Universidade Federal de Santa Maria

Centro de Ciências Naturais e Exatas

Departamento de Estatística

Notas de Aula

STC1071 - Probabilidade

Prof. Dr. Cleber Bisognin

Santa Maria - Maio/2021

Sumário

Capítulo 1

Introdução

O que é Estatística?

- Conjunto de métodos especialmente apropriado ao tratamento de dados numéricos, afetados por uma multiplicidade de causas;
- Estes métodos fazem uso da Matemática, e especialmente do cálculo de probabilidades.

Como Ciência:

- Permite organizar, descrever, analisar, e interpretar dados;
- Utiliza-se da Teoria da Probabilidade para modelar a aleatoriedade e a incerteza associada aos fenômenos naturais, econômicos, sociais, $\cdots$;
- Auxilia a tirar conclusões sobre as características das fontes de onde os dados foram retirados, para melhor compreende-los;
- Indispensável para a tomada de decisões sob condições de incerteza, sob o menor risco possível.

A estatística tem sido utilizada para:

- Otimização de recursos econômicos;
- Aumento da qualidade e produtividade;
- Análise de decisões judiciais;
- Previsões (climáticas, econômicas, $\cdots$)

Por que estudar Estatística?

- Impossibilidade de estudar a população;
- Aumento da capacidade de registro de dados que precisam ser compreendidos;
- Expansão do conhecimento científico, das áreas de pesquisa e dos instrumentos de investigação;
- Necessidade de compreensão dos fenômenos naturais e sociais, de otimização de recursos, planejamento de atividades, redução de riscos, de previsão de resultados para correta tomada de decisão.

A Estatística pode ser pensada como a **ciência de aprendizagem a partir dos dados**.

Áreas da Estatística

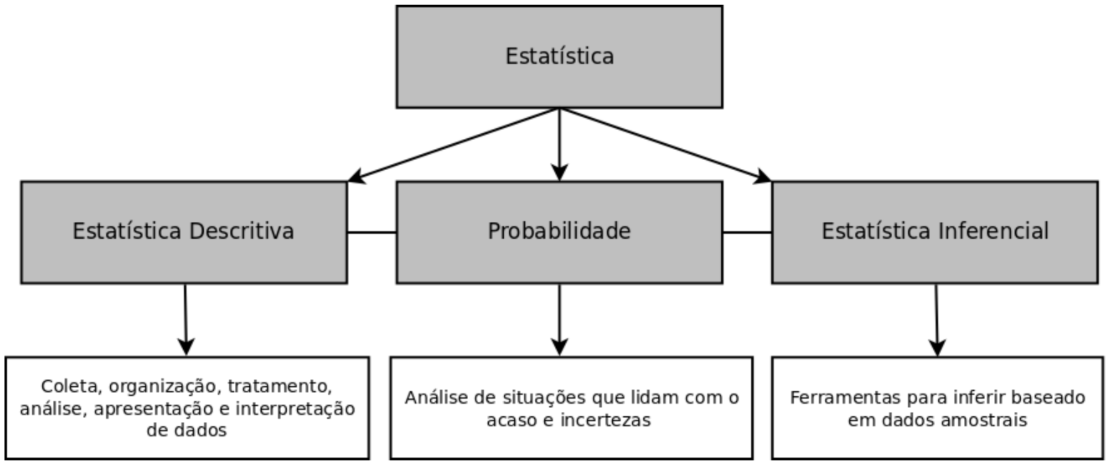


Figura 1.1: Organograma da Estatística.

Estatística Descritiva: etapa inicial de qualquer análise. É um conjunto de técnicas destinadas a descrever e resumir os dados, que auxiliam a descrever características de interesse.
$\Longrightarrow$ *Conheça seus dados*

Probabilidade: é a ferramenta matemática utilizada pela Estatística para se estudar a incerteza oriunda de fenômenos aleatórios.
$\Longrightarrow$ *Qual a incerteza associada aos dados?*

Estatística Inferencial: é um conjunto de técnicas que possibilita tirar conclusões sobre uma população, a partir de um subconjunto de valores (amostra).
$\Longrightarrow$ *Quais conclusões podemos tirar a partir destes dados?*

Fonte: `http://bit.do/eG5U7` e `http://bit.do/eG5Ve`.

Capítulo 2

Probabilidade

2.1 Introdução

Um sistema de comunicação formado por n antenas aparentemente idênticas que deve ser alinhadas em sequência. O sistema resultante será capaz de receber qualquer sinal - e estará funcionando - desde que duas antenas consecutivas não apresentem defeito. Se exatamente m das n antenas apresentam defeito, qual será a probabilidade de que o sistema esteja funcionando. Por exemplo, considerando $n = 4$ e $m = 2$, há seis configurações possíveis para o sistema, a saber,

$$\begin{array}{cccc} 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 \end{array}$$

onde 1 significa que a antena funciona e 0, que ela está com defeito. O sistema funciona nos 3 primeiros arranjos e não funciona nos 3 últimos. Assim parece razoável tomar $3/6 = 0,5$ como a probabilidade desejada. No caso de m e n quaisquer, podemos calcular de forma similar a probabilidade do sistema funcionar. Ou seja, podemos contar o número de configurações que resultam no funcionamento do sistema e dividir pelo número total de configurações possíveis.

Muitos problemas da teoria da probabilidade poderiam ser resolvidos contando-se o número de diferentes maneiras pelas quais certo evento pode ocorrer. Na teoria matemática da contagem é formalmente conhecida como *análise combinatória*.

2.2 Princípio da Contagem

Dito de forma simples, o Princípio Básico da Contagem diz que se um experimento pode levar a qualquer um dos m possíveis resultados e se outro experimento pode resultar em qualquer um dos n possíveis resultados, então os dois experimentos possuem mn resultados possíveis.

Definição 2.1. [Princípio Básico da Contagem] Suponha a realização de dois experimentos. Se o experimento 1 pode gerar qualquer um dos m resultados possíveis e se, para cada um dos resultados do experimento 1, houver n resultados possíveis para o experimento 2, então os dois experimentos possuem conjuntamente mn diferentes resultados possíveis.

Podemos verificar o resultado enumerando todos os possíveis resultados dos dois experimentos, ou seja,

$$\begin{array}{cccc}
(1,1) & (1,2) & \cdots & (1,n) \\
(2,1) & (2,2) & \cdots & (2,n) \\
\vdots & & & \\
(m,1) & (m,2) & \cdots & (m,n)
\end{array}$$

onde o resultado é (i,j) se o experimento 1 resultar ao seu i-ésimo resultado possível e o experimento 2 levar ao j-ésimo resultado possível. Portanto, o conjunto dos resultados possíveis consiste em m linhas, cada uma contendo n elementos. Logo temos mn possíveis resultados.

Exemplo 2.1. Uma pequena comunidade é composta por 10 mulheres, cada uma com 3 filhos. Se uma mulher e um dos seus filhos devem ser escolhidos como mãe e filho do ano, quantas escolhas diferentes são possíveis?

Solução: Supondo a escolha da mulher como o resultado do primeiro experimento, e subsequente escolha de um dos seus filhos como resultado do segundo experimento, temos $10 \times 3 = 30$ escolhas possíveis.

□

Havendo mais de 2 experimentos a serem realizados, podemos generalizar o Princípio Básico da Contagem.

Definição 2.2. [Generalização do Princípio Básico da Contagem] Se r experimentos são tais que o primeiro experimento pode levar a qualquer um dos n_1 resultados possíveis; e se, para cada um dos n_1 resultados houver n_2 resultados possíveis para o segundo experimento; e se, para cada um dos possíveis resultados dos dois primeiros experimentos houver n_3 resultados possíveis para o terceiro experimento; e se $\cdots$, então haverá um total de $n_1 \cdot n_2 \cdots n_r$ resultados possíveis para os r experimentos.

Exemplo 2.2. Quantas diferentes placas de automóveis com 7 caracteres são possíveis se os três primeiros campos forem ocupados por letras e os 4 campos finais por números?

Solução: Sabemos que o alfabeto latino possui 26 letras. Pelo versão generalizada do Princípio Básico da Contagem, a resposta é

$$26 \times 26 \times 26 \times 10 \times 10 \times 10 \times 10 = 175.760.000.$$

□

Exemplo 2.3. Quantas funções definidas em n pontos são possíveis se cada valor da função for igual a 0 ou 1?

Solução: Suponha que os pontos sejam $1, 2, \cdots, n$. Como $f(i)$ deve ser igual a 0 ou 1 para cada $i = 1, 2, \cdots, n$, tem-se 2^n possíveis funções.

□

2.2.1. Permutações

Quantos diferentes arranjos ordenados das letras a, b e c são possíveis? Por enumeração direta vemos que são 6, ou seja,

$$abc, acb, bac, bca, cab, cba$$

Cada combinação é conhecida como uma permutação. Assim, vê-se que o conjunto de 3 objetos permite 6 permutações possíveis. Assim, pelo Princípio Básico da Contagem, temos $3 \cdot 2 \cdot 1 = 6$.

Suponha que temos n objetos. Pelo mesmo raciocínio temos

$$n(n-1)(n-2)\cdots 3\cdot 2\cdot 1 = n!$$

permutações diferentes dos n objetos.

Exemplo 2.4. Quantas diferentes ordens de rebatedores são possíveis em um time de beisebol formado por 9 jogadores?

Solução: Há $9! = 362.880$ ordens de rebatedores possíveis.

□

Exemplo 2.5. Uma turma de Teoria da Probabilidade é formada por 6 homens e 4 mulheres. Aplica-se uma prova e os estudantes são classificados de acordo com o seu desempenho. Suponha que nenhum dos estudantes tenha tirado a mesma nota.

(a) Quantas diferentes classificações são possíveis?

(b) Se os homens forem classificados apenas entre si e as mulheres entre si, quantas diferentes classificações são possíveis?

Solução:

(a) Como cada classificação corresponde a um arranjo particular das 10 pessoas, a resposta é $10! = 3.628.800$.

(b) Como há 6! possíveis classificações dos homens entre si e 4! classificações possíveis das mulheres entre si, segue do Princípio Básico da Contagem que há $(6!)(4!) = (720)(24) = 17.280$ classificações possíveis neste caso.

□

Exemplo 2.6. Quantos diferentes arranjos de letras podem ser formados a partir da palavra PEPPER?

Solução: Primeiramente notamos que as letras $P_1E_1P_2P_3E_2R$ permitem 6! permutações quando 3 P's e os 2 E's são diferentes uns dos outros. Entretanto, considere uma destas permutações - por exemplo, $P_1P_2E_1P_3E_2R$. Se agora permutarmos os P's e os E's entre si, então o arranjo resultante continuará a ser PPEPER. Isto é, todas as 3!2! permutações

$$\begin{array}{ll}
P_1P_2E_1P_3E_2R & P_1P_2E_2P_3E_1R \\
P_1P_3E_1P_2E_2R & P_1P_3E_2P_2E_1R \\
P_2P_1E_1P_3E_2R & P_2P_1E_2P_3E_1R \\
P_2P_3E_1P_1E_2R & P_2P_3E_2P_1E_1R \\
P_3P_1E_1P_2E_2R & P_3P_1E_2P_2E_1R \\
P_3P_2E_1P_1E_2R & P_3P_2E_2P_1E_1R
\end{array}$$

são da forma PPEPER. Portanto, há $6!/(3!2!) = 60$ arranjos possíveis das letras PEPPER.

□

Generalizando temos

$$\frac{n!}{n_1!n_2!\cdots n_r!}$$

permutações diferentes de n objetos, dos quais n_1 são iguais, n_2 são iguais, $\cdots$, n_r são iguais.

Exemplo 2.7. Um torneio de xadrez tem 10 competidores, dos quais 4 são russos, 3 dos estados Unidos, 2 da Inglaterra e 1 do Brasil. Se o resultado do torneio listar apenas a nacionalidade dos jogadores em sua ordem de colocação, quantos resultados serão possíveis?

Solução: Há

$$\frac{10!}{4!3!2!1!} = 12.600$$

□

2.2.2. Combinações

Estamos frequentemente interessados em determinar o número de grupos diferentes de r objetos que podem ser formados a partir de um total de n objetos. Por exemplo, quantos diferentes grupos de 3 podem ser selecionados dos 5 itens A,B,C,D e E? Temos 5 maneiras diferentes de selecionar o primeiro item, 4 de selecionar o segundo item e 3 de selecionar o item final, há portanto, $5 \cdot 4 \cdot 3$ maneiras de selecionar o grupo de 3 quando a ordem de seleção dos itens for relevante. Entretanto, cada grupo de 3 - por exemplo, o grupo formado pelos itens A, B e C - será contado 6 vezes (isto é, todas as permutações ABC, ACB, BAC, BCA, CAB e CBA serão contadas quando a ordem da seleção for relevante), tem-se que o número total de grupos que podem ser formados é igual a

$$\frac{5 \cdot 4 \cdot 3}{3 \cdot 2 \cdot 1} = 10.$$

Em geral, como $n(n-1)\cdots(n-r+1)$ representa o número de diferentes maneiras pelas quais um grupo de r itens pode ser selecionado a partir de n itens quando a ordem de seleção é relevante, e como cada grupo de r itens será considerado $r!$ vezes, tem-se que o número de grupos diferentes de r itens que podem ser formados a partir de um conjunto de n itens é

$$\frac{n(n-1)\cdots(n-r+1)}{r!} = \frac{n!}{(n-r)!r!}$$

Notação e Terminologia: Definimos $\binom{n}{r}$, para $r \leqslant n$, como

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

E dizemos que $\binom{n}{r}$ representa o número de combinações possíveis de n objetos em grupos de r elementos de cada vez.

Por convenção, define-se $0! = 1$. Com isso, $\binom{n}{0} = \binom{n}{n} = 1$. Além disso, assume-se que $\binom{n}{i} = 0$ quando $i < 0$ ou $i > n$.

Assim, $\binom{n}{r}$ representa o número de grupos diferentes com r elementos que podem ser selecionados de um conjunto de n objetos quando a ordem da seleção não é considerada relevante.

Exemplo 2.8. Um comitê de 3 pessoas deve ser formado a partir de um grupo de 20 pessoas. Quantos comitês diferentes são possíveis?

Solução: Há

$$\binom{20}{3} = \frac{20!}{(20-3)!3!} = \frac{20 \cdot 19 \cdot 18}{3 \cdot 2 \cdot 1} = 1140$$

possíveis comitês.

$\square$

Exemplo 2.9. De um grupo de 5 mulheres e 7 homens, quantos comitês diferentes formados por 2 mulheres e 3 homens podem ser formados? E se dois dos homens estiverem brigados e se recusarem a trabalhar juntos?

Solução: Como há $\binom{5}{2}$ grupos possíveis de mulheres e $\binom{7}{3}$ grupos possíveis de homens, temos que há $\binom{5}{2}\binom{7}{3} = \left(\frac{5 \cdot 4}{2 \cdot 1}\right)\left(\frac{7 \cdot 6 \cdot 5}{3 \cdot 2 \cdot 1}\right) = 350$ comitês possíveis formados por 2 mulheres e 3 homens.

Suponha agora que 2 dos homens recusem a trabalhar juntos. Como um total de $\binom{2}{2}\binom{5}{1} = 5$ dos $\binom{7}{3} = 35$ grupos possíveis de 3 homens contém os dois homens brigados, tem-se $35 - 5 = 30$ grupos não contendo os homens brigados. Como ainda $\binom{5}{2} = 10$ maneiras de escolher as 2 mulheres, há $30 \cdot 10 = 300$ comitês possíveis neste caso.

$\square$

Na próxima seção iremos definir o conceito de Probabilidade. O termo Probabilidade refere-se ao estudo da aleatoriedade e da incerteza.

2.3 Definições Básicas

Definição 2.3. *Modelo Determinístico:* Neste modelo as condições sob as quais o experimento é executado, determinam o resultado do experimento.

Exemplo 2.10. A Lei de Ohm, V = I.R. Se R e I forem conhecidos, então V estará precisamente determinado. Um litro de água pura ferve sempre a 100^0C.

$\square$

Definição 2.4. *Modelo Não - Determinístico ou Probabilístico:* É um modelo em que de antemão não é possível explicitar ou definir um resultado particular. Este modelo é especificado através de uma distribuição de probabilidade. É utilizado quando se tem um grande número de variáveis influenciando o resultado e estas variáveis não podem ser controladas.

Exemplo 2.11. O lançamento de um dado onde se tenta prever o número da face que irá sair, a retirada de uma carta de um baralho, etc.

$\square$

O Modelo Estocástico é caracterizado como um modelo probabilístico que depende ou varia com o tempo.

Definição 2.5. *Experimento Aleatório (Não - Determinístico):* É qualquer atividade ou processo cujo resultado está sujeito a incerteza. Os exemplos dados são de fenômenos para os quais os modelos probabilísticos são adequados e que por simplicidade, são denominados de experimentos aleatórios.

Definição 2.6. *Espaço Amostral:* o espaço amostral de um experimento, denotado por Ω, é o conjunto de todos os resultados possíveis do experimento.

Assim, $\Omega = \{w_1, w_2, \cdots, w_n, \cdots\}$, onde cada w_i é um possível resultado do experimento aleatório.

Exemplo 2.12. Exemplos de Espaço Amostral:

(i) Analisar se um fusível está funcionando ou com defeito. $\Omega = \{N, D\}$, onde N representa sem defeito e D com defeito.

(ii) Jogar um percevejo e observar se a ponta cai para cima (C) ou para abaixo (B). $\Omega = \{C, B\}$.

(iii) Observar se o sexo da próxima criança nascida no HUSM. $\Omega = \{F, M\}$.

□

Exemplo 2.13. Se examinarmos três fusíveis em sequência e anotarmos o resultado de cada exame, o resultado é qualquer sequência de N e D de 3 elementos.

$$\Omega = \{NNN, NND, NDN, DNN, DND, DDN, NDD, DDD\}.$$

□

Exemplo 2.14. Dois postos de gasolina estão localizados em um determinado cruzamento. Cada um possui 6 bombas. Considere o experimento em que o número de bombas em uso em determinada hora do dia é determinado por cada posto. O espaço amostral Ω do experimento é dado a seguir.

		Segundo Posto						
		0	1	2	3	4	5	6
	0	(0,0)	(0,1)	(0,2)	(0,3)	(0,4)	(0,5)	(0,6)
	1	(1,0)	(1,1)	(1,2)	(1,3)	(1,4)	(1,5)	(1,6)
	2	(2,0)	(2,1)	(2,2)	(2,3)	(2,4)	(2,5)	(2,6)
Primeiro Posto	3	(3,0)	(3,1)	(3,2)	(3,3)	(3,4)	(3,5)	(3,6)
	4	(4,0)	(4,1)	(4,2)	(4,3)	(4,4)	(4,5)	(4,6)
	5	(5,0)	(5,1)	(5,2)	(5,3)	(5,4)	(5,5)	(5,6)
	6	(6,0)	(6,1)	(6,2)	(6,3)	(6,4)	(6,5)	(6,6)

□

Exemplo 2.15. O espaço amostral Ω de um experimento em que um dado é lançado duas vezes é dados a seguir.

(1,1)	(1,2)	(1,3)	(1,4)	(1,5)	(1,6)
(2,1)	(2,2)	(2,3)	(2,4)	(2,5)	(2,6)
(3,1)	(3,2)	(3,3)	(3,4)	(3,5)	(3,6)
(4,1)	(4,2)	(4,3)	(4,4)	(4,5)	(4,6)
(5,1)	(5,2)	(5,3)	(5,4)	(5,5)	(5,6)
(6,1)	(6,2)	(6,3)	(6,4)	(6,5)	(6,6)

□

Observação: Um espaço amostral pode ser classificado em:

(i) Finito: exemplos anteriores.

(ii) Infinito:

(a) Enumeráveis (ou contáveis): $\{1, 2, 3, 4, 5, \cdots\}$;

(b) Não-enumeráveis (ou não contáveis): $\{t \in \mathbb{R}; t \geqslant 0\}$.

2.3.1. Eventos

No estudo de probabilidade, estamos interessados não apenas nos resultados individuais de Ω, como também em qualquer grupo de resultados de Ω.

Definição 2.7. Um *evento* A é qualquer grupo (subconjunto) de resultados contidos no espaço amostral Ω. Um evento é simples se ele consistir em exatamente em um único resultado e composto se ele consistir em mais de um resultado. ($A \subseteq \Omega$).

Eventos Especiais:

(1) Evento Certo: é aquele evento que ocorre toda vez que se realiza o experimento, portanto, esse evento é o próprio Ω.

$$\{(x, y) \in \mathbb{R}^2 | x^2 + y^2 \geqslant 0\}.$$

(2) Evento Impossível: é aquele evento que nunca irá ocorrer, é também conhecido como o conjunto vazio ($\varnothing$). ($\varnothing \in \Omega$).

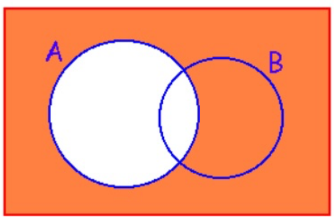
$$\{(x, y) \in \mathbb{R}^2 | x^2 + y^2 < 0\} = \varnothing.$$

(3) Evento Elementar: é qualquer resultado particular de um experimento aleatório. $A = \{a\}$.

2.3.2. Operações com Conjuntos:

(1) O complemento de um conjunto A, denotado por A^c ou A' ou $\bar{A}$, é o conjunto de todos os elementos de Ω que não são elementos de A, ou seja,

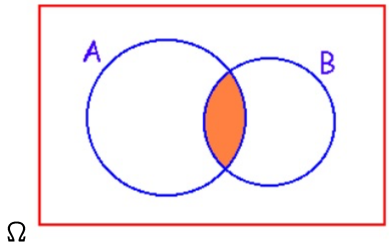
$$A^c = \{x \in \mathbb{R} | x \in \Omega \ e \ x \notin A\}.$$



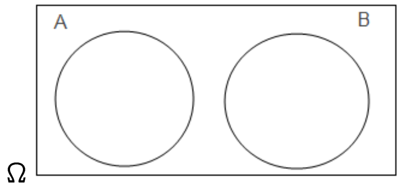
Ω

(2) A intersecção dos conjuntos A e B, denotada por $A \cap B$, é o conjunto formado pelos elementos comuns de A e B, ou seja,

$$A \cap B = \{x \in \mathbb{R} | x \in A \ e \ x \in B\}.$$

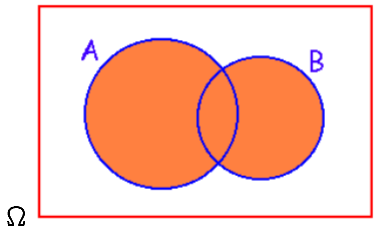


(3) Dois conjuntos A e B que não possuem elementos em comum, isto é, tais que $A \cap B = \varnothing$ são denominados *conjuntos disjuntos*.



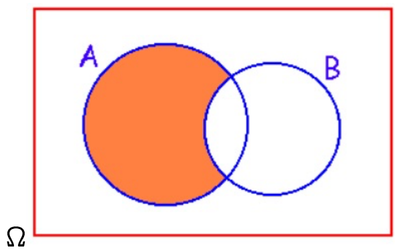
(4) A união de dois conjuntos A e B, denotado por $A \cup B$, é o conjunto de todos os elementos pertencentes tanto a A quanto a B, ou seja,

$$A \cup B = \{x \in \mathbb{R} | x \in A \text{ ou } x \in B\}.$$



(5) A diferença entre os conjuntos A e B, denotada por $A - B$, é o conjunto de todos os elementos que pertencem ao conjunto A e não ao B, ou seja,

$$A - B = \{x \in \mathbb{R} | x \in A \text{ e } x \notin B\} = A \cap B^c.$$



Observação 2.1. Ao escrever um conjunto que contém vários elementos, a ordem em que os elementos aparecem não é relevante. Por exemplo, $\{5, 1\} = \{1, 5\}$. No entanto, existem muitas situações na Matemática onde a ordem de dois ou mais objetos é importante. Isto leva a ideia de par ordenado. Quando escrever um par ordenado use parênteses ao invés de chaves que são reservadas para escrever conjuntos.

2.3.3. Aplicações dos Diagramas de Venn

Os diagramas de Venn podem ser usados para ilustrar propriedades das operações entre conjuntos. Por exemplo, verificar que a operação entre conjuntos $A - B$ é igual a $A \cap B^c$.

Outras propriedades que podem ser verificadas através dos diagramas são:

Leis de Morgan:

(1) $(A \cap B)^c = A^c \cup B^c$

(2) $(A \cup B)^c = A^c \cap B^c$

Propriedade comutativa:

(3) $A \cup B = B \cup A$

(4) $A \cap B = B \cap A$

Propriedade associativa:

(5) $(A \cup B) \cup C = A \cup (B \cup C)$

(6) $(A \cap B) \cap C = A \cap (B \cap C)$

Propriedade distributiva:

(7) $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$

(8) $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$

Propriedades da identidade:

(9) $A \cup \varnothing = A$

(10) $A \cap \Omega = A$

Exemplo 2.16. (Continuação do Exemplo 2.15):

Para o experimento em que é observado o número de bombas em uso em um posto de gasolina de 6 bombas, considere: $A = \{0, 1, 2, 3, 4\}$, $B = \{3, 4, 5, 6\}$ e $C = \{1, 3, 5\}$. Então:

$$A^c = \{5, 6\}, \quad A \cup B = \{0, 1, 2, 3, 4, 5, 6\} = \Omega, \quad A \cup C = \{0, 1, 2, 3, 4, 5\},$$

$$A \cap B = \{3, 4\}, \quad A \cap C = \{1, 3\}, \quad (A \cap C)^c = \{0, 2, 4, 5, 6\}.$$

□

2.3.4. Conceitos de Probabilidade

Definição Clássica

Definição 2.8. Seja E um experimento aleatório e Ω o espaço amostral associado formado por n resultados igualmente prováveis. Seja $A \subseteq \Omega$ um evento com m elementos. A Probabilidade de A, denotada por $\mathbb{P}(A)$, é definida como:

$$\mathbb{P}(A) = \frac{m}{n}.$$

Ou seja, a probabilidade do evento A é o quociente entre o número m de casos favoráveis e o número n de casos possíveis.

Problema: Não pode ser aplicada quando o espaço amostral é infinito.

Distribuição de Frequência

Quando se estuda uma variável, o maior interesse é conhecer o comportamento dessa variável, analisando a ocorrência de suas possíveis realizações.

Tabela 2.1: Informações sobre Estado Civil, grau de instrução, número de filhos, salário(expresso como fração de salário mínimo), idade(medida em anos e meses) e procedência de 36 empregados de uma certa empresa.

N	EstadoCivil	Grau de Instrução	Nº de Filhos	Salário($\times$ Sal. Mín.)	Idade(anos)	Idade(meses)	Região
1	Solteiro	Ensino Fundamental	0	4,00	26	3	Interior
2	Casado	Ensino Fundamental	1	4,56	32	10	Capital
3	Casado	Ensino Fundamental	2	5,25	36	5	Capital
4	Solteiro	Ensino Médio	0	5,73	20	10	Outra
5	Solteiro	Ensino Fundamental	0	6,26	40	7	Outra
6	Casado	Ensino Fundamental	0	6,66	28	0	Interior
7	Solteiro	Ensino Fundamental	0	6,86	41	0	Interior
8	Solteiro	Ensino Fundamental	0	7,39	43	4	Capital
9	Casado	Ensino Médio	1	7,59	34	10	Capital
10	Solteiro	Ensino Médio	0	7,44	23	6	Outra
11	Casado	Ensino Médio	2	8,12	33	6	Interior
12	Solteiro	Ensino Fundamental	0	8,46	27	11	Capital
13	Solteiro	Ensino Médio	0	8,74	37	5	Outra
14	Casado	Ensino Fundamental	3	8,95	44	2	Outra
15	Casado	Ensino Médio	0	9,13	30	5	Interior
16	Solteiro	Ensino Médio	0	9,35	38	8	Outra
17	Casado	Ensino Médio	1	9,77	31	7	Capital
18	Casado	Ensino Fundamental	2	9,80	39	7	Outra
19	Solteiro	Superior	0	10,53	25	8	Interior
20	Solteiro	Ensino Médio	0	10,76	37	4	Interior
21	Casado	Ensino Médio	1	11,06	30	9	Outra
22	Solteiro	Ensino Médio	0	11,59	34	2	Capital
23	Solteiro	Ensino Fundamental	0	12,00	41	0	Outra
24	Casado	Superior	0	12,79	26	1	Outra
25	Casado	Ensino Médio	2	13,23	32	5	Interior
26	Casado	Ensino Médio	2	13,60	35	0	Outra
27	Solteiro	Ensino Fundamental	0	13,85	46	7	Outra
28	Casado	Ensino Médio	0	14,69	29	8	Interior
29	Casado	Ensino Médio	5	14,71	40	6	Interior
30	Casado	Ensino Médio	2	15,99	35	10	Capital
31	Solteiro	Superior	0	16,22	31	5	Outra
32	Casado	Ensino Médio	1	16,61	36	4	Interior
33	Casado	Superior	3	17,26	43	7	Capital
34	Solteiro	Superior	0	18,75	33	7	Capital
35	Casado	Ensino Médio	2	19,40	48	11	Capital
36	Casado	Superior	3	23,30	42	2	Interior

Fonte: Bussab e Morettin (2006).

Exemplo 2.17. A Tabela 2.2 apresenta a *distribuição de frequência* da variável grau de instrução, utilizando os dados apresentados na Tabela 2.1.

Tabela 2.2: Frequências e porcentagens dos 36 empregados de uma certa empresa segundo o grau de instrução.

Grau de Instrução	Frequência Absoluta (f)	Frequência Relativa (f_{re})	Porcentagem ($100 \times f_{re}$)
Ensino Fundamental	12	0,33	33,33
Ensino Médio	18	0,50	50,00
Superior	6	0,17	16,67
Total	36	1,00	100,00

Fonte: Bussab e Morettin (2006).

Observando os resultados, percebemos que dos 36 empregados, 12 têm ensino fundamental, 18 ensino médio e 6 superior.

Uma medida bastante útil na interpretação de tabelas de frequência é a proporção de cada uma das realizações em relação ao total. Assim, $6/36 = 0,17$ dos empregados tem instrução superior. Na última coluna da Tabela 2.2 são apresentadas as porcentagens para cada realização da variável grau de instrução. Como notação usaremos f para indicar *frequência* (absoluta) de cada classe, ou categoria, da variável, e a notação $f_{re} = f/n$ para indicar *proporção* (ou *frequência relativa*) de cada classe, sendo n o número total de observações. As proporções são muito mais úteis para comprar resultados de duas pesquisas distintas.

A construção de tabelas de frequência (ver Seção ??) para variáveis contínuas necessita de certo cuidado. Por exemplo, a construção da tabela de frequência para a variável salário, utilizando o procedimento acima não resumirá as 36 observações num grupo menor, pois não existem observações iguais.A solução empregada é agrupar os dados por faixas de salário.

A Tabela 2.3 fornece a distribuição de frequência dos salários dos ,6 empregados por faixa de salário.

Tabela 2.3: Frequências e porcentagens dos 36 empregados por faixa de salário.

Classes de Salário	Frequência Absoluta(f)	Frequência Relativa(f_{re})	Porcentagem ($100\times f_{re}$)
4,00|- 8,00	10	0,28	27,78
8,00|-12,00	12	0,33	33,33
12,00|-16,00	8	0,22	22,22
16,00|-20,00	5	0,14	13,89
20,00|-24,00	1	0,03	2,78
Total	36	1,00	100,00

Fonte: Bussab e Morettin (2006).

Note que estamos utilizando a notação $a|\text{-}b$ para o intervalo contendo o extremo a mas não contendo o extremo b. Podemos utilizar também a notação $[a, b)$ para designar o mesmo intervalo.

Frequência Relativa

Na prática acontece que nem sempre é possível determinar a probabilidade de um evento. Neste caso é necessário ter um método de aproximação desta probabilidade. Um dos métodos utilizados é a experimentação que objetiva estimar o valor da probabilidade de um evento A com base em valores reais. A probabilidade avaliada através deste processo é denominada de probabilidade empírica.

Definição 2.9. Seja E um experimento e A um evento de um espaço amostra associado ao experimento E. Suponha-se que E seja repetido n vezes e seja m o número de vezes que A ocorre nas n repetições de E. Então a frequência relativa do evento A, anotada por $f_{re}(A)$, é o quociente:

$$f_{re}(A) = \frac{m}{n} = \frac{\text{Número de vezes que A ocorre}}{\text{Número de vezes que E é repetido}}$$

Exemplo 2.18. Exemplos de probabilidade relativa:

(1) Uma moeda é lançada 200 vezes e forneceu 102 caras. Então a frequência relativa de caras é

$$\mathbb{P}(caras) = f_{re}(caras) = \frac{102}{200} = 0.51.$$

(2) Um dado foi lançado 100 vezes e a face 6 apareceu 18 vezes. Então a frequência relativa do evento $A = \{\text{face } 6\}$ é:

$$\mathbb{P}(A) = f_{re}(A) = \frac{18}{100} = 0.18.$$

□

Propriedades da frequência relativa

(1) $0 \leqslant f_{re}(A) \leqslant 1$, isto é, a frequência relativa do evento A é um número que varia entre 0 e 1.

(2) $f_{re}(A) = 1$, se e somente se, A ocorre em todas as n repetições de E.

(3) $f_{re}(A) = 0$, se e somente se, A nunca ocorre nas n repetições de E.

(4) $f_{re}(A \cup B) = f_{re}(A) + f_{re}(B)$ se A e B forem eventos mutuamente excludentes.

Definição de Probabilidade como Frequência Relativa

Definição 2.10. Seja E um experimento e A um evento de um espaço amostra associado Ω. Suponhamos que E é repetido n vezes e seja $f_{re}(A)$ a frequência relativa do evento. Então a probabilidade de A é definida como sendo o limite de $f_{re}(A)$ quando n tende ao infinito. Ou seja:

$$\mathbb{P}(A) = \lim_{n\to\infty} f_{re}(A)$$

Deve-se notar que a frequência relativa do evento A é uma aproximação da probabilidade de A. As duas se igualam apenas no limite. Em geral, para um valor de n, razoavelmente grande a $f_{re}(A)$ é uma boa aproximação de $\mathbb{P}(A)$.

Problema: Esta definição, embora útil na prática, apresenta dificuldades matemáticas, pois o limite pode não existir. Em virtude dos problemas apresentados pela definição clássica e pela definição frequencial, foi desenvolvida uma teoria moderna, na qual a probabilidade é um conceito indefinido, como o ponto e a reta o são na geometria.

Definição Axiomática de Probabilidade

Definição 2.11. Seja E um experimento aleatório com um espaço amostra associado Ω. A cada evento A, tal que $A \subseteq \Omega$ associa-se um número real, representado por $\mathbb{P}(A)$ que satisfaz as seguintes propriedades (axiomas):

(i) $0 \leqslant \mathbb{P}(A) \leqslant 1$

(ii) $\mathbb{P}(\Omega) = 1$

(iii) $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B)$, se A e B forem eventos mutuamente excludentes.

(iv) Se $A_1, A_2, A_3, \cdots, A_n, \cdots$, forem, dois a dois, eventos mutuamente excludentes, então

$$\mathbb{P}\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} \mathbb{P}(A_i).$$

Propriedades:

(i) $\mathbb{P}(\varnothing) = 0$

(ii) $\mathbb{P}(A) = 1 - \mathbb{P}(A^c)$

(iii) Se $A \subseteq B$, então $\mathbb{P}(A) \leqslant \mathbb{P}(B)$

(iv) Se A e B são dois eventos quaisquer, então

$$\mathbb{P}(A - B) = \mathbb{P}(A \cap B^c) = \mathbb{P}(A) - \mathbb{P}(A \cap B).$$

(v) Se A e B são dois eventos quaisquer de Ω, então

$$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B).$$

(vi) Se A, B e C são três eventos quaisquer de Ω, então

$$\mathbb{P}(A \cup B \cup C) = \mathbb{P}(A) + \mathbb{P}(B) + \mathbb{P}(C) - \mathbb{P}(A \cap B) - \mathbb{P}(A \cap C) - \mathbb{P}(B \cap C) + \mathbb{P}(A \cap B \cap C).$$

Exemplo 2.19. Considere testar baterias que saem de uma linha de montagem, uma por uma, até ser encontrada uma que tenha voltagem dentro dos limites prescritos. Os eventos simples são $E_1 = \{S\}$, $E_2 = \{FS\}$, $E_3 = \{FFS\}$, $E_4 = \{FFFS\}$, $\cdots$. Suponha que a probabilidade de qualquer bateria particular ser satisfatória seja 0.99. Então, pode-se mostrar que a $\mathbb{P}(E_1) = 0.99$, $\mathbb{P}(E_2) = 0.01 \times 0.99$, $\mathbb{P}(E_3) = 0.01 \times 0.01 \times 0.99 = (0.01)^2(0.99)$, $\mathbb{P}(E_4) = (0.01)^3(0.99)$, $\cdots$ é uma distribuição de probabilidade aos eventos simples que satisfaz os axiomas. Em particular, pelo fato de os E_i serem disjuntos e $\Omega = E_1 \cup E_2 \cup E_3 \cup \cdots$, temos

$$\begin{aligned} 1 &= \mathbb{P}(\Omega) = \mathbb{P}(E_1) + \mathbb{P}(E_2) + \mathbb{P}(E_3) + \cdots \\ &= 0.99[1 + 0.01 + (0.01)^2 + (0.01)^3 + \cdots] \end{aligned}$$

Pela soma de uma série geométrica, temos que

$$a + ar + ar^2 + ar^3 + \cdots = \frac{a}{1-r}.$$

□

Exemplo 2.20. Considere um sistema com 5 componentes idênticos ligados em série, conforme a Figura 2.1.

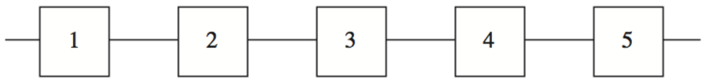


Figura 2.1: Sistema com 5 componentes idênticos ligados em série.

Representamos um componente que falha por F e o que não falha por S. Representamos por A o evento em que o sistema falha. Para que A ocorra, ao menos um dos componentes individuais deve falhar. Assim,

$$A = \{FSSSS, SFSSS, SSFSS, \cdots\}.$$

Há, na verdade, 32 resultados. O evento A compreende 31 resultados. O evento $A^c = \{SSSSS\}$, ou seja, o sistema funciona. Se considerarmos $\mathbb{P}(S) = 0.9$, temos que, $\mathbb{P}(A^c) = \mathbb{P}(\{SSSSS\}) = (0.9)^5 = 0.59$. Portanto, $\mathbb{P}(A) = 1 - \mathbb{P}(A^c) = 1 - 0.59 = 0.41$.

□

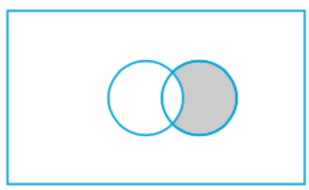
Exemplo 2.21. Em um certo residencial, 60% de todas as residências recebem internet da companhia a cabo local, 80% recebem o serviço de televisão desta companhia e 50% recebem ambos os serviços. Se uma residência for selecionada aleatoriamente, qual a probabilidade de ela receber ao menos um desses dois serviços da companhia e qual é a probabilidade de ela receber exatamente um desses serviços da companhia?

Defina: A={recebe serviço de internet}, B={recebe serviço de TV}. Assim, $\mathbb{P}(A) = 0.6$, $\mathbb{P}(B) = 0.8$ e $\mathbb{P}(A \cap B) = 0.5$.

$$\mathbb{P}(\text{Receber ao menos um dos serviços}) =$$

$$= \mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B) = 0.6 + 0.8 - 0.5 = 0.9.$$

O evento em que a residência recebe apenas o serviço de TV pode ser definido por $A^c \cap B$, ou seja [(sem internet) e (com TV)]. Como vemos na figura a seguir



ou seja,

$$\mathbb{P}(A^c \cap B) = \mathbb{P}(A \cup B) - \mathbb{P}(A) = 0.9 - 0.6 = 0.3.$$

Tabela 2.4: Distribuição de alunos segundo o sexo e a escolha do curso.

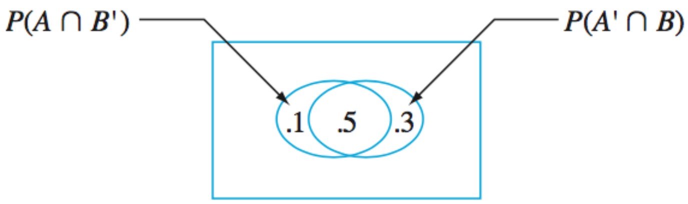
Curso	Homens (H)	Mulheres (M)	Total
Matemática Pura (P)	70	40	110
Matemática Aplicada(A)	15	15	30
Estatística (E)	10	20	30
Computação (C)	20	10	30
Total	115	85	200

Analogamente, o evento em que a residência recebe apenas o serviço de internet pode ser definido por $A \cap B^c$, ou seja [(com internet) e (sem TV)].

$$\mathbb{P}(A \cap B^c) = \mathbb{P}(A \cup B) - \mathbb{P}(B) = 0.9 - 0.8 = 0.1.$$

Logo,

$$\mathbb{P}(\text{Receber exatamente um dos serviços}) = \mathbb{P}(A^c \cap B) + \mathbb{P}(A \cap B^c) = 0.4.$$



□

Exemplo 2.22. Na Tabela 2.4 temos dados referentes a alunos matriculados em quatro cursos de uma universidade em dado ano.

Algumas probabilidades: $\mathbb{P}(E) = \frac{30}{200}$, $\mathbb{P}(H) = \frac{115}{200}$, $\mathbb{P}(A \cap H) = \frac{15}{200}$.

Queremos calcular $\mathbb{P}(A \cup H)$.

$$\mathbb{P}(A \cup H) = \mathbb{P}(A) + \mathbb{P}(H) - \mathbb{P}(A \cap H) = \frac{30}{200} + \frac{115}{200} - \frac{15}{200} = \frac{130}{200}.$$

Considere os eventos A e C, temos que $\mathbb{P}(A) = \frac{30}{200}$ e $\mathbb{P}(C) = \frac{30}{200}$ e $\mathbb{P}(A \cup C) = \frac{60}{200} = \mathbb{P}(A) + \mathbb{P}(C)$. Neste caso, os eventos A e C são disjuntos ou mutuamente exclusivos, pois se A ocorre, então C não ocorre e vice-versa. Aqui, $A \cap C = \varnothing$ e $\mathbb{P}(A \cap C) = \mathbb{P}(\varnothing) = 0$.

□

Exemplo 2.23. Considere um experimento aleatório e os eventos A e B associados, tais que $\mathbb{P}(A) = \frac{1}{2}$ e $\mathbb{P}(B) = \frac{1}{3}$ e $\mathbb{P}(A \cap B) = \frac{1}{4}$. Então:

(i) $\mathbb{P}(A^c) = 1 - \mathbb{P}(A) = \frac{1}{2}$ e $\mathbb{P}(B^c) = 1 - \mathbb{P}(B) = \frac{2}{3}$;

(ii) $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B) = \frac{1}{2} + \frac{1}{3} - \frac{1}{4} = \frac{7}{12}$;

(iii) $\mathbb{P}(A^c \cap B^c) = \mathbb{P}((A \cup B)^c) = 1 - \mathbb{P}(A \cup B) = 1 - \frac{7}{12} = \frac{5}{12}$;

(iv) $\mathbb{P}(A^c \cup B^c) = \mathbb{P}((A \cap B)^c) = 1 - \mathbb{P}(A \cap B) = 1 - \frac{1}{4} = \frac{3}{4}$;

(v) Calcule $\mathbb{P}(A^c \cap B)$, ou seja, a probabilidade de que B ocorra e não ocorra A. Temos que

$$B = (A \cap B) \cup (A^c \cap B),$$

lembrando que $(A \cap B)$ e $(A^c \cap B)$ são disjuntos, ou seja $(A \cap B) \cap (A^c \cap B) = \varnothing$. Assim,

$$\mathbb{P}(B) = \mathbb{P}(A \cap B) + \mathbb{P}(A^c \cap B),$$

ou seja,

$$\mathbb{P}(A^c \cap B) = \mathbb{P}(B) - \mathbb{P}(A \cap B) = \frac{1}{3} - \frac{1}{4} = \frac{1}{12}.$$

□

2.4 Probabilidade Condicional

Exemplo 2.24. Componentes complexos são montados em em uma fábrica com duas linhas de montagem A_1 e A_2. A linha A_1 usa equipamentos mais antigos que a linha A_2, de forma que é mais lenta e menos eficiente. Suponha, que em determinado dia, a linha A_1 tenha montado 8 componentes, dos quais 2 foram identificados como defeituosos B_1 e 6 como não defeituosos B_2, enquanto a linha A_2 produziu 1 defeituoso e 9 não defeituosos. As informações estão resumidas na tabela abaixo.

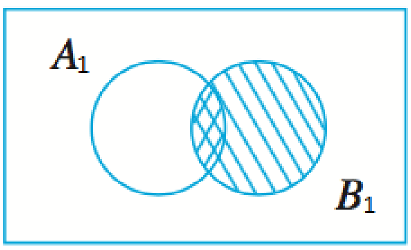
	Condição	
Linha	B_1	B_2
A_1	2	6
A_2	1	9

Não tendo estas informações, foi selecionado aleatoriamente 1 dos 18 componentes para uma demonstração.

$$\mathbb{P}(\text{Componentes selecionado da linha } A_1) = \mathbb{P}(A_1) = \frac{N(A_1)}{n} = \frac{8}{18} = 0.44.$$

Agora, se o componente escolhido for defeituoso, o evento B_1 terá ocorrido, de forma que o componente deve ser 1 dos 3 da coluna B_1 da tabela. Como estes componentes são igualmente prováveis entre si após a ocorrência de B_1,

$$\mathbb{P}(A_1|B_1) = \frac{2}{3} = \frac{\frac{2}{18}}{\frac{3}{18}} = \frac{\mathbb{P}(A_1 \cap B_1)}{\mathbb{P}(B_1)}.$$



Dado que B_1 ocorreu, o espaço amostral relevante não é mais Ω, mas consiste em resultados contidos em B_1. A_1 ocorre se, e somente se, um dos resultados da intersecção tiver ocorrido, de forma que a probabilidade condicional de A_1 dado B_1 é proporcional a $\mathbb{P}(A_1 \cap B_1)$. A proporção $1/\mathbb{P}(B_1)$ é utilizada para assegurar que a probabilidade $\mathbb{P}(B_1|B_1)$ do novo espaço amostral de B_1 seja igual a 1.

