$

Além disso,

$$\mathrm{E}(X) = \frac{k}{\lambda} \quad \text{e} \quad \mathrm{Var}(X) = \frac{k}{\lambda^2}.$$

Distribuição Qui-quadrado

Quando, na distribuição gama, o parâmetro de forma α é igual a $\frac{n}{2}$, com n inteiro positivo, e o parâmetro λ é $\frac{1}{2}$ a distribuição é chamada de **distribuição Qui-quadrado** com n graus de liberdade e a sua função densidade, para $x > 0$ é definida por:

$$f_X(x) = \frac{1}{\Gamma(n/2)} x^{(n/2)-1} \left(\frac{1}{2}\right)^{n/2} e^{-(1/2)x} = \frac{1}{\Gamma(n/2) 2^{n/2}} x^{(n/2)-1} e^{-x/2}.$$

Ou seja, se X tem distribuição Qui-quadrado com n graus de liberdade (notação: $X \sim \chi^2_n$), a sua função densidade é

$$f_X(x) = \begin{cases} \frac{1}{\Gamma(n/2) 2^{n/2}} x^{(n/2)-1} e^{-x/2} & , \text{se } x > 0 \\ 0 & , \text{caso contrário.} \end{cases} \tag{5.11}$$

Além disso,

$$\mathrm{E}(X) = \frac{n/2}{1/2} = n \quad \text{e} \quad \mathrm{Var}(X) = \frac{n/2}{(1/2)^2} = 2n.$$

5.4 Distribuição Beta

A Função Beta

Antes de apresentarmos a distribuição Beta precisamos conhecer a função Beta.

Definição 5.13 Função Beta
Chama-se função $Beta$ a função de duas variáveis, estritamente positiva, definida por:

$$Beta(\alpha, \beta) = \int_0^1 x^{\alpha-1}(1-x)^{\beta-1} dx \qquad , \alpha > 0 \text{ e } \beta > 0.$$

Proposição 5.14 **Relação entre função Beta e função Gama**

$$Beta(\alpha, \beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha + \beta)}.$$

Demonstração:
A demonstração desta proposição utiliza integração dupla e como esse conceito ainda não foi visto pela maioria dos alunos, ela será omitida.

□

Como consequência imediata da Proposição 5.14, temos os resultados apresentados no Corolário 5.15 a seguir.

Corolário 5.15 *Seja $Beta$ a função apresentada na Definição 5.13. Então,*

(i) $Beta$ é uma função real, isto é, $\forall\ \alpha > 0$ e $\beta > 0$ tem-se $Beta(\alpha, \beta) < \infty$.

(ii) $Beta(\alpha, \beta) = Beta(\beta, \alpha)$, isto é, $Beta$ é uma função comutativa.

Função Densidade

Definição 5.16 **Distribuição Beta**
Uma variável aleatória contínua X tem distribuição beta com parâmetros $\alpha > 0$ e $\beta > 0$ se sua função densidade é dada por

$$f_X(x) = \begin{cases} \dfrac{1}{Beta(\alpha, \beta)} x^{\alpha-1}(1-x)^{\beta-1} & \text{, se } 0 < x < 1 \\ 0 & \text{, caso contrário.} \end{cases}$$

Os parâmetros da distribuição Beta são α e β, e o espaço paramétrico é definido por: $\alpha > 0$ e $\beta > 0$. A notação usada para indicar que uma variável aleatória X segue a distribuição Beta é $X \sim beta(\alpha, \beta)$.

Veja que se $X \sim beta(\alpha, \beta)$ os valores que essa variável aleatória pode assumir são: $Im(X) = (0, 1)$. Além disso, se $X \sim beta(1, 1)$ então $X \sim U(0, 1)$.

Note que f_X apresentada na Definição 5.16 é de fato função densidade, uma vez que $f_X(x) \geq 0\ \forall\ x \in \mathbb{R}$ e

$$\begin{aligned} \int_{-\infty}^{\infty} f_X(x)dx &= \int_0^1 \frac{1}{Beta(\alpha, \beta)} x^{\alpha-1}(1-x)^{\beta-1}dx \\ &= \frac{1}{Beta(\alpha, \beta)} \int_0^1 x^{\alpha-1}(1-x)^{\beta-1}dx \\ &= \frac{1}{Beta(\alpha, \beta)} Beta(\alpha, \beta) = 1. \end{aligned}$$

Usando a relação entre as funções Γ e $Beta$ podemos reescrever a função densidade da distribuição Beta por:

$$f_X(x) = \begin{cases} \dfrac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1} & \text{, se } 0 < x < 1 \\ 0 & \text{, caso contrário.} \end{cases}$$

Esperança e Variância

Para os cálculos a seguir vamos usar muito a relação entre as funções Beta e Γ, além das propriedades da função Γ, já vistas na Propriedade 5.11. Vamos começar calculando $\mathrm{E}(X)$.

$$
\begin{aligned}
\mathrm{E}(X) &= \int_{-\infty}^{\infty} x f_X(x) dx = \int_0^1 x \frac{1}{Beta(\alpha,\beta)} x^{\alpha-1}(1-x)^{\beta-1} dx = \frac{1}{Beta(\alpha,\beta)} \int_0^1 x^{\alpha}(1-x)^{\beta-1} dx \\
&= \frac{1}{Beta(\alpha,\beta)} Beta(\alpha+1,\beta) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\cancel{\Gamma(\beta)}} \frac{\Gamma(\alpha+1)\cancel{\Gamma(\beta)}}{\Gamma(\alpha+\beta+1)} = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)} \frac{\Gamma(\alpha+1)}{\Gamma(\alpha+\beta+1)} \\
&= \frac{\cancel{\Gamma(\alpha+\beta)}}{\cancel{\Gamma(\alpha)}} \frac{\alpha\cancel{\Gamma(\alpha)}}{(\alpha+\beta)\cancel{\Gamma(\alpha+\beta)}} = \frac{\alpha}{(\alpha+\beta)}.
\end{aligned}
$$

Portanto,

$$
X \sim beta(\alpha,\beta) \qquad \Rightarrow \qquad \mathrm{E}(X) = \frac{\alpha}{\alpha+\beta}. \tag{5.12}
$$

O cálculo para $\mathrm{E}(X^2)$ é bem parecido com o que acabamos de fazer para $\mathrm{E}(X)$.

$$
\begin{aligned}
\mathrm{E}(X^2) &= \int_{-\infty}^{\infty} x^2 f_X(x) dx = \int_0^1 x^2 \frac{1}{Beta(\alpha,\beta)} x^{\alpha-1}(1-x)^{\beta-1} dx \\
&= \frac{1}{Beta(\alpha,\beta)} \int_0^1 x^{\alpha+1}(1-x)^{\beta-1} dx = \frac{1}{Beta(\alpha,\beta)} Beta(\alpha+2,\beta) \\
&= \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\cancel{\Gamma(\beta)}} \frac{\Gamma(\alpha+2)\cancel{\Gamma(\beta)}}{\Gamma(\alpha+\beta+2)} = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)} \frac{\Gamma(\alpha+2)}{\Gamma(\alpha+\beta+2)} \\
&= \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)} \frac{(\alpha+1)\alpha\Gamma(\alpha)}{(\alpha+\beta+1)(\alpha+\beta)\Gamma(\alpha+\beta)} = \frac{\cancel{\Gamma(\alpha+\beta)}}{\cancel{\Gamma(\alpha)}} \frac{(\alpha+1)\alpha\cancel{\Gamma(\alpha)}}{(\alpha+\beta+1)(\alpha+\beta)\cancel{\Gamma(\alpha+\beta)}} \\
&= \frac{(\alpha+1)\alpha}{(\alpha+\beta+1)(\alpha+\beta)}
\end{aligned}
$$

Então,

$$
\begin{aligned}
\mathrm{Var}(X) &= \mathrm{E}(X^2) - \mathrm{E}(X)^2 = \frac{(\alpha+1)\alpha}{(\alpha+\beta+1)(\alpha+\beta)} - \frac{\alpha^2}{(\alpha+\beta)^2} \\
&= \frac{(\alpha+1)\alpha(\alpha+\beta) - \alpha^2(\alpha+\beta+1)}{(\alpha+\beta+1)(\alpha+\beta)^2} = \frac{\alpha\left(\alpha^2+\alpha\beta+\alpha+\beta-\alpha^2-\alpha\beta-\alpha\right)}{(\alpha+\beta+1)(\alpha+\beta)^2} \\
&= \frac{\alpha\beta}{(\alpha+\beta+1)(\alpha+\beta)^2}.
\end{aligned}
$$

Portanto,

$$
X \sim beta(\alpha,\beta) \qquad \Rightarrow \qquad \mathrm{Var}(X) = \frac{\alpha\beta}{(\alpha+\beta+1)(\alpha+\beta)^2}. \tag{5.13}
$$

Exemplo 5.17
A percentagem de impurezas por lote, em um determinado produto químico, é uma variável aleatória com distribuição beta de parâmetros $\alpha = 3$ e $\beta = 2$. Um lote com mais de 40% de impurezas não pode ser vendido.

(a) Qual é a probabilidade de que um lote selecionado ao acaso não poder ser vendidos por causa do excesso de impurezas?

(b) Quantos lotes em média são selecionados ao acaso até que se encontre um que não pode ser vendidos por causa do excesso de impurezas?

(c) Qual a porcentagem média de impurezas nos lotes desse produto químico?

Solução:

(a) Seja Y =percentagem de impurezas em um lote. Queremos encontrar $P(Y > 0,4)$.

Pelo enunciado temos que $Y \sim beta(3,2)$. Então, a função densidade de Y, para $0 < y < 1$, é definida por:

$$\begin{aligned} f_Y(y) &= \frac{1}{Beta(3,2)} y^{3-1}(1-y)^{2-1} = \frac{\Gamma(3+2)}{\Gamma(3)\Gamma(2)} y^2(1-y) \\ &= \frac{4!}{2! \times 1!} y^2(1-y) = 12y^2(1-y), \quad 0 < y < 1. \end{aligned}$$

Então,

$$\begin{aligned} P(Y > 0,4) &= \int_{0,4}^{\infty} f_Y(y)dy = \int_{0,4}^{1} 12y^2(1-y)dy = 12\int_{0,4}^{1} y^2 - y^3 dy \\ &= 12 \left. \frac{y^3}{3} - \frac{y^4}{4} \right|_{0,4}^{1} = 12\left(\frac{1}{3} - \frac{1}{4} - \frac{0,4^3}{3} + \frac{0,4^4}{4}\right) = 0,8208. \end{aligned}$$

(b) Seja W = números de lotes selecionados ao acaso até que se encontre um que não pode ser vendido por excesso de impureza. Queremos $E(W)$. Veja que $W \sim geo(p = 0,8208)$. Então, $E(W) = \frac{1}{p} = 1,218$.

(c) Nesse item queremos $E(Y)$. Como $Y \sim beta(3,2)$ sabemos que $E(Y) = \frac{3}{5}0,6$. Ou seja, a porcentagem média de impurezas nos lotes desse produto químico é de 60%.

♦♦

5.5 Distribuição de Weibull

Assim como no caso das distribuições anteriores, a definição de uma variável aleatória com distribuição de Weibull será feita a partir da definição da sua função densidade.

Função Densidade

Definição 5.18 Distribuição de Weibull
Uma variável aleatória X tem distribuição de Weibull com parâmetros $\alpha > 0$ e $\beta > 0$ se sua função densidade de probabilidade é dada por

$$f(x) = \frac{\alpha}{\beta^\alpha} x^{\alpha-1} e^{-\left(\frac{x}{\beta}\right)^\alpha} \quad , \; x > 0 \tag{5.14}$$

Logo, de acordo com a Definição 5.18, a distribuição de Weibull tem dois parâmetros, α e β, e o espaço paramétrico é $\alpha > 0$ e $\beta > 0$. Além disso, Se X tem distribuição de Weibull com parâmetros α e β ($X \sim Weibull(\alpha, \beta)$), os valores que X pode assumir estão no intervalo $(0, \infty)$, isto é, $Im(X) = (0, \infty)$.

Alguns autores (ver Rohatgi (2008), por exemplo) usam um novo parâmetro η em vez de β^α ou λ em vez de $1/\beta^\alpha$ (ver Magalhães (2011), por exemplo).

Para mostrar que (5.14) define uma densidade, vamos mostrar que sua integral é 1. Para tal, note que podemos reescrever 5.14 como

$$f(x) = \frac{\alpha}{\beta}\left(\frac{x}{\beta}\right)^{\alpha-1} e^{-\left(\frac{x}{\beta}\right)^\alpha} \qquad , \ x > 0 \tag{5.15}$$

Fazendo a mudança de variável:

$$u = \left(\frac{x}{\beta}\right)^\alpha \Longrightarrow du = \frac{\alpha}{\beta}\left(\frac{x}{\beta}\right)^{\alpha-1} \qquad \text{, então} \quad x = 0 \Longrightarrow u = 0 \text{ e } x = \infty \Longrightarrow u = \infty.$$

Assim obtemos

$$\int_0^\infty \frac{\alpha}{\beta}\left(\frac{x}{\beta}\right)^{\alpha-1} e^{-\left(\frac{x}{\beta}\right)^\alpha} dx = \int_0^\infty e^{-u} du = 1.$$

Função de Distribuição

No caso da distribuição de Weibull é possível encontrar a sua função de distribuição também parametrizada por α e β.

Por definição,

$$F(x) = \int_0^x \frac{\alpha}{\beta}\left(\frac{t}{\beta}\right)^{\alpha-1} e^{-\left(\frac{t}{\beta}\right)^\alpha} dt$$

Fazendo a mudança de variável

$$u = \left(\frac{t}{\beta}\right)^\alpha \Longrightarrow du = \frac{\alpha}{\beta}\left(\frac{t}{\beta}\right)^{\alpha-1} \qquad \text{, então} \quad t = 0 \Longrightarrow u = 0 \text{ e } t = x \Longrightarrow u = \left(\frac{x}{\beta}\right)^\alpha.$$

Logo,

$$F(x) = \int_0^x \frac{\alpha}{\beta}\left(\frac{t}{\beta}\right)^{\alpha-1} e^{-\left(\frac{t}{\beta}\right)^\alpha} dt = \int_0^{\left(\frac{x}{\beta}\right)^\alpha} e^{-u} du = -e^{-u}\Big|_0^{\left(\frac{x}{\beta}\right)^\alpha} = 1 - e^{-\left(\frac{x}{\beta}\right)^\alpha}.$$

Portanto,

$$X \sim Weibull(\alpha, \beta) \qquad \Rightarrow \qquad F_X(x) = \begin{cases} 0 & , \ x \leq 0 \\ 1 - e^{-(x/\beta)^\alpha} & , \ x > 0 \end{cases} \tag{5.16}$$

Esperança e Variância

Para encontrar a esperança e a variância de $X \sim Weibull(\alpha, \beta)$ vamos calcular $\mathrm{E}(X^r)$ para qualquer r inteiro positivo. Assim, para achar $\mathrm{E}(X)$ basta fazer $r = 1$ e para encontrar $\mathrm{E}(X^2)$ basta fazer $r = 2$.

$$\mathrm{E}(X^r) = \int_0^\infty \frac{\alpha}{\beta}\left(\frac{x}{\beta}\right)^{\alpha-1} e^{-\left(\frac{x}{\beta}\right)^\alpha} x^r dx$$

Fazendo $u = \frac{x}{\beta}$, resulta que $x = \beta u$ e $dx = \beta du$; logo

$$\begin{aligned}
\mathrm{E}(X^r) &= \int_0^\infty \frac{\alpha}{\beta}\left(\frac{x}{\beta}\right)^{\alpha-1} e^{-\left(\frac{x}{\beta}\right)^\alpha} x^r dx = \int_0^\infty \frac{\alpha}{\beta} u^{\alpha-1} e^{-u^\alpha} \beta^r u^r \beta du \\
&= \int_0^\infty \alpha u^{\alpha-1} e^{-u^\alpha} \beta^r u^r du
\end{aligned}$$

Fazendo $u^\alpha = t$ resulta que $u = t^{1/\alpha}$ e $\alpha u^{\alpha-1} du = dt$; logo,

$$\begin{aligned}
\mathrm{E}(X^r) &= \int_0^\infty e^{-t} \beta^r \left(t^{1/\alpha}\right)^r dt = \beta^r \int_0^\infty t^{r/\alpha} e^{-t} dt \\
&= \beta^r \int_0^\infty t^{r/\alpha+1-1} e^{-t} dt = \beta^r \int_0^\infty t^{\frac{r+\alpha}{\alpha}-1} e^{-t} dt = \beta^r \Gamma\left(\frac{r+\alpha}{\alpha}\right)
\end{aligned}$$

Fazendo $r = 1$, obtemos $\mathrm{E}(X)$:

$$X \sim Weibull(\alpha, \beta) \qquad \Rightarrow \qquad \mathrm{E}(X) = \beta\Gamma\left(\frac{\alpha+1}{\alpha}\right) \tag{5.17}$$

Fazendo $r = 2$ obtemos $\mathrm{E}(X^2)$ e, consequentemente $\mathrm{Var}(X)$.

$$\mathrm{E}(X^2) = \beta^2\Gamma\left(\frac{\alpha+2}{\alpha}\right)$$

e, portanto,

$$X \sim Weibull(\alpha, \beta) \qquad \Rightarrow \qquad \mathrm{Var}(X) = \beta^2\left[\Gamma\left(\frac{\alpha+2}{\alpha}\right) - \Gamma^2\left(\frac{\alpha+1}{\alpha}\right)\right] \tag{5.18}$$

5.6 Distribuição de Pareto

Novamente, a definição de uma variável aleatória com distribuição de Pareto será feita a partir da definição da sua função densidade.

Função Densidade

Definição 5.19 Densidade de Pareto
Uma variável aleatória X tem distribuição de Pareto com parâmetros $\alpha > 0$ e $b > 0$ se sua função densidade de probabilidade é dada por

$$f(x) = \begin{cases} \dfrac{\alpha}{b}\left(\dfrac{b}{x}\right)^{\alpha+1} & , \text{ se } x \geq b \\ 0 & , \text{ se } x < b \end{cases}$$

Logo, de acordo com a Definição 5.19, a distribuição de Pareto tem dois parâmetros, α e b, e o espaço paramétrico é $\alpha > 0$ e $b > 0$. Além disso, Se X tem distribuição de Pareto com parâmetros α e b ($X \sim Pareto(\alpha, b)$), os valores que X pode assumir estão no intervalo (b, ∞), isto é, $Im(X) = (b, \infty)$.

Para mostrar que $f(x)$ realmente define uma função densidade de probabilidade resta provar que a integral é 1, uma vez que $f(x) \geq 0$.

$$\int_b^\infty \frac{\alpha}{b}\left(\frac{b}{x}\right)^{\alpha+1} dx = \alpha b^\alpha \int_b^\infty x^{-\alpha-1} dx = \alpha b^\alpha \frac{x^{-\alpha}}{-\alpha}\Big|_b^\infty$$

Essa integral converge apenas se $-\alpha < 0$ ou equivalentemente, $\alpha > 0$, pois nesse caso $\lim_{x\to\infty} x^{-\alpha} = \lim_{x\to\infty} \frac{1}{x^\alpha} = 0$. Satisfeita esta condição, temos que

$$\alpha b^\alpha \frac{x^{-\alpha}}{-\alpha}\Big|_b^\infty = 0 - \alpha b^\alpha \frac{b^{-\alpha}}{-\alpha} = 1$$

Na **Figura 5.10(a)** ilustra-se a função densidade de Pareto para $\alpha = 3$ e $b = 2$.

Função de Distribuição

No caso da distribuição de Pareto também é possível encontrar a sua função de distribuição parametrizada por α e β.

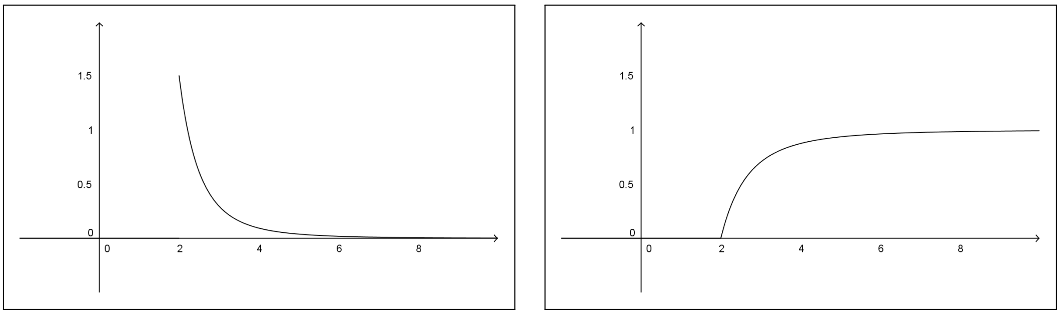
Por definição, $F(x) = \mathrm{P}(X \leq x) = 0$ se $x < b$. Para $x \geq b$,

$$\begin{aligned} F(x) &= \mathrm{P}(X \leq x) = \int_b^x \frac{\alpha}{b}\left(\frac{b}{t}\right)^{\alpha+1} dt = \alpha b^\alpha \int_b^x t^{-\alpha-1} dx = \alpha b^\alpha \frac{t^{-\alpha}}{-\alpha}\Big|_b^x \\ &= -b^\alpha\left(x^{-\alpha} - b^{-\alpha}\right) = 1 - \left(\frac{b}{x}\right)^\alpha. \end{aligned}$$

Portanto,

$$X \sim Pareto(\alpha, b) \qquad \Rightarrow \qquad F_X(x) = \begin{cases} 0 & \text{, se } x < b \\ 1 - \left(\frac{b}{x}\right)^\alpha & \text{, se } x \geq b \end{cases} \tag{5.19}$$

Na Figura 5.10(b) ilustra-se a função de distribuição de Pareto com parâetros $\alpha = 3$ e $b = 2$.



(a) Função Densidade (b) Função de Distribuição

Figura 5.10 – Distribuição de Pareto com parâmetros $\alpha = 3$ e $b = 2$.

Esperança e Variância

Se $X \sim Pareto(\alpha, b)$ então

$$\mathrm{E}(X) = \int_b^\infty x\frac{\alpha}{b}\left(\frac{b}{x}\right)^{\alpha+1} dx = \alpha b^\alpha \int_b^\infty x^{-\alpha} dx = \alpha b^\alpha \frac{x^{-\alpha+1}}{-\alpha+1}\Big|_b^\infty.$$

Para que essa integral convirja, temos que ter $-\alpha + 1 < 0$, ou seja, $\alpha > 1$. Satisfeita esta condição,

$$\mathrm{E}(X) = \alpha b^{\alpha} \left(0 - \frac{b^{-\alpha+1}}{-\alpha+1} \right) = \frac{-\alpha b^{\alpha-\alpha+1}}{1-\alpha} = \frac{\alpha b}{\alpha - 1}.$$

Portanto,

$$X \sim Pareto(\alpha, b) \qquad \Rightarrow \qquad \mathrm{E}(X) = \begin{cases} \dfrac{\alpha b}{\alpha - 1} & , \text{se } \alpha > 1 \\ \text{não existe} & , \text{se } \alpha \leq 1 \end{cases} \tag{5.20}$$

Vejamos agora as contas para $\mathrm{E}(X^2)$.

$$\mathrm{E}(X^2) = \int_b^\infty x^2 \frac{\alpha}{b} \left(\frac{b}{x} \right)^{\alpha+1} dx = \alpha b^{\alpha} \int_b^\infty x^{-\alpha+1} dx = \alpha b^{\alpha} \left. \frac{x^{-\alpha+2}}{-\alpha+2} \right|_b^\infty .$$

Para que essa integral convirja, temos que ter $-\alpha + 2 < 0$, ou $\alpha > 2$. Satisfeita esta condição,

$$\mathrm{E}(X) = \alpha b^{\alpha} \left(0 - \frac{b^{-\alpha+2}}{-\alpha+2} \right) = \frac{-\alpha b^{\alpha-\alpha+2}}{2-\alpha} = \frac{\alpha b^2}{\alpha - 2}.$$

Logo, se $\alpha > 2$,

$$\begin{aligned} \mathrm{Var}(X) &= \frac{\alpha b^2}{\alpha - 2} - \left(\frac{\alpha b}{\alpha - 1} \right)^2 = \frac{\alpha b^2 (\alpha - 1)^2 - \alpha^2 b^2 (\alpha - 2)}{(\alpha - 1)^2 (\alpha - 2)} \\ &= \frac{\alpha b^2 \left[\alpha^2 - 2\alpha + 1 - \alpha(\alpha - 2) \right]}{(\alpha - 1)^2 (\alpha - 2)} = \frac{\alpha b^2 \left[\alpha^2 - 2\alpha + 1 - \alpha^2 + 2\alpha \right]}{(\alpha - 1)^2 (\alpha - 2)} \\ &= \frac{\alpha b^2}{(\alpha - 1)^2 (\alpha - 2)} \end{aligned}$$

Portanto,

$$X \sim Pareto(\alpha, b) \qquad \Rightarrow \qquad \mathrm{Var}(X) = \begin{cases} \dfrac{\alpha b^2}{(\alpha - 1)^2 (\alpha - 2)} & , \text{se } \alpha > 2 \\ \text{não existe} & , \text{se } \alpha \leq 1 \end{cases} \tag{5.21}$$

Capítulo 6

Funções de Variáveis Aleatórias Contínuas

Dada uma variável aleatória contínua X com função densidade f_X, muitas vezes estamos interessados em conhecer a densidade de uma outra variável aleatória $Y = g(X)$ definida como uma função de X. Este é o problema que vamos resolver nesse capítulo.

Esse problema já foi abordado na Seção 2.2 para o caso de X ser variáveis aleatórias discretas. Na ocasião vimos que se X é variável aleatória discreta então $Y = g(X)$ também será variável aleatória discreta e o problema era resolvido definindo a função de probabilidade de Y a partir da função de probabilidade de X.

No caso de X ser variáveis aleatória contínua a solução será um pouco diferente. Primeiro, não necessariamente $Y = g(X)$ será variável aleatória contínua, Y por ser contínua, discreta e até mesmo uma variável aleatória que não seja nem discreta nem contínua. Então buscar a função de probabilidade de Y não é uma solução adequada. Neste caso iremos buscar a função de distribuição de Y, $F_Y(y) = \mathrm{P}(Y \leq)$, que está bem definida qualquer que seja o tipo da variável Y, e o primeiro método considerado para tratar este tipo de problema será o denominado *Método da Função de Distribuição*.

6.1 Método da Função de Distribuição

Esse método consiste em encontrar a função de distribuição da variável transformada $Y = g(X)$ a partir da função de distribuição de X. Se Y for variável aleatória continua podemos então obter a sua função densidade a partir da derivada da função de distribuição, já encontrada.

Para facilitar o desenvolvimento do método veja o passo-a-passo a seguir.

Seja X variável aleatória contínua, F_X a sua função de distribuição, f_X a sua função densidade e $Y = g(X)$ a transformação que define Y como função de X. Então, para encontrar F_Y a partir do método da Função de Distribuição siga os seguintes passos:

Passo 1) Faça um esboço do gráfico de g. O eixo horizontal será X e o eixo vertical Y.

Pbsso 2) Encontre $Im(X)$ e marque esse conjunto no eixo X do gráfico de g.

Pcsso 3) A partir do gráfico de g encontre $Im(Y)$. Marque esse conjunto no eixo Y do gráfico de g.

Pdsso 4) Escreva $F_Y(y) = \mathrm{P}(Y \leq y) = \mathrm{P}(g(X) \leq y)$ e, a partir de manipulações algébricas, encontre a relação entre F_Y e F_X. Use o gráfico de g já feito para te ajudar. Pode ser que você tenha que considerar diferentes regiões para $y \in \mathbb{R}$.

Pesso 5) Se você conhece a expressão de F_X, fim. Substitua e encontre a expressão para F_Y. Se quiser encontrar f_Y, derive F_Y. Se você conhece a expressão de F_X continue.

Pfsso 6) Uma vez encontrada a relação entre F_Y e F_X, derive os dois lados da equação para encontrar uma relação entre f_Y e f_X.

Pgsso 7) Substitua f_X conhecida e encontre a expressão para f_Y.

Dois pontos importantes que devem ser verificados depois de encontrada F_Y ou f_Y, a fim de detectar erros no desenvolvimento.

- A $Im(Y)$ encontrada no Passo Pcsso 3) está de acordo com a $Im(Y)$ encontrada a partir de F_Y ou f_Y?
- A função F_Y ou f_Y segue as propriedades de função de distribuição ou função densidade, respectivamente?

6.1.1 Caso em que g é inversível

Quando a função g que relaciona as variáveis X e Y é inversível, não teremos muito problema em encontrar a função de distribuição de Y ou a sua função densidade. Para esses casos as contas seguem sem preocupação.

Exemplo 6.1
Seja $X \sim exp(\lambda)$ e $Y = X + a$, com $a \in \mathbb{R}$ e $\lambda > 0$. Encontre f_Y e esboce seu gráfico.

Solução:
Veja que $Im(X) = (0, \infty)$. Então $Im(Y) = (a, \infty)$.

Veja que,

$$F_Y(y) = \mathrm{P}(Y \leq y) = \mathrm{P}(X + a \leq y) = \mathrm{P}(X \leq y - a) = F_X(y - a).$$

Assim encontramos uma relação entre F_Y e F_X:

$$F_Y(y) = F_X(y - a).$$

Então, uma vez que

$$F_X(x) = \begin{cases} 1 - e^{-\lambda x} & , \text{se} x > 0 \\ 0 & , \text{se } x \geq 0, \end{cases}$$

podemos escrever,

$$\begin{aligned} F_Y(y) = F_X(y-a) &= \begin{cases} 1 - e^{-\lambda(y-a)} & , \text{se} y - a > 0 \\ 0 & , \text{se } y - a \geq 0, \end{cases} \\ &= \begin{cases} 1 - e^{-\lambda(y-a)} & , \text{se} y > a \\ 0 & , \text{se } y \geq a. \end{cases} \end{aligned}$$

Se quisermos encontrar f_Y podemos derivar F_Y.

$$f_Y(y) = \frac{d}{dy}F_Y(y) = \begin{cases} \lambda e^{-\lambda(y-a)} & \text{, se} y > a \\ 0 & \text{, se } y \geq a. \end{cases}$$

♦♦

Exemplo 6.2
Se $X \sim \mathcal{U}(-1, 1)$, calcule a função densidade de $Z = \eta(X) = e^X$.

Solução:
As função densidade e de distribuição de $X \sim \mathcal{U}(-1, 1)$ são

$$f_X(x) = \begin{cases} \frac{1}{2} & \text{, se } -1 < x < 1 \\ 0 & \text{, caso contrário} \end{cases} \qquad \text{e} \qquad F_X(x) = \begin{cases} 0 & \text{, se } x \leq -1 \\ \frac{1}{2}(x+1) & \text{, se } -1 < x < 1 \\ 1 & \text{, se } x \geq 1. \end{cases}$$

Veja como ficam as imagens de $Z = \eta(X)$.

$$-1 < x < 1 \Rightarrow e^{-1} < z = \eta(x) < e^x.$$

Vamos às contas.

$$\begin{aligned} F_Z(z) = \text{P}(Z \leq z) &= \text{P}(e^X \leq z) = \text{P}(X \leq \ln z) = F_X(\ln z) \\ &= \begin{cases} 0 & z \leq e^{-1} \\ \dfrac{1}{2}(\ln z + 1) & e^{-1} < z < e^{+1} \\ 1 & z \geq e^1 \end{cases} \end{aligned}$$

Portanto

$$f_Z(z) = F_Z'(z) = \begin{cases} \dfrac{1}{2z} & e^{-1} < z < e^{+1} \\ 0 & \text{caso contrário.} \end{cases}$$

♦♦

Esse método também nos permite demonstrar alguns resultados importantes, como o apresentado na Proposição 6.3 a seguir.

Proposição 6.3
Seja $X \sim gama(\alpha, \lambda)$, $Y = cx$ e $c \in \mathbb{R}$. Então, $Y \sim gama(\alpha, \lambda/c)$.

Demonstração:
A demonstração será feita simplesmente aplicando o Método da Função de Distribuição para obter f_Y. Veja que podemos escrever F_Y em termos de F_X

$$F_Y(y) = \text{P}(Y \leq y) = \text{P}(cX \leq y) = \text{P}\left(X \leq \frac{y}{c}\right) = F_X\left(\frac{y}{c}\right).$$

Como não conhecemos F_X, e sim f_X, vamos derivar para encontrar uma relação entre f_Y e f_X.

$$f_Y(y) = \frac{d}{dy}F_Y(y) = \frac{d}{dy}F_X\left(\frac{y}{c}\right) = \frac{1}{c}f_X\left(\frac{y}{c}\right).$$

Sabemos que

$$f_X(x) = \frac{\lambda^\alpha}{\Gamma(\alpha)}x^{\alpha-1}e^{-\lambda x}, \text{ para } x > 0,$$

então,

$$\begin{aligned} f_X\left(\frac{y}{c}\right) &= \frac{\lambda^\alpha}{\Gamma(\alpha)}\left(\frac{y}{c}\right)^{\alpha-1} e^{-\lambda\frac{y}{c}}, \ \frac{y}{c} > 0. \\ &= \left(\frac{1}{c}\right)^{\alpha-1}\frac{\lambda^\alpha}{\Gamma(\alpha)} y^{\alpha-1} e^{-\frac{\lambda}{c}y}, \ y > 0. \end{aligned}$$

Portanto,

$$\begin{aligned} f_Y(y) = \frac{1}{c} f_X\left(\frac{y}{c}\right) &= \frac{1}{c}\left(\frac{1}{c}\right)^{\alpha-1}\frac{\lambda^\alpha}{\Gamma(\alpha)} y^{\alpha-1} e^{-\frac{\lambda}{c}y}, \ y > 0 \\ &= \frac{(\lambda/c)^\alpha}{\Gamma(\alpha)} y^{\alpha-1} e^{-\frac{\lambda}{c}y}, \ y > 0. \end{aligned}$$

Com isso concluímos que $Y \sim gama(\alpha, \lambda/c)$.

□

Corolário 6.4 *Se $X \sim exp(\lambda)$ e $c \in \mathbb{R}$, então $Y = cX \sim exp(\lambda/c)$.*

Exemplo 6.5
Seja X variável aleatória cuja função densidade é definida por:

$$f_X(x) = \begin{cases} (1-x)/2 & , se\ 0 < x \leq 1 \\ 3(x-1)/2 & , se\ 1 < x < 2 \end{cases}$$

Encontre f_Y para $Y = 1 - X$. Em seguida, esboce os gráficos de f_X e f_Y.

Solução:
Sempre é bom primeiro encontrar $Im(X)$ e $Im(Y)$. Veja que $Im(X) = (0, 2)$, logo $Im(Y) = (-1, 1)$. A partir do esboço do gráfico de $g(x) = 1 - x$ fica fácil ver que o conjunto $(0, 2)$ é levando no conjunto $(-1, 1)$ pela função g.

Vamos agora buscar a relação entre F_Y e F_X:

$$F_Y(y) = \text{P}(Y \leq y) = \text{P}(1 - X \leq y) = \text{P}(X \geq 1 - y) = 1 - F_X(1 - y).$$

Derivando temos a relação entre f_Y e f_X:

$$f_Y(y) = \frac{d}{dy} F_Y(y) = \frac{d}{dy}\left(1 - F_X(1-y)\right) = f_X(1-y).$$

Então,

$$\begin{aligned} f_Y(y) = f_X(1-y) &= \begin{cases} (1-(1-y))/2 & , \text{se } 0 < (1-y) \leq 1 \\ 3((1-y)-1)/2 & , \text{se } 1 < (1-y) < 2. \end{cases} \\ &= \begin{cases} y/2 & , \text{se } 0 \leq y < 1 \\ -3y/2 & , \text{se } -1 < y < 0. \end{cases} \\ &= \begin{cases} -3y/2 & , \text{se } -1 < y < 0. \\ y/2 & , \text{se } 0 \leq y < 1 \end{cases} \end{aligned}$$

♦♦

6.1.2 Caso em que g não é inversível

Quando a função g, que relaciona as variáveis X e Y, não é inversível, precisamos redobrar o cuidado ao fazer as contas. Nesse caso, esboçar e conhecer bem a função g pode ajudar bastante. Vejamos alguns exemplos.

Exemplo 6.6
Se $X \sim \mathcal{U}(-1,1)$, calcule a função densidade das seguintes variáveis aleatórias

(a) $Y = g(X) = X^2$

(b) $W = h(X) = |X|$

Solução:
Como já vimos, as função densidade e de distribuição de $X \sim \mathcal{U}(-1,1)$ são:

$$f_X(x) = \begin{cases} \frac{1}{2} & \text{, se } -1 < x < 1 \\ 0 & \text{, caso contrário} \end{cases} \qquad \text{e} \qquad F_X(x) = \begin{cases} 0 & \text{, se } x \leq -1 \\ \frac{1}{2}(x+1) & \text{, se } -1 < x < 1 \\ 1 & \text{, se } x \geq 1. \end{cases}$$

Veja como ficam as imagens de Y e W.

$$-1 < x < 1 \Rightarrow \begin{cases} 0 \leq y = g(x) < 1 \\ 0 \leq w = h(x) < 1 \end{cases}$$

(a) Para calcular a densidade de $Y = g(X) = X^2$ devemos notar que

$$\begin{aligned} F_Y(y) &= \mathrm{P}(Y \leq y) = \mathrm{P}(X^2 \leq y) \\ &= \begin{cases} 0 & \text{se } y \leq 0 \\ \mathrm{P}\left(-\sqrt{y} \leq X \leq \sqrt{y}\right) & \text{se } 0 < y < 1 \\ 1 & \text{se } y \geq 1 \end{cases} \\ &= \begin{cases} 0 & \text{se } y \leq 0 \\ F_X\left(\sqrt{y}\right) - F_X\left(-\sqrt{y}\right) & \text{se } 0 < y < 1 \\ 1 & \text{se } y \geq 1 \end{cases} \end{aligned}$$

e, portanto

$$\begin{aligned} f_Y(y) &= \begin{cases} \dfrac{d}{dy}\left[F_X\left(\sqrt{y}\right)\right] - \frac{d}{dy}\left[F_X\left(-\sqrt{y}\right)\right] & \text{se } 0 < y < 1 \\ 0 & \text{caso contrário} \end{cases} \\ &= \begin{cases} F_X'\left(\sqrt{y}\right)\frac{1}{2\sqrt{y}} - F_X'\left(-\sqrt{y}\right)\left(-\frac{1}{2\sqrt{y}}\right) & \text{se } 0 < y < 1 \\ 0 & \text{caso contrário} \end{cases} \\ &= \begin{cases} f_X\left(\sqrt{y}\right)\frac{1}{2\sqrt{y}} + f_X\left(-\sqrt{y}\right)\frac{1}{2\sqrt{y}} & \text{se } 0 < y < 1 \\ 0 & \text{caso contrário} \end{cases} \end{aligned}$$

Como $0 \leq \sqrt{y} < 1$ e $-1 < -\sqrt{y} \leq 0$, resulta que $f_X\left(\sqrt{y}\right) = f_X\left(-\sqrt{y}\right) = \frac{1}{2}$. Logo

$$f_Y(y) = \begin{cases} \dfrac{1}{2\sqrt{y}} & \text{se } 0 < y < 1 \\ 0 & \text{caso contrário} \end{cases}$$

(b) De modo análogo, para $0 \leq w < 1$

$$\begin{aligned}
F_W(w) &= \mathrm{P}(W \leq w) = \mathrm{P}(|X| \leq w) \\
&= \begin{cases} 0 & \text{se } w \leq 0 \\ \mathrm{P}(-w \leq X \leq w) & \text{se } 0 < w < 1 \\ 1 & \text{se } w \geq 1 \end{cases} \\
&= \begin{cases} 0 & \text{se } w \leq 0 \\ F_X(w) - F_X(-w) & \text{se } 0 < w < 1 \\ 1 & \text{se } w \geq 1 \end{cases}
\end{aligned}$$

e, portanto

$$\begin{aligned}
f_W(w) = F'_W(w) &= \begin{cases} F'_X(w) - F'_X(-w)(-1) & \text{se } 0 < w < 1 \\ 0 & \text{caso contrário} \end{cases} \\
&= \begin{cases} f_X(w) + f_X(-w) & \text{se } 0 < w < 1 \\ 0 & \text{caso contrário} \end{cases}
\end{aligned}$$

Como $0 \leq w < 1$ e $-1 < -w \leq 0$, resulta que $f_X(w) = f_X(-w) = \frac{1}{2}$. Logo

$$f_W(w) = \begin{cases} 1 & \text{se } 0 < w < 1 \\ 0 & \text{caso contrário} \end{cases}$$

Note que $W \sim \mathcal{U}(0, 1)$.

♦♦

Os dois exemplos a seguir mostram que, mesmo X sendo variável aleatória contínua, $Y = g(X)$ não necessariamente é contínua.

Exemplo 6.7
Seja $X \sim U(0, 1)$ e

$$Y = \begin{cases} 1 & , \text{se } X \leq p \\ 0 & , \text{se } X > p \end{cases}$$

para algum $0 < p < 1$. Encontre a função distribuição de Y e classifique esta variável aleatória.

Solução:
Veja que $Im(Y) = \{0, 1\}$. Logo, já de início, sabemos que Y será variável aleatória discreta, pois a sua imagem é um conjunto finito.

Nesse caso podemos encontrar primeiro $\mathrm{P}(Y = 0)$ e $\mathrm{P}(Y = 1)$, depois encontramos a função de distribuição F_Y.

$$\mathrm{P}(Y = 0) = \mathrm{P}(X > p) = 1 - F_X(p) = 1 - p \quad \text{e} \quad \mathrm{P}(Y = 1) = \mathrm{P}(X \leq p) = F_X(p) = p.$$

Veja então que $Y \sim Bernoulli(p)$ e, portanto, a sua função de distribuição é definida por:

$$F_Y(y) = \begin{cases} 0 & , \text{se } y < 0 \\ 1 - p & , \text{se } 0 \leq y < 1 \\ 1 & , \text{se } y \geq 1. \end{cases}$$

♦♦

Exemplo 6.8 James (2004) – ex.9a pág.89
Seja X uma v.a. contínua com função densidade

$$f_X(x) = \begin{cases} \frac{1}{(1+x)^2}, & x > 0 \\ 0 & , \text{caso contrário.} \end{cases}$$

Defina $Y = \max\{X, c\}$, em que $c > 0$, $c \in \mathbb{R}$. Ache a função de distribuição de Y e classifique esta variável aleatória como discreta, contínua ou nem discreta e nem contínua.

Solução:
Primeiro faça o gráfico da função $g(X) = \max\{X, c\}$, veja Figura 6.1 abaixo. A partir do gráfico

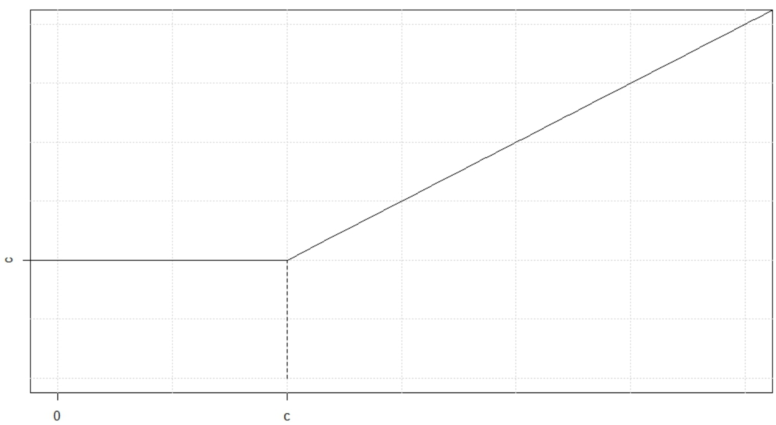


Figura 6.1 – Gráfico da função $g(X) = \max\{X, c\}$.

de g é possível concluir que se $Im(X) = (0, \infty)$, então $Im(Y) = [c, \infty)$.

Queremos encontrar F_Y. Veja que se $y < c$, $F_Y(y) = \mathrm{P}(Y \leq y) = 0$. Já se $y \geq c$, de acordo com o gráfico,

$$F_Y(y) = \mathrm{P}(Y \leq y) = \mathrm{P}(X \leq y) = F_X(y).$$

Então,

$$F_Y(y) = \begin{cases} 0 & , \text{se } y < c \\ F_X(y) & , \text{se } y \geq c. \end{cases}$$

Vamos encontrar a função F_X para então definir a função F_Y.

$$F_X(x) = \int_{-\infty}^{x} f_X(t)dt = \begin{cases} 0 & , \text{se } x < 0 \\ \int_0^x \frac{1}{(1+t)^2} dt & , \text{se } x \geq 0. \end{cases}$$

Veja que, para $x > 0$,

$$\int_0^x \frac{1}{(1+t)^2} dt = \int_1^{1+x} \frac{1}{u^2} du = -\frac{1}{u}\bigg|_1^{1+x} = -1 + \frac{1}{1+x} = \frac{x}{1+x}.$$

Então,

$$F_X(x) = \begin{cases} 0 & , \text{se } x < 0 \\ \frac{x}{1+x} & , \text{se } x \geq 0. \end{cases}$$

Assim podemos encontrar F_Y.

$$F_Y(y) = \begin{cases} 0 & \text{, se } y < c \\ \frac{y}{1+y} & \text{, se } y \geq c. \end{cases}$$

Podemos verificar que F_Y é uma função que satisfaz as propriedades da função de distribuição (Proposição 1.16). Mas F_Y não é uma função contínua, logo Y não é variável aleatória contínua. Também podemos ver que $Im(Y)$ não é um conjunto enumerável, logo Y não é variável aleatória discreta. Então, Y não é variável aleatória discreta nem contínua.

♦♦

Exemplo 6.9
Seja X variável aleatória tal que $f(x) = x^3/64$, $0 \leq x \leq 4$. Seja $Y = \min\{\sqrt{X}, 2 - \sqrt{X}\}$. Encontre a função densidade de Y.

Solução:
Primeiro faça o gráfico da função g e encontre, a partir da $Im(X)$ o conjunto $Im(Y)$. Veja na Figura 6.2 abaixo o gráfico da função g. Veja que como $Im(X) = (0, 4)$, então $Im(Y) = (0, 1)$.

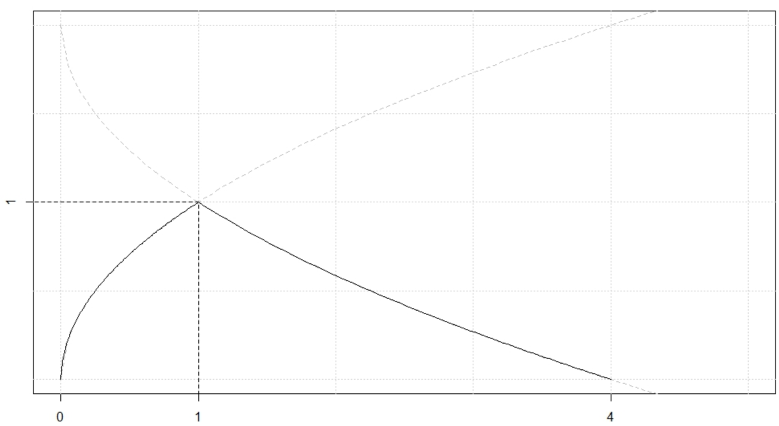


Figura 6.2 – Gráfico da função $g(X) = \min\{\sqrt{X}, 2 - \sqrt{X}\}$.

Queremos encontrar F_Y. Vejamos primeiro os casos triviais: se $y < 0 \Rightarrow F_Y(y) = 0$; e se $y > 1 \Rightarrow F_Y(y) = 1$.

Agora vamos analisar o caso em que $0 \leq y \leq 1$.

$$\begin{aligned} F_Y(y) &= P(Y \leq y) = P\left(X \leq y^2 \text{ ou } X \geq (2-y)^2\right) \\ &= P\left(X \leq y^2\right) + P\left(X \geq (2-y)^2\right) \\ &= F_X\left(y^2\right) + 1 - F_X\left((2-y)^2\right). \end{aligned}$$

Assim,

$$F_Y(y) = \begin{cases} 0 & \text{, se } y < 0 \\ F_X\left(y^2\right) + 1 - F_X\left((2-y)^2\right) & \text{, se } 0 \leq y \leq 1 \\ 1 & \text{, se } y > 1. \end{cases}$$

Para encontrar a função densidade podemos derivar em relação à y.

$$f_Y(y) = \begin{cases} f_X\left(y^2\right)2y + f_X\left((2-y)^2\right)2(2-y) & \text{, se } 0 \leq y \leq 1 \\ 0 & \text{, caso contrário.} \end{cases}$$

Substituindo pela expressão de f_X, chegamos em:

$$f_Y(y) = \begin{cases} \frac{1}{32}\left(y^7 + (2-y)^7\right) & \text{, se } 0 \leq y \leq 1 \\ 0 & \text{, caso contrário.} \end{cases}$$

♦♦

Exemplo 6.10
Seja X variável aleatória tal que

$$f(x) = \begin{cases} \frac{x+1}{4} & \text{, se } -1 < x \leq 1 \\ \frac{3-x}{4} & \text{, se } 1 < x < 3 \\ 0 & \text{, caso contrário.} \end{cases}$$

Defina $Y = |X|$ e encontre f_Y.

Solução:
Primeiro faça o gráfico da função g e encontre, a partir da $Im(X)$ o conjunto $Im(Y)$. Veja na Figura 6.3 abaixo o gráfico da função módulo. Como $Im(X) = (-1, 3)$, então $Im(Y) = (0, 3)$.

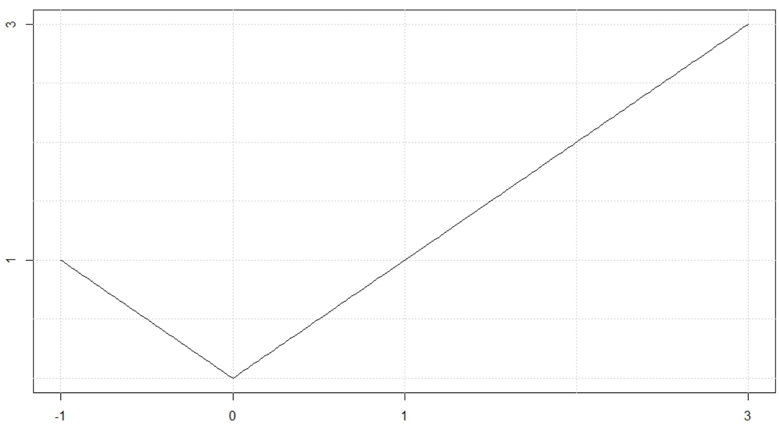


Figura 6.3 – Gráfico da função $g(X) = |X|$.

Queremos encontrar F_Y. Vejamos primeiro os casos triviais: se $y < 0 \Rightarrow F_Y(y) = 0$; e se $y > 3 \Rightarrow F_Y(y) = 1$.

Agora vamos analisar o caso em que $0 \leq y \leq 3$, para isso vamos pensar separadamente em $0 \leq y \leq 1$ e $1 < y \leq 3$. Se $0 \leq y \leq 1$,

$$F_Y(y) = P(Y \leq y) = P(-y \leq X \leq y) = F_X(y) - F_X(-y).$$

Agora, se $1 < y \leq 3$,

$$F_Y(y) = P(Y \leq y) = P(X \leq y) = F_X(y).$$

Assim,

$$F_Y(y) = \begin{cases} 0 & \text{, se } y < 0 \\ F_X(y) - F_X(-y) & \text{, se } 0 \leq y \leq 1 \\ F_X(y) & \text{, se } 1 < y \leq 3 \\ 1 & \text{, se } y > 3. \end{cases}$$

Para encontrar a função densidade podemos derivar em relação à y.

$$f_Y(y) = \begin{cases} f_X(y) + f_X(-y) & \text{, se } 0 \leq y \leq 1 \\ f_X(y) & \text{, se } 1 < y \leq 3 \\ 0 & \text{, caso contrário.} \end{cases}$$

Substituindo pela expressão de f_X chegamos na resposta,

$$f_Y(y) = \begin{cases} 1/2 & \text{, se } 0 \leq y \leq 1 \\ (3-y)/4 & \text{, se } 1 < y \leq 3 \\ 0 & \text{, caso contrário.} \end{cases}$$

Veja que essa função realmente define uma função densidade.

♦♦

6.2 Método do Jacobiano

Quando a função g é inversível, é possível obter uma expressão para a função densidade de $Y = g(X)$ através do resultado apresentado no Teorema 6.11 a seguir.

Teorema 6.11 Método do Jacobiano
Seja X variável aleatória contínua com função densidade f_X. Seja também $g : \mathbb{R} \to \mathbb{R}$ uma função estritamente monótona e diferenciável no conjunto $Im(X)$. Se $Y = g(X)$, então:

(i) $Im(Y) = g(Im(X)) = \{y \in \mathbb{R} \mid \exists\, x \in Im(X) \text{ com } y = g(x)\}$;

(ii) Y é variável aleatória contínua.

(iii) $f_Y(y) = f_X\left(g^{-1}(y)\right)\left|\frac{d}{dy}g^{-1}(y)\right|$, *se* $y \in Im(Y)$;

Demonstração:
O item (i) concluímos que $Im(Y) = g(Im(X))$ uma vez que Y é a composição da função X com a função g. Então, os valores que Y pode assumir são os valores que g pode assumir partindo do conjunto $Im(X)$.

A demonstração do item (ii) será omitida.

A demonstração do item (iii) será feita seguindo o Método da Função de Distribuição. Queremos encontrar $F_Y(y) = \mathrm{P}(Y \leq y)$. Vamos pensar separadamente no caso em que g é estritamente crescente e estritamente decrescente.

Se g é estritamente crescente, existe g^{-1} e esta também é uma função estritamente crescente. Então, nesse caso,

$$F_Y(y) = \mathrm{P}(Y \leq y) = \mathrm{P}(g(X) \leq y) = \mathrm{P}(X \leq g^{-1}(y)) = F_X(g^{-1}(y)).$$

Seguindo, derivando em relação à y, encontramos a relação entre f_Y e f_X:

$$f_Y(y) = \frac{d}{dy} F_X(g^{-1}(y)) = f_X(g^{-1}(y)) \frac{d}{dy} g^{-1}(y).$$

Se g é estritamente decrescente, existe g^{-1} e esta também é uma função estritamente decrescente. Então, nesse caso,

$$F_Y(y) = \mathrm{P}(Y \leq y) = \mathrm{P}(g(X) \leq y) = \mathrm{P}(X \geq g^{-1}(y)) = 1 - F_X(g^{-1}(y)).$$

Seguindo, derivando em relação à y, encontramos a relação entre f_Y e f_X:

$$f_Y(y) = \frac{d}{dy}\left(1 - F_X(g^{-1}(y))\right) = -f_X(g^{-1}(y)) \frac{d}{dy} g^{-1}(y).$$

Veja que no caso de g ser estritamente crescente, $\frac{d}{dy} g^{-1}(y) > 0$ para todo y, então podemos escrever $\frac{d}{dy} g^{-1}(y) = \left|\frac{d}{dy} g^{-1}(y)\right|$.

Já quando g é estritamente decrescente, $\frac{d}{dy} g^{-1}(y) < 0$ para todo y, então podemos escrever $\frac{d}{dy} g^{-1}(y) = -\left|\frac{d}{dy} g^{-1}(y)\right|$.

Dessa forma, qualquer que seja g estritamente monótona,

$$f_Y(y) = f_X(g^{-1}(y)) \left|\frac{d}{dy} g^{-1}(y)\right|.$$

□

Exemplo 6.12
Seja $X \sim U(0,1)$. Obtenha a função densidade de $Y = -\ln X$.

Solução:
A função $g(x) = -\ln x$ é estritamente decrescente e podemos aplicar o Teorema 6.11. Então, como $0 < x < 1$, segue que $0 < Y = -\ln X < \infty$ (ver Figura 6.4).

Por outro lado, a inversa de $y = g(x) = -\ln x$ é $g^{-1}(y) = e^{-y}$ e, portanto,

$$\frac{dg^{-1}(y)}{dy} = -e^{-y}$$

Como $0 < y < \infty$, então $0 < e^{-y} < 1$ e a função densidade de Y é

$$f_Y(y) = f_X\left[e^{-y}\right] \times \left|-e^{-y}\right| = 1 \times e^{-y} \Rightarrow f_Y(y) = e^{-y} \qquad y \in (0,\infty)$$

uma vez que $f_X(x) = 1$ no intervalo $(0,1)$. Note que $Y \sim exp(1)$

♦♦

A partir do Método do Jacobiano podemos encontrar um resultado geral sempre que a transformação que define Y como função de X for linear. Veja a Proposição 6.13 a seguir.

Proposição 6.13 Transformação Linear
Seja X variável aleatória contínua com função densidade f_X e $a, b \in \mathbb{R}$. Defina $Y = aX + b$. Então,

$$f_Y(y) = f_X\left(\frac{y-b}{a}\right)\left|\frac{1}{a}\right|.$$

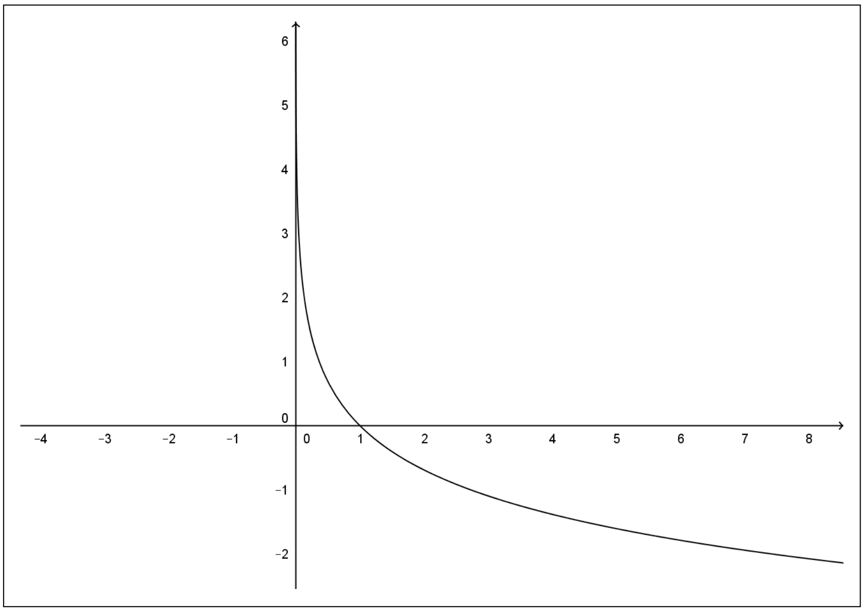


Figura 6.4 – Gráfico da função $g(x) = -\ln x$

Demonstração:
Consideremos a transformação $Y = aX + b$, que define uma reta. Se X é uma variável aleatória contínua com densidade f_X, então podemos aplicar o Teorema 6.11 para calcular a função densidade de $Y = g(X) = aX + b$. Nesse caso, então a função inversa é

$$X = g^{-1}(Y) = \frac{Y - b}{a}$$

cuja derivada é

$$\frac{dg^{-1}(y)}{dy} = \frac{1}{a}$$

Logo, aplicando o Método do Jacobiano, a função densidade de Y é

$$f_Y(y) = f_X\left(\frac{y - b}{a}\right)\left|\frac{1}{a}\right|.$$

$\square$

Exemplo 6.14
Se a função de densidade da variável aleatória X é dada por

$$f(x) = \begin{cases} 3x^2 & \text{, se } -1 \leq x \leq 0 \\ 0 & \text{, caso contrário,} \end{cases}$$

calcule a função densidade de $Y = 2X - \frac{3}{5}$, bem como sua esperança e sua variância.

Solução:
Temos que $a = 2$ e $b = -0,6$. Como $-1 \leq x \leq 0$, resulta que $-2,6 \leq y \leq -0,6$. Logo,

$$f_Y(y) = f_X\left(\frac{y + 0,6}{2}\right) \times \frac{1}{2} \qquad \text{se } -2,6 \leq y \leq -0,6$$

ou seja

$$f_Y(y) = 3\left(\frac{y + 0,6}{2}\right)^2 \times \frac{1}{2} = \frac{3}{8}(y + 0,6)^2 \qquad \text{se } -2,6 \leq y \leq -0,6$$

$$
\begin{aligned}
\mathrm{E}(Y) &= \int_{-2,6}^{-0,6} y \frac{3}{8}(y+0,6)^2\,dy = \frac{3}{8}\int_{-2,6}^{-0,6}\left(y^3+1,2y^2+0,36y\right)dy \\
&= \frac{3}{8}\left[\frac{y^4}{4}+1,2\frac{y^3}{3}+0,36\frac{y^2}{2}\right]_{-2,6}^{-0,6} \\
&= \frac{3}{8}\left[0,25\times(-0,6)^4+0,4\times(-0,6)^3+0,18\times(-0,6)^2\right]- \\
&\ \frac{3}{8}\left[0,25\times(-2,6)^4+0,4\times(-2,6)^3+0,18\times(-2,6)^2\right] \\
&= \frac{3}{8}[0,0108-5,6108] = -2,10
\end{aligned}
$$

$$
\begin{aligned}
\mathrm{E}(Y^2) &= \int_{-2,6}^{-0,6} y^2 \frac{3}{8}(y+0,6)^2\,dy = \frac{3}{8}\int_{-2,6}^{-0,6}\left(y^4+1,2y^3+0,36y^2\right)dy \\
&= \frac{3}{8}\left[\frac{y^5}{5}+1,2\frac{y^4}{4}+0,36\frac{y^3}{3}\right]_{-2,6}^{-0,6} \\
&= \frac{3}{8}\left[0,2\times(-0,6)^5+0,3\times(-0,6)^4+0,12\times(-0,6)^3\right]- \\
&\ \frac{3}{8}\left[0,2\times(-2,6)^5+0,3\times(-2,6)^4+0,12\times(-2,6)^3\right] \\
&= \frac{3}{8}[-0,00259-(-12,162592)] = \frac{3}{8}\times 12,16 = 4,56 \\
\mathrm{Var}(Y) &= 4,56-2,1^2 = 0,15
\end{aligned}
$$

Para o cálculo da esperança e da variância de Y, poderíamos ter usado o fato de que $\mathrm{E}(Y) = 2\,\mathrm{E}(X) - \frac{3}{5}$ e $\mathrm{Var}(Y) = 4\,\mathrm{Var}(X)$. Temos que

$$
\begin{aligned}
\mathrm{E}(X) &= \int_{-1}^{0} x3x^2dx = \frac{3}{4}x^4\Big|_{-1}^{0} = -\frac{3}{4} = -0,75 \\
\mathrm{E}(X^2) &= \int_{-1}^{0} x^2 3x^2dx = \frac{3}{5}x^5\Big|_{-1}^{0} = \frac{3}{5} = 0,6 \\
\mathrm{Var}(X) &= 0,6-0,5625 = 0,0375
\end{aligned}
$$

e isso nos dá

$$
\begin{aligned}
\mathrm{E}(Y) &= 2\,\mathrm{E}(X) - \frac{3}{5} = 2\times(-0,75)-0,6 = -2,1 \\
\mathrm{Var}(Y) &= 4\,\mathrm{Var}(X) = 4\times 0,0375 = 0,15
\end{aligned}
$$

mesmos resultados obtidos anteriormente.

♦♦

Exemplo 6.15

Seja X variável aleatória cuja função densidade é

$$
f_X(x) = \begin{cases} 1/3 & \text{, se } 0 < x \le 1 \\ 4(2-x)/3 & \text{, se } 1 < x < 2 \end{cases}
$$

Seja $Y = X^2$. Encontre a função densidade de Y a partir do Método da Função de Distribuição e também a partir do Método do Jacobiano.

Solução:
Vamos primeiro resolver usando o Método do Jacobiano. Veja que, apesar de $g(x) = x^2$ não ser monótona em toda a reta, para $x > 0$ a função é estritamente crescente. Como $Im(X) = (0, 2) \in (0, \infty)$, podemos aplicar o Método do Jacobiano.

Primeiro veja que $Im(Y) = (0, 4)$, uma vez que $Im(X) = (0, 2)$. Para $x \in Im(X)$, $g^{-1}(y) = \sqrt{y}$ e $\frac{d}{dy} g^{-1}(y) = \frac{1}{2\sqrt{y}}$. Então, de acordo com o Teorema 6.11,

$$f_Y(y) = f_X\left(\sqrt{y}\right) \frac{1}{2\sqrt{y}} \text{ , se } 0 < y < 4.$$

Veja que,

$$f_X\left(\sqrt{y}\right) = \begin{cases} 1/3 & \text{, se } 0 < \sqrt{y} \leq 1 \\ 4(2-\sqrt{y})/3 & \text{, se } 1 < \sqrt{y} < 2 \end{cases} = \begin{cases} 1/3 & \text{, se } 0 < y \leq 1 \\ 4(2-\sqrt{y})/3 & \text{, se } 1 < y < 4. \end{cases}$$

Então,

$$f_Y(y) = \begin{cases} \dfrac{1}{6\sqrt{y}} & \text{, se } 0 < y \leq 1 \\ \\ \dfrac{4(2-\sqrt{y})}{6\sqrt{y}} & \text{, se } 1 < y < 4. \end{cases}$$

Agora vamos resolver o problema usando o Método da Função de Distribuição. Veja que se $y < 0$ temos $F_Y(y) = 0$ e se $y > 4$ temos $F_Y(y) = 1$. Se $0 \leq y \leq 4$ temos,

$$F_Y(y) = \mathrm{P}(Y \leq y) = \mathrm{P}(X^2 \leq y) = \mathrm{P}(X \leq \sqrt{y}) = F_X(\sqrt{y}).$$

Logo,

$$F_Y(y) = \begin{cases} 0 & \text{, se } y < 0 \\ F_X(\sqrt{y}) & \text{, se } 0 \leq y \leq 4 \\ 1 & \text{, se } y > 4 \end{cases}$$

Então, derivando com relação à y temos,

$$f_Y(y) = \begin{cases} f_X(\sqrt{y})\frac{1}{2\sqrt{y}} & \text{, se } 0 \leq y \leq 4 \\ 0 & \text{, caso contrário.} \end{cases}$$

Veja que essa expressão é a mesma que chegamos pelo Método do Jacobiano. Então, substituindo pela expressão de $f_X(\sqrt{y})$ chegamos à mesma resposta.

♦♦

6.2.1 Generalização do Método do Jacobiano

Se $Y = g(X)$ não é monótona em $\mathbb{R}$, podemos aplicar o Teorema 6.11 em cada um dos intervalos em que g é monótona, da seguinte forma:

1. Defina uma partição de $\mathbb{R}_X$ formada pelos intervalos $I_1, I_2, ..., I_k$ tais que a função g é monótona em cada I_j, $j = 1, ..., k$.

2. Aplique o Teorema ?? a I_j, obtendo

$$f_j(y) = f_X[g^{-1}(y)] \left| \frac{dg^{-1}(y)}{dy} \right|$$

3. Finalmente, obtenha
$$f_Y(y) = f_1(y) + f_2(y) + \cdots + f_k(y)$$

Exemplo 6.16 Exemplo 6.2 revisitado
Para $X \sim \mathcal{U}(-1,1)$ vamos calcular a função densidade de $Y = X^2$ e $Z = e^X$.

Solução:
Definindo a função indicadora de um conjunto A como

$$I_A(x) = \begin{cases} 1 & \text{se } x \in A \\ 0 & \text{se } x \notin A \end{cases} \tag{6.1}$$

podemos escrever a função densidade de X como $f_X(x) = \frac{1}{2} I_{(-1,1)}(x)$.

Como $Z = e^X$ é função estritamente crescente, podemos aplicar o Teorema ?? diretamente, obtendo

$$f_Z(z) = f_X(\ln z)\left|\frac{d}{dz}\ln z\right| = \begin{cases} \frac{1}{2}\frac{1}{z} = \frac{1}{2z} & e^{-1} < z < e^{+1} \\ 0 & \text{caso contrário} \end{cases}$$

Por outro lado, $Y = X^2$ não é monótona em $(-1,1)$, mas o é em $(-1,0)$ e $(0,1)$. Assim, aplicamos o Teorema ?? a cada subintervalo e somamos os resultados parciais:

- $(-1,0)$ $\quad X = -\sqrt{Y}$:

$$f_I(y) = f_X(-\sqrt{y})\left|\frac{d(-\sqrt{y})}{dy}\right| = \begin{cases} \frac{1}{2}\left|\frac{1}{2\sqrt{y}}\right| = \frac{1}{4\sqrt{y}} & 0 < y < 1 \\ 0 & \text{caso contrário} \end{cases}$$

- $(0,1)$ $\quad X = \sqrt{Y}$:

$$f_{II}(y) = f_X(\sqrt{y})\left|\frac{d\sqrt{y}}{dy}\right| = \begin{cases} \frac{1}{2}\left|\frac{1}{2\sqrt{y}}\right| = \frac{1}{4\sqrt{y}} & 0 < y < 1 \\ 0 & \text{caso contrário} \end{cases}$$

- $(0,1)$

$$f_Y(y) = f_I(y) + f_{II}(y) = \frac{1}{2\sqrt{y}} I_{(0,1)}(y)$$

♦♦

Capítulo 7

A Distribuição Normal

7.1 Distribuição Normal Padrão

7.1.1 Função Densidade

Analisando a equação (A.34), vemos que a função $\frac{1}{\sqrt{2\pi}} \exp\left(-\frac{t^2}{2}\right)$ satisfaz as condições para ser uma função de densidade. De fato, essa é uma importante distribuição probabilística, denominada *distribuição normal padrão*.

Definição 7.1 Densidade Normal Padrão
Diz-se que uma variável aleatória X tem ***densidade normal padrão*** *se sua função densidade é dada por*

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) \qquad -\infty < x < \infty \tag{7.1}$$

Vamos denotar por $N(0;1)$ a densidade normal padrão e, se uma variável aletária tem tal distribuição, é comum representá-la pela letra Z, de modo que estabelecemos a notação $Z \sim N(0;1)$.

Analisando a expressão de $\varphi(z)$, podemos ver que ela é simétrica em torno de 0, ou seja, ela é uma função par: $\varphi(z) = \varphi(-z)$. Na Figura 7.1 temos o gráfico de $\varphi(z)$.

7.1.2 Esperança e Variância

Seja $Z \sim N(0,1)$. Como $\varphi(z)$ é simétrica em torno do ponto $x = 0$, sabemos, por (??), que $\mathrm{E}(Z) = 0$.

Como $\mathrm{E}(Z) = 0$, então

$$\mathrm{Var}(Z) = \mathrm{E}(Z^2) = \int_{-\infty}^{+\infty} z^2 \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2}\right) dz = \frac{2}{\sqrt{2\pi}} \int_0^{+\infty} z^2 \exp\left(-\frac{z^2}{2}\right) dz$$

uma vez que o integrando é par. Esta integral é calculada usando o método de integração por partes. Fazendo:

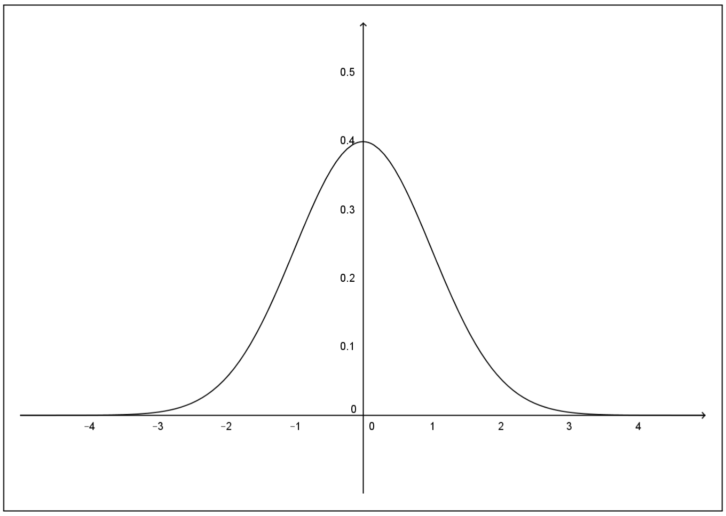


Figura 7.1 – Gráfico da densidade normal padrão $\varphi(z)$

- $z \exp\left(-\frac{z^2}{2}\right) dz = dv \Rightarrow v = -\exp\left(-\frac{z^2}{2}\right)$
- $z = u \Rightarrow dz = du$

resulta que:

$$-z \exp\left(-\frac{z^2}{2}\right)\Bigg|_0^\infty = \int_0^\infty \left[-\exp\left(-\frac{z^2}{2}\right)\right] dz + \int_0^\infty z^2 \exp\left(-\frac{z^2}{2}\right) dz \tag{7.2}$$

Pelos resultados (A.25) e (A.33)

$$0 = -\sqrt{\frac{\pi}{2}} + \int_0^\infty z^2 \exp\left(-\frac{z^2}{2}\right) dz \Longrightarrow \int_0^\infty z^2 \exp\left(-\frac{z^2}{2}\right) dz = \sqrt{\frac{\pi}{2}}$$

Logo,

$$\operatorname{Var}(Z) = \frac{2}{\sqrt{2\pi}} \times \sqrt{\frac{\pi}{2}} \Rightarrow \operatorname{Var}(Z) = 1 \tag{7.3}$$

Resumindo:

$$Z \sim N(0;1) \quad \Rightarrow \quad \begin{cases} \operatorname{E}(Z) = 0 \\ \\ \operatorname{Var}(Z) = 1 \end{cases} \tag{7.4}$$

7.1.3 Função de distribuição

A função de distribuição acumulada de qualquer variável aleatória X é definida por $F_X(X) = \operatorname{P}(X \le x)$. No caso da densidade normal padrão, essa função é dada pela integral

$$\Phi(z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}t^2\right) dt \tag{7.5}$$

para a qual não existe uma antiderivada em forma de função elementar. Assim, cálculos com a distribuição acumulada da normal padrão requerem integração numérica. Todos os pacotes estatísticos possuem rotinas especiais para esse cálculo. No EXCEL, a função DIST.NORMP calcula $\operatorname{P}(Z \le z)$ para qualquer z, onde $Z \sim N(0;1)$.

7.2 Cálculo de Probabilidades da Normal Padrão

Vimos anteriormente que o cálculo de probabilidades associadas a variáveis aleatórias contínuas envolve cálculo de integrais da função densidade:

$$P(a \leq X \leq b) = \int_a^b f_X(x)dx$$

Isso, obviamente, continua valendo para a densidade normal. A diferença está no fato de que o cálculo de tais integrais no caso da densidades normais requer métodos numéricos e, para facilitar esses cálculos, podemos usar uma tabela em que alguns valores já se encontram calculados. A tabela básica fornece probabilidades associadas à densidade normal padrão. Vamos estudar essa tabela e depois veremos como generalizar para uma normal qualquer.