□

Definição 2.12. Para quaisquer dois eventos A e B, com $\mathbb{P}(B) > 0$, a *probabilidade condicional de A dado que B ocorreu* é definida por

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}.$$

Exemplo 2.25. Suponha que, de todos os indivíduos que compram um câmera digital, 60% incluem um cartão de memória opcional na compra, 40% incluem uma pilha extra e 30% incluem um cartão e uma pilha. Considere a seleção aleatória de um comprador e seja A={compra de cartão de memória} e B={compra de pilha}. Dessa forma, $\mathbb{P}(A) = 0.6$, $\mathbb{P}(B) = 0.4$ e $\mathbb{P}(A \cap B) = 0.3$. Dado que o indivíduo selecionado comprou uma pilha extra, a probabilidade de compra de um cartão opcional é:

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} = \frac{0.3}{0.4} = 0.75.$$

Ou seja, de todos os que compraram uma pilha extra, 75% compraram um cartão de memória extra. De forma análoga,

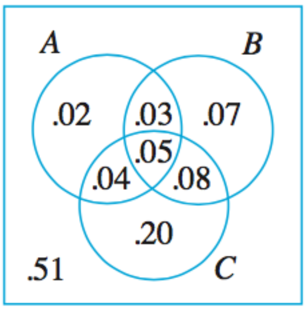
$$\mathbb{P}(\text{pilha}|\text{cartão de memória}) = \mathbb{P}(B|A) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(A)} = \frac{0.3}{0.6} = 0.5.$$

Observe que $\mathbb{P}(A|B) \neq \mathbb{P}(A)$ e $\mathbb{P}(B|A) \neq \mathbb{P}(B)$.

□

Exemplo 2.26. (Devore, 2017) Uma revista publica 3 colunas, intituladas “Arte"(A), “Livros"(B) e “Cinema"(C). Os hábitos de leitura de um eleitor lecionado aleatoriamente em relação a essas colunas são:

Lê regularmente	A	B	C	$A \cap B$	$A \cap C$	$B \cap C$	$A \cap B \cap C$
Probabilidade	0.14	0.23	0.37	0.08	0.09	0.13	0.05



Assim temos:

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} = \frac{0.08}{0.23} = 0.348.$$

$$\mathbb{P}(A|(B \cup C)) = \frac{\mathbb{P}(A \cap (B \cup C))}{\mathbb{P}(B \cup C)} = \frac{0.04 + 0.05 + 0.03}{0.47} = \frac{0.12}{0.47} = 0.255.$$

$$\begin{aligned}
\mathbb{P}(A|\text{lê pelo menos uma}) &= \mathbb{P}(A|(A \cup B \cup C)) = \frac{\mathbb{P}(A \cap (A \cup B \cup C))}{\mathbb{P}(A \cup B \cup C)} \\
&= \frac{\mathbb{P}(A)}{\mathbb{P}(A \cup B \cup C)} = \frac{0.14}{0.49} = 0.286.
\end{aligned}$$

$$\mathbb{P}((A \cup B)|C) = \frac{\mathbb{P}((A \cup B) \cap C)}{\mathbb{P}(C)} = \frac{0.04 + 0.05 + 0.08}{0.37} = 0.459.$$

□

Propriedades:

(i) $0 \leq \mathbb{P}(A|B) \leq 1$;

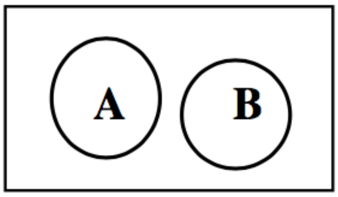
(ii) $\mathbb{P}(\Omega|A) = 1$;

(iii) $\mathbb{P}((B_1 \cup B_2)|A) = \mathbb{P}(B_1|A) + \mathbb{P}(B_2|A)$, se $B_1 \cap B_2 = \emptyset$.

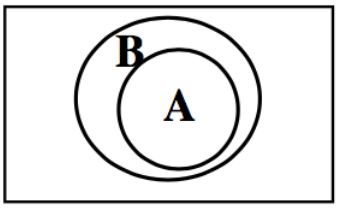
(iv) $\mathbb{P}((B_1 \cup B_2 \cup \cdots)|A) = \mathbb{P}(B_1|A) + \mathbb{P}(B_2|A) + \cdots$, se $B_i \cap B_j = \emptyset$, para $i \neq j$.

Observações:

(i) Se A e B são disjuntos, então $\mathbb{P}(A|B) = 0$;



(ii) Se $A \subseteq B$, temos que $\mathbb{P}(A) \leq \mathbb{P}(B)$ e assim, $\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} = \frac{\mathbb{P}(A)}{\mathbb{P}(B)} \geq \mathbb{P}(A)$.

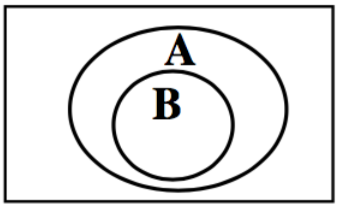


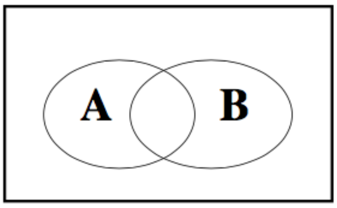
(iii) Se $B \subset A$, temos que $\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} = \frac{\mathbb{P}(B)}{\mathbb{P}(B)} = 1 \geq \mathbb{P}(A)$.

(iv) Caso Geral: nada podemos afirmar sobre $\mathbb{P}(A|B)$ e $\mathbb{P}(A)$.

Regra da Multiplicação:

$$\mathbb{P}(A \cap B) = \mathbb{P}(A|B)\mathbb{P}(B).$$





Exemplo 2.27. Quatro pessoas respondem a uma solicitação de um banco de sangue para doação. Nenhum doou sangue anteriormente, de forma que os quatro tipos sanguíneos são desconhecidos. Suponha que apenas o O^+ seja desejado, e apenas uma das 4 pessoas possua este tipo sanguíneo. Se os doadores em potencial forem selecionados aleatoriamente, qual será a probabilidade de que pelo menos 3 pessoas tenham de ser testadas para a obtenção do tipo sanguíneo desejado?

Defina: B={primeiro tido não O^+} e A={segundo tido não O^+}. Temos que $\mathbb{P}(B) = \frac{3}{4}$. Dado que o primeiro tipo não é O^+, dois das três pessoas restantes não serão O^+, de forma que $\mathbb{P}(A|B) = \frac{2}{3}$. Assim,

$$\begin{aligned} \mathbb{P}(\text{Pelo menos 3 pessoas foram testadas}) &= \mathbb{P}(A \cap B) \\ &= \mathbb{P}(A|B)\mathbb{P}(B) \\ &= \frac{2}{3} \times \frac{3}{4} = \frac{6}{12} = 0.5. \end{aligned}$$

□

A Regra da Multiplicação é mais útil quando o experimento consistem em diversas etapas. Assim,

$$\begin{aligned} \mathbb{P}(A_1 \cap A_2 \cap A_3) &= \mathbb{P}(A_3|(A_2 \cap A_1))\mathbb{P}(A_2 \cap A_1) \\ &= \mathbb{P}(A_3|(A_2 \cap A_1))\mathbb{P}(A_2|A_1)\mathbb{P}(A_1), \end{aligned}$$

onde A_1 ocorre primeiro, seguido por A_2 e finalmente por A_3.

Exemplo 2.28. Para o Exemplo 2.27, calcule a probabilidade da terceira pessoa ser do tipo O^+.

$$\begin{aligned} \mathbb{P}(\text{terceito tipo é } O^+) &= \mathbb{P}(\text{terceiro é}|\text{primeiro não é} \cup \text{segundo não é}) \\ & \quad \times \mathbb{P}(\text{segundo não é}|\text{primeiro não é}) \times \mathbb{P}(\text{primeiro não é}) \\ &= \frac{1}{2} \times \frac{2}{3} \times \frac{3}{4} = \frac{1}{4} = 0.25. \end{aligned}$$

□

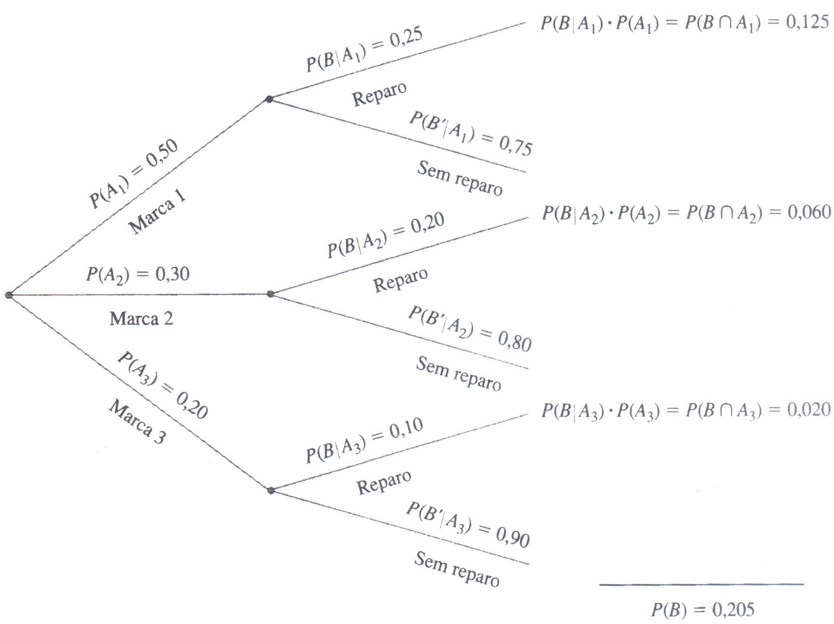
Exemplo 2.29. Uma cadeia de loja de vídeos vende 3 marcas diferentes de dvd. Dessas vendas, 50% são da marca 1 (a mais barata), 30% são da marca 2 20% da marca 3. Cada fabricante oferece um ano de garantia para peças e mão-de-obra. É sabido que 25% dos dvd's da marca 1 necessitam de reparos de garantia, enquanto os percentuais correspondentes para as marcas 2 e 3 são 20% e 10%, respectivamente.

1. Qual é a probabilidade de que um comprador selecionado aleatoriamente compre um dvd da marca 1 que precise de reparo durante a garantia?

2. Qual é a probabilidade de que um comprador selecionado aleatoriamente possua um aparelho que necessite de reparos durante a garantia?

3. Se um cliente voltar à loja com um dvd que precise de reparos de garantia, qual é a probabilidade de ele ser da marca 1? E da marca 2? E da marca 3?

Seja $A_i = \{\text{compra da marca } i\}$, para $i \in \{1, 2, 3\}$. Então, $\mathbb{P}(A_1) = 0.5$, $\mathbb{P}(A_2) = 0.3$ e $\mathbb{P}(A_3) = 0.2$
Seja $B = \{\text{precisa de reparo}\}$ $B' = \{\text{não precisa de reparo}\}$. Assim, $\mathbb{P}(B|A_1) = 0.25$, $\mathbb{P}(B|A_2) = 0.2$ e $\mathbb{P}(B|A_3) = 0.10$



Pergunta 1:

$$\mathbb{P}(A_1 \cap B) = \mathbb{P}(B|A_1)\mathbb{P}(A_1) = 0.125.$$

Pergunta 2:

$$\begin{aligned}
\mathbb{P}(B) &= \mathbb{P}[(\text{marca 1 e reparo}) \text{ ou } (\text{marca 2 e reparo}) \text{ ou } (\text{marca 3 e reparo})] \\
&= \mathbb{P}(A_1 \cap B) + \mathbb{P}(A_2 \cap B) + \mathbb{P}(A_3 \cap B) \\
&= 0.125 + 0.06 + 0.02 = 0.205.
\end{aligned}$$

Pergunta 3:

$$\mathbb{P}(A_1|B) = \frac{\mathbb{P}(A_1 \cap B)}{\mathbb{P}(B)} = \frac{0.125}{0.250} = 0.61$$

$$\mathbb{P}(A_2|B) = \frac{\mathbb{P}(A_2 \cap B)}{\mathbb{P}(B)} = \frac{0.06}{0.250} = 0.29$$

$$\mathbb{P}(A_3|B) = 1 - \mathbb{P}(A_1|B) - \mathbb{P}(A_2|B) = 0.1$$

Observe que a probabilidade inicial da marca 1 é 0.5, enquanto, depois que se sabe que o dvd selecionado precisa de reparos, a probabilidade posterior da marca 1 aumenta para 0.61. Isso ocorre porque os dvd's da marca 1 são mais prováveis de precisar de reparos do que os das outras marcas. A probabilidade posterior da marca 3 é $\mathbb{P}(A_3|B) = 0.1$, que é muito menor que a probabilidade anterior $\mathbb{P}(A_3) = 0.2$.

□

2.5 Teorema de Bayes

O cálculo de uma probabilidade posterior $\mathbb{P}(A_j|B)$ a partir das probabilidades anteriores dadas $\mathbb{P}(A_j)$ e das probabilidades condicionais $\mathbb{P}(B|A_j)$ ocupa uma posição central na teoria da probabilidade.

O conceito de probabilidade condicionada pode ser utilizado para calcular a probabilidade de um evento simples A ao invés da probabilidade da interseção de dois eventos A e B. Para tanto é necessário o conceito de partição de um espaço amostral.

Definição 2.13. Sejam os conjuntos $A_1, A_2, \cdots, A_n$ eventos de um mesmo espaço amostral Ω. Os conjuntos $A_1, A_2, \cdots, A_n$ formam uma *partição* deste espaço se:

(i) $A_i \cap A_j = \emptyset$, para todo $i \neq j$;

(ii) $A_1 \cup A_2 \cup \cdots \cup A_n = \Omega$;

(iii) $\mathbb{P}(A_j) > 0$, para todo $j = 1, \cdots, n$.

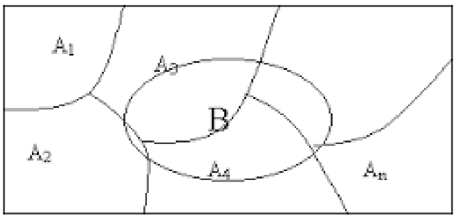


Figura 2.2: Exemplo de partição.

Teorema 2.1. Teorema da Probabilidade Total: *Sejam $A_1, A_2, \cdots, A_n$ eventos que formam uma partição de Ω. Então, para qualquer outro evento B, temos*

$$\begin{aligned} \mathbb{P}(B) &= \mathbb{P}(B|A_1)\mathbb{P}(A_1) + \cdots + \mathbb{P}(B|A_n)\mathbb{P}(A_n) \\ &= \sum_{j=1}^{n} \mathbb{P}(B|A_j)\mathbb{P}(A_j). \end{aligned}$$

Como os eventos $A_1, A_2, \cdots, A_n$ formam uma partição Ω, se B ocorrer será com exatamente um dos A_j, ou seja, $B = (A_1$ e $B)$ ou $(A_2$ e $B)$ ou $\cdots$ ou $(A_n$ e $B)$. Assim, $B = (A_1 \cap B) \cup (A_2 \cap B) \cup \cdots \cup (A_n \cap B)$, onde os eventos $(A_j \cap B)$ são mutuamente exclusivos. A partição é ilustrada na Figura 2.2 acima. Assim,

$$\mathbb{P}(B) = \sum_{j=1}^{n} \mathbb{P}(A_j \cap B) = \sum_{j=1}^{n} \mathbb{P}(B|A_j)\mathbb{P}(A_j).$$

Exemplo 2.30. Uma determinada peça é manufaturada por 3 fábricas: A, B e C. Sabe-se que A produz o dobro de peças que B e que B e C produzem o mesmo número de peças. Sabe-se ainda que 2% das peças produzidas por A e por B são defeituosas, enquanto que 4% das produzidas por C são defeituosas. Todas as peças produzidas são misturadas e colocadas em um depósito. Se do depósito for retirada uma peça ao acaso, qual a probabilidade de que ela seja defeituosa?

Defina os eventos: D={A peça é defeituosa}, A={A peça provém da fábrica A}, B={A peça provém da máquina B} e C={A peça provém da máquina C}.

Logo temos que, $\mathbb{P}(A) = 0.5$ e $\mathbb{P}(B) = \mathbb{P}(C) = 0.25$. Temos que $\mathbb{P}(D|A) = \mathbb{P}(D|B) = 0.02$ e $\mathbb{P}(D|C) = 0.04$.

Pela Teorema da Probabilidade Total pode-se escrever que:

$$\begin{aligned} \mathbb{P}(D) &= \mathbb{P}(D|A)\mathbb{P}(A) + \mathbb{P}(D|B)\mathbb{P}(B) + \mathbb{P}(D|C)\mathbb{P}(C) \\ &= 0.5 \times 0.02 + 0.25 \times 0.02 + 0.05 \times 0.04 = 0.025, \end{aligned}$$

pois, A, B e C formam uma partição do espaço amostral Ω.

$\square$

Suponha-se que no exemplo acima, uma peça é retirada do depósito e se verifica que é defeituosa. Qual a probabilidade de que tenha sido produzida pela fábrica A? ou B? ou ainda C?

Neste caso, o que se quer calcular é a probabilidade condicionada $\mathbb{P}(A|D)$.

Generalizando, a questão que se está interessado é em obter é a probabilidade de ocorrência de um dos A_i dado que B ocorreu, isto é, o que se quer é saber o valor de $\mathbb{P}(A_i|B)$, onde os eventos $A_1, A_2, \cdots, A_n$ formam uma partição Ω e B é um evento qualquer de Ω.

Teorema 2.2. Teorema de Bayes: *Sejam $A_1, A_2, \cdots, A_n$ eventos que formam uma partição de Ω. Então, para qualquer outro evento B, temos*

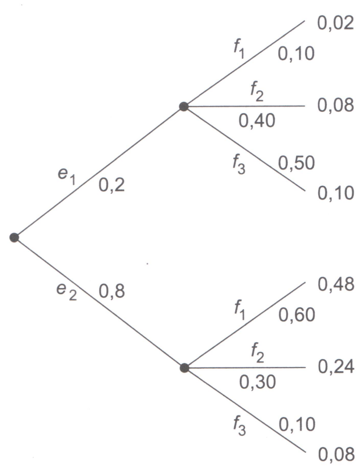
$$\mathbb{P}(A_i|B) = \frac{\mathbb{P}(A_i \cap B)}{\mathbb{P}(B)} = \frac{\mathbb{P}(B|A_j)\mathbb{P}(A_j)}{\sum_{j=1}^{n} \mathbb{P}(B|A_j)\mathbb{P}(A_j)}, \text{ para todo } j = 1, \cdots, n.$$

Exemplo 2.31. Suponha que um geólogo, com base em observações superficiais, tenha estabelecido que a probabilidade de que haja petróleo, em certo lugar, é 0.2. Então os eventos “há petróleo"e ‘não há petróleo", que denotaremos por E_1 e E_2, respectivamente, possuem probabilidade $\mathbb{P}(E_1) = 0.2$ e $\mathbb{P}(E_2) = 0.8$. Após o estabelecimento dessas probabilidades, o geólogo realiza um teste sísmico. Consideremos, que o teste pode resultar em um dos 3 eventos mutuamente exclusivos: F_1, F_2 e F_3. O resultado do teste está relacionado a existência ou não de petróleo. Suponhamos que as probabilidades condicionais sejam:

$$\mathbb{P}(F_1|E_1) = 0.1, \quad \mathbb{P}(F_2|E_1) = 0.4, \quad \mathbb{P}(F_3|E_1) = 0.5$$

$$\mathbb{P}(F_1|E_2) = 0.6, \quad \mathbb{P}(F_2|E_2) = 0.3, \quad \mathbb{P}(F_3|E_2) = 0.1$$

Se o teste resultar em F_3, o geólogo deve determinar a probabilidade de que haja petróleo, dado que F_3 ocorreu, ou seja, deve calcular $\mathbb{P}(E_1|F_3)$.



Temos que

$$\mathbb{P}(E_1|F_3) = \frac{\mathbb{P}(E_1 \cap F_3)}{\mathbb{P}(F_3)}$$

onde, pela definição de probabilidade condicional, temos que

$$\mathbb{P}(E_1 \cap F_3) = \mathbb{P}(F_3|E_1)\mathbb{P}(E_1) = 0.5 \times 0.2 = 0.1$$

e pelo Teorema da Probabilidade Total, temos que

$$\mathbb{P}(F_3) = \mathbb{P}(F_3|E_1)\mathbb{P}(E_1) + \mathbb{P}(F_3|E_2)\mathbb{P}(E_2) = 0.5 \times 0.2 + 0.1 \times 0.8 = 0.1 + 0.08 = 0.18.$$

Portanto,

$$\mathbb{P}(E_1|F_3) = \frac{\mathbb{P}(E_1 \cap F_3)}{\mathbb{P}(F_3)} = \frac{0.1}{0.18} = 0.5556.$$

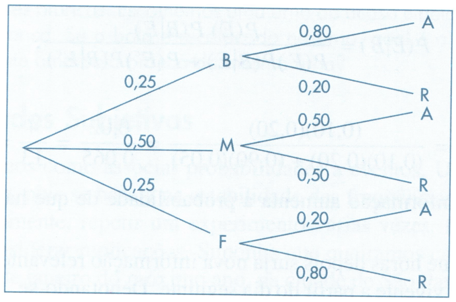
Analogamente,

$$\mathbb{P}(E_1|F_1) = 0.04, \quad \mathbb{P}(E_1|F_2) = 0.25.$$

□

Exemplo 2.32. Para selecionar seus funcionários uma empresa fornece aos seus candidatos um curso de treinamento por uma semana. No final do curso, eles são submetidos a uma prova e 25% são classificados como bons (B), 50% como médios (M) e os restantes como fracos (F). Para facilitar a seleção, a empresa promete substituir o treinamento por um teste contendo questões referentes a conhecimentos gerais e específicos. Para isso, gostaria de conhecer qual a probabilidade de um indivíduo aprovado no teste ser considerado fraco, caso fizesse o curso. Assim, neste ano, antes do curso, os candidatos foram submetidos ao teste e recebera o conceito aprovado (A) ou reprovado (R), No final do curso, obtiveram-se as seguintes probabilidades condicionais:

$$\mathbb{P}(A|B) = 0.8, \quad \mathbb{P}(A|M) = 0.5, \quad \mathbb{P}(A|F) = 0.2.$$



Queremos encontrar $\mathbb{P}(F|A)$. Pelo Teorema de Bayes, essa probabilidade é dada por

$$\begin{aligned}\mathbb{P}(F|A) &= \frac{\mathbb{P}(A\cap F)}{\mathbb{P}(A)} = \frac{\mathbb{P}(A|F)\mathbb{P}(F)}{\mathbb{P}(A|B)\mathbb{P}(B)+\mathbb{P}(A|M)\mathbb{P}(M)+\mathbb{P}(A|F)\mathbb{P}(F)}\\ &= \frac{0.2\times 0.25}{0.8\times 0.25+0.5\times 0.5+0.2\times 0.25} = 0.1.\end{aligned}$$

Então, apenas 10% dos aprovados é que seriam classificados como fracos durante o curso. De modo análogo, podemos calcular $\mathbb{P}(B|A) = 0.4$ e $\mathbb{P}(M|A) = 0.5$.

□

Exemplo 2.33. [Problema de Monte Hall] O Problema de Monty Hall é um problema matemático que surgiu a partir de um programa de TV chamado *Let's Make a Deal*, que era exibido nos EUA. Você está em um jogo em um programa de TV, e precisa escolher uma entre três portas: por trás de uma está um carro, e as outras duas portas estão vazias. Você escolhe um delas - digamos, a número 1 - e o apresentador (que sabe o que está por trás de cada uma delas) abre outra porta - digamos a número 3 - que está vazia. Você então tem a opção de continuar com a que escolheu, ou mudar para a outra - a número 2, no nosso exemplo. A questão é: você deve mudar a sua escolha?

Solução: A partição do espaço amostral (no momento em que eu já escolhi a porta 1) é a seguinte:

- P_1: eu escolhi a porta 1 e o carro está na porta 1;

- P_2: eu escolhi a porta 1 e o carro está na porta 2;

- P_3: eu escolhi a porta 1 e o carro está na porta 3.

Os a prioris são iguais: $\mathbb{P}(P_i) = 1/3$, para todo i. O evento B é o Monty Hall abrir a porta 2. Os condicionais são:

- $\mathbb{P}(B|P_1) = 1/2$ (Nesse caso Monty Hall abre umas das portas ao acaso);

- $\mathbb{P}(B|P_2) = 0$ (Nesse caso Monty Hall não pode abrir a porta 2);

- $\mathbb{P}(B|P_3) = 1$ (Nesse caso Monty Hall tem que abrir a porta 2).

Pelo Teorema de Bayes, temos

$$\mathbb{P}(P_3|B)\frac{1\times 1/3}{1/2\times 1/3+0\times 1/3+1\times 1/3} = \frac{2}{3}.$$

□

2.6 Independência

Há situações em que a possibilidade de ocorrência do evento A não é afetada pelo conhecimento da ocorrência de B, de forma que $\mathbb{P}(A|B) = \mathbb{P}(A)$. É natural, pensar em A e B como eventos independentes, indicando que a ocorrência ou não-ocorrência desses eventos não afeta a possibilidade de ocorrer o outro evento.

Definição 2.14. Sejam A e B dois eventos de um mesmo espaço amostral Ω. A e B são ditos independentes se a probabilidade de um dele ocorrer não afeta a probabilidade do outro ocorrer, ou seja, se um dos seguintes casos ocorrer:

(i) $\mathbb{P}(A|B) = \mathbb{P}(A)$;

(ii) $\mathbb{P}(B|A) = \mathbb{P}(B)$;

(iii) $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$.

Qualquer uma das 3 relações acima pode ser usada como definição de independência.

Exemplo 2.34. Sabe-se que 30% das máquinas na linha de produção A de uma determinada empresa requerem manutenção enquanto estiverem na garantia e somente 10% das máquinas na linha de produção B precisam de manutenção. Se alguém comprar uma máquina de cada linha de produção, qual a probabilidade de que ambas as máquinas precisem de manutenção?

Seja A={máquinas produzida na linha de produção A} e B={máquinas produzida na linha de produção B}. Então $\mathbb{P}(A) = 0.3$ e $\mathbb{P}(B) = 0.1$. Assumindo que as máquinas funcionem de forma independente uma da outra, a probabilidade desejada será

$$\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B) = 0.3 \times 0.1 = 0.03.$$

A probabilidade que nenhuma máquina precise do serviço é

$$\mathbb{P}(A^c \cap B^c) = \mathbb{P}(A^c)\mathbb{P}(B^c) = 0.7 \times 0.9 = 0.63.$$

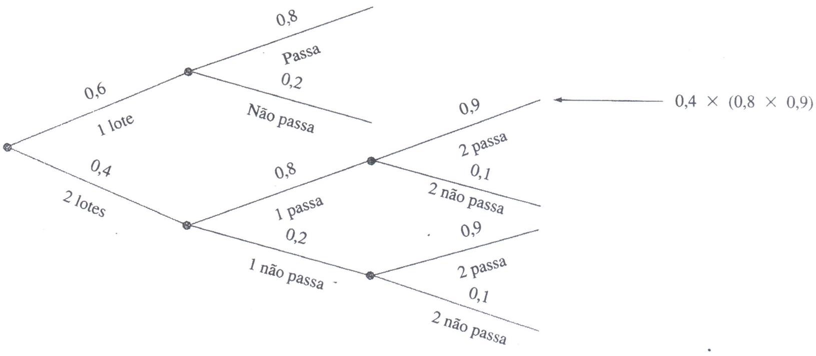
□

Exemplo 2.35. Cada dia, de segunda a sexta-feira, chega em uma determinada instalação de inspeção um lote de componentes enviados por um fornecedor. Duas vezes por semana, chega um lote de um segundo fornecedor. 80% de todos os lotes do fornecedor 1 passam na inspeção, assim como 90% dos lotes do fornecedor 2. Qual a probabilidade de, em um dia selecionado, dois, lotes passarem na inspeção?

Vamos considerar que, em dias em que os dois lotes são testados, os fatos de o primeiro ou de o segundo lote passarem ou não na inspeção são independentes.

$$\begin{aligned}
\mathbb{P}(\text{ambos passam}) &= \mathbb{P}(\text{dois recebidos} \cap \text{ambos passam}) \\
&= \mathbb{P}(\text{ambos passam}|\text{dois recebidos})\mathbb{P}(\text{dois recebidos}) \\
&= [0.8 \times 0.9] \times 0.4 = 0.288.
\end{aligned}$$

□



> **Definição 2.15.** Os eventos $A_1, \cdots, A_n$ são *mutuamente independentes* se para cada conjunto de índices $i_1, \cdots, i_k$, temos
>
> $$\mathbb{P}(A_{i_1} \cap \cdots \cap A_{i_k}) = \mathbb{P}(A_{i_1}) \times \cdots \times \mathbb{P}(A_{i_k}).$$

Os eventos são mutuamente independentes se a probabilidade da intersecção de qualquer subconjunto dos n eventos for igual ao produto das probabilidades individuais.

Exemplo 2.36. O artigo ‘Reliability evaluation of solar photovoltaic arrays"(Solar Energy, 2002: 129-141) apresenta diversas configurações de matrizes fotovoltaicas solares formadas por células solares de silício cristalino. Considere o sistema na figura a seguir.

Há dois sistema ligados em paralelo, cada um com 3 células. Para que o sistema funcione, pelo menos dois subsistemas paralelos devem funcionar. Em cada subsistema, as 3 células estão ligadas em série, de forma que um subsistema só funciona se todas as sus células estiverem funcionando. Considere a vida útil t_0 e suponha que se queira determinar a probabilidade de a vida útil do sistema exceder t_0. Seja A_i o evento na qual a vida útil da célula i excede t_0 ($i = 1, \cdots, 6$). Suponha que os eventos A_is sejam independentes (se alguma célula durar mais de t_0 horas, não tem nenhuma relação com a duração das outras células) e que $\mathbb{P}(A_i) = 0.9$ para cada i, já que as células são idênticas. Então,

$$\begin{aligned}
\mathbb{P}(\text{vida útil do sistema excede } t_0) &= \mathbb{P}[(A_1 \cap A_2 \cap A_3) \cup (A_4 \cap A_5 \cap A_6)] \\
&= \mathbb{P}(A_1 \cap A_2 \cap A_3) + \mathbb{P}(A_4 \cap A_5 \cap A_6) \\
&\quad - \mathbb{P}[(A_1 \cap A_2 \cap A_3) \cap (A_4 \cap A_5 \cap A_6)] \\
&= \mathbb{P}(A_1)\mathbb{P}(A_2)\mathbb{P}(A_3) + \mathbb{P}(A_4)\mathbb{P}(A_5)\mathbb{P}(A_6) \\
&\quad - \mathbb{P}(A_1)\mathbb{P}(A_2)\mathbb{P}(A_3)\mathbb{P}(A_4)\mathbb{P}(A_5)\mathbb{P}(A_6)] \\
&= (0.9)^3 + (0.9)^3 - (0.9)^6 \\
&= 0.927.
\end{aligned}$$

De forma análoga,

$$\begin{aligned}
\mathbb{P}(\text{vida útil do sistema excede } t_0) &= 1 - \mathbb{P}(\text{vida útil de ambos os sistema são} \leq t_0) \\
&= 1 - [\mathbb{P}(\text{vida útil do subsistema é} \leq t_0)]^2 \\
&= 1 - [1 - \mathbb{P}(\text{vida útil do subsistema é} > t_0)]^2 \\
&= 1 - [1 - (0.9)^3]^2 = 0.927.
\end{aligned}$$

Considerando o segundo sistema. O sistema falhará se houver falha de uma coluna e a vida útil do sistema excederá t_0 apenas se a vida útil de cada coluna também exceder t_0. Portanto,

$$\begin{aligned}
\mathbb{P}(\text{vida útil do sistema é no mínimo } t_0) &= [\mathbb{P}(\text{vida útil da coluna excede } t_0)]^3 \\
&= [1 - \mathbb{P}(\text{vida útil da coluna é } \leq t_0)]^3 \\
&= [1 - \mathbb{P}(\text{ambas as células da coluna têm vida útil } \leq t_0)]^3 \\
&= [1 - (1 - 0.9)^2]^3 = 0.970.
\end{aligned}$$

□

Exemplo 2.37. Um lote contém 10 peças, sendo 7 boas (B) e 3 defeituosas (D). Retiramos duas peças, ao acaso, para inspeção. Qual a probabilidade de se obter duas peças defeituosas?
Solução:

Com Reposição:
O experimento de realizar a primeira retirada tem como espaço amostral $\Omega_1 = \{D_1, B_1\}$ e a segunda retirada tem como espaço amostral $\Omega_2 = \{D_2, B_2\}$, em que D_i significa que retiramos uma peça defeituosa na i-ésima retirada e B_i significa que retiramos uma peça boa na i-ésima retirada, para $i = 1, 2$. Como as duas peças são retiradas ao acaso e com reposição, isto é, após retirarmos a primeira peça esta é colocada novamente no lote para que possamos efetuar a segunda retirada, temos que

$$\mathbb{P}(D_1) = \mathbb{P}(D_2) = \frac{3}{10} \qquad \text{e} \qquad \mathbb{P}(B_1) = \mathbb{P}(B_2) = \frac{7}{10}.$$

Associamos ao experimento de retirar duas peças ao acaso e com reposição o seguinte espaço amostral

$$\Omega = \{(D_1, B_2); (B_1, D_2); (D_1, D_2); (B_1, B_2)\}.$$

Queremos encontrar a probabilidade de se obter duas peças defeituosas, ou seja, a probabilidade das peças na primeira retirada e na segunda retirada serem defeituosas. Assim, desde que a primeira e a segunda retirada sejam executadas de forma independente, temos que

$$\mathbb{P}\left((D_1, D_2)\right) = \mathbb{P}(D_1 \cap D_2) = \mathbb{P}(D_1)\mathbb{P}(D_2) = \frac{3}{10} \cdot \frac{3}{10} = \frac{9}{100}.$$

Vamos examinar melhor a diferença entre extrair uma peça de um lote, ao acaso, com reposição ou sem reposição. Como vimos neste exemplo, se a retirada for feita com reposição, então

$$\mathbb{P}(D_1) = \mathbb{P}(D_2) = \frac{3}{10} \qquad \text{e} \qquad \mathbb{P}(B_1) = \mathbb{P}(B_2) = \frac{7}{10}$$

pois cada vez que extraímos peças do lote, sempre existirão 3 peças defeituosas e 7 peças boas num total de 10. No entanto, se estivermos extraindo sem reposição, o resultado é diferente. É ainda verdade, naturalmente, que

$$\mathbb{P}(D_1) = \frac{3}{10} \qquad \text{e} \qquad \mathbb{P}(B_1) = \frac{7}{10},$$

mas as probabilidades de sair uma peça defeituosa ou de sair uma peça boa na segunda retirada não serão as mesmas. Para calcularmos essas probabilidades devemos conhecer a composição do lote no momento de se extrair a segunda peça. Por exemplo, para calcularmos a probabilidade de extrairmos uma peça defeituosa na segunda retirada, D2, temos que saber se ocorreu D_1 ou B_1. Caso tenha ocorrido D_1,

$$\mathbb{P}(D_2) = \frac{2}{9}$$

e, se ocorreu B1,

$$\mathbb{P}(D_2) = \frac{3}{9}.$$

Sem Reposição:

Agora as retiradas serão feitas sem reposição, isto é, a primeira peça retirada não volta ao lote para retirarmos a segunda peça. Qual a probabilidade de se retirar duas peças defeituosas?

A probabilidade de sair uma peça defeituosa na primeira retirada é $\mathbb{P}(D_1) = \frac{3}{10}$. Além disso, $\mathbb{P}(D_2|D_1) = \frac{2}{9}$. Assim,

$$\mathbb{P}[(D_1, D_2)] = \mathbb{P}(D_1 \cap D_2) = \mathbb{P}(D_1) \times \mathbb{P}(D_2|D_1) = \frac{3}{10} \times \frac{2}{9} = \frac{6}{90} = \frac{1}{15}.$$

□

Exemplo 2.38. Em uma prova de múltipla escolha, cada questão tem 5 alternativas, sendo apenas uma delas correta. Ao não saber a resposta, o aluno *chuta* aleatoriamente uma resposta qualquer entre as possíveis escolhas. Levando-se em conta um aluno mediano, que saiba 60% do conteúdo, qual será a chance de ele acertar uma das 5 questões escolhida aleatoriamente? E qual a chance de ele acertar exatamente 3 questões?

Solução: Este problema consiste em calcular a probabilidade incondicional de um aluno acertar uma questão qualquer. Isto é, sem saber se ele domina ou não o conteúdo, qual é a chance de acertar uma questão?

Para resolver este exercício, portanto, você deve aplicar o teorema da Probabilidade Total.

Considere os eventos: A = acertar, B = saber o conteúdo e B^c = não saber o conteúdo e, portanto, *chutar* uma alternativa.

Assuma que, se o aluno sabe o conteúdo, ele tem 100% de probabilidade de acertar a questão considerada. Se ele não domina o assunto, chutará uma resposta, com 20% de chance de acertar ? pois há 5 alternativas possíveis.

Logo,

$$\mathbb{P}(A) = \mathbb{P}(B)\mathbb{P}(A|B) + \mathbb{P}(B^c)\mathbb{P}(A|B^c)$$

Então:

$$\mathbb{P}(A) = 0.6 \times 1 + 0.4 \times 0.2 = 0.68$$

Para calcular a chance de o aluno acertar exatamente 3 questões, vamos utilizar a informação obtida na primeira parte do exercício.

Essa chance será calculada multiplicando a chance de ele acertar 3 questões multiplicada pela chance de errar 2 questões multiplicada pelo número possível de combinações com 3 questões certas e 2 erradas. Utilizando o que você aprendeu sobre regra do produto e os conhecimentos que já possuía sobre combinações temos:

$$(0.68)^3 \times (1 - 0.68)^2 \times \binom{5}{3} = 0.314432 \times 0.1024 \times 10 = 0.321978$$

□

Exemplo 2.39. A probabilidade condicional também é uma probabilidade ($P(\cdot|B)$, para B um subconjunto fixo de Ω), ou seja a probabilidade condicional satisfaz os três axiomas de probabilidade.

Mostremos primeiramente que $\mathbb{P}(\Omega|B) = 1$ e que $\mathbb{P}(\emptyset|B) = 0$. De fato, note que

$$\mathbb{P}(\Omega|B) = \frac{\mathbb{P}(\Omega \cap B)}{\mathbb{P}(B)} = \frac{\mathbb{P}(B)}{\mathbb{P}(B)} = 1$$

e que

$$\mathbb{P}(\emptyset|B) = \frac{\mathbb{P}(\emptyset \cap B)}{\mathbb{P}(B)} = \frac{\mathbb{P}(\emptyset)}{\mathbb{P}(B)} = \frac{0}{\mathbb{P}(B)} = 0$$

o que demonstra o primeiro axioma.

O segundo axioma diz que $0 \leq \mathbb{P}(A|B) \leq 1$, para qualquer $A \subset \Omega$. Observe que $\mathbb{P}(A|B) = \frac{\mathbb{P}(A\cap B)}{\mathbb{P}(B)}$, e como $A \cap B \subset B$. Temos que por P4 que $0 \leq \mathbb{P}(A \cap B) \leq \mathbb{P}(B)$, o que implica que $0 \leq \mathbb{P}(A|B) \leq 1$.

O terceiro e último axioma diz que para qualquer sequência de eventos mutuamente exclusivos $A_1, A_2, \ldots$, temos que

$$\mathbb{P}(\bigcup_{n=1}^{\infty} A_n|B) = \sum_{n=1}^{\infty} \mathbb{P}(A_n|B).$$

Observamos que:

$$\mathbb{P}\left(\bigcup_{n=1}^{\infty} A_n|B\right) = \frac{\mathbb{P}(B \cap \bigcup_{n=1}^{\infty} A_n)}{\mathbb{P}(B)} = \frac{\mathbb{P}(\bigcup_{n=1}^{\infty} A_n \cap B)}{\mathbb{P}(B)} = \frac{\sum_{n=1}^{\infty} \mathbb{P}(A_n \cap B)}{\mathbb{P}(B)} = \sum_{n=1}^{\infty} \mathbb{P}(A_n|B).$$

Logo, a probabilidade condicional satisfaz todos os axiomas da probabilidade, o que implica que a probabilidade condicional também é uma probabilidade. Assim sendo, todas as propriedades de probabilidade também são válidas.

□

Exemplo 2.40. Considere novamente o Exemplo 1.4.3 onde A_i é o evento de ter um adversário do tipo i e

$$\mathbb{P}(A_1) = 0,5; \qquad \mathbb{P}(A_2) = 0,25; \qquad \mathbb{P}(A_3) = 0,25.$$

Além disso, B é evento vencer uma partida e

$$\mathbb{P}(B|A_1) = 0,3; \qquad \mathbb{P}(B|A_2) = 0,4; \qquad \mathbb{P}(B|A_3) = 0,5.$$

Suponha que o jogador disputou uma partida e venceu. Qual a probabilidade $\mathbb{P}(A_1|B)$ dele ter jogado contra um adversário do tipo 1?
Solução: Usando o teorema de Bayes, temos que

$$\mathbb{P}(A_1|B) = \frac{P(\mathbb{A}_1)\mathbb{P}(B|A_1)}{\mathbb{P}(A_1)\mathbb{P}(B|A_1) + \mathbb{P}(A_2)\mathbb{P}(B|A_2) + \mathbb{P}(A_3)\mathbb{P}(B|A_3)} = 0,4.$$

Ou seja, a probabilidade do jogador ter disputado uma partida contra um adversário do tipo 1, dado que ele venceu a partida é de 40%.

□

Exemplo 2.41. Um teste de laboratório detecta uma doença quando ela está presente em 95% dos casos. No entanto, o teste também fornece um resultado "falso positivo"para 1% das pessoas saudáveis testadas. (Isto é, se uma pessoa saudável faz o teste, então, com probabilidade $0,01$, o resultado do teste dirá que ela possui a doença.) Se $0,5\%$ da população tem a doença, qual é a probabilidade de uma pessoa ter a doença dado que o resultado do teste é positivo?

Resposta: Para resolver este problema, consideramos D o evento de a pessoa testada ter a doença e E o evento de que o resultado do teste é positivo. Então, a probabilidade desejada $\mathbb{P}(D|E)$ é obtida por

$$\mathbb{P}(D|E) = \frac{\mathbb{P}(D \cap E)}{\mathbb{P}(E)} = \frac{\mathbb{P}(E|D)\mathbb{P}(D)}{\mathbb{P}(E|D)\mathbb{P}(D) + \mathbb{P}(E|D^c)\mathbb{P}(D^c)},$$

ou seja,

$$\mathbb{P}(D|E) = \frac{(0,95)(0,005)}{(0,95)(0,005) + (0,01)(0,995)} \approx 0,323.$$

Assim, apenas 32% das pessoas cujos resultados do teste deram positivo realmente possuem a doença.

□

Exemplo 2.42. Em um teste de múltipla escolha, ou um estudante sabe a resposta ou arrisca uma das alternativas. Seja p a probabilidade do estudante saber a resposta e $1-p$ a probabilidade do estudante arriscar adivinhá-la. Assuma que um estudante que arrisca a resposta acerta a resposta correta com probabilidade $1/m$, onde m é o número de alternativas de múltipla escolha. Qual é a probabilidade condicional de que um estudante soubesse a resposta da questão, dado que ele ou ela respondeu corretamente?
Solução: Seja C o evento de que o estudante responde a questão corretamente e K o evento de que ele saiba a resposta. Então

$$\mathbb{P}(K|C) = \frac{\mathbb{P}(K \cap C)}{\mathbb{P}(C)} = \frac{\mathbb{P}(C|K)\mathbb{P}(K)}{\mathbb{P}(C|K)\mathbb{P}(K) + \mathbb{P}(C|K^c)\mathbb{P}(K^c)},$$

ou seja,

$$\mathbb{P}(K|C) = \frac{p}{p + (1/m)(1-p)} = \frac{mp}{1+(m-1)p}.$$

Por exemplo, se $m = 5$ e $p = 1/2$, então a probabilidade de que um estudante saber a resposta de uma questão que ele respondeu corretamente é $5/6$.

□

Exemplo 2.43. Determinado veículo pode ter problemas mecânicos ou elétricos. Se ele tiver problemas mecânicos, não para, mas se tiver problema elétrico tem de parar imediatamente. A probabilidade de esse veículo ter problemas mecânicos é de 0.2. Já a probabilidade do mesmo veículo ter problemas elétricos é de 0.15 se não houve problema mecânico precedente, e de 0.25 se houve problema mecânico precedente. Calcule:

(a) Qual é a probabilidade de o veículo parar em determinado dia?

(b) Se o veículo parou em certo dia, qual a chance de que tenha havido defeito mecânico?

(c) Qual é a probabilidade de que tenha havido defeito mecânico em determinado dia se o veículo não parou nesse dia?

Solução: Considere os eventos: M = ter problema mecânico E = ter problema elétrico. São dadas as informações: $\mathbb{P}(M) = 0.2$, $\mathbb{P}(E|M^c) = 0.15$ e $\mathbb{P}(E|M) = 0.25$.