7.2.1 Tabela 1: $P(0 \leq Z \leq z)$

A Tabela 1 será usada para calcular probabilidades associadas a uma variável aleatória normal padrão Z. Assim, com essa tabela, poderemos calcular probabilidades do tipo $P(Z > 1), P(Z \leq 3), P(-1 \leq Z \leq 2)$, etc.

Vamos analisar cuidadosamente esta tabela. A partir do cabeçalho e do gráfico na tabela, podemos ver que as entradas no corpo da tabela fornecem probabilidades do tipo $P(0 \leq Z \leq z)$. Com relação à abscissa z, seus valores são apresentados na tabela ao longo da coluna lateral à esquerda em conjunto com a linha superior, ambas sombreadas de cinza. Na coluna à esquerda, temos a casa inteira e a primeira casa decimal; na linha superior, temos a segunda casa decimal. Por exemplo, ao longo da primeira linha da tabela, temos probabilidades associadas às abscissas 0,00; 0,01; 0,02, . . . , 0,09; na segunda linha da tabela, temos probabilidades associadas às abscissas 0,10; 0,11; 0,12; . . . , 0,19; na última linha da tabela, temos probabilidades associadas às abscissas 4,00; 4,01; 4,02; . . . ; 4,09.

A entrada 0,00000 no canto superior esquerdo da tabela corresponde à seguinte probabilidade: $P(0 \leq Z \leq 0,00)$, ou seja, $P(Z = 0)$ e, como visto, essa probabilidade é nula, uma vez que, para qualquer variável aleatória contínua X, $P(X = x_0) = 0$. A segunda entrada na primeira linha, 0,00399, corresponde a $P(0 \leq Z \leq 0,01)$, que é a área sob a curva de densidade normal padronizada compreendida entre os valores 0 e 0,01 (veja o gráfico na tabela).

Note que esta tabela apresenta probabilidades correspondentes a abscissas positivas, ou seja, esta tabela trata de área sob a curva no lado positivo do eixo. Para calcular áreas no lado negativo, teremos que usar o fato de a curva da densidade normal ser simétrica. Sempre faça um esboço da curva de densidade, sombreando a área correspondente à probabilidade desejada; isso lhe ajudará no cálculo da probabilidade. Vamos terminar esta seção apresentando vários exemplos de cálculos de probabilidades de uma v.a. Z com distribuição normal padrão, ou seja, no que segue, $Z \sim N(0; 1)$. Os exemplos apresentados cobrem todas as situações possíveis. Assim, é importante que você entenda bem a situação ilustrada por cada um dos exemplos, para poder aplicar o método de solução adequado aos novos exercícios.

Para simplificar a solução dos exercícios, vamos adotar a seguinte notação.

! Entradas da Tabela 1 do Apêndice B

Vamos representar por $tab(z)$ as entradas da Tabela 1 do Apêndice B, ou seja,

$$tab(z) = P(0 \leq Z \leq z)$$

Exemplo 7.2
A partir da Tabela 1 do Apêndice B calcule $P(0 \leq Z \leq 1,22)$.

Solução:
Veja as Figuras 7.2 e 7.3. Essa probabilidade é dada diretamente na Tabela 1, utilizando a entrada correspondente à linha 1,2 e à coluna com o valor 2. O resultado é

$$P(0 \leq Z \leq 1,22) = tab(1,22) = 0,3888$$

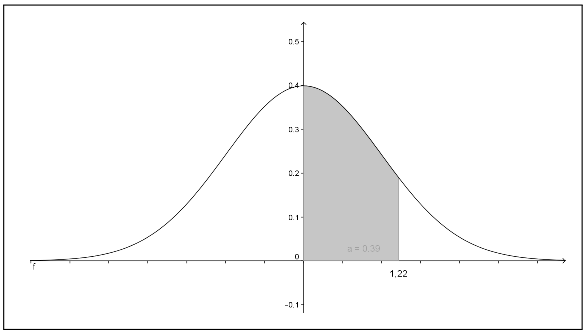


Figura 7.2 – $P(0 \leq Z \leq 1,22)$ como área

e 1ª. Decimal	0	1	2	3
0,0	0,0000	0,0040	0,0080	0,0120
0,1	0,0398	0,0438	0,0478	0,0517
0,9	0,3159	0,3186	0,3212	0,3238
1,0	0,3413	0,3438	0,3461	0,3485
1,1	0,3643	0,3665	0,3686	0,3708
1,2	0,3849	0,3869	**0,3888**	0,3907
1,3	0,4032	0,4049	0,4066	0,4082

Figura 7.3 – $P(0 \leq Z \leq 1,22)$ - Uso da tabela

♦♦

Exemplo 7.3
A partir da Tabela 1 do Apêndice B calcule $P(1 \leq Z \leq 2)$.

Solução:
Note que este exemplo trata da área (probabilidade) entre duas abscissas positivas. Na Figura 7.4 ilustra-se a área (probabilidade) desejada. Note que esta área pode ser obtida pela diferença entre as áreas das Figuras 7.5 e 7.7, cujos valores são encontrados na Tabela 1 conforme ilustram as Figuras 7.6 e 7.8.

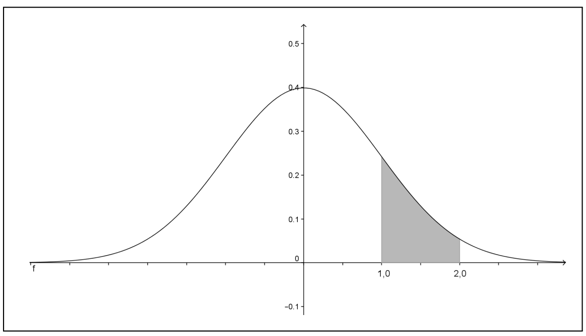


Figura 7.4 – $P(1 \leq Z \leq 2)$ como área

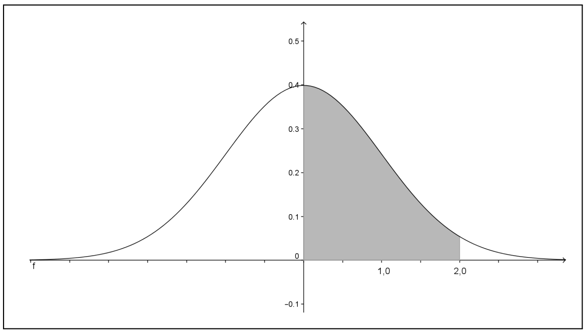


Figura 7.5 – $P(0 \leq Z \leq 2)$ como área

e 1ª. Decimal	0	1	2	3	4
0,0	0,0000	0,0040	0,0080	0,0120	0,0160
0,1	0,0398	0,0438	0,0478	0,0517	0,0557
0,2	0,0793	0,0832	0,0871	0,0910	0,0948
1,8	0,4641	0,4649	0,4656	0,4664	0,4671
1,9	0,4713	0,4719	0,4726	0,4732	0,4738
2,0	**0,4772**	0,4778	0,4783	0,4788	0,4793
2,1	0,4821	0,4826	0,4830	0,4834	0,4838

Figura 7.6 – $P(0 \leq Z \leq 2)$ - Uso da tabela

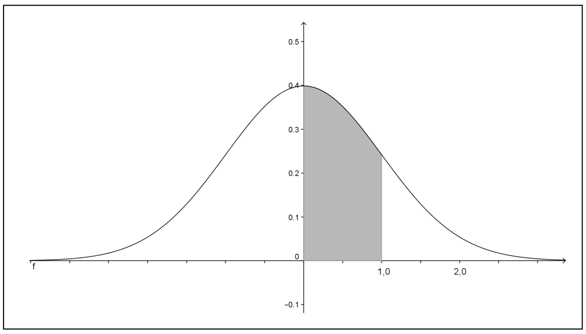


Figura 7.7 – $P(0 \leq Z \leq 1)$ como área

e 1ª. Decimal	0	1	2	3	4
0,0	0,0000	0,0040	0,0080	0,0120	0,0160
0,1	0,0398	0,0438	0,0478	0,0517	0,0557
0,2	0,0793	0,0832	0,0871	0,0910	0,0948
0,8	0,2881	0,2910	0,2939	0,2967	0,2995
0,9	0,3159	0,3186	0,3212	0,3238	0,3264
1,0	**0,3413**	0,3438	0,3461	0,3485	0,3508
1,1	0,3643	0,3665	0,3686	0,3708	0,3729

Figura 7.8 – $P(0 \leq Z \leq 1)$ - Uso da tabela

Concluimos, então, que

$$P(1 \leq Z \leq 2) = P(0 \leq Z \leq 2) - P(0 \leq Z \leq 1) = tab(2,0) - tab(1,0) = 0,4772 - 0,3413 = 0,1359$$

♦♦

Exemplo 7.4
A partir da Tabela 1 do Apêndice B calcule $P(Z \geq 1)$.

Solução:
Note que este exemplo trata da área (probabilidade) à direita ($\geq$) de uma abscissa positiva. Na Figura 7.9 ilustra-se a área (probabilidade) desejada. Note que esta área pode ser obtida pela diferença entre as áreas das Figuras 7.10 e 7.11. A primeira área corresponde à probabilidade $P(Z \geq 0)$ e é igual a 0,5, pois a média $\mu = 0$ é o eixo de simetria e a área total é 1. Logo, $P(Z \geq 0) = P(Z \leq 0) = 0,5$. A segunda área vem direto da Tabela 1.

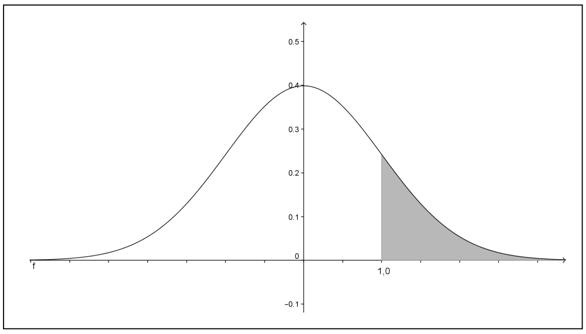


Figura 7.9 – $P(Z \geq 1)$

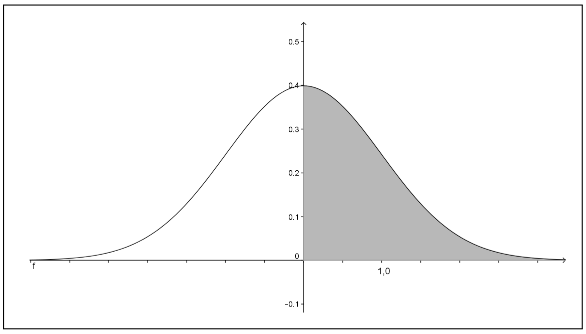


Figura 7.10 – $\mathrm{P}(Z \geq 0)$

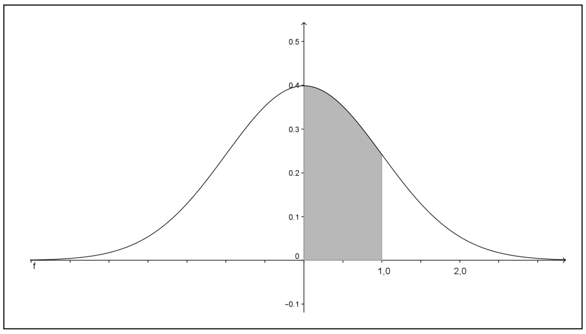


Figura 7.11 – $\mathrm{P}(0 \leq Z \leq 1)$

Concluimos, então, que

$$\mathrm{P}(Z \geq 1) = \mathrm{P}(Z \geq 0) - \mathrm{P}(0 \leq Z \leq 1) = 0,5 - tab(1,0) = 0,5 - 0,3413 = 0,1587$$

♦♦

Exemplo 7.5
A partir da Tabela 1 do Apêndice B calcule $\mathrm{P}(Z \leq 1)$.

Solução:
Note que este exemplo trata da área (probabilidade) à esquerda ($\leq$) de uma abscissa positiva. Na Figura 7.12 ilustra-se a área (probabilidade) desejada. Note que esta área pode ser obtida pela soma das áreas das Figuras 7.13 e 7.14. A primeira área corresponde à probabilidade $\mathrm{P}(Z \leq 0)$ e é igual a 0,5, conforme visto no exemplo anterior.

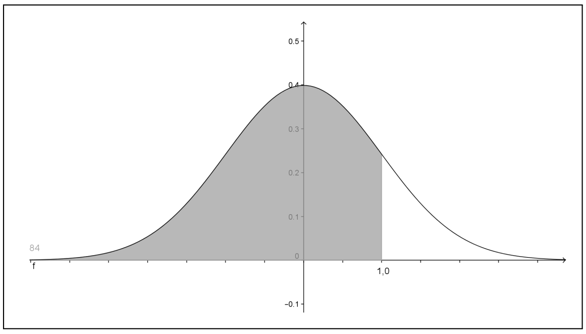


Figura 7.12 – $\mathrm{P}(Z \leq 1)$

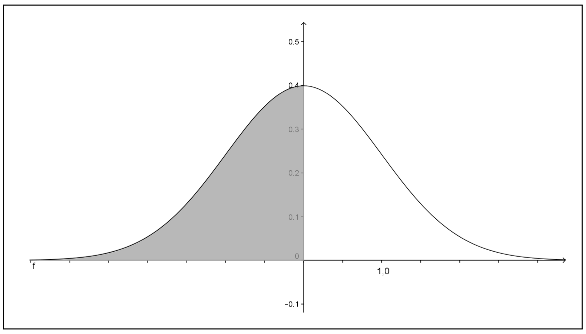


Figura 7.13 – $\mathrm{P}(Z \leq 0)$

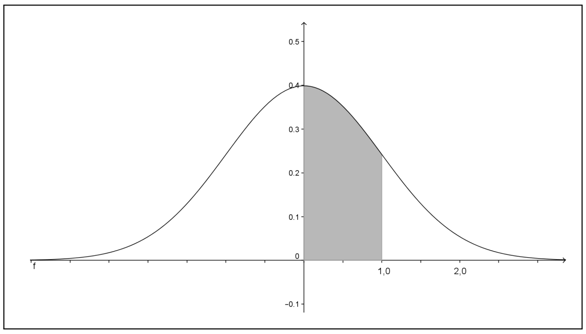


Figura 7.14 – $\mathrm{P}(0 \leq Z \leq 1)$

Concluimos, então, que

$$\mathrm{P}(Z \geq 1) = \mathrm{P}(Z \geq 0) + \mathrm{P}(0 \leq Z \leq 1) = 0,5 + tab(1,0) = 0,5 + 0,3413 = 0,8413$$

♦♦

Exemplo 7.6
A partir da Tabela 1 do Apêndice B calcule $\mathrm{P}(Z \leq -0,5)$

Solução:
Note que este exemplo trata da área (probabilidade) à esquerda ($\leq$) de uma abscissa negativa e, agora, começamos a trabalhar com abscissas negativas. Na Figura 7.15 ilustra-se a área (probabilidade) desejada. Pela simetria da curva de densidade normal, resulta que essa área é igual à área ilustrada na Figura 7.16.

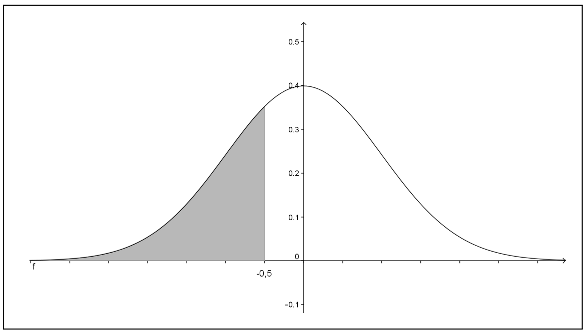


Figura 7.15 – $\mathrm{P}(Z \leq -0,5)$

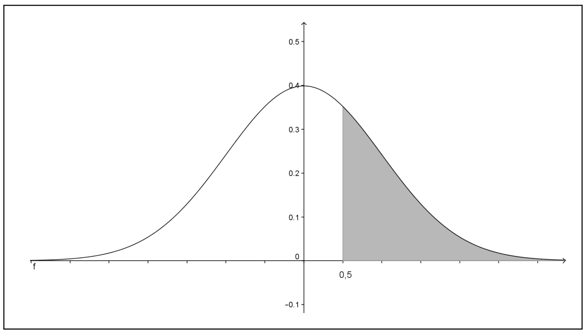


Figura 7.16 – $\mathrm{P}(Z \geq 0,5)$

Concluimos, então, que

$$\mathrm{P}(Z \leq -0,5) = \mathrm{P}(Z \geq 0,5) = 0,5 - \mathrm{P}(0 \leq Z < 0,5) = 0,5 - tab(0,5) = 0,5 - 0,1915 = 0,3085$$

♦♦

Exemplo 7.7
A partir da Tabela 1 do Apêndice B calcule $\mathrm{P}(Z \geq -0,5)$

Solução:
Note que este exemplo trata da área (probabilidade) à direira ($\geq$) de uma abscissa negativa. Na Figura 7.17 ilustra-se a área (probabilidade) desejada. Essa área é a soma das áreas representadas nas Figuras 7.18 e 7.19. Essa última área, por sua vez, é igual à area representada na Figura 7.20, pela simetria da curva de densidade.

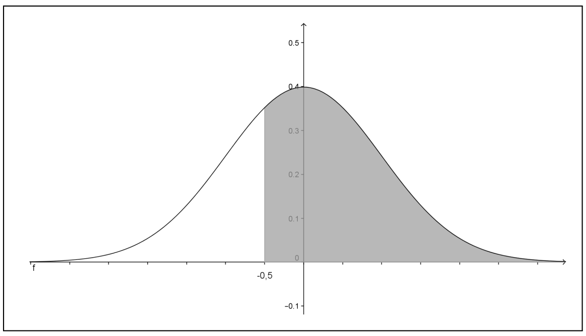


Figura 7.17 – $\mathrm{P}(Z \geq -0,5)$

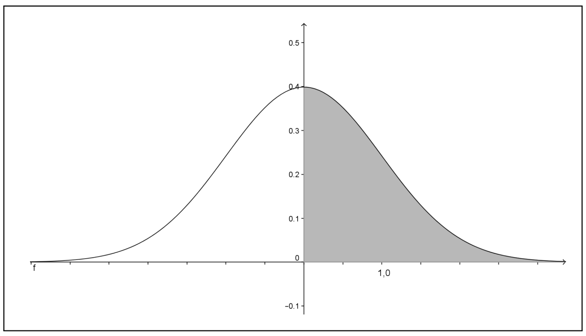


Figura 7.18 – $\mathrm{P}(Z \geq 0)$

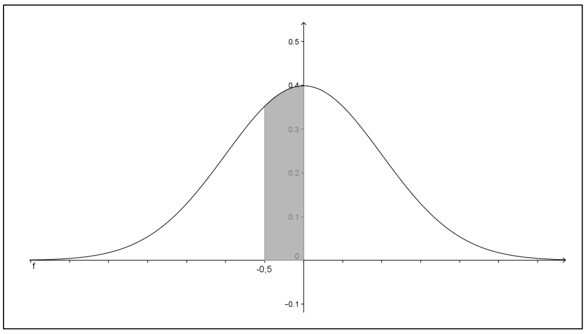


Figura 7.19 – $P(-0,5 \leq Z \leq 0)$

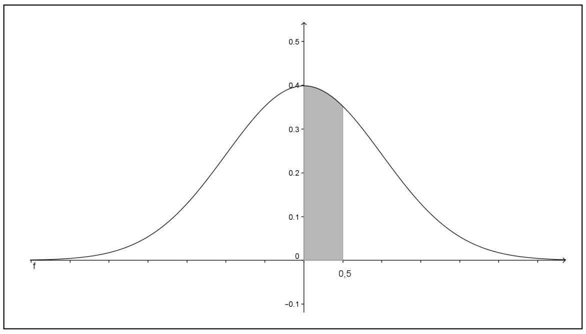


Figura 7.20 – $P() \leq Z \leq 0,5)$

Concluimos, então, que

$$
\begin{aligned}
P(Z \geq -0,5) &= P(-0,5 \leq Z \leq 0) + P(Z \geq 0) \\
&= P(0 \leq Z < 0,5) + 0,5 = tab(0,5) + 0,5 = 0,1915 + 0,5 = 0,6915
\end{aligned}
$$

♦♦

Exemplo 7.8
A partir da Tabela 1 do Apêndice B calcule $P(-2,1 \leq Z \leq -1,4)$

Solução:
Note que este exemplo trata da área (probabilidade) entre duas abscissas negativas. Na Figura 7.21 ilustra-se a área (probabilidade) desejada. Por simetria, essa área é igual à área ilustrada na Figura 7.22, já analisada no Exemplo 7.3.

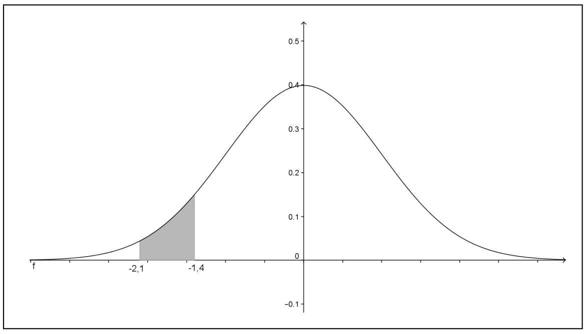


Figura 7.21 – $P(-2,1 \leq Z \leq -1,4)$

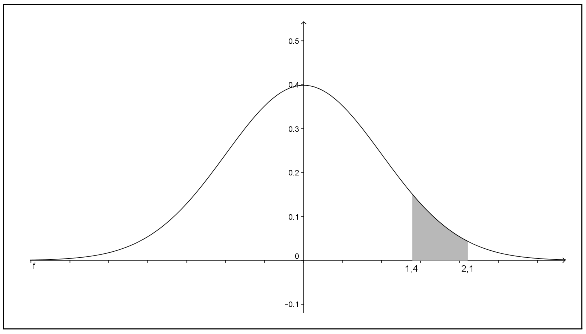


Figura 7.22 – $P(1,4 \leq Z \leq -2,1)$

Concluimos, então, que

$$
\begin{aligned}
P(-2,1 \leq Z \leq -1,4) &= P(1,4 \leq Z \leq 2,1) = P(0 \leq Z \leq 2,1) - P(0 \leq Z \leq 1,4) \\
&= tab(2,1) - tab(1,4) = 0,4821 - 0,4192 = 0,0629
\end{aligned}
$$

♦♦

Exemplo 7.9
A partir da Tabela 1 do Apêndice B calcule $P(-2,1 \leq Z \leq 1,4)$

Solução:
Note que este exemplo trata da área (probabilidade) entre duas abscissas, uma negativa e outra positiva. Na Figura 7.23 ilustra-se a área (probabilidade) desejada. Essa área é a soma das áreas representadas nas Figuras 7.24 e 7.25. Por simetria, essa última área é igual à

área sombreada na Figura 7.26, o que nos leva à cocnlusão de que

$$
\begin{aligned}
P(-2,1 \leq Z \leq 1,4) &= P(0 \leq Z \leq 1,4) + P(-2,1 \leq Z \leq 0) \\
&= P(0 \leq Z \leq 1,4) + P(0 \leq Z \leq 2,1) = tab(1,4) + tab(2,1) \\
&= 0,4821 + 0,4192 = 0,9013
\end{aligned}
$$

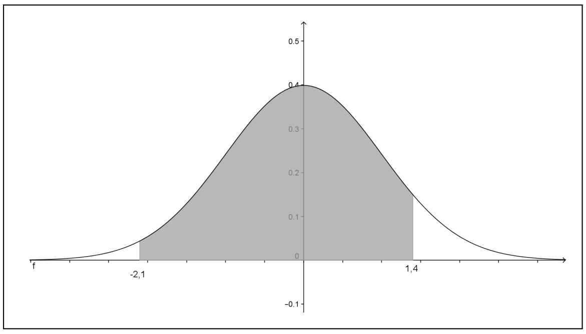


Figura 7.23 – $P(-2,1 \leq Z \leq 1,4)$

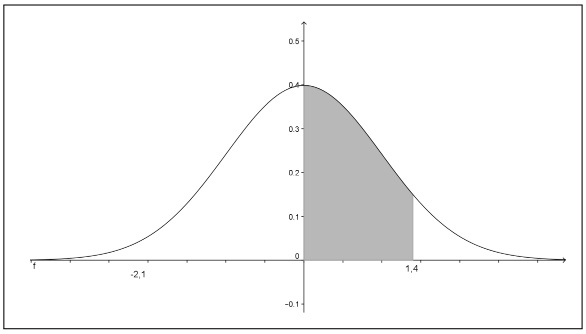


Figura 7.24 – $P(0 \leq Z \leq 1,4)$

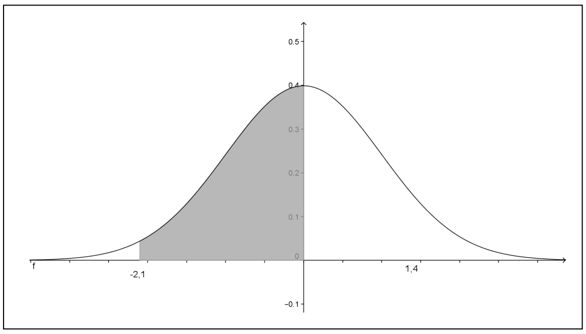


Figura 7.25 – $P(-2,1 \leq Z \leq 0)$

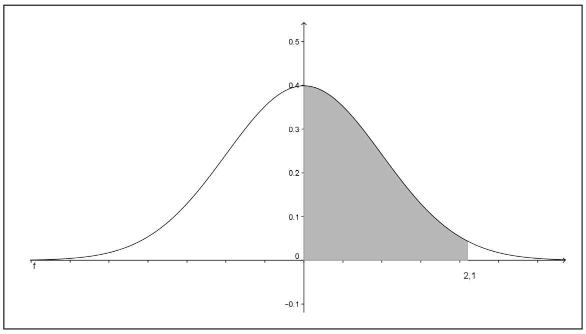


Figura 7.26 – $P(0 \leq Z \leq 2,1)$

♦♦

7.2.2 Tabela 2: $\Phi(z) = P(Z \leq z)$

Muitos livros trabalham com a tabela da função de distribuição da normal padrão, que, como vimos, representamos pela letra grega fi maiúscula, Φ:

$$\Phi(z) = P(Z \leq z).$$

A Tabela 2 do Apêndice B apresenta os valores de $\Phi(z)$ para $z \geq 0$. Vamos usar essa tabela para refazer os exemplos vistos anteriormente, que serão apresentados em uma ordem diferente, mais didaticamente apropriada para esse contexto.

Exemplo 7.10
A partir da Tabela 2 do Apêndice B calcule $P(Z \leq 1)$

Solução:
Essa probabilidade resulta diretamente da definição de distribuição acumulada:

$$P(Z \leq 1) = \Phi(1,0) = 0,8413$$

♦♦

Exemplo 7.11
A partir da Tabela 2 do Apêndice B calcule $P(Z \geq 1)$

Solução:
Pela lei do complementar, temos que

$$P(Z \geq 1) = 1 - P(Z < 1)$$

Mas, como Z é uma variável aleatória contínua, sabemos que $P(Z = z) = 0$. Logo

$$P(Z < z) = P(Z \leq z)$$

Logo,

$$P(Z \geq 1) = 1 - P(Z < 1) = 1 - P(Z \leq 1) = 1 - \Phi(1,0) = 1 - 0,8413 = 0,1587$$

♦♦

Exemplo 7.12
A partir da Tabela 2 do Apêndice B calcule $P(Z \leq -0,5)$

Solução:
Vimos, no Exemplo 7.6, que

$$P(Z \leq -0,5) = P(Z \geq 0,5)$$

Logo,

$$P(Z \leq -0,5) = P(Z \geq 0,5) = 1 - P(Z < 0,5) = 1 - P(Z \leq 0,5) = 1 - \Phi(0,5) = 1 - 0,6915 = 0,3085$$

♦♦

Exemplo 7.13
A partir da Tabela 2 do Apêndice B calcule $P(Z \geq -0,5)$

Solução:
Veja as Figuras 7.27 e 7.28.

$$\begin{aligned} P(Z \geq -0,5) &= 1 - P(Z < -0,5) = 1 - P(Z > 0,5) = 1 - [1 - P(Z \leq 0,5)] \\ &= P(Z \leq 0,5) = \Phi(0,5) = 0,6915 \end{aligned}$$

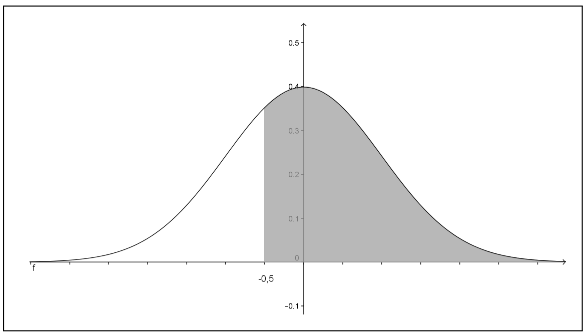


Figura 7.27 – $P(Z \geq -0,5)$

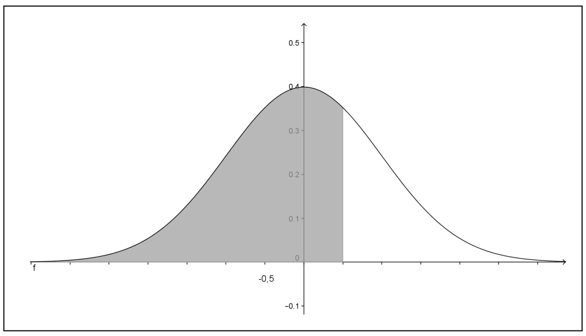


Figura 7.28 – $P(Z \leq 0,5)$

♦♦

Exemplo 7.14
A partir da Tabela 2 do Apêndice B calcule $P(0 \leq Z \leq 1,22)$

Solução:
Veja as Figuras 7.2, 7.29 e 7.30.

$$P(0 \leq Z \leq 1,22) = P(Z \leq 1,22) - P(Z \leq 0) = \Phi(1,22) - 0,5 = 0,8888 - 0,5 = 0,3888$$

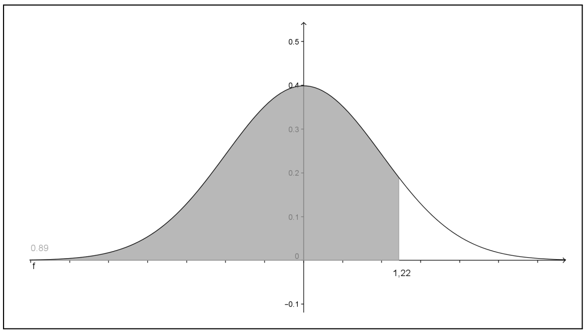


Figura 7.29 – $P(Z \leq 1,22)$

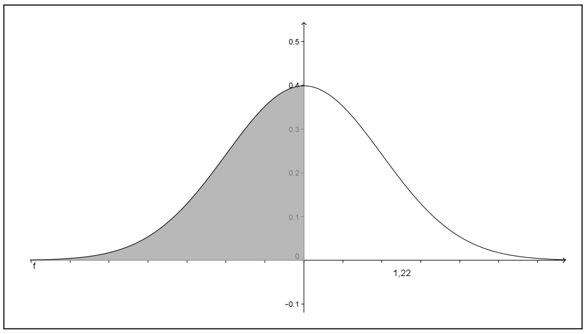


Figura 7.30 – $P(Z \leq 0)$

♦♦

Exemplo 7.15
A partir da Tabela 2 do Apêndice B calcule $P(1 \leq Z \leq 2)$

Solução:
Veja as Figuras 7.4, 7.31 e 7.32.

$$\begin{aligned} P(1 \leq Z \leq 2) &= P(Z \leq 2) - P(Z < 1) = P(Z \leq 2) - P(Z \leq 1) = \Phi(2,0) - \Phi(1,0) \\ &= 0,9772 - 0,8413 = 0,1359 \end{aligned}$$

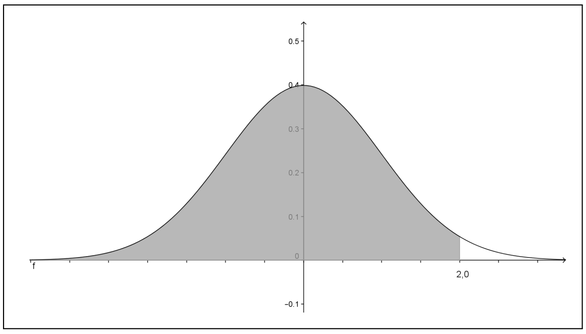


Figura 7.31 – $P(Z \geq -0,5)$

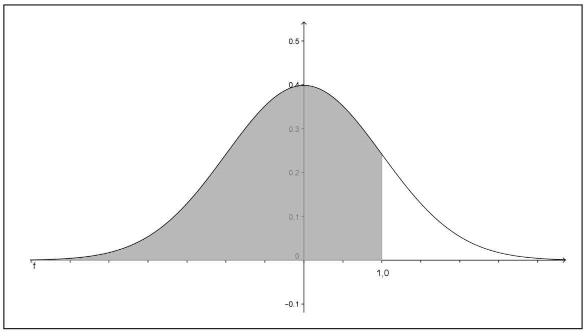


Figura 7.32 – $P(Z \leq 0,5)$

♦♦

Exemplo 7.16
A partir da Tabela 2 do Apêndice B calcule $P(-2,1 \leq Z \leq -1,4)$

Solução:
Usando os resultados do Exemplo 7.15, temos que

$$P(-2,1 \leq Z \leq -1,4) = P(1,4 \leq Z \leq 2,1) = \Phi(2,1) - \Phi(1,4) = 0,9821 - 0,9192 = 0,0629$$

♦♦

Exemplo 7.17
A partir da Tabela 2 do Apêndice B calcule $P(-2,1 \leq Z \leq 1,4)$

Solução:
Usando os resultados do Exemplo 7.12, temos que

$$\begin{aligned} P(-2,1 \leq Z \leq 1,4) &= \Phi(1,4) - P(Z < -2,1) = \Phi(1,4) - \Phi(-2,1) \\ &= \Phi(1,4) - [1 - \Phi(2,1)] = 0,9192 - [1 - 0,9821] = 0,9013 \end{aligned}$$

♦♦

7.3 Distribuição Normal

Seja $Z \sim N(0;1)$ e vamos definir uma nova variável aleatória por

$$X = g(Z) = \mu + \sigma Z,$$

em que $\sigma > 0$. Usando a Proposição 6.13 temos que

$$f_X(x) = f_Z\left(\frac{x-\mu}{\sigma}\right) \times \frac{1}{\sigma} = \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right] \times \frac{1}{\sigma}$$

ou ainda:

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right]$$

e essa é a densidade normal $N(\mu;\sigma^2)$.

Definição 7.18 Distribuição Normal
Diz-se que, uma variável aleatória contínua X, definida para todos os valores em $\mathbb{R}$, tem ***distribuição normal*** *com parâmetros μ e σ^2, onde $-\infty < \mu < \infty$ e $0 < \sigma^2 < \infty$, se sua função densidade de probabilidade é dada por*

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right] \qquad -\infty < x < \infty \qquad .$$

Usaremos a seguinte notação para indicar que uma v.a. X tem distribuição normal com parâmetros μ e σ^2 : $X \sim N(\mu, \sigma^2)$.

7.3.1 Características da Função Densidade Normal

1. Simétrica em torno de μ; note que $f(\mu - x) = f(\mu + x)$.
2. Assíntotas: $\lim\limits_{x\to-\infty} f(x) = \lim\limits_{x\to\infty} f(x) = 0$; esse resultado segue diretamente dos resultados sobre a função exponencial dados em (A.16) e (A.17) e do fato de que $e^{-x} = \dfrac{1}{e^x}$.
3. Ponto de máximo

   Para calcular a primeira e segunda derivadas de $f(x)$, devemos lembrar que $(e^x)' = e^x$ e, pela regra da cadeia, $(e^{g(x)})' = e^{g(x)}g'(x)$. Aplicando esses resultados à densidade normal, obtemos que:

$$f'(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right]\left[-\frac{1}{2\sigma^2}2(x-\mu)\right] = -f(x)\left(\frac{x-\mu}{\sigma^2}\right) \qquad (7.6)$$

Derivando novamente, obtemos:

$$\begin{aligned} f''(x) &= -f'(x)\left(\frac{x-\mu}{\sigma^2}\right) - f(x)\frac{1}{\sigma^2} = -\left[-f(x)\left(\frac{x-\mu}{\sigma^2}\right)\right]\left[\frac{x-\mu}{\sigma^2}\right] - f(x)\frac{1}{\sigma^2} = \\ &= f(x)\left[\frac{(x-\mu)^2}{\sigma^4}\right] - f(x)\frac{1}{\sigma^2} = f(x)\left[\frac{(x-\mu)^2-\sigma^2}{\sigma^4}\right] \end{aligned} \tag{7.7}$$

Analisando a equação (7.6) e lembrando que $f(x) > 0$, pode-se ver que:

$$f'(x) = 0 \Leftrightarrow x = \mu$$

e assim, $x = \mu$ é um ponto crítico. Como $f'(x) > 0$ para $x < \mu$ e $f'(x) < 0$ para $x > \mu$, então f é crescente à esquerda de μ e decrescente à direita de μ. Segue, então, que $x = \mu$ é um ponto de máximo e nesse ponto

$$f(\mu) = \frac{1}{\sqrt{2\pi\sigma^2}} \tag{7.8}$$

4. Pontos de inflexão

   Analisando a segunda derivada dada por (7.7), tem-se que:

$$f''(x) = 0 \Leftrightarrow (x-\mu)^2 = \sigma^2 \Leftrightarrow |x-\mu| = \sigma \Leftrightarrow \begin{cases} x = \mu + \sigma \\ x = \mu - \sigma \end{cases} \tag{7.9}$$

   Além disso,

$$\begin{aligned} f''(x) &> 0 \Leftrightarrow (x-\mu)^2 > \sigma^2 \Leftrightarrow |x-\mu| > \sigma \Leftrightarrow \\ &\Leftrightarrow x-\mu > \sigma \quad \text{ou} \quad \mu - x > \sigma \\ \Leftrightarrow &\quad x > \mu+\sigma \quad \text{ou} \quad x < \mu - \sigma \end{aligned} \tag{7.10}$$

   e

$$\begin{aligned} f''(x) &< 0 \Leftrightarrow (x-\mu)^2 < \sigma^2 \Leftrightarrow |x-\mu| < \sigma \Leftrightarrow \\ &\Leftrightarrow \begin{cases} x-\mu < \sigma \\ \mu - x < \sigma \end{cases} \Leftrightarrow \mu - \sigma < x < \mu + \sigma \end{aligned} \tag{7.11}$$

   Logo, $f(x)$ é côncava para cima se $x > \mu+\sigma$ ou $x < \mu-\sigma$ e é côncava para baixo quando $\mu - \sigma < x < \mu + \sigma$.

   Na Figura 7.33 é apresentado o gráfico da densidade normal no caso em que $\mu = 3$ e $\sigma^2 = 1$. Aí a linha pontilhada central representa o eixo de simetria e as linhas pontilhadas laterais passam pelos pontos de inflexão 3 ± 1.

7.3.2 Parâmetros Distribuição Normal

Se $X \sim N\left(\mu; \sigma^2\right)$, então $X = \mu + \sigma Z$, em que $Z \sim N(0;1)$. Das propriedades de média e variância, resulta

$$\mathrm{E}(X) = \mu + \sigma\,\mathrm{E}(Z) = \mu + 0 \Rightarrow \mathrm{E}\,(X) = \mu \tag{7.12}$$

e

$$\mathrm{Var}\,(X) = \sigma^2\,\mathrm{Var}(Z) == \sigma^2 \times 1 \Rightarrow \mathrm{Var}(X) = \sigma^2 \tag{7.13}$$

Resumindo:

$$X \sim N\left(\mu; \sigma^2\right) \Longrightarrow \begin{cases} \mathrm{E}(X) = \mu \\ \mathrm{Var}(X) = \sigma^2 \end{cases} \tag{7.14}$$

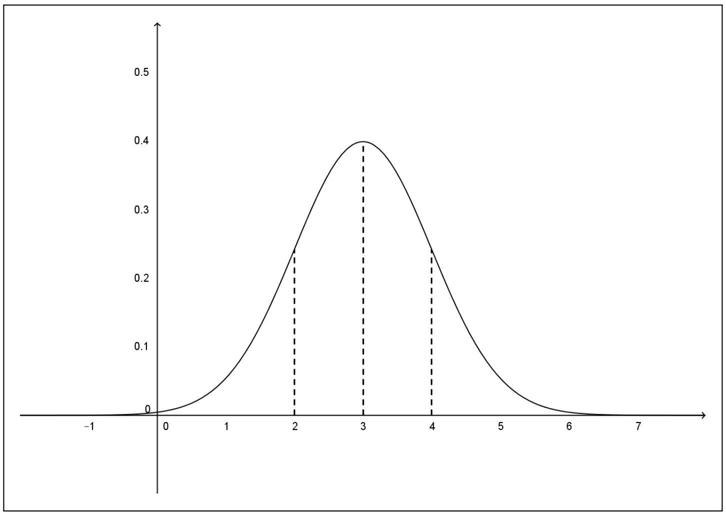


Figura 7.33 – Densidade normal com média $\mu = 3$ e variância $\sigma^2 = 1$

Os parâmetros da densidade normal são, então, a média e a variância, que são medidas de posição e dispersão, respectivamente. Valores diferentes de μ deslocam o eixo de simetria da curva e valores diferentes de σ^2 mudam a dispersão da curva. Quanto maior σ^2, mais "espalhada" é a curva, mas o ponto de máximo, dado pela equação (7.8), é inversamente proporcional a σ^2. Logo, quanto maior σ^2, mais "espalhada" e mais "achatada" é a curva. O importante a observar é que a forma é sempre a de um "sino". Na Figura 7.34 temos exemplos de densidades normais com a mesma variância, mas com médias diferentes. O efeito é o "delocamento" do eixo de simetria da densidade. Já na Figura 7.35, temos duas densidades com a mesma média, mas variâncias diferentes. O efeito é que a densidade com maior variância é mais dispersa e achatada.

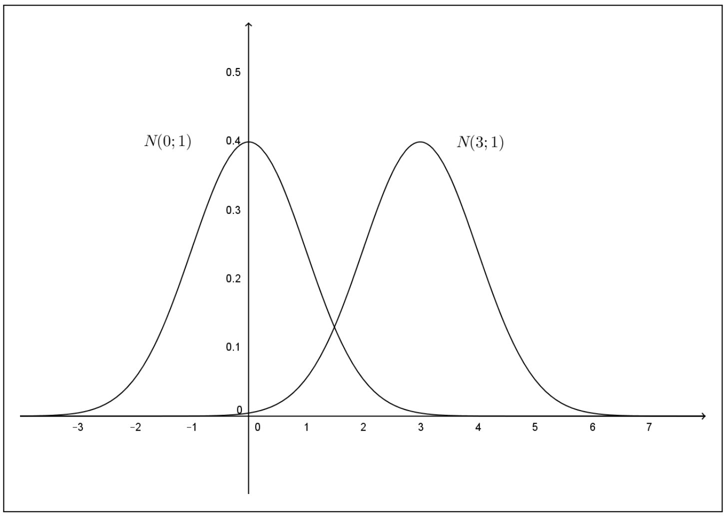


Figura 7.34 – Densidades normais com mesma variância e médias diferentes

7.3.3 Função de Distribuição

A função de distribuição da densidade normal é dada pela integral

$$F(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2\right] dt \qquad (7.15)$$

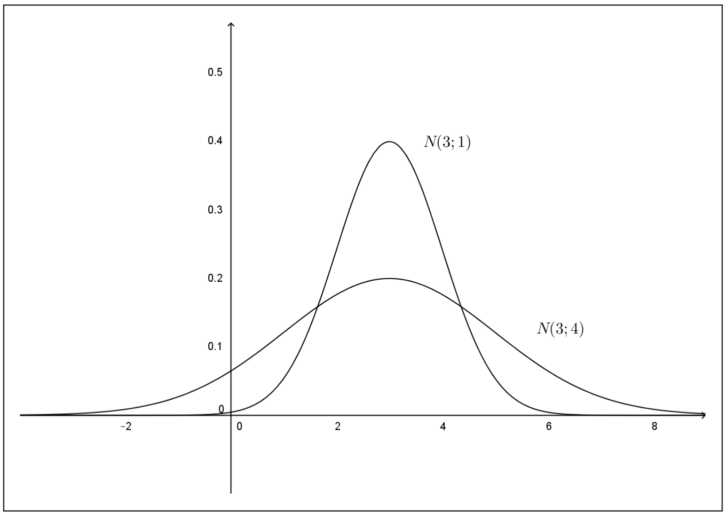


Figura 7.35 – Densidades normais com mesma média e variâncias diferentes

Na Figura 7.36 apresentamos função de distribuição associada às densidades $N(0;1)$, $N(3;1)$ e $N(3;4)$. Note que, pela simetria da densidade em torno da média μ, sempre teremos $F(\mu) = 0,5$. Esse fato é ilustrado com as linhas pontilhadas na figura.

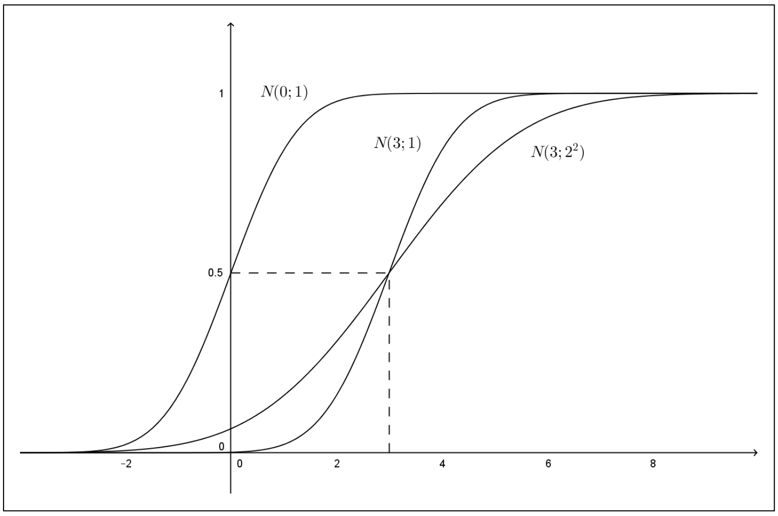


Figura 7.36 – Função de distribuição da $N(0;1)$, $N(3;1)$ e $N(3;4)$

7.4 Cálculo de Probabilidades da Normal

Já foi visto que se $X \sim N\left(\mu, \sigma^2\right)$, então $X = \mu + \sigma Z$, onde $Z \sim N(0,1)$. Vamos ver como utilizar esse resultado para calcular probabilidades da normal. Temos que

$$\mathrm{P}\left(X \leq x\right) = \mathrm{P}\left(\mu + \sigma Z \leq x\right) = \mathrm{P}\left(Z \leq \frac{x-\mu}{\sigma}\right) = \Phi\left(\frac{x-\mu}{\sigma}\right) \tag{7.16}$$

Nas Figuras 7.37 e 7.38 ilustra-se esse fato, utilizando as densidades $Z \sim N(0;1)$ e $X \sim N(3;4)$. No gráfico à esquerda, a área sombreada representa $\mathrm{P}(X \leq 5)$ e no gráfico à direita, a área sombreada representa a probabilidade equivalente:

$$\mathrm{P}(X \leq 5) = \mathrm{P}\left(\frac{X-3}{2} \leq \frac{5-3}{2}\right) = \mathrm{P}(Z \leq 1)$$

O que o resultado diz é que essas áreas (probabilidades) são iguais.

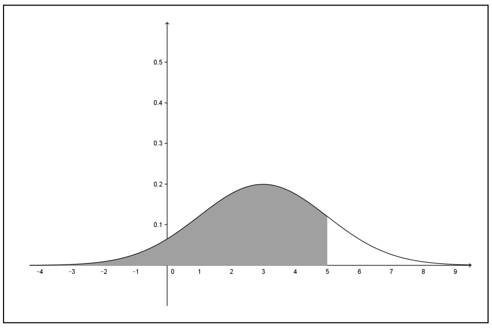


Figura 7.37 – $X \sim N(;3^2) - -\mathrm{P}(X \leq 5)$

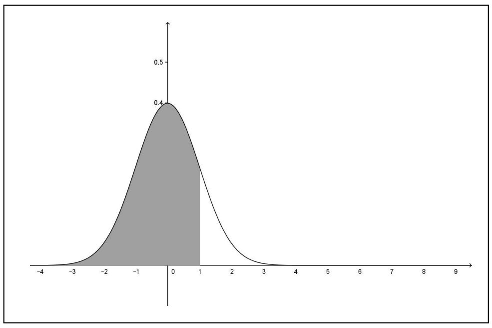


Figura 7.38 – $Z \sim N(0;1) - -\mathrm{P}(Z \leq 1,0)$

Isso significa que probabilidades de uma $N(\mu;\sigma^2)$ podem ser calculadas a partir da operação de padronização.

! **Padronização da distribuição normal** $N(\mu;\sigma^2)$

Se $X \sim N\left(\mu;\sigma^2\right)$, então

$$Z = \frac{X - \mu}{\sigma} \tag{7.17}$$

tem distribuição $N(0;1)$.

É interessante lembrar que a transformação dada na equação (7.4) corresponde ao cálculo do escore padronizado associado à abscissa x. Assim, cálculos de probabilidades de v.a. normais sempre envolverão o cálculo do escore padronizado da(s) abscissa(s) de interesse.

Exemplo 7.19

Seja $X \sim N(3;9)$, *clacule* $\mathrm{P}(-1 \leq X \leq 4)$

Solução:

$$\begin{aligned}
\mathrm{P}(-1 \leq X \leq 4) &= \mathrm{P}\left(\frac{-1-3}{\sqrt{9}} \leq \frac{X-3}{\sqrt{9}} \leq \frac{4-3}{\sqrt{9}}\right) \\
&= \mathrm{P}\left(-1,33 \leq Z \leq 0,33\right) \\
&= \Phi(0,33) - \Phi(-1,33) = 0,62930 - 0,09176 \\
&= tab(0,33) + tab(1,33) = 0,12930 + 0,40824 \\
&= 0,53754
\end{aligned}$$

♦♦

Exemplo 7.20

Seja $X \sim N(2;5)$, *calcule* $\mathrm{P}(1 \leq X \leq 7)$

Solução:

$$\begin{aligned}
\mathrm{P}(1 \leq X \leq 7) &= \mathrm{P}\left(\frac{1-2}{\sqrt{5}} \leq \frac{X-2}{\sqrt{5}} \leq \frac{7-2}{\sqrt{5}}\right) \\
&= \mathrm{P}\left(-0,45 \leq Z \leq 2,24\right) \\
&= \Phi(2,24) - \Phi(-0,45) = \Phi(2,24) - [1 - \Phi(0,45)] = 0,9875 - [1 - 0,6700] \\
&= tab(2,24) + tab(0,45) = 0,4875 + 0,1700 \\
&= 0,6575
\end{aligned}$$

♦♦

Exemplo 7.21
Seja $X \sim N(5, 1)$*, calcule* $\text{P}(X > 7)$

Solução:

$$\begin{aligned}
\text{P}(X > 7) &= \text{P}\left(\frac{X-5}{1} > \frac{7-5}{1}\right) \\
&= \text{P}(Z > 2) \\
&= 1,0 - \Phi(2,0) = 1,0 - 0,97725 \\
&= 0,5 - tab(2,0) = 0,5 - 0,47725 \\
&= 0,02275
\end{aligned}$$

♦♦

Exemplo 7.22 A regra 68-95-99,7
Seja $X \sim N(\mu; \sigma^2)$*. Calcule* $\text{P}(\mu - k\sigma < X < \mu + k\sigma)$ *, para* $k = 1, 2, 3$*.*

Solução:
Note que essa probabilidade corresponde à probabilidade de X estar a uma distância de k desvios-padrão da média.

$$\begin{aligned}
\text{P}(\mu - k\sigma \le X \le \mu + k\sigma) &= \text{P}\left(\frac{\mu - k\sigma - \mu}{\sigma} \le \frac{X-\mu}{\sigma} \le \frac{\mu + k\sigma - \mu}{\sigma}\right) \\
&= \text{P}(-k \le Z \le k)
\end{aligned}$$

Note que chegamos a uma probabilidade que não depende de μ ou σ, ou seja, esse resultado vale qualquer que seja a distribuição normal.

- $k = 1$

$$\text{P}(\mu - \sigma \le X \le \mu + \sigma) = \text{P}(-1 \le Z \le 1) = 2 \times tab(1,0) = 2 \times 0,3414 = 0,6828$$

- $k = 2$

$$\text{P}(\mu - 2\sigma \le X \le \mu + 2\sigma) = \text{P}(-2 \le Z \le 2) = 2 \times tab(2,0) = 2 \times 0,4772 = 0,9544$$

- $k = 3$

$$\text{P}(\mu - 3\sigma \le X \le \mu + 3\sigma) = \text{P}(-3 \le Z \le 3) = 2 \times tab(3,0) = 2 \times 0,4987 = 0,9974$$

Essas probabilidades nos dizem que, para *qualquer* distribuição normal, 68,28% dos valores estão a um desvio-padrão da média, 95,44% estão a dois desvios-padrão e 99,73% dos valores estão a três desvios-padrão da média. Veja a Figura 7.39 para uma ilustração desses resultados.

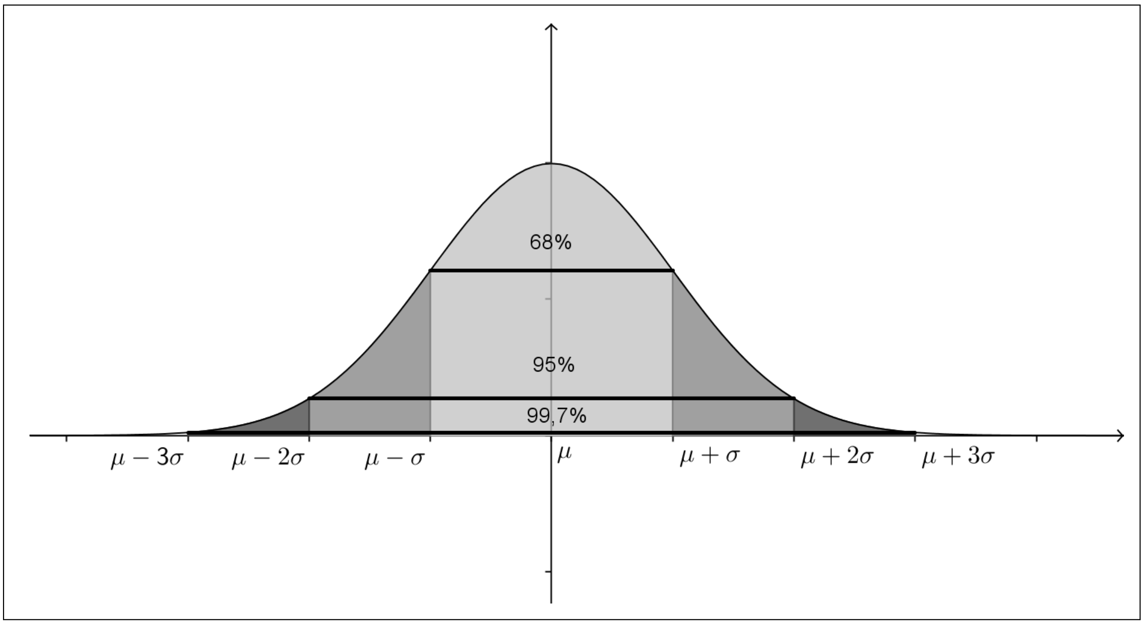


Figura 7.39 – Ilustração da regra 68-95-99,7

♦♦

7.5 Encontrando a Abscissa da Normal para uma Probabilidade Específica

Nos exemplos vistos até o momento, consideramos situações em que tínhamos uma abscissa de uma distribuição normal e queríamos a probabilidade associada a essa abscissa. Agora, vamos lidar com a situação inversa: dada uma probabilidade, qual é a abscissa correspondente? Eis algumas situações que envolvem esse tipo de problema:

- Em uma turma de Estatística, os 10% melhores alunos receberão um livro de presente.

- Em uma comunidade, as famílias com as 15% piores rendas irão receber um auxílio da prefeitura.

Como antes, vamos apresentar vários exemplos que ilustram essa situação.

Exemplo 7.23
Seja $Z \sim N(0; 1)$*, determine o valor de* k *tal que* $\mathrm{P}(Z \leq k) = 0,90$.

Solução:
Vamos "traduzir" esse problema: queremos encontrar a abscissa k da normal padrão com $0,90$ de área (probabilidade) à esquerda dela. Como 0,9 é a área à esquerda de k, resulta que k tem que ser maior que zero, isto é, temos que ter $k > 0$. Veja a Figura 7.40: á esquerda de k temos área 0,90 e à esquerda de 0 temos área 0,5. Logo, entre 0 e k temos que ter área 0,40.

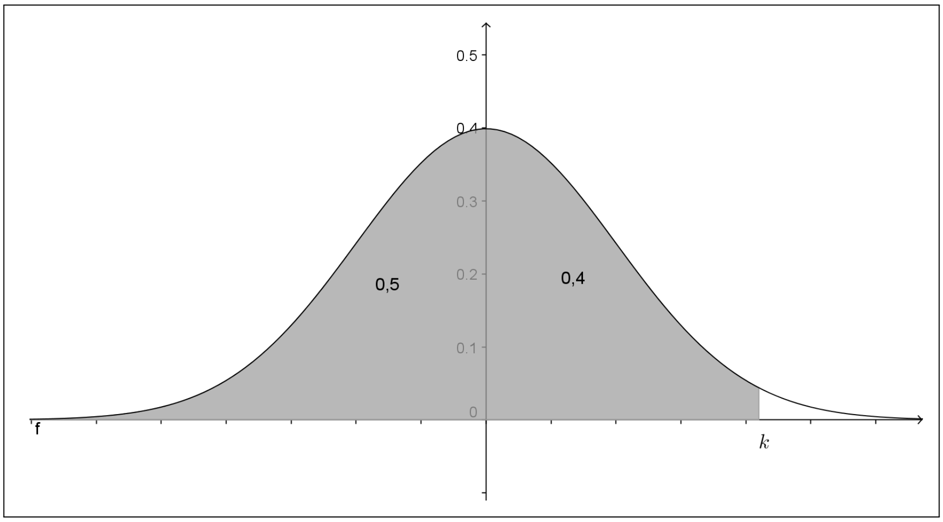


Figura 7.40 – Determinação de k tal que $P(Z \leq k) = 0,90$

Escrevendo essas observações em termos de probabilidade, temos:

$$
\begin{aligned}
&P(Z \leq k) = 0,90 \Leftrightarrow \\
&P(Z \leq 0) + P(0 < Z \leq k) = 0,90 \Leftrightarrow \\
&0,5 + P(0 < Z \leq k) = 0,90 \Leftrightarrow \\
&P(0 < Z \leq k) = 0,40 \Leftrightarrow \\
&\text{tab}(k) = 0,40
\end{aligned}
$$

Esta última igualdade nos diz que k é a abscissa correspondente ao valor 0,40 na Tabela 1. Para identificar k, temos que buscar no corpo dessa tabela, o valor mais próximo de 0,40. Na linha correspondente ao valor 1,2 encontramos as entradas 0,39973 e 0,40147. Como a primeira está mais próxima de 0,40, olhamos qual é a abscissa correspondente: a linha é 1,2 e a coluna é 8, o que nos dá a abscissa de 1,28, ou seja, $k = 1,28$ e $P(Z \leq 1,28) = 0,90$, completando a solução.

♦♦

Agora vamos olhar o mesmo exemplo, mas para uma distribuição normal qualquer.

Exemplo 7.24
Seja $X \sim N(3; 4)$, determine o valor de k tal que $P(X \leq k) = 0,90$.

Solução:
Com a probabilidade à esquerda de k é maior que 0,5, resulta que k tem de ser maior que a média. O primeiro passo na solução é escrever a probabilidade dada em termos da normal

padrão.

$$
\begin{aligned}
&\mathrm{P}(X \leq k) = 0,90 \Leftrightarrow \\
&\mathrm{P}\left(\frac{X-3}{2} \leq \frac{k-3}{2}\right) = 0,90 \Leftrightarrow \\
&\mathrm{P}\left(Z \leq \frac{k-3}{2}\right) = 0,90 \Leftrightarrow \\
&\mathrm{P}(Z \leq 0) + \mathrm{P}\left(0 \leq Z \leq \frac{k-3}{2}\right) = 0,90 \Leftrightarrow \\
&0,5 + \mathrm{P}\left(0 \leq Z \leq \frac{k-3}{2}\right) = 0,90 \Leftrightarrow \\
&\mathrm{P}\left(0 \leq Z \leq \frac{k-3}{2}\right) = 0,40 \Leftrightarrow \\
&tab\left(\frac{k-3}{2}\right) = 0,40 \Leftrightarrow \\
&\frac{k-3}{2} = 1,28 \Leftrightarrow k = 5,56
\end{aligned}
$$

♦♦

Exemplo 7.25
Seja $X \sim N(3;4)$, determine o valor de k tal que $\mathrm{P}(X \leq k) = 0,05$.

Solução:
À esquerda de k temos 5% da área total; logo, k tem que estar no lado esquerdo, ou seja, temos de ter $k < 3$ e a abscissa padronizada correspondente tem de ser negativa. Vamos escrever a probabilidade dada em termos da normal padronizada:

$$
\begin{aligned}
&\mathrm{P}(X \leq k) = 0,05 \Leftrightarrow \\
&\mathrm{P}\left(\frac{X-3}{2} \leq \frac{k-3}{2}\right) = 0,05 \Leftrightarrow \\
&\mathrm{P}\left(Z \leq \frac{k-3}{2}\right) = 0,05
\end{aligned}
$$

Como a área (probabilidade) à esquerda de $\frac{k-3}{2}$ é menor que $0,5$, isso significa que $\frac{k-3}{2}$ tem de ser negativo. Veja a Figura 7.41. Para nos adequarmos à tabela disponível, temos de rabalhar com áreas na metade direita da curva de densidade, ou seja, temos qde usar a simetria da curva. Veja a Figura 7.42 e note que a abscissa simétrica a $\dfrac{k-3}{2}$ é $-\dfrac{k-3}{2} = \dfrac{3-k}{2}$.

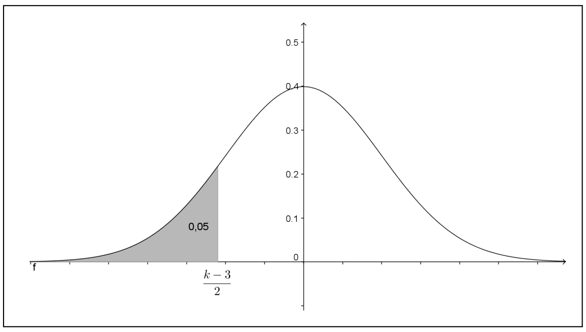


Figura 7.41 – $\mathrm{P}(Z \leq \frac{k-3}{2}) = 0,05$

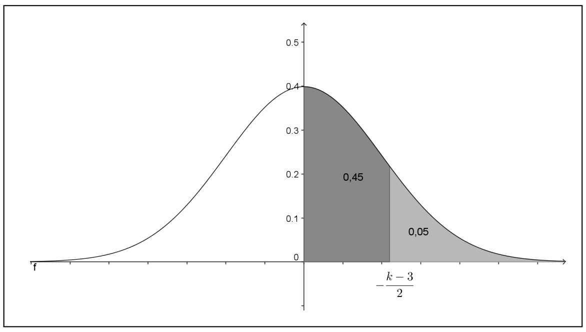


Figura 7.42 – $\mathrm{P}(Z \geq -\frac{k-3}{2}) = 0,05$

Temos, então, as seguintes probabilidades equivalentes:

$$
\begin{aligned}
&\mathrm{P}\left(Z \leq \frac{k-3}{2}\right) = 0,05 \Leftrightarrow \\
&\mathrm{P}\left(Z \geq -\frac{k-3}{2}\right) = 0,05 \Leftrightarrow \\
&\mathrm{P}\left(0 \leq Z \leq -\frac{k-3}{2}\right) = 0,45 \Leftrightarrow \\
&\mathrm{tab}\left(-\frac{k-3}{2}\right) = 0,45
\end{aligned}
$$

O valor mais próximo de 0,45 no corpo da Tabela 1 é 0,4495 que corresponde à abscissa 1,64, e isso nos dá que

$$-\frac{k-3}{2} = 1,64 \Rightarrow k = -0,28$$

♦♦

Exemplo 7.26
Seja $X \sim N(3;4)$, determine o valor de k tal que $\mathrm{P}(|X-3| \leq k) = 0,95$.

Solução:
Vamos relembrar as propriedades da função módulo. Para isso, veja a Figura 7.43. As duas retas que definem a função $f(x) = |x|$ são $y = x$ e $y = -x$. Os segmentos traçados no gráfico mostram que $|x| = k \Leftrightarrow x = k$ ou $x = -k$. Os valores de y abaixo do segmento horizontal correspondem a valores de $y = |x| \leq k$ e esses valores de y estão associados a valores x tais que $-k \leq x \leq k$. De forma análoga, valores de y acima do segmento horizontal correspondem a valores de $y = |x| > k$ e esses valores de y estão associados a valores x tais que $x > k$ ou $x < -k$. Resumindo:

$$|x| = k \Leftrightarrow \begin{cases} \text{x=k} \\ \text{ou} \\ \text{x=-k} \end{cases} \tag{7.18}$$

$$|x| < k \Leftrightarrow -k < x < k \tag{7.19}$$

$$|x| > k \Leftrightarrow \begin{cases} \text{x>k} \\ \text{ou} \\ \text{x<-k} \end{cases} \tag{7.20}$$

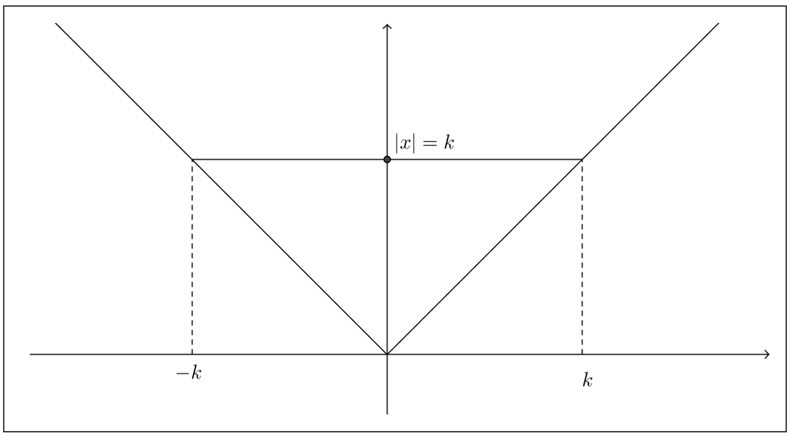


Figura 7.43 – Determinação de k tal que $\mathrm{P}(Z \leq k) = 0,90$

Agora, vamos usar esses resultados para resolver o exemplo.

$$
\begin{aligned}
&\mathrm{P}(|X-3| \leq k) = 0,95 \Leftrightarrow \\
&\mathrm{P}(-k \leq X-3 \leq k) = 0,95 \Leftrightarrow \\
&\mathrm{P}(3-k \leq X \leq k+3) = 0,95 \Leftrightarrow \\
&\mathrm{P}\left(\frac{3-k-3}{2} \leq \frac{X-3}{2} \leq \frac{k+3-3}{2}\right) = 0,95 \Leftrightarrow \\
&\mathrm{P}\left(-\frac{k}{2} \leq Z \leq \frac{k}{2}\right) = 0,95
\end{aligned}
$$

Veja a Figura 7.44 para entender que

$$
\begin{aligned}
&\mathrm{P}\left(-\frac{k}{2} \leq Z \leq \frac{k}{2}\right) = 0,95 \Leftrightarrow \\
&\mathrm{P}\left(-\frac{k}{2} \leq Z \leq 0\right) + \mathrm{P}\left(0 \leq Z \leq \frac{k}{2}\right) = 0,95 \Leftrightarrow \\
&2 \times \mathrm{P}\left(0 \leq Z \leq \frac{k}{2}\right) = 0,95 \Leftrightarrow \\
&\mathrm{P}\left(0 \leq Z \leq \frac{k}{2}\right) = 0,475 \Leftrightarrow \\
&\mathrm{tab}\left(\frac{k}{2}\right) = 0,475 \Leftrightarrow \\
&\frac{k}{2} = 1,96 \Leftrightarrow k = 3,92
\end{aligned}
$$

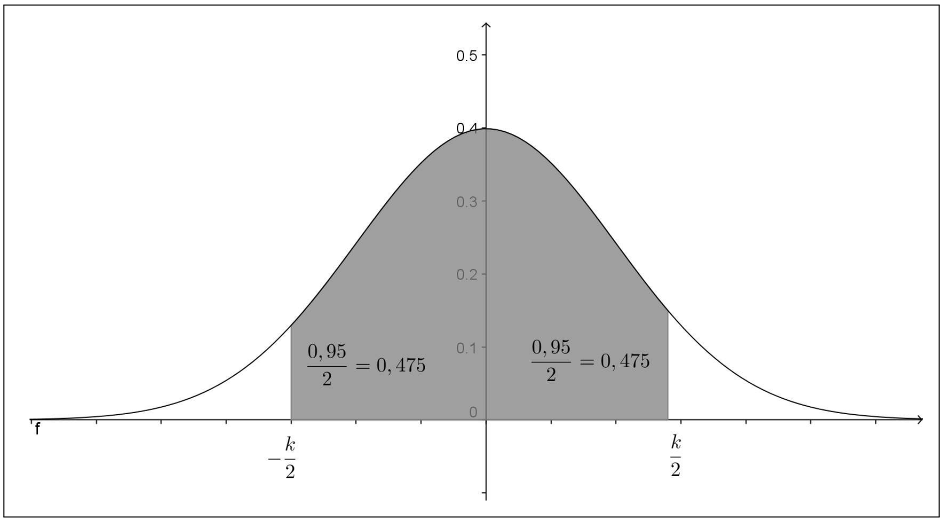


Figura 7.44 – Determinação de k tal que $\mathrm{P}(|ZZ \leq \frac{k}{2}) = 0,95$

♦♦

7.6 Exemplos de aplicação da distribuição Normal

A distribuição normal é um modelo probabilístico que se aplica a diversas situações práticas. Vamos finalizar este capítulo com alguns exemplos práticos, mas, na última parte do curso, você verá mais aplicações no contexto da inferência estatística, em que decisões têm de ser tomadas com base nos resultados obtidos a partir de uma amostra.

Exemplo 7.27 Saldo bancário
O saldo médio dos clientes de um banco é uma v.a. normal com média R\$ 2.000, 00 e desvio-padrão R\$ 250,00. Os clientes com os 10% maiores saldos médios recebem tratamento VIP, enquanto aqueles com os 5% menores saldos médios receberão propaganda extra para estimular maior movimentação da conta.

(a) *Quanto você precisa de saldo médio para se tornar um cliente VIP?*

(b) *Abaixo de qual saldo médio o cliente receberá a propaganda extra?*

Solução:
Seja X = "saldo médio"; é dado que $X \sim N(2000; 250^2)$.

(a) Temos que determinar o valor de k tal que $\mathrm{P}(X \geq k) = 0,10$. Note que isso equivale a calcular o 90º percentil da distribuição. A área à esquerda de k tem de ser 0,90; logo, k tem de ser maior que a média.

$$\begin{aligned}
\mathrm{P}(X \geq k) &= 0,10 \Leftrightarrow \\
\mathrm{P}\left(\frac{X-2000}{250} \geq \frac{k-2000}{250}\right) &= 0,10 \Leftrightarrow \\
\mathrm{P}\left(\frac{X-2000}{250} \leq \frac{k-2000}{250}\right) &= 0,90 \Leftrightarrow \\
\mathrm{P}\left(Z \leq \frac{k-2000}{250}\right) &= 0,90 \Leftrightarrow \\
\mathrm{P}(Z \leq 0) + \mathrm{P}\left(0 \leq Z \leq \frac{k-2000}{250}\right) &= 0,90 \Leftrightarrow \\
\mathrm{P}\left(0 \leq Z \leq \frac{k-2000}{250}\right) &= 0,90 - 0,50 \Leftrightarrow \\
\mathrm{tab}\left(\frac{k-2000}{250}\right) &= 0,40 \Leftrightarrow \\
\frac{k-2000}{250} &= 1,28 \Leftrightarrow k = 2320
\end{aligned}$$

Os clientes com saldo médio maior ou igual a R\$ 2.320, 00 terão tratamento VIP.

(b) Temos de determinar o valor de k tal que $\mathrm{P}(X \leq k) = 0,05$. Note que isso equivale a calcular o 5º percentil da distribuição. A área à esquerda de k tem de ser 0,05; logo, k tem de ser menor que a média. Na solução, teremos que usar a simetria da distribuição, invertendo o sinal da abscissa, para lidarmos com área na metade direita da função de densidade.

$$
\begin{aligned}
\mathrm{P}(X \leq k) = 0,05 &\Leftrightarrow \\
\mathrm{P}\left(\frac{X-2000}{250} \leq \frac{k-2000}{250}\right) = 0,05 &\Leftrightarrow \\
\mathrm{P}\left(Z \geq -\frac{k-2000}{250}\right) = 0,05 &\Leftrightarrow \\
\mathrm{P}\left(Z \geq \frac{2000-k}{250}\right) = 0,05 &\Leftrightarrow \\
\mathrm{tab}\left(\frac{2000-k}{250}\right) = 0,45 &\Leftrightarrow \\
\frac{2000-k}{250} = 1,64 \Leftrightarrow k = 1590 &
\end{aligned}
$$

Os clientes com saldo médio inferior a R$ 1.590,00 receberão a propaganda extra.

Na Figura 7.45 ilustra-se a solução do exercício.