(a) O veículo somente vai parar se tiver problema elétrico. Então, precisamos calcular a Probabilidade Total de ocorrer defeito elétrico, independentemente de ter havido ou não defeito mecânico.

$$\mathbb{P}(E) = \mathbb{P}(M)\mathbb{P}(E|M) + \mathbb{P}(M^c)\mathbb{P}(E|M^c)$$

$$\mathbb{P}(E) = 0.2 \times 0.25 + 0.8 \times 0.15 = 0.17$$

(b) Devemos calcular a probabilidade de ter havido defeito mecânico condicionada ao fato de sabermos que o veículo parou (lembre-se que o veículo para quando há defeito elétrico). Isso é feito por meio do Teorema de Bayes.

$$\mathbb{P}(M|E) = \frac{\mathbb{P}(M)\mathbb{P}(E|M)}{\mathbb{P}(E} = \frac{0.2 \times 0.25}{0.17} = 0.294$$

(c) Mais uma vez, vamos utilizar o Teorema de Bayes para calcular a probabilidade de que tenha havido problema mecânico, dado que não houve defeito elétrico.

$$\mathbb{P}(M|E^c) = \frac{\mathbb{P}(M)\mathbb{P}(E^c|M)}{\mathbb{P}(E^c}$$

A probabilidade de não haver defeito elétrico é dada pela propriedade do evento complementar: $\mathbb{P}(E^c) = 1 - 0.17 = 0.83$.

Agora vamos calcular a probabilidade de não haver defeito elétrico, dado que houve defeito mecânico. Considerando o espaço amostral de todos os eventos que podem ocorrer, dado que houve defeito mecânico, sabemos que a chance de haver defeito elétrico é $\mathbb{P}(E|M) = 0.25$. A chance de não haver defeito elétrico será, portanto, o complementar do evento E em relação a este espaço amostral.

$$\mathbb{P}(E^c|M) = 1 - \mathbb{P}(E|M) = 1 - 0.25 = 0.75$$

Substituindo na expressão do Teorema de Bayes, temos:

$$\mathbb{P}(M|E^c) = \frac{\mathbb{P}(M)\mathbb{P}(E^c|M)}{\mathbb{P}(E^c} = \frac{0.2 \times 0.75}{0.83} = 0.181$$

□

Capítulo 3

Variáveis Aleatórias

Variável aleatória pode ser entendida como o resultado numérico de operar um mecanismo não determinístico ou de fazer uma experiência não determinística para gerar resultados aleatórios.

Definição 3.1. A variável aleatória é uma função de um espaço amostral Ω nos números reais, isto é:

$$X : \Omega \to \mathbb{R} \quad \Leftrightarrow \quad [X \leq x] \in A, \forall x \in \mathbb{R}, \text{ para } A \text{ } \sigma\text{-álgebra em } \Omega. \tag{3.1}$$

A função X de Ω em $\mathbb{R}$ será uma variável aleatória se, e somente se, para todo x que a função assumir, o conjunto X dos valores menores ou iguais a x pertencer ao sigma-álgebra, para qualquer x pertencente ao conjunto dos números Reais.

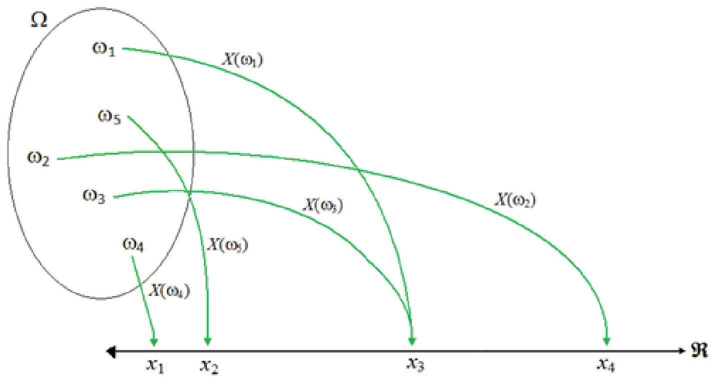


Figura 3.1: Fonte: Wikipédia

Variáveis aleatórias podem ser discretas ou contínuas.

3.1 Variáveis Aleatórias Discretas

São discretas todas as variáveis cujo espaço amostral Ω_X é enumerável infinito ou finito. Se X é uma variável aleatória discreta, então Ω_X é um subconjunto dos inteiros.

Exemplo 3.1. Lançamento de uma moeda honesta até que ocorra a face cara e observação das faces que ocorrem.

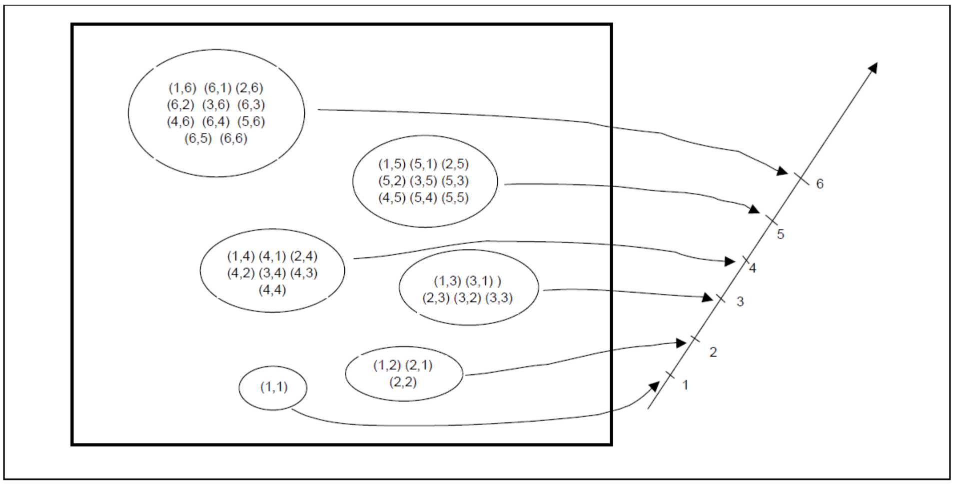


Figura 3.2: Variável aleatória: máximo de 2 dados.

Solução:

$$\Omega = \{k, ck, cck, ccck, cccck, \dots\},$$

Fonte: http://bit.do/eGFtD.

□

X = num. de coroas até que ocorra cara;
$$\Omega_X = \{0, 1, 2, 3, 4, \dots\},$$
$$X : \Omega \to \Omega_X$$
$$X(k) = 0,$$
$$X(ck) = 1;$$

Ω = num. de lançamentos até que ocorra cara;
$$\Omega_Y = \{1, 2, 3, 4, 5 \dots\},$$
$$Y : \Omega \to \Omega_Y$$
$$Y(k) = 1,$$
$$Y(ck) = 2;$$

Exemplo 3.2. Um homem possui 4 chaves em seu bolso. Como está escuro, ele não consegue ver qual a chave correta para abrir a porta de sua casa. Ele testa cada uma das chaves até encontrar a correta.

(a) Defina um espaço amostral para esse experimento.

(b) Defina a v.a. X = número de chaves experimentadas até conseguir abrir a porta (inclusive a chave correta). Quais são os valores de X?

Solução:

(a) Vamos designar por C a chave da porta e por E_1, E_2 e E_3 as outras chaves. Se ele para de testar as chaves depois que acha a chave correta, então o espaço amostral é:

$$\Omega=\{\ C, E_1C, E_2C, E_3C, E_1E_2C, E_2E_1C,\ E_1E_3C,\ E_3E_1C,\ E_2E_3C, E_3E_2C$$
$$E_1E_2E_3C, E_1E_3E_2C, E_2E_1E_3C, E_2E_3E_1C, E_3E_1E_2C, E_3E_2E_1C\ \}$$

(b) $\Omega_X = \{1, 2, 3, 4\}$.

□

Exemplo 3.3. Dentre os 5 alunos de um curso com coeficiente de rendimento (CR) superior 8.5, dois serão sorteados para receber uma bolsa de estudos. Os CRs desses alunos são: 8.8; 9.2; 8.9; 9.5; 9.0.

(a) Designando por A, B, C, D e E os alunos, defina um espaço amostral para esse experimento.

(b) Seja X = CR médio dos alunos sorteados. Liste os possíveis valores de X.

(c) Liste o evento $\{X \geqslant 9.0\}$.

Solução:

(a) Note que aqui a ordem não importa; logo, $\#\Omega = \binom{5}{2} = 10$. Mais especificamente,

$$\Omega = \{(A, B), (A, C), (A, D), (A, E), (B, C), (B, D), (B, E), (C, D), (C, E), (D, E)\}$$

(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)		9.00	8.85	9.15	8.90
B (9.2)			9.05	9.35	9.10
C (8.9)				9.20	8.95
D (9.5)					9.25
E (9.0)					

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

□

Exemplo 3.4. Numa urna há 7 bolas brancas e 4 bolas verdes. Cinco bolas são extraídas dessa urna. Defina a v.a. X = número de bolas verdes. Quais são os possíveis valores de X se as extrações são feitas:

(a) sem reposição;

(b) com reposição.

Solução:

(a) Como há apenas 4 verdes, os valores de X são 0, 1, 2, 3, 4. Note que temos bolas brancas em quantidade suficiente para que X = 0 (isto é, podemos tirar todas brancas).

(b) Se as extrações são feitas com reposição, em cada extração podemos tirar bola branca. Logo, os possíveis valores de X são 0, 1, 2, 3, 4, 5.

□

3.1.1. Função Massa de Probabilidade

Os valores de uma v.a. discreta são definidos a partir do espaço amostral de um experimento aleatório. Sendo assim, é natural perguntarmos *qual é a probabilidade do valor x?* No exemplo do máximo das 2 faces de um dado da Figura 3.2, por exemplo, o valor 6 da v.a. é imagem de 11 pontos do espaço amostral, enquanto o valor 2 é imagem de apenas 3 pontos. Sendo assim, é de se esperar que o valor 6 seja mais provável que o valor 2. Na verdade, temos a seguinte equivalência de eventos: se chamamos de X a v.a. *máximo das 2 faces*, então

$$\{X = 6\} \equiv \{(6,1),(6,2),(6,3),(6,4),(6,5),(6,6),(1,6),((2,6),((3,6),((4,6),((5,6)\}$$

e, assim

$$\mathbb{P}(X = 6) = \mathbb{P}\{(6,1) \cup (6,2) \cup (6,3) \cup (6,4) \cup (6,5) \cup (6,6) \cup (1,6) \cup (2,6) \cup (3,6) \cup (4,6) \cup (5,6)\}$$

Como os eventos expressão acima são mutuamente exclusivos e igualmente prováveis, resulta que

$$\mathbb{P}(X = 6) = 11 \times \frac{1}{36} = \frac{11}{36}.$$

De maneira análogo, temos

$$\mathbb{P}(X = 1) = \frac{1}{36}, \quad \mathbb{P}(X = 2) = \frac{3}{36} \quad \mathbb{P}(X = 3) = \frac{5}{36}$$
$$\mathbb{P}(X = 4) = \frac{7}{36} \quad \mathbb{P}(X = 5) = \frac{9}{36} \quad \mathbb{P}(X = 6) = \frac{11}{36}.$$

Definição 3.2. Seja X uma variável aleatória discreta e Ω_X o seu espaço amostral. A função massa de probabilidade $P(X = x)$, ou simplesmente $p_X(x)$, será a função que associa a cada valor de X a sua probabilidade de ocorrência, desde que atenda duas condições:

Função Massa de probabilidade:

1) $0 \leqslant p_X(x) \leqslant 1, \quad \forall x \in \Omega_X;$

2) $\displaystyle\sum_{x \in \Omega_X} p_X(x) = 1.$

O cálculo da fdp de uma v.a. X qualquer se dá em três etapas:

(i) primeiro, temos que identificar todos os possíveis valores x da v.a.X;

(ii) segundo, temos que identificar os resultados que dão origem a cada valor x e suas respectivas probabilidades;

(iii) finalmente, temos que somar todas essas probabilidades para obter $p_X(x)$.

Exemplo 3.5. Considerando novamente a v.a. definida na Figura 3.2 (máximo das duas faces), podemos resumir a fmp da variável em questão na seguinte tabela:

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

□

Exemplo 3.6. Consideremos novamente o lançamento de dois dados mas agora vamos definir a seguinte v.a. $X =$ *soma das 2 faces*. Para facilitar a solução desse problema, vamos construir uma tabela de duas entradas, onde cada dimensão representa o resultado de um dado e em cada cela temos a soma das duas faces.

	1	2	3	4	5	6
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

Como cada ponto do espaço amostral é equiprovável, a fmp de X é:

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

□

A função de massa de probabilidade de uma v.a. discreta X que assume um número finito de valores pode ser representada por um gráfico de colunas, onde a cada valor de X corresponde uma coluna cuja altura representa a probabilidade do respectivo valor. Na Figura 3.3 ilustra-se a fmp da v.a. X do Exemplo 3.6.

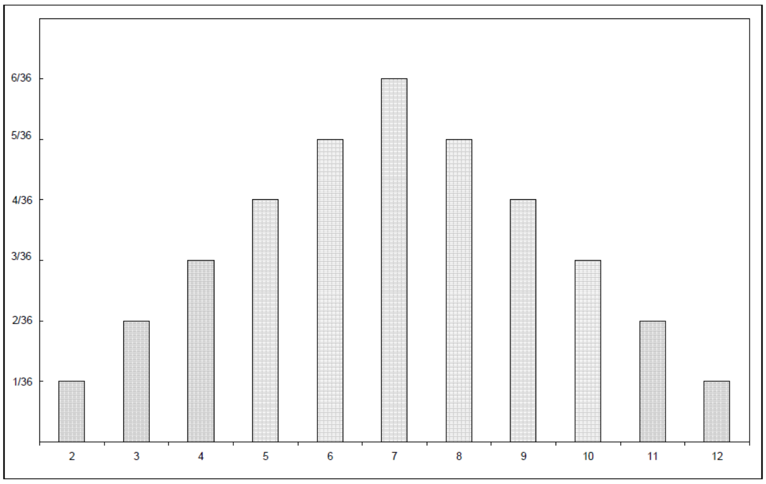


Figura 3.3: Função Massa de Probabilidade da v.a. X = *soma das faces de dois dados*.

Exemplo 3.7. Suponha que uma moeda é lançada 10 vezes e vamos definir a v.a. $X =$ *número de caras*. Suponhamos que a probabilidade de cara seja p e, por conseguinte, a probabilidade de coroa é $1 - p$. Os possíveis valores de X são $0, 1, 2, ..., 10$.

Solução: Vamos agora calcular a probabilidade de cada um desses valores, estabelecendo a equivalência dos eventos envolvidos. Para isso vamos usar a notação K_i = cara no i-ésimo lançamento e C_i = coroa no i-ésimo lançamento.

$$\{X = 0\} = \{\text{coroa nos 10 lançamentos}\} = \{C_1 \cap \cdots \cap C_{10}\}$$

Podemos considerar os lançamentos da moeda como eventos independentes. Logo,

$$\mathbb{P}(X=0)=\mathbb{P}(C_1)\times\cdots\times\mathbb{P}(C_{10})=(1-p)^{10}$$

O evento $\{X=1\}$ corresponde à ocorrência de 1 cara e 9 coroas. Uma sequência possível de resultados é $KCCCCCCCCC$ e a probabilidade é

$$\mathbb{P}(KCCCCCCCCC)=p(1-p)^9$$

Mas a sequência $CKCCCCCCCC$ também resulta em $\{X=1\}$. Na verdade existem $\binom{10}{1}$ tais sequências, todas com a mesma probabilidade. Logo

$$\mathbb{P}(X=1)=\binom{10}{1}p(1-p)^9.$$

Analogamente, o evento $\{X=2\}$ corresponde à ocorrência de 2 caras e 8 coroas; uma sequência possível é $KKCCCCCCCC$, que tem probabilidade

$$\mathbb{P}(KKCCCCCCCC)=p^2(1-p)^8$$

Mas, existem $\binom{10}{2}$ maneiras de colocar caras numa sequência de 10 lançamentos e todas tem a mesma probabilidade. Portanto,

$$\mathbb{P}(X=2)=\binom{10}{2}p^2(1-p)^8.$$

Em geral, para qualquer valor de x temos

$$\mathbb{P}(X=x)=\binom{10}{x}p^x(1-p)^{1-x},\quad x=0,1,2,\cdots,10.$$

Na Figura 3.3 apresentamos a fmp para diferentes valores de p.

□

Exemplo 3.8. Considere uma urna com 10 bolas, das quais 6 são vermelhas e 4 brancas. Dessa urna retiram-se 3 bolas sem reposição e conta-se o número de bolas brancas retiradas. Qual é a distribuição dessa variável aleatória?

Solução: Os possíveis valores de X são 0,1,2,3. Para calcular a probabilidade de cada um desses valores, devemos notar inicialmente que o espaço amostral tem $\binom{10}{3}$ eventos elementares. O evento $\{X=0\}$ corresponde à união dos eventos (sequências) onde não aparece nenhuma bola branca ou, equivalentemente, onde todas as bolas são vermelhas; o número de tais sequências é $\binom{6}{3}\binom{4}{0}=\binom{6}{3}$. Logo,

$$\mathbb{P}(X=0)=\frac{\binom{6}{3}\binom{4}{0}}{\binom{10}{3}}=\frac{20}{120}.$$

Analogamente, o evento $\{X=1\}$ corresponde à união dos eventos onde aparece 1 bola branca e 2 vermelhas. O número de tais sequências é $\binom{6}{2}\binom{4}{1}$ e, logo

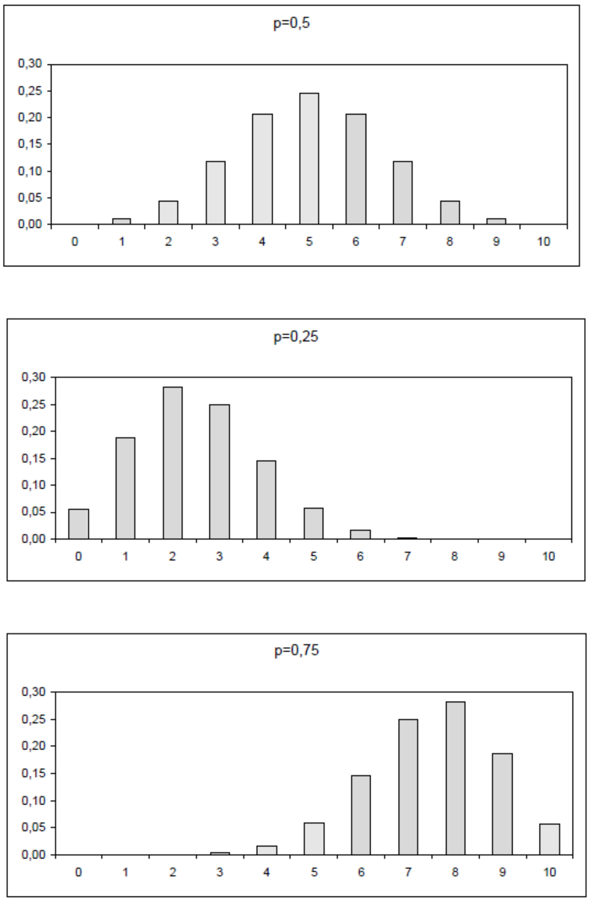


Figura 3.4: Função Massa de Probabilidade da v.a. X = *número de caras em 10 lançamentos*.

$$\mathbb{P}(X=1) = \frac{\binom{6}{2}\binom{4}{1}}{\binom{10}{3}} = \frac{60}{120}.$$

Analogamente,

$$\mathbb{P}(X=2) = \frac{\binom{6}{1}\binom{4}{2}}{\binom{10}{3}} = \frac{36}{120},$$

$$\mathbb{P}(X=3) = \frac{\binom{6}{0}\binom{4}{3}}{\binom{10}{3}} = \frac{4}{120}.$$

Portanto,

x	0	1	2	3
$p_X(x)$	$\frac{1}{6}$	$\frac{1}{2}$	$\frac{3}{10}$	$\frac{1}{30}$

□

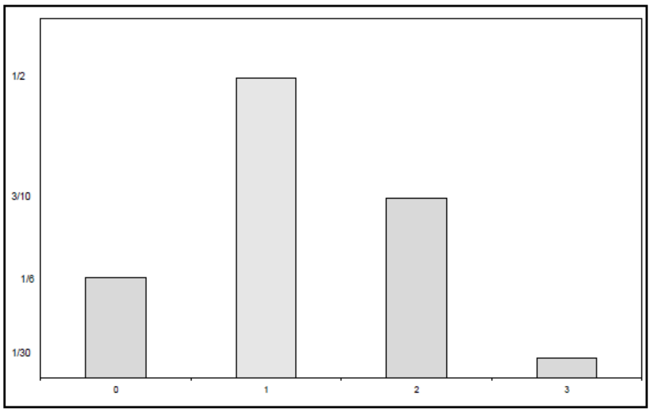


Figura 3.5: Função Massa de Probabilidade da v.a. X = *Número de bolas brancas em 3 extrações de uma urna com 6 vermelhas e 4 brancas.*

Exemplo 3.9. Considere a função dada na tabela abaixo.

x	0	1	2	3
$g(x)$	a	$3a$	$3a$	a

Estabeleça condições sobre a, de modo que a função $g(x)$ seja uma função massa de probabilidade.

Solução: Para que $g(x)$ seja uma função massa de probabilidade, temos que ter $0 \leqslant a \leqslant 1$ e $0 \leqslant 3a \leqslant 1$. Além disso,

$$0 + 3a + 3a + a = 1 \quad \to 8a = 1 \quad \to a = \frac{1}{8}$$

Portanto,

x	0	1	2	3
$g(x)$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

Assim, $g(x)$ é uma função massa de probabilidade.

□

3.1.2. Função de Distribuição ou Função de Probabilidade Acumulada

A partir da função massa de probabilidades de uma v.a. discreta X é possível calcular a probabilidade de qualquer evento associado a ela. Por exemplo, para a fmp da Figura 3.5, temos que

$$\mathbb{P}(X \geqslant 2) = \mathbb{P}(\{X = 2\} \cup \{X = 3\}) = \mathbb{P}(X = 2) + \mathbb{P}(X = 3) = \frac{1}{3}.$$

$$\mathbb{P}(X \leqslant 1) = \mathbb{P}(\{X = 0\} \cup \{X = 1\}) = \mathbb{P}(X = 0) + \mathbb{P}(X = 1) = \frac{2}{3}.$$

Então, podemos dizer que a fmp de uma variável aleatória discreta X nos dá toda a informação sobre X. Existe uma outra função com tal característica, que é a função de distribuição acumulada de X, cuja definição apresentamos a seguir.

Definição 3.3. Seja X uma variável aleatória discreta e Ω_X o seu espaço amostral. A função de distribuição, denotada por $F_X(x)$ ou $P(X \leqslant x)$, é a função que associa a cada valor de X a probabilidade $P(X \leqslant x)$. Desta forma, temos

$$F_X(x) = P(X \leqslant x) = \sum_{X \leqslant x} P(X = x) = \sum_{X \leqslant x} p_X(x). \tag{3.2}$$

Exemplo 3.10. Voltando ao Exemplo 3.5, temos que a fmp da v.a. $X =$ *máximo das duas faces* é dada por

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

Para calcular a fda de X, notemos inicialmente que nenhum valor menor que 1 é possível. Logo,

$$F_X(x) = 0, \text{ para todo } x < 1.$$

Para $x = 1$ devemos notar que

$$F_X(1) = P(X \leqslant 1) = \mathbb{P}(X < 1) + \mathbb{P}(X = 1) = 0 + \frac{1}{36} = \frac{1}{36}.$$

Para qualquer valor de x tal que $1 < x < 2$, temos que $p_X(x) = 0$. Logo,

$$F_X(x) = P(X \leqslant 1) + \mathbb{P}(1 < X < x) = F_X(1) + 0 = F_X(1) = \frac{1}{36} \text{ para todo } 1 \leqslant x < 2.$$

Analogamente, temos

$$F_X(2) = P(X \leqslant 2) = \mathbb{P}(X \leqslant 1) + \mathbb{P}(1 < X < 2) + \mathbb{P}(X = 2) = \frac{1}{36} + 0 + \frac{3}{36} = \frac{4}{36}.$$

$$F_X(x) = P(X \leqslant x) = \frac{4}{36} \text{ para todo } 2 \leqslant x < 3,$$

Seguindo o mesmo raciocínio temos

$$F_X(x) = P(X \leqslant x) = \frac{9}{36} \text{ para todo } 3 \leqslant x < 4,$$

$$F_X(x) = P(X \leqslant x) = \frac{16}{36} \text{ para todo } 4 \leqslant x < 5$$

$$F_X(x) = P(X \leqslant x) = \frac{25}{36} \text{ para todo } 5 \leqslant x < 6.$$

Para $x \geqslant 6$ devemos notar que o evento $\{X \leqslant x\}$ corresponde ao espaço amostral completo. Logo

$$F_X(x) = P(X \leqslant x) = 1, \text{ para todo } x \geqslant 6.$$

Resumindo,

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

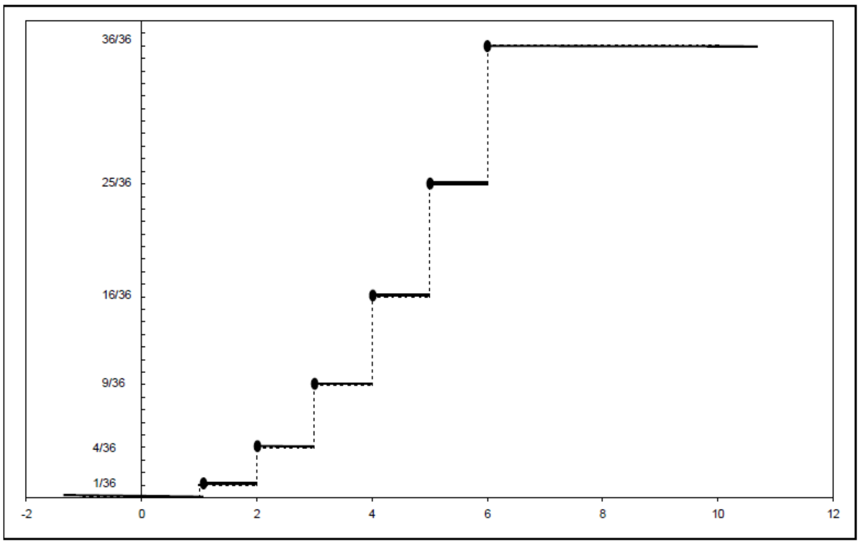


Figura 3.6: Função Distribuição Acumulada v.a. X = *máximo das duas faces.*

□

Proposição 3.1. *Seja X uma variável aleatória (discreta ou contínua) definida em $(\Omega, \mathcal{F}, \mathbb{P})$. A função de distribuição (ou função acumulada) desta variável aleatória, denotada por $F_X(\cdot)$, satisfaz as seguintes propriedades.*

(i) $0 \leqslant F_X(x) \leqslant 1$;

(ii) $\lim\limits_{x \to \infty} F_X(x) = 1$;

(iii) $\lim\limits_{x \to -\infty} F_X(x) = 0$;

(iv) *$F_X(x)$ é uma função não decrescente, isto é, para qualquer $x_1 < x_2$, temos $F_X(x_1) \leqslant F_X(x_2)$;*

Se A é um evento definido como $A = \{X \leqslant x_1\}$ e B é um evento definido como $B = \{X \leqslant x_2\}$, então, como $x_1 \leqslant x_2$, tem-se $A \subseteq B$, Portanto $\mathbb{P}(A) \leqslant \mathbb{P}(B)$.

(v) *$F_X(x)$ uma função contínua à direita.*

Observação 3.1. No caso de variáveis aleatórias discretas valem as seguintes relações:

(a) $\mathbb{P}(a < X \leqslant b) = F_X(b) - F_X(a)$

(b) $\mathbb{P}(a \leqslant X \leqslant b) = F_X(b) - F_X(a) + \mathbb{P}(X = a)$

(c) $\mathbb{P}(a < X < b) = F_X(b) - F_X(a) - \mathbb{P}(X = a)$ b

No caso de variáveis aleatórias ~~contínuas~~ valem as seguintes relações;

(a) $\mathbb{P}(X > a) = 1 - F_X(a)$

~~(b) $\mathbb{P}(a \leqslant X \leqslant b) = F_X(b) - F_X(a)$~~

Exemplo 3.11. Considere a v.a. X cuja fmp 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

Encontre a função de distribuição acumulada e faça o seu gráfico.

Solução: A função de distribuição da variável aleatória X cuja função massa de probabilidade é dada por

x	-2	-1	0	1	2	3
$p_X(x)$	0.1	0.2	0.2	0.3	0.1	0.1

é dada por

$$F_X(x) = \begin{cases} 0, & \text{se} \quad x < -2; \\ 0.1, & \text{se} \quad -2 \leqslant x < -1; \\ 0.3, & \text{se} \quad -1 \leqslant x < 0; \\ 0.5, & \text{se} \quad 0 \leqslant x < 1; \\ 0.8, & \text{se} \quad 1 \leqslant x < 2; \\ 0.9, & \text{se} \quad 2 \leqslant x < 3; \\ 1, & \text{se} \quad x \geqslant 3. \end{cases}$$

□

Exemplo 3.12. A variável X tem função de distribuição Acumulada dada por:

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

(a) Obtenha a correspondente função massa de probabilidade.

(b) Expresse $\mathbb{P}(X \geqslant 0)$ e $\mathbb{P}(X > 0)$ em termos de $F_X(x)$ e calcule seus valores.

(c) Expresse $\mathbb{P}(X \geqslant -1)$ e $\mathbb{P}(X > -1)$ em termos de $F_X(x)$ e calcule seus valores. Comente sobre as diferenças em relação aos resultados de **(b)**.

Solução:

(a) A função massa de probabilidade é dada por

x	-1	$\frac{1}{2}$	2
$p_X(x)$	$\frac{1}{2} - 0 = \frac{1}{2}$	$\frac{3}{4} - \frac{1}{2} = \frac{1}{4}$	$1 - \frac{3}{4} = \frac{1}{4}$

ou seja,

x	-1	$\frac{1}{2}$	2
$p_X(x)$	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{4}$

(b) Neste item iremos utilizar o resultado: $\mathbb{P}(A) = 1 - \mathbb{P}(A^c)$.

$$\begin{aligned}\mathbb{P}(X \geqslant 0) &= 1 - \mathbb{P}(X < 0) = 1 - \mathbb{P}(X < 0) - \mathbb{P}(X = 0) + \mathbb{P}(X = 0)\\ &= 1 - \mathbb{P}(X \leqslant 0) + \mathbb{P}(X = 0) = 1 - F_X(0) + \mathbb{P}(X = 0).\end{aligned}$$

Então,

$$\mathbb{P}(X \geqslant 0) = 1 - \frac{1}{2} - 0 = \frac{1}{2}.$$

Da mesma forma,

$$\mathbb{P}(X > 0) = 1 - \mathbb{P}(X \leqslant 0) = 1 - F_X(0) = 1 - \frac{1}{2} = \frac{1}{2}$$

(c) Neste item iremos utilizar o resultado: $\mathbb{P}(A) = 1 - \mathbb{P}(A^c)$.

$$\begin{aligned}\mathbb{P}(X \geqslant -1) &= 1 - \mathbb{P}(X < -1) = 1 - \mathbb{P}(X < -1) - \mathbb{P}(X = -1) + \mathbb{P}(X = -1)\\ &= 1 - \mathbb{P}(X \leqslant -1) + \mathbb{P}(X = -1) = 1 - F_X(-1) + \mathbb{P}(X = -1).\end{aligned}$$

Então,

$$\mathbb{P}(X \geqslant -1) = 1 - \frac{1}{2} - 0 = \frac{1}{2}.$$ Corrigir

Da mesma forma,

$$\mathbb{P}(X > -1) = 1 - \mathbb{P}(X \leqslant -1) = 1 - F_X(-1) = 1 - \frac{1}{2} = \frac{1}{2}.$$

□

3.2 Variáveis Aleatórias Contínuas

Definição 3.4. Seja E um experimento e Ω um espaço amostral associado. Se X é uma variável aleatória definida em Ω tal que $X(\Omega) = \Omega_X$ seja infinito não-enumerável, isto é, Ω_X seja um intervalo de números reais, então X é dita uma variável aleatória contínua.

Se X é uma variável aleatória contínua, X pode assumir qualquer valor num intervalo $[a, b]$ ou no intervalo $(-\infty; +\infty)$.

O espaço Ω_X será sempre definido como um intervalo do conjunto dos reais, sendo, portanto, um conjunto infinito.

Exemplos:

(i) tempo de vida de um animal;

(ii) vida útil de um componente eletrônico;

(iii) peso de uma pessoa;

(iv) quantidade de chuva que ocorre numa região;

(v) tempo de espera até a chegada do próximo cliente.

3.2.1. Função densidade de probabilidade

Os valores de uma variável aleatória contínua são definidos a partir do espaço amostral de um experimento aleatório. Sendo assim, é natural o interesse na probabilidade de obtenção de diferentes valores dessa variável. O comportamento probabilístico de uma variável aleatória contínua será descrito pela sua função de densidade de probabilidade.

Inicialmente apresentamos a definição da função de densidade de probabilidade utilizando a noção de área, para seguir a apresentação inicial que considerou um histograma de uma variável contínua.

Definição 3.5. Seja X uma variável aleatória contínua e Ω_X o seu espaço associado. Uma função $f_X(\cdot)$ associada a variável X é denominada função densidade de probabilidade (fdp) se satisfizer duas condições:

(i) $f_X(x) \geqslant 0, \quad$ para todo $x \in \Omega_X$;

(ii) $\displaystyle\int_{-\infty}^{\infty} f_X(x)dx = 1.$

Observação 3.2. Dada uma função $f_X(x)$ satisfazendo as propriedades acima, então $f_X(x)$ representa função densidade de probabilidade de alguma variável aleatória contínua X, de modo que $\mathbb{P}(a \leqslant X \leqslant b)$ é a área sob a curva limitada pelos pontos a e b, conforme a Figura 1 a seguir.

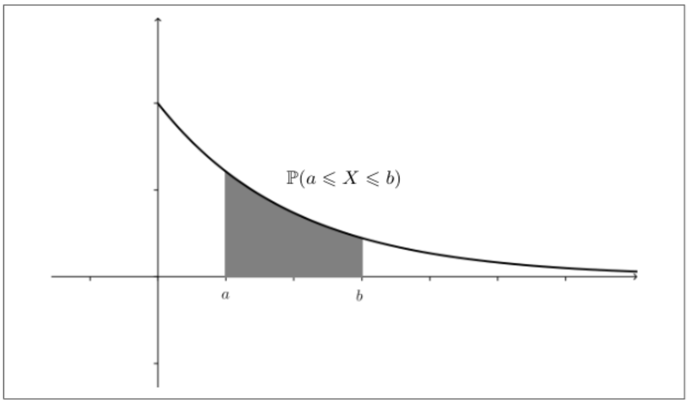


Figura 3.7: Probabilidade calculada atráves da função desnidade de probabilidade.

Para obter a probabilidade da variável aleatória estar em um certo intercalo $[a, b]$, fazemos a integral da função densidade de probabilidade no intervalo. Assim,

$$\mathbb{P}(a \leqslant X \leqslant b) = \int_a^b f_X(x).\mathrm{dx}$$

Uma primeira observação importante que resulta da interpretação geométrica de probabilidade como área sob a curva de densidade de probabilidade é a seguinte: se X é uma variável aleatória contínua, então a probabilidade do evento $[X = a]$ é zero, ou seja, a probabilidade de X ser exatamente igual a um valor específico é nula.

$$\mathbb{P}(X = a) = \int_a^a f_X(x)dx = 0$$

Como consequência, temos as seguintes igualdades:

$$\mathbb{P}(a \leqslant X \leqslant b) = \mathbb{P}(a \leqslant X < b) = \mathbb{P}(a < X \leqslant b) = \mathbb{P}(a < X < b) = \int_a^b f_X(x)dx.$$

Exemplo 3.13. Seja a função $f_X(x) = 2x$, no intervalo $\Omega_X = [0, 1]$. Verifique se a função abaixo é uma função densidade de probabilidade.

$$f_X(x) = \begin{cases} 2x, & \text{se } 0 \leqslant x \leqslant 1; \\ 0, & \text{c.c.} \end{cases} \tag{3.3}$$

Solução:

(i) $f_X(x) \geqslant 0, \quad$ para todo $x \in \Omega_X$;

(ii) $\displaystyle\int_{-\infty}^{\infty} f_X(x)dx = \int_{-\infty}^{0} 0dx + \int_0^1 2xdx + \int_1^{\infty} 0dx = 1.$

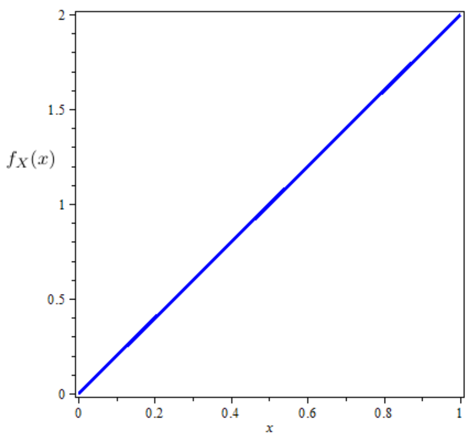


Figura 3.8: Função densidade de probabilidade dada pela equação (3.3).

□

Exemplo 3.14. Vamos avaliar para que valores da constante $c \in \mathbb{R}$, a função abaixo representa uma função densidade de probabilidade.

$$f(x) = \begin{cases} c(1-x)^2, & \text{se } 0 \leqslant x < \frac{1}{2}; \\ \frac{1}{(c+1)}, & \text{se } \frac{1}{2} \leqslant x \leqslant 1; \\ 0, & \text{c.c.} \end{cases} \tag{3.4}$$

Verifique as condições para que a função acima seja uma função densidade de probabilidade.

Solução:

As condições que precisam ser satisfeitas pela função $f(\cdot)$ são $f(x) \geqslant 0$ e $\displaystyle\int_{-\infty}^{\infty} f(x)dx = 1$. Observe que, se $c \geqslant 0$, temos $f(x)$ não negativa. Agora devemos obter os valores de c que satisfaçam a segunda condição. Assim, temos que

$$\int_{-\infty}^{\infty} f(x)dx = \int_{-\infty}^{0} 0dx + \int_{0}^{\frac{1}{2}} c(1-x)^2dx + \int_{\frac{1}{2}}^{1} \frac{1}{(c+1)}dx + \int_{1}^{\infty} 0dx = 1,$$

o que resulta em

$$c\frac{-(1-x)^3}{3}\Bigg|_0^{\frac{1}{2}} + \frac{x}{(c+1)}\Bigg|_{\frac{1}{2}}^{1} = 1 \quad \rightarrow \quad 7c^2 - 17c - 12 = 0.$$

Da mesma forma

$$\int_{0}^{\frac{1}{2}} c(1-x)^2dx + \int_{\frac{1}{2}}^{1} \frac{1}{(c+1)}dx = 1$$

$$\int_{0}^{\frac{1}{2}} c(1-2x+x^2)dx + \int_{\frac{1}{2}}^{1} \frac{1}{(c+1)}dx = 1$$

$$c\left[\int_{0}^{\frac{1}{2}} 1dx - 2\int_{0}^{\frac{1}{2}} xdx + \int_{0}^{\frac{1}{2}} x^2dx\right] + \int_{\frac{1}{2}}^{1} \frac{1}{(c+1)}dx = 1$$

$$c\left[x\Bigg|_0^{\frac{1}{2}} - 2\frac{x^2}{2}\Bigg|_0^{\frac{1}{2}} + \frac{x^3}{3}\Bigg|_0^{\frac{1}{2}}\right] + \frac{x}{(c+1)}\Bigg|_{\frac{1}{2}}^{1} = 1$$

$$c\left[\frac{1}{2} - \frac{1}{4} + \frac{1}{24}\right] + \frac{1}{2(c+1)} = 1$$

$$\frac{7c}{24} + \frac{1}{2(c+1)} = 1 \rightarrow 1 - \frac{7c}{24} = \frac{1}{2(c+1)} \rightarrow \frac{-7c+24}{24} = \frac{1}{2(c+1)}$$

$$-14c^2 + 48c - 14c + 48 = 24 \rightarrow -14c^2 + 34c + 24 = 0 \rightarrow 14c^2 - 34c - 24 = 0$$

Logo precisamos encontrar as raízes da equação $7c^2 - 17c - 12 = 0$.
A solução negativa dessa equação de 2 grau é descartada e obtemos $c = 3$.

$$f_X(x) = \begin{cases} 3(1-x)^2, & \text{se } 0 \leqslant x < \frac{1}{2}; \\ \frac{1}{4}, & \text{se } \frac{1}{2} \leqslant x \leqslant 1; \\ 0, & \text{c.c.} \end{cases} \tag{3.5}$$

□

Exemplo 3.15. A quantidade de tempo em horas que um computador funciona sem estragar é uma variável aleatória contínua com função densidade de probabilidade dada por

$$f(x) = \begin{cases} \lambda e^{-x/100}, & \text{se } x \geqslant 0; \\ 0, & \text{c.c. } x<0 \end{cases} \tag{3.6}$$

Qual a probabilidade de que:

(a) o computador funcione entre 50 e 100 horas antes de estragar?

(b) ele funcione menos de 100 horas?

Solução:

(a) Primeiramente precisamos encontra o valor de λ tal que a função acima seja uma função densidade de probabilidade. A primeira condição é que $f(x) \geqslant 0$. Para que essa condição seja satisfeita, temo que ter $\lambda > 0$. Para a segunda condição temos que ter

$$1 = \int_{-\infty}^{\infty} f(x)dx = \int_{-\infty}^{0} 0dx + \int_{0}^{\infty} \lambda e^{-x/100} dx.$$

Logo,

$$1 = -100\lambda e^{-x/100}\Big|_0^{\infty} \lambda \quad \rightarrow \quad 1 = 100\lambda \quad \rightarrow \quad \lambda = \frac{1}{100}.$$

Logo,

$$f(x) = \begin{cases} \frac{1}{100} e^{-x/100}, & \text{se } x \geqslant 0; \\ 0, & \text{c.c.} \end{cases} \tag{3.7}$$

Portanto, a probabilidade de que o computador funcione entre 50 e 100 horas antes de estragar antes de estragar é dada por

$$\mathbb{P}(50 < X < 100) = \int_{50}^{100} \frac{1}{100} e^{-x/100} dx = -e^{-x/100}\Big|_{50}^{100} = e^{-1/2} - e^{-3/2} \approx 0.384.$$

(b) Da mesma forma temos

$$\mathbb{P}(X < 100) = \int_{0}^{100} \frac{1}{100} e^{-x/100} dx = -e^{-x/100}\Big|_{0}^{100} = 1 - e^{-1} \approx 0.633.$$

□

3.2.2. Função de distribuição de Probabilidade ou Função Acumulada

Definição 3.6. Seja X uma variável aleatória contínua e Ω_X o seu espaço associado. A função de distribuição, denotada por $F_X(x) = \mathbb{P}(X \leqslant x)$, é a função que associa a cada ponto $x \in \Omega_X$ a probabilidade $P(X \leqslant x)$. Desta forma, tem-se:

$$F_X(x) = P(X \leqslant x) = \int_{-\infty}^{x} f_X(y)dy. \tag{3.8}$$

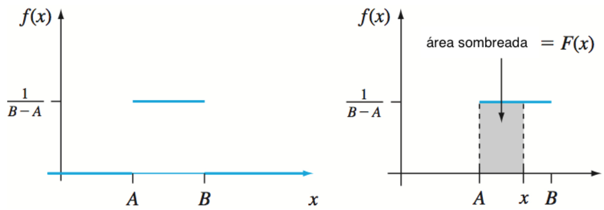


Figura 3.9: Função densidade de probabilidade dada pela equação (3.9).

Exemplo 3.16. Seja X a espessura de uma determinada chapa de metal, com função densidade de probabilidade dada por

$$f_X(x) = \frac{1}{B-A} I_{[A,B]}(x) \tag{3.9}$$

Encontre a função de distribuição.

Solução: Para $x < A$, temos

$$F_X(x) = \int_{-\infty}^{x} f_X(y)dy = \int_{-\infty}^{x} 0dy = 0.$$

Para $A \leqslant x < B$, temos

$$F_X(x) = \int_{-\infty}^{x} f_X(y)dy = \int_{A}^{x} \frac{1}{B-A}dy = \frac{1}{B-A}\, y\Big|_A^x = \frac{x-A}{B-A}.$$

Para $x \geqslant B$, temos

$$F_X(x) = \int_{-\infty}^{x} f_X(y)dy = \int_{A}^{B} \frac{1}{B-A}dy = 1.$$

Portanto a Função de distribuição acumulada é dada por

$$F_X(x) = \begin{cases} 0, & \text{se } x \leqslant A; \\ \frac{x-A}{B-A}, & \text{se } A \leqslant x < B; \\ 1, & \text{se } x \geqslant B. \end{cases} \tag{3.10}$$

O gráfico da função de distribuição dado pela equação (3.10) é dada pela Figura 3.10.

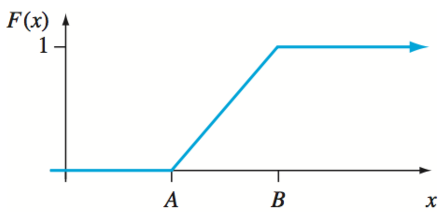


Figura 3.10: Função de Distribuição dada pela equação (3.10).

□

Exemplo 3.17. Seja X uma variável aleatória contínua que significa o tempo em minutos de um teste. A função densidade de probabilidade da variável aleatória X é dada por:

$$f_X(x) = \begin{cases} \frac{1}{40}(x-4), & \text{se } 8 \leqslant x < 10; \\ \frac{3}{20}, & \text{se } 10 \leqslant x \leqslant 15; \\ 0, & \text{c.c.} \end{cases} \tag{3.11}$$

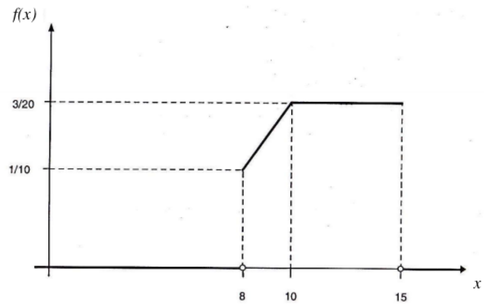


Figura 3.11: Gráfico da função de densidade de probabilidade dada pela equação (3.11).

Gráfico da função densidade de probabilidade dada pela equação (3.11).
Encontre a função de distribuição e calcule $\mathbb{P}(9 < X \leqslant 12)$.

Solução: Vamos encontrar a função de distribuição. Sabemos que se $x < 8$, temos $F_X(x) = \int_{-\infty}^{x} 0dy = 0$. Para $8 \leqslant x < 10$,

$$F_X(x) = \int_8^x \frac{1}{40}(y-4)dy = \frac{1}{40}\left(\frac{y^2}{2} - 4y\right)\Bigg|_8^x = \frac{1}{40}\left(\frac{x^2}{2} - 4x\right)$$

Para $10 \leqslant x < 15$,

$$F_X(x) = \int_8^{10} \frac{1}{40}(y-4)dy + \int_{10}^{x} \frac{3}{20}dy = \frac{1}{40}\left(\frac{y^2}{2} - 4y\right)\Bigg|_8^{10} + \frac{3y}{20}\Bigg|_{10}^{x} = \frac{3x}{20} - \frac{5}{4}$$

Para $x \geqslant 15$,

$$F_X(x) = \int_8^{10} \frac{1}{40}(y-4)dy + \int_{10}^{15} \frac{3}{20}dy = 1.$$

Portanto a função de distribuição acumulada da variável aleatória X é dada por

$$F_X(x) = \begin{cases} 0, & \text{se } x < 8; \\ \frac{1}{40}(\frac{x^2}{2} - 4x), & \text{se } 8 \leqslant x < 10; \\ \frac{3x}{20} - \frac{5}{4}, & \text{se } 10 \leqslant x < 15; \\ 1, & \text{se } x \geqslant 15. \end{cases} \tag{3.12}$$

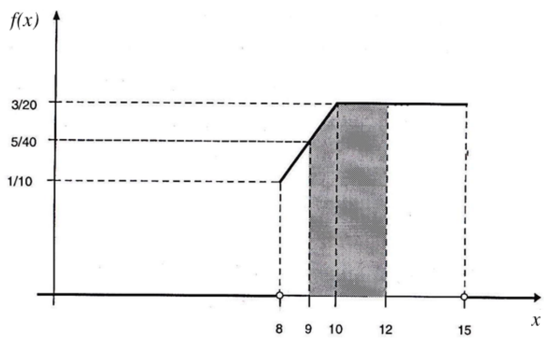


Figura 3.12: Gráfico da função densidade dada pela equação (3.11) para o cálculo da $\mathbb{P}(9 < X \leqslant 12)$.

Assim,

$$\mathbb{P}(9 < X \leqslant 12) = F_X(12) - F_X(9) = \int_8^{10} \frac{1}{40}(y-4)dy + \int_{10}^{12} \frac{3}{20}dy = \frac{7}{16}.$$

□

3.2.3. Obtendo $f_X(x)$ a partir de $F_X(x)$

Para X uma variável aleatória discreta a f.m.p. é obtida a partir da função de distribuição calculando-se a diferença entre dois valores da $F_X(x)$. O análogo contínuo de uma diferença é a derivada. O resultado a seguir é uma consequência do Teorema Fundamental do Cálculo.

Proposição 3.2. *Seja X uma variável aleatória contínua com função densidade $f_X(x)$ e função de distribuição $F_X(x)$, então, em cada x cuja derivada $F'_X(x)$ existe, $F'_X(x) = \frac{d}{dx}F_X(x) = f_X(x)$.*

Exemplo 3.18. Do Exemplo 2.24 temos a seguinte função de distribuição acumulada

$$F_X(x) = \begin{cases} 0, & \text{se } x \leqslant A; \\ \frac{x-A}{B-A}, & \text{se } A \leqslant x < B; \\ 1, & \text{se } x \geqslant B. \end{cases} \tag{3.13}$$

Solução: Derivando temos

$$f_X(x) = \frac{1}{B-A} I_{[A,B]}(x) \tag{3.14}$$

□

3.2.4. Percentis de uma distribuição contínua

Quando dizemos que a pontuação de um indivíduo estava no 85^0 percentil da população, queremos dizer que 85% de todas as pontuações da população estavam abaixo daquela pontuação e 15% acima dela.

Definição 3.7. Seja p um número entre 0 e 1. O $(100p)-$ésimo percentil da distribuição de uma variável aleatória X, representado por $\eta(p)$, é definido por:

$$p = F_X(\eta(p)) = \mathbb{P}(X \leqslant \eta(p)) = \int_{-\infty}^{\eta(p)} f_X(y)dy. \tag{3.15}$$

Pela equação (3.15), $\eta(p)$ é o valor no eixo x tal que $100(p)\%$ da área sob o gráfico de $f_X(x)$ encontra-se à esquerda de $\eta(p)$ e $100(1-p)\%$ encontra-se à direita.

Exemplo 3.19 (Devore, 2016). A distribuição da quantidade de cascalho (em toneladas) vendida para uma determinada loja de materiais de construção em uma determinada semana é uma v.a. contínua X com função densidade de probabilidade dada por

$$f_X(x) = \frac{3}{4}(1-x^2)I_{[0,1]}(x). \tag{3.16}$$

Solução: A função de distribuição da v.a. é dada por

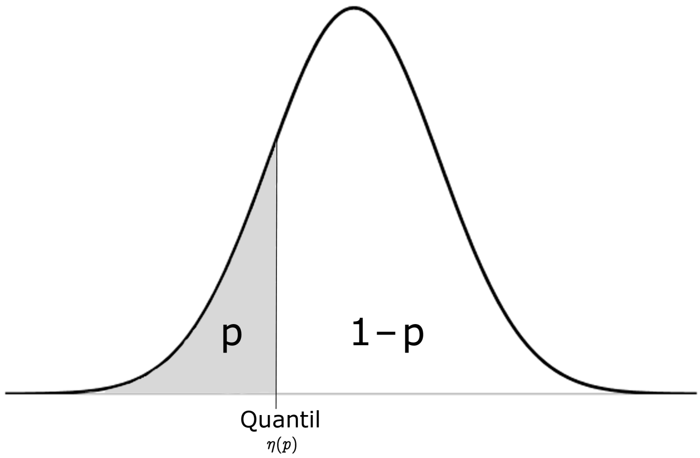


Figura 3.13: Quantil de probabilidade $100(p)\%$ da distribuição de uma variável aleatória contínua X.

$$F_X(x) = \begin{cases} 0, & \text{se } x \leqslant 0; \\ \frac{3}{2}\left(x - \frac{x^3}{3}\right), & \text{se } 0 \leqslant x < 1; \\ 1, & \text{se } x \geqslant 1. \end{cases} \tag{3.17}$$

Os gráficos das f.d.p. e da f.d. dados, respectivamente, pelas equações (3.11) e (3.10) são apresentados na Figura 3.14. O $100(p)-$ésimo percentil dessa distribuição satisfaz a equação

$$p = F_X(\eta(p)) = \frac{3}{2}\left(\eta(p) - \frac{(\eta(p))^3}{3}\right)$$

ou seja,

$$(\eta(p))^3 - 3\eta(p) + 2p = 0.$$

Para encontrarmos o 50^0 percentil, assumimos $p = 0.5$, e a equação a ser resolvida é $(\eta(0.5))^3 - 3\eta(0.5) + 1 = 0$; a solução é $\eta(0.5) = 0.347$. Se a distribuição continuar de semana para semana, no longo prazo, 50% de todas as semanas resultarão em vendas de menos de 0.347 toneladas e 50% em mais de 0.347 toneladas.

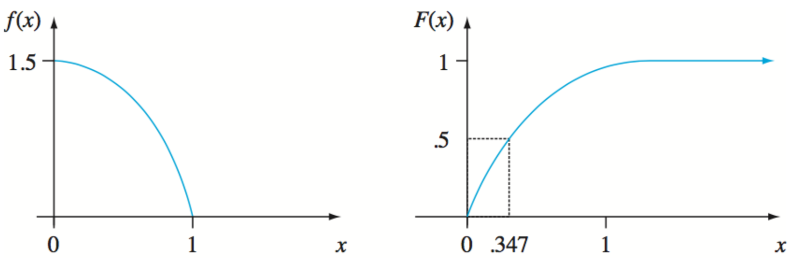


Figura 3.14: A f.d.p. e da f.d. dados, respectivamente, pelas equações (3.11) e (3.10) .

□

3.3 Funções de Variável Aleatória

Com frequência, conhecemos a distribuição de probabilidade de uma variável aleatória e estamos interessados em determinar a distribuição de alguma função dessa variável. Por exemplo, suponha que conheçamos a distribuição de X e queiramos conhecer a distribuição de $g(X)$. Para fazer isso, é necessário expressar o evento em que $g(X) \leqslant y$ em termos de X em um conjunto. Isso é ilustrado nos exemplos a seguir.

Inicialmente iremos estudar as funções de uma variável aleatória, isto é, funções do tipo X^n, e^X, etc, ou seja, funções $g(X)$, para alguma função $g : \mathbb{R} \to \mathbb{R}$.

Resultado: Sendo X uma variável aleatória em $(\Omega, \mathcal{F}, \mathbb{P})$, a função $g : \mathbb{R} \to \mathbb{R}$ também será uma variável aleatória no mesmo espaço de probabilidade.

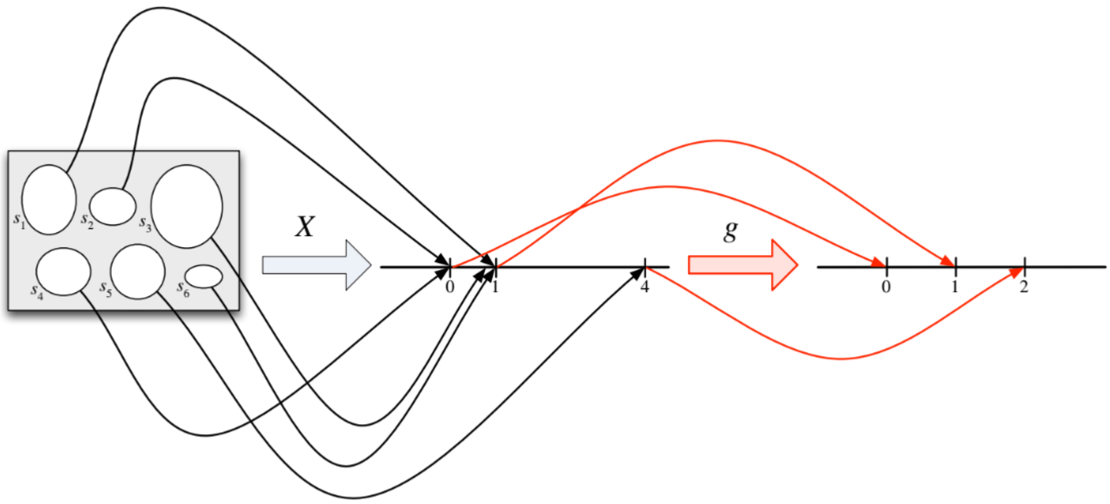


Figura 3.15: Funções de Variáveis Aleatórias.

Fonte: Blitzstein e Hwang (2019).

Conhecendo a função de distribuição, função massa de probabilidade ou função densidade de probabilidade de X, desejamos obter o comportamento de $g(X)$. Em termos matemáticos, dizemos que $Y = g(X)$ é uma transformação de X.

Para obtermos o comportamento probabilístico de transformações uma técnica muito conveniente, principalmente para o caso discreto, é chamado de *método direto*. Ele consiste em realizar operações algébricas simples, aplicando a definição da transformação diretamente na expressão da função de distribuição (ou função densidade ou de massa de probabilidade conforme o caso).

Variável Aleatória Discreta: Para X uma variável aleatória discreta, com função de massa de probabilidade conhecida, como podemos encontrar a função massa de probabilidade da função $Y = g(X)$? No caso de $g(\cdot)$ uma função bijetiva, a resposta é direta: a imagem de Y é o conjunto de todos os $g(x)$ com $x \in \Omega_X$, e

$$\mathbb{P}(Y = y) = \mathbb{P}(g(X) = g(x)) = \mathbb{P}(X = x).$$

O caso onde $Y = g(X)$ é uma bijeção está ilustrado nas tabelas a seguir.

x	$P(X=x)$
x_1	p_1
x_2	p_2
x_3	p_3
$\vdots$	$\vdots$

y	$P(Y=y)$
$g(x_1)$	p_1
$g(x_2)$	p_2
$g(x_3)$	p_3
$\vdots$	$\vdots$

Fonte: Blitzstein e Hwang (2019).

Resultado: Dada uma variável aleatória X com função de distribuição (ou função massa ou densidade de probabilidades) conhecida, a distribuição de uma variável aleatória $Y = g(X)$, onde h é uma função Borel-mensurável é determinada por

$$\mathbb{P}(Y \leqslant y) = \mathbb{P}(X \in g^{-1}(-\infty, y]).$$

Exemplo 3.20. Seja X uma variável aleatória com função massa de probabilidade dada por

x	0	1	2
$p_X(x)$	1/3	1/3	1/3

Encontre a função massa de probabilidade da função $Y = 3X + 1$.

Solução: Temos que $A_x = \{0, 1, 2\}$ e $A_y = \{1, 4, 7\}$. Então,

$$\mathbb{P}(Y = 1) = \mathbb{P}(X = 0) = \frac{1}{3}$$

$$\mathbb{P}(Y = 4) = \mathbb{P}(X = 1) = \frac{1}{3}$$

$$\mathbb{P}(Y = 7) = \mathbb{P}(X = 2) = \frac{1}{3}$$

Ou seja,

y	1	4	7
$p_Y(y)$	1/3	1/3	1/3

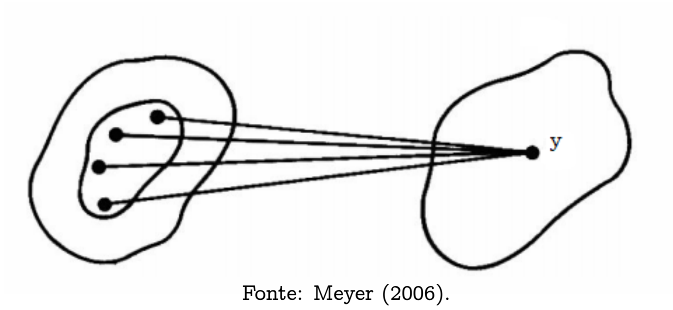
□

Se $g(\cdot)$ não é uma bijeção, então para um dado y podem existir múltiplos valores de x tal que $g(x) = y$. Para calcular $\mathbb{P}(G(X) = y)$, é necessário somar a probabilidade de X para todos os valores de X tal que $\mathbb{P}(G(X) = y)$.

Teorema 3.1. *Seja X uma variável aleatória discreta e $g : \mathbb{R} \to \mathbb{R}$. Então a imagem de $g(X)$ é o conjunto de todos os y tal que $g(x) = y$ para pelo menos um x na imagem de X, e a função massa de probabilidade de $Y = g(X)$ é dada por*

$$\mathbb{P}(Y = y) = \mathbb{P}(g(X) = y) = \sum_{x : g(x) = y} \mathbb{P}(X = x),$$

para todo y pertencente a imagem de $g(X)$.



Fonte: Meyer (2006).

Exemplo 3.21. Seja X uma variável aleatória com função massa de probabilidade dada por

x	-2	-1	0	1	2
$p_X(x)$	1/5	1/6	1/5	1/15	11/30

Encontre a função massa de probabilidade da função $Y = X^2$.

Solução: Temos que $A_x = \{-2, -1, 0, 1, 2\}$ e $A_y = \{0, 1, 4\}$. Então,

$$\mathbb{P}(Y = 0) = \mathbb{P}(X = 0) = \frac{1}{5}$$

$$\mathbb{P}(Y = 1) = \mathbb{P}(X = -1) + \mathbb{P}(X = 1) = \frac{1}{6} + \frac{1}{15} = \frac{7}{30}$$

$$\mathbb{P}(Y = 4) = \mathbb{P}(X = -2) + \mathbb{P}(X = 2) = \frac{1}{5} + \frac{11}{30} = \frac{17}{30}$$

Ou seja,

y	0	1	4
$p_Y(y)$	1/5	7/30	17/30

□

Exemplo 3.22. Seja X uma variável aleatória com função massa de probabilidade dada por

$$f_X(x) = \mathbb{P}(X = x) = \left(\frac{1}{2}\right)^n, \quad \text{para } x \in \{1, 2, 3, \cdots\}.$$

Encontre a função massa de probabilidade da função

$$Y = 1, \text{ se } X \text{ for par},$$

$$Y = -1, \text{ se } X \text{ for ímpar},$$

Solução: Temos que $Y = 1$ para $A = \{2, 4, 6, 8, \cdots\}$ e $Y = -1$ para $A = \{1, 3, 5, 7, \cdots\}$. Então, pelo Teorema 3.1, temos

$$\mathbb{P}(Y = 1) = \frac{1}{2^2} + \frac{1}{2^4} + \frac{1}{2^6} + \cdots = \sum_{n \geqslant 1} \frac{1}{4}\left(\frac{1}{4}\right)^{n-1}$$

Ou seja, temos uma progressão geométrica. Logo

$$\mathbb{P}(Y = 1) = \frac{a_1}{1 - q} = \frac{1/4}{1 - (1/4)} = \frac{1}{3}.$$

Como $\Omega_Y = \{-1, 1\}$, $\mathbb{P}(Y = -1) + \mathbb{P}(Y = 1) = 1$. Portanto,

$$\mathbb{P}(Y = -1) = 1 - \mathbb{P}(Y = 1) = 1 - \frac{1}{3} = \frac{2}{3}.$$

□

Exemplo 3.23. Seja X uma variável aleatória com distribuição Poisson, ou seja,

$$p_X(x) = \mathbb{P}(X = x) = \frac{e^{-\lambda}\lambda^x}{x!}, \quad x = 0, 1, 2, \cdots, \lambda > 0,$$

Considere a função $Y = X^2 + 3$. Encontra a função massa de probabilidade da variável aleatória Y.

Solução: Temos que $y = h(x) = x^2 + 3$, com $A_x = \{0, 1, 2, \cdots\}$ com $A_y = \{3, 4, 7, 12, 19, 28, \cdots\}$. A função inversa é dada por $x = \sqrt{y-3}$, e como não há valores negativos em A, tomamos a raiz quadrada positiva de $y - 3$. Assim,

$$\mathbb{P}(Y = y) = \mathbb{P}(X = \sqrt{y-3}) = \frac{e^{-\lambda}\lambda^{\sqrt{y-3}}}{\sqrt{y-3!}}, y \in A_y$$

e $\mathbb{P}(Y = y) = 0$ caso contrário.

□

Variável Aleatória Contínua: Agora considere X uma variável aleatória contínua. Segue o seguinte resultado.

Resultado: Seja X uma variável aleatória contínua com função densidade de probabilidade $f_X(x)$. Seja $y = g(x)$ uma função estritamente monótona (crescente ou decrescente) e diferenciável para todo x (portanto contínua). Então a variável aleatória $Y = g(X)$ é também contínua.
Se g é diferenciável para todo x e $g(x) > 0$ para todo x, então $g(\cdot)$ é contínua e estritamente crescente e a função inversa $x = g^{-1}(y)$ existe e é estritamente crescente e é diferenciável. Então a função de distribuição de $Y = g(X)$ é dada por

$$F_Y(y) = \mathbb{P}(Y \leqslant y) = \mathbb{P}(g(X) \leqslant y) = \mathbb{P}(X \leqslant g^{-1}(y)) = F_X(g^{-1}(y)).$$

A função densidade de probabilidade $Y = g(X)$ é obtida diferenciando a função de distribuição, ou seja,

$$f_Y(y) = \frac{d}{dy}F_Y(y) = \frac{d}{dy}F_X(g^{-1}(y)) = f_X(g^{-1}(y))\frac{d}{dy}(g^{-1}(y)) \tag{3.18}$$

Da mesma forma, se a derivada de g for negativa, então g é estritamente decrescente e temos

$$F_Y(y) = \mathbb{P}(Y \leqslant y) = \mathbb{P}(g(X) \leqslant y) = \mathbb{P}(X \geqslant g^{-1}(y)) = 1 - \mathbb{P}(X \leqslant g^{-1}(y)) = 1 - F_X(g^{-1}(y)).$$

Derivando temos

$$f_Y(y) = \frac{d}{dy}F_Y(y) = \frac{d}{dy}[1 - F_X(g^{-1}(y))] = -f_X(g^{-1}(y))\frac{d}{dy}(g^{-1}(y)). \tag{3.19}$$

Como g e g^{-1} são ambas estritamente decrescentes, $\frac{d}{dy}(g^{-1}(y))$ é negativa e o resultado dado pela equação 3.18 segue.

Exemplo 3.24. [Meyer, 2006] Seja X uma variável aleatória com função densidade de probabilidade dada por

$$f_X(x) = \begin{cases} 2x, & \text{se } 0 < x < 1; \\ 0, & \text{caso contrário;} \end{cases}$$

Considere a função $Y = 3X+1$. Encontra a função densidade de probabilidade da variável aleatória Y.

Solução: No caso de variáveis aleatórias contínuas começamos encontrando a função de distribuição (acumulada) da variável aleatória X.

c

Aplicando o resultado acima temos, onde $g^{-1}(y) = \frac{y-1}{3}$, temos

$$f_Y(y) = f_X(g^{-1}(y))\frac{d}{dy}(g^{-1}(y)) = 2\left(\frac{y-1}{3}\right)\frac{1}{3} = \frac{2}{9}(y-1).$$

Desde que $f_X(x) > 0$ para $0 < x < 1$, encontramos que $f_Y(y) > 0$ para $1 < y < 4$. Portanto,

$$f_Y(y) = \begin{cases} \frac{2}{9}(y-1), & \text{se } 1 < y < 4; \\ 0, & \text{caso contrário;} \end{cases}$$

Existe uma maneira ligeiramente diferente de obtermos o mesmo resultado. Considere a função de distribuição das variáveis aleatórias. Então,

$$\begin{aligned} F_Y(y) &= \mathbb{P}(Y \leqslant y) = \mathbb{P}(3X+1 \leqslant y) \\ &= \mathbb{P}\left(X \leqslant \frac{y-1}{3}\right) \\ &= F_X\left(\frac{y-1}{3}\right) \\ &= \int_0^{\frac{y-1}{3}} 2x dx = \left(\frac{y-1}{3}\right)^2. \end{aligned}$$

Portanto,

$$F_Y(y) = \begin{cases} 0, & \text{se } y < 1; \\ \left(\frac{y-1}{3}\right)^2 & \text{se } 1 \leqslant y < 4; \\ 1, & \text{se} y \geqslant 4; \end{cases}$$

□

Exemplo 3.25. [Ross, 2010] Suponha que X seja uniformemente distribuída, ou seja, sua função densidade de probabilidade é dada por

$$f_X(x) = \begin{cases} \frac{1}{b-a}, & \text{se } a < x < b, \quad a < b, \text{ com } a, b \geqslant 0; \\ 0, & \text{caso contrário;} \end{cases} \tag{3.20}$$

Queremos obter a distribuição da variável aleatória Y, definida como $Y = X^n$.

Solução: Como X é uma variável aleatória contínua, temos que $Y = g(X) = X^n$ também é uma variável aleatória contínua. Assim, primeiramente temos que encontrar a função de distribuição. Assim,

$$\begin{aligned} F_Y(y) &= \mathbb{P}(Y \leqslant y) = \mathbb{P}(g(X) \leqslant y) = \mathbb{P}(X^n \leqslant y) \\ &= \mathbb{P}(X \leqslant y^{\frac{1}{n}}) = F_X\left(y^{\frac{1}{n}}\right). \end{aligned} \tag{3.21}$$

Temos dois caminhos.

1) Derivamos a equação (3.21), em relação a y e obtemos a função densidade. Ou seja,

$$f_Y(y) = \frac{d}{dy} F_Y(y) = \frac{d}{dy} F_X\left(y^{\frac{1}{n}}\right) = f_X\left(y^{\frac{1}{n}}\right) \frac{y^{\frac{1}{n-1}}}{n}$$

Então,

$$f_Y(y) = \frac{1}{b-a} \frac{y^{\frac{1}{n-1}}}{n}.$$

Qual o intervalo $f_Y(y)$ é a função acima e qual o intervalo $f_Y(y)$ é nula?

Temos que $y = x^n$, ou seja, $x = \sqrt[n]{y}$. Assim,

$$a < x < b \quad \rightarrow \quad a < \sqrt[n]{y} < b \quad \rightarrow \quad a^n < y < b^n$$

Caso $a = 1$ e $b = 2$, então $a^n < y < b^n \rightarrow (1)^n < y < (2)^n$.

2) Encontramos a função de distribuição e após é calculada a derivada.

A função de distribuição da variável aleatória X, cuja função densidade é dada pela equação (3.20), é dada por

$$F_X(x) = \begin{cases} 0, & \text{se } x < a; \\ \frac{x-a}{b-a}, & \text{se } a \leqslant x < b; \\ 1, & \text{se } x \geqslant b; \end{cases} \tag{3.22}$$

Então,

$$F_Y(y) = F_X\left(y^{\frac{1}{n}}\right) = \begin{cases} 0, & \text{se } y < a^n; \\ \frac{y^{\frac{1}{n}}-a}{b-a}, & \text{se } a^n \leqslant y < b^n; \\ 1, & \text{se } y \geqslant b^n; \end{cases}$$

Assim,

$$f_Y(y) = \frac{d}{dy} F_Y(y) = \frac{d}{dy} F_X\left(y^{\frac{1}{n}}\right) = \frac{1}{b-a} \frac{y^{\frac{1}{n-1}}}{n} I_{[a^n;b^n]}(y)$$

□

Exemplo 3.26. Seja X uma variável aleatória contínua com função densidade de probabilidade $f_X(x)$. Encontre a distribuição de $Y = X^2$.

Solução: Como X é uma variável aleatória contínua, temos que $Y = g(X) = X^2$ também é uma variável aleatória contínua. Assim, primeiramente temos que encontrar a função de distribuição. Assim,

$$\begin{aligned} F_Y(y) &= \mathbb{P}(Y \leqslant y) = \mathbb{P}(g(X) \leqslant y) = \mathbb{P}(X^2 \leqslant y) \\ &= \mathbb{P}(-\sqrt{y} \leqslant X \leqslant \sqrt{y}) \\ &= \mathbb{P}(X \leqslant \sqrt{y}) - \mathbb{P}(X < -\sqrt{y}) - \mathbb{P}(X = -\sqrt{y}) + \mathbb{P}(X = -\sqrt{y}) \\ &= F_X(\sqrt{y}) - F_X(-\sqrt{y}) + \mathbb{P}(X = -\sqrt{y}). \end{aligned} \tag{3.23}$$

Como estamos trabalhando com variáveis aleatórias contínuas, $\mathbb{P}(X = -\sqrt{y}) = 0$, então

$$F_Y(y) = F_X\left(\sqrt{y}\right) - F_X\left(-\sqrt{y}\right). \tag{3.24}$$

Derivamos a equação (3.24), em relação a y e obtemos a função densidade. Ou seja,

$$\begin{aligned} f_Y(y) &= \frac{d}{dy}F_Y(y) = \frac{d}{dy}F_X\left(\sqrt{y}\right) - \frac{d}{dy}F_X\left(-\sqrt{y}\right) \\ &= \frac{1}{2\sqrt{y}}f_X\left(\sqrt{y}\right) - \frac{-1}{2\sqrt{y}}f_X\left(-\sqrt{y}\right) \\ &= \frac{1}{2\sqrt{y}}\left[f_X\left(\sqrt{y}\right) + f_X\left(-\sqrt{y}\right)\right] \end{aligned} \tag{3.25}$$

Suponha agora que $f_X(x) = \frac{1}{2}I_{(-1,1)}(x)$. Então, $f_Y(y)$, onde $Y = X^2$ é dada por

$$f_Y(y) = \frac{1}{2\sqrt{y}}\left[f_X\left(\sqrt{y}\right) + f_X\left(-\sqrt{y}\right)\right]$$

Como $Y = g(X) = X^2$, temos que $x = \sqrt{y}$

$$-1 < x < 1 \quad \to \quad -1 < \sqrt{y} < 1 \quad \to \quad 0 < y < 1$$

assim,

$$f_Y(y) = \frac{1}{2\sqrt{y}}\left[f_X\left(\sqrt{y}\right) + f_X\left(-\sqrt{y}\right)\right]$$

$$f_Y(y) = \frac{1}{2\sqrt{y}}\left[\frac{1}{2} + \frac{1}{2}\right]$$

$$f_Y(y) = \frac{1}{2\sqrt{y}}I_{(0,1)}(y)$$

$\square$

Exemplo 3.27. [Meyer, 2006] Seja X uma variável aleatória com função densidade de probabilidade dada por

$$f_X(x) = \frac{1}{2}I_{[1,3]}(x).$$

Considere a função $Y = e^X$. Encontra a função acumulada e densidade de probabilidade da variável aleatória Y.

Solução: Observe que Y é uma variável aleatória contínua e com valores no intervalo $[e, e^3]$. Sendo $F_Y(\cdot)$ dua função de distribuição, temos de imediato que $F_Y(y) = 0$, se $y < e$, e $F_Y(y) = 1$, se $y \geqslant e^3$. Para $e \leqslant Y < e^3$,

$$\begin{aligned} F_Y(y) &= \mathbb{P}(Y \leqslant y) = \mathbb{P}(e^X \leqslant y) \\ &= \mathbb{P}(X \leqslant \log(y)) = \int_{-\infty}^{\log(y)} \frac{1}{2}I_{[1,3]}(x)dx \\ &= \frac{\log(y) - 1}{2}. \end{aligned}$$

Dessa forma,

$$F_Y(y) = \begin{cases} 0, & \text{se } y < e; \\ \frac{\log(y)-1}{2} & \text{se } e \leqslant Y < e^3; \\ 1, & \text{se} y \geqslant e^3; \end{cases}$$

Derivando a função de distribuição obtemos a função densidade de probabilidade,

$$f_Y(y) = \frac{1}{2y} I_{[e,e^3]}(y).$$

□

Exemplo 3.28. Seja X uma variável aleatória contínua com função densidade de probabilidade

$$f_X(x) = \frac{1}{2} I_{(-1,1)}(x)$$

Encontre a distribuição de $Y = 4 - X^2$.

Solução: Como X é uma variável aleatória contínua, temos que $Y = g(X) = 4 - X^2$ também é uma variável aleatória contínua. Assim, primeiramente temos que encontrar a função de distribuição. Assim,

$$\begin{aligned} F_Y(y) &= \mathbb{P}(Y \leqslant y) = \mathbb{P}(g(X) \leqslant y) = \mathbb{P}(4 - X^2 \leqslant y) \\ &= \mathbb{P}(-X^2 \leqslant y - 4) = \mathbb{P}(X^2 \geqslant 4 - y) = 1 - \mathbb{P}(X^2 \leqslant 4 - y) \\ &= 1 - \mathbb{P}(-\sqrt{4-y} \leqslant X \leqslant \sqrt{4-y}) \\ &= 1 - \left[\mathbb{P}(X \leqslant \sqrt{4-y}) - \mathbb{P}(X < -\sqrt{4-y}) - \mathbb{P}(X = -\sqrt{4-y}) + \mathbb{P}(X = -\sqrt{4-y})\right] \\ &= 1 - \left[\mathbb{P}(X \leqslant \sqrt{4-y}) - \mathbb{P}(X \leqslant -\sqrt{4-y}) + \mathbb{P}(X = -\sqrt{4-y})\right] \\ &= 1 - \left[F_X(\sqrt{4-y}) - F_X(-\sqrt{4-y}) + \mathbb{P}(X = -\sqrt{4-y})\right]. \end{aligned} \tag{3.26}$$

Por estarmos tratando com uma variável aleatória contínua, $\mathbb{P}(X = -\sqrt{4-y}) = 0$. Assim, derivamos a equação (3.26), em relação a y e obtemos a função densidade. Ou seja,

$$\begin{aligned} f_Y(y) &= \frac{d}{dy} F_Y(y) = -\frac{d}{dy} F_X\left(\sqrt{4-y}\right) + \frac{d}{dy} F_X\left(-\sqrt{4-y}\right) \\ &= -\frac{-1}{2\sqrt{4-y}} f_X\left(\sqrt{4-y}\right) + \frac{1}{2\sqrt{4-x}} f_X\left(-\sqrt{4-y}\right) \\ &= \frac{1}{2\sqrt{4-y}} \left[f_X\left(\sqrt{4-y}\right) + f_X\left(-\sqrt{4-y}\right)\right] \end{aligned} \tag{3.27}$$

Suponha agora que $f_X(x) = \frac{1}{2} I_{(-1,1)}(x)$. Então, $f_Y(y)$, onde $Y = 4 - X^2$ é dada por

$$f_Y(y) = \frac{1}{2\sqrt{4-y}} \left[f_X\left(\sqrt{4-y}\right) + f_X\left(-\sqrt{4-y}\right)\right]$$

Como $Y = g(X) = 4 - X^2$, temos que $x = \sqrt{4-y}$

$$-1 < x < 1 \quad \rightarrow \quad -1 < \sqrt{4-y} < 1 \quad \rightarrow \quad 3 < y < 4$$

assim,

$$f_Y(y) = \frac{1}{2\sqrt{4-y}} \left[f_X(\sqrt{4-y}) + f_X\left(-\sqrt{4-y}\right)\right]$$

$$f_Y(y) = \frac{1}{2\sqrt{4-y}}\left[\frac{1}{2}+\frac{1}{2}\right]$$

$$f_Y(y) = \frac{1}{2\sqrt{4-y}} I_{(0,1)}(y) \quad [3,4]$$

□

Exemplo 3.29. Seja X uma variável aleatória com função massa de probabilidade ou função densidade de probabilidade. Encontre a distribuição de $Y = |X|$.

Solução: Como X é uma variável aleatória, temos que $Y = g(X) = |X|$ também é uma variável aleatória. Assim, primeiramente iremos encontrar a função de distribuição. Assim, considerando $y > 0$,

$$\begin{aligned}
F_Y(y) &= \mathbb{P}(Y \leqslant y) = \mathbb{P}(g(X) \leqslant y) = \mathbb{P}(|X| \leqslant y) \\
&= \mathbb{P}(-y \leqslant X \leqslant y) \\
&= \mathbb{P}(X \leqslant y) - \mathbb{P}(X < -y) - \mathbb{P}(X = -y) + \mathbb{P}(X = -y) \\
&= \mathbb{P}(X \leqslant y) - \mathbb{P}(X \leqslant -y) + \mathbb{P}(X = -y) \\
&= F_X(y) - F_X(-y) + \mathbb{P}(X = -y). \qquad (3.28)
\end{aligned}$$

Se considerarmos X uma variável aleatória contínua, derivamos a equação (3.28), em relação a y e obtemos a função densidade. Ou seja,

$$\begin{aligned}
f_Y(y) &= \frac{d}{dy}F_Y(y) = \frac{d}{dy}F_X(y) - \frac{d}{dy}F_X(-y) \\
&= f_X(y) + f_X(-y) \qquad (3.29)
\end{aligned}$$

□

Exemplo 3.30. Seja X uma variável aleatória com função massa de probabilidade ou função densidade de probabilidade. Então $aX + b$ (com $a \neq 0$ e b, são constantes) e $|X|^\alpha$ ($\alpha > 0$) são variáveis aleatórias. Defina

$$X^+ = \begin{cases} X, & \text{se } X \geqslant 0 \\ 0, & \text{se } X < 0, \end{cases}$$

e

$$X^- = \begin{cases} X, & \text{se } X \leqslant 0 \\ 0, & \text{se } X < 0, >0 \end{cases}$$

Então X^+ e X^- são variáveis aleatórias. Encontre as funções de distribuição destas variáveis aleatórias.

Solução: Considere $Y = g(X) = aX + b$ (com $a \neq 0$ e b, são constantes), então

$$\begin{aligned}
F_Y(y) &= \mathbb{P}(Y \leqslant y) = \mathbb{P}(g(X) \leqslant y) = \mathbb{P}(aX + b \leqslant y) \\
&= \mathbb{P}(aX \leqslant y - b) = \begin{cases} \mathbb{P}\left(X \leqslant \frac{y-b}{a}\right), & \text{se } a > 0, \\ \mathbb{P}\left(X \geqslant \frac{y-b}{a}\right), & \text{se } a < 0. \end{cases}
\end{aligned}$$

Considere $Y = g(X) = |X|^\alpha$ ($\alpha > 0$), então

$$
\begin{aligned}
F_Y(y) &= \mathbb{P}(Y \leqslant y) = \mathbb{P}(g(X) \leqslant y) = \mathbb{P}(|X|^\alpha \leqslant y) = \mathbb{P}\left(|X| \leqslant y^{\frac{1}{\alpha}}\right) \\
&= \mathbb{P}\left(-y^{\frac{1}{\alpha}} \leqslant X \leqslant y^{\frac{1}{\alpha}}\right) \\
&= \mathbb{P}\left(X \leqslant y^{\frac{1}{\alpha}}\right) - \mathbb{P}\left(X < -y^{\frac{1}{\alpha}}\right) - \mathbb{P}\left(X = -y^{\frac{1}{\alpha}}\right) + \mathbb{P}\left(X = -y^{\frac{1}{\alpha}}\right) \\
&= \mathbb{P}\left(X \leqslant y^{\frac{1}{\alpha}}\right) - \mathbb{P}\left(X \leqslant -y^{\frac{1}{\alpha}}\right) + \mathbb{P}\left(X = -y^{\frac{1}{\alpha}}\right) \\
&= F_X\left(y^{\frac{1}{\alpha}}\right) - F_X\left(-y^{\frac{1}{\alpha}}\right) + \mathbb{P}\left(X = -y^{\frac{1}{\alpha}}\right).
\end{aligned}
$$

Considere $Y = g(X) = X^+$, então

$$
F_Y(y) = \mathbb{P}(Y \leqslant y) = \mathbb{P}(g(X) \leqslant y) = \mathbb{P}(X^+ \leqslant y) = \begin{cases} 0, & \text{se } y < 0, \\ \mathbb{P}(X \leqslant 0), & \text{se } y = 0, \\ \mathbb{P}(X < 0) + \mathbb{P}(0 \leqslant X \leqslant y), & \text{se } y > 0, \end{cases}
$$

Considere $Y = g(X) = X^-$, então

$$
F_Y(y) = \mathbb{P}(Y \leqslant y) = \mathbb{P}(g(X) \leqslant y) = \mathbb{P}(X^- \leqslant y) = \begin{cases} \mathbb{P}(X \leqslant y), & \text{se } y < 0, \\ 1, & \text{se } y \geqslant 0. \end{cases}
$$

□

Capítulo 4

Características de Variáveis Aleatórias

4.1 Esperança Matemática

Se repetirmos um experimento aleatório muitas vezes, o que acontece em média? Para responder essa pergunta não é necessário realizar um experimento, basta conhecer a distribuição de probabilidade da variável aleatória e calcular a sua Esperança Matemática.

Definição 4.1. Seja X é uma variável aleatória discreta com função massa de probabilidade $p_X(x)$. Então o valor esperado (ou Esperança Matemática ou Média de X), denotada por $\mathbb{E}(X)$, é definida por

$$\mathbb{E}(X) = \sum_{x \in \Omega_X} x p_X(x) = \sum_{x \in \Omega_X} x \mathbb{P}(X = x). \tag{4.1}$$

onde o somatório se estende por todos os valores possíveis de X e se a série $\sum_{x \in \Omega_X} x p_X(x)$ converge absolutamente, ou seja $\sum_{x \in \Omega_X} |x| p_X(x) < \infty$.

Observação 4.1. O conceito de Esperança na probabilidade é análogo ao conceito físico de *Centro de Gravidade* de uma distribuição de massas. Considere uma variável aleatória discreta X com função massa de probabilidade $p_X(x_i)$, com $i \geqslant 1$. Se agora imaginamos uma haste sem peso na qual os pesos com massa $p_X(x_i)$, com $i \geqslant 1$ esteja localizados nos pontos x_i, com $i \geqslant 1$ (ver Figura 4.1), então o ponto no qual a haste estaria em equilíbrio é conhecido como centro de gravidade. Por exemplo, seja G o centro de gravidade, então para encontrarmos o valor de G temos que resolver

$$\sum_i (x_i - G) p_X(x_i) = 0 \quad \rightarrow \quad \sum_i (x_i p_X(x_i) - G p_X(x_i)) = 0$$

$$\sum_i x_i p_X(x_i) - \sum_i G p_X(x_i) = 0 \quad \rightarrow \quad \sum_i x_i p_X(x_i) - G \sum_i p_X(x_i) = 0$$

$$\sum_i x_i p_X(x_i) - G = 0 \quad \rightarrow \quad G = \sum_i x_i p_X(x_i)$$

Pela Definição 4.2, temos que $G = \mathbb{E}(X)$.

Exemplo 4.1. [Ross, 2010] Uma Turma com 120 estudantes é levada em três ônibus para a apresentação de uma orquestra sinfônica. Há 36 estudantes em um dos ônibus, 40 no outro e 44 no terceiro ônibus. Quando os ônibus chegam, um dos 120 estudantes é escolhido aleatoriamente. Suponha X

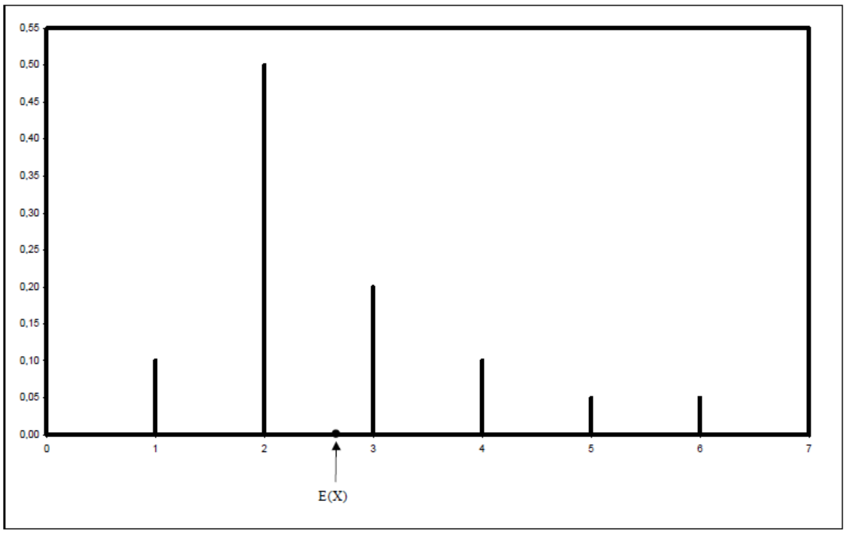


Figura 4.1: Interpretação da esperança Matemática como centro de gravidade da distribuição.

represente o número de estudantes que chegaram no mesmo ônibus do estudante escolhido e determine $\mathbb{E}(X)$.

Solução: Como o estudante escolhido aleatoriamente pode ser, com mesma probabilidade, qualquer um dos 120 estudantes, tem-se que

$$\mathbb{P}(X=36)=\frac{36}{120}, \quad \mathbb{P}(X=40)=\frac{40}{120}, \quad \mathbb{P}(X=44)=\frac{44}{120}.$$

Com isso,

$$\mathbb{E}(X)=36\frac{36}{120}+40\frac{40}{120}+44\frac{44}{120}=\frac{1208}{30}=40,2667.$$

Entretanto, o número médio de estudantes em um ônibus é $120/3=40$, o que mostra que o número esperado de estudantes no ônibus de onde foi escolhido aleatoriamente um estudante é maior do que o número médio de estudantes em um ônibus. Este é um fenômeno geral, que ocorre porque, quanto mais estudantes houver dentro de um ônibus, mais provável é que um estudante escolhido aleatoriamente esteja naquele ônibus. Como resultado, ônibus com muitos estudantes recebem um peso maior do que aqueles com menos estudantes.

□

Exemplo 4.2. Em determinado setor de uma loja de departamentos, o número de produtos vendidos em um dia pelos funcionários é uma variável aleatória X com a seguinte distribuição de probabilidades (esses números foram obtidos dos resultados de vários anos de estudo) - $p_X(x)$ probabilidade da venda:

x	0	1	2	3	4	5	6
$p_X(x)$	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 2 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. Qual é o número médio de produtos vendidos por cada vendedor e qual a comissão média de cada um deles?

Solução: O número médio de vendas por funcionário é:

$$\mathbb{E}(X)=0\times 0.1+1\times 0.4+2\times 0.2+3\times 0.1+4\times 0.1+5\times 0.05+6\times 0.05=2.05.$$

Com relação à comissão, vamos construir sua fmp:

x	0	1	2	3	4	5	6
C	0	10	20	70	120	170	220
$p_X(x)$	0.1	0.4	0.2	0.1	0.1	0.05	0.05

$$\mathbb{E}(C) = 0 \times 0.1 + 10 \times 0.4 + 20 \times 0.2 + 70 \times 0.1 + 120 \times 0.1 + 170 \times 0.05 + 220 \times 0.05 = 46.5.$$

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

□

Definição 4.2. Seja X é uma variável aleatória contínua com função densidade de probabilidade $f_X(x)$. Então o valor esperado (ou Esperança Matemática ou Média de X), denotada por $\mathbb{E}(X)$, é definida por

$$\mathbb{E}(X) = \int_{-\infty}^{\infty} x f_X(x) dx. \tag{4.2}$$

desde que a integral esteja bem definida.

Exemplo 4.3. [Magalhães, 2006] Seja X uma variável aleatória contínua com função densidade de probabilidade dada por

$$f_X(x) = \frac{1}{4} I_{(-2,0)}(x) + \frac{1}{4} I_{(2,4)}(x).$$

Calcule $\mathbb{E}(X)$.