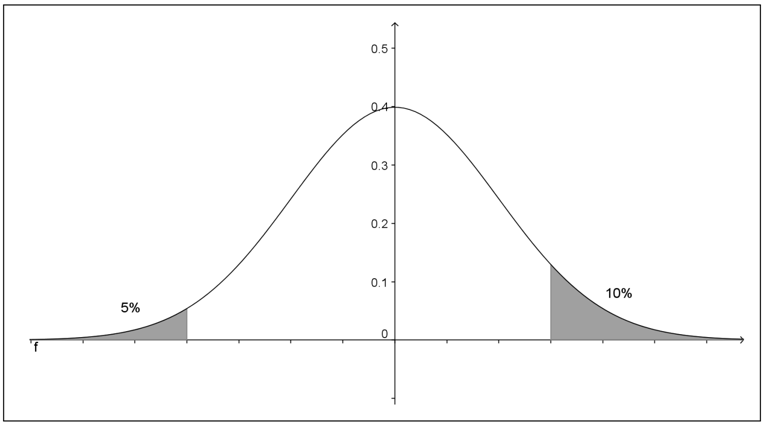


Figura 7.45 – Solução do Exemplo 7.27

♦♦

Exemplo 7.28 Regulagem de máquinas
Uma máquina de empacotar determinado produto oferece variações de peso que se distribuem segundo uma distribuição normal com desvio padrão de 20 gramas. Em quanto deve ser regulado o peso médio desses pacotes para que apenas 10% deles tenham menos que 500 gramas?

Solução:
Esse é um exemplo clássico de aplicação da distribuição normal. Seja X o peso dos pacotes em gramas. Então, $X \sim N(\mu; 400)$. Temos de ter $\mathrm{P}(X \leq 500) = 0,10$. Note que o peso médio

tem de ser superior a 500 g. Note, na solução, a inversão do sinal da abscissa!

$$
\begin{aligned}
\mathrm{P}(X \leq 500) = 0,10 &\Leftrightarrow \\
\mathrm{P}\left(\frac{X-\mu}{20} \leq \frac{500-\mu}{20}\right) = 0,10 &\Leftrightarrow \\
\mathrm{P}\left(Z \leq \frac{500-\mu}{20}\right) = 0,10 &\Leftrightarrow \\
\mathrm{P}\left(Z \geq -\frac{500-\mu}{20}\right) = 0,10 &\Leftrightarrow \\
\mathrm{P}\left(Z \geq \frac{\mu-500}{20}\right) = 0,10 &\Leftrightarrow \\
\mathrm{tab}\left(\frac{\mu-500}{20}\right) = 0,40 &\Leftrightarrow \\
\frac{\mu-500}{20} = 1,28 \Leftrightarrow \mu = 525,6 &
\end{aligned}
$$

A máquina tem de ser regulada com um peso médio de 525,6g para que apenas 10% dos pacotes tenham peso inferior a 500g. Veja a Figura 7.46.

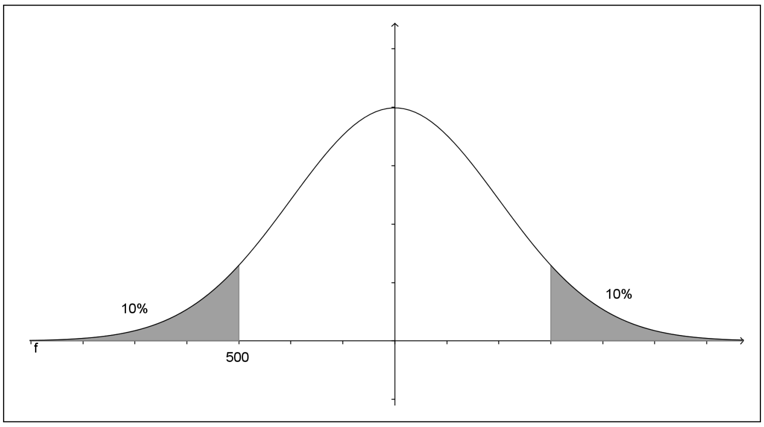


Figura 7.46 – Solução do Exemplo 7.28

♦♦

Exemplo 7.29 Mais sobre regulagem de máquinas
Uma máquina fabrica tubos metálicos cujos diâmetros podem ser considerados uma variável aleatória normal com média 200mm e desvio-padrão 2mm. Verifica-se que 15% dos tubos estão sendo rejeitados como grandes e 10% como pequenos.

(a) Quais são as tolerâncias de especificação para esse diâmetro?

(b) Mantidas essas especificações, qual deverá ser a regulagem média da máquina para que a rejeição por diâmetro grande seja praticamente nula? Nesse caso, qual será a porcentagem de rejeição por diâmetro pequeno?

Solução:
Seja D = diâmetro dos tubos. Então $D \sim N(200, 2^2)$.

(a) Sejam k_I e k_S as especificações inferior e superior, respectivamente. Isso significa que tubos com diâmetro menor que k_I são rejeitados como pequenos e tubos com diâmetro maior que k_S são rejeitados como grandes.

$$
\begin{aligned}
\mathrm{P}(D < k_I) = 0,10 &\Rightarrow \\
\mathrm{P}\left(\frac{D-200}{2} < \frac{k_I - 200}{2}\right) = 0,10 &\Rightarrow \\
\mathrm{P}\left(Z < \frac{k_I - 200}{2}\right) = 0,10 &\Rightarrow \\
\mathrm{P}\left(Z > -\frac{k_I - 200}{2}\right) = 0,10 &\Rightarrow \\
\mathrm{P}\left(Z > \frac{200 - k_I}{2}\right) = 0,10 &\Rightarrow \\
\mathrm{P}\left(0 \le Z < \frac{200 - k_I}{2}\right) = 0,40 &\Rightarrow \\
\mathrm{tab}\left(\frac{200 - k_I}{2}\right) = 0,40 &\Rightarrow \\
\frac{200 - k_I}{2} = 1,28 \Rightarrow k_I = 197,44
\end{aligned}
$$

$$
\begin{aligned}
\mathrm{P}(D > k_S) = 0,15 &\Rightarrow \\
\mathrm{P}\left(\frac{D-200}{2} > \frac{k_S - 200}{2}\right) = 0,15 &\Rightarrow \\
\mathrm{P}\left(0 \le Z < \frac{k_S - 200}{2}\right) = 0,35 &\Rightarrow \\
\mathrm{tab}\left(\frac{k_S - 200}{2}\right) = 0,35 &\Rightarrow \\
\frac{k_S - 200}{2} = 1,03 \Rightarrow k_S = 202,06
\end{aligned}
$$

Logo, tubos com diâmetro menor que 197,44 cm são rejeitados como pequenos e tubos com diâmetros maiores que 202,06 cm são rejeitados como grandes.

(b) Com a nova regulagem, temos que $D \sim N(\mu; 2^2)$ e μ deve ser tal que

$$
\begin{aligned}
\mathrm{P}(D > 202,06) = 0 &\Rightarrow \\
\mathrm{P}\left(\frac{D-\mu}{2} > \frac{202,06 - \mu}{2}\right) = 0 &\Rightarrow \\
\mathrm{P}\left(Z > \frac{202,06 - \mu}{2}\right) = 0 &\Rightarrow \\
\mathrm{P}\left(0 \le Z \le \frac{202,06 - \mu}{2}\right) = 0,5 &\Rightarrow \\
\mathrm{tab}\left(\frac{202,06 - \mu}{2}\right) = 0,5 &\Rightarrow \\
\frac{202,06 - \mu}{2} \simeq 4,5 \Rightarrow \mu \simeq 193,06
\end{aligned}
$$

Com essa média, a porcentagem de rejeição por diâmetro pequeno é

$$\begin{aligned}
P(D < 197,44) &= P\left(\frac{D - 193,06}{2} < \frac{197,44 - 193,06}{2}\right) \\
&= P(Z < 2,19) \\
&= P(Z \le 0) + P(0 < Z < 2,19) \\
&= 0,5 + tab(2,19) = 0,9857
\end{aligned}$$

Com essa nova regulagem, a rejeição por diâmetro grande é nula, mas a rejeição por diâmetro pequeno é muito alta! Veja as Figuras 7.47 e 7.48, nas quais ficam claros os resultados obtidos.

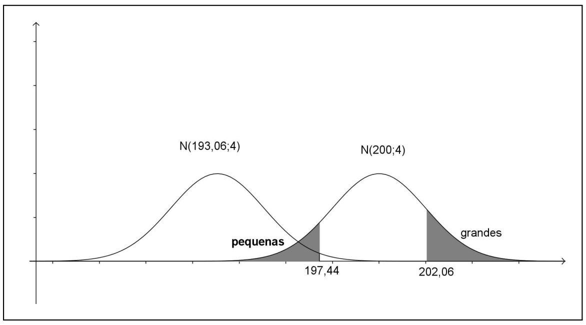


Figura 7.47 – Exemplo 7.29 - regulagem original

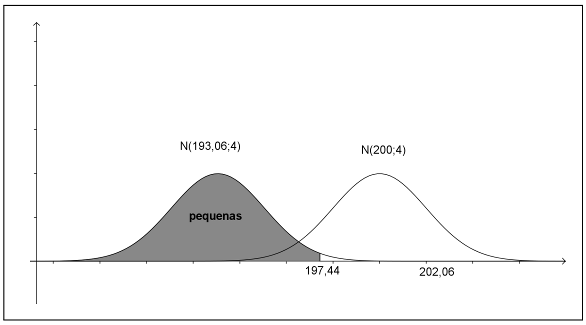


Figura 7.48 – Exemplo 7.29 - regulagem com 0% de tubos grandes

♦♦

Exemplo 7.30 Troca de lâmpadas

Em um grande complexo industrial, o departamento de manutenção tem instruções para substituir as lâmpadas antes que se queimem. Os registros indicam que a duração das lâmpadas, em horas, tem distribuição normal, com média de 900 horas e desvio-padrão de 75 horas. Quando devem ser trocadas as lâmpadas, de modo que no máximo 5% delas queimem antes de serem trocadas?

Solução:

Seja T = "tempo de duração (em horas) das lâmpadas"; então, $T \sim N(900; 75^2)$. Temos que determinar t tal que $P(T \le t) = 0,05$.

$$\begin{aligned}
P(T \le t) = 0,05 &\Leftrightarrow \\
P\left(\frac{T - 900}{75} \le \frac{t - 900}{75}\right) = 0,05 &\Leftrightarrow \\
P\left(Z \ge -\frac{t - 900}{75}\right) = 0,05 &\Leftrightarrow \\
P\left(Z \ge \frac{900 - t}{75}\right) = 0,05 &\Leftrightarrow \\
P\left(0 \le Z \le \frac{900 - t}{75}\right) = 0,45 &\Leftrightarrow \\
tab\left(\frac{900 - t}{75}\right) = 0,45 &\Leftrightarrow \\
\frac{900 - t}{75} = 1,64 &\Leftrightarrow t = 777
\end{aligned}$$

As lâmpadas devem ser trocadas com 777 horas de uso para que apenas 5% se queimem antes da troca.

♦♦

Aqui cabe a seguinte observação: em geral, não é apropriado utilizar-se a distribuição normal para modelar o tempo de sobrevivência de lâmpadas ou equipamentos em geral. Modelos tipo exponencial são mais adequados, pois atribuem probabilidade alta de sobrevivência no início da vida do equipamento e probabilidade decrescente à medida que o equipamento envelhece.

Exemplo 7.31 Regulagem de máquinas – controle da variabilidade

Uma enchedora automática enche garrafas de acordo com uma distribuição normal de média 1.000 ml. Deseja-se que no máximo uma garrafa em cada 100 saia com menos de 990ml. Qual deve ser o maior desvio padrão tolerável?

Solução:

Seja X = "conteúdo da garrafa (em ml)", então $X \sim N(1000; \sigma^2)$.

Queremos que $P(X < 990) \leq 0,01$.

Seja σ_0 o valor do desvio padrão de X tal que $P(X < 990) = 0,01$. Então, qualquer valor de σ tal que $\sigma < \sigma_0$ resulta em $P(X < 990) < 0,01$. Veja a Figura 7.49.

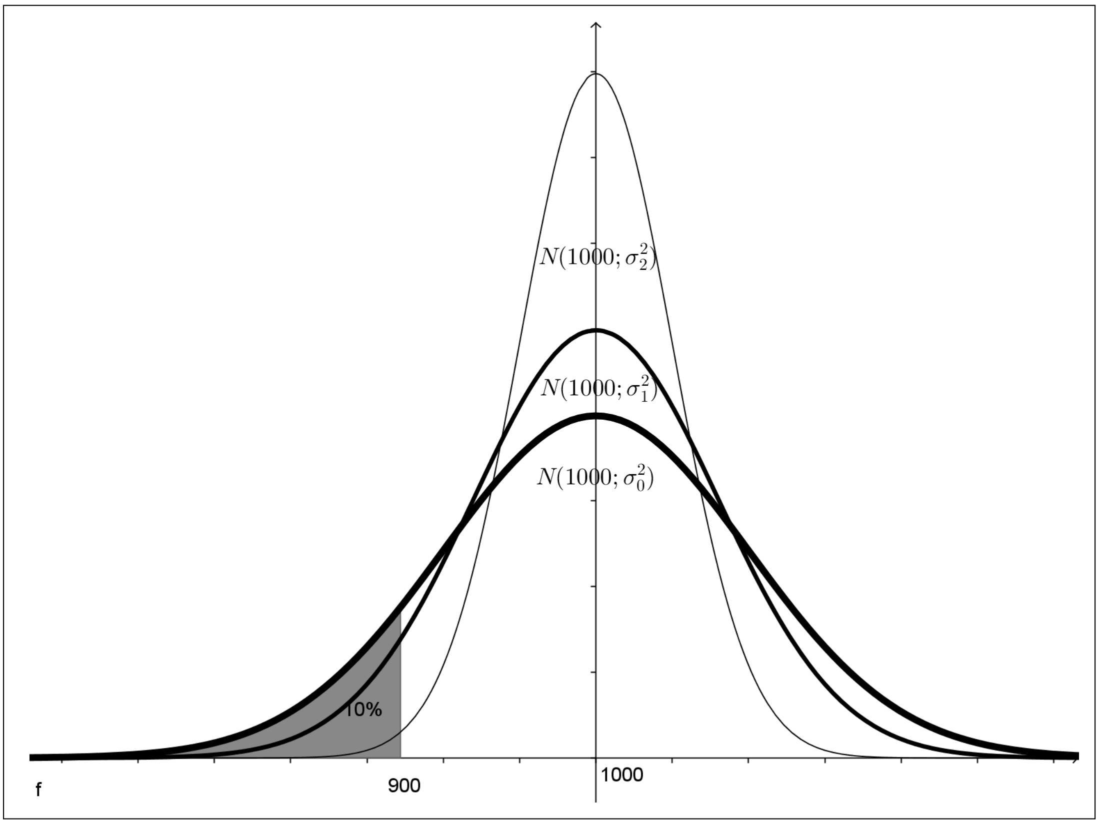


Figura 7.49 – Solução do Exemplo 7.31

A área sombreada corresponde a $P(X < 990) = 0,10$ quando $X \sim N(1000; \sigma_0^2)$ (curva de densidade mais espessa). As duas outras densidades correspondem a distribuições normais com desvios-padrão menores. Note que para essas distribuições, $P(X < 990) < 0,01$. Assim, o desvio-padrão máximo tolerável é tal que:

$$
\begin{aligned}
\mathrm{P}(X < 990) \leq 0,01 &\Leftrightarrow \\
\mathrm{P}\left(Z < \frac{990-1000}{\sigma}\right) \leq 0,01 &\Leftrightarrow \\
\mathrm{P}\left(Z > -\frac{990-1000}{\sigma}\right) \leq 0,01 &\Leftrightarrow \\
\mathrm{P}\left(Z > \frac{10}{\sigma}\right) \leq 0,01 &\Leftrightarrow \\
0,5 - tab\left(Z > \frac{10}{\sigma}\right) \leq 0,01 &\Leftrightarrow \\
tab\left(\frac{10}{\sigma}\right) \geq 0,49 &\Leftrightarrow \\
\frac{10}{\sigma} \geq 2,33 \Leftrightarrow \sigma \leq \frac{10}{2,33} &= 4,2918
\end{aligned}
$$

♦♦

7.7 A distribuição log-normal

7.7.1 Definição

Definição 7.32 Distribuição log-normal
Seja $X \sim N(\mu, \sigma^2)$. Se $Y = e^X$ então a variável aleatória Y tem distribuição log-normal com parâmetros μ e σ^2. Reciprocamente, se Y tem distribuição log-normal, então $X = \ln Y$ tem distribuição $N(\mu, \sigma^2)$.

Vamos calcular a função densidade de probabilidade de uma variável aleatória log-normal a partir de sua função de distribuição acumulada. Note que Y só pode assumir valores positivos. Temos que:

$$
\begin{aligned}
F_Y(y) &= \mathrm{P}(Y \leq y) = \mathrm{P}\left(e^X \leq y\right) = \mathrm{P}(X \leq \ln y) = \\
&= \mathrm{P}\left(\frac{X-\mu}{\sigma} \leq \frac{\ln y - \mu}{\sigma}\right) = \mathrm{P}\left(Z \leq \frac{\ln y - \mu}{\sigma}\right) \\
&= \Phi\left(\frac{\ln y - \mu}{\sigma}\right) \qquad y > 0
\end{aligned}
$$

Sabemos que $f_Y(y) = F'(y)$ e, também, no caso da normal padrão, $\Phi'(z) = \varphi(z)$. Logo, pela regra da cadeia,

$$
\begin{aligned}
f_Y(y) &= \Phi'\left(\frac{\ln y - \mu}{\sigma}\right) \times \frac{1}{\sigma y} = \varphi\left(\frac{\ln y - \mu}{\sigma}\right) \times \frac{1}{\sigma y} = \\
&= \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{\ln y - \mu}{\sigma}\right)^2\right] \times \frac{1}{\sigma y}
\end{aligned}
$$

ou ainda:

$$
f_Y(y) = \frac{1}{y\sqrt{2\pi\sigma^2}} \exp\left[-\frac{1}{2}\left(\frac{\ln y - \mu}{\sigma}\right)^2\right] \qquad y > 0
$$

7.7.2 Esperança

A esperança de Y é:

$$\mathrm{E}(Y) = \int_0^\infty y \frac{1}{y\sqrt{2\pi\sigma^2}} \exp\left[-\frac{1}{2}\left(\frac{\ln y - \mu}{\sigma}\right)^2\right] dy$$

Fazendo a mudança de variável $t = \ln y$ temos que

$$\begin{aligned}
dt &= \frac{1}{y} dy \\
y &= e^t \\
y &= 0 \Rightarrow t = -\infty \\
y &= \infty \Rightarrow t = \infty
\end{aligned}$$

e, portanto

$$\begin{aligned}
\mathrm{E}(Y) &= \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^\infty e^t \exp\left[-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2\right] dt = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^\infty \exp\left[-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2 + t\right] dt \\
&= \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^\infty \exp\left[-\frac{t^2 - 2t\mu + \mu^2 - 2\sigma^2 t}{2\sigma^2}\right] dt = \\
&= \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^\infty \exp\left[-\frac{t^2 - 2t\left(\mu + \sigma^2\right) + \mu^2}{2\sigma^2}\right] dt \\
&= \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^\infty \exp\left[-\frac{t^2 - 2t\left(\mu + \sigma^2\right)}{2\sigma^2}\right] \exp\left(-\frac{\mu^2}{2\sigma^2}\right) dt = \\
&= \exp\left(-\frac{\mu^2}{2\sigma^2}\right) \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^\infty \exp\left[-\frac{t^2 - 2t\left(\mu + \sigma^2\right) + \left(\mu + \sigma^2\right)^2 - \left(\mu + \sigma^2\right)^2}{2\sigma^2}\right] dt \\
&= \exp\left(-\frac{\mu^2}{2\sigma^2}\right) \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^\infty \exp\left\{-\frac{\left[t - \left(\mu + \sigma^2\right)\right]^2}{2\sigma^2}\right\} \exp\left(\frac{\left(\mu + \sigma^2\right)^2}{2\sigma^2}\right) dt \\
&= \exp\left(-\frac{\mu^2}{2\sigma^2} + \frac{\left(\mu + \sigma^2\right)^2}{2\sigma^2}\right) \left[\frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^\infty \exp\left\{-\frac{\left[t - \left(\mu + \sigma^2\right)\right]^2}{2\sigma^2}\right\} dt\right]
\end{aligned}$$

Mas o termo entre os colchetes externos é a integral de uma densidade normal com média $\lambda = \left(\mu + \sigma^2\right)$ e variância σ^2; logo, essa integral é 1 e, portanto:

$$\mathrm{E}(Y) = \exp\left(-\frac{\mu^2}{2\sigma^2} + \frac{\left(\mu + \sigma^2\right)^2}{2\sigma^2}\right) = \exp\left(\frac{-\mu^2 + \mu^2 + 2\mu\sigma^2 + \sigma^4}{2\sigma^2}\right) \Rightarrow$$

$$\mathrm{E}(Y) = \exp\left(\mu + \frac{\sigma^2}{2}\right) \tag{7.21}$$

7.7.3 Variância

Vamos calcular de modo análogo $\mathrm{E}(Y^2)$, usando a mesma transformação:

$$\begin{aligned}
\mathrm{E}(Y^2) &= \int_0^\infty y^2 \frac{1}{y\sqrt{2\pi\sigma^2}} \exp\left[-\frac{1}{2}\left(\frac{\ln y-\mu}{\sigma}\right)^2\right] dy \\
&= \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^\infty e^{2t} \exp\left[-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2\right] dt = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^\infty \exp\left[-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2 + 2t\right] dt \\
&= \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^\infty \exp\left[-\frac{t^2-2t\mu+\mu^2-4\sigma^2 t}{2\sigma^2}\right] dt = \\
&= \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^\infty \exp\left[-\frac{t^2-2t\left(\mu+2\sigma^2\right)+\mu^2}{2\sigma^2}\right] dt \\
&= \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^\infty \exp\left[-\frac{t^2-2t\left(\mu+2\sigma^2\right)}{2\sigma^2}\right] \exp\left(-\frac{\mu^2}{2\sigma^2}\right) dt = \\
&= \exp\left(-\frac{\mu^2}{2\sigma^2}\right) \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^\infty \exp\left[-\frac{t^2-2t\left(\mu+2\sigma^2\right)+\left(\mu+2\sigma^2\right)^2-\left(\mu+2\sigma^2\right)^2}{2\sigma^2}\right] dt \\
&= \exp\left(-\frac{\mu^2}{2\sigma^2}\right) \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^\infty \exp\left\{-\frac{\left[t-\left(\mu+2\sigma^2\right)\right]^2}{2\sigma^2}\right\} \exp\left(\frac{\left(\mu+2\sigma^2\right)^2}{2\sigma^2}\right) dt \\
&= \exp\left(-\frac{\mu^2}{2\sigma^2}+\frac{\left(\mu+2\sigma^2\right)^2}{2\sigma^2}\right) \left[\frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^\infty \exp\left\{-\frac{\left[t-\left(\mu+2\sigma^2\right)\right]^2}{2\sigma^2}\right\} dt\right]
\end{aligned}$$

Como antes, o termo entre os colchetes externos é 1 porque é a integral de uma densidade normal com média $\mu + 2\sigma^2$ e variância σ^2. Logo,

$$\begin{aligned}
\mathrm{E}(Y^2) &= \exp\left(-\frac{\mu^2}{2\sigma^2}+\frac{\left(\mu+2\sigma^2\right)^2}{2\sigma^2}\right) = \exp\left(\frac{-\mu^2+\mu^2+4\mu\sigma^2+4\sigma^4}{2\sigma^2}\right) \Rightarrow \\
&\mathrm{E}(Y^2) = \exp\left(2\mu+2\sigma^2\right)
\end{aligned}$$

e

$$\begin{aligned}
\mathrm{Var}(Y) &= \exp\left(2\mu+2\sigma^2\right) - \left[\exp\left(\mu+\frac{\sigma^2}{2}\right)\right]^2 = \\
&= \exp\left(2\mu+2\sigma^2\right) - \exp\left[2\left(\mu+\frac{\sigma^2}{2}\right)\right] = \\
&= \exp\left(2\mu+2\sigma^2\right) - \exp\left(2\mu+\sigma^2\right) \\
&= \exp\left(2\mu+\sigma^2\right)\left[\frac{\exp\left(2\mu+2\sigma^2\right)}{\exp\left(2\mu+\sigma^2\right)} - 1\right] = \\
&= \exp\left(2\mu+\sigma^2\right)\left[\exp\left(2\mu+2\sigma^2-2\mu-\sigma^2\right) - 1\right] = \\
&= \exp\left(2\mu+\sigma^2\right)\left[\exp\sigma^2 - 1\right]
\end{aligned}$$

Definindo $m = \mathrm{E}(X) = \exp\left(\mu+\frac{\sigma^2}{2}\right)$, temos que $m^2 = \exp\left(2\mu+\sigma^2\right)$. Logo,

$$\mathrm{Var}(Y) = m^2\left[e^{\sigma^2}-1\right] \tag{7.22}$$

Capítulo 8

Momentos e sua Função Geradora

8.1 Momentos

Na definição de variância aparece a esperança do quadrado, seja dos desvios em torno da média, ou da própria variável. Esses são casos particulares de uma caracterítica mais geral de uma variável aleatória, definida a seguir.

Definição 8.1 Momentos
O $k-$ésimo ***momento*** *de uma variável aleatória X, denotado por μ'_k, é definido como*

$$\mu'_k = \mathrm{E}(X^k) \tag{8.1}$$

desde que essa quantidade exista. O $k-$ésimo ***momento central*** *de uma variável aleatória X, denotado por μ_k, é definido como*

$$\mu_k = \mathrm{E}[X - \mathrm{E}(X)]^k \tag{8.2}$$

sempre que essa quantidade existir.

Note que $\mathrm{E}(X) = \mu'_1$ e $\mathrm{Var}(X) = \mu_2 = \mu'_2 - (\mu'_1)^2$. Temos também que, para qualquer variável aleatória, $\mu_1 = 0$.

Exemplo 8.2 Momentos da distribuição gama
Calcule o k-ésimo momento de $X \sim Gama(\alpha, \beta)$.

Solução:

$$\begin{aligned}
\mathrm{E}(X^k) &= \int_0^\infty x^k \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x} dx \\
&= \int_0^\infty \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha+k-1} e^{-\beta x} dx \\
&= \frac{\Gamma(\alpha+k)}{\beta^{\alpha+k}} \frac{\beta^\alpha}{\Gamma(\alpha)} \int_0^\infty \frac{\beta^{\alpha+k}}{\Gamma(\alpha+k)} x^{\alpha+k-1} e^{-\beta x} dx \\
&= \frac{\Gamma(\alpha+k)}{\beta^{\alpha+k}} \frac{\beta^\alpha}{\Gamma(\alpha)} = \frac{\Gamma(\alpha+k)}{\beta^k \Gamma(\alpha)} = \frac{\alpha(\alpha+1)\cdots(\alpha+k-1)}{\beta^k}
\end{aligned}$$

♦♦

Exemplo 8.3 Momentos da distribuição de Weibull
Calcule o k-ésimo momento de $X \sim Weibull(\alpha, \beta)$.

Solução:

$$\begin{aligned}
\mathrm{E}(X^k) &= \int_{-\infty}^{\infty} x^k f_X(x)dx \\
&= \int_0^{+\infty} x^k \frac{\alpha}{\beta}\left(\frac{x}{\beta}\right)^{\alpha-1} e^{-\left(\frac{x}{\beta}\right)^\alpha} dx \\
&\underbrace{=}_{u=\left(\frac{x}{\beta}\right)^\alpha} \int_0^{+\infty} (\beta u^{1/\alpha})^k \frac{\alpha}{\beta}\left(u^{1/\alpha}\right)^{\alpha-1} e^{-u}\frac{\beta}{\alpha}\left(u^{1/\alpha-1}\right)du \\
&= \int_0^{+\infty} (\beta u^{1/\alpha})^k e^{-u} du \\
&= \beta^k \int_0^{+\infty} u^{(\frac{k}{\alpha}+1)-1} e^{-u} du \\
&= \beta^k \Gamma\left(\frac{k}{\alpha}+1\right) = \beta^k \Gamma\left(\frac{\alpha+k}{\alpha}\right)
\end{aligned}$$

♦♦

Proposição 8.4
Seja X variável aleatória com densidade simétrica em torno de μ. Então,

$$\mathrm{E}((X-\mu)^k) = 0$$

sempre que k for ímpar.

Demonstração:
Primeiro veja que se X tem densidade simétrica em torno de μ, então as variáveis aleatórias $Y = (X-\mu)$ e $W = -(X-\mu)$ tem mesma distribuição.

$F_Y(Y) = \mathrm{P}(Y \leq y) = \mathrm{P}(X-\mu \leq y) = \mathrm{P}(X \leq \mu + y) = F_X(\mu+y) \Rightarrow f_Y(y) = f_X(\mu+y)$.

$F_W(w) = \mathrm{P}(W \leq w) = \mathrm{P}(-(X-\mu) \leq w) = \mathrm{P}(X \geq \mu - w) = 1 - F_X(\mu-w) \Rightarrow f_W(w) = f_X(\mu-w)$.

Como f_X é simétrica em torno de μ podemos afirmar que $f_Y(a) = f_X(\mu+a) = f_X(\mu-a) = f_W(a)$, logo Y e W tem mesma distribuição.

Se $(X-\mu)$ e $-(X-\mu)$ tem mesma distribuição, então $(X-\mu)^k$ e $(-(X-\mu))^k$ também tem mesma distribuição. Logo,

$$\mathrm{E}\left((X-\mu)^k\right) = \mathrm{E}\left((-(X-\mu))^k\right).$$

Então, se k for ímpar,

$$\mathrm{E}\left((X-\mu)^k\right) = \mathrm{E}\left((-(X-\mu))^k\right) = -\mathrm{E}\left((X-\mu)^k\right) \Rightarrow \mathrm{E}\left((X-\mu)^k\right) = 0.$$

□

8.2 Alguns Coeficientes e suas Interpretações

Já vimos interpretações para o primeiro momento, $\mathrm{E}(X)$, e para o segundo momento central, $\mathrm{Var}(X)$. Dessa forma os dois primeiros momentos de uma variável aleatória X guardam informações sobre a localização e sobre a variação em torno da média dessa variável. Nessa seção veremos a definição de dois novos coeficientes, que dependem do 3º e do 4º momento central. Com eles veremos como esses dois outros momentos podem nos passar informações sobre a variável aleatória.

Definição 8.5 Coeficiente de Assimetria
Seja X variável aleatória com $\mathrm{E}(X) = \mu$ e $\mathrm{Var}(X) = \sigma^2$. O coeficiente de assimetria de X, denotado por α_3, é definido por

$$\alpha_3 = \frac{\mathrm{E}[(X-\mu)^3]}{\sigma^3}$$

supondo a existência dos momentos envolvidos.

O coeficiente de assimetria, além de quantificar o grau de assimetria de uma distribuição, indica também o tipo de assimetria. Nas Figuras 8.1 e 8.2 os histogramas representam uma distribuição assimétrica à direita e uma assimétrica à esquerda, respectivamente. Em termos do coeficiente de assimetria, temos a seguinte caracerização:

$$\begin{aligned}
\alpha_3 > 0 &\quad \Rightarrow \quad \text{assimetria positiva ou à direita} \\
\alpha_3 = 0 &\quad \Rightarrow \quad \text{assimetria nula ou simetria} \\
\alpha_3 < 0 &\quad \Rightarrow \quad \text{assimetria negativa ou à esquerda}
\end{aligned}$$

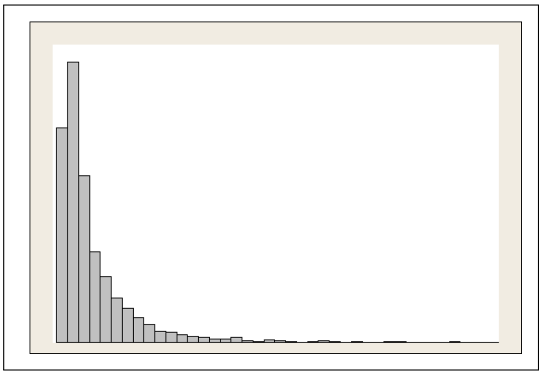


Figura 8.1 – Assimetria à direita

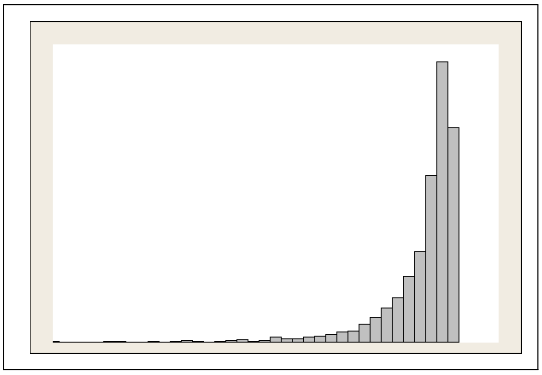


Figura 8.2 – Assimetria à esquerda

Já vimos que se X é variável aleatória com densidade simétrica em torno de μ, todos os momentos centrais ímpares são nulos. Em particular, também será nulo o 3º momento central. Logo, $\alpha_3 = 0$, indicando que a variável aleatória é simétrica, como esperávamos.

Vejamos agora a definição de outro coeficiente, que depende do 4º momento central.

Definição 8.6 Coeficiente de Curtose
Seja X variável aleatória com $\mathrm{E}(X) = \mu$ e $\mathrm{Var}(X) = \sigma^2$. O coeficiente de curtose de X, denotado por α_4, é definido por

$$\alpha_4 = \frac{\mathrm{E}[(X-\mu)^4]}{\sigma^4} - 3$$

supondo a existência dos momentos envolvidos.

O coeficiente de curtose mede afastamentos em relação à distribuição normal, tanto em termos de pico, quanto em termos da espessura das caudas da distribuição. A subtração de 3 no coeficiente tem a ver com o fato de o coeficiente de curtose da distribuição normal ser 3; assim, um coeficiente 0 indica total semelhança com a distribuição normal. Coeficientes positivos indicam distribuições leptocúrticas, que têm picos mais acentuados que a normal e coeficientes negativos indicam uma distribuição platicúrtica, que é mais "achatada" que a normal. Nas Figuras 8.3 a 8.6 ilustram algumas dessas distribuições. [1]

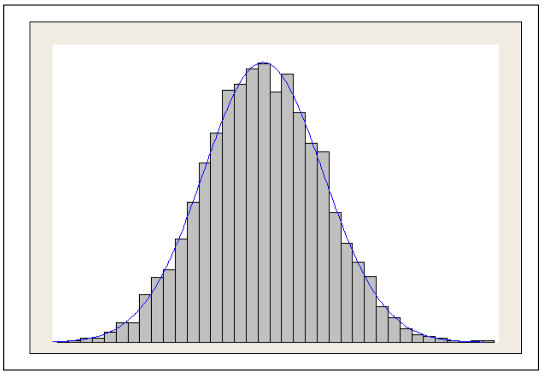


Figura 8.3 – Mesocúrtica - $\alpha_4 = 0.0$

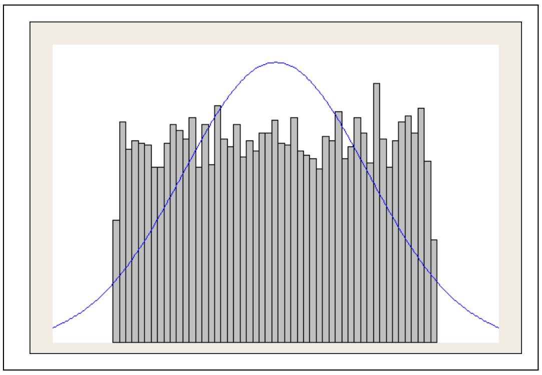


Figura 8.4 – Platicúrtica - $\alpha_4 = -1.2$

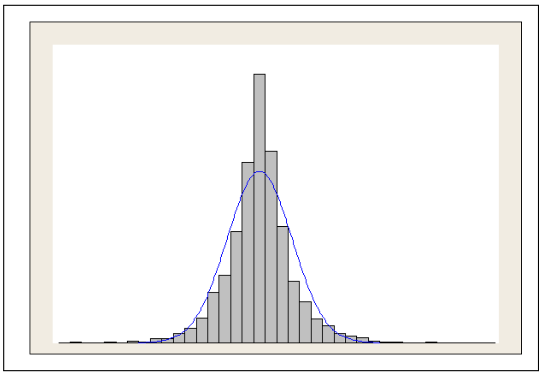


Figura 8.5 – Leptocúrtica - $\alpha_4 = 3.0$

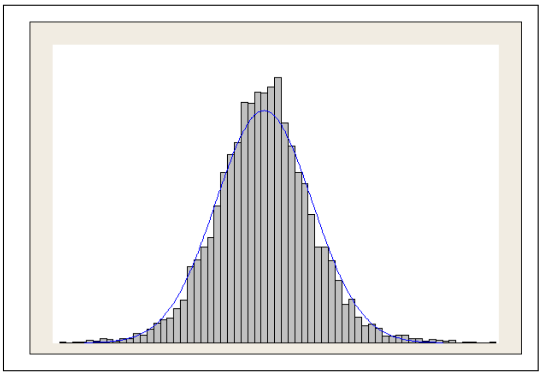


Figura 8.6 – Leptocúrtica - $\alpha_4 = 1.2$

8.3 Função Geradora de Momentos

Definição 8.7 Função Geradora de Momentos
*Seja X uma variável aleatória qualquer. A **função geradora de momentos** (fgm) de X, denotada por M_X, é definida por*

$$M_X(t) = \mathrm{E}\left(e^{tX}\right),$$

desde que essa esperança exista para todo t em alguma vizinhança de 0.

Importante: a função geradora de momentos é função de t. Para ela existir basta que exista $\epsilon > 0$ tal que $\mathrm{E}\left(e^{tX}\right)$ esteja bem definida para qualquer $t \in (-\epsilon, \epsilon)$.

[1] Os histogramas foram gerados a partir de 5000 observações simuladas das seguintes distribuições: $t(200)$, $Uni(0,1)$, Laplace e logística, ambas com parâmetro de escla igual a 1 e parâmetro de localização 0. todas com parâmetro de escala 1.

Veja que a definição de função geradora de momentos é feita independente do tipo de variável, mas a forma de encontrá-la depende se X é discreta ou contínua. Isto é,

$$\text{Se } X \text{ é discreta } \Rightarrow M_X(t) = \mathrm{E}\left(e^{tX}\right) = \sum_{\forall x \in Im(X)} e^{tx} p_X(x).$$

$$\text{Se } X \text{ é contínua } \Rightarrow M_X(t) = \mathrm{E}\left(e^{tX}\right) = \int_{-\infty}^{\infty} e^{tx} f_X(x) dx.$$

Lembrando da Definição 8.1, do momento de uma variável aleatória, temos o Teorema 8.8 a seguir, que justifica o nome da função M_X.

Teorema 8.8 *Seja X variável aleatória com função geradora de momentos M_X , então*

$$\mathrm{E}\left(X^k\right) = \left.\frac{d^k}{dt^k} M_x(t)\right|_{t=0}$$

ou seja, o k-ésimo momento de X é igual à derivada de ordem k de M_X avaliada em $t = 0$.

A demonstração do Teorema 8.8 acima será omitida pois precisamos de resultados ainda não vistos nesse curso. Mas, a título de curiosidade, ela pode ser encontrada no Teorema 5.10 de Magalhães (2011).

Antes de seguir com os exemplos veja que, qualquer que seja a variável aleatória X,

$$M_X(0) = \mathrm{E}(e^{0X}) = 1.$$

Esse resultado pode ser útil para verificarmos se a função geradora de momentos encontrada está correta.

Exemplo 8.9 Distribuição de Bernoulli
Seja $X \sim Bern(p)$. Encontre a sua função geradora de momentos e a partir dela calcule $\mathrm{E}(X)$ *e* $\mathrm{Var}(X)$.

Solução:
Por definição $M_X(t) = \mathrm{E}(e^{tX})$. Se $X \sim Bern(p)$,

$$M_X(t) = \sum_{x=0}^{1} e^{tx} p^x (1-p)^{1-x} = \sum_{x=0}^{1} \left(pe^t\right)^x (1-p)^{1-x} = 1 - p + pe^t$$

Logo,

$$\text{Se } X \sim Bern(p) \qquad \Rightarrow \qquad M_X(t) = 1 - p + pe^t. \tag{8.3}$$

Veja que $M_X(0) = 1$. Calculando as derivadas primeira e segunda, obtemos que

$$M'_X(t) = M''_X(t) = pe^t$$

e, portanto

$$\begin{aligned} \mathrm{E}(X) &= M'_X(0) = p \\ \mathrm{E}(X^2) &= M''_X(0) = p \\ \mathrm{Var}(X) &= \mathrm{E}(X^2) - [\mathrm{E}(X)]^2 = p - p^2 = p(1-p) \end{aligned}$$

♦♦

Exemplo 8.10 Distribuição Binomial
Seja $X \sim B(n,p)$. Encontre a sua função geradora de momentos e a partir dela calcule $\mathrm{E}(X)$ *e* $\mathrm{Var}(X)$.

Solução:
Por definição,

$$\begin{aligned} M_X(t) &= \mathrm{E}(e^{tX}) = \sum_{k=0}^{n} e^{tk}\,\mathrm{P}(X=k) \\ &= \sum_{k=0}^{n} e^{tk}\tbinom{n}{k}p^k(1-p)^{n-k} = \sum_{k=0}^{n} \tbinom{n}{k}\left(pe^t\right)^k(1-p)^{n-k} \\ &= \left(pe^t + (1-p)\right)^n \end{aligned}$$

pela fórmula do binômio de Newton.

Logo,

$$\text{Se } X \sim B(n,p) \qquad \Rightarrow \qquad M_X(t) = \left(pe^t + (1-p)\right)^n. \tag{8.4}$$

Veja que $M_X(0) = 1$. Para encontrar os dois primeiros momentos precisamos das duas primeiras derivadas.

$$\begin{aligned} M_X'(t) &= npe^t\left[pe^t + (1-p)\right]^{n-1} \\ M_X''(t) &= npe^t\left[pe^t + (1-p)\right]^{n-1} + n(n-1)pe^t\left[pe^t + (1-p)\right]^{n-2}pe^t \\ &= npe^t\left[pe^t + (1-p)\right]^{n-1} + n(n-1)p^2e^{2t}\left[pe^t + (1-p)\right]^{n-2} \end{aligned}$$

e, portanto

$$\mathrm{E}(X) = M_X'(0) = np\,[p + (1-p)]^{n-1} = np$$

$$\begin{aligned} \mathrm{E}(X^2) &= M_X''(0) = np\,[p+(1-p)]^{n-1} + n(n-1)p^2\,[p+(1-p)]^{n-2} \\ &= np + n(n-1)p^2 = np + n^2p^2 - np^2 \end{aligned}$$

$$\mathrm{Var}(X) = \mathrm{E}(X^2) - [\mathrm{E}(X)]^2 = np + n^2p^2 - np^2 - n^2p^2 = np - np^2 = np(1-p)$$

♦♦

Exemplo 8.11 Distribuição Geométrica
Seja $X \sim geo(p)$. Encontre a sua função geradora de momentos e a partir dela calcule $\mathrm{E}(X)$ *e* $\mathrm{Var}(X)$.

Solução:
Por definição,

$$\begin{aligned} M_X(t) &= \mathrm{E}(e^{tX}) = \sum_{k=1}^{\infty} e^{tk}p(1-p)^{k-1} = \sum_{j=0}^{\infty} e^{t(j+1)}p(1-p)^j = pe^t\sum_{j=0}^{\infty} e^{tj}(1-p)^j \\ &= pe^t\sum_{j=0}^{\infty}\left[e^t(1-p)\right]^j \end{aligned}$$

Então, se $e^t(1-p) < 1$, isto é, para $t < \ln\left(\frac{1}{1-p}\right)$ temos que

$$M_X(t) = pe^t\frac{1}{1-e^t(1-p)} = \frac{pe^t}{1-e^t(1-p)}.$$

Logo,

$$\text{Se } X \sim geo(p) \qquad \Rightarrow \qquad M_X(t) = \frac{pe^t}{1 - e^t(1-p)} \;, \quad t < \ln\left(\frac{1}{1-p}\right). \tag{8.5}$$

Veja que $M_X(0) = 1$. Calculando as derivadas primeira e segunda, temos que

$$M_X'(t) = \frac{pe^t\left[1 - e^t(1-p)\right] - pe^t\left[-e^t(1-p)\right]}{\left[1 - e^t(1-p)\right]^2} = \frac{pe^t - pe^{2t}(1-p) + pe^{2t}(1-p)}{\left[1 - e^t(1-p)\right]^2} = \frac{pe^t}{\left[1 - e^t(1-p)\right]^2}$$

$$\begin{aligned} M_X''(t) &= \frac{pe^t\left[1 - e^t(1-p)\right]^2 - 2pe^t\left[1 - e^t(1-p)\right]\left[-e^t(1-p)\right]}{\left[1 - e^t(1-p)\right]^4} \\ &= \frac{pe^t\left[1 - e^t(1-p)\right]^2 + 2pe^{2t}\left[1 - e^t(1-p)\right](1-p)}{\left[1 - e^t(1-p)\right]^4} \end{aligned}$$

Logo,

$$\mathrm{E}(X) = M_X'(0) = \frac{p}{\left[1 - (1-p)\right]^2} = \frac{p}{p^2} = \frac{1}{p}$$

e

$$\begin{aligned} \mathrm{E}(X^2) &= M_X''(0) = \frac{p\left[1 - (1-p)\right]^2 + 2p\left[1 - (1-p)\right](1-p)}{\left[1 - (1-p)\right]^4} \\ &= \frac{p^3 + 2p^2(1-p)}{p^4} = \frac{2p^2 - p^3}{p^4} = \frac{2-p}{p^2} \end{aligned}$$

e, portanto

$$\mathrm{Var}(X) = \frac{2-p}{p^2} - \frac{1}{p^2} = \frac{1-p}{p^2}$$

mesmos resultados obtidos anteriormente.

♦♦

Veja no Exemplo 8.11 que a função geradora de momentos de $X \sim geo(p)$ não está definida para todo $t \in \mathbb{R}$, mas está bem definida para $t < \ln(1/(1-p))$. E como $\ln(1/(1-p)) > 0$, a função geradora de momentos está bem definida para t em uma vizinhança de zero.

Proposição 8.12 Transformações Lineares
Seja X uma variável aleatória com função geradora de momentos M_X. Seja $Y = aX + b$. Então, a função geradora de momentos de Y é $M_Y(t) = e^{bt}M_X(at)$.

Demonstração:

$$M_Y(t) = \mathrm{E}(e^{tY}) = \mathrm{E}\left(e^{t(aX+b)}\right) = \mathrm{E}\left(e^{atX}e^{bt}\right) = e^{bt}\,\mathrm{E}\left(e^{atX}\right) = e^{bt}\,\mathrm{E}\left(e^{(at)X}\right) = e^{bt}M_X(at)$$

□

O resultado apresentado na Proposição 8.12 acima pode ser usado para encontrar a função geradora de momentos de uma variável aleatória definida como transformação linear de outra, da qual já conhecemos a função geradora de momentos. Veja uma aplicação desse resultado no Exemplo 8.13 a seguir.

Exemplo 8.13 Distribuição Geométrica Alternativa
Seja X a variável aleatória definida como o número de fracassos até a ocorrência do 1º sucesso em ensaios de Bernoulli independentes, todos com probabilidade de sucesso p. Veja que X segue a forma alternativa da distribuição Geométrica, definida na Seção 3.6. Encontre a sua função geradora de momentos de X.

Solução:
Para facilitar a resolução do exercício vamos definir X como transformação linear de uma geométrica e então usar a Proposição 8.12 e o resultado do Exemplo 8.11.

Podemos escrever $X = W - 1$, onde W é o número de tentativas até a ocorrência do 1º sucesso em ensaios de Bernoulli independentes e todos com probabilidade de sucesso p. Veja que $W \sim geo(p)$ e, de acordo com o Exercício 8.11,

$$M_W(t) = \frac{pe^t}{1 - e^t(1-p)} \ , \quad \text{para } t < \ln \frac{1}{1-p}.$$

Usando o resultado da Proposição 8.12, $M_X(t) = e^{-t} M_W(t)$. Logo,

$$M_X(t) = e^{-t} \frac{pe^t}{1 - e^t(1-p)} = \frac{p}{1 - e^t(1-p)} \ , \quad \text{para } t < \ln \frac{1}{1-p}.$$

♦♦

Exemplo 8.14 Distribuição de Poisson
Seja $X \sim Poisson(\lambda)$. Encontre a sua função geradora de momentos e a partir dela calcule $\mathrm{E}(X)$ *e* $\mathrm{Var}(X)$.

Solução:
Por definição,

$$M_X(t) = \mathrm{E}(e^{tX}) = \sum_{k=0}^{\infty} e^{tk} \frac{\lambda^k}{k!} e^{-\lambda} = e^{-\lambda} \sum_{k=0}^{\infty} e^{tk} \frac{\lambda^k}{k!} = e^{-\lambda} \sum_{k=0}^{\infty} \frac{\left(\lambda e^t\right)^k}{k!} = e^{-\lambda} e^{\lambda e^t}.$$

Logo,

$$\text{Se } X \sim Poisson(\lambda) \quad \Rightarrow \quad M_X(t) = e^{\lambda\left(e^t - 1\right)} \tag{8.6}$$

Veja que $M_X(0) = 1$. Vamos calcular as derivadas de ordem primeira e ordem segunda para calcular os respectivos momentos:

$$\begin{aligned} M_X'(t) &= e^{\lambda\left(e^t-1\right)} \lambda e^t = \lambda e^t e^{\lambda\left(e^t-1\right)} = \lambda e^{t+\lambda\left(e^t-1\right)} \\ M_X''(t) &= \lambda e^{t+\lambda\left(e^t-1\right)} \left(1 + \lambda e^t\right) \end{aligned}$$

Logo,

$$\begin{aligned} \mathrm{E}(X) &= M_X'(0) = \lambda e^{0+\lambda\left(e^0-1\right)} = \lambda \\ \mathrm{E}(X^2) &= M_X''(0) = \lambda e^{0+\lambda\left(e^0-1\right)} \left(1 + \lambda e^0\right) = \lambda(1+\lambda) = \lambda^2 + \lambda \end{aligned}$$

o que nos leva ao mesmo resultado anterior: $\mathrm{Var}(X) = \lambda$.

♦♦

Exemplo 8.15 Distribuição Exponencial
Seja $X \sim exp(\lambda)$. Encontre a sua função geradora de momentos e a partir dela calcule $\mathrm{E}(X)$ *e* $\mathrm{Var}(X)$.

Solução:
Por definição

$$M_X(t) = \mathrm{E}(e^{tX}) = \int_0^{\infty} e^{tx} \lambda e^{-\lambda x} dx = \int_0^{\infty} \lambda e^{-(\lambda - t)x} dx = \lambda \left. \frac{e^{-(\lambda-t)x}}{-(\lambda - t)} \right|_0^{\infty}.$$

Para que esse limite seja real, isto é, a integral exista, é necessário que $\lambda - t > 0$, ou seja, $t < \lambda$. Sob essa condição podemos seguir e concluir que

$$M_X(t) = \lambda \left.\frac{e^{-(\lambda-t)x}}{-(\lambda-t)}\right|_0^\infty = 0 - \lambda\frac{1}{-(\lambda-t)} = \frac{\lambda}{\lambda-t}.$$

Logo,

$$\text{Se } X \sim exp(\lambda) \qquad \Rightarrow \qquad M_X(t) = \frac{\lambda}{\lambda-t}\,, \qquad t < \lambda. \tag{8.7}$$

Veja que $M_X(0) = 1$. As derivadas primeira e segunda são

$$M_X'(t) = \frac{\lambda}{(\lambda-t)^2} \qquad \text{e} \qquad M_X''(t) = 2(\lambda-t)\frac{\lambda}{(\lambda-t)^4} = \frac{2\lambda}{(\lambda-t)^3}$$

e os momentos de ordem 1 e 2 são:

$$\mathrm{E}(X) = M_X'(0) = \frac{1}{\lambda} \qquad \text{e} \qquad \mathrm{E}(X^2) = M_X''(0) = \frac{2}{\lambda^2}.$$

o que nos dá, como antes, $\mathrm{Var}(X) = \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}$.

♦♦

Exemplo 8.16 Distribuição Gama
Seja $X \sim Gama(\alpha, \lambda)$. Encontre a sua função geradora de momentos.

Solução:
Por definição

$$M_X(t) = \mathrm{E}(e^{tX}) = \frac{\lambda^\alpha}{\Gamma(\alpha)}\int_0^\infty e^{tx}x^{\alpha-1}e^{-\lambda x}dx = \frac{\lambda^\alpha}{\Gamma(\alpha)}\underbrace{\int_0^\infty x^{\alpha-1}e^{-(\lambda-t)x}dx}.$$

Veja que a integral destacada com as chaves pode ser considerara uma gama modificada somente se $\lambda - t > 0$, ou seja, $t < \lambda$. Sob essa condição podemos seguir e concluir que

$$M_X(t) = \frac{\lambda^\alpha}{\Gamma(\alpha)}\frac{\Gamma(\alpha)}{(\lambda-t)^\alpha} = \frac{\lambda^\alpha}{(\lambda-t)^\alpha}.$$

Logo,

$$\text{Se } X \sim Gama(\alpha,\lambda) \qquad \Rightarrow \qquad M_X(t) = \frac{\lambda^\alpha}{(\lambda-t)^\alpha}\,, \quad t < \lambda. \tag{8.8}$$

Veja que $M_X(0) = 1$. Veja também que quando $\alpha = 1$ o resultado da Equações 8.8 coincide com a Equação 8.7, que apresenta a função geradora de momentos da exponencial.

♦♦

Veja que assim como a função geradora de momentos da geométrica, no caso da exponencial e da gama a função geradora de momentos não está definida para todo $t \in \mathbb{R}$, mas para $t < \lambda$. Logo, ela está bem definida para t em uma vizinhança de zero.

Para encontrar a função geradora de momentos de uma variável aleatória Normal vamos primeiro encontrar a função geradora de momentos da Normal Padrão. Em seguida vamos usar a Proposição 8.12 para encontrar a função geradora de momentos de uma normal qualquer. Veja os Exemplos 8.17 e 8.18 a seguir.

Exemplo 8.17 Distribuição Normal Padrão
Seja $X \sim N(0,1)$. Encontre a sua função geradora de momentos.

Solução:
Por definição,

$$\begin{aligned}
M_Z(t) &= \int_{-\infty}^{\infty} e^{tx} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left[-\left(\frac{x^2-2tx}{2}\right)\right] dx \\
&= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left[-\left(\frac{x^2-2tx+t^2-t^2}{2}\right)\right] dx \\
&= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left[-\left(\frac{x^2-2tx+t^2}{2}\right)+\left(\frac{t^2}{2}\right)\right] dx \\
&= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left[-\left(\frac{x^2-2tx+t^2}{2}\right)\right] e\left(\frac{t^2}{2}\right) dx \\
&= \exp\left(\frac{t^2}{2}\right) \underbrace{\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{(x-t)^2}{2}\right] dx}_{1}
\end{aligned}$$

Logo,

$$\text{Se } Z \sim N(0,1) \qquad \Rightarrow \qquad M_Z(t) = \exp\left(\frac{t^2}{2}\right). \tag{8.9}$$

Veja que $M_X(0) = 1$.

♦♦

Exemplo 8.18 Distribuição Normal
Seja $X \sim N(\mu, \sigma^2)$. Encontre a sua função geradora de momentos e a partir dela calcule: $\mathrm{E}(X)$, $\mathrm{Var}(X)$, *α_3 e α_4.*

Solução:
Se $X \sim N(\mu;\sigma^2)$ então $X = \sigma Z + \mu$, onde $Z \sim N(0;1)$ *com* função geradora de momentos dada por $M_Z(t) = e^{t^2/2}$.

Como X é uma transformação linear de Z, resulta da Proposição 8.12 que a função geradora de momentos de X é dada por

$$M_X(t) = e^{\mu t} M_Z(\sigma t) = e^{\mu t} e^{\sigma^2 t^2/2}.$$

Logo,

$$\text{Se } X \sim N(\mu,\sigma^2) \qquad \Rightarrow \qquad M_X(t) = \exp\left(\mu t + \frac{\sigma^2 t^2}{2}\right). \tag{8.10}$$

Veja que $M_X(0) = 1$.

- Derivada primeira (primeiro momento – esperança)

$$M_X'(t) = M_X(t)\left(\mu + t\sigma^2\right)$$

e, portanto,

$$\mathrm{E}(X) = M_X'(0) = M_X(0)\,(\mu) = \mu$$

- Derivada segunda (segundo momento – variância)

$$M''_X(t) = M'_X(t)\left(\mu + t\sigma^2\right) + \sigma^2 M_X(t)$$

Logo,

$$\mathrm{E}(X^2) = M''_X(0) = M'_X(0)\,(\mu) + \sigma^2 M_X(0) = \mu^2 + \sigma^2$$

e, portanto

$$\mathrm{Var}(X) = \mathrm{E}(X^2) - [\mathrm{E}(X)]^2 = \mu^2 + \sigma^2 - \mu^2 = \sigma^2$$

- Derivada terceira (terceiro momento – assimetria)

$$M'''_X(t) = M''_X(t)\left(\mu + t\sigma^2\right) + M'_X(t)\sigma^2 + M'_X(t)\sigma^2 = M''_X(t)\left(\mu + t\sigma^2\right) + 2M'_X(t)\sigma^2$$

Logo,

$$\mathrm{E}(X^3) = M'''_X(0) = \mu M''_X(0) + 2M'_X(0)\sigma^2 = \mu(\mu^2 + \sigma^2) + 2\mu\sigma^2 = \mu^3 + 3\mu\sigma^2$$

e, portanto, o coeficiente de assimetria é

$$\alpha_3 = \frac{\mathrm{E}(X-\mu)^3}{\sigma^3} = \frac{\mathrm{E}(X^3 - 3X^2\mu + 3X\mu^2 - \mu^3)}{\sigma^3} = \frac{\mu^3 + 3\mu\sigma^2 - 3\mu(\mu^2+\sigma^2) + 3\mu^2\mu - \mu^3}{\sigma^3} = 0$$

- Derivada quarta (quarto momento – curtose)

$$M^{(iv)}_X(t) = M'''_X(t)\left(\mu + t\sigma^2\right) + M''_X(t)\sigma^2 + 2M''_X(t)\sigma^2 = M'''_X(t)\left(\mu + t\sigma^2\right) + 3M''_X(t)\sigma^2$$

Logo,

$$\mathrm{E}(X^4) = M^{(iv)}_X(0) = M'''_X(0)\mu + 3M''_X(0)\sigma^2 = (\mu^3 + 3\mu\sigma^2)\mu + 3\sigma^2(\mu^2+\sigma^2) = \mu^4 + 6\mu^2\sigma^2 + 3\sigma^4$$

e, portanto, o coeficiente de curtose é

$$\begin{aligned}
\alpha_4 &= \frac{\mathrm{E}(X-\mu)^4}{\sigma^4} - 3 = \frac{\mathrm{E}(X^4 - 4X^3\mu + 6X^2\mu^2 - 4X\mu^3 + \mu^4)}{\sigma^4} - 3 \\
&= \frac{\mu^4 + 6\mu^2\sigma^2 + 3\sigma^4 - 4\mu(\mu^3 + 3\mu\sigma^2) + 6\mu^2(\mu^2+\sigma^2) - 4\mu^3\mu + \mu^4}{\sigma^4} - 3 \\
&= 3\frac{\sigma^4}{\sigma^4} - 3 = 0
\end{aligned}$$

♦♦

Além do fato de ser possível encontrar todos os momentos de uma variável aleatória a partir da sua função geradora de momentos, essa função tem mais uma característica importante: ela também define por completo a distribuição da variável aleatória. Veja Teorema 8.19 a seguir.

Teorema 8.19 *Se duas variáveis aleatórias têm funções geradoras de momentos que existem, e são iguais, então elas têm a mesma função de distribuição.*

A demonstração do Teorema 8.19 será omitida, pois ela usa conceitos ainda não estudados neste curso. Veja que o Teorema 8.19 nos mostra que se duas variáveis aleatórias tem mesma função geradora de momentos então estas variáveis aleatórias são identicamente distribuídas. Isso significa que podemos identificar a distribuição de uma variável aleatória a partir da sua função geradora de momentos, assim como identificamos a distribuição da variável aleatória a partir da sua função de distribuição, função densidade ou função de probabilidade.

Veja uma aplicação do Teorema 8.19 no Exemplo 8.20 a seguir. Nele é apresentada uma alternativa para a demonstração da Proposição 6.3.

Exemplo 8.20
Seja $X \sim gama(\alpha, \lambda)$, $Y = cx$ e $c \in \mathbb{R}$. Mostre, a partir da função geradora de momentos, que $Y \sim gama(\alpha, \lambda/c)$.

Solução:
Como $X \sim gama(\alpha, \lambda)$ já sabemos que a sua função geradora de momentos é definida por:

$$M_X(t) = \frac{\lambda^\alpha}{(\lambda - t)^\alpha}\ , \quad t < \lambda.$$

Vamos, a partir de M_X, encontrar a função geradora de momentos de $Y = cX$.

$$M_Y(t) = \mathrm{E}(e^{tY}) = \mathrm{E}(e^{tcX}) = M_X(tc) = \frac{\lambda^\alpha}{(\lambda - tc)^\alpha}\ , \quad tc < \lambda$$

Ou seja,

$$M_Y(t) = \frac{(\lambda/c)^\alpha}{(\lambda/c - t)^\alpha}\ , \quad t < \lambda/c.$$

Veja que trata-se da função geradora de momentos de uma variável aleatória com distribuição $gama(\alpha, \lambda/c)$. Logo, $Y \sim gama(\alpha, \lambda/c)$.

♦♦

Vamos usar o Teorema 8.19 para mostrar que a transformação linear de uma variável aleatória normal também é variável aleatória normal. Veja a Proposição 8.21 a seguir.

Proposição 8.21 Transformação Linear de $N(\mu, \sigma)$
Seja $X \sim N(\mu, \sigma)$ e $Y = aX + b$, com $a, b \in \mathbb{R}$. Então, $Y \sim N(a\mu + b, a^2\sigma^2)$.

Demonstração:
Como $X \sim N(\mu, \sigma)$ sabemos, pela Equação 8.10, que

$$M_X(t) = \exp\left(\mu t + \frac{\sigma^2 t^2}{2}\right).$$

Usando o resultado da Proposição 8.12, podemos concluir que

$$\begin{aligned}
M_Y(t) &= e^{tb} M_X(ta) = e^{tb} \exp\left(\mu ta + \frac{\sigma^2 (ta)^2}{2}\right) \\
&= \exp\left(tb + \mu ta + \frac{\sigma^2 (ta)^2}{2}\right) \\
&= \exp\left((b + a\mu)t + \frac{a^2\sigma^2 t^2}{2}\right)
\end{aligned}$$

Veja que trata-se da função geradora de momentos de uma variável aleatória com distribuição Normal de média $b + a\mu$ e variância $a^2\sigma^2$. Logo, $Y \sim N(b + a\mu, a^2\sigma^2)$.

□

8.4 Algumas Desigualdades Importantes

Se conhecemos a função distribuição de uma variável aleatória, sabemos calcular qualquer probabilidade envolvendo essa varável aleatória. Se conhecemos a função geradora de momentos, sabemos calcular qualquer momento de uma variável aleatória (e podemos inclusive identificar a sua distribuição). Mas se conhecemos apenas alguns momentos de X não conseguimos encontrar a sua distribuição e nem a sua função geradora de momentos. Consequentemente, não sabemos calcular probabilidades envolvendo X. Apesar disso, o conhecimento do 1º e do 2º momento já nos dá alguma informação sobre o comportamento de X, como veremos nessa seção.

Proposição 8.22 Desigualdade de Markov (ou Básica de Chebyshev)
Seja X uma variável aleatória tal que $\mathrm{P}(X \geq 0) = 1$, *ou seja, X assume apenas valores não negativos. Então,* $\forall\ t > 0$,

$$\mathrm{P}(X \geq t) \leq \frac{\mathrm{E}(X)}{t}.$$

Demonstração:
Primeiro vamos mostrar que a desigualdade vale quando X é v.a. discreta.

$$\begin{aligned}
\mathrm{E}(X) &= \sum_x x p_X(x) = \underbrace{\sum_{x<t} x p_X(x)}_{\geq 0,\ \text{pois } X \geq 0} + \sum_{x \geq t} x p_X(x) \\
&\geq \sum_{x \geq t} x p_X(x) \geq \sum_{x \geq t} t p_X(x) = t \sum_{x \geq t} p_X(x) = t\,\mathrm{P}(X \geq t).
\end{aligned}$$

A demonstração é análoga quando X é v.a. contínua . Veja,

$$\begin{aligned}
\mathrm{E}(X) &= \int x f_X(x) dx = \underbrace{\int_{x<t} x f_X(x) dx}_{\geq 0,\ \text{pois } X \geq 0} + \int_{x \geq t} x f_X(x) dx \\
&\geq \int_{x \geq t} x f_X(x) dx \geq \int_{x \geq t} t f_X(x) dx = t \int_{x \geq t} f_X(x) dx = t\,\mathrm{P}(X \geq t).
\end{aligned}$$

□

Exemplo 8.23 (Magalhães (2011) , Exemplo 4.27)
Numa empresa com 100 funcionários o número médio de usuários simultâneos de Internet, num certo período do dia, é de aproximadamente 30. Atualmente, existem 30 linhas telefônicas disponíveis para conexão. Avalie a necessidade de aumentar esse número.

Solução:
Seja X = número de usuários simultâneos na Internet. Veja que se temos 30 linhas disponíveis, a probabilidade de um usuário ficará sem Internet é $\mathrm{P}(X > 30)$, que não sabemos calcular. Se esse valor for grande, o aumento no número de linhas disponíveis deve ser adotado.

Apesar de não sabermos calcular $\mathrm{P}(X > 30)$ podemos tentar usar a Desigualdade de Markov para encontrar uma cota superior. Veja que X é variável aleatória não negativa, além disso sabemos que $\mathrm{E}(X) = 30$. Então, usando a desigualdade,

$$\mathrm{P}(X > 30) = \mathrm{P}(X \geq 31) \leq \mathrm{E}(X)/31 = 0,9677.$$

Nesse caso a desigualdade não ajudou muito. Mas se aumentarmos o número de linhas disponíveis para 40 a probabilidade de um usuário ficar sem Internet será $\mathrm{P}(X > 40)$. Mesmo sem conhecer a distribuição de X podemos usar a Desigualdade de Markov e garantir que

$$\mathrm{P}(X > 40) = \mathrm{P}(X \geq 41) \leq \mathrm{E}(X)/41 = 0,7317.$$

Veja que já temos alguma informação! Vamos calcular as cotas superiores para a probabilidade de um usuário ficar sem Internet para outros números de linhas disponíveis:

Nº de linhas disponíveis	Desigualdade de Markov	Cota superior
30	$P(X \geq 31) \leq 0,9677$	97%
40	$P(X \geq 41) \leq 0,7317$	74%
50	$P(X \geq 51) \leq 0,5882$	59%
60	$P(X \geq 61) \leq 0,4918$	50%
70	$P(X \geq 71) \leq 0,4225$	43%
80	$P(X \geq 81) \leq 0,3704$	38%

Baseado nessas informações, você faria algum investimento para aumentar a quantidade de linhas disponíveis?

♦♦

Proposição 8.24 Desigualdade Clássica de Chebyshev
Seja X variável aleatória qualquer tal que $E(X) = \mu$ *e* $Var(X) = \sigma^2$ *existem e são finitas. Então,*

$$P(|X - \mu| \geq t) \leq \frac{\sigma^2}{t^2}.$$

Demonstração:
Primeiro veja que

$$P(|X - \mu| \geq t) = P(|X - \mu|^2 \geq t^2),$$

pois a função $f(x) = x^2$ é estritamente crescente para $x > 0$. Além disso, $|X - \mu|^2$ é variável aleatória não negativa, então podemos aplicar a Desigualdade de Markov e concluir que

$$P(|X - \mu|^2 \geq t^2) \leq \frac{E(|X - \mu|^2}{t^2} = \frac{\sigma^2}{t^2}.$$

Juntando,

$$P(|X - \mu| \geq t) \leq \frac{\sigma^2}{t^2}.$$

□

Corolário 8.25 *Seja X variável aleatória qualquer tal que* $E(X) = \mu$ *e* $Var(X) = \sigma^2$ *existem e são finitas. Então,*

$$P(|X - \mu| < t) \geq 1 - \frac{\sigma^2}{t^2}.$$

Demonstração:
Primeiro veja que

$$P(|X - \mu| \geq t) \leq \frac{\sigma^2}{t^2} \Rightarrow -P(|X - \mu| \geq t) \geq -\frac{\sigma^2}{t^2}.$$

Com isso podemos concluir que,

$$P(|X - \mu| < t) = 1 - P(|X - \mu| \geq t) \geq 1 - \frac{\sigma^2}{t^2}.$$

□

Exemplo 8.26
Continuando o Exemplo 8.23, sabendo que o desvio padrão no número de usuários simultâneos na Internet é 10, melhore as aproximações encontradas no Exemplo 8.23.

Solução:
Nesse caso temos mais uma informações, o desvio padrão. Com isso podemos usar a Desigualdade Clássica de Chebyshev.

Veja que a probabilidade de um usuário ficar sem Internet se temos k linhas disponíveis é

$$P(X > k) = P(X \geq k+1) = P(X-30 \geq k-29) \leq P(|X-30| \geq k-29) \leq \frac{\sigma^2}{(k-29)^2} = \frac{100}{(k-29)^2}.$$

Logo,

$$P(X \geq k+1) \leq \frac{100}{(k-29)^2}.$$

Com isso podemos recalcular as cotas superiores.

Nº de linhas disponíveis	Desigualdade	Cota superior
30	$P(X \geq 31) \leq 25$	100%
40	$P(X \geq 41) \leq 0,8264$	83%
50	$P(X \geq 51) \leq 0,2267$	23%
60	$P(X \geq 61) \leq 0,1040$	11%
70	$P(X \geq 71) \leq 0,0595$	6%
80	$P(X \geq 81) \leq 0,0384$	4%

♦♦

Exemplo 8.27
Seja X é uma v.a. com média $\mu < \infty$ e variância $\sigma^2 < \infty$. Encontre uma cota superior para $P(|X-\mu| \geq 2\sigma)$ (ou uma cota inferior para $P(|X-\mu| \leq 2\sigma)$). Encontre também uma cota superior para $P(|X-\mu| \geq 3\sigma)$ (ou uma cota inferior para $P(|X-\mu| \leq 3\sigma)$).

Solução:
Se X é uma v.a. com média $\mu < \infty$ e variância $\sigma^2 < \infty$, então

- $$P(|X-\mu| \geq 2\sigma) \leq \frac{1}{2^2} = 0,25 \Rightarrow P(|X-\mu| \leq 2\sigma) \geq 0,75$$

  ou seja, para qualquer v.a. com média e variância finitas, a porcentagem de valores a uma distância de 2 desvios-padrão da média é de pelo menos 75%.

- $$P(|X-\mu| \geq 3\sigma) \leq \frac{1}{3^2} = 0,1111 \Rightarrow P(|X-\mu| \leq 3\sigma) \geq 0,8889$$

  ou seja, para qualquer v.a. com média e variância finitas, a porcentagem de valores a uma distância de 3 desvios-padrão da média é de pelo menos 88,89%.

Compare esses valores com aqueles obtidos no Exemplo 7.22 e ilustrados na Figura 7.39 para uma distribuição normal.

♦♦

Apêndice A

Alguns Resultados de Cálculo

A.1 Séries geométricas

Convergência

Vamos estudar a convergência da série geométrica

$$\sum_{i=0}^{\infty} ar^i \tag{A.1}$$

considerando os possíveis valores da razão r.

- $r = 1$

  Nesse caso, a série é $a + a + a + \cdots$ e a $n-$ésima soma parcial é

  $$s_n = a + a + \cdots + a = (n+1)a$$

  para a qual temos $\lim\limits_{n\to\infty} s_n = \lim\limits_{n\to\infty}(n+1)a = \pm a$ dependendo do sinal de a. Resulta, então, que a série geométrica (A.1) diverge.

- $r = -1$

  Nesse caso, a série é $a - a + a - a + a - a + \cdots$ e a sequência das somas parciais é $a, 0, a, 0, a, \cdots$ que, obviamente, diverge. Logo, a série geométrica diverge se $r = -1$.

- $|r| \neq 1$

  A $n-$ésima soma parcial é

  $$s_n = a + ar + ar^2 + \ldots + ar^n$$

  Multiplicando ambos os lados por r, obtemos que

  $$rs_n = ar + ar^2 + ar^3 + \ldots + ar^n + ar^{n+1}$$

  e, portanto,

  $$s_n - rs_n = (a + ar + ar^2 + \ldots + ar^n) - (ar + ar^2 + ar^3 + \ldots + ar^n + ar^{n+1}) = a - ar^{n+1}$$

  Logo,

  $$s_n(1-r) = a(1-r^{n+1})$$

  Como $r \neq 1$,podemos escrever

  $$s_n = \frac{a(1-r^{n+1})}{1-r} \tag{A.2}$$

⋆ $|r| < 1$ $(-1 < r < 1)$

Nesse caso, $\lim_{n\to\infty} r^{n+1} = 0$ e $\lim_{n\to\infty} s_n = \frac{a}{1-r}$, ou seja,

$$\sum_{i=0}^{\infty} ar^i = \frac{a}{1-r} \qquad \text{se } |r| < 1. \tag{A.3}$$

⋆ $r > 1$

Nesse caso, $\lim_{n\to\infty} r^{n+1} = \infty$ o que implica que a sequência das somas parciais diverge, o mesmo ocorrendo com a série geométrica.

⋆ $r < -1$

Nesse caso, r^{n+1} oscila entre valores negativos e positivos de magnitude crescente, o que implica novamente que a sequência das somas parciais diverge, o mesmo ocorrendo com a série geométrica.

Resumindo, temos o seguinte resultado:

! Série geométrica

A série geométrica $\sum_{i=0}^{\infty} ar^i$ diverge se $|r| \geq 1$ e converge se $|r| < 1$. Temos também que

$$\sum_{i=0}^{\infty} ar^i = \frac{a}{1-r} \qquad \text{se } |r| < 1. \tag{A.4}$$

Note que podemos escrever (atenção aos índices dos somatórios!):

$$\sum_{i=0}^{\infty} ar^i = a + \sum_{i=1}^{\infty} ar^i \Longrightarrow \frac{a}{1-r} = a + \sum_{i=1}^{\infty} ar^i \quad \text{se } |r| < 1$$

ou

$$\sum_{i=1}^{\infty} ar^i = \frac{a}{1-r} - a = \frac{ar}{1-r} \qquad \text{se } |r| < 1 \tag{A.5}$$

Derivadas

Vamos considerar agora o caso particular em que $a = 1$.