Solução: Utilizando a definição de esperança para variáveis aleatórias contínuas, temos

$$\mathbb{E}(X) = \int_{-\infty}^{\infty} x f_X(x) dx = \int_{-2}^{0} x \frac{1}{4} dx + \int_{2}^{4} x \frac{1}{4} dx = 1.$$

Pelo gráfico de $f_X(x)$ (ver Figura 4.2), notamos que o valor esperado é o centro de gravidade da distribuição.

Por exemplo, considere que as massas de probabilidade correspondem aos intervalos $[-2, 0]$ e $[2, 4]$ são deslocadas, respectivamente, aos intervalos $[-3, -1]$ e $[3, 5]$. O centro de gravidade permanecerá inalterado e, portanto, também o valor esperado. Claro que existem modificações que alterariam o valor esperado.

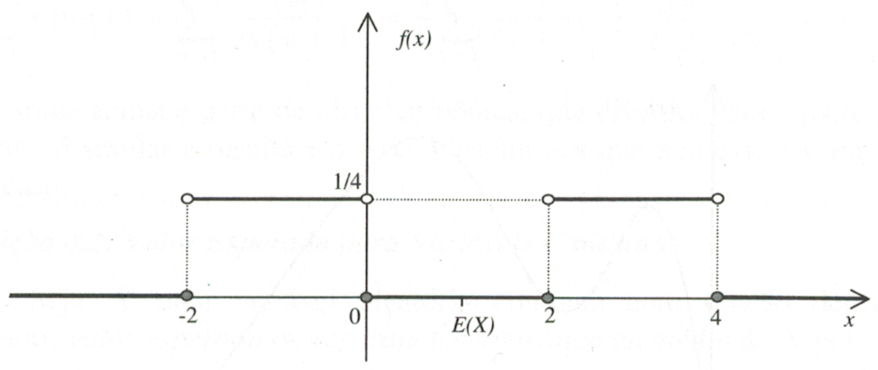


Figura 4.2: Esperança Matemática como centro de gravidade da distribuição. Fonte: Magalhães (2006).

□

Exemplo 4.4. [Magalhães, 2006] Seja X uma variável aleatória contínua com função densidade de probabilidade dada por

$$f_X(x) = \alpha e^{-\alpha x} I_{(0,\infty)}(x), \quad \alpha > 0.$$

Calcule $\mathbb{E}(X)$.

Solução: Utilizando a definição de esperança para variáveis aleatórias contínuas, temos

$$\mathbb{E}(X) = \int_{-\infty}^{\infty} x f_X(x)dx = \int_0^{\infty} x\alpha e^{-\alpha x}dx.$$

Aplicando integração por partes temos

$$\begin{aligned} u &= x & dv &= \alpha e^{-\alpha x} \\ du &= 1 & v &= -e^{-\alpha x} \end{aligned}$$

Logo,

$$\mathbb{E}(X) = -xe^{-\alpha x}\Big|_0^{\infty} - \int_0^{\infty} (-e^{-\alpha x})dx.$$

ou

$$\mathbb{E}(X) = \int_0^{\infty} (e^{-\alpha x})dx = -\frac{e^{-\alpha x}}{\alpha}\Big|_0^{\infty} = \frac{1}{\alpha}.$$

□

Exemplo 4.5. [Magalhães, 2006] Seja X uma variável aleatória contínua com função de distribuição de probabilidade dada por

$$F_X(x) = \begin{cases} 0, & x < 0; \\ \frac{x^2}{4}, & 0 \leqslant x < 1 \\ \frac{2x-1}{4}, & 1 \leqslant x < 2 \\ \frac{-(x^2-6x+5)}{4}, & 2 \leqslant x < 3 \\ 1, & x \geqslant 3. \end{cases}$$

Calcule $\mathbb{E}(X)$.

Solução: Primeiramente precisamos derivar a função de distribuição para obtermos a função densidade

$$f_X(x) = \frac{x}{2} I_{[0,1)}(x) + \frac{1}{2} I_{[1,2)}(x) - \frac{(x-3)}{2} I_{[2,3]}(x).$$

Utilizando a definição de esperança para variáveis aleatórias contínuas, temos

$$\mathbb{E}(X) = \int_{-\infty}^{\infty} x f_X(x)dx = \int_0^1 x\frac{x}{2}dx + \int_1^2 x\frac{1}{2}dx - \int_2^3 x\frac{(x-3)}{2}dx = \frac{3}{2}.$$

□

Exemplo 4.6. [Magalhães, 2006] Seja X uma variável aleatória contínua com distribuição Cauchy com parâmetros $\alpha = 0$ e $\beta = 1$, ou seja, sua função densidade de probabilidade dada por

$$f_X(x) = \frac{1}{\pi(1+x^2)}, -\infty < x < \infty.$$

Calcule $\mathbb{E}(X)$.

Solução: Utilizando a definição de esperança para variáveis aleatórias contínuas, e partindo a integral em duas partes

$$\mathbb{E}(X) = \int_{-\infty}^{\infty} x f_X(x)dx = \int_{-\infty}^{0} x f_X(x)dx + \int_{0}^{\infty} x f_X(x)dx$$

Temos,

$$\int_{0}^{\infty} x f_X(x)dx = \int_{0}^{\infty} x\frac{1}{\pi(1+x^2)}dx = \int_{0}^{\infty} x\frac{2}{2\pi(1+x^2)}dx = \frac{\ln\left(1+x^2\right)}{2\pi}\Bigg|_{0}^{\infty} = \infty.$$

De modo análogo, temos que

$$\int_{-\infty}^{0} x f_X(x)dx = \infty.$$

Desta forma, o valor esperado não existe. Observe que a densidade Cauchy, com parâmetros $\alpha = 0$ e $\beta = 1$, é simétrica ao redor de zero, o que poderia induzir a tomar o zero como valor esperado. Entretanto, a definição de esperança exige que pelo menos uma das partes, em que foi repartida $\mathbb{E}(X)$, seja finita.

□

Teorema 4.1. *Seja X uma variável aleatória com função de distribuição $F_X(\cdot)$ e cujo valor esperado existe. Então,*

$$\mathbb{E}(X) = \int_{0}^{\infty} [1 - F_X(x)]dx - \int_{-\infty}^{0} F_X(x)dx.$$

Demonstração. Ver Magalhães (2006), Teorema 4.5, página 214.

□

Exemplo 4.7. [Magalhães, 2006] Seja X uma variável aleatória com função de distribuição dada por

$$F_X(x) = \begin{cases} 0, & x < 1; \\ \frac{1}{18}(x^2 + x - 2), & 1 \leqslant x < 4 \\ 1, & x \geqslant 4. \end{cases}$$

Calcule $\mathbb{E}(X)$.

Solução: Pelo Teorema 4.1 temos que

$$\begin{aligned}
\mathbb{E}(X) &= \int_{0}^{\infty} [1 - F_X(x)]dx - \int_{-\infty}^{0} F_X(x)dx \\
&= \int_{0}^{\infty} [1 - F_X(x)]dx \\
&= \int_{0}^{1} (1-0)dx + \int_{1}^{4} \left(1 - \frac{1}{18}(x^2 + x - 2)\right) dx + \int_{4}^{\infty} (1-1)dx \\
&= \int_{0}^{1} 1dx - \int_{1}^{4} \frac{1}{18}(x^2 + x - 20)dx + \int_{4}^{\infty} 0dx \\
&= x\Big|_{0}^{1} - \frac{1}{18}\left(\frac{x^3}{3} + \frac{x^2}{2} - 20x\right)\Big|_{1}^{4} = \frac{11}{4}.
\end{aligned}$$

Note que, apesar de que $F_X(x)$ ser nula no intervalo $[0,1]$, a expressão alternativa para o cálculo do valor esperado recebe contribuição positiva, pois integramos $1-F_X(x)$, no referido intervalo.

□

Exemplo 4.8. [Magalhães, 2006] Seja X uma variável aleatória com função de distribuição dada por

$$F_X(x)=\begin{cases} 0, & x<-2;\\ \frac{1}{24}(x^2+4x+4), & -2\leqslant x<0\\ \frac{1}{6}(x+2), & 0\leqslant x<1\\ \frac{1}{6}(x+3), & 1\leqslant x<2\\ \frac{1}{24}(-x^2+8x+8), & 2\leqslant x<4\\ 1, & x\geqslant 4. \end{cases}$$

Calcule $\mathbb{E}(X)$.

Solução: Pelo Teorema 4.1 temos que

$$\begin{aligned}
\mathbb{E}(X) &= \int_0^{\infty}[1-F_X(x)]dx-\int_{-\infty}^{0}F_X(x)dx\\
&= \int_0^1\left(1-\frac{1}{6}(x+2)\right)dx+\int_1^2\left(1-\frac{1}{6}(x+3)\right)dx+\int_2^4\left(1-\frac{1}{24}(-x^2+8x+8)\right)dx\\
&\quad -\int_{-2}^{0}\frac{1}{24}(x^2+4x+4)dx\\
&= \int_0^1\frac{(4-x)}{6}dx+\int_1^2\frac{(3-x)}{6}dx+\int_2^4\frac{(16+x^2-8x)}{24}dx-\int_{-2}^{0}\frac{(x^2+4x+4)}{24}dx\\
&= \left[\frac{4x}{6}-\frac{x^2}{12}\Big|_0^1\right]+\left[\frac{3x}{6}-\frac{x^2}{12}\Big|_1^2\right]+\left[\frac{16x}{24}+\frac{x^3}{72}-\frac{8x^2}{78}\Big|_2^4\right]-\left[\frac{x^3}{72}+\frac{4x^2}{48}+\frac{4x}{24}\Big|_{-2}^{0}\right]\\
&= \frac{5}{6}.
\end{aligned}$$

□

4.2 Esperança de uma Função de Variável Aleatória

Suponha que conheçamos uma variável aleatória discreta e sua função massa de probabilidade e que queiramos calcular o valor esperado de uma função de X, digamos $Y=g(X)$. Como Y é uma variável aleatória discreta, ela tem uma função massa de probabilidade, que pode ser determinada a partir da função massa de probabilidade de X. Uma vez que tenhamos determinado a função massa de probabilidade de Y, podemos calcular $\mathbb{E}(Y)=\mathbb{E}(g(X))$ utilizando a definição de valor esperado.

Exemplo 4.9. Seja X uma variável aleatória que pode receber os valores $\Omega_X=\{-1,0,1\}$ com probabilidade

x	-1	0	1
$p_X(x)$	0,2	0,5	0,3

Calcule $\mathbb{E}(X^2)$.

Solução: Seja $Y=g(X)=X^2$. Então a função massa de probabilidade de Y é dada por

$$\mathbb{P}(Y=1)=\mathbb{P}(X=-1)+\mathbb{P}(X=1)=0,5$$

$$\mathbb{P}(Y=0)=\mathbb{P}(X=0)=0,5$$

Logo,

$$\mathbb{E}(X^2)=\mathbb{E}(Y)=1(0,5)+0(0,5).$$

Observe que

$$0,5=\mathbb{E}(X^2)\neq(\mathbb{E}(X))^2=0,01.$$

□

Embora o procedimento anterior sempre nos permita calcular o valor esperado de qualquer função de X a partir do conhecimento da função massa de probabilidade de X, há uma outra maneira de obter $\mathbb{E}(g(X))$: já que $g(X)$ será igual a $g(x)$ sempre que X for igual a x, parece razoável que $\mathbb{E}(g(X))$ deva ser uma média ponderada dos valores $g(x)$, com $g(x)$sendo ponderado pela probabilidade de que X seja igual a x. Isto é, o resultado a seguir é bastante intuitivo.

Proposição 4.1. *Seja X é uma variável aleatória discreta com função massa de probabilidade $p_X(x)$. Defina $Y=g(X)$ uma função da variável aleatória X. Então o valor esperado (ou Esperança Matemática ou Média de $Y=g(X)$), denotada por $\mathbb{E}(Y)=\mathbb{E}(g(X))$, é definida por*

$$\mathbb{E}(Y)=\mathbb{E}(g(X))=\sum_{x\in\Omega_X}g(x)p_X(x)=\sum_{x\in\Omega_X}g(x)\mathbb{P}(X=x), \tag{4.3}$$

onde o somatório se estende por todos os valores possíveis de X e se a série $\sum_{x\in\Omega_X}g(x)p_X(x)$ converge absolutamente, ou seja $\sum_{x\in\Omega_X}|g(x)|p_X(x)<\infty$.

A equação (4.3) é equivalente a:

$$\mathbb{E}(Y)=\mathbb{E}(g(X))=\sum_i g(x_i)p_X(x_i)=\sum_i g(x_i)\mathbb{P}(X=x_i). \tag{4.4}$$

Antes de demostrar a Proposição 4.1, vamos verificar se ela está de acordo com os resultados do Exemplo 4.9. Assim, aplicado os resultado temos

$$\mathbb{E}(X^2)=\sum_{x\in\Omega_X}g(x)p_X(x)=\sum_{x\in\Omega_X}x^2p_X(x)=(-1)^2(0,2)+0^2(0,5)+1^2(0,3)=0,5,$$

o que está de acordo com o resultado do Exemplo 4.9.

Demonstração. A demonstração segue, assim como na verificação anterior, com a agrupamento de todos os valores de $\sum_i g(x_i)p_X(x_i)$ com o mesmo valor de $g(x_i)$. Especificamente. suponha que y_j, para $j\leqslant 1$, represente os diferentes valores de $g(x_i)$, para $i\leqslant 1$. Então, o agrupamento de todos os $g(x_i)$ com todos os valores iguais resulta em

$$
\begin{aligned}
\sum_i g(x_i)p_X(x_i) &= \sum_j \sum_{i:g(x_i)=y_j} g(x_i)p_X(x_i) \\
&= \sum_j \sum_{i:g(x_i)=y_j} y_j p_X(x_i) \\
&= \sum_j y_j \sum_{i:g(x_i)=y_j} p_X(x_i) \\
&= \sum_j y_j \mathbb{P}(g(X)=y_j) \\
&= \sum_j y_j \mathbb{P}(Y=y_j) \\
&= \mathbb{E}(Y) = \mathbb{E}(g(X)).
\end{aligned}
$$

□

Exemplo 4.10. Seja X uma variável aleatória com função de distribuição dada por

$$
F_X(x) = \begin{cases} 0, & \text{se } x < -2; \\ 1/2, & \text{se } -2 \leqslant x < 0; \\ 5/8, & \text{se } 0 \leqslant x < 1; \\ 7/8, & \text{se } 1 \leqslant x < 2; \\ 1, & \text{se } x \geqslant 2. \end{cases}
$$

(i) Determine $\mathbb{E}(X)$;

(ii) Determine $\mathbb{E}(Y)$, onde $Y = 3X + 1$;

(iii) Determine $\mathbb{E}(Y)$, onde $Y = e^X$.

Solução: Primeiramente temos que encontrar a função massa de probabilidade da variável aleatória X. Assim,

x	-2	0	1	2
$p_X(x)$	1/2	1/8	2/8	1/8

Logo,

(i) $\mathbb{E}(X) = \sum_{x \in \Omega_X} x p_X(x) = (-2) \times \frac{1}{2} + 0 \times \frac{1}{8} + 1 \times \frac{2}{8} + 2 \times \frac{1}{8} = -\frac{1}{2}$.

(ii) $\mathbb{E}(Y) = \mathbb{E}(g(X)) = \sum_{x \in \Omega_X} g(x) p_X(x) = \sum_{x \in \Omega_X} (3x+1) p_X(x) = (3(-2)+1) \times \frac{1}{2} + (3(0)+1) \times \frac{1}{8} + (3(1)+1) \times \frac{2}{8} + (3(2)+1) \times \frac{1}{8} = -\frac{1}{2}$.

(iii) $\mathbb{E}(Y) = \mathbb{E}(g(X)) = \sum_{x \in \Omega_X} g(x) p_X(x) = \sum_{x \in \Omega_X} e^x p_X(x) = e^{-2} \times \frac{1}{2} + e^0 \times \frac{1}{8} + e^1 \times \frac{2}{8} + e^2 \times \frac{1}{8} = 1,79587$.

□

Teorema 4.2. *Seja X uma variável aleatória contínua cujo valor esperado existe. Considere $Y = g(X)$, uma função de X que também é uma variável aleatória no mesmo espaço de probabilidade. Então,*

$$
\mathbb{E}(Y) = \mathbb{E}(g(X)) = \int_{-\infty}^{\infty} g(x) f_X(x) dx.
$$

Demonstração. Ver Magalhães (2006), página 211, Exemplo 4.21.

□

Exemplo 4.11. Seja X uma variável aleatória com função densidade de probabilidade dada por

$$f_X(x) = \begin{cases} \frac{1}{4}(x+2), & -2 \leqslant x < 0 \\ \frac{1}{2}, & 0 \leqslant x < 1 \\ 0, & \text{caso contrário} \end{cases}$$

Calcule $\mathbb{E}(X^2)$.

Solução: Considere $Y = g(X) = X^2$. Então o valor esperado de $g(X)$ é calculado da seguinte forma:

$$\begin{aligned} \mathbb{E}(Y) &= \mathbb{E}(g(X)) = \int_{-\infty}^{\infty} g(x) f_X(x) dx \\ &= \int_{-2}^{0} x^2 \frac{1}{4}(x+2) dx + \int_{0}^{1} x^2 \frac{1}{2} dx \\ &= \left[\frac{1}{4} \left(\frac{x^4}{4} + \frac{2x^3}{3} \right) \Bigg|_{-2}^{0} \right] + \left[\frac{x^3}{6} \Bigg|_{0}^{1} \right] \\ &= \frac{1}{2}. \end{aligned}$$

□

Exemplo 4.12. [Ross, 2010] Seja X uma v.a.a contínua com função densidade de probabilidade dada por

$$f_X(x) = I_{[0,1]}(x) \tag{4.5}$$

Calcule $\mathbb{E}\left(e^X\right)$.

Solução:

$$\mathbb{E}\left(e^X\right) = \int_{-\infty}^{\infty} g(x) f_X(x) dx = \int_0^1 e^x dx = e^x \Bigg|_0^1 = e - 1.$$

□

Exemplo 4.13. [Magalhães, 2004] Seja X uma v.a.a contínua com função de distribuição dada por

$$F_X(x) = \begin{cases} 0, & \text{se } x < 0; \\ \frac{x^2}{4}, & \text{se } 0 \leqslant x < 1; \\ \frac{(2x-1)}{4}, & \text{se } 1 \leqslant x < 2; \\ \frac{-(x^2-6x+5)}{4}, & \text{se } 2 \leqslant x < 3; \\ 1, & \text{se } x \geqslant 3. \end{cases} \tag{4.6}$$

Calcule $\mathbb{E}(X)$ e $\mathbb{E}\left(X^2\right)$.

Solução: Primeiramente precisamos encontrar a função densidade de probabilidade. Para isso precisamos derivar $F_X(x)$ em relação a x. Assim,

$$f_X(x) = \begin{cases} \frac{x}{2}, & \text{se } 0 \leqslant x < 1; \\ \frac{1}{2}, & \text{se } 1 \leqslant x < 2; \\ \frac{-(x-3)}{2}, & \text{se } 2 \leqslant x < 3. \end{cases} \tag{4.7}$$

Analogamente,

$$f_X(x) = \frac{x}{2} I_{[0,1)}(x) + \frac{1}{2} I_{[1,2)}(x) - \frac{(x-3)}{2} I_{[2,3]}(x).$$

Logo,

$$\mathbb{E}(X) = \int_{-\infty}^{\infty} x f_X(x) dx = \int_0^1 x \frac{x}{2} dx + \int_1^2 x \frac{1}{2} dx - \int_2^3 x \frac{(x-3)}{2} dx = \frac{3}{2}.$$

Da mesma forma, pela Teorema 4.2, temos que

$$\mathbb{E}\left(X^2\right) = \int_{-\infty}^{\infty} x^2 f_X(x) dx = \int_0^1 x^2 \frac{x}{2} dx + \int_1^2 x^2 \frac{1}{2} dx - \int_2^3 x^2 \frac{(x-3)}{2} dx = \frac{8}{3}.$$

$\square$

Proposição 4.2. *Seja X uma variável aleatória com a e b constantes reais diferentes de zero.*

(i) $\mathbb{E}(a) = a$;

(ii) $\mathbb{E}(aX) = a\mathbb{E}(X)$;

(iii) $\mathbb{E}(aX + b) = a\mathbb{E}(X) + b$;

(iv) $\mathbb{E}(X + Y) = \mathbb{E}(X) + \mathbb{E}(Y)$;

(v) *Se $X \geqslant Y$ em Ω, então $\mathbb{E}(X) \geqslant \mathbb{E}(Y)$.*

Demonstração. Provaremos o caso em que X é uma variável aleatória discreta. O caso contínuo é idêntico. Seja X uma variável aleatória discreta com a e b constantes reais diferentes de zero. Então:

(i) Neste caso, $X = a$, ou seja, $\mathbb{P}(X = a) = 1$. Logo,

$$\mathbb{E}(X) = \sum_{x \in \Omega_X} x p_X(x) = a \times \mathbb{P}(X = a) = a \times 1 = a.$$

(ii) Seja $Y = g(X) = aX$. Então pela Proposição 4.1 temos que

$$X \text{ é v.a.d.} \rightarrow \mathbb{E}(Y) = \mathbb{E}(g(X)) = \sum_{x \in \Omega_X} g(x) p_X(x) = \sum_{x \in \Omega_X} (ax) p_X(x) = a \sum_{x \in \Omega_X} x p_X(x) = a\mathbb{E}(X).$$

$$X \text{ é v.a.c.} \rightarrow \mathbb{E}(Y) = \mathbb{E}(g(X)) = \int_{-\infty}^{\infty} g(x) f_X(x) dx = \int_{-\infty}^{\infty} (ax) f_X(x) dx = a \int_{-\infty}^{\infty} x f_X(x) dx = a\mathbb{E}(X).$$

(iii) Considere X uma variável aleatória discreta. Logo $Y = g(X) = aX + b$ também é uma variável aleatória discreta. Então pela Proposição 4.1 temos que

$$\mathbb{E}(Y) = \mathbb{E}(g(X)) = \sum_{x \in \Omega_X} g(x) p_X(x) = \sum_{x \in \Omega_X} (ax + b) p_X(x) = \sum_{x \in \Omega_X} ax p_X(x) + \sum_{x \in \Omega_X} b p_X(x)$$

$$\mathbb{E}(Y) = \mathbb{E}(g(X)) = a\sum_{x\in\Omega_X} xp_X(x) + b\sum_{x\in\Omega_X} p_X(x) = a\mathbb{E}(X) + b.$$

Considere agora X uma variável aleatória contínua. Logo $Y = g(X) = aX + b$ também é uma variável aleatória contínua. Então pelo Teorema 4.2 temos que

$$\mathbb{E}(Y) = \mathbb{E}(g(X)) = \int_{-\infty}^{\infty} g(x)f_X(x)dx = \int_{-\infty}^{\infty}(ax+b)f_X(x)dx = \int_{-\infty}^{\infty} axf_X(x)dx + \int_{-\infty}^{\infty} bf_X(x)dx$$

$$\mathbb{E}(Y) = \mathbb{E}(g(X)) = a\int_{-\infty}^{\infty} xf_X(x)dx + b\int_{-\infty}^{\infty} f_X(x)dx = a\mathbb{E}(X) + b.$$

(iv) Considere $(A_i)_{i=1}^k$ e $(B_j)_{j=1}^r$, partições de Ω. Partições de um conjunto (em Probabilidade) satisfazem as seguintes propriedades:

(i) $\cup_{i=1}^k A_j = \Omega$;

(ii) $A_i \cap A_\ell = \emptyset$, para $\ell, i = 1, \cdots k$, com $i \neq \ell$;

(iii) $\mathbb{P}(A_i) > 0$, para $i = 1, \cdots k$.

Idem para $(B_j)_{j=1}^r$. A sequência $(A_i)_{i=1}^k$ é tal que a variável aleatória X vale x_i em A_i, para $i = 1, \cdots k$, ou seja, temos

$$\mathbb{P}(X = x_i) = p_X(x_i) = \mathbb{P}(A_i), \quad i = 1, \cdots k,$$

sendo que os valores de X não precisam ser distintos entre si e podem ter qualquer sinal.

Da mesmo forma, a sequência $(B_j)_{j=1}^r$ é tal que a variável aleatória Y vale y_j em B_j, para $j = 1, \cdots r$, ou seja, temos

$$\mathbb{P}(Y = y_j) = p_Y(y_j) = \mathbb{P}(B_j), \quad j = 1, \cdots r,$$

sendo que os valores de Y não precisam ser distintos entre si e podem ter qualquer sinal.

Assim as variáveis aleatórias possuem a seguinte representação

$$X(w) = \sum_{i=1}^k x_i I_{A_i}(w) \quad Y(w) = \sum_{j=1}^r y_j I_{B_i}(w),$$

ou

$$X(w) = \sum_{i=1}^k \sum_{j=1}^r x_i I_{A_i\cap B_j}(w) \quad Y(w) = \sum_{j=1}^r \sum_{i=1}^k y_j I_{B_i\cap A_i}(w).$$

Para simplificar a notação, omitiremos w das expressões. Primeiramente observamos que $X+Y$ também é uma variável aleatória discreta com valores $x_i + y_j$ em $A_i \cap B_j$. Assim,

$$\begin{aligned}
\mathbb{E}(X+Y) &= \sum_{i=1}^k \sum_{j=1}^r (x_i + y_j)\mathbb{P}(A_i \cap B_j) \\
&= \sum_{i=1}^k \sum_{j=1}^r x_i \mathbb{P}(A_i \cap B_j) + \sum_{i=1}^k \sum_{j=1}^r y_j \mathbb{P}(A_i \cap B_j) \\
&= \mathbb{E}(X) + \mathbb{E}(Y).
\end{aligned}$$

(v) Observe que $x_i \geqslant y_j$ sempre que $A_i \cap B_j \neq \emptyset$, então,

$$\mathbb{E}(X) = \sum_{i=1}^{k}\sum_{j=1}^{r} x_i \mathbb{P}(A_i \cap B_j) \geqslant \sum_{i=1}^{k}\sum_{j=1}^{r} y_j \mathbb{P}(A_i \cap B_j) = \mathbb{E}(Y).$$

□

Proposição 4.3. *Sejam $X_1, X_2, \cdots, X_n$ variáveis aleatórias cujo valor esperado existe. Então, se a esperança da soma dessas variáveis existir, temos:*

$$\mathbb{E}(X_1 + X_2 + \cdots + X_n) = \mathbb{E}(X_1) + \mathbb{E}(X_2) + \cdots + \mathbb{E}(X_n).$$

Demonstração. Iremos demonstrar utilizando indução matemática. Pela Proposição 4.3, item (iv), temos que a proposição vale para $n = 2$. Vamos supor que o resultado vale para n variáveis aleatórias. Vamos provar para $n+1$. Definindo $Y = X_1 + X_2 + \cdots + X_n$, a soma das das $n+1$ variáveis aleatórias torna-se a soma de Y e X_{n+1}. Assim,

$$\mathbb{E}(X_1 + X_2 + \cdots + X_n + X_{n+1}) = \mathbb{E}(Y + X_{n+1}) = \mathbb{E}(Y) + \mathbb{E}(X_{n+1}),$$

uma vez que já provamos que o resultado vale para duas variáveis aleatórias.

Aplicando a hipótese de indução, substituímos $\mathbb{E}(Y)$ pela soma $\mathbb{E}(X_1) + \mathbb{E}(X_2) + \cdots + \mathbb{E}(X_n)$, Assim,

$$\mathbb{E}(X_1 + X_2 + \cdots + X_n + X_{n+1}) = \mathbb{E}(Y) + \mathbb{E}(X_{n+1}) = \mathbb{E}(X_1) + \mathbb{E}(X_2) + \cdots + \mathbb{E}(X_n) + \mathbb{E}(X_{n+1}).$$

Logo a proposição está demonstrada. Observe que não necessitamos da independência entre as variáveis aleatórias. □

Exemplo 4.14. [Magalhães, 2006] Sejam $X_1, X_2, \cdots, X_n$ variáveis aleatórias cujo valor esperado é p. Calcule o valor esperado de $X_1 + X_2 + \cdots + X_n$.

Solução: Pela Proposição 4.3, temos que

$$\mathbb{E}(X_1 + X_2 + \cdots + X_n) = \mathbb{E}(X_1) + \mathbb{E}(X_2) + \cdots + \mathbb{E}(X_n).$$

Portanto,

$$\mathbb{E}(X_1 + X_2 + \cdots + X_n) = p + p + \cdots + p = np.$$

□

Teorema 4.3. *Seja X uma variável aleatória não negativa e considere $a > 0$. Então,*

$$\mathbb{P}(X \geqslant a) \leqslant \frac{\mathbb{E}(X)}{a}.$$

A desigualdade acima é chamada de Desigualdade Básica de Chebyshev.

Demonstração. Como X é uma variável aleatória não negativa, a existência de $\mathbb{E}(X)$ está garantida. Considere o evento $A = \{X \geqslant a\}$ e uma variável aleatória discreta $Y = aI_A$. Ou seja, Yé tal que, para todo $w \in \Omega$, temos

$$Y(w) = \begin{cases} a, & X(w) \geqslant a; \\ 0, & 0 \leqslant X(w) < a. \end{cases}$$

Obsere que por definição temos $X \geqslant Y$. Aplicando as propriedades da esperança, obtemos

$$\mathbb{E}(X) \geqslant \mathbb{E}(Y) = a\mathbb{P}(A) = a\mathbb{P}(X \geqslant a).$$

Assumindo que o evento A tem probabilidade positiva, a desigualdade fica verificada, ou seja,

$$\mathbb{P}(X \geqslant a) \leqslant \frac{\mathbb{E}(X)}{a}.$$

□

Exemplo 4.15. [Magalhães, 2006] Em uma empresa de 100 funcionários, o número médio de usuários simultâneos de Internet, num período do dia, é de aproximadamente 30. Atualmente, existem 30 pontos disponíveis para conexão via cabo. Deseja-se avaliar a a necessidade de aumentar este número.

Em uma avaliação mais realista, os custos de expansão deveriam sem comparados com as possíveis vantagens para a empresa. Aqui, faremos um estudo preliminar, calculando algumas probabilidades.

Vamos admitir que 30 é o valor esperado do número de conexões simultâneas, mas nada sabemos sobre a distribuição de probabilidade dessa variável aleatória.

Dessa forma, uma alternativa conveniente é usar a desigualdade de Chebychev para construir a tabela abaixo com o valor máximo de probabilidades, em função do número mínimo de usuários conectados simultaneamente. Denotaremos o limite fornecido pela desigualdade de Chebyshev por $p(k)$.

k	$p(k)$
30	1
40	0,75
50	0,60
60	0,50
70	0,43
80	0,38
90	0,33

Apesar de podermos obter estimativas mais precisas por outros métodos, a desigualdade de Chebyshev fornece uma avaliação probabilística que, combinadas com outros fatores, dá subsídios à tomada de decisões.

Por exemplo, da tabela acima, podemos afirmar que a probabilidade de haver até 70 usuários simultâneos é superior a $0,57$. A despeito de outros fatores, se for aceitável um índice mínimo e 57% para a possibilidade de conexão, a empresa deveria ampliar suas linhas telefônicas para 70.

□

Observação 4.2. Algumas observações.

(i) [https://cutt.ly/WiPa2CD] Dizemos que uma função f definida em um intervalo I é convexa , se o segmento de reta secante que passa pelos pontos (p, f(p)) e (q, f(q)) sempre está acima ou coincide com o gráfico de f para qualquer escolha de pontos p e q em I.

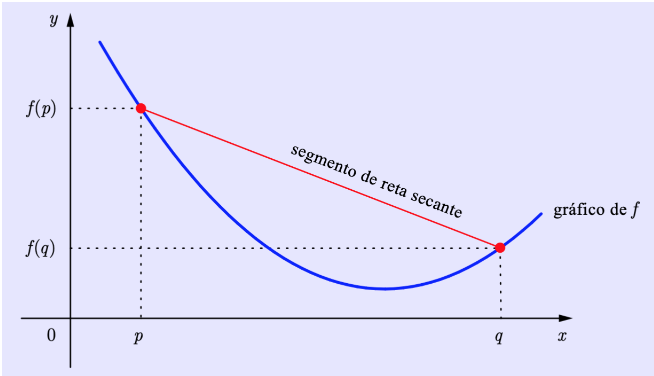


Figura 4.3: Exemplo de Função Convexa.
Fonte: https://cutt.ly/WiPa2CD

(ii) [https://cutt.ly/ci009G4] Seja $I \subseteq \mathbb{R}$ um intervalo (limitado ou não). Uma função $f : I \longrightarrow \mathbb{R}$ é dita convexa se tiver a seguinte propriedade: Dados dois pontos A e B no gráfico de f, a corda que une estes dois pontos está sempre acima do gráfico de f. Dados $x_1 < x < x_2$ em I, como na figura, chamando de $\mu = \dfrac{x - x_1}{x_2 - x_1}$, temos $0 \leq \mu \leq 1$

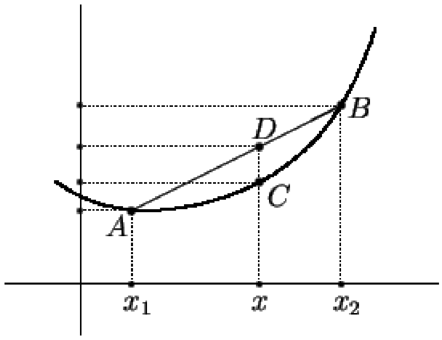


Figura 4.4: Exemplo de Função Convexa.
Fonte: https://cutt.ly/ci009G4

$$x = x_1 + (x - x_1) = x_1 + \mu\,(x_2 - x_1) = (1 - \mu)\,x_1 + \mu\,x_2 \ ,$$

ou ainda, chamando $\lambda = 1 - \mu$, $x = \lambda\,x_1 + \mu\,x_2$ com $\lambda \geq 0$, $\mu \geq 0$ e $\lambda + \mu = 1$.
Os pontos C e D da figura têm coordenadas

$$C = (\lambda\,x_1 + \mu\,x_2, f(\lambda\,x_1 + \mu\,x_2)) \text{ e } D = (\lambda\,x_1 + \mu\,x_2, \lambda\,f(x_1) + \mu\,f(x_2))$$

A função f é convexa quando o ponto D está sempre acima de C. Isto se expressa como

$$\forall x_1, x_2 \in I, \quad \forall \lambda, \mu \geq 0 \text{ com } \lambda + \mu = 1,$$

vale

$$f(\lambda\,x_1 + \mu\,x_2)) \leq \lambda\,f(x_1) + \mu\,f(x_2) \ .$$

(iii) Algumas funções Convexas.

(a) A função $f(x) = x^2$ é convexa.

(b) A função $f(x) = e^x$ é convexa.

(c) O valor absoluto é uma função convexa ($f(x) = |x|$)

(iv) Dizemos que uma função f definida em um intervalo I é côncava (ou côncava para baixo), se o segmento de reta secante que passa pelos pontos $(p, f(p))$ e $(q, f(q))$ sempre está abaixo ou coincide com o gráfico de f para qualquer escolha de pontos p e q em I.

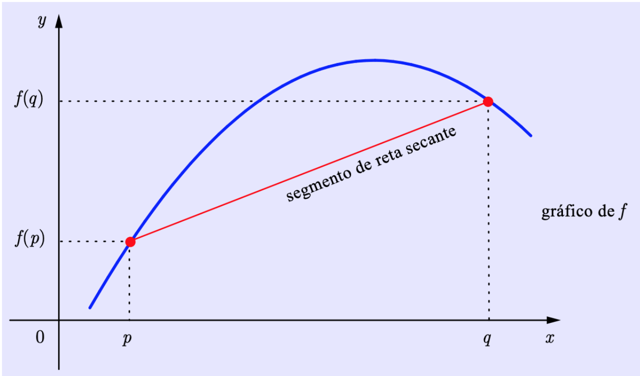


Figura 4.5: Exemplo de Função Côncava.
Fonte: https://cutt.ly/WiPa2CD

(v) Algumas funções Côncavas.

(a) A função $f(x) = \sqrt{x}$ é côncava.

(b) A função $f(x) = \frac{1}{x}$ é convexa em $\mathbb{R}^+$ e côncava em $\mathbb{R}^-$.

Proposição 4.4. [Magalhães, 2006] *Seja X uma variável aleatória com esperança finita e f uma função convexa. Então,*

$$\mathbb{E}(f(X)) \geqslant f(\mathbb{E}(X)).$$

Demonstração. Sendo f uma função convexa, existe uma reta que passa no ponto $(x_0, f(x_0))$ e fica abaixo do gráfico da função f.

Para o ponto $(x_0, f(x_0)) = (\mathbb{E}(X), f(\mathbb{E}(X)))$, seja $h(x) = ax + b$ a reta mencionada acima. Dessa forma, (veja Figura 4.6), no ponto $(\mathbb{E}(X), f(\mathbb{E}(X)))$, temos $h = f$ e, nos demais, $h(x) \leqslant f(x)$.

Pelo item (v) da Proposição 4.3 temos que

$$f(\mathbb{E}(X)) = h(\mathbb{E}(X)) = a\mathbb{E}(X) + bE(aX + b) = \mathbb{E}(h(X)) \leqslant \mathbb{E}(f(X)),$$

em que a suposição de esperança finita, permitiu a aplicação das funções f e g no valor esperado de X.

□

Observação 4.3. Se na Proposição 4.4 f for uma função côncava, então

$$\mathbb{E}(f(X)) \leqslant f(\mathbb{E}(X)).$$

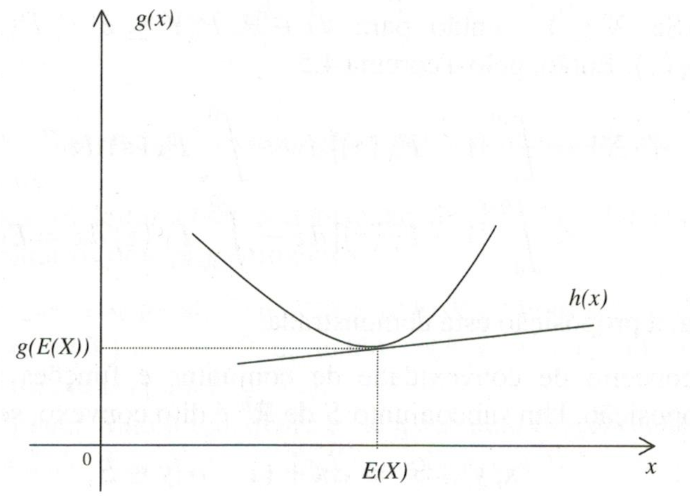


Figura 4.6: Função Convexa.Fonte: Magalhães (2006).

Teorema 4.4. *Seja I um intervalo contido no domínio de uma função f. Suponha que f, f' e f'' sejam contínuas em I.*

(i) Se $f''(x) > 0$ para todo $x \in I$, então f é uma função convexa no intervalo I.

(ii) Se $f''(x) < 0$ para todo $x \in I$, então f é uma função côncava no intervalo I.

4.2.1. Momentos e Variância de uma Variável Aleatória

Segundo Magalhães (2006), diversas características importantes das variáveis aleatórias podem ser qualificadas através do valor esperado de potências de variáveis aleatórias. Assim, nesta seção vamos estudar algumas delas. A seguir definimos os momentos de uma variável aleatória.

Definição 4.3. [Magalhães, 2006] Para $k = 1, 2, \cdots$, *o momento de ordem k* da variável aleatória X é definido por $\mathbb{E}(X^k)$, desde que essa quantidade exista (seja finita). Se $\mathbb{E}(X) = \mu < \infty$, definimos o *momento central de ordem k* por $\mathbb{E}[(X-\mu)^k]$, sempre que essa quantia exista. De modo similar, o *momento absoluto de ordem k* da variável aleatória X é definido por $\mathbb{E}(|X|^k)$.

Exemplo 4.16. [Ross, 2010] Sejam X variáveis aleatórias com distribuição $P(\lambda)$, cuja função massa de probabilidade é dada por

$$p_X(x) = \frac{e^{-\lambda}\lambda^x}{x!}, \quad x = 0, 1, 2, 3, \cdots, \lambda > 0.$$

Encontre $\mathbb{E}(X^k)$, para $k = 1, 2$.

Solução: Para calcularmos $\mathbb{E}(X)$ utilizamos a definição de valor esperado. Então,

$$\mathbb{E}(X) = \sum_{x=0}^{\infty} x \frac{e^{-\lambda}\lambda^x}{x!} = \sum_{x=0}^{\infty} x \frac{e^{-\lambda}\lambda^x}{x(x-1)!} = \lambda \sum_{x=1}^{\infty} \frac{e^{-\lambda}\lambda^{(x-1)}}{(x-1)!}.$$

Fazendo a mudança de variável $y = x - 1$, temos

$$\mathbb{E}(X) = \lambda \sum_{y=0}^{\infty} \frac{e^{-\lambda}\lambda^y}{y!} = \mathbb{E}(X) = \lambda e^{-\lambda} \sum_{y=0}^{\infty} \frac{\lambda^y}{y!} = \lambda e^{-\lambda} e^{\lambda} = \lambda,$$

pois $\sum_{y=0}^{\infty} \frac{\lambda^y}{y!} = e^{\lambda}$.

Pela Proposição 4.1, temos que

$$\mathbb{E}(X^2) \;=\; \sum_{x=0}^{\infty} x^2 \frac{e^{-\lambda}\lambda^x}{x!} = \sum_{x=0}^{\infty} x \frac{e^{-\lambda}\lambda^x}{(x-1)!} = \lambda \sum_{x=1}^{\infty} x \frac{e^{-\lambda}\lambda^{(x-1)}}{(x-1)!}$$

Fazendo a mudança de variável $y = x - 1$, temos

$$\begin{aligned}
\mathbb{E}(X^2) &= \lambda \sum_{x=1}^{\infty} x \frac{e^{-\lambda}\lambda^{(x-1)}}{(x-1)!} = \lambda \sum_{y=0}^{\infty} (y+1) \frac{e^{-\lambda}\lambda^y}{y!} \\
&= \lambda \sum_{y=0}^{\infty} \left[y \frac{e^{-\lambda}\lambda^y}{y!} + \frac{e^{-\lambda}\lambda^y}{y!} \right] \\
&= \lambda \left[\sum_{y=0}^{\infty} y \frac{e^{-\lambda}\lambda^y}{y!} + e^{-\lambda} \sum_{y=0}^{\infty} \frac{\lambda^y}{y!} \right] \\
&= \lambda \left[\underbrace{\sum_{y=0}^{\infty} y \frac{e^{-\lambda}\lambda^y}{y!}}_{\mathbb{E}(X)} + e^{-\lambda} \underbrace{\sum_{y=0}^{\infty} \frac{\lambda^y}{y!}}_{e^{\lambda}} \right] = \lambda(\lambda + 1).
\end{aligned}$$

□

Exemplo 4.17. [Magalhães, 2006] Sejam X variáveis aleatórias com distribuição $\Gamma(\alpha, \beta)$, cuja função densidade de probabilidade é dada por

$$f_X(x) = \frac{\beta^{\alpha}}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}, \quad x \leqslant 0, \alpha, \beta > 0.$$

Encontre $\mathbb{E}(X^k)$, para $k = 1, 2, \cdots$.

Solução: Pelo Teorema 4.2, temos que

$$\begin{aligned}
\mathbb{E}(X^k) &= \mathbb{E}(g(X)) = \int_{-\infty}^{\infty} g(x) f_X(x) dx = \int_{-\infty}^{\infty} x^k f_X(x) dx \\
&= \int_0^{\infty} x^k \frac{\beta^{\alpha}}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x} dx \\
&= \frac{\beta^{\alpha}}{\Gamma(\alpha)} \int_0^{\infty} x^{k+\alpha-1} e^{-\beta x} dx.
\end{aligned}$$

Fazendo $u = \beta x$ $(u > 0)$ temos que $x = \frac{u}{\beta}$, $\frac{du}{dx} = \beta$ e $dx = \frac{du}{\beta}$. Então,

$$
\begin{aligned}
\mathbb{E}(X^k) &= \frac{\beta^\alpha}{\Gamma(\alpha)}\int_0^\infty \left(\frac{u}{\beta}\right)^{k+\alpha-1} e^{-u}\frac{du}{\beta} \\
&= \frac{\beta^\alpha}{\Gamma(\alpha)\beta^{k+\alpha}}\int_0^\infty u^{k+\alpha-1}e^{-u}du.
\end{aligned}
$$

Como

$$\Gamma(\alpha) = \int_0^\infty u^{\alpha-1}e^{-u}du \quad \alpha > 0,$$

temos que

$$
\begin{aligned}
\mathbb{E}(X^k) &= \frac{\beta^\alpha}{\Gamma(\alpha)\beta^{k+\alpha}}\Gamma(k+\alpha) = \frac{\Gamma(k+\alpha)}{\Gamma(\alpha)\beta^k} \\
&= \frac{1}{\beta^k}\frac{(\alpha+k-1)(\alpha+k-2)\cdots(\alpha+1)\alpha\Gamma(\alpha)}{\Gamma(\alpha)} \\
&= \frac{(\alpha+k-1)(\alpha+k-2)\cdots(\alpha+1)\alpha}{\beta^k}
\end{aligned}
$$

pois $\Gamma(\alpha+1) = \alpha\Gamma(\alpha)$. Se $k = 1$, temos que

$$\mathbb{E}(X) = \frac{\Gamma(1+\alpha)}{\Gamma(\alpha)\beta} = \frac{\alpha\Gamma(\alpha)}{\Gamma(\alpha)\beta} = \frac{\alpha}{\beta}.$$

Se $k = 2$, temos que

$$\mathbb{E}(X^2) = \frac{\Gamma(2+\alpha)}{\Gamma(\alpha)\beta^2} = \frac{(\alpha+1)\alpha\Gamma(\alpha)}{\Gamma(\alpha)\beta^2} = \frac{\alpha(\alpha+1)}{\beta^2}.$$

Se $\alpha = 1$, X tem distribuição exponencial de parâmetro β e assim,

$$\mathbb{E}(X^k) = \frac{\Gamma(k+\alpha)}{\Gamma(\alpha)\beta^k} = \frac{\Gamma(k+1)}{\Gamma(1)\beta^k} = \frac{k!}{\beta^k}.$$

$\square$

A esperança de uma variável aleatória X é uma medida de posição. No entanto, é possível que duas variáveis bem diferentes tenham a mesma esperança, como é o caso das duas distribuições apresentadas na Figura 4.7.