- Derivada primeira

$$
\begin{aligned}
\frac{d}{dr}\left(\sum_{i=0}^{\infty} r^i\right) &= \frac{d}{dr}(1 + r + r^2 + r^3 + r^4 + \cdots + r^n + \cdots) \\
&= 1 + 2r + 3r^2 + 4r^3 + \cdots + nr^{n-1} + \cdots = \\
&= 1 + (r + r) + (r^2 + 2r^2) + (r^3 + 3r^3) + \cdots \left[r^{n-1} + (n-1)r^{n-1} + \cdots\right] \\
&= (1 + r + r^2 + r^3 + \cdots + r^{n-1} + \cdots) + \left[r + 2r^2 + 3r^3 + \cdots + (n-1)r^{n-1} + \cdots\right] \\
&= \sum_{i=0}^{\infty}(r^i + ir^i) \\
&= \sum_{i=0}^{\infty}(1 + i)r^i
\end{aligned}
$$

Como a série geométrica é convergente para $|\, r \,| < 1$, podemos igualar as derivadas, isto é:

$$
\sum_{i=0}^{\infty}(1+i)r^i = \frac{d}{dr}\left(\sum_{i=0}^{\infty} r^i\right) = \frac{d}{dr}\left(\frac{1}{1-r}\right) \qquad \text{se } |\, r \,| < 1
$$

ou seja,

$$
\sum_{i=0}^{\infty}(1+i)r^i = \frac{1}{(1-r)^2} \qquad \text{se } |\, r \,| < 1 \tag{A.6}
$$

Mas podemos escrever

$$
(1+i) = \frac{(1+i) \times i!}{i!} = \frac{(1+i)!}{i!\,1!} = \binom{1+i}{i}
$$

resultando que

$$
\sum_{i=0}^{\infty}\binom{1+i}{i} r^i = \frac{1}{(1-r)^2} \qquad \text{se } |\, r \,| < 1 \tag{A.7}
$$

- Derivada segunda

$$
\begin{aligned}
&\frac{d^2}{dr^2}\left(\sum_{i=0}^{\infty} r^i\right) = \frac{d^2}{dr^2}(1 + r + r^2 + r^3 + r^4 + \cdots + r^{n-1} + r^n + r^{n+1} \cdots) = \\
&= \frac{d}{dr}\Big(1 + 2r + 3r^2 + 4r^3 \cdots + (n-1)r^{n-2} + nr^{n-1} + (n+1)r^n + \cdots\Big) = \\
&= 2 \times 1 + 3 \times 2 \times r + 4 \times 3 \times r^2 + \cdots + (n-1)(n-2)r^{n-3} + n(n-1)r^{n-2} + (n+1)nr^{n-1} + \cdots \\
&= \sum_{i=0}^{\infty}(i+2)(i+1)r^i
\end{aligned}
$$

Igualando as derivadas, obtemos:

$$
\sum_{i=0}^{\infty}(i+2)(i+1)r^i = \frac{d^2}{dr^2}\left(\frac{1}{1-r}\right) = \frac{d}{dr}\left(\frac{1}{(1-r)^2}\right) = \frac{-2 \times (1-r)(-1)}{(1-r)^4}
$$

ou

$$
\sum_{i=0}^{\infty}(i+2)(i+1)r^i = \frac{2}{(1-r)^3} \tag{A.8}
$$

Esse resultado pode ser escrito de outra forma:

$$\begin{aligned}
\sum_{i=0}^{\infty}(i+2)(i+1)r^i &= \frac{2}{(1-r)^3} \Leftrightarrow \\
\sum_{i=0}^{\infty}\frac{(i+2)(i+1)}{2}r^i &= \frac{1}{(1-r)^3} \Leftrightarrow \\
\sum_{i=0}^{\infty}\frac{(i+2)(i+1)i!}{2\times i!}r^i &= \frac{1}{(1-r)^3} \Leftrightarrow \\
\sum_{i=0}^{\infty}\frac{(i+2)!}{i!\times 2!}r^i &= \frac{1}{(1-r)^3} \Leftrightarrow \\
\sum_{i=0}^{\infty}\binom{i+2}{2}r^i &= \frac{1}{(1-r)^3}
\end{aligned}$$

Pela propriedade das relações complementares, sabemos que $\binom{i+2}{2} = \binom{i+2}{i}$ e, portanto, resulta que

$$\sum_{i=0}^{\infty}\binom{i+2}{2}r^i = \sum_{i=0}^{\infty}\binom{i+2}{i}r^i = \frac{1}{(1-r)^3} \tag{A.9}$$

- Terceira derivada

$$\begin{aligned}
& \frac{d^3}{dr^3}\left(\sum_{i=0}^{\infty}r^i\right) = \frac{d^3}{dr^3}(1+r+r^2+r^3+r^4+r^5+\cdots+r^{n-1}+r^n+r^{n+1}\cdots) = \\
= & \frac{d^2}{dr^2}\left(1+2r+3r^2+4r^3+5r^4+\cdots+(n-1)r^{n-2}+nr^{n-1}+(n+1)r^n+\cdots\right) = \\
= & \frac{d}{dr}\left[\begin{array}{c} 2\times 1+3\times 2\times r+4\times 3\times r^2+5\times 4\times r^3+\cdots+(n-1)(n-2)r^{n-3} \\ +n(n-1)r^{n-2}+(n+1)nr^{n-1}+\cdots \end{array}\right] \\
= & 3\times 2\times 1+4\times 3\times 2\times r+5\times 4\times 3\times r^2+\cdots+(n-1)(n-2)(n-3)r^{n-4}+\cdots \\
& +n(n-1)(n-2)r^{n-3}+(n+1)n(n-1)r^{n-2}+\cdots \\
= & \sum_{i=0}^{\infty}(i+3)(i+2)(i+1)r^i
\end{aligned}$$

Igualando as derivadas:

$$\begin{aligned}
\sum_{i=0}^{\infty}(i+3)(i+2)(i+1)r^i = \frac{d^3}{dr^3}\left(\frac{1}{1-r}\right) &= \frac{d}{dr}\left(\frac{2}{(1-r)^3}\right) \\
&= \frac{-2\times 3\times (1-r)^2\times(-1)}{(1-r)^6} = \frac{3!}{(1-r)^4}
\end{aligned}$$

Logo,

$$\sum_{i=0}^{\infty}\frac{(i+3)(i+2)(i+1)}{3!}r^i = \frac{1}{(1-r)^4} \Rightarrow \sum_{i=0}^{\infty}\frac{(i+3)(i+2)(i+1)i!}{3!\times i!}r^i = \frac{1}{(1-r)^4} \Rightarrow$$

$$\sum_{i=0}^{\infty}\binom{3+i}{i}r^i = \sum_{i=0}^{\infty}\binom{3+i}{3}r^i = \frac{1}{(1-r)^4} \tag{A.10}$$

- Resultado geral

  Continuando a derivar (note que podemos tomar derivadas de qualquer ordem!), obtém-se o seguinte resultado geral:

$$\sum_{i=0}^{\infty} \binom{k+i}{i} r^i = \sum_{i=0}^{\infty} \binom{k+i}{k} r^i = \frac{1}{(1-r)^{k+1}} \tag{A.11}$$

A.2 A função logarítmica natural

Definição e propriedades

Definição A.1 Função logarítmica natural

A função logarítmica natural, denotada por ln, *é definida como*

$$\ln(x) = \int_1^x \frac{1}{t} dt \qquad x > 0 \tag{A.12}$$

A condição $x > 0$ é necessária pois, se tivéssemos $x \leq 0$, o domínio de integração incluiria $t = 0$, onde a função $f(t) = \frac{1}{t}$ não está definida.

Pelo Teorema Fundamental do Cálculo, vemos que $\ln(x)$ é uma antiderivada de $\frac{1}{x}$, isto é

$$\frac{d}{dx} \ln(x) = \frac{1}{x} \qquad x > 0$$

Isso significa que $\ln(x)$ é diferenciável e sua derivada é positiva no seu domínio de definição. Resulta, então, que $\ln(x)$ é uma função contínua (porque ela é diferenciável) e crescente (porque sua derivada é sempre positiva).

Da definição da função $\ln(x)$ seguem as seguintes propriedades:

1. $\ln(1) = 0$

2. Se $p > 0$ e $q > 0$, então $\ln(pq) = \ln(p) + \ln(q)$

   Demonstração

$$\frac{d}{dx} \ln(px) = \frac{1}{px} . p = \frac{1}{x}$$

   Isso significa que $\ln(px)$ e $\ln(x)$ são ambas antiderivadas de $\dfrac{1}{x}$ e, portanto, diferem por uma constante, ou seja:

$$\ln(px) = \ln(x) + C$$

   Fazendo $x = 1$, resulta que $\ln(p) = 0 + C$ e, portanto

$$\ln(px) = \ln(x) + \ln(p) \qquad x > 0, \quad p > 0$$

   Fazendo $x = q$, prova-se o resultado desejado. ■

3. Se $p > 0$, então $\ln\left(\frac{1}{p}\right) = -\ln(p)$

   Demonstração

   Fazendo $q = \frac{1}{p}$ no resultado anterior obtemos:

$$\ln\left(p\frac{1}{p}\right) = \ln(1) = \ln(p) + \ln\left(\frac{1}{p}\right)$$

   Como $\ln(1) = 0$, segue o resultado. ■

4. Se $p > 0$ e $q > 0$, então $\ln\left(\frac{p}{q}\right) = \ln(p) - \ln(q)$

   Demonstração

   Usando os resultados anteriores temos que

$$\ln\left(\frac{p}{q}\right) = \ln\left(p\frac{1}{q}\right) = \ln(p) + \ln\left(\frac{1}{q}\right) = \ln(p) - \ln(q)$$

■

5. Se $p > 0$ e r é um número racional qualquer, então $\ln(p^r) = r\ln p$

   Demonstração

   Usando a regra da cadeia, temos que

$$\frac{d}{dp}\ln(p^r) = \frac{1}{p^r}rp^{r-1} = r\left(\frac{1}{p}\right) = r\frac{d}{dp}\ln(p) = \frac{d}{dp}(r\ln p)$$

   Daí vê-se que $\ln(p^r)$ e $r\ln(p)$ têm a mesma derivada; logo

$$\ln(p^r) = r\ln(p) + C$$

   Fazendo $p = 1$ obtém-se que $C = 0$ e, portanto

$$\ln(p^r) = r\ln(p)$$

■

Gráfico

Como $\ln(x)$ é contínua e crescente, e $\ln(1) = 0$, resulta que para $x > 1$, $\ln x > 0$ e para $0 < x < 1$, $\ln x < 0$. Além disso,

$$\frac{d^2}{dx^2}\ln(x) = \frac{d}{dx}\left(\frac{1}{x}\right) = -\frac{1}{x^2} < 0$$

Logo, a função $\ln(x)$ é côncava para baixo. Vamos ver, agora, o comportamento assintótico de $\ln(x)$.

Para $x > 0$, a função $f(x) = \frac{1}{x}$ é positiva e, portanto, a integral definida pode ser interpretada como área sob a curva. Analisando a Figura A.1, podemos ver que a área sob a curva de f entre 1 e 4 é a área dos retângulos em cinza escuro mais a área em cinza claro. Os

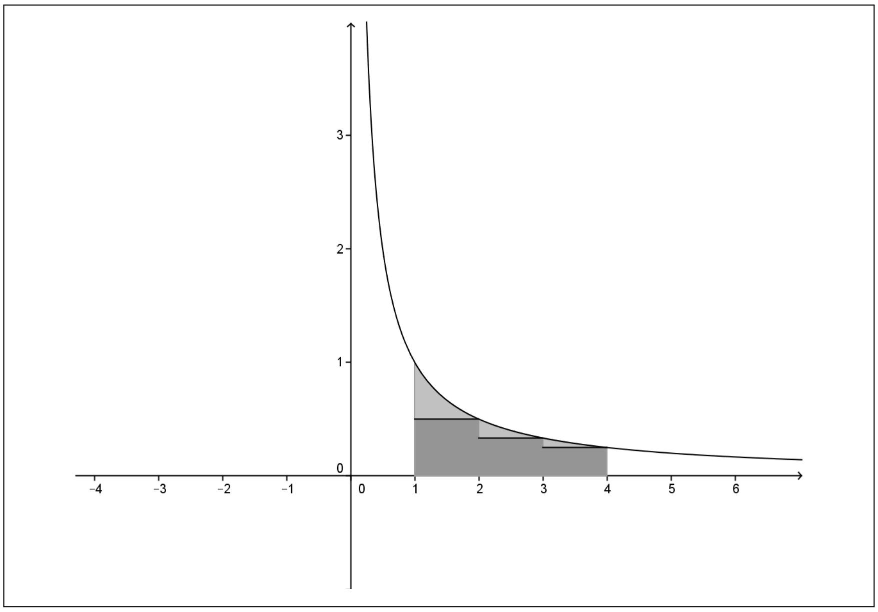


Figura A.1 – Gráfico da função $f(x) = \dfrac{1}{x}$

três retângulos (cinza escuro) têm vértices superiores à direita sobre a curva. Isso significa que a área sob a curva entre 1 e 4 é maior que a área dos três retângulos, ou seja,

$$\int_1^4 \frac{1}{t}dt > \frac{1}{2} + \frac{1}{3} + \frac{1}{4} = \frac{13}{12} > 1$$

Logo,

$$\ln(4) > 1$$

e se M é um número racional positivo qualquer,

$$M \ln(4) > M \Longrightarrow \ln\left(4^M\right) > M$$

Tomando $x > 4^M$, como $\ln(x)$ é uma função crescente, teremos

$$\ln(x) > \ln\left(4^M\right) > M$$

Como M é qualquer racional positivo, isso significa que podemos fazer $\ln(x)$ tão grande quanto desejado, bastando para isso tomar x suficientemente grande, donde se conclui que

$$\lim_{x\to\infty} (\ln x) = \infty$$

Como $\ln\left(\dfrac{1}{x}\right) = -\ln(x)$, segue que $\ln x = -\ln\left(\dfrac{1}{x}\right)$ e, portanto,

$$\lim_{x\to 0^+} (\ln x) = \lim_{x\to 0^+} \left(-\ln\frac{1}{x}\right)$$

Mas quando $x \to 0^+$, $\dfrac{1}{x} \to \infty$ e, pelo resultado acima, $\ln\frac{1}{x} \to \infty$ e, portanto, $-\ln\left(\dfrac{1}{x}\right) \to -\infty$. Isso significa que

$$\lim_{x\to 0^+} (\ln x) = -\infty$$

Na Figura A.2 temos o gráfico da função $f(x) = \ln(x)$.

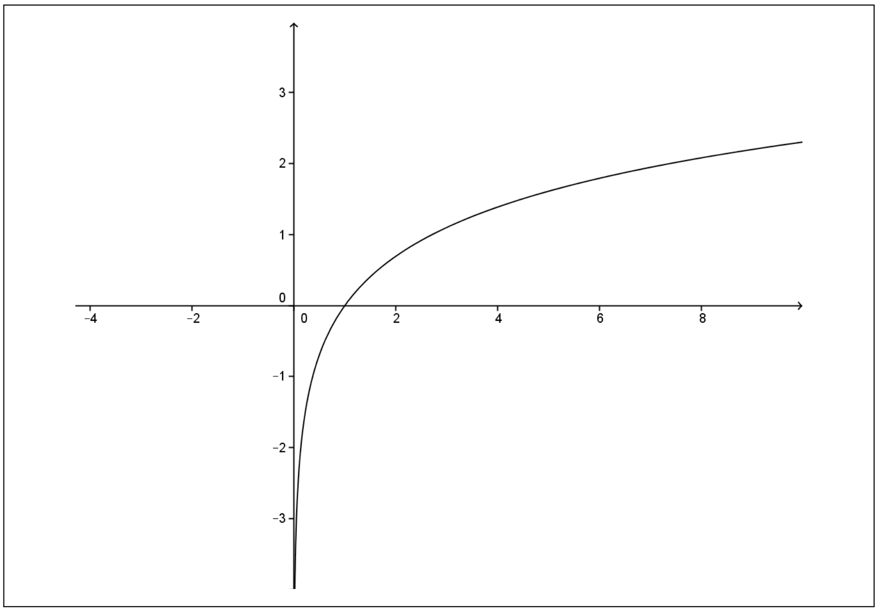


Figura A.2 – Gráfico da função $f(x) = \ln(x)$

A.3 A função exponencial natural

Definindo $f(x) = \ln(x)$, vimos que $f : \mathbb{R}^+ \to \mathbb{R}$ é uma função contínua estritamente crescente tal que $\lim\limits_{x\to\infty} \ln(x) = \infty$ e $\lim\limits_{x\to 0^+} \ln(x) = -\infty$. Com base nesses resultados, podemos provar o seguinte teorema:

Teorema A.2
A cada número real x existe um único número real positivo y tal que $\ln y = x$.

Demonstração:

□ **Demonstração**

- Se $x = 0$, então $\ln(1) = 0$ e 1 é o único valor de y para o qual $\ln(y) = 0$, uma vez que ln é estritamente crescente.

- Se $x > 0$, como $\lim\limits_{x\to\infty} \ln(x) = \infty$, podemos escolher um número b tal que $\ln(1) < x < \ln(b)$. Como ln é contínua, o Teorema do Valor Intermediário garante a existência de um número y entre 1 e b tal que $\ln(y) = x$ e esse número é único porque ln é crescente.

- Se $x < 0$, como $\lim\limits_{x\to 0^+} \ln(x) = -\infty$, existe um número $b > 0$ tal que $\ln(b) < x < \ln(1)$. Como antes, o Teorema do Valor Intermediário e a continuidade de ln garantem a existência de um único número y entre b e 1 tal que $\ln(y) = x$.

■

O teorema acima nos permite definir uma função de $\mathbb{R}$ em $\mathbb{R}^+$, uma vez que a cada $x \in \mathbb{R}$ podemos associar um único número $y \in \mathbb{R}^+$.

Definição A.3 A função exponencial natural
A função exponencial natural, denotada por exp, *é definida como*

$$\exp(x) = y \quad \textit{se e somente se} \quad \ln(y) = x \tag{A.13}$$

para todo x *e* $y > 0$.

Propriedades da função exponencial

- A função $f(x) = \exp(x)$ é biunívoca.

  Demonstração

  Se $\exp(x_1) = \exp(x_2) = b$, então, por definição, $\ln b = x_1$ e $\ln b = x_2$ e, portanto, $x_1 = x_2$. ∎

- As funções exp e ln são uma a inversa da outra.

  Demonstração

$$\begin{aligned}
\exp(\ln(x)) &= y \iff \ln(y) = \ln(x) \iff y = x \iff \exp(\ln(x)) = x \\
\ln(\exp(x)) &= y \iff \exp(y) = \exp(x) \iff y = x \iff \ln(\exp x) = x
\end{aligned}$$

∎

O número neperiano

O número neperiano é também conhecido como número de Euler, em homenagem ao matemátco suiço Leonard Euler.

Definição A.4 O número neperiano
O número neperiano e é definido como o o único número real positivo tal que

$$\ln e = 1$$

Usando a regra do trapezoidal, pode-se mostrar que

$$\int_1^{2,7} \frac{1}{x} dx < 1 < \int_1^{2,8} \frac{1}{x} dx$$

o que significa que

$$\ln(2,7) < \ln(e) < \ln(2,8)$$

ou ainda,

$$2,7 < e < 2,8$$

uma vez que ln é crescente,

Se r é qualquer número racional, então

$$\ln(e^r) = r \ln(e) = r$$

Comparando esse resultado com a definição da função exponencial, concluímos que $e^r = \exp(r)$, o que nos leva a uma outra definição.

Definição A.5
Se x é qualquer número real, então e^x é o único número real y tal que $ln(y) = x$.

Comparando as definições, concluímos que

$$\exp(x) = e^x \tag{A.14}$$

e isso nos diz que

$$e^x = y \iff \ln(y) = x \tag{A.15}$$

Mais propriedades da função exponencial

As propriedades a seguir seguem das definições e do fato de a função ln ser bijetora.

- Se p e q são números reais, então $e^p e^q = e^{p+q}$

  Demonstração

$$\ln(e^p e^q) = \ln(e^p) + \ln(e^q) = p + q = \ln(e^{p+q}) \Rightarrow e^p e^q = e^{p+q}$$

■

- Se p e q são números reais, então $\dfrac{e^p}{e^q} = e^{p-q}$

  Demonstração

$$\ln\left(\frac{e^p}{e^q}\right) = \ln(e^p) - \ln(e^q) = p - q = \ln(e^{p-q}) \Rightarrow \frac{e^p}{e^q} = e^{p-q}$$

■

- Se p é um número real e r é um número racional, então $(e^p)^r = e^{pr}$.

  Demonstração

$$\ln(e^p)^r = r\ln e^p = rp = \ln(e^{pr}) \Rightarrow (e^p)^r = e^{pr}$$

■

Derivada e gráfico

Teorema A.6 Teorema da função inversa
Se f é uma função diferenciável com inversa g e se $f'(g(c)) \neq 0$, então g é diferenciável em c e $g'(c) = \dfrac{1}{f'(g(c))}$.

Demonstração:

▼

□

Vamos aplicar este resultado para a função $f(x) = \ln(x)$. Sabemos que f é diferenciável com derivada $f'(x) = \dfrac{1}{x}$, positiva para todo $x > 0$. Vimos também que $f^{-1}(x) = g(x) = e^x$. Pelo teorema A.6, resulta que a função $g(x) = e^x$ também é diferenciável e sua derivada é dada por

$$g'(x) = \frac{1}{f'(g(x))} = \frac{1}{\frac{1}{g(x)}} = \frac{1}{\frac{1}{e^x}} = e^x$$

ou seja, a função exponencial é sua própria derivada! Como $e^x > 0$, segue que as derivadas primeira e segunda são positivas; logo, e^x é uma função crescente com concavidade para cima. Além disso,

$$x \to \infty \Longrightarrow \ln(\exp x) \to \infty \Longrightarrow \exp x \to \infty \Longrightarrow \lim_{x\to\infty} \exp x = \infty \qquad \text{(A.16)}$$

$$x \to -\infty \Longrightarrow \ln(\exp x) \to -\infty \Longrightarrow \exp x \to 0 \Longrightarrow \lim_{x\to-\infty} \exp x = 0 \qquad \text{(A.17)}$$

O gráfico da função $g(x) = e^x$ está na Figura A.3.

Note que, como a função e^x é a inversa da função $\ln(x)$, o gráfico da função exponencial é uma reflexão do gráfico da função logarítmica através da reta $y = x$. Veja a Figura A.4.

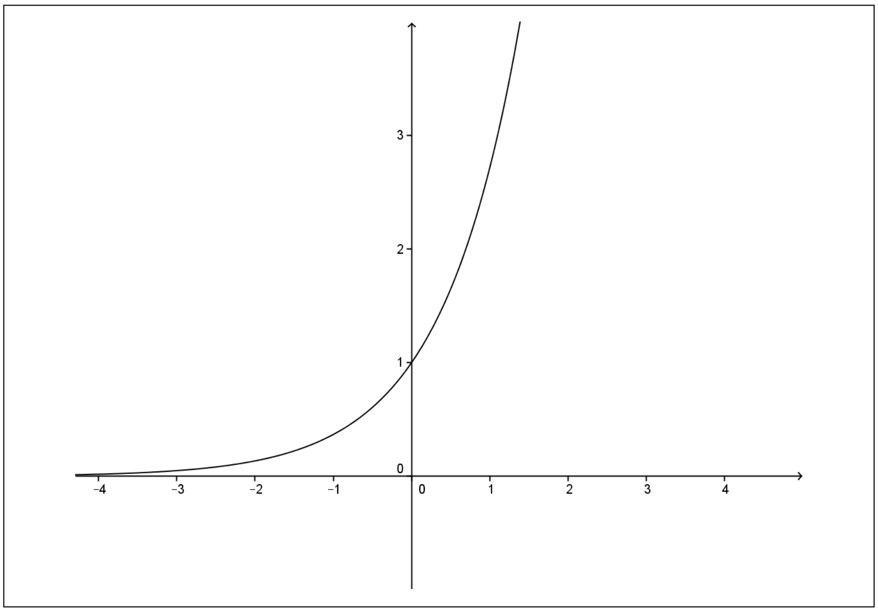


Figura A.3 – Gráfico da função $f(x) = e^x$

A.4 Função exponencial de base a

Vimos que, se r é um número racional e $a > 0$, então

$$\ln a^r = r \ln a \Longrightarrow a^r = e^{r \ln a}$$

Iremos usar este resultado para definir a função exponencial com base a.

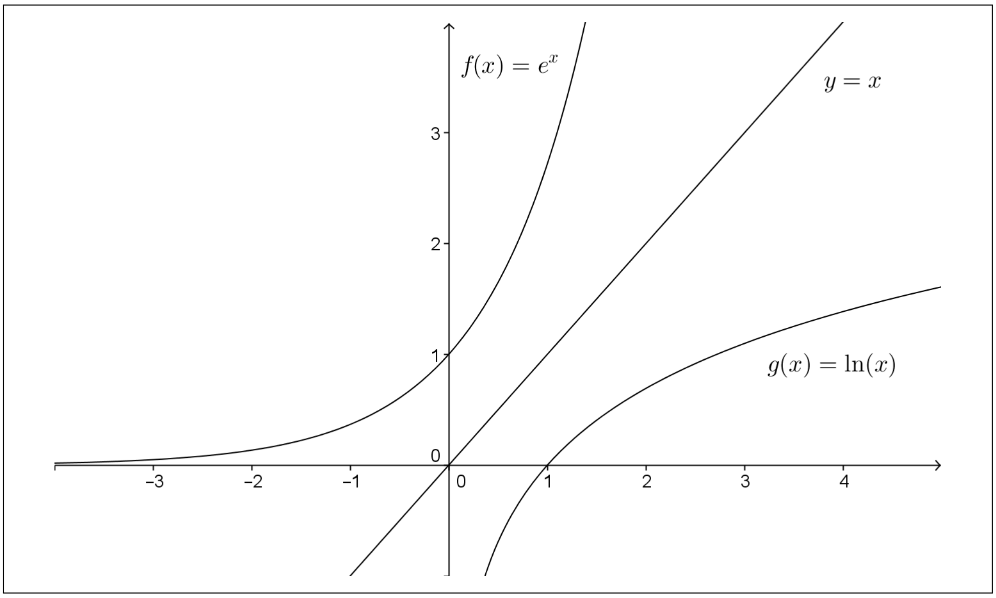


Figura A.4 – Funções logarítmica e exponencial naturais - reflexão através a reta $y = x$

Definição A.7 Função exponencial com base a
Se x *é qualquer número real e* a *é qualquer número real define-se*

$$a^x = e^{x \ln a}$$

A função $f(x) = a^x$ *é chamada a função exponencial com base* a.

Com essa definição, podemos generalizar as propriedades da funçã exponencial natural vista anteriormente.

Propriedades da função exponencial

Todas as propriedades a seguir são decorrência da definição da função exponencial com base a e propriedades já vistas da função exponencial natural.

1. Se u é um número real qualquer e $a > 0$, então $\ln a^u = u \ln a$

   Demonstração

   $$\ln a^u = \ln e^{u \ln a} = u \ln a$$

   ■

2. Se u e v são números reais quaisquer e $a > 0$, então $a^u a^v = a^{u+v}$.

   Demonstração

   $$a^u a^v = e^{u \ln a} e^{v \ln a} = e^{u \ln a + v \ln a} = e^{(u+v) \ln a} = a^{u+v}$$

   ■

3. Se u e v são números reais quaisquer e $a > 0$, então $\frac{a^u}{a^v} = a^{u-v}$.

   Demonstração

$$\frac{a^u}{a^v} = \frac{e^{u \ln a}}{e^{v \ln a}} = e^{u \ln a - v \ln a} = e^{(u-v) \ln a} = a^{u-v}$$

■

4. Se u e v são números reais quaisquer e $a > 0$, então $(a^u)^v = a^{uv}$.

   Demonstração

$$(a^u)^v = e^{v \ln a^u} = e^{vu \ln a} = a^{uv}$$

■

Derivada e gráfico da função exponencial

Para calcular a derivada da função exponencial com base a, vamos usar a regra da cadeia e o fato de a derivada da função exp ser a própria função:

$$\frac{d}{dx}(a^x) = \frac{d}{dx}\left(e^{x \ln a}\right) = e^{x \ln a} \frac{d}{dx}(x \ln a) = (\ln a)\, e^{x \ln a}$$

ou seja

$$\frac{d}{dx}(a^x) = a^x \ln a \tag{A.18}$$

Se $a > 1$, $\ln a > 0$ e a^x é uma função crescente para $x \in \mathbb{R}$ (note que $a^x > 0$ para todo a, pela definição). Se $0 < a < 1$, $\ln a < 0$ e a^x é uma função decrescente para $x \in \mathbb{R}$.

Note também que, se $a > 1$

$$\begin{aligned} \lim_{x\to\infty} a^x &= \lim_{x\to\infty} e^{x \ln a} = \infty \\ \lim_{x\to-\infty} a^x &= \lim_{x\to-\infty} e^{x \ln a} = 0 \end{aligned}$$

e se $0 < a < 1$

$$\begin{aligned} \lim_{x\to\infty} a^x &= \lim_{x\to\infty} e^{x \ln a} = 0 \\ \lim_{x\to-\infty} a^x &= \lim_{x\to-\infty} e^{x \ln a} = \infty \end{aligned}$$

Na Figura A.5 ilustra-se o gráfico de $f(x) = a^x$ para $a = 2$ e $a = \frac{1}{2}$. O gráfico da função exponencial com base a sempre terá essas formas, dependendo se $a > 1$ ou $0 < a < 1$.

A.5 A função logarítmica de base a

Se $a \neq 1$, a função $f(x) = a^x$ é bijetora e, portanto, admite uma inversa. Essa inversa será denotada por $\log_a$, o que nos leva à seguinte definição.

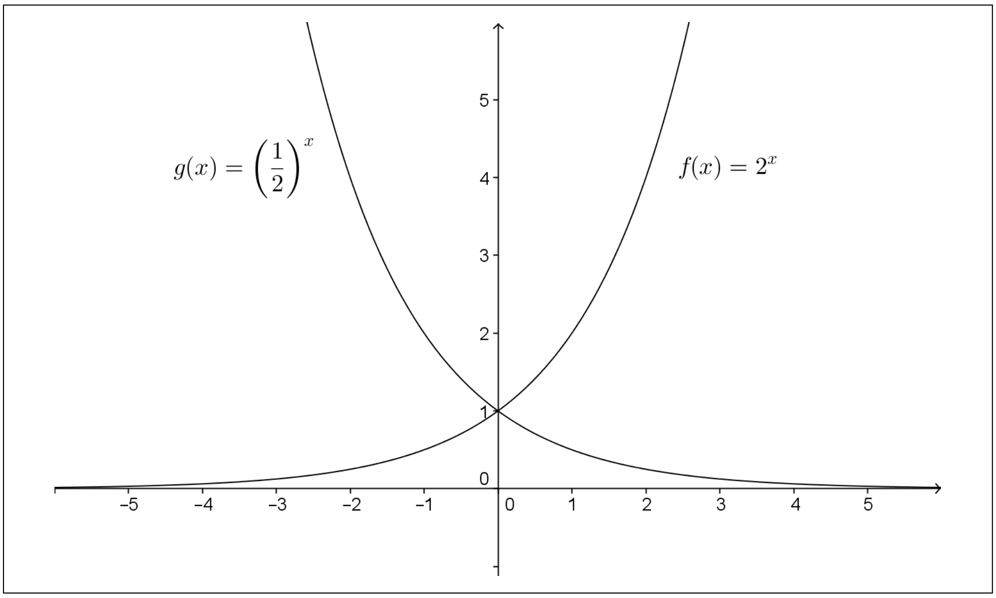


Figura A.5 – Gráfico de $f(x) = a^x$ para $a = 2$ e $a = \dfrac{1}{2}$

Definição A.8 Função logarítmica de base a
A função logarítmica de base a, $a \neq 1$, denotada por log_a é a inversa da função exponencial de base a e, portanto,

$$\log_a x = y \Longleftrightarrow x = a^y \tag{A.19}$$

Note que, quando $a = e$, temos a função logarítmica natural.

Propriedades da função logarítmica

- Relação entre as funções logarítmica natural e de base a

$$\log_a x = y \Leftrightarrow x = a^y \Leftrightarrow \ln x = y \ln a \Longleftrightarrow y = \frac{\ln x}{\ln a}$$

Como $y = \log_a x$, segue que

$$\log_a x = \frac{\ln x}{\ln a} \tag{A.20}$$

- Se $p > 0$ e $q > 0$ são números reais, então $\log_a pq = \log_a p + \log_a q$.

Demonstração

$$\log_a pq = \frac{\ln pq}{\ln a} = \frac{\ln p + \ln q}{\ln a} = \frac{\ln p}{\ln a} + \frac{\ln q}{\ln a} = \log_a p + \log_a q$$

■

- Se $p > 0$ e $q > 0$ são números reais, então $\log_a \frac{p}{q} = \log_a p - \log_a q$.

Demonstração

$$\log_a \frac{p}{q} = \frac{\ln \frac{p}{q}}{\ln a} = \frac{\ln p - \ln q}{\ln a} = \frac{\ln p}{\ln a} - \frac{\ln q}{\ln a} = \log_a p - \log_a q$$

■

- Se $p > 0$ e x é número real qualquer, então $\log_a p^x = x \log_a p$.

  Demonstração

$$\log_a p^x = \frac{\ln p^x}{\ln a} = \frac{x \ln p}{\ln a} = x \log_a p$$

■

Derivada e gráfico

Vamos calcular a derivada de $f(x) = \log_a x$.

$$\log_a x = \frac{\ln x}{\ln a} \Longrightarrow \frac{d}{dx}\left(\log_a x\right) = \frac{d}{dx}\left(\frac{\ln x}{\ln a}\right) = \frac{1}{\ln a}\frac{1}{x}$$

ou seja,

$$\frac{d}{dx}\left(\log_a x\right) = \frac{1}{x \ln a} \tag{A.21}$$

Como no caso das funções exponencial e logarítimica naturais, os gráficos de $f(x) = a^x$ e $g(x) = \log_a x$ são um a reflexão do utro através da reta $y = x$, conforme ilustrado na Figura A.6.

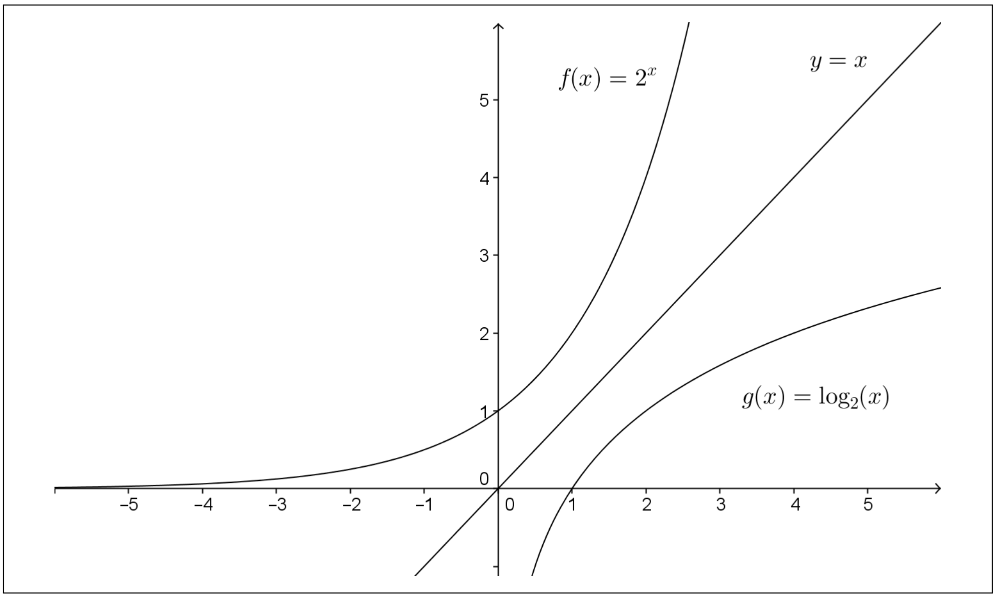


Figura A.6 – Funções logarítmica e exponencial de base a - reflexão através a reta $y = x$

A.6 Limites envolvendo as funções exponencial e logarítmica

- $\lim_{h \to 0} (1 + h)^{\frac{1}{h}} = e$

  Demonstração

  Da definição de derivada da função $f(x) = \ln x$, temos que

$$
\begin{aligned}
f'(x) &= \lim_{h\to 0} \frac{\ln(x+h) - \ln(x)}{h} \\
&= \lim_{h\to 0} \frac{1}{h} \ln\left(\frac{x+h}{x}\right) \\
&= \lim_{h\to 0} \ln\left(\frac{x+h}{x}\right)^{\frac{1}{h}} \\
&= \lim_{h\to 0} \ln\left(1+\frac{h}{x}\right)^{\frac{1}{h}}
\end{aligned}
$$

Como $f'(x) = \dfrac{1}{x}$, para $x = 1$ temos $f'(1) = 1$ e, portanto,

$$1 = \lim_{h\to 0} \ln(1+h)^{\frac{1}{h}}$$

Mas

$$(1+h)^{\frac{1}{h}} = \exp\left[\ln(1+h)^{\frac{1}{h}}\right]$$

Como a função exponencial é contínua,

$$
\begin{aligned}
\lim_{h\to 0}(1+h)^{\frac{1}{h}} &= \lim_{h\to 0} \exp\ln(1+h)^{\frac{1}{h}} \\
&= \exp\lim_{h\to 0} \ln(1+h)^{\frac{1}{h}} \\
&= \exp(1) \\
&= e
\end{aligned}
$$

ou seja,

$$\lim_{h\to 0}(1+h)^{\frac{1}{h}} = e \tag{A.22}$$

Fazendo $\frac{1}{h} = n$ na expressão acima, obtemos que

$$\lim_{n\to\infty}\left(1+\frac{1}{n}\right)^n = e \tag{A.23}$$

■

- $\lim_{n\to\infty}\left(1+\dfrac{x}{n}\right)^n = e^x$

Demonstração

Fazendo $t = \dfrac{x}{n}$, resulta que

$$\lim_{n\to\infty}\left(1+\frac{x}{n}\right)^n = \lim_{t\to 0}(1+t)^{\frac{x}{t}} = \lim_{t\to 0}\left[(1+t)^{1/t}\right]^x$$

Mas a função exponencial é contínua; logo,

$$\lim_{t\to 0}\left[(1+t)^{1/t}\right]^x = \left[\lim_{t\to 0}(1+t)^{1/t}\right]^x = e^x$$

Logo,

$$\lim_{n\to\infty}\left(1+\frac{x}{n}\right)^n = e^x \tag{A.24}$$

■

- $\lim_{x\to\infty} x^k e^{-x} = 0$

Demonstração

Vamos mostrar esse resultado usando demonstração por indução e usando, também, a regra de L'Hôpital.