Como já visto no caso da Estatística Descritiva, é necessário mensurar outros aspectos da distribuição, entre eles a dispersão dos dados. Esta será medida através da distância quadrática de cada valor à média da distribuição.

Variância pode ser interpretada como uma medida de variabilidade em torno da média da variável aleatória e é definida como

Definição 4.4. Seja X uma variável aleatória com $\mathbb{E}(X) < \infty$. Definimos a *variância* de X como sendo o momento central de ordem 2, isto é,

$$\mathrm{Var}(X) = \sigma^2 = \mathbb{E}[X - \mathbb{E}(X)]^2. \tag{4.8}$$

É comum denotar $\mathrm{Var}(X)$ por σ^2. A raiz quadrada da variância é definida de *desvio-padrão*

$$\sigma = \sqrt{\mathrm{Var}(X)}. \tag{4.9}$$

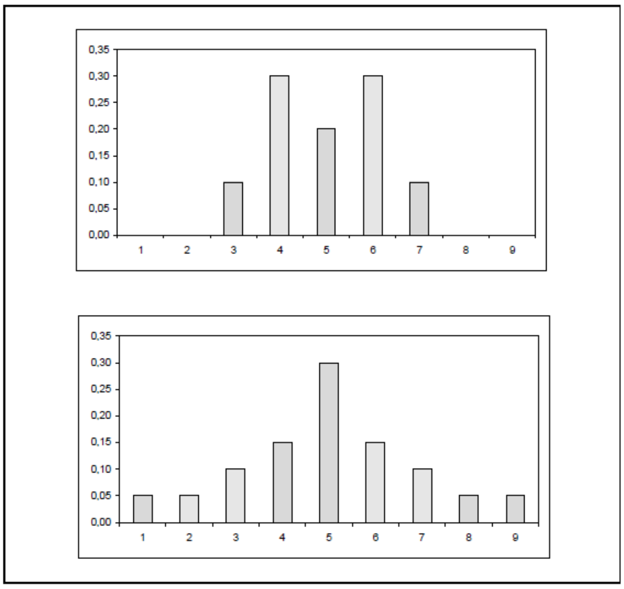


Figura 4.7: Distribuições com mesma esperança e diferentes dispersões.

O desvio-padrão tem a mesma unidade de medida da variável aleatória original e, assim, é útil para avaliar a amplitude dos possíveis valores da variável.

Proposição 4.5. [Magalhães, 2006] *Sejam a e b constantes quaisquer, então*

$$Var(aX+b) = a^2\, Var(X).$$

Demonstração. Denote $Y = aX+b$, então pelas propriedades de valor esperado, $\mathbb{E}(Y) = \mathbb{E}(aX+b) = a\mathbb{E}(X)+b$. Assim,

$$\begin{aligned} \mathrm{Var}(aX+b) &= \mathrm{Var}(Y) = \mathbb{E}[Y-\mathbb{E}(Y)]^2 = \mathbb{E}[aX+b-\mathbb{E}(aX+b)]^2 \\ &= \mathbb{E}[aX+b-a\mathbb{E}(X)-b]^2 = \mathbb{E}[aX-a\mathbb{E}(X)]^2 \\ &= a^2\mathbb{E}[X-\mathbb{E}(X)]^2 = a^2\mathrm{Var}(X). \end{aligned}$$

□

Pela Proposição 4.5, o acréscimo de uma constante não interfere na dispersão dos valores da variável. De fato, se todos os valores são igualmente deslocados, a média se desloca e a dispersão entre todos os valores fica inalterada.

A multiplicação por uma constante , no entanto, tem seu efeito. Teremos a variância original multiplicada pelo quadrado da constante. Se a constante é muito maior que 1, a variabilidade pode sofrer um acréscimo expressivo.

Proposição 4.6. [Magalhães, 2006] *Expressão alternativa para a variância*

$$Var(X) = \mathbb{E}[X-\mathbb{E}(X)]^2 = \mathbb{E}(X^2) - [\mathbb{E}(X)]^2$$

Demonstração. Seja X uma variável aleatória qualquer. Então

$$\mathrm{Var}(X) = \mathbb{E}[X-\mathbb{E}(X)]^2 = \mathbb{E}[X^2-2X\mathbb{E}(X)+\mathbb{E}^2(X)] = \mathbb{E}(X^2)-2\mathbb{E}(X)\mathbb{E}(X)+\mathbb{E}^2(X) = \mathbb{E}(X^2)-[\mathbb{E}(X)]^2.$$

□

Exemplo 4.18. Um lojista mantém extensos registros das vendas diárias de um certo aparelho. O quadro a seguir dá a distribuição de probabilidades do número de aparelhos vendidos em uma semana. Se é de R$500,00 o lucro por unidade vendida, qual o lucro esperado em uma semana? Qual é o desvio padrão do lucro?

x	0	1	2	3	4	5
$p_X(x)$	0.1	0.1	0.2	0.3	0.2	0.1

Solução: Seja X = número de aparelhos vendidos em uma semana e seja L o lucro semanal. Então, $L = 500X$.

$$\mathbb{E}(X) = \sum_{x\in\Omega_X} x\mathbb{P}(X = x) = 0\times 0.1 + 1\times 0.1 + 2\times 0.2 + 3\times 0.3 + 4\times 0.2 + 5\times 0.1 = 2.7 \text{ aparelhos}$$

$$\mathbb{E}(X^2) = \sum_{x\in\Omega_X} x^2\mathbb{P}(X = x) = 0^2\times 0.1+1^2\times 0.1+2^2\times 0.2+3^2\times 0.3+4^2\times 0.2+5^2\times 0.1 = 9.3 \text{ aparelhos}^2$$

$$\text{Var}(X) = \mathbb{E}(X^2) - [\mathbb{E}(X)]^2 = 9.3 - (2.7)^2 = 2.01 \text{ aparelhos}^2 \quad \text{DP}(X) = \sqrt{2.01} = 1.4177 \text{ aparelhos}$$

Com relação ao lucro semanal, temos que

$$\mathbb{E}(L) = 500\mathbb{E}(X) = R\$1350.00 \quad \text{Var}(L) = (500)^2\text{Var}(X) = 502500 \quad \text{DP}(X) = \sqrt{502500} = R\$708.8723$$

□

Exemplo 4.19. A variável aleatória discreta X assume apenas os valores $0, 1, 2, 3, 4$ e 5. A função massa de probabilidade de X é dada por

$$P(X = 0) = P(X = 1) = P(X = 2) = P(X = 3) = a$$

$$P(X = 4) = P(X = 5) = b$$

$$P(X = 2) = 3P(X = 4).$$

Temos que $\mathbb{E}[.]$ e $\text{Var}[.]$ denotam, respectivamente, esperança e variância. Verifique as seguintes afirmativas.

(i) Para que a função densidade de probabilidade seja válida, $a = 3/14$ e $b = 1/14$.

(ii) $\mathbb{E}[X] = 1,928571$.

(iii) $\mathbb{E}[X^2] = 5,928571$.

(iv) $\text{Var}[X] = 2,209184$.

(v) Defina $Z = 3 + 4X$. Então $\mathbb{E}(Z) = 10,71428$.

Solução:

(i) Para que a função densidade de probabilidade seja válida temos que ter $4a + 2b = 1$ e $a = 3b$, assim

$$12b + 2b = 1 \rightarrow 14b = 1 \rightarrow b = \frac{1}{14} \rightarrow a = 3b = \frac{3}{14}$$

(ii) $\mathbb{E}[X] = \sum_{x=0}^{5} x\mathbb{P}(X = x) = 1,928571$.

(iii) $\mathbb{E}[X^2] = \sum_{x=0}^{5} x^2\mathbb{P}(X = x) = 5,928571$.

(iv) $\text{Var}[X] = \mathbb{E}(X^2) - [\mathbb{E}(X)]^2 = 2,209184$.

(v) Como $Z = 3 + 4X$, então $\mathbb{E}(Z) = \mathbb{E}(3 + 4X) = 3 + 4\mathbb{E}(X) = 10,71428$.

□

Exemplo 4.20. [Magalhães, 2004] Seja X uma v.a.a contínua com função de distribuição dada por

$$F_X(x) = \begin{cases} 0, & \text{se } x < 0; \\ \frac{x^2}{4}, & \text{se } 0 \leqslant x < 1; \\ \frac{(2x-1)}{4}, & \text{se } 1 \leqslant x < 2; \\ \frac{-(x^2-6x+5)}{4}, & \text{se } 2 \leqslant x < 3; \\ 1, & \text{se } x \geqslant 3. \end{cases} \tag{4.10}$$

(a) Calcule $\text{Var}(X)$.

(b) Calcule $\text{Var}(Y)$, com $Y = 4X + 5$.

(c) Calcule $\text{Var}(Y)$, com $Y = aX^2 + b$. Depois substitua $a = -2$ e $b = 3$ e calcule a variância.

Solução: Primeiramente precisamos encontrar a função densidade de probabilidade. Para isso precisamos derivar $F_X(x)$ em relação a x. Assim,

$$f_X(x) = \begin{cases} \frac{x}{2}, & \text{se } 0 \leqslant x < 1; \\ \frac{1}{2}, & \text{se } 1 \leqslant x < 2; \\ \frac{-(x-3)}{2}, & \text{se } 2 \leqslant x < 3. \end{cases} \tag{4.11}$$

Analogamente,

$$f_X(x) = \frac{x}{2} I_{[0,1)}(x) + \frac{1}{2} I_{[1,2)}(x) - \frac{(x-3)}{2} I_{[2,3]}(x).$$

Logo,

$$\mathbb{E}(X) = \int_{-\infty}^{\infty} x f_X(x)dx = \int_0^1 x\frac{x}{2}dx + \int_1^2 x\frac{1}{2}dx - \int_2^3 x\frac{(x-3)}{2}dx = \frac{3}{2}.$$

Da mesma forma, pelo Teorema 4.2, temos que

$$\mathbb{E}\left(X^2\right) = \int_{-\infty}^{\infty} x^2 f_X(x)dx = \int_0^1 x^2\frac{x}{2}dx + \int_1^2 x^2\frac{1}{2}dx - \int_2^3 x^2\frac{(x-3)}{2}dx = \frac{8}{3}.$$

Portanto,

(a) $\text{Var}(X) = \mathbb{E}(X^2) - [\mathbb{E}(X)]^2 = \frac{8}{3} - \left[\frac{3}{2}\right]^2 = \frac{5}{12}$.

(b) $\text{Var}(Y) = \text{Var}(4X + 5) = 16\text{Var}(X) = 16 \times \frac{5}{12} = \frac{20}{3}$.

(c) $\text{Var}(Y) = \text{Var}(aX^2+b) = a^2\text{Var}(X^2) = a^2\left[\mathbb{E}(X^4) - \left(\mathbb{E}(X^2)\right)^2\right]$.

Sabemos que $\mathbb{E}\left(X^2\right) = \frac{8}{3}$. Precisamos calcular $\mathbb{E}(X^4)$. Assim, pelo Teorema 4.2, temos que

$$\mathbb{E}(X^4) = \int_{-\infty}^{\infty} x^4 f_X(x)dx = \int_0^1 x^4\frac{x}{2}dx + \int_1^2 x^4\frac{1}{2}dx - \int_2^3 x^4\frac{(x-3)}{2}dx = \frac{1}{12} + \frac{31}{10} - \left(-\frac{473}{60}\right) = \frac{664}{60}.$$

Logo, para $a = -2$ e $b = 3$, temos

$$\text{Var}(Y) = \text{Var}(aX^2+b) = a^2\text{Var}(X^2) = 4\left[\mathbb{E}(X^4) - \left(\mathbb{E}(X^2)\right)^2\right] = 4\left[\frac{664}{60} - \frac{64}{9}\right] = 4\frac{712}{180} = \frac{712}{45}$$

□

Exemplo 4.21. Voltando ao Exemplo 4.17 Sejam X variáveis aleatórias com distribuição $\Gamma(\alpha,\beta)$, cuja função densidade de probabilidade é dada por

$$f_X(x) = \frac{\beta^\alpha}{\Gamma(\alpha)}x^{\alpha-1}e^{-\beta x}, \quad x \leqslant 0, \alpha, \beta > 0.$$

Encontramos $\mathbb{E}(X) = \frac{\alpha}{\beta}$ e $\mathbb{E}(X^2) = \frac{\alpha(\alpha+1)}{\beta^2}$.
Portanto

$$\text{Var}(X) = \mathbb{E}(X^2) - [\mathbb{E}(X)]^2 = \frac{\alpha(\alpha+1)}{\beta^2} - \frac{\alpha^2}{\beta^2} = \frac{\alpha}{\beta^2}.$$

Se $\alpha = 1$, X tem distribuição exponencial de parâmetro β e assim,
$\mathbb{E}(X) = \frac{1}{\beta}$ e $\mathbb{E}(X^2) = \frac{2}{\beta^2}$.

$$\text{Var}(X) = \mathbb{E}(X^2) - [\mathbb{E}(X)]^2 = \frac{2}{\beta^2} - \frac{1}{\beta^2} = \frac{1}{\beta^2}.$$

□

Exemplo 4.22. Seja X uma variável aleatória discreta com distribuição $B(n,p)$, ou seja, sua função massa de probabilidade é dada por

$$\mathbb{P}(X = x) = \binom{n}{x}p^x(1-p)^{n-x}, \quad 0 \leqslant p \leqslant 1, x = 0, 1, 2, \cdots, n,$$

onde $\binom{n}{x} = \frac{n!}{x!(n-x)!}$ é o Binômio de Newton. Calcule $\text{Var}(X)$.

Solução: Iniciamos calculando $\mathbb{E}(X)$.

$$\begin{aligned}
\mathbb{E}(X) &= \sum_{x=0}^{n} x\binom{n}{x}p^x(1-p)^{n-x} = \sum_{x=0}^{n} x\frac{n!}{x!(n-x)!}p^x(1-p)^{n-x} \\
&= \sum_{x=0}^{n} n\frac{(n-1)!}{(x-1)!(n-x)!}p^{x-1}p(1-p)^{n-1-(x-1)}
\end{aligned}$$

Como $\frac{(n-1)!}{(x-1)!(n-x)!} = \binom{n-1}{x-1}$ e $\binom{k}{n} = 0$ se $k < 0$ ou $k > n$, temos que

$$\mathbb{E}(X) = np\sum_{x=1}^{n}\binom{n-1}{x-1}p^{x-1}(1-p)^{n-1-(x-1)}. \tag{4.12}$$

Fazendo $k = x-1$ na equação (5.3), temos

$$\mathbb{E}(X) = np\sum_{k=0}^{n-1}\binom{n-1}{k}p^{k}(1-p)^{n-1-(k)},$$

onde $\binom{n-1}{k}p^k(1-p)^{n-1-(k)}$ é a função massa de probabilidade de uma variável aleatória discreta com $k = 0, 1, 2, \cdots, n-1$. Portanto,

$$\sum_{k=0}^{n-1}\binom{n-1}{k}p^{k}(1-p)^{n-1-(k)} = 1$$

e

$$\mathbb{E}(X) = np.$$

Como próximo passo precisamos calcular $\mathbb{E}(X^2)$.

$$\begin{aligned}
\mathbb{E}(X^2) &= \sum_{x=0}^{n}x^2\binom{n}{x}p^x(1-p)^{n-x} = \sum_{x=0}^{n}(x^2-x+x)\binom{n}{x}p^x(1-p)^{n-x} \\
&= \sum_{x=0}^{n}(x(x-1)+x)\binom{n}{x}p^x(1-p)^{n-x} \\
&= \sum_{x=0}^{n}x(x-1)\binom{n}{x}p^x(1-p)^{n-x} + \underbrace{\sum_{x=0}^{n}+x\binom{n}{x}p^x(1-p)^{n-x}}_{\mathbb{E}(X)}.
\end{aligned}$$

Como

$$x(x-1)\binom{n}{x} = x(x-1)\frac{n!}{x!(n-x)!} = \frac{n!}{(x-2)!(n-x)!} = n(n-1)\frac{(n-2)!}{(x-2)!(n-x)!} = n(n-1)\binom{n-2}{x-2}$$

e $\binom{k}{n} = 0$ se $k < 0$ ou $k > n$, temos que

$$\begin{aligned}
\mathbb{E}(X^2) &= n(n-1)\sum_{x=2}^{n}\binom{n-2}{x-2}p^x(1-p)^{n-x} + \mathbb{E}(X) \\
&= n(n-1)p^2\sum_{x=2}^{n}\binom{n-2}{x-2}p^{x-2}(1-p)^{n-2-(x-2)} + \mathbb{E}(X). \qquad (4.13)
\end{aligned}$$

Fazendo $k = x-2$ na equação (5.4), temos

$$\mathbb{E}(X^2) = n(n-1)p^2\sum_{k=0}^{n-2}\binom{n-2}{k}p^{k}(1-p)^{n-2-(k)} + \mathbb{E}(X),$$

onde $\binom{n-2}{k}p^k(1-p)^{n-2-(k)}$ é a função massa de probabilidade de uma variável aleatória discreta com $k = 0, 1, 2, \cdots, n-2$. Portanto,

$$\sum_{k=0}^{n-2} \binom{n-2}{k} p^k(1-p)^{n-2-(k)} = 1$$

e

$$\mathbb{E}(X^2) = n(n-1)p^2 + \mathbb{E}(X) = n(n-1)p^2 + np = (np)^2 + np(1-p).$$

Agora podemos calcular a variância desta variável aleatória

$$\text{Var}(X) = \mathbb{E}(X^2) - [\mathbb{E}(X)]^2 = (np)^2 + np(1-p) - (np)^2 = np(1-p).$$

□

Os próximos resultados auxiliam no estudo do comportamento das variáveis. As desigualdades fornecem limites, em função da variância e de momentos, para a probabilidade da variável aleatória exceder um certo valor de interesse.

Proposição 4.7. *Seja X uma variável aleatória com média μ. Então, para qualquer $t > 0$, temos*

$$\mathbb{P}(|X-\mu| \geqslant t) \leqslant \frac{Var(X)}{t^2}.$$

A desigualdade acima é chamada de Desigualdade Clássica de Chebyshev.

Demonstração. Sejam $Y = |X - \mu|$ e $A = \{Y \geqslant t\}$. Assim, $Y^2 = (X-\mu)^2$ e também $Y^2 \geqslant Y^2 I_A$. Logo, pelo item (v), da Proposição 4.3,

$$\begin{aligned} \text{Var}(X) &= \mathbb{E}[(X-\mu)^2] = \mathbb{E}(Y^2) \geqslant \mathbb{E}(Y^2 I_A) \\ &= \int_t^{\infty} y^2 f_Y(y)dy \geqslant \int_t^{\infty} t^2 f_Y(y)dy = t^2\mathbb{P}(A) \\ &= t^2\mathbb{P}(|X-\mu| \geqslant t) \end{aligned}$$

e, portanto, segue o resultado.

□

Proposição 4.8. *Seja X uma variável aleatória qualquer. Então, para qualquer $t, k > 0$, temos*

$$\mathbb{P}(|X| \geqslant t) \leqslant \frac{\mathbb{E}(|X|^k)}{t^k}.$$

A desigualdade acima é chamada de Desigualdade de Markov.

Demonstração. Pelo Teorema 4.3 (Desigualdade Básica de Chebyshev), temos que

$$\mathbb{P}(|X| \geqslant t) = \mathbb{P}(|X|^k \geqslant t^k) \leqslant \frac{\mathbb{E}(|X|^k)}{t^k}.$$

□

Se a partir de uma amostra pudermos obter estimativas de momentos de uma variável (população), as desigualdades acima podem ajudar a estabelecer limites em probabilidades que tivermos interesse. Elas podem ser úteis em termos teóricos também. Por exemplo, suponha que desejamos avaliar $\mathbb{P}(|X - \mu| < 2\sigma)$, sendo μ e σ finitos, sendo a média e o desvio-padrão da variável aleatória X, respectivamente. Com a aplicação da Desigualdade Clássica de Chebyshev,

$$\mathbb{P}(|X - \mu| < 2\sigma) = 1 - \mathbb{P}(|X - \mu| \geqslant 2\sigma) \geqslant 1 - \frac{\text{Var}(X)}{4\sigma^2} = 1 - \frac{\sigma^2}{4\sigma^2} = 1 - \frac{1}{4} = \frac{3}{4}.$$

Ou seja, a porcentagem de valores, num raio de 2 desvios-padrões da média, é de pelo menos 75%. Note que nada foi mencionado sobre a distribuição da variável aleatória e, portanto, esse resultado vale em geral, contanto que a média e a variância sejam finitas.

Capítulo 5

Principais Modelos Discretos

A distribuição de probabilidade de uma variável aleatória descreve como as probabilidades de distribuem sobre os valores da variável aleatória.

No caso de uma variável aleatória discreta, a distribuição de probabilidade é definida por uma função de probabilidade $p_X(\cdot)$.

Neste capítulo são apresentadas algumas distribuições discretas. Mais distribuições discretas podem ser encontradas em https://bit.ly/3kfvEYK.

5.1 Modelo Bernoulli

Qualquer experimento aleatório com somente dois resultados possíveis "fracasso"e "sucesso". Seja p a probabilidade de sucesso e q a probabilidade de fracasso, com $p + q = 1$ [$p = \mathbb{P}(\text{"sucesso"})$, $q = \mathbb{P}(\text{"fracasso"})$].

Definição 5.1. Seja X o número de sucessos em uma única tentativa do experimento. A variável aleatória X segue o *modelo Bernoulli* se assume apenas dois valores 0 e 1.

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

A função massa de probabilidade da variável aleatória X é dada por

$$p_X(x) = \mathbb{P}(X = x) = p^x q^{1-x} = \begin{cases} q, & \text{se } x = 0; \\ p, & \text{se } x = 1. \end{cases} \tag{5.1}$$

Notação: $X \sim \text{Ber}(p)$, $0 \leqslant p \leqslant 1$. [A variável aleatória discreta X segue distribuição Bernoulli com parâmetro p.]

Corolário 5.1. *Seja $X \sim \text{Ber}(p)$, com $0 \leqslant p \leqslant 1$. Então $p_X(\cdot)$ dada pela equação (5.1) satisfaz as propriedades de uma função massa de probabilidade, ou seja, é uma função massa de probabilidade.*

Demonstração. Seja $X \sim \text{Ber}(p)$, com $0 \leqslant p \leqslant 1$. Temos que $p_X(\cdot)$, dada pela equação (5.1), é $0 \leqslant p_X(x) \leqslant 1$, para todo $x \in \{0, 1\}$ e

$$\sum_{x=0}^{1} p_X(x) = \mathbb{P}(X = 0) + \mathbb{P}(X = 1) = p + q = p + 1 - p = 1.$$

Então, $p_X(\cdot)$ satisfaz as propriedades de uma função massa de probabilidade.

$\square$

Resultado 5.1. Seja X uma variável aleatória discreta com distribuição $\text{Ber}(p)$, então sua função acumulada de probabilidade é dada por

$$F_X(x) = \mathbb{P}(X \leqslant x) = \begin{cases} 0, & \text{se } x < 0; \\ q = 1 - p, & \text{se } 0 \leqslant x < 1; \\ 1, & \text{se } x \geqslant 1. \end{cases}$$

$\square$

Exemplo 5.1. Considere o lançamento de uma moeda honesta sendo que, o sucesso é a ocorrência de cara, e o fracasso é a ocorrência de coroa. Defina X a variável aleatória

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

Encontre a função massa de probabilidade e função de distribuição da variável aleatória discreta X.

Solução: Como estamos trabalhando com uma moeda honesta temos que

$$p = \mathbb{P}(\text{sucesso}) = \mathbb{P}(\text{cara}) = \mathbb{P}(X = 1) = 0,5$$

$$q = 1 - p = \mathbb{P}(\text{fracasso}) = \mathbb{P}(\text{coroa}) = \mathbb{P}(X = 0) = 0,5$$

Então a função massa de probabilidade de X é dada por

x	0	1
$p_X(x) = \mathbb{P}(X = x)$	0,5	0,5

A sua função de distribuição é dada por

$$F_X(x) = \mathbb{P}(X \leqslant x) = \begin{cases} 0, & \text{se } x < 0; \\ 0,5, & \text{se } 0 \leqslant x < 1; \\ 1, & \text{se } x \geqslant 1. \end{cases}$$

$\square$

Exemplo 5.2. Uma urna contem 30 bolas brancas e 20 verdes. Retira-se uma bola dessa urna. Seja X o número de bolas verdes. Encontre a função de probabilidade e a distribuição da variável aleatória X.

Solução:

$$X = \begin{cases} 1, & \text{bola verde;} \\ 0, & \text{bola branca.} \end{cases}$$

$$p_X(x) = \mathbb{P}(X = x) = \left(\frac{2}{5}\right)^x \left(\frac{3}{5}\right)^{1-x}, \quad x \in \{0, 1\}.$$

Ou seja, $X \sim \text{Ber}(p)$, onde $p = \frac{2}{5}$.

$\square$

Proposição 5.1. *Seja $X \sim \text{Ber}(p)$, então*

(i) $\mathbb{E}(X) = p$.

(ii) $\mathbb{E}(X^2) = p$.

(iii) $\text{Var}(X) = p(1-p)$.

Demonstração. Seja $X \sim \text{Ber}(p)$, então

(i) Por definição de valor esperado temos que

$$\mathbb{E}(X) = \sum_{x=0}^{1} x p_X(x) = (1-p) \times 0 + p \times 1 = p.$$

(ii) Por definição, temos que $\text{Var}(X) = \mathbb{E}(X^2) - [\mathbb{E}(X)]^2$. Pelo item (i), temos que $\mathbb{E}(X) = p$. Temos que calcular $\mathbb{E}(X^2)$. Portanto,

$$\mathbb{E}(X^2) = \sum_{x=0}^{1} x^2 p_X(x) = (1-p) \times 0^2 + p \times 1^2 = p.$$

(iii) Assim,

$$\text{Var}(X) = \mathbb{E}(X^2) - [\mathbb{E}(X)]^2 = p - p^2 = p(1-p).$$

□

5.2 Modelo Binomial

Considere agora n repetições independentes de um mesmo experimento aleatório. Cada repetição admite apenas dois resultados: sucesso (com probabilidade p) e fracasso (com probabilidade $1-p$). As probabilidades de sucesso e fracasso são as mesmas em todas as repetições do experimento.

Esta descrição de repetições independentes de um experimento com as mesmas características citadas acima é, na verdade, a descrição de n repetições independentes de eventos de Bernoulli.

Sendo assim, considere o exemplo a seguir.

Exemplo 5.3. Suponha que uma moeda é lançada 10 vezes e vamos definir a variável aleatória $X =$ *número de caras*. Suponhamos que a probabilidade de cara seja p e, por conseguinte, a probabilidade de coroa é $1-p$. Os possíveis valores de X são 0, 1, 2, . . . , 10. Vamos agora calcular a probabilidade de cada um desses valores, estabelecendo a equivalência dos eventos envolvidos. Para isso vamos usar a notação K_i = cara no i-ésimo lançamento e C_i = coroa no i-ésimo lançamento.

$$\{X = 0\} = \{\text{coroa nos 10 lançamentos}\} = \{C_1 \cap \cdots \cap C_{10}\}$$

Podemos considerar os lançamentos da moeda como eventos independentes. Logo,

$$\mathbb{P}(X = 0) = \mathbb{P}(C_1) \times \cdots \times \mathbb{P}(C_{10}) = (1-p)^{10}$$

O evento $\{X = 1\}$ corresponde à ocorrência de 1 cara e 9 coroas. Uma sequência possível de resultados é $KCCCCCCCCC$ e a probabilidade é

$$\mathbb{P}(KCCCCCCCCC) = p(1-p)^9$$

Mas a sequência $CKCCCCCCCC$ também resulta em $\{X = 1\}$. Na verdade existem $\binom{10}{1}$ tais sequências, todas com a mesma probabilidade. Logo

$$\mathbb{P}(X = 1) = \binom{10}{1} p(1-p)^9.$$

Analogamente, o evento $\{X = 2\}$ corresponde à ocorrência de 2 caras e 8 coroas; uma sequência possível é $KKCCCCCCCC$, que tem probabilidade

$$\mathbb{P}(KKCCCCCCCC) = p^2(1-p)^8$$

Mas, existem $\binom{10}{2}$ maneiras de colocar caras numa sequência de 10 lançamentos e todas tem a mesma probabilidade. Portanto,

$$\mathbb{P}(X = 2) = \binom{10}{2} p^2(1-p)^8.$$

Em geral, para qualquer valor de x temos

$$\mathbb{P}(X = x) = \binom{10}{x} p^x(1-p)^{1-x}, \quad x = 0, 1, 2, \cdots, 10.$$

□

O experimento Binomial é um experimento de Bernoulli repetido n vezes, independentemente.

Consideremos n tentativas independentes de um mesmo experimento aleatório. Cada tentativa admite apenas dois resultados: fracasso com probabilidade q e sucesso com probabilidade p, onde $p + q = 1$. Ou seja, um experimento de Bernoulli é repetido n vezes, independentemente. As probabilidades de sucesso e fracasso são as mesmas para cada tentativa.

Definição 5.2. Seja X a variável aleatória número de sucessos nas n repetições independentes. Diremos que X segue o *modelo Binomial* com parâmetros n e p e sua função de probabilidade é dada por

$$p_X(x) = \mathbb{P}(X = x) = \binom{n}{x} p^x(1-p)^{n-x}, \quad x \in \{0, 1, \cdots, n\}, \tag{5.2}$$

onde $\binom{n}{x} = \frac{n!}{x!(n-x)!}$ é o Binômio de Newton.

Notação: $X \sim \mathrm{B}(n,p)$, $0 < p \leqslant 1$. [A variável aleatória discreta X segue distribuição Binomial com parâmetros (n,p), com p a probabilidade de sucesso nos eventos de Bernoulli.]

Corolário 5.2. *Seja* $X \sim \mathrm{B}(n,p)$, $0 < p \leqslant 1$. *Então* $p_X(\cdot)$ *dada pela equação* (5.2) *satisfaz as propriedades de uma função massa de probabilidade, ou seja, é uma função massa de probabilidade.*

Demonstração. Vamos verificar se $p_X(\cdot)$ é função de probabilidade. Temos que $0 \leqslant p_X(x) \leqslant 1$, para todo $x \in \{0, 1, \cdots, n\}$ e

$$\sum_{x=0}^{n} p_X(x) = \sum_{x=0}^{n} \binom{n}{x} p^x (1-p)^{n-x} = (p + (1-p))^n = 1,$$

pois

$$\sum_{j=0}^{n} \binom{n}{j} a^j (b)^{n-j} = (a+b)^n.$$

Logo $p_X(\cdot)$ é função massa de probabilidade.

$\square$

Resultado 5.2. Seja X uma variável aleatória discreta com distribuição $\mathrm{B}(n,p)$, então sua função acumulada de probabilidade é dada por

$$F_X(x) = \mathbb{P}(X \leqslant x) = \sum_{j=0}^{x} \mathbb{P}(X = j) = \begin{cases} 0, & \text{se } x < 0; \\ \sum_{j=0}^{[x]} \binom{n}{j} p^j (1-p)^{n-j}, & \text{se } 0 \leqslant x < n; \\ 1, & x \geqslant n. \end{cases}$$

onde $[x]$ é a parte inteira de x.

$\square$

Nas Figuras 5.1 e 5.2 apresentamos a função massa de probabilidade e acumulada para diferentes valores de p.

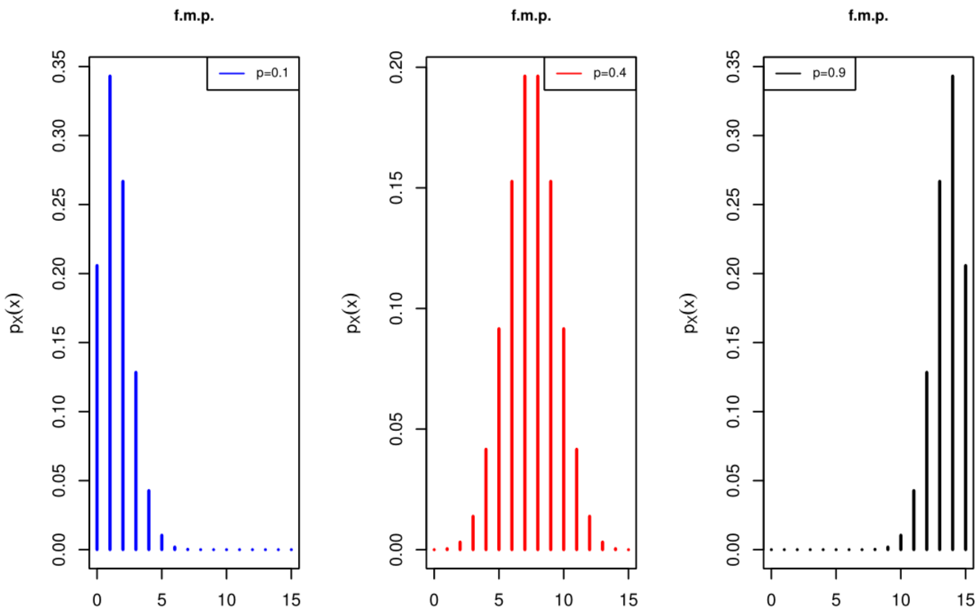


Figura 5.1: Função Massa de Probabilidade da variável aleatória $X \sim \mathrm{B}(15, p)$, com $p \in \{0,1; 0,4; 0,9\}$.

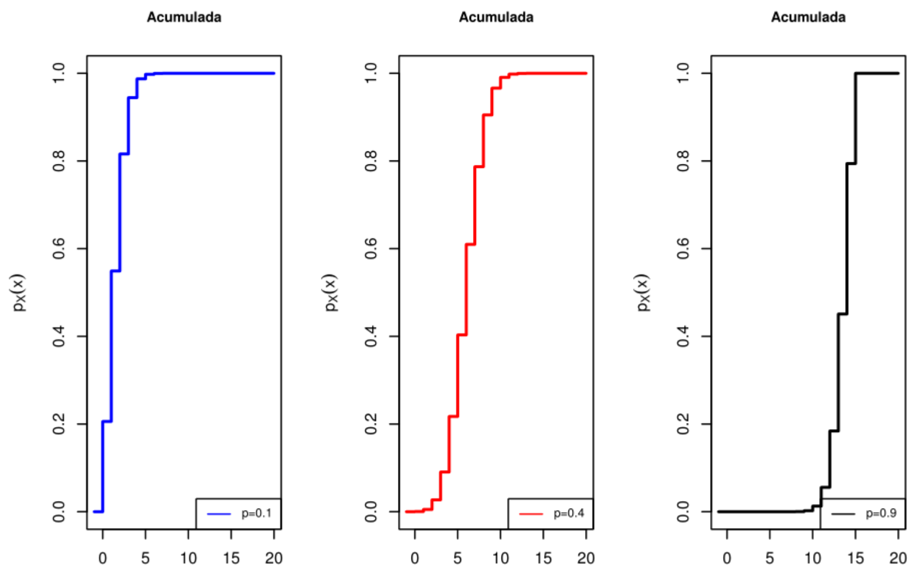


Figura 5.2: Função de Distribuição (ou acumulada) da variável aleatória $X \sim \mathrm{B}(15, p)$, com $p \in \{0,1; 0,4; 0,9\}$.

Proposição 5.2. *Seja* $X \sim \mathrm{B}(n,p)$, *então*

(i) $\mathbb{E}(X) = np$.

(ii) $\mathbb{E}(X^2) = (np)^2 + np(1-p)$.

(iii) $\mathrm{Var}(X) = npq$. q=1-p

Demonstração. Seja X uma variável aleatória discreta com distribuição $\mathrm{B}(n,p)$, ou seja, sua função massa de probabilidade é dada por

$$\mathbb{P}(X = x) = \binom{n}{x} p^x (1-p)^{n-x}, \quad 0 \leqslant p \leqslant 1, x = 0, 1, 2, \cdots, n,$$

onde $\binom{n}{x} = \frac{n!}{x!(n-x)!}$ é o Binômio de Newton.

(i) Iniciamos calculando $\mathbb{E}(X)$.

$$\begin{aligned}
\mathbb{E}(X) &= \sum_{x=0}^{n} x \binom{n}{x} p^x (1-p)^{n-x} = \sum_{x=0}^{n} x \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x} \\
&= \sum_{x=0}^{n} n \frac{(n-1)!}{(x-1)!(n-x)!} p^{x-1} p (1-p)^{n-1-(x-1)}
\end{aligned}$$

Como $\frac{(n-1)!}{(x-1)!(n-x)!} = \binom{n-1}{x-1}$ e $\binom{k}{n} = 0$ se $k < 0$ ou $k > n$, temos que

$$\mathbb{E}(X) = np \sum_{x=1}^{n} \binom{n-1}{x-1} p^{x-1} (1-p)^{n-1-(x-1)}. \tag{5.3}$$

Fazendo $k = x - 1$ na equação (5.3), temos

$$\mathbb{E}(X) = np\sum_{k=0}^{n-1}\binom{n-1}{k}p^k(1-p)^{n-1-(k)},$$

onde $\binom{n-1}{k}p^k(1-p)^{n-1-(k)}$ é a função massa de probabilidade de uma variável aleatória discreta com $k = 0, 1, 2, \cdots, n-1$. Portanto,

$$\sum_{k=0}^{n-1}\binom{n-1}{k}p^k(1-p)^{n-1-(k)} = 1$$

e

$$\mathbb{E}(X) = np.$$

(ii) Por definição, temos que $\mathrm{Var}(X) = \mathbb{E}(X^2) - [\mathbb{E}(X)]^2$. Pelo item (i), temos que $\mathbb{E}(X) = p$. Temos que calcular $\mathbb{E}(X^2)$. Portanto,

$$\begin{aligned}
\mathbb{E}(X^2) &= \sum_{x=0}^{n}x^2\binom{n}{x}p^x(1-p)^{n-x} = \sum_{x=0}^{n}(x^2 - x + x)\binom{n}{x}p^x(1-p)^{n-x}\\
&= \sum_{x=0}^{n}(x(x-1)+x)\binom{n}{x}p^x(1-p)^{n-x}\\
&= \sum_{x=0}^{n}x(x-1)\binom{n}{x}p^x(1-p)^{n-x} + \underbrace{\sum_{x=0}^{n}x\binom{n}{x}p^x(1-p)^{n-x}}_{\mathbb{E}(X)}.
\end{aligned}$$

Como

$$x(x-1)\binom{n}{x} = x(x-1)\frac{n!}{x!(n-x)!} = \frac{n!}{(x-2)!(n-x)!} = n(n-1)\frac{(n-2)!}{(x-2)!(n-x)!} = n(n-1)\binom{n-2}{x-2}$$

e $\binom{k}{n} = 0$ se $k < 0$ ou $k > n$, temos que

$$\begin{aligned}
\mathbb{E}(X^2) &= n(n-1)\sum_{x=2}^{n}\binom{n-2}{x-2}p^x(1-p)^{n-x} + \mathbb{E}(X)\\
&= n(n-1)p^2\sum_{x=2}^{n}\binom{n-2}{x-2}p^{x-2}(1-p)^{n-2-(x-2)} + \mathbb{E}(X). \qquad (5.4)
\end{aligned}$$

Fazendo $k = x - 2$ na equação (5.4), temos

$$\mathbb{E}(X^2) = n(n-1)p^2\sum_{k=0}^{n-2}\binom{n-2}{k}p^k(1-p)^{n-2-(k)} + \mathbb{E}(X),$$

onde $\binom{n-2}{k}p^k(1-p)^{n-2-(k)}$ é a função massa de probabilidade de uma variável aleatória discreta com $k = 0, 1, 2, \cdots, n-2$. Portanto,

$$\sum_{k=0}^{n-2} \binom{n-2}{k} p^k (1-p)^{n-2-(k)} = 1$$

e

$$\mathbb{E}(X^2) = n(n-1)p^2 + \mathbb{E}(X) = n(n-1)p^2 + np = (np)^2 + np(1-p).$$

(iii) Agora podemos calcular a variância desta variável aleatória

$$\text{Var}(X) = \mathbb{E}(X^2) - [\mathbb{E}(X)]^2 = (np)^2 + np(1-p) - (np)^2 = np(1-p).$$

□

Exemplo 5.4. Uma moeda honesta é lançada 20 vezes. Qual a probabilidade de saírem 8 caras?

Solução: Temos que X é o número de sucessos (caras).

$$p = \mathbb{P}(\text{sucesso}) = \frac{1}{2}.$$

Logo, $X \sim \text{B}\left(20, \frac{1}{2}\right)$

$$\mathbb{P}(X = x) \quad = \quad \binom{20}{x}\left(\frac{1}{2}\right)^x\left(\frac{1}{2}\right)^{20-x}, \quad x \in \{0, 1, 2, 3, \cdots, 20\}.$$

Se $x = 8$, temos

$$\begin{aligned}\mathbb{P}(X = 8) &= \binom{20}{8}\left(\frac{1}{2}\right)^8\left(\frac{1}{2}\right)^{20-8} \\ &= \text{dbinom(x=8,size=20,prob=0.5)} \\ &= 0.1201344\end{aligned}$$

□

Exemplo 5.5. Uma prova tipo teste tem 50 questões independentes. Cada questão tem 5 alternativas. Apenas uma delas é a correta. Se um aluno resolve a prova respondendo a esmo a questão, qual a probabilidade de tirar nota 5?

Solução: A variável aleatória X é o número de acertos, $x \in \{0, 1, \cdots, n\}$. A probabilidade de acerto $p = \mathbb{P}(\text{acerto}) = \frac{1}{5}$. Logo, $X \sim \text{B}(50, \frac{1}{5})$.

Portanto a função de probabilidade é dada por

$$\mathbb{P}(X = x) = \binom{50}{x}\left(\frac{1}{5}\right)^x\left(\frac{4}{5}\right)^{50-x}.$$

Logo,

$$\begin{aligned}
\mathbb{P}(X = 25) &= \binom{50}{25}\left(\frac{1}{5}\right)^{25}\left(\frac{4}{5}\right)^{50-25} \\
&= \text{dbinom(x=25,size=50,prob=1/5)} \\
&= 0,000002
\end{aligned}$$

Probabilidade de tirar uma nota menor ou igual a 5:

$$\begin{aligned}
\mathbb{P}(X \leqslant 25) &= \sum_{x=0}^{25}\binom{50}{x}\left(\frac{1}{5}\right)^{x}\left(\frac{4}{5}\right)^{50-x} \\
&= \text{pbinom(q=25,size=50,prob=1/5)} \\
&= 0,9999995
\end{aligned}$$

□

Exemplo 5.6. Um atirador acerta no alvo, 20% dos tiros. Se ele dá 10 tiros, qual a probabilidade de ele acertar na mosca no máximo 1 vez?

Solução: Podemos pensar os tiros como experimentos de Bernoulli independentes, onde a probabilidade de sucesso é $p = 0,20$. Então, o problema pede $\mathbb{P}(X \leqslant 1)$, onde $X = \{$número de acertos em 10 tiros$\}$. Logo, $X \sim \mathrm{B}(k, p)$, com $k = 10$ e $p = 0,20$.

Então,

$$\begin{aligned}
\mathbb{P}(X \leqslant 1) &= \mathbb{P}(X = 0) + \mathbb{P}(X = 1) = \binom{10}{0}0,2^{0}0,8^{10-0} + \binom{10}{1}0,2^{1}0,8^{10-1} \\
&= \text{pbinom(q=1,size=10,prob=0.2)} \\
&= 0.3758096.
\end{aligned}$$

□

Exemplo 5.7. Dois adversários A e B disputam uma série de 8 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). Assumindo a independência das provas, se definimos $X = \{$número de vitórias de A$\}$, então $X \sim \mathrm{B}(8; 0,6)$ e o problema é calcular $\mathbb{P}(X \geqslant 5)$, isto é A ganha mais partidas que B.

$$\begin{aligned}
\mathbb{P}(X \geqslant 5) &= \sum_{x=5}^{8}\mathbb{P}(X = x) \\
&= \binom{8}{5}0,6^{5}0,4^{8-5} + \binom{8}{6}0,6^{6}0,4^{8-6} + \binom{8}{7}0,6^{7}0,4^{8-7} + \binom{8}{8}0,6^{8}0,4^{8-8} \\
&= \text{sum(dbinom(x=5:8,size=8,prob=0.6))} \\
&= \text{pbinom(q=4,size=8,prob=0.6,lower.tail=FALSE)} \\
&= 0,5940864.
\end{aligned}$$

□

5.3 Modelo Poisson

Na distribuição Binomial, a variável de interesse era o número de sucessos em um intervalo discreto (n repetições de um experimento 0-1). Muitas vezes, entretanto, o interesse reside no número de sucessos em um intervalo contínuo, que pode ser o tempo, comprimento, etc.

A probabilidade de ocorrência de um sucesso no intervalo é proporcional ao intervalo. A probabilidade de mais de um sucesso neste intervalo é bastante pequena com relação à probabilidade de um sucesso.

Definição 5.3. Seja X o número de sucessos em um intervalo. A variável aleatória X segue o modelo Poisson de parâmetro λ, $\lambda > 0$, se a sua função massa de probabilidade for dada por

$$p_X(x) = \mathbb{P}(X = x) = \frac{e^{-\lambda}\lambda^x}{x!}, \quad x = 0, 1, 2, \cdots. \tag{5.5}$$

Notação: $X \sim \mathrm{P}(\lambda)$, $\lambda > 0$. [A variável aleatória discreta X segue distribuição Poisson com parâmetros λ, com λ é a média de ocorrências em um intervalo em que a variável aleatória discreta X é contada.]

O parâmetro λ indica o número esperado de sucessos no intervalo (a taxa de ocorrência para uma unidade de medida).

A distribuição de Poisson é largamente utilizada quando se deseja contar o número de sucessos que ocorrem em intervalos de tempo, ou superfície ou volume. Por exemplo

(i) carros que passam por um cruzamento por minuto, durante uma certa hora do dia;

(ii) erro tipográficos por página, em um material impresso;

(iii) defeitos por unidade (m^2, m^3, m etc.) por peça fabricada;

(iv) colônia de bactérias numa dada cultura por 0.01 mm^2, numa plaqueta de microscópio;

(v) mortes por ataque do coração por ano, numa cidade;

(vi) em problemas de filas em geral;

(vii) número de chamadas recebidas por um telefone durante cinco minutos;

(viii) número de falhas de um computador num dia de operação;

(ix) número de relatórios de acidentes enviados a uma companhia de seguros numa semana.

Corolário 5.3. *Seja $X \sim P(\lambda)$, $\lambda > 0$. Então $p_X(\cdot)$ dada pela equação (5.5) satisfaz as propriedades de uma função massa de probabilidade, ou seja, é uma função massa de probabilidade.*

Demonstração. Vamos verificar se $p_X(\cdot)$ é função de probabilidade. Temos que $0 \leqslant p_X(x) \leqslant 1$, para todo $x = 0, 1, \cdots$ e

$$\sum_{x \geqslant 0} p_X(x) = \sum_{x \geqslant 0} \frac{e^{-\lambda}\lambda^x}{x!} = e^{-\lambda} \sum_{x \geqslant 0} \frac{\lambda^x}{x!} = e^{-\lambda} e^{\lambda} = 1,$$

pois $\sum_{k \geqslant 0} \frac{\lambda^k}{k!} = e^{\lambda}$. Logo, $p_X(\cdot)$ é função de probabilidade.

$\square$

Observação 5.1. Seja X a variável aleatória definida como o número de eventos que ocorrem sobre um período de tempo t. Substituímos λ na função massa de probabilidade por $t\lambda$. Dessa forma,

$$p_X(x) = \mathbb{P}(X = x) = \frac{e^{-t\lambda}(t\lambda)^x}{x!}, \quad x = 0, 1, 2, \cdots .$$

Resultado 5.3. Seja X uma variável aleatória discreta com distribuição $\mathrm{P}(\lambda)$, então sua função acumulada de probabilidade é dada por

$$F_X(x) = \mathbb{P}(X \leqslant x) = \sum_{j=0}^{x} \mathbb{P}(X = j) = \begin{cases} 0, & \text{se } x < 0; \\ \sum_{j=0}^{[x]} \frac{e^{-\lambda}\lambda^j}{j!}, & \text{se } x \geqslant 0; \end{cases}$$

onde $[x]$ é a parte inteira de x.

□

Nas Figuras 5.3 e 5.4 apresentamos a função massa de probabilidade e acumulada para diferentes valores de p.

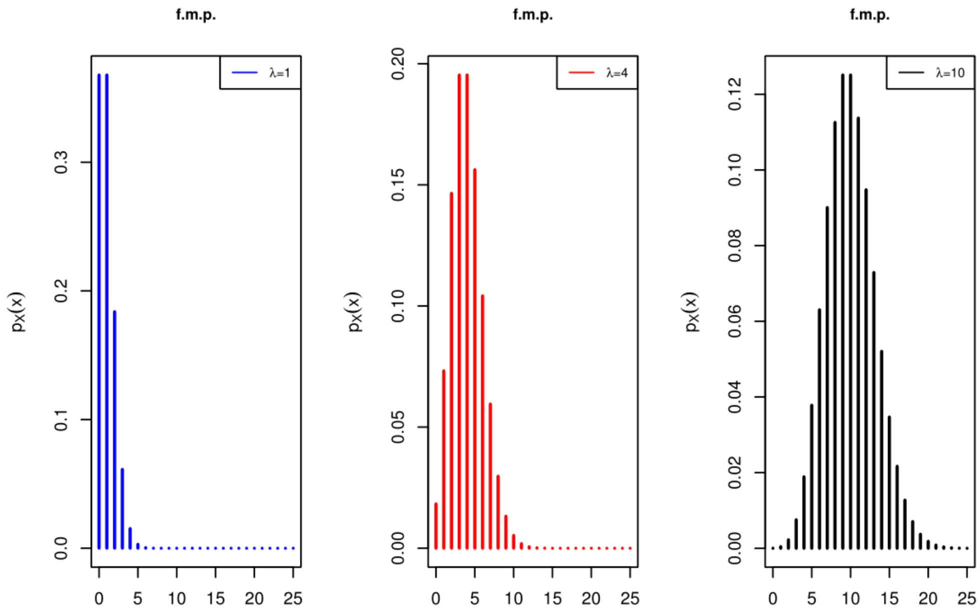


Figura 5.3: Função Massa de Probabilidade da variável aleatória $X \sim \mathrm{P}(\lambda)$.

Proposição 5.3. *Seja* $X \sim \mathrm{P}(\lambda)$, *então*

(i) $\mathbb{E}(X) = \lambda$.

(ii) $\mathbb{E}(X^2) = \lambda^2 + \lambda$.

(iii) $\mathrm{Var}(X) = \lambda$.

Demonstração. Seja $X \sim \mathrm{P}(\lambda)$, com $\lambda > 0$.

(i) Por definição, temos que

$$\mathbb{E}(X) = \sum_{x=0}^{\infty} x \frac{e^{-\lambda}\lambda^x}{x!} = \sum_{x=0}^{\infty} x \frac{e^{-\lambda}\lambda^x}{x(x-1)!} = \lambda \sum_{x=1}^{\infty} \frac{e^{-\lambda}\lambda^{(x-1)}}{(x-1)!}.$$

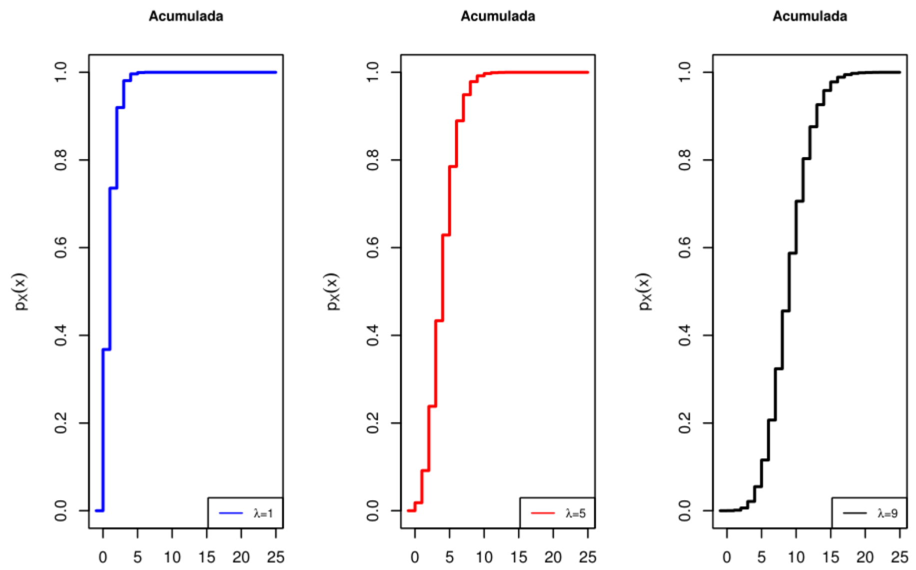


Figura 5.4: Função de Distribuição (ou acumulada) da variável aleatória $X \sim \mathrm{P}(\lambda)$.

Fazendo a mudança de variável $y = x - 1$, temos

$$\mathbb{E}(X) = \lambda \sum_{y=0}^{\infty} \frac{e^{-\lambda}\lambda^y}{y!} = \mathbb{E}(X) = \lambda e^{-\lambda} \sum_{y=0}^{\infty} \frac{\lambda^y}{y!} = \lambda e^{-\lambda} e^{\lambda} = \lambda,$$

pois $\sum_{y=0}^{\infty} \frac{\lambda^y}{y!} = e^{\lambda}$.

(ii) Pela Proposição 4.1, temos que

$$\mathbb{E}(X^2) \;=\; \sum_{x=0}^{\infty} x^2 \frac{e^{-\lambda}\lambda^x}{x!} = \sum_{x=0}^{\infty} x \frac{e^{-\lambda}\lambda^x}{(x-1)!} = \lambda \sum_{x=1}^{\infty} x \frac{e^{-\lambda}\lambda^{(x-1)}}{(x-1)!}$$

Fazendo a mudança de variável $y = x - 1$, temos

$$\begin{aligned}
\mathbb{E}(X^2) &= \lambda \sum_{x=1}^{\infty} x \frac{e^{-\lambda}\lambda^{(x-1)}}{(x-1)!} = \lambda \sum_{y=0}^{\infty} (y+1) \frac{e^{-\lambda}\lambda^y}{y!} \\
&= \lambda \sum_{y=0}^{\infty} \left[y \frac{e^{-\lambda}\lambda^y}{y!} + \frac{e^{-\lambda}\lambda^y}{y!} \right] \\
&= \lambda \left[\sum_{y=0}^{\infty} y \frac{e^{-\lambda}\lambda^y}{y!} + e^{-\lambda} \sum_{y=0}^{\infty} \frac{\lambda^y}{y!} \right] \\
&= \lambda \left[\underbrace{\sum_{y=0}^{\infty} y \frac{e^{-\lambda}\lambda^y}{y!}}_{\mathbb{E}(X)} + e^{-\lambda} \underbrace{\sum_{y=0}^{\infty} \frac{\lambda^y}{y!}}_{e^{\lambda}} \right] = \lambda(\lambda + 1).
\end{aligned}$$

(iii) Por definição, temos que

$$\text{Var}(X) = \mathbb{E}(X^2) - [\mathbb{E}(X)]^2 = \lambda(\lambda + 1) - \lambda^2 = \lambda.$$

□

Exemplo 5.8. Em um livro de 800 páginas há 800 erros de impressão. Qual a probabilidade de que uma página contenha pelo menos 3 erros?

Solução: Seja X o número de erros por página. Temos que $\lambda = 1$. Logo,

$$\begin{aligned}
\mathbb{P}(X \geqslant 3) &= 1 - \mathbb{P}(X < 3) = 1 - \{\mathbb{P}(X = 0) + \mathbb{P}(X = 1) + \mathbb{P}(X = 2)\} \\
&= 1 - \left\{ \frac{e^{-1} \cdot 1^0}{0!} + \frac{e^{-1} \cdot 1^1}{1!} + \frac{e^{-1} \cdot 1^2}{2!} \right\} \\
&= 1 - \{0.367879 + 0.367879 + 0.183940\} \\
&= 1 - \text{ppois(q=2,lambda=1)} \\
&= \text{ppois(q=2,lambda=1,lower.tail=FALSE)} \\
&= 1 - 0.9196986 = 0.0803014.
\end{aligned}$$

□

Exemplo 5.9. Em uma central telefônica chegam 300 telefonemas por hora. Qual a probabilidade de que

(i) num minuto não haja nenhuma chamada?

(ii) em 2 minutos haja 2 chamados?

(iii) em t minutos não haja chamados?

Solução:

(i) Seja X a variável aleatória número de chamadas por minuto. Então, $\lambda = 5$.

$$\mathbb{P}(X = 0) = \frac{e^{-5} \cdot 5^0}{0!} = \text{dpois(x=0,lambda=5)} = 0.006737947$$

(ii) Em dois minutos, $\lambda = 10$. Então,

$$\mathbb{P}(X = 2) = \frac{e^{-10} \cdot 10^2}{2!} = \text{dpois(x=2,lambda=10)} = 0.002269996$$

(iii) Em t minutos, $\lambda = 5t$. Então,

$$\mathbb{P}(X = 0) = \frac{e^{-5t} \cdot (5t)^0}{0!} = e^{-5t}.$$

□

Exemplo 5.10. 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 no máximo dois cortes? Pelo menos dois cortes?

Solução: Seja $Y = \{\text{número de cortes num rolo de 4000 pés}\}$. Então, $Y \sim \mathrm{P}(2)$.

Logo,

$$\begin{aligned}
\mathbb{P}(\text{no máximo 2 cortes}) &= \mathbb{P}(X \leqslant 2) = \sum_{x=0}^{2} \mathbb{P}(X = x) \\
&= \frac{e^{-2}2^0}{0!} + \frac{e^{-2}2^1}{1!} + \frac{e^{-2}2^2}{2!} \\
&= \texttt{sum(dpois(x=0:2,lambda=2))} \\
&= \texttt{ppois(q=2,lambda=2)} \\
&= 0.6766764.
\end{aligned}$$

$$\begin{aligned}
\mathbb{P}(\text{pelo menos 2 cortes}) &= \mathbb{P}(X \geqslant 2) = 1 - \mathbb{P}(X < 2) = 1 - [\mathbb{P}(X = 0) + \mathbb{P}(X = 1)] \\
&= 1 - \left[\frac{e^{-2}2^0}{0!} + \frac{e^{-2}2^1}{1!}\right] \\
&= \texttt{ppois(q=1,lambda=2,lower.tail=FALSE)} \\
&= 0.5939942.
\end{aligned}$$

□

5.4 Modelo Geométrico

5.4.1. Número de Ensaios até o Primeiro Sucesso

[https://bit.ly/3hdKcFK] Consideramos repetições sucessivas e independentes de um mesmo experimento de Bernoulli. E seja X o número de tentativas necessárias até o primeiro sucesso. Logo X assume os seguintes valores.

$X = 1$ - corresponde ao sucesso na primeira tentativa: $\mathbb{P}(X = 1) = p$;

$X = 2$ - corresponde ao fracasso na primeira tentativa e sucesso na segunda tentativa:

$$\mathbb{P}(X = 2) = \mathbb{P}(\mathrm{F,S}) = \mathbb{P}(\mathrm{F}) \times \mathbb{P}(\mathrm{S}) = (1 - p)p;$$

$X = 3$ - corresponde ao fracasso na primeira tentativa, fracasso na segunda tentativa e sucesso na terceira tentativa:

$$\mathbb{P}(X = 3) = \mathbb{P}(\mathrm{F,F,S}) = \mathbb{P}(\mathrm{F}) \times \mathbb{P}(\mathrm{F}) \times \mathbb{P}(\mathrm{S}) = (1 - p)^2 p;$$

Seguindo analogamente, temos a seguinte definição.

Definição 5.4. Uma variável aleatória discreta X segue o *modelo Geométrico* com parâmetro p, onde $0 < p < 1$, se a sua função massa de probabilidade é dada por

$$p_X(x) = \mathbb{P}(X = x) = p(1 - p)^{x-1}, \quad x = 1, 2, \cdots . \tag{5.6}$$

A variável aleatória X é o número de repetições necessárias até o aparecimento do primeiro sucesso e $\mathbb{P}(X = x)$ é a probabilidade de fracasso nos primeiros $x - 1$ experimentos e sucesso no $x-$ ésimo experimento.

Notação: $X \sim \text{Geo}(p)$, $0 < p < 1$. [A variável aleatória discreta X segue distribuição Geométrica com parâmetros p, com p a probabilidade de sucesso nos eventos de Bernoulli.]

Corolário 5.4. *Seja $X \sim \text{Geo}(p)$, com $0 < p < 1$. Então $p_X(\cdot)$, dada pela equação (5.6), satisfaz as propriedades de uma função massa de probabilidade, ou seja, é uma função massa de probabilidade.*

Demonstração. Seja $X \sim \text{Geo}(p)$, com $0 < p < 1$. Temos que $p_X(\cdot)$, dada pela equação (5.6), é $0 \leqslant p_X(x) \leqslant 1$, para todo $x = 1, 2, \cdots$. Como próximo passo, temos que verificar se

$$\sum_{x \geqslant 1} \mathbb{P}(X = x) = 1.$$

De fato, temos que

$$\sum_{x \geqslant 1} \mathbb{P}(X = x) = \sum_{x \geqslant 1} p(1-p)^{x-1} = p \sum_{x \geqslant 1} (1-p)^{x-1}.$$

Fazendo $y = x - 1$, então $y = 0, 1, 2, \cdots$ temos

$$\sum_{x \geqslant 1} \mathbb{P}(X = x) = p \sum_{x \geqslant 1} (1-p)^{x-1} = p \sum_{y \geqslant 0} (1-p)^{y}.$$

Como $\sum_{y \geqslant 0} (1-p)^y$ é uma série geométrica de razão $1 - p$ $(0 < p < 1)$ e sabemos que uma série geométrica de razão r converge quando $|r| < 1$ (https://bit.ly/2DTf00n) para

$$\sum_{y \geqslant 0} r^y = \frac{1}{1-r}.$$

Assim,

$$\sum_{x \geqslant 1} \mathbb{P}(X = x) = p \sum_{y \geqslant 0} (1-p)^{y} = p \frac{1}{1-(1-p)} = \frac{p}{p} = 1.$$

Portanto, $p_X(\cdot)$, dada pela equação (5.6), satisfaz as propriedades de uma função massa de probabilidade .

□

Resultado 5.4. Seja X uma variável aleatória discreta com distribuição $\text{Geo}(p)$, com $0 < p < 1$, então sua função acumulada de probabilidade é dada por

$$F_X(x) = \mathbb{P}(X \leqslant x) = \sum_{j=1}^{x} \mathbb{P}(X = j) = \sum_{j=1}^{x} p(1-p)^{j-1}.$$

Tomando $i = j - 1$ na expressão acima, temos

$$F_X(x) = p \sum_{i=0}^{x-1} (1-p)^i = p \frac{1-(1-p)^{x-1+1}}{1-(1-p)} = 1 - (1-p)^{[x]},$$

pois $\sum_{i=0}^{n} (a)^i = \frac{1-a^{n+1}}{1-a}$, para $0 < a < 1$ e $[x]$ é a parte inteira de $x \geqslant 1$. Logo,

$$F_X(x) = \begin{cases} 0, & \text{se } x < 1; \\ 1 - (1-p)^{[x]}, & \text{se } x \geqslant 1. \end{cases}$$

$\square$

Nas Figuras 5.5 e 5.6 apresentamos a função massa de probabilidade e acumulada para diferentes valores de p.

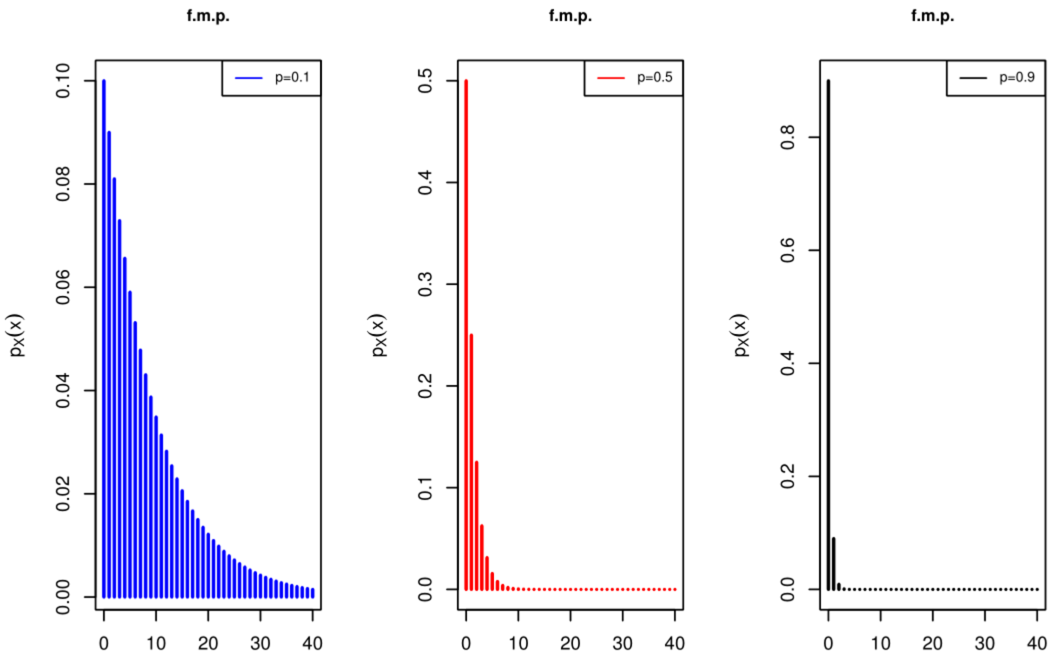


Figura 5.5: Função Massa de Probabilidade da variável aleatória $X \sim Geo(p)$.

Proposição 5.4. *Seja* $X \sim \text{Geo}(p)$, *então*

(i) $\mathbb{E}(X) = \frac{1}{p}$.

(ii) $\mathbb{E}(X^2) = \frac{2-p}{p^2}$

(iii) $\text{Var}(X) = \frac{1-p}{p^2}$.

Demonstração. Seja $X \sim \text{Geo}(p)$.

(i) Iniciamos calculando $\mathbb{E}(X)$.

$$\mathbb{E}(X) = \sum_{x \geqslant 1} x \mathbb{P}(X = x) = \sum_{x \geqslant 1} x p (1-p)^{x-1} = p \sum_{x \geqslant 1} x (1-p)^{x-1}$$

Fazendo a mudança de variável $y = x - 1$, temos $x = y + 1$. Logo,

$$\mathbb{E}(X) = p \sum_{y \geqslant 0} (y+1)(1-p)^y = p \left[\frac{1}{[1-(1-p)]^2} \right],$$

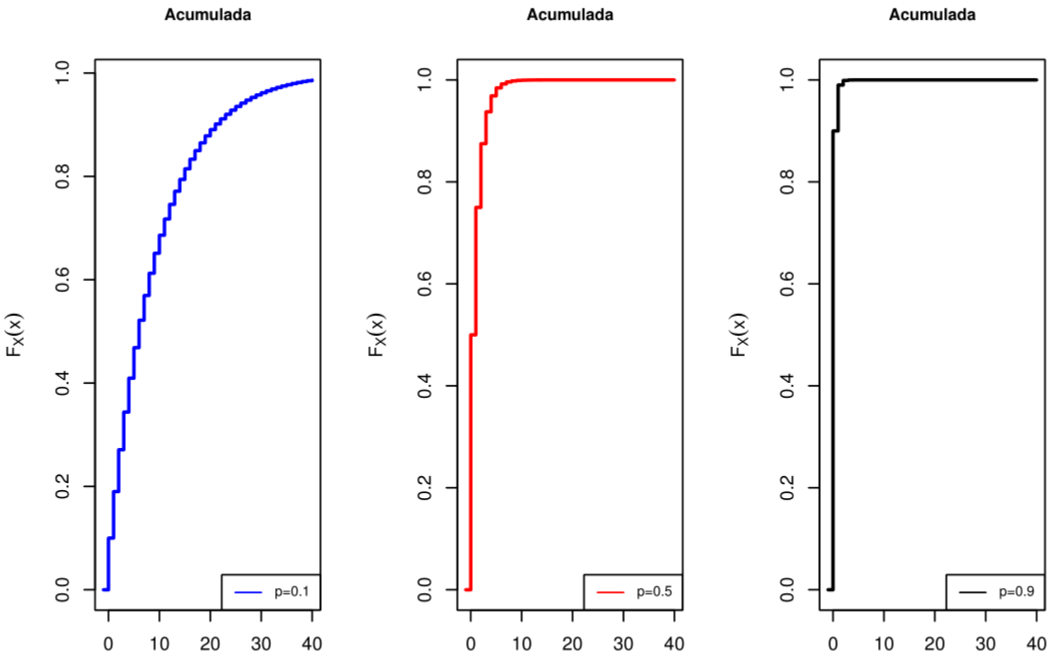


Figura 5.6: Função Distribuição (acumulada) da variável aleatória $X \sim Geo(p)$.

pois, $\sum_{i \geqslant 0} (i+1) r^i = \dfrac{1}{[1-r]^2}$, onde $0 < r < 1$ é a razão da série geométrica. Assim,

$$\mathbb{E}(X) = p \left[\frac{1}{[1-(1-p)]^2} \right] = p \frac{1}{(-p)^2} = \frac{1}{p}.$$

(ii) Vamos calcular $\mathbb{E}(X^2)$.

$$\mathbb{E}(X^2) = \sum_{x \geqslant 1} x^2 \mathbb{P}(X = x) = \sum_{x \geqslant 1} x^2 p (1-p)^{x-1}$$

Transformando $x^2 = x(x-1) + x$, temos

$$\begin{aligned}
\mathbb{E}(X^2) &= \sum_{x \geqslant 1} x^2 p(1-p)^{x-1} = \sum_{x \geqslant 1} [x(x-1) + x] p (1-p)^{x-1} \\
&= \sum_{x \geqslant 1} x(x-1) p (1-p)^{x-1} + \underbrace{\sum_{x \geqslant 1} x p (1-p)^{x-1}}_{\mathbb{E}(X)} \\
&= \sum_{x \geqslant 2} x(x-1) p (1-p)^{x-1} + \frac{1}{p} \\
&= p(1-p) \sum_{x \geqslant 2} x(x-1)(1-p)^{x-2} + \frac{1}{p}.
\end{aligned}$$

Fazendo a mudança de variável $y = x - 2$ $(x = y + 2)$ e utilizando a série geométrica $\sum_{i \geqslant 0} (i + 2)(i+1) r^i = \dfrac{2}{[1-r]^3}$, onde $0 < r < 1$ é a razão da série geométrica, temos

$$\begin{aligned}
\mathbb{E}(X^2) &= p(1-p)\sum_{x\geqslant 2} x(x-1)(1-p)^{x-2} + \frac{1}{p} \\
&= p(1-p)\sum_{y\geqslant 0}(y+2)[(y+2)-1](1-p)^{y} + \frac{1}{p} \\
&= p(1-p)\sum_{y\geqslant 0}(y+2)(y+1)(1-p)^{y} + \frac{1}{p} \\
&= p(1-p)\frac{2}{[1-(1-p)]^3} + \frac{1}{p} = \frac{2p(1-p)}{p^3} + \frac{1}{p} \\
&= \frac{2p(1-p)+p^2}{p^3} = \frac{2(1-p)+p}{p^2} = \frac{2-p}{p^2}.
\end{aligned}$$

(iii) Pela definição de variância temos

$$\mathrm{Var}(X) = \mathbb{E}(X^2) - [\mathbb{E}(X)]^2 = \frac{2-p}{p^2} - \left[\frac{1}{p}\right]^2 = \frac{1-p}{p^2}.$$

□

Exemplo 5.11. Um atirador acerta na mosca do alvo, 20% dos tiros. Qual a probabilidade de ele acertar na mosca pela primeira vez no 10º tiro?

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 desse número ser 10. Seja $X = \{$número de tiros até primeiro acerto$\}$. Então, $X \sim \mathrm{Geo}(0,20)$. Queremos calcular $\mathbb{P}(X = 10)$. Logo,

$$\mathbb{P}(X = 10) = 0,2 \times 0,8^9 = 0,02684.$$

□

Exemplo 5.12. As cinco primeiras repetições de um experimento custam R$10,00 cada. Todas as repetições subsequentes custam R$ 5,00 cada. Suponha que o experimento seja repetido até que o primeiro sucesso ocorra. Se a probabilidade de sucesso de uma repetição é igual a 0.9, e se as repetições independentes, qual é o custo esperado?

Solução: Seja X=\{o número de tentativas\} e C=\{o custo dos experimentos\}. Sabemos que iremos realizar o experimento até que o primeiro sucesso ocorra ou seja, temos uma Distribuição Geométrica, ou seja $X \sim \mathrm{Geo}(p)$, com $p = 0,9$.

Queremos o custo esperado. O cálculo do custo é feito da seguinte forma:
- 5 primeiras tentativas: R$10,00 cada.
- Outras tentativas: R$5,00 cada.

Assim podemos escrever que:

$$\mathbb{E}(C) = 10\sum_{k=1}^{5}\mathbb{P}(X = k) + 5\sum_{k=6}^{\infty}\mathbb{P}(X = k).$$

Como $\sum_{i=0}^{n}(a)^i = \frac{1-a^{n+1}}{1-a}$, para $0 < a < 1$, temos que,

$$\begin{aligned} E(C) &= 10 \times 0,9(0,1^0 + 0,1^1 + 0,1^2 + 0,1^3 + 0,1^4) \\ &\quad +5 \times \left[1 - 0,9(0,1^0 + 0,1^1 + 0,1^2 + 0,1^3 + 0,1^4)\right] \\ &= 10 \times 0,9 \times 1,1111 + 5[1 - 0,9 \times 1,1111] = 9,99995. \end{aligned}$$

□

Exemplo 5.13. Em seu caminho matinal, você se aproxima de um determinado sinal de trânsito, que está verde em 20% do tempo. Suponha que cada manhã represente um tentativa independente.

(a) Qual é a probabilidade de que a primeira manhã que a luz esteja verde seja a quarta manhã que você se aproxima?

(b) Qual é a probabilidade de que a luz não esteja verde durante exatamente 10 manhãs consecutivas?

Solução:

(a) Qual é a probabilidade de que a primeira manhã que a luz esteja verde seja a quarta manhã que você se aproxima?

A probabilidade de pegar sinal verde é de 20%. Queremos a probabilidade de que a primeira vez que eu pegue o sinal verde seja na quarta tentativa. Ou seja estamos realizando repetições até obter o sucesso.

Se X é a v.a que representa o número de tentativas até o primeiro sucesso, então $X \sim \text{Geo}(0,2)$ e queremos calcular $\mathbb{P}(X = 4)$. Assim,

$$\mathbb{P}(X = 4) = 0,2(1 - 0,2)^{4-1} = 0,1024.$$

(b) Qual é a probabilidade de que a luz não esteja verde durante exatamente 10 manhãs consecutivas?

Pra que a luz não esteja verde dentro das 10 primeiras manhãs consecutivas, concorda que a gente tem que calcular a probabilidade de a gente só encontrar ela verde na décima primeira manhã pela primeira vez? Vamos utilizar a mesma fórmula só que agora pra $k = 11$.

$\mathbb{P}(X = 11) = 0,2(1 - 0,2)^{11-1} = 0,02147.$

□

Exemplo 5.14. No Call center de uma empresa distribuidora de telefonia, apenas 35% das chamadas são relacionadas a reclamações sobre erros nas faturas emitidas pela empresa. Pede-se:

(a) Qual a probabilidade da primeira reclamação sobre erro na fatura emitida da conta, ocorrer até a 2ª chamada.

(b) A média, desvio padrão desta variável aleatória.

Solução:

(a) Qual a probabilidade da primeira reclamação sobre erro na fatura emitida da conta, ocorrer até a 2ª chamada.

Queremos calcular a probabilidade do primeiro sucesso (reclamação sobre erro na fatura) acontecer até a segunda chamada.

Temos duas situações então: a primeira reclamação acontecer na primeira chamada ou a primeira reclamação acontecer na segunda chamada. Nos dois casos, vamos usar a distribuição geométrica com $p = 0,35$, isto por que queremos calcular a probabilidade do primeiro sucesso acontecer em determinada tentativa.

Situações: na primeira $k = 1$ e na segunda $k = 2$ que é o número da tentativa em que acontece o primeiro sucesso.

$$\mathbb{P}(X = 1) = 0,35 \cdot 0,65^0 = 0,35 \quad \mathbb{P}(X = 2) = 0,35 \cdot 0,65^1 = 0,2275$$

A probabilidade que queremos é a soma das duas.

$$\mathbb{P}(X \leqslant 2) = 0,35 + 0,2275 = 0,5775.$$

(b) A média, desvio padrão desta variável aleatória.

Para o caso de uma distribuição geométrica conhecemos a fórmula da média e da variância

$$\mathbb{E}(X) = \frac{1}{p} = \frac{1}{0,35} = 2,86$$

$$\text{Var}(X) = \frac{(1-p)}{p^2} = \frac{0,65}{0,35^2} = 5,31.$$

O desvio padrão é a raiz quadrada da variância:

$$DP(X) = \sqrt{5,31} = 2,30.$$

□

5.4.2. Número de Falhas até o Primeiro Sucesso

Definição 5.5. Seja X a variável aleatória que fornece o número de falhas até o primeiro sucesso. A variável X tem distribuição Geométrica com parâmetro p, $0 < p < 1$, se sua função de probabilidade é dada por

$$\mathbb{P}(X = x) = (1-p)^x p, \quad x = 0, 1, \ldots \tag{5.7}$$

O evento $[X = x]$ ocorre se, e somente se, ocorrem somente falhas nos x primeiros ensaios e sucesso no $(x+1)$-ésimo ensaio.

Notação: $X \sim \text{Geo}(p)$, $0 < p < 1$. [Seja X a variável aleatória que fornece o número de falhas até o primeiro sucesso. A variável aleatória X possui distribuição Geométrica com parâmetro p.]

A distribuição geométrica tem uma propriedade que serve para caracterizá-la no conjunto das distribuições discretas:

Corolário 5.5. *Seja $X \sim \text{Geo}(p)$, com $0 < p < 1$. Então $p_X(\cdot)$, dada pela equação (5.7), satisfaz as propriedades de uma função massa de probabilidade, ou seja, é uma função massa de probabilidade.*

Demonstração. Seja $X \sim \text{Geo}(p)$, com $0 < p < 1$. Temos que $p_X(\cdot)$, dada pela equação (5.7), é $0 \leqslant p_X(x) \leqslant 1$, para todo $x = 0, 1, 2, \cdots$. Como próximo passo, temos que verificar se

$$\sum_{x \geqslant 0} \mathbb{P}(X = x) = 1.$$

De fato, temos que

$$\sum_{x \geqslant 0} \mathbb{P}(X = x) = \sum_{x \geqslant 0} p(1-p)^x = p \sum_{x \geqslant 0} (1-p)^x.$$

Como $\sum_{x \geqslant 0}(1-p)^x$ é uma série geométrica de razão $1 - p$ $(0 < p < 1)$ e sabemos que uma série geométrica de razão r converge quando $|r| < 1$ (https://bit.ly/2DTf00n) para

$$\sum_{x \geqslant 0} r^x = \frac{1}{1-r}.$$

Assim,

$$\sum_{x \geqslant 0} \mathbb{P}(X = x) = p \sum_{x \geqslant 0} (1-p)^x = p \frac{1}{1-(1-p)} = \frac{p}{p} = 1.$$

Portanto, $p_X(\cdot)$, dada pela equação (5.7), satisfaz as propriedades de uma função massa de probabilidade .

□

Resultado 5.5. Seja X uma variável aleatória discreta com distribuição $\text{Geo}(p)$, com $0 < p < 1$, então sua função acumulada de probabilidade é dada por

$$F_X(x) = \mathbb{P}(X \leqslant x) = \sum_{j=0}^{[x]} \mathbb{P}(X = j) = \sum_{j=0}^{[x]} p(1-p)^j = 1 - (1-p)^{[x]},$$

pois $\sum_{i=0}^{n}(a)^i = \frac{1-a^{n+1}}{1-a}$, para $0 < a < 1$ e $[x]$ é a parte inteira de $x \geqslant 0$. Logo,

$$F_X(x) = \begin{cases} 0, & \text{se } x < 0; \\ 1 - (1-p)^{[x]}, & \text{se } x \geqslant 0. \end{cases}$$

□

Proposição 5.5. *Seja* $X \sim \text{Geo}(p)$, *então*

(i) $\mathbb{E}(X) = \frac{1-p}{p}$.

(ii) $\mathbb{E}(X^2) = \frac{2-3p+p^2}{p^2}$

(iii) $\text{Var}(X) = \frac{1-p}{p^2}$.

Demonstração. Seja $X \sim \text{Geo}(p)$.

i) Iniciamos calculando $\mathbb{E}(X)$.

$$\mathbb{E}(X) = \sum_{x \geqslant 0} x\mathbb{P}(X = x) = \sum_{x \geqslant 1} x\mathbb{P}(X = x) = \sum_{x \geqslant 1} xp(1-p)^x = p(1-p) \sum_{x \geqslant 1} x(1-p)^{x-1}.$$

Como $\sum_{i \geqslant 1} ir^{i-1} = \frac{1}{[1-r]^2}$, onde $|r| < 1$ é uma série geométrica [https://bit.ly/2ZLAXa4], temos que

$$\mathbb{E}(X) = p(1-p)\left[\frac{1}{[1-(1-p)]^2}\right] = p(1-p)\frac{1}{(-p)^2} = \frac{1-p}{p}.$$

ii) Vamos calcular $\mathbb{E}(X^2)$.

$$\mathbb{E}(X^2) = \sum_{x \geqslant 0} x^2\mathbb{P}(X = x) = \sum_{x \geqslant 0} x^2 p(1-p)^x$$

Transformando $x^2 = x(x-1) + x$, temos

$$\begin{aligned}
\mathbb{E}(X^2) &= \sum_{x \geqslant 1} x^2 p(1-p)^x = \sum_{x \geqslant 1}[x(x-1)+x]p(1-p)^x \\
&= \sum_{x \geqslant 2} x(x-1)p(1-p)^x + \underbrace{\sum_{x \geqslant 1} xp(1-p)^x}_{\mathbb{E}(X)} \\
&= p(1-p)^2 \sum_{x \geqslant 2} x(x-1)(1-p)^{x-2} + \frac{1-p}{p} \\
&= p(1-p)^2 \frac{2}{(1-(1-p))^3} + \frac{1-p}{p} \\
&= \frac{2(1-p)^2}{p^2} + \frac{1-p}{p} = \frac{2-3p+p^2}{p^2}
\end{aligned}$$

pois utilizamos a série geométrica $\sum_{i \geqslant 2} i(i-1)r^{i-2} = \frac{2}{[1-r]^3}$, onde $|r| < 1$ [https://bit.ly/2ZLAXa4].

iii) Pela definição de variância temos

$$\text{Var}(X) = \mathbb{E}(X^2) - [\mathbb{E}(X)]^2 = \frac{2-3p+p^2}{p^2} - \left[\frac{(1-p)}{p}\right]^2 = \frac{1-p}{p^2}.$$

□

Exemplo 5.15. Considere o experimento em que uma moeda viciada é lançada sucessivas vezes, até que ocorra a primeira cara. Seja X a variável aleatória que conta o número de coroas obtidos no experimento (ou seja, a quantidade de lançamentos anteriores a obtenção da primeira cara). Sabendo que a probabilidade de cara é de $0,4$, qual é a probabilidade de $\mathbb{P}(2 \leqslant X < 4)$ e a probabilidade de $\mathbb{P}(X > 1 | X \leqslant 2)$.

Solução: Observamos que

$$\mathbb{P}(2 \leqslant X < 4) = \mathbb{P}(X = 2) + \mathbb{P}(X = 3) = 0,6^2 \cdot 0,4 + 0,6^3 \cdot 0,4 = 0,2304.$$

Vamos calcular agora $\mathbb{P}(X > 1 | X \leqslant 2)$.

$$\mathbb{P}(X > 1 | X \leqslant 2) = \frac{\mathbb{P}(X > 1 \cap X \leqslant 2)}{\mathbb{P}(X \leqslant 2)} = \frac{\mathbb{P}(X = 2)}{\mathbb{P}(X = 0) + \mathbb{P}(X = 1) + \mathbb{P}(X = 2)} = \frac{0,144}{0,784} = 0,18367.$$

□

Exemplo 5.16. Considerando o Exemplo 5.15, qual a probabilidade de que $X \geqslant 1$.

Solução:

$$\mathbb{P}(X \geqslant 1) = 1 - \mathbb{P}(X < 1) = 1 - \mathbb{P}(X = 0) = 1 - 0,4 = 0,6.$$

□

Exemplo 5.17. Um pesquisador está realizando um experimentos químico independentes e sabe que a probabilidade de que cada experimento apresente uma reação positiva é $0,3$. Qual é a probabilidade de que menos de 5 reações negativas ocorram antes da primeira positiva?

Solução: Para resolver este problema, considere X como sendo a variável aleatória que representa o número de reações negativas até a ocorrência da primeira positiva. Neste caso, temos que $X \sim \text{Geo}(0,3)$ e então

$$\mathbb{P}(X < 5) = \sum_{x=0}^{4} \mathbb{P}(X = x) = 0,3 + 0,7 \cdot 0,3 + 0,7^2 \cdot 0,3 + 0,7^3 \cdot 0,3 + 0,7^4 \cdot 0,3 = 0,83193.$$

□

5.4.3. Falta de Memória da Distribuição Geométrica

Seja $X \sim \text{Geo}(p)$ e vamos verificar que, para quaisquer números inteiros positivos m e n, vale a seguinte relação

$$\mathbb{P}(X \geqslant m + n | X \geqslant m) = \mathbb{P}(X \geqslant n).$$

Esta propriedade é conhecida como *falta de memória*. Ela indica a maneira como a variável aleatória incorpora a informação anterior, ou seja, podemos considerar que a variável “lembra” do presente, mas “esqueceu” do que ocorreu no passado.

Por exemplo, se X representa a espera em dias para a ocorrência de um certo evento, a probabilidade condicional acima representaria a probabilidade de espera de pelo menos $m + n$ dias, sabendo que o evento não ocorreu antes de m dias. A propriedade de falta de memória estabelece que essa probabilidade é a mesma de esperar pelo menos n dias.

Para verificar a propriedade, observe que

$$\mathbb{P}(X \geqslant m + n | X \geqslant m) = \frac{\mathbb{P}([X \geqslant m + n] \cap [X \geqslant m])}{\mathbb{P}(X \geqslant m)} = \frac{\mathbb{P}([X \geqslant m + n]}{\mathbb{P}(X \geqslant m)}.$$

Como a função massa de probabilidade de X é dada por $p_X(x) = \mathbb{P}(X = x) = p(1-p)^x$, obtemos

$$\mathbb{P}(X \geqslant m + n | X \geqslant m) = \frac{\mathbb{P}([X \geqslant m + n]}{\mathbb{P}(X \geqslant m)} = \frac{(1-p)^{m+n}}{(1-p)^m} = (1-p)^n = \mathbb{P}(X \geqslant n).$$

Pode-se verificar que o modelo Geométrico é o único modelo discreto com a propriedade de falta de memória.

5.5 Modelo Hipergeométrico

Considere uma população de tamanho N dividida em 2 classes, uma composta de r *sucessos* e a outra composta de $N - r$ *fracassos*. Dessa população, vamos extrair uma amostra de tamanho n sem reposição. Veja Figura 5.7.

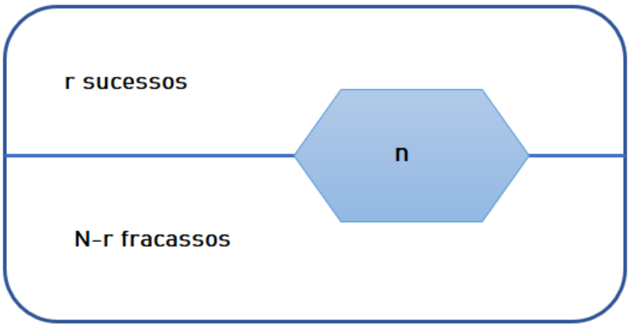


Figura 5.7: Ilustração do experimento que define uma variável aleatória hipergeométrica.

Vamos considerar a seguinte variável aleatória para esse experimento:

$$X = \text{número de sucessos em uma amostra retirada sem reposição.}$$

Possíveis Valores de X

Para determinar os valores possíveis de X, vamos considerar o problema em termos de uma urna com bolas verdes (Sucessos) e brancas (Fracassos), só para facilitar. Então, temos que considerar as seguintes situações:

→ Se eu tiver bolas suficientes de ambas as cores (sucessos e fracassos), isto é, se $r \geqslant n$ e $N - r \geqslant n$, então os possíveis valores de X variam de 0 (todas as bolas na amostra podem ser brancas, fracassos) a n (todas as bolas na amostra podem ser verdes, sucessos).

→ Se eu não tiver bolas brancas suficientes (fracassos), isto é, se $N - r < n$, então teremos que ter pelo menos $n - (N - r)$ bolas verdes (sucessos) na amostra. Juntando esse resultado com a observação anterior, conclui-se que o valor mínimo de X é

$$\max\{0, n - (N - r)\}.$$

Exemplo: $N = 6$, $r = 4$ e $n = 3$: $N - r = 2 < 3$; como só há duas bolas brancas, temos que ter pelo menos 1 bola verde na amostra, isto é, o valor mínimo de X é $1 = \max\{0, 3 - 2\}$.

→ Se eu não tiver bolas verdes suficientes (sucessos), isto é, se $r < n$, então o máximo de bolas brancas na amostra será r. Juntando esse resultado com o resultado da primeira observação, conclui-se que o valor máximo de X é

$$\min\{n, r\}.$$

Exemplo: $N = 6$, $r = 2$ e $n = 3$; como só há duas bolas verdes, o número máximo possível de bolas brancas na amostra é $2 = \min\{3, 2\}$.

Cálculo das Probabilidades

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).

Consideremos inicialmente a situação em que existem bolas suficientes de ambas as cores, isto é, $r \geqslant n$ e $N - r \geqslant n$. A presença de k bolas verdes (sucessos) na amostra implica na presença de $n - k$ bolas brancas (fracassos) e o número de tais amostras é

$$\binom{r}{x} \times \binom{N-r}{n-x},$$

pelo princípio fundamental da multiplicação. Logo,

$$p_X(x) = \mathbb{P}(X = x) = \frac{\binom{r}{x} \times \binom{N-r}{n-x}}{\binom{N}{n}}, \quad x \in \{0, 1, 2, \cdots, n\}, x \leqslant n.$$

Definição 5.6. No modelo hipergeométrico, a população é constituída de N elementos, dos quais r são classificados como sucessos e $n - r$ de fracassos. É selecionado, ao acaso e sem reposição, uma amostra de tamanho $n < N$.
Seja X a variável aleatória número de sucessos na amostra de tamanho n. Assim,

$$\mathbb{P}(sucesso) = \frac{r}{N} \qquad \mathbb{P}(fracasso) = 1 - \frac{r}{N} = \frac{N-r}{N}.$$

Podemos tirar $\binom{N}{n}$ amostras sem reposição. Os sucessos na amostra podem ocorrer de $\binom{r}{x}$ maneiras e fracassos $\binom{N-r}{n-x}$ modos. Então as probabilidades de uma variável aleatória hipergeométrica podem ser avaliadas por

$$p_X(x) = \mathbb{P}(X = x) = \frac{\binom{r}{x}\binom{N-r}{n-x}}{\binom{N}{n}}, \quad x \in \{0, 1, \cdots, min(r, n)\} \text{ ou } 0 \leqslant x \leqslant n \text{ e } x \leqslant r, \quad (5.8)$$

onde $\binom{n}{x} = \frac{n!}{x!(n-x)!}$ é o Binômio de Newton.
O evento $[X = x]$ significa que foram retirados x elementos dentre os r sucessos e foram retirados $n - x$ dentre os $n - r$ fracassos.