⋆ $k = 1$

$$\lim_{x\to\infty} xe^{-x} = \lim_{x\to\infty} \frac{x}{e^x}$$

que tem a forma $\frac{\infty}{\infty}$ e, portanto, podemos aplicar L'Hôpital, que diz que

$$\lim_{x\to\infty} xe^{-x} = \lim_{x\to\infty} \frac{x}{e^x} = \lim_{x\to\infty} \frac{x'}{(e^x)'} = \lim_{x\to\infty} \frac{1}{e^x} = 0$$

Logo, o resultado vale para $k = 1$.

⋆ Suponhamos verdadeiro para qualquer k; vamos mostrar que vale para $k + 1$.

$$\begin{aligned}\lim_{x\to\infty} x^{k+1}e^{-x} &= \lim_{x\to\infty} \frac{x^{k+1}}{e^x} = \lim_{x\to\infty} \frac{\left(x^{k+1}\right)'}{(e^x)'} = \lim_{x\to\infty} \frac{(k+1)\,x^k}{e^x} \\ &= (k+1) \lim_{x\to\infty} \frac{x^k}{e^x} = (k+1)\times 0 = 0\end{aligned}$$

pela hipótese de indução. ■

A interpretação desse resultado pode ser vista na Figura A.7 para o caso de x^3: para valores de x maiores que a abscissa do ponto de interseção representado em verde, o valor de e^x é sempre maior que x^3 e, portanto o limite de $\frac{x^3}{e^x}$ é zero. Dito de outra forma, a função exponencial cresce muito mais rapidamente que qualquer função polinomial.

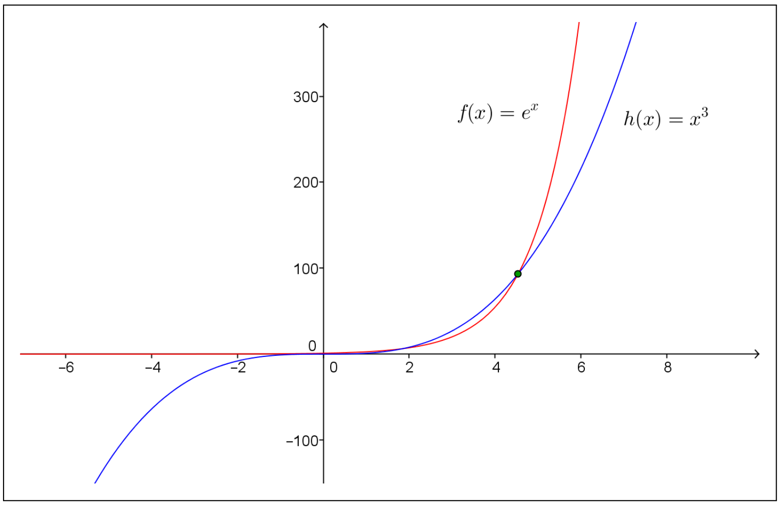


Figura A.7 – Ilustração do resultado $\lim_{x\to\infty} \frac{x^3}{e^x} = 0$

De maneira análoga, prova-se um resultado mais geral dado por:

$$\lim_{x\to\infty} x^k e^{-\lambda x} = 0 \qquad \forall k > 0 \text{ e } \lambda > 0 \tag{A.25}$$

Fórmula de Taylor

Vamos ver, agora, a fórmula de Taylor para a função $f(x) = e^x$. Lembrando que a derivada de $f(x)$ é a própria $f(x)$ e que $f(0) = 1$, podemos escrever:

$$e^x = 1 + x + \frac{x^2}{2!} + \cdots + \frac{x^n}{n!} + R_{n+1}$$

onde

$$R_{n+1} = \frac{e^z}{(n+1)!}x^{n+1}$$

para algum z entre 0 e x. Se $x > 0$, então $0 < z < x$, o que implica que $e^z < e^x$ e, portanto

$$0 < R_{n+1} < \frac{e^x}{(n+1)!}x^{n+1}$$

Mas

$$\lim_{n\to\infty} \frac{e^x}{(n+1)!}x^{n+1} = e^x \lim_{n\to\infty} \frac{x^{n+1}}{(n+1)!} = e^x \times 0 = 0$$

e, pelo teorema do sanduíche, $\lim R_{n+1} = 0$. Se $x < 0$, como z está entre x e 0, segue que $z < 0$; logo, $e^z < 1$ e

$$0 < |R_{n+1}| < \left|\frac{x^{n+1}}{(n+1)!}\right| \to 0$$

Logo, também neste caso $\lim\limits_{n\to\infty} R_{n+1} = 0$.

Segue, então, que $f(x) = e^x$ é representada pela série de Taylor, ou seja, podemos escrever

$$e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!} \tag{A.26}$$

A.7 Funções pares e ímpares

Definição A.9 Funções pares e ímpares
Uma função $f(x)$ é par se

$$f(-x) = f(x) \tag{A.27}$$

Uma função $f(x)$ é ímpar se

$$f(-x) = -f(x) \tag{A.28}$$

Teorema A.10 Integral de funções pares e ímpares

- *Se $f(x)$ é par, então*

$$\int_{-\infty}^{\infty} f(x)dx = 2\int_0^{\infty} f(x)dx \tag{A.29}$$

- *Se $f(x)$ é ímpar, então*

$$\int_{-\infty}^{\infty} f(x)dx = 0 \tag{A.30}$$

Demonstração:

☐ **Demonstração**

Podemos escrever

$$\int_{-\infty}^{\infty} f(x)dx = \int_{-\infty}^{0} f(x)dx + \int_{0}^{\infty} f(x)dx$$

Fazendo $x = -t$ na primeira integral, tem-se que:

$$\begin{aligned}\int_{-\infty}^{\infty} f(x)dx &= \int_{\infty}^{0} f(-t)(-dt) + \int_{0}^{\infty} f(x)dx = -\int_{\infty}^{0} f(-t)dt + \int_{0}^{\infty} f(x)dx \\ &= \int_{0}^{\infty} f(-t)dt + \int_{0}^{\infty} f(x)dx\end{aligned}$$

- Se a função é par, resulta que:

$$\int_{-\infty}^{\infty} f(x)dx = \int_{0}^{\infty} f(-t)dt + \int_{0}^{\infty} f(x)dx = \int_{0}^{\infty} f(t)dt + \int_{0}^{\infty} f(x)dx = 2\int_{0}^{\infty} f(x)dx$$

- Se a função é ímpar, resulta que:

$$\begin{aligned}\int_{-\infty}^{\infty} f(x)dx &= \int_{0}^{\infty} f(-t)dt + \int_{0}^{\infty} f(x)dx = \int_{0}^{\infty} -f(t)dt + \int_{0}^{\infty} f(x)dx \\ &= -\int_{0}^{\infty} f(t)dt + \int_{0}^{\infty} f(x)dx = 0\end{aligned}$$

■

A.8 Integrais duplas com coordenadas polares

No cálculo de integrais duplas, a utilização de coordenadas polares é bastante útil. No sistema cartesiano, cada ponto do plano é representado pela sua abscissa e sua ordenada. No sistema de coordenadas polares, em vez dos eixos x e y, trabalha-se com um *polo* ou *origem*, que é um ponto fixo do plano, e um *eixo polar*, que é uma semi-reta orientada, conforme ilustrado na Figura A.8. As coordenadas polares de um ponto P qualquer no plano são a distância $r = d(O, P)$ e o ângulo θ determinado pelo eixo polar e o segmento OP. Em geral, a coordenada θ é considerada positiva se é gerada por uma rotação anti-horária e negativa, caso contrário.

As coordenadas polares de um ponto não são únicas, pois há vários ângulos com o mesmo lado terminal OP. Por exemplo, $\left(3, \frac{\pi}{4}\right)$, $\left(3, \frac{9\pi}{4}\right)$ e $\left(3, -\frac{7\pi}{4}\right)$ representam o mesmo ponto. Permite-se também que a coordenada r seja negativa.

Se uma função de uma variável f está definida em um intervalo $[a, b]$, a definição de integral (simples) requer a partição do intervalo $[a, b]$ em subintervalos de comprimento Δx_i, conforme ilustrado na Figura A.9. Se denotamos por P a partição $\{x_0, x_1, \ldots, x_n\}$, então a integral definida de Riemann é definida como

$$\int_{a}^{b} f(x)dx = \lim_{\|P\| \to 0} \sum f(w_i)\Delta x_i$$

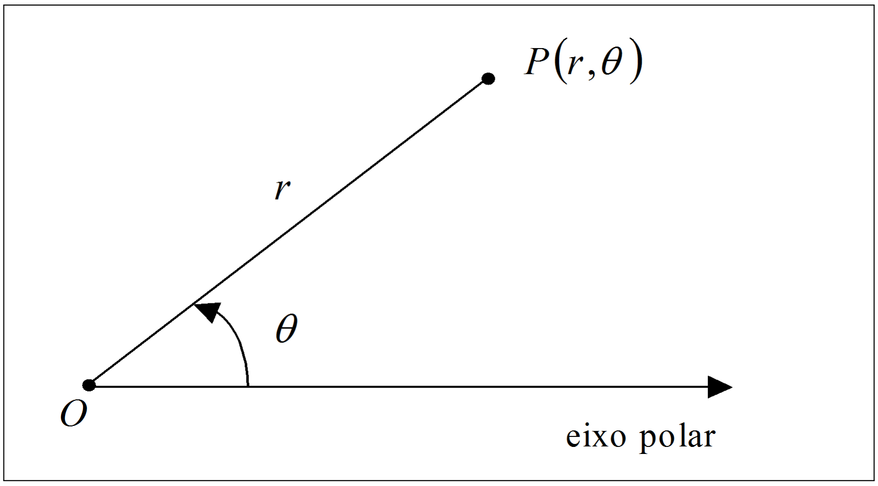


Figura A.8 – Sistema de coordenadads polares

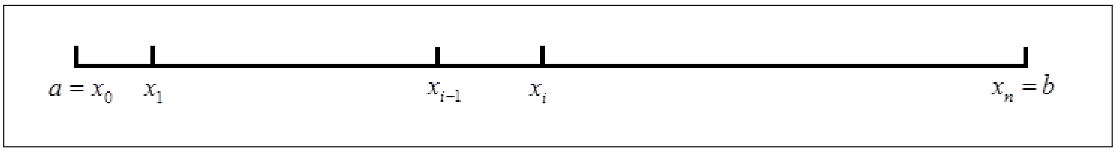


Figura A.9 – Partição do intervalo $[a, b]$

Para integrais duplas de funções definidas em coordenadas cartesianas, temos que ter a partição da região de definição R da função de duas variáveis $f(x, y)$, conforme ilustrado na Figura A.10. Nesse caso, a integral dupla é definida como

$$\int\int_R f(x, y) dA = \lim_{\|P\| \to 0} f(u_i, v_i) \Delta A_i$$

e é calculada através de integrais parciais, isto é, calcula-se primeiro a integral com respeito a y, mantendo x constante, e depois integra-se o resultado com relação a x :

$$\int_a^b \int_c^d f(x, y) dy dx = \int_a^b \left[\int_c^d f(x, y) dy \right] dx$$

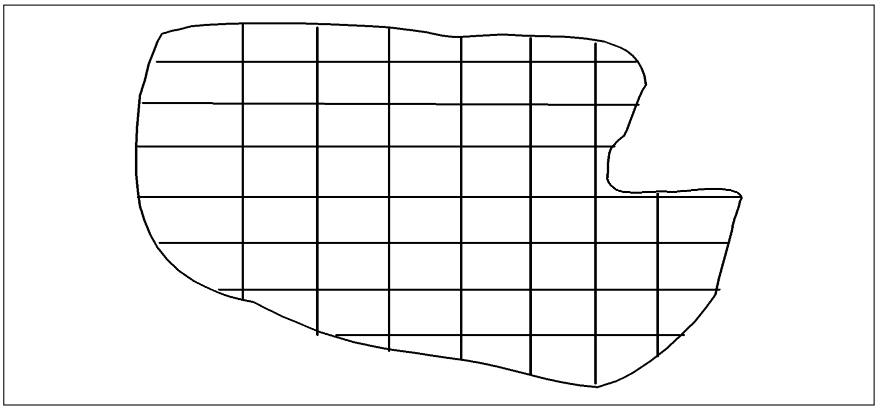


Figura A.10 – Partição da região R – coordenadas cartesianas

O cálculo de integrais duplas de funções parametrizadas em termos de coordenadas polares requer que a partição da região de integração seja definida em termos de tais coordenadas, conforme ilustrado na Figura A.11.

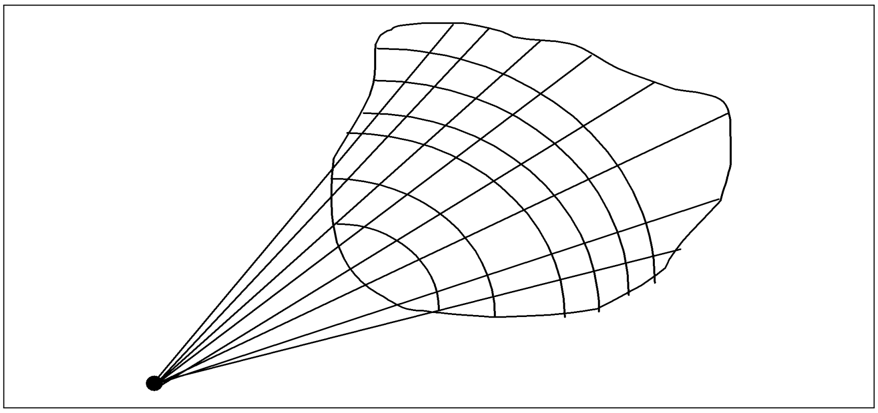


Figura A.11 – Partição da região R – coordenadas polares

Vamos considerar uma *região polar elementar* (ver Figura A.12). A área total de um

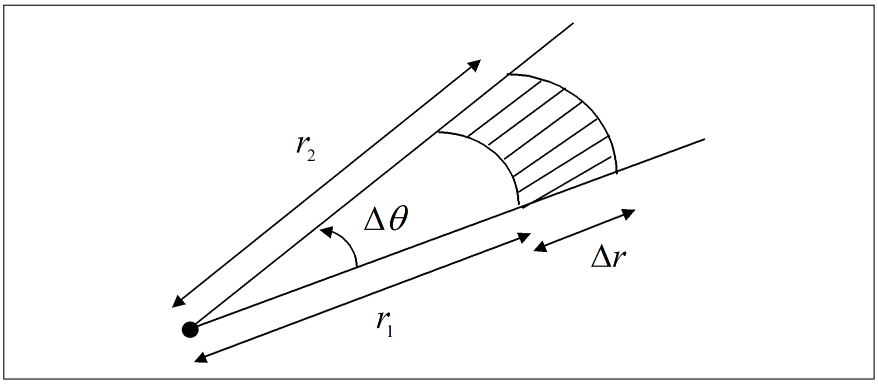


Figura A.12 – Região polar elementar

círculo é dada por

$$\pi r^2 = \frac{2\pi r^2}{2} = \frac{1}{2}(2\pi)\, r^2 \quad ;$$

logo, a área de um setor circular correspondente a um ângulo $\Delta\theta$ e raio r é

$$\frac{1}{2}(\Delta\theta)\, r^2$$

Então, a área da região polar elementar é

$$\Delta A = \frac{1}{2}(\Delta\theta)\, r_2^2 - \frac{1}{2}(\Delta\theta)\, r_1^2 = \frac{1}{2}(\Delta\theta)\left(r_2^2 - r_1^2\right) = \frac{1}{2}(\Delta\theta)(r_1 + r_2)(r_2 - r_1)$$

Definindo

$$\begin{aligned} \overline{r} &= \frac{1}{2}(r_1 + r_2) \\ \Delta r &= r_2 - r_1 \end{aligned}$$

podemos escrever:

$$\Delta A = \overline{r}\Delta\theta\Delta r$$

A integral em termos das coordenadas polares é, então:

$$\int\int_R f(r,\theta)dA = \lim_{\|P\|\to 0} \sum f(r_i,\theta_i)r_i\Delta\theta_i\Delta r_i$$

e calculada como

$$\int\int_R f(r,\theta)dA = \int\int_R f(r,\theta)r\,dr\,d\theta \tag{A.31}$$

Exemplo A.11 $e^{-t^2/2}$

Solução:
A

♦♦ função $f(t) = \exp\left(-\frac{t^2}{2}\right)$ desempenha papel importante em probabilidade. Para calcular sua integral, não podemos aplicar o Teorema Fundamental do Cálculo pois $f(t)$ não possui antiderivada. Mas podemos ver que $f(t)$ é uma função par. Sabemos, então, que

$$\int_{-\infty}^{\infty} \exp\left(-\frac{t^2}{2}\right) dt = 2\int_0^{\infty} \exp\left(-\frac{t^2}{2}\right) dt$$

Temos, também, que

$$\begin{aligned}\left[\int_0^{\infty} \exp\left(-\frac{t^2}{2}\right) dt\right]^2 &= \left(\int_0^{\infty} \exp\left(-\frac{t^2}{2}\right) dt\right)\left(\int_0^{\infty} \exp\left(-\frac{u^2}{2}\right) du\right) = \\ &= \int_0^{\infty}\int_0^{\infty} \exp\left(-\frac{t^2+u^2}{2}\right) dt\,du\end{aligned}$$

Nessa integral dupla, a região de integração é o primeiro quadrante e a relação entre as coordenadas cartesianas (u,t) e as coordenadas polares (r,θ) nesse quadrante é (veja a Figura A.13)

$$\begin{array}{lll} u = r\cos\theta & u > 0 & r > 0 \\ t = r\sin\theta & t > 0 & 0 < \theta < \frac{\pi}{2} \end{array}$$

Temos também que

$$t^2 + u^2 = r^2$$

e, então, podemos calcular a integral dupla usando coordenadas polares da seguinte forma:

$$\int_0^{\infty}\int_0^{\infty} \exp\left(-\frac{t^2+u^2}{2}\right) dt\,du = \int_0^{\infty}\int_0^{\frac{\pi}{2}} \exp\left(-\frac{r^2}{2}\right) r\,d\theta\,dr$$

Fazendo $\frac{r^2}{2} = w$, então $r\,dr = dw$ e w varia de 0 a ∞. Logo,

$$\begin{aligned}\int_0^{\infty}\int_0^{\frac{\pi}{2}} \exp\left(-\frac{r^2}{2}\right) r\,d\theta\,dr &= \int_0^{\infty}\int_0^{\frac{\pi}{2}} \exp(-w)\,d\theta\,dw = \int_0^{\infty}\left[\int_0^{\frac{\pi}{2}} d\theta\right]\exp(-w)\,dw = \\ &= \int_0^{\infty} \frac{\pi}{2}\exp(-w)dw = \frac{\pi}{2}\left(-e^{-w}\right)\Big|_0^{\infty} = \frac{\pi}{2}\end{aligned}$$

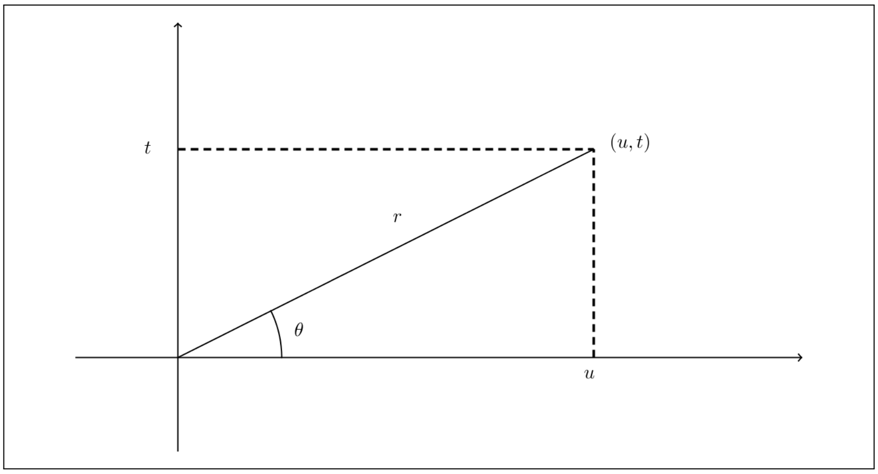


Figura A.13 – Relação entre coordenadas cartesianas e polares no primeiro quadrante

Provamos assim que

$$\left[\int_0^\infty \exp\left(-\frac{t^2}{2}\right)dt\right]^2 = \frac{\pi}{2} \Rightarrow$$
$$\int_0^\infty \exp\left(-\frac{t^2}{2}\right)dt = \sqrt{\frac{\pi}{2}} \tag{A.32}$$

e, portanto,

$$\int_{-\infty}^\infty \exp\left(-\frac{t^2}{2}\right)dt = \sqrt{2\pi} \tag{A.33}$$

ou ainda

$$\frac{1}{\sqrt{2\pi}}\int_{-\infty}^\infty \exp\left(-\frac{t^2}{2}\right)dt = 1 \tag{A.34}$$

A.9 A função gama

A função gama é definida pela seguinte integral:

$$\Gamma(\alpha) = \int_0^\infty e^{-x}x^{\alpha-1}dx \qquad \alpha \geq 1 \tag{A.35}$$

A função gama tem a seguinte propriedade recursiva: $\Gamma(\alpha+1) = \alpha\Gamma(\alpha)$. Para demonstrar esse resultado, iremos usar integração por partes.

$$\Gamma(\alpha+1) = \int_0^\infty e^{-x}x^{\alpha}dx$$

Fazendo

- $u = x^\alpha \Rightarrow du = \alpha x^{\alpha-1}$

- $dv = e^{-x}dx \Rightarrow v = -e^{-x}$

Logo,

$$\begin{aligned} \Gamma(\alpha+1) &= -x^{\alpha}e^{-x}\Big|_0^{\infty} - \int_0^{\infty} -e^{-x}\alpha x^{\alpha-1}dx \Longrightarrow \\ \Gamma(\alpha+1) &= 0 + \alpha\int_0^{\infty} e^{-x}x^{\alpha-1}dx \Longrightarrow \\ \Gamma(\alpha+1) &= \alpha\Gamma(\alpha) \end{aligned} \tag{A.36}$$

Aqui usamos o resultado dado em (A.25).

Vamos trabalhar, agora, com $\alpha = n$ inteiro.

$$\Gamma(1) = \int_0^{\infty} e^{-x}x^{1-1}dx = \int_0^{\infty} e^{-x}dx = 1$$

$$\begin{aligned} \Gamma(2) &= 1 \times \Gamma(1) = 1 = 1! \\ \Gamma(3) &= 2 \times \Gamma(2) = 2 \times 1 = 2! \\ \Gamma(4) &= 3 \times \Gamma(3) = 3 \times 2 \times 1 = 3! \\ \Gamma(5) &= 4 \times \Gamma(4) = 4 \times 3 \times 2 \times 1 = 4! \end{aligned}$$

Em geral, se n é inteiro,

$$\Gamma(n) = (n-1)! \tag{A.37}$$

Exemplo A.12 $\Gamma\left(\frac{1}{2}\right)$

Solução:
P

♦♦ or definição,

$$\Gamma\left(\frac{1}{2}\right) = \int_0^{\infty} e^{-x}x^{1/2-1}dx = \int_0^{\infty} \frac{e^{-x}}{\sqrt{x}}dx$$

Usando a seguinte transformação de variável:

$$x = \frac{t^2}{2}$$

obtemos

$$\begin{aligned} dx &= tdt \\ x &= 0 \Rightarrow t = 0 \\ x &\to \infty \Rightarrow t \to \infty \end{aligned}$$

Logo,

$$\Gamma\left(\frac{1}{2}\right) = \int_0^{\infty}\left(\frac{e^{-t^2/2}}{\sqrt{\frac{t^2}{2}}}\right)tdt = \sqrt{2}\int_0^{\infty} e^{-t^2/2}dt = \sqrt{2} \times \sqrt{\frac{\pi}{2}}$$

ou seja:

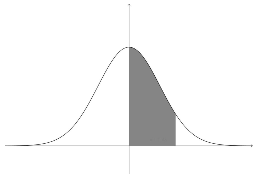
$$\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi} \tag{A.38}$$

Apêndice B

Tabelas da Distribuição Normal

- Tabela 1: Tabela da normal padrão – $p = \mathrm{P}(0 \leq Z \leq z)$
- Tabela 2: Tabela da distribuição acumulada da normal padrão – $\Phi(z) = \mathrm{P}(Z \leq z),\ z \geq 0$

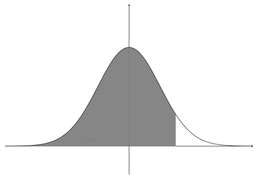
Tabela 1: $Z \sim N(0;1)$
Valores de p
$p = \mathrm{P}(0 \leq Z \leq z)$



Casa inteira e 1ª Decimal	2ª decimal									
	0	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
0,0	0,0000	0,0040	0,0080	0,0120	0,0160	0,0199	0,0239	0,0279	0,0319	0,0359
0,1	0,0398	0,0438	0,0478	0,0517	0,0557	0,0596	0,0636	0,0675	0,0714	0,0753
0,2	0,0793	0,0832	0,0871	0,0910	0,0948	0,0987	0,1026	0,1064	0,1103	0,1141
0,3	0,1179	0,1217	0,1255	0,1293	0,1331	0,1368	0,1406	0,1443	0,1480	0,1517
0,4	0,1554	0,1591	0,1628	0,1664	0,1700	0,1736	0,1772	0,1808	0,1844	0,1879
0,5	0,1915	0,1950	0,1985	0,2019	0,2054	0,2088	0,2123	0,2157	0,2190	0,2224
0,6	0,2257	0,2291	0,2324	0,2357	0,2389	0,2422	0,2454	0,2486	0,2517	0,2549
0,7	0,2580	0,2611	0,2642	0,2673	0,2704	0,2734	0,2764	0,2794	0,2823	0,2852
0,8	0,2881	0,2910	0,2939	0,2967	0,2995	0,3023	0,3051	0,3078	0,3106	0,3133
0,9	0,3159	0,3186	0,3212	0,3238	0,3264	0,3289	0,3315	0,3340	0,3365	0,3389
1,0	0,3413	0,3438	0,3461	0,3485	0,3508	0,3531	0,3554	0,3577	0,3599	0,3621
1,1	0,3643	0,3665	0,3686	0,3708	0,3729	0,3749	0,3770	0,3790	0,3810	0,3830
1,2	0,3849	0,3869	0,3888	0,3907	0,3925	0,3944	0,3962	0,3980	0,3997	0,4015
1,3	0,4032	0,4049	0,4066	0,4082	0,4099	0,4115	0,4131	0,4147	0,4162	0,4177
1,4	0,4192	0,4207	0,4222	0,4236	0,4251	0,4265	0,4279	0,4292	0,4306	0,4319
1,5	0,4332	0,4345	0,4357	0,4370	0,4382	0,4394	0,4406	0,4418	0,4429	0,4441
1,6	0,4452	0,4463	0,4474	0,4484	0,4495	0,4505	0,4515	0,4525	0,4535	0,4545
1,7	0,4554	0,4564	0,4573	0,4582	0,4591	0,4599	0,4608	0,4616	0,4625	0,4633
1,8	0,4641	0,4649	0,4656	0,4664	0,4671	0,4678	0,4686	0,4693	0,4699	0,4706
1,9	0,4713	0,4719	0,4726	0,4732	0,4738	0,4744	0,4750	0,4756	0,4761	0,4767
2,0	0,4772	0,4778	0,4783	0,4788	0,4793	0,4798	0,4803	0,4808	0,4812	0,4817
2,1	0,4821	0,4826	0,4830	0,4834	0,4838	0,4842	0,4846	0,4850	0,4854	0,4857
2,2	0,4861	0,4864	0,4868	0,4871	0,4875	0,4878	0,4881	0,4884	0,4887	0,4890
2,3	0,4893	0,4896	0,4898	0,4901	0,4904	0,4906	0,4909	0,4911	0,4913	0,4916
2,4	0,4918	0,4920	0,4922	0,4925	0,4927	0,4929	0,4931	0,4932	0,4934	0,4936
2,5	0,4938	0,4940	0,4941	0,4943	0,4945	0,4946	0,4948	0,4949	0,4951	0,4952
2,6	0,4953	0,4955	0,4956	0,4957	0,4959	0,4960	0,4961	0,4962	0,4963	0,4964
2,7	0,4965	0,4966	0,4967	0,4968	0,4969	0,4970	0,4971	0,4972	0,4973	0,4974
2,8	0,4974	0,4975	0,4976	0,4977	0,4977	0,4978	0,4979	0,4979	0,4980	0,4981
2,9	0,4981	0,4982	0,4982	0,4983	0,4984	0,4984	0,4985	0,4985	0,4986	0,4986
3,0	0,4987	0,4987	0,4987	0,4988	0,4988	0,4989	0,4989	0,4989	0,4990	0,4990
3,1	0,4990	0,4991	0,4991	0,4991	0,4992	0,4992	0,4992	0,4992	0,4993	0,4993
3,2	0,4993	0,4993	0,4994	0,4994	0,4994	0,4994	0,4994	0,4995	0,4995	0,4995
3,3	0,4995	0,4995	0,4995	0,4996	0,4996	0,4996	0,4996	0,4996	0,4996	0,4997
3,4	0,4997	0,4997	0,4997	0,4997	0,4997	0,4997	0,4997	0,4997	0,4997	0,4998
3,5	0,4998	0,4998	0,4998	0,4998	0,4998	0,4998	0,4998	0,4998	0,4998	0,4998
3,6	0,4998	0,4998	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999
3,7	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999
3,8	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999
3,9	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000
4,0	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000

Para abscissas maiores que 4,09, use a probabilidade 0,5000.

Tabela 2: $Z \sim N(0; 1)$
Valores de p
$p = \Phi(z) = P(Z \leq z)$



Casa inteira e 1ª Decimal	2ª decimal									
	0	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
0,0	0,5000	0,5040	0,5080	0,5120	0,5160	0,5199	0,5239	0,5279	0,5319	0,5359
0,1	0,5398	0,5438	0,5478	0,5517	0,5557	0,5596	0,5636	0,5675	0,5714	0,5753
0,2	0,5793	0,5832	0,5871	0,5910	0,5948	0,5987	0,6026	0,6064	0,6103	0,6141
0,3	0,6179	0,6217	0,6255	0,6293	0,6331	0,6368	0,6406	0,6443	0,6480	0,6517
0,4	0,6554	0,6591	0,6628	0,6664	0,6700	0,6736	0,6772	0,6808	0,6844	0,6879
0,5	0,6915	0,6950	0,6985	0,7019	0,7054	0,7088	0,7123	0,7157	0,7190	0,7224
0,6	0,7257	0,7291	0,7324	0,7357	0,7389	0,7422	0,7454	0,7486	0,7517	0,7549
0,7	0,7580	0,7611	0,7642	0,7673	0,7704	0,7734	0,7764	0,7794	0,7823	0,7852
0,8	0,7881	0,7910	0,7939	0,7967	0,7995	0,8023	0,8051	0,8078	0,8106	0,8133
0,9	0,8159	0,8186	0,8212	0,8238	0,8264	0,8289	0,8315	0,8340	0,8365	0,8389
1,0	0,8413	0,8438	0,8461	0,8485	0,8508	0,8531	0,8554	0,8577	0,8599	0,8621
1,1	0,8643	0,8665	0,8686	0,8708	0,8729	0,8749	0,8770	0,8790	0,8810	0,8830
1,2	0,8849	0,8869	0,8888	0,8907	0,8925	0,8944	0,8962	0,8980	0,8997	0,9015
1,3	0,9032	0,9049	0,9066	0,9082	0,9099	0,9115	0,9131	0,9147	0,9162	0,9177
1,4	0,9192	0,9207	0,9222	0,9236	0,9251	0,9265	0,9279	0,9292	0,9306	0,9319
1,5	0,9332	0,9345	0,9357	0,9370	0,9382	0,9394	0,9406	0,9418	0,9429	0,9441
1,6	0,9452	0,9463	0,9474	0,9484	0,9495	0,9505	0,9515	0,9525	0,9535	0,9545
1,7	0,9554	0,9564	0,9573	0,9582	0,9591	0,9599	0,9608	0,9616	0,9625	0,9633
1,8	0,9641	0,9649	0,9656	0,9664	0,9671	0,9678	0,9686	0,9693	0,9699	0,9706
1,9	0,9713	0,9719	0,9726	0,9732	0,9738	0,9744	0,9750	0,9756	0,9761	0,9767
2,0	0,9772	0,9778	0,9783	0,9788	0,9793	0,9798	0,9803	0,9808	0,9812	0,9817
2,1	0,9821	0,9826	0,9830	0,9834	0,9838	0,9842	0,9846	0,9850	0,9854	0,9857
2,2	0,9861	0,9864	0,9868	0,9871	0,9875	0,9878	0,9881	0,9884	0,9887	0,9890
2,3	0,9893	0,9896	0,9898	0,9901	0,9904	0,9906	0,9909	0,9911	0,9913	0,9916
2,4	0,9918	0,9920	0,9922	0,9925	0,9927	0,9929	0,9931	0,9932	0,9934	0,9936
2,5	0,9938	0,9940	0,9941	0,9943	0,9945	0,9946	0,9948	0,9949	0,9951	0,9952
2,6	0,9953	0,9955	0,9956	0,9957	0,9959	0,9960	0,9961	0,9962	0,9963	0,9964
2,7	0,9965	0,9966	0,9967	0,9968	0,9969	0,9970	0,9971	0,9972	0,9973	0,9974
2,8	0,9974	0,9975	0,9976	0,9977	0,9977	0,9978	0,9979	0,9979	0,9980	0,9981
2,9	0,9981	0,9982	0,9982	0,9983	0,9984	0,9984	0,9985	0,9985	0,9986	0,9986
3,0	0,9987	0,9987	0,9987	0,9988	0,9988	0,9989	0,9989	0,9989	0,9990	0,9990
3,1	0,9990	0,9991	0,9991	0,9991	0,9992	0,9992	0,9992	0,9992	0,9993	0,9993
3,2	0,9993	0,9993	0,9994	0,9994	0,9994	0,9994	0,9994	0,9995	0,9995	0,9995
3,3	0,9995	0,9995	0,9995	0,9996	0,9996	0,9996	0,9996	0,9996	0,9996	0,9997
3,4	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9998
3,5	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998
3,6	0,9998	0,9998	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999
3,7	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999
3,8	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999
3,9	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000
4,0	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000

Para abscissas maiores que 4,09, use a probabilidade 1,0000.

Bibliografia

Bussab, W. O. e Morettin, P. A. (2002) *Estatística Básica*. Editora Saraiva, 5a. ed.

Casella, G. e Berger, R. L. (1990) *Statistical Inference*. Duxbury Press, 1a. ed.

DeGroot, M. H. e Schervish, M. J. (2012) *Probability and Statistics*. Pearson, 4a. ed.

James, B. R. (2004) *Probabilidade: Um curso em nível intermediário*. IMPA, 3a. ed.

Magalhães, M. N. (2011) *Probabilidade e Variáveis Aleatórias*. Edusp.

Meyer, P. L. (2011) *Probabilidade: Aplicações a Estatística*. Editora LTC.

Rohatgi, V. K. (2008) *An Introduction To Probability And Statistics*. Wiley, 2a. ed.

Swokowski, E. W. (1979) *Calculus with Analytic Geometry*. Prindle, Weber & Schmidt, 2a. ed.

Índice Remissivo