Notação: $X \sim \mathrm{H}(N, n, r)$, com $n < N$ e $r < N$. [$[X = x]$ significa que foram retirados x elementos dentre os r sucessos e foram retirados $n - x$ dentre os $n - r$ fracassos.]

Corolário 5.6. *Seja $X \sim \mathrm{H}(N, n, r)$, com $n < N$ e $r < N$. Então $p_X(\cdot)$ dada pela equação (5.8) satisfaz as propriedades de uma função massa de probabilidade, ou seja, é uma função massa de probabilidade.*

Demonstração. Para provar que (5.8) realmente define uma função de distribuição de probabilidade, note inicialmente que $\mathbb{P}(X = x) \geqslant 0$. Como próximo passo temos que provar que $\sum_{x=0}^{n} \mathbb{P}(X = x) = 1$. Assim no caso da distribuição hipergeométrica, temos que

$$\sum_{x=0}^{n} \mathbb{P}(X = x) = \sum_{x=0}^{n} \frac{\binom{r}{x}\binom{N-r}{n-x}}{\binom{N}{n}} = \frac{1}{\binom{N}{n}} \sum_{x=0}^{n} \binom{r}{x}\binom{N-r}{n-x}.$$

Considerando a Fórmula de Euler

$$\sum_{j=0}^{m} \binom{a}{j}\binom{b}{m-j} = \binom{a+b}{m},$$

e fazendo $j = x$, $m = n$, $a = r$ e $b = N - r$, temos

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

Portanto,

$$\sum_{x=0}^{n} \mathbb{P}(X = x) \frac{1}{\binom{N}{n}} \sum_{x=0}^{n} \binom{r}{x}\binom{N-r}{n-x} = \frac{1}{\binom{N}{n}} \binom{N}{n} = 1.$$

Desta forma, $p_X(\cdot)$, dada pela equação (5.8), realmente define uma função de distribuição de probabilidade.

□

Proposição 5.6. *Seja* $X \sim \mathrm{H}(N, n, r)$, *então*

(i) $\mathbb{E}(X) = \frac{nr}{N}$.

(ii) $\mathbb{E}(X^2) = \frac{rn}{N}\left[\frac{(n-1)(r-1)+(N-1)}{(N-1)}\right]$

(iii) $\mathrm{Var}(X) = r\left[\frac{n}{N}\right]\left[\frac{N-r}{N}\right]\left[\frac{N-n}{N-1}\right]$.

Demonstração. Seja X uma variável aleatória discreta com distribuição $\mathrm{H}(N, n, r)$, ou seja, sua função massa de probabilidade é dada por

$$p_X(x) = \mathbb{P}(X = x) = \frac{\binom{r}{x}\binom{N-r}{n-x}}{\binom{N}{n}}, \quad x \in \{0, 1, \cdots, min(r, n)\} \text{ ou } 0 \leqslant x \leqslant n \text{ e } x \leqslant r,$$

onde $\binom{n}{x} = \frac{n!}{x!(n-x)!}$ é o Binômio de Newton.

(i) Iniciamos calculando $\mathbb{E}(X)$.

$$
\begin{aligned}
\mathbb{E}(X) = \sum_{x=0}^{n} x\mathbb{P}(X=x) &= \sum_{x=0}^{n} x \frac{\binom{r}{x}\binom{N-r}{n-x}}{\binom{N}{n}} = \frac{1}{\binom{N}{n}} \sum_{x=1}^{n} x \binom{r}{x}\binom{N-r}{n-x} \quad \text{(note o índice)} \\
&= \frac{1}{\binom{N}{n}} \sum_{x=1}^{n} x \frac{r(r-1)!}{x(x-1)!(r-x)!} \binom{N-r}{n-x} \quad (\text{pois } x \neq 0) \\
&= \frac{r}{\binom{N}{n}} \sum_{x=1}^{n} \frac{(r-1)!}{(x-1)!(r-x)!} \binom{N-r}{n-x} \\
&= \frac{r}{\binom{N}{n}} \sum_{x=1}^{n} \frac{(r-1)!}{(x-1)!((r-1)-(x-1))!} \binom{N-r}{n-x}
\end{aligned}
$$

Fazendo a mudança de variável $j = x - 1$, temos

$$
\begin{aligned}
\mathbb{E}(X) &= \frac{r}{\binom{N}{n}} \sum_{x=1}^{n} \frac{(r-1)!}{(x-1)!((r-1)-(x-1))!} \binom{N-r}{n-x} \\
&= \frac{r}{\binom{N}{n}} \sum_{j=0}^{n-1} \frac{(r-1)!}{j!((r-1)-j)!} \binom{N-r}{n-(j+1)} \\
&= \frac{r}{\binom{N}{n}} \sum_{j=0}^{n-1} \binom{r-1}{j} \binom{N-r}{n-(j+1)} \\
&= \frac{r}{\binom{N}{n}} \sum_{j=0}^{n-1} \binom{r-1}{j} \binom{(N-1)-(r-1)}{n-1-j}
\end{aligned}
$$

Considerando a Fórmula de Euler

$$
\sum_{j=0}^{m} \binom{a}{j}\binom{b}{m-j} = \binom{a+b}{m},
$$

e fazendo $j = j$, $m = n - 1$, $a = r - 1$ e $b = (N - 1) - (r - 1)$, temos

$$
\sum_{j=0}^{n-1} \binom{r-1}{j}\binom{(N-1)-(r-1)}{n-1-j} = \binom{(r-1)+(N-1)-(r-1)}{n-1} = \binom{N-1}{n-1}.
$$

Assim,

$$
\begin{aligned}
\mathbb{E}(X) &= \frac{r}{\binom{N}{n}} \sum_{j=0}^{n-1} \binom{r-1}{j} \binom{(N-1)-(r-1)}{n-1-j} \\
&= \frac{r}{\binom{N}{n}} \binom{N-1}{n-1} = r\frac{n!(N-n)!}{N!} \frac{(N-1)!}{(n-1)!(N-n)!} \\
&= r\frac{n}{N} = \frac{nr}{N}.
\end{aligned}
$$

(ii) Agora vamos calcular $\mathbb{E}(X^2)$. Portanto,

$$
\begin{aligned}
\mathbb{E}(X^2) = \sum_{x=0}^{n} x^2 \mathbb{P}(X=x) &= \sum_{x=0}^{n} x^2 \frac{\binom{r}{x}\binom{N-r}{n-x}}{\binom{N}{n}} = \frac{1}{\binom{N}{n}} \sum_{x=1}^{n} x^2 \binom{r}{x}\binom{N-r}{n-x} \quad \text{(note o índice)} \\
&= \frac{1}{\binom{N}{n}} \sum_{x=1}^{n} x^2 \frac{r(r-1)!}{x(x-1)!(r-x)!} \binom{N-r}{n-x} \quad (\text{pois } x \neq 0) \\
&= \frac{r}{\binom{N}{n}} \sum_{x=1}^{n} x \frac{(r-1)!}{(x-1)!(r-x)!} \binom{N-r}{n-x} \\
&= \frac{r}{\binom{N}{n}} \sum_{x=1}^{n} x \frac{(r-1)!}{(x-1)!((r-1)-(x-1))!} \binom{N-r}{n-x}
\end{aligned}
$$

Fazendo a mudança de variável $j = x - 1$, temos

$$
\begin{aligned}
\mathbb{E}(X^2) &= \frac{r}{\binom{N}{n}} \sum_{x=1}^{n} x \frac{(r-1)!}{(x-1)!((r-1)-(x-1))!} \binom{N-r}{n-x} \\
&= \frac{r}{\binom{N}{n}} \sum_{j=0}^{n-1} (j+1) \frac{(r-1)!}{j!((r-1)-j)!} \binom{N-r}{n-(j+1)} \\
&= \frac{r}{\binom{N}{n}} \sum_{j=0}^{n-1} (j+1) \binom{r-1}{j} \binom{N-r}{n-(j+1)} \\
&= \frac{r}{\binom{N}{n}} \sum_{j=0}^{n-1} (j+1) \binom{r-1}{j} \binom{(N-1)-(r-1)}{n-1-j} \\
&= \frac{r}{\binom{N}{n}} \left[\sum_{j=0}^{n-1} j \binom{r-1}{j} \binom{(N-1)-(r-1)}{n-1-j} + \sum_{j=0}^{n-1} \binom{r-1}{j} \binom{(N-1)-(r-1)}{n-1-j} \right] \\
&= \frac{r\binom{N-1}{n-1}}{\binom{N}{n}} \left[\sum_{j=0}^{n-1} j \frac{\binom{r-1}{j} \binom{(N-1)-(r-1)}{n-1-j}}{\binom{N-1}{n-1}} + \sum_{j=0}^{n-1} \frac{\binom{r-1}{j} \binom{(N-1)-(r-1)}{n-1-j}}{\binom{N-1}{n-1}} \right]
\end{aligned}
$$

- $\rightarrow$ O primeiro somatório é a esperança de uma hipergeométrica com parâmetros $N-1$, $n-1$ e $r-1$ [$\mathrm{H}(N-1, n-1, r-1)$], dada por $\mathbb{E}(X) = \frac{(n-1)(r-1)}{(N-1)}$;
- $\rightarrow$ O segundo somatório é a soma das probabilidades no espaço amostral de uma hipergeométrica com os mesmos parâmetros, ou seja, 1.

Segue, então, que

$$
\begin{aligned}
\mathbb{E}(X^2) &= \frac{r\binom{N-1}{n-1}}{\binom{N}{n}} \left[\sum_{j=0}^{n-1} j \frac{\binom{r-1}{j} \binom{(N-1)-(r-1)}{n-1-j}}{\binom{N-1}{n-1}} + \sum_{j=0}^{n-1} \frac{\binom{r-1}{j} \binom{(N-1)-(r-1)}{n-1-j}}{\binom{N-1}{n-1}} \right] \\
&= \frac{r\binom{N-1}{n-1}}{\binom{N}{n}} \left[\frac{(n-1)(r-1)}{(N-1)} + 1 \right] \\
&= r \frac{(N-1)!}{(n-1)!(N-n)!} \frac{n!(N-n)!}{N!} \left[\frac{(n-1)(r-1)+(N-1)}{(N-1)} \right] \\
&= \frac{rn}{N} \left[\frac{(n-1)(r-1)+(N-1)}{(N-1)} \right]
\end{aligned}
$$

(iii) Agora podemos calcular a variância desta variável aleatória

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

□

Exemplo 5.18. Pequenos motores são guardados em caixas de 50 unidades. Um inspetor de qualidade examina cada caixa, antes da posterior remessa, testando 5 motores. Se nenhum motor for defeituoso, a caixa é aceita. Se pelo menos um for defeituoso, todos os 50 motores são testados. Há 6 motores defeituosos numa caixa. Qual a probabilidade de que seja necessário examinar todos os motores dessa caixa?

Resolução: Seja X o número de motores defeituosos da amostra. Temos que $N = 50$, $r = 6$ e $n = 5$. Então,

$$
\begin{aligned}
\mathbb{P}(X \geqslant 1) &= 1 - \mathbb{P}(X < 1) = 1 - \mathbb{P}(X = 0) \\
&= 1 - \frac{\binom{6}{0}\binom{50-6}{5-0}}{\binom{50}{5}} = 1 - 0.5126 = 0.4874.
\end{aligned}
$$

□

Exemplo 5.19. De um baralho de 52 cartas, retiram-se 8 cartas ao acaso, sem reposição. Qual a probabilidade de que 4 sejam figuras?

Resolução: Seja X o número de figuras em 8 cartas. Temos que $N = 52$, $r = 12$ e $n = 8$. Então,

$$
\mathbb{P}(X = 4) = \frac{\binom{12}{4}\binom{52-12}{8-4}}{\binom{52}{8}} = 0.0601.
$$

□

Exemplo 5.20. Uma firma compra lâmpadas por centenas. Examina sempre uma amostra de 15 lâmpadas para verificar se estão boas. Se uma centena inclui 12 lâmpadas queimadas, qual a probabilidade de se escolher uma amostra com pelo menos uma lâmpada queimada?

Resolução: Seja X o número de lâmpadas queimadas na amostra. Temos que $N = 100$, $r = 12$ e $n = 15$. Então,

$$\begin{aligned}\mathbb{P}(X \geqslant 1) &= 1 - \mathbb{P}(X < 1) = 1 - \mathbb{P}(X = 0)\\ &= 1 - \frac{\binom{12}{0}\binom{100-12}{15-0}}{\binom{100}{15}} = 0.8747.\end{aligned}$$

□

Exemplo 5.21. Um caçador, após um dia de caça, verificou que matou 5 andorinhas e 2 aves de uma espécie rara, proibida de ser caçada. Como todos os espécimes tinham o mesmo tamanho, ele os colocou na mesma bolsa, pensando em dificultar o trabalho dos fiscais. No posto de fiscalização há dois fiscais, Manoel e Pedro, que adotam diferentes métodos de inspeção. Manoel retira três espécimes de cada bolsa dos caçadores. Pedro retira um espécime, classifica-o e o repõe na bolsa, retirando em seguida um segundo espécime. Em qualquer caso, o caçador é multado se é encontrado pelo menos um espécime proibido. Qual dos dois fiscais é mais favorável para o caçador em questão?

Resolução: Seja X = número de aves proibidas (sucessos) encontradas por um fiscal. No caso de Manoel, temos que $X \sim \mathrm{H}(7,3,2)$ e no caso do fiscal Pedro, $X \sim \mathrm{B}\left(2, \frac{2}{7}\right)$. Queremos calcular

$$\mathbb{P}(\text{multa}) = \mathbb{P}(X \geqslant 1) = 1 - \mathbb{P}(X = 0).$$

$$\text{Manoel}: \mathbb{P}(\text{multa}) = 1 - \mathbb{P}(X=0) = 1 - \frac{\binom{2}{0}\binom{5}{3}}{\binom{7}{3}} = 1 - \frac{2}{7} = \frac{5}{7} = \frac{35}{49}$$

$$\text{Pedro}: \mathbb{P}(\text{multa}) = 1 - \mathbb{P}(X=0) = 1 - \binom{2}{0}\left(\frac{2}{7}\right)^0\left(1-\frac{2}{7}\right)^{2-0} = 1 - \frac{25}{49} = \frac{24}{49}$$

Logo, a probabilidade de multa é maior no caso do fiscal Manoel, e, portanto, Pedro é o fiscal mais favorável para o caçador.

□

5.6 Modelo Binomial Negativo

Considere uma sequencia de ensaios de Bernoulli independentes com probabilidade de sucesso p e defina a variável aleatória X como o {número de ensaios de Bernoulli necessários até o r-ésimo sucesso}, $r \geqslant 1$.

Para definir os valores de X, precisamos notar que para ter r sucessos, são necessários no mínimo r repetições. Logo, os possíveis valores de X são $\Omega_X = \{r, r+1, r+2, \cdots\}$. O evento $X = k$ indica que foram necessários k ensaios de Bernoulli para obtermos r sucessos e $k - r$ fracasso. Pela definição da variável aleatória X, o último ensaio resultou em sucesso e os outros $r - 1$ sucessos podem estar em quaisquer das $k - 1$ repetições restantes conforme pode ser observado na Figura 5.8.

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: $S_1 \cdots S_{r-1} F_r \cdots F_{k-1} S_k$. A probabilidade de tal sequência é dada pelo produto das probabilidades, já que as repetições são independentes, isto é:

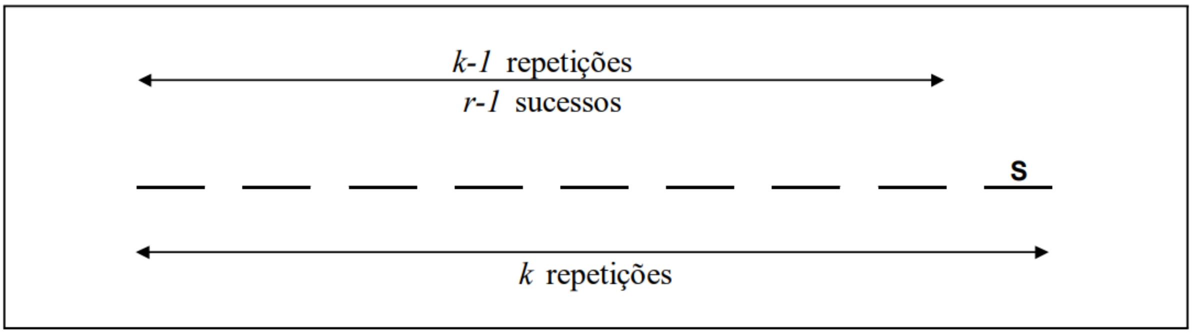


Figura 5.8: Ilustração do evento $\{X = k\}$ para a variável aleatória Binomial Negativa.

$$
\begin{aligned}
& \mathbb{P}(S_1 \cap S_2 \cap \cdots S_{r-1} \cap F_r \cap \cdots \cap F_{k-1} \cap S_k) \\
= \; & \mathbb{P}(S_1) \times \mathbb{P}(S_2) \times \cdots \mathbb{P}(S_{r-1}) \times \mathbb{P}(F_r) \times \cdots \times \mathbb{P}(F_{k-1}) \times \mathbb{P}(S_k) \\
= \; & \underbrace{p \times \cdots \times p}_{r \text{ sucesso}} \times \underbrace{(1-p) \times \cdots \times (1-p)}_{k-r \text{ fracassos}} \\
= \; & p^r (1-p)^{k-r}.
\end{aligned}
$$

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 $\mathbb{P}(X = k)$, é a soma das probabilidades, ou seja

$$\mathbb{P}(X = k) = p^r(1-p)^{k-r} + p^r(1-p)^{k-r} + \cdots + p^r(1-p)^{k-r},$$

onde o número de possíveis resultados é $\binom{k-1}{r-1}$. Seta forma temos a definição da Distribuição Binomial Negativa dada a seguir.

Definição 5.7. Considere uma sequencia de ensaios de Bernoulli independentes e defina X como o número de ensaios de Bernoulli necessários até o r-ésimo sucesso. A variável aleatória X segue o *modelo Binomial Negativo* com parâmetros r e p com $0 < p < 1$ e $r > 0$ e possui função massa de probabilidade dada por

$$p_X(x) = \mathbb{P}(X = x) = \binom{x-1}{r-1} p^r (1-p)^{x-r}, \quad x \geqslant r, \quad r \geqslant 1. \tag{5.9}$$

sendo $\binom{n}{x} = \frac{n!}{x!(n-x)!}$ o Binômio de Newton.

Notação: $X \sim \mathrm{BN}(r, p)$. [Seja X a variável aleatória que fornece o número de ensaios de Bernoulli necessários até o r-ésimo sucesso. A variável aleatória X possui distribuição Binomial Negativa com parâmetros (r, p).]

Corolário 5.7. *Seja $X \sim \mathrm{BN}(r, p)$, $0 < p < 1$. Então $p_X(\cdot)$ dada pela equação (5.9) satisfaz as propriedades de uma função massa de probabilidade, ou seja, é uma função massa de probabilidade.*

Demonstração. Vamos verificar se $p_X(\cdot)$ é função de probabilidade. Temos que $0 \leqslant p_X(x) \leqslant 1$, para todo $x \geqslant r$, com $r \geqslant 1$. Temos que verificar a segunda propriedade. Assim,

$$\sum_{x \geqslant r} p_X(x) = \sum_{x \geqslant r} \binom{x-1}{r-1} p^r (1-p)^{x-r}.$$

Fazendo a mudança de variável $k = x - r$, temos

$$\sum_{x \geqslant r} p_X(x) = \sum_{k \geqslant 0} \binom{k+r-1}{r-1} p^r (1-p)^k = p^r \sum_{k \geqslant 0} \binom{k+r-1}{r-1} (1-p)^k.$$

Como

$$\sum_{k \geqslant 0} \binom{k+b}{b} a^k = \frac{1}{(1-a)^{b+1}}, \qquad |a| < 1,$$

temos que

$$\sum_{x \geqslant r} p_X(x) = p^r \sum_{k \geqslant 0} \binom{k+r-1}{r-1} (1-p)^k = p^r \frac{1}{(1-(1-p))^{r-1+1}} = \frac{p^r}{p^r} = 1$$

Logo $p_X(\cdot)$ é função massa de probabilidade.

□

Resultado 5.6. Seja X uma variável aleatória discreta com distribuição $\mathrm{B}(n,p)$, então sua função acumulada de probabilidade é dada por

$$F_X(x) = \mathbb{P}(X \leqslant x) = \sum_{j=r}^{[x]} \mathbb{P}(X = j) = \begin{cases} 0, & \text{se } x < r; \\ \sum_{j=r}^{[x]} \binom{j-1}{r-1} p^r (1-p)^{j-r}, & \text{se } x \geqslant r; \end{cases}$$

onde $r \geqslant 1$ e $[x]$ é a parte inteira de x.

□

Proposição 5.7. *Seja* $X \sim \mathrm{BN}(r,p)$, $0 < p < 1$, *então*

(i) $\mathbb{E}(X) = \frac{r}{p}$.

(ii) $\mathbb{E}(X^2) = \frac{r^2 + r(1-p)}{p^2}$.

(iii) $\mathrm{Var}(X) = \frac{r(1-p)}{p^2}$.

Demonstração. Seja X uma variável aleatória discreta com distribuição $\mathrm{BN}(r,p)$, ou seja, sua função massa de probabilidade é dada por

$$p_X(x) = \mathbb{P}(X = x) = \binom{x-1}{r-1} p^r (1-p)^{x-r}, \quad x \geqslant r, \quad r \geqslant 1$$

onde $\binom{n}{x} = \frac{n!}{x!(n-x)!}$ é o Binômio de Newton.

(i) Iniciamos calculando $\mathbb{E}(X)$.

$$\mathbb{E}(X) = \sum_{x \geqslant r} x \mathbb{P}(X = x) = \sum_{x \geqslant r} x \binom{x-1}{r-1} p^r (1-p)^{x-r}.$$

Fazendo $j = x - r$, temos

$$\begin{aligned}
\mathbb{E}(X) &= \sum_{j\geqslant 0}(j+r)\binom{j+r-1}{r-1}p^r(1-p)^j \\
&= p^r\sum_{j\geqslant 0} j\binom{j+r-1}{r-1}(1-p)^j + rp^r\sum_{j\geqslant 0}\binom{j+r-1}{r-1}(1-p)^j \\
&= p^r\sum_{j\geqslant 0} j\binom{j+r-1}{r-1}(1-p)^j + \frac{rp^r}{[1-(1-p)]^{r-1+1}} \\
&= p^r\sum_{j\geqslant 0} j\binom{j+r-1}{r-1}(1-p)^j + r
\end{aligned} \tag{5.10}$$

pois

$$\sum_{k\geqslant 0}\binom{k+b}{b}a^k = \frac{1}{(1-a)^{b+1}}, \quad |a| < 1. \tag{5.11}$$

Agora, se $j = 0$, $j\binom{j+r-1}{r-1}(1-p)^j = 0$, então

$$\begin{aligned}
\sum_{j\geqslant 0} j\binom{j+r-1}{r-1}(1-p)^j &= \sum_{j\geqslant 1} j\binom{j+r-1}{r-1}(1-p)^j \\
&= \sum_{j\geqslant 1} j\frac{(j+r-1)!}{(r-1)!(j+r-1-(r-1))!}(1-p)^j \\
&= \sum_{j\geqslant 1} j\frac{(j+r-1)!}{(r-1)!j!}(1-p)^j \\
&= \sum_{j\geqslant 1} \frac{(j+r-1)!}{(r-1)!(j-1)!}(1-p)^j \\
&= (1-p)\sum_{j\geqslant 1} \frac{(j+r-1)!}{(r-1)!(j-1)!}(1-p)^{j-1}
\end{aligned}$$

Fazendo a mudança de variável $j - 1 = n$, obtemos

$$\begin{aligned}
\sum_{j\geqslant 0} j\binom{j+r-1}{r-1}(1-p)^j &= (1-p)\sum_{j\geqslant 1} \frac{(j+r-1)!}{(r-1)!(j-1)!}(1-p)^{j-1} \\
&= (1-p)\sum_{n\geqslant 0} \frac{(r+n)!}{(r-1)!n!}(1-p)^n.
\end{aligned} \tag{5.12}$$

Mas,

$$\frac{(r+n)!}{(r-1)!n!} = \frac{r(r+n)!}{r(r-1)!n!} = r\frac{(r+n)!}{r!n!} = r\frac{(r+n)!}{r!(r+n-r)!} = r\binom{r+n}{n}. \tag{5.13}$$

Substituindo a equação (5.13) na equação (5.12), obtemos

$$\begin{aligned}\sum_{j\geqslant 0} j\binom{j+r-1}{r-1}(1-p)^j &= (1-p)\sum_{n\geqslant 0}\frac{(r+n)!}{(r-1)!n!}(1-p)^n \\ &= (1-p)\sum_{n\geqslant 0} r\binom{r+n}{n}(1-p)^n \\ &= r(1-p)\sum_{n\geqslant 0}\binom{r+n}{n}(1-p)^n \\ &= r(1-p)\frac{1}{[1-(1-p)]^{r+1}} = \frac{r(1-p)}{p^{r+1}}. \end{aligned} \tag{5.14}$$

Onde o resultado da equação (5.14), segue aplicando o resultado da equação (5.11). Substituindo a equação (5.14) na equação (5.10), obtemos

$$\mathbb{E}(X) = p^r\sum_{j\geqslant 0} j\binom{j+r-1}{r-1}(1-p)^j + r = p^r \times \frac{r(1-p)}{p^{r+1}} + r = \frac{r(1-p)}{p} + r = \frac{r}{p}.$$

(ii) Agora vamos calcular $\mathbb{E}(X^2)$. Portanto,

$$\mathbb{E}(X^2) = \sum_{x\geqslant r} x^2\mathbb{P}(X=x) = \sum_{x\geqslant r} x^2\binom{x-1}{r-1}p^r(1-p)^{x-r}.$$

Fazendo $j = x - r$, temos

$$\begin{aligned}\mathbb{E}(X^2) &= \sum_{j\geqslant 0}(j+r)^2\binom{j+r-1}{r-1}p^r(1-p)^j \\ &= p^r\sum_{j\geqslant 0}(r^2+2jr+j^2)\binom{j+r-1}{r-1}(1-p)^j \\ &= r^2p^r\sum_{j\geqslant 0}\binom{j+r-1}{r-1}(1-p)^j + 2rp^r\sum_{j\geqslant 0} j\binom{j+r-1}{r-1}(1-p)^j \\ &\quad + p^r\sum_{j\geqslant 0} j^2\binom{j+r-1}{r-1}(1-p)^j \end{aligned} \tag{5.15}$$

Vamos analisar cada um dos somatórios da equação (5.15) individualmente. Então:

(a) Pelo resultado da equação (5.11), temos que

$$\sum_{j\geqslant 0}\binom{j+r-1}{r-1}(1-p)^j = \frac{1}{[1-(1-p)]^{r-1+1}} = \frac{1}{p^r}.$$

(b) Pela equação (5.14), temos que

$$\sum_{j\geqslant 0} j\binom{j+r-1}{r-1}(1-p)^j = \frac{r(1-p)}{p^{r+1}}.$$

(c) Como próximo passo iremos analisar o terceiro somatório. Se $j = 0$, a primeira parcela do somatório é zero, logo

$$\begin{aligned}
\sum_{j\geqslant 0} j^2 \binom{j+r-1}{r-1}(1-p)^j &= \sum_{j\geqslant 1} j^2 \binom{j+r-1}{r-1}(1-p)^j \\
&= \sum_{j\geqslant 1} j^2 \frac{(j+r-1)!}{(r-1)!(j+r-1-(r-1))!}(1-p)^j \\
&= \sum_{j\geqslant 1} j^2 \frac{(j+r-1)!}{(r-1)!j(j-1)!}(1-p)^j \\
&= (1-p)\sum_{j\geqslant 1} j \frac{(j+r-1)!}{(r-1)!(j-1)!}(1-p)^{j-1}.
\end{aligned}$$

Fazendo a mudança de variável $n = j - 1$ na equação acima temos

$$\sum_{j\geqslant 0} j^2 \binom{j+r-1}{r-1}(1-p)^j = (1-p)\sum_{n\geqslant 0}(n+1)\frac{(r+n)!}{(r-1)!n!}(1-p)^n.$$

Mas,

$$\frac{(r+n)!}{(r-1)!n!} = \frac{r(r+n)!}{r(r-1)!n!} = r\frac{(r+n)!}{r!n!} = r\binom{r+n}{r},$$

então,

$$\begin{aligned}
\sum_{j\geqslant 0} j^2 \binom{j+r-1}{r-1}(1-p)^j &= (1-p)\sum_{n\geqslant 0}(n+1)\frac{(r+n)!}{(r-1)!n!}(1-p)^n \\
&= r(1-p)\sum_{n\geqslant 0}(n+1)\binom{r+n}{r}(1-p)^n \\
&= r(1-p)\left[\sum_{n\geqslant 0} n\binom{r+n}{r}(1-p)^n + \sum_{n\geqslant 0}\binom{r+n}{r}(1-p)^n\right]
\end{aligned} \tag{5.16}$$

- Pela equação (5.14), substituindo r por $r+1$, obtemos

$$\sum_{n\geqslant 0} n\binom{r+n}{r}(1-p)^n = \frac{(r+1)(1-p)}{[1-(1-p)]^{r+1+1}} = \frac{(r+1)(1-p)}{[1-(1-p)]^{r+2}} = \frac{(r+1)(1-p)}{p^{r+2}}.$$

- Pelo resultado da equação (5.11), temos que

$$\sum_{n\geqslant 0}\binom{r+n}{r}(1-p)^n = \frac{1}{[1-(1-p)]^{r+1}} = \frac{1}{p^{r+1}}.$$

Voltando a equação (5.16), então

$$
\begin{aligned}
\sum_{j\geqslant 0} j^2 \binom{j+r-1}{r-1}(1-p)^j &= r(1-p)\left[\sum_{n\geqslant 0} n\binom{r+n}{r}(1-p)^n + \sum_{n\geqslant 0}\binom{r+n}{r}(1-p)^n\right]\\
&= r(1-p)\left[\frac{(r+1)(1-p)}{p^{r+2}} + \frac{1}{p^{r+1}}\right]\\
&= \left[\frac{(1-p)^2 r(r+1)}{p^{r+2}} + \frac{(1-p)r}{p^{r+1}}\right]\\
&= \left[\frac{(1-2p+p^2)(r^2+r)}{p^{r+2}} + \frac{(1-p)r}{p^{r+1}}\right]\\
&= \frac{(1-2p+p^2)(r^2+r)+rp(1-p)}{p^{r+2}}\\
&= \frac{r^2-2pr^2+p^2r^2+r-rp}{p^{r+2}}\\
&= \frac{r^2(1-2p+p^2)+r(1-p)}{p^{r+2}} \qquad (5.17)
\end{aligned}
$$

Substituindo os itens (a) a (c) na equação (5.15), temos que

$$
\begin{aligned}
\mathbb{E}(X^2) &= r^2p^r\sum_{j\geqslant 0}\binom{j+r-1}{r-1}(1-p)^j + 2rp^r\sum_{j\geqslant 0} j\binom{j+r-1}{r-1}(1-p)^j\\
&\quad + p^r\sum_{j\geqslant 0} j^2\binom{j+r-1}{r-1}(1-p)^j\\
&= r^2p^r\times\frac{1}{p^r} + 2rp^r\times\frac{r(1-p)}{p^{r+1}} + p^r\times\frac{r^2(1-2p+p^2)+r(1-p)}{p^{r+2}}\\
&= r^2 + \frac{2r^2(1-p)}{p} + \frac{r^2(1-2p+p^2)+r(1-p)}{p^2}\\
&= \frac{r^2p^2+2r^2p(1-p)+r^2(1-2p+p^2)+r(1-p)}{p^2}\\
&= \frac{r^2p^2+2r^2p-2r^2p^2+r^2-2r^2p+r^2p^2+r-rp}{p^2}\\
&= \frac{r^2+r-rp}{p^2} = \frac{r^2+r(1-p)}{p^2}
\end{aligned}
$$

(iii) Agora podemos calcular a variância desta variável aleatória

$$\mathrm{Var}(X) = \mathbb{E}(X^2) - [\mathbb{E}(X)]^2 = \frac{r^2+r(1-p)}{p^2} - \left[\frac{r}{p}\right]^2 = \frac{r(1-p)}{p^2}.$$

□

Maiores detalhes ver https://bit.ly/32TGb5o. Os exemplos a seguir foram retirados de https://bit.ly/2D6e348.

Exemplo 5.22. Suponha que, para se ganhar um jogo de dados seja necessário obter 3 vezes a face voltada para cima do dado com o número de 1. Sendo que o número de lançamentos devem ser 6 e devemos obter a face 1 voltada para cima pela terceira vez no sexto lançamento. Supondo que o dado seja honesto, qual será a probabilidade de vencermos o jogo.

Solução: De fato, como o dado é honesto temos que a probabilidade de sair a face 1 é $1/6$, portanto temos que $X \sim \mathrm{BN}(3, 1/6)$ e queremos calcular

$$\mathbb{P}(X=6) = \begin{pmatrix} x-1 \\ k-1 \end{pmatrix} p^k(1-p)^{x-k} = \begin{pmatrix} 6-1 \\ 3-1 \end{pmatrix} \frac{1}{6}^3 \left(1-\frac{1}{6}\right)^3 = \frac{1250}{46656} \approx 0,0267918.$$

□

Exemplo 5.23. Suponha que um vendedor de automóveis tem, além de seu salário, uma bonificação de 200 reais por cada automóvel que ele venda. Suponha também que ele necessite vender, no mínimo, 5 automóveis por mês para que ele não seja despedido. Qual a probabilidade do vendedor ser despedido no mês de fevereiro, dado que ele trabalha todos os dias, inclusive finais de semana e feriados. Suponha que ele consiga vender no máximo 1 automóvel por dia com probabilidade de $0,2$.

Solução: Para que ele seja despedido no mês de fevereiro ele precisa vender, no máximo, 4 automóveis durante todo o mês. Desta forma, considere X_i a variável aleatória que representa a quantidade de dias necessários para vender i carros. Neste caso, temos que $X_i \sim \mathrm{BN}(0,2; i)$ e então

então

$$\mathbb{P}(X_i = 28) = \begin{pmatrix} 28-1 \\ i-1 \end{pmatrix} 0,2^i(1-0,2)^{28-i} \text{ para } i = 0,1,2,3,4.$$

Portanto,

$$\mathbb{P}(X_0 = 28) = (0,8)^{28} = 0,001934.$$

$$\mathbb{P}(X_1 = 28) = \begin{pmatrix} 27 \\ 0 \end{pmatrix} 0,2(0,8)^{27} \approx 0,0004835.$$

$$\mathbb{P}(X_2 = 28) = \begin{pmatrix} 27 \\ 1 \end{pmatrix} 0,2^2(0,8)^{26} \approx 0,0032640.$$

$$\mathbb{P}(X_3 = 28) = \begin{pmatrix} 27 \\ 2 \end{pmatrix} 0,2^3(0,8)^{25} \approx 0,010608.$$

$$\mathbb{P}(X_4 = 28) = \begin{pmatrix} 27 \\ 3 \end{pmatrix} 0,2^4(0,8)^{24} \approx 0,0221006.$$

Portanto, a probabilidade dele ser mandado embora no mês de fevereiro é dada por

$$\mathbb{P}(\text{Ser demitido}) = \sum_{i=0}^{4} \mathbb{P}(X_i = 28) \approx 0,0364561.$$

□

Exemplo 5.24. Suponha que em uma fábrica produz resistência para chuveiros, com uma taxa de defeitos de 2%. Qual a probabilidade de que em uma inspeção de 10 resistências se tenha 3 resistências defeituosas sendo que a terceira defeituosa seja exatamente a décima inspecionada.

Solução: Seja X o número de resistências inspecionadas até que encontremos a terceira defeituosa. Então temos que $X \sim \mathrm{BN}(0,02; 3)$ e

$$\mathbb{P}(X = 10) = \binom{9}{2} (0,02)^3 (0,98)^7 \approx 0,00025.$$

□

5.6.1. Forma Alternativa da Distribuição Binomial Negativa

Assim como a Distribuição Geométrica, é possível definir a distribuição Binomial Negativa de forma alternativa.

Definição 5.8. Seja Y a contagem do número de fracassos antes de r-ésimo sucesso. Isso significa que a última repetição corresponde ao r-ésimo sucesso e antes dele temos y fracassos e $r-1$ sucessos. Logo, antes do último sucesso, temos um total de $y+r-1$ repetições. Veja Figura 5.9. Neste caso, $Y = g(X) = X - r$ é uma função de X. A função massa de probabilidade de Y é dada por

$$\begin{aligned} p_Y(y) &= \mathbb{P}(Y = y) = \mathbb{P}(g(X) = y) = \mathbb{P}(X - r = y) = \mathbb{P}(X = y + r) = p_X(y + r) \\ &= \binom{y+r-1}{r-1} p^r (1-p)^{y+r-r} \\ &= \binom{y+r-1}{r-1} p^r (1-p)^y, \end{aligned}$$

ou seja,

$$p_Y(y) = \binom{y+r-1}{r-1} p^r (1-p)^y, \quad y = 0, 1, 2, \cdots, \quad r \geqslant 1, \tag{5.18}$$

onde $\binom{n}{x} = \frac{n!}{x!(n-x)!}$ é o Binômio de Newton.

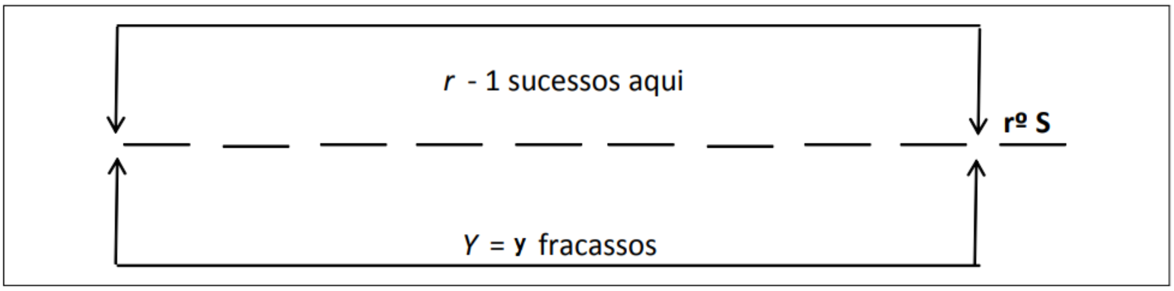


Figura 5.9: Ilustração do evento $\{Y = y\}$ para a variável aleatória Binomial Negativa deslocada.

Notação: $Y \sim \text{BN}(r, p)$. [Seja Y a contagem do número de fracassos antes de r-ésimo sucesso. A variável aleatória Y possui distribuição Binomial Negativa com parâmetros (r, p).]

Corolário 5.8. *Seja $Y \sim \text{BN}(r, p)$, $0 < p < 1$. Então $p_X(\cdot)$ dada pela equação (5.18) satisfaz as propriedades de uma função massa de probabilidade, ou seja, é uma função massa de probabilidade.*

Demonstração. Vamos verificar se $p_Y(\cdot)$ é função de probabilidade. Temos que $0 \leqslant p_Y(y) \leqslant 1$, para todo $y \geqslant 0$, com $r \geqslant 1$. Como $Y = X - r$, Y é uma função de X, na Definição 5.8 construímos sua função massa de probabilidade.

Logo $p_X(\cdot)$ é função massa de probabilidade.

□

Resultado 5.7. Seja Y uma variável aleatória discreta com distribuição $\mathrm{B}(n, p)$, então sua função acumulada de probabilidade é dada por

$$F_Y(y) = \mathbb{P}(Y \leqslant y) = \sum_{j=0}^{[y]} \mathbb{P}(Y = j) = \begin{cases} 0, & \text{se } Y < 0; \\ \sum_{j=0}^{[y]} \binom{y+r-1}{r-1} p^r (1-p)^y, & \text{se } y \geqslant 0; \end{cases}$$

onde $r \geqslant 1$ e $[y]$ é a parte inteira de y.

□

Proposição 5.8. *Seja* $Y \sim \mathrm{BN}(r, p)$, $0 < p < 1$, *então*

(i) $\mathbb{E}(X) = \frac{r(1-p)}{p}$.

(ii) $\mathbb{E}(X^2) = \frac{r^2(1-p)^2 + r(1-p)}{p^2}$.

(iii) $\mathrm{Var}(X) = \frac{r(1-p)}{p^2}$.

Demonstração. Seja $Y \sim \mathrm{BN}(r, p)$, $0 < p < 1$, então $Y = g(X) = X - r$, onde X é dada pela Definição 5.7). Pela Proposição 5.7, temos que

(i) $\mathbb{E}(X) = \frac{r}{p}$.

(ii) $\mathbb{E}(X^2) = \frac{r^2 + r(1-p)}{p^2}$.

(iii) $\mathrm{Var}(X) = \frac{r(1-p)}{p^2}$.

Como $Y = g(X) = X - r$, obtemos

(i) $\mathbb{E}(Y) = \mathbb{E}(X - r) = \mathbb{E}(X) - r = \frac{r}{p} - r = \frac{r - rp}{p} = \frac{r(1-p)}{p}$.

(ii) $\mathbb{E}(Y^2) = \mathbb{E}[(X - r)^2] = \mathbb{E}(X^2) - 2r\mathbb{E}(X) + r^2 = \frac{r^2 + r(1-p)}{p^2} - 2r\frac{r}{p} + r^2$

$$\mathbb{E}(Y^2) = \frac{r^2 + r(1-p) - 2r^2p + r^2p^2}{p^2} = \frac{r^2(1 - 2p + p^2) + r(1-p)}{p^2} = \frac{r^2(1-p)^2 + r(1-p)}{p^2}.$$

(iii) $\mathrm{Var}(Y) = \mathrm{Var}(X - r) = \mathrm{Var}(X) = \frac{r(1-p)}{p^2}$.

□

Os exemplos a seguir forma retirados de https://bit.ly/39pDN7Q.

Exemplo 5.25. Lançamos repetidas vezes uma moeda. Seja Y o número de caras até que consigamos sete coroas. Qual é a probabilidade de que o número de caras seja igual a cinco até que consigamos as sete coroas [$\mathbb{P}(Y = 5)$]?

Solução: Seja Y o número de caras até que consigamos sete coroas, com $p = 0.5$. Portanto $Y \sim \text{BN}(7, 0.5)$. Assim temos que

$$\mathbb{P}(Y = 5) = \binom{7+5-1}{7-1} 0.5^7(1-0.5)^5 = 0.112793.$$

□

5.7 Relação entre Binomial e Poisson

Muitas vezes no uso da distribuição Binomial, acontece que n é muito grande ($n \to \infty$) e p é muito pequeno. Nestes casos, podemos utilizar uma aproximação da Distribuição Binomial pela Distribuição Poisson.

Corolário 5.9. *Seja $X \sim \text{B}(n, p)$ e considere*

$$\begin{cases} n \to \infty \\ p \to 0 \\ 0 < np \leqslant 10. \end{cases}$$

Sob estas condições, $\mathbb{P}(X = x)$ pode ser aproximada por $\mathbb{P}(Y = x)$, onde $Y \sim \text{P}(\lambda)$, com $\lambda = np$, ou seja,

$$\mathbb{P}(X = x) = \binom{n}{x} p^x (1-p)^{n-x} \simeq \frac{e^{-\lambda}\lambda^x}{x!} = \mathbb{P}(Y = x),$$

ou equivalentemente

$$\lim_{n\to\infty} \mathbb{P}(X = x) = \lim_{n\to\infty} \binom{n}{x} p^x (1-p)^{n-x} = \frac{e^{-\lambda}\lambda^x}{x!} = \mathbb{P}(Y = x)$$

onde $\binom{n}{x} = \frac{n!}{x!(n-x)!}$ é o Binômio de Newton.

Demonstração. Considere as variáveis aleatórias $X \sim B(n, p)$ e $Y \sim P(\lambda)$. Então,

$$\begin{aligned} \mathbb{P}(X = x) &= \binom{n}{x} p^x (1-p)^{n-x} = \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x} \\ &= \frac{n(n-1)(n-2)\cdots(n-x+1)(n-x)!}{x!(n-x)!} p^x (1-p)^{n-x} \\ &= \frac{n(n-1)(n-2)\cdots(n-x+1)}{x!} p^x (1-p)^{n-x}. \end{aligned}$$

Por hipótese, temos que $\lambda = np$, ou seja, $p = \frac{\lambda}{n}$. Assim,

$$\begin{aligned}
\mathbb{P}(X = x) &= \frac{n(n-1)(n-2)\cdots(n-x+1)}{x!}\left(\frac{\lambda}{n}\right)^x\left(1-\frac{\lambda}{n}\right)^{n-x} \\
&= \frac{1}{x!}\left[\frac{n}{n}\frac{(n-1)}{n}\frac{(n-2)}{n}\cdots\frac{(n-x+1)}{n}\right]\lambda^x\left(1-\frac{\lambda}{n}\right)^{n-x} \\
&= \frac{\lambda^x}{x!}\left[\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\cdots\left(1-\frac{x+1}{n}\right)\right]\left(1-\frac{\lambda}{n}\right)^{n}\left(1-\frac{\lambda}{n}\right)^{-x}.
\end{aligned}$$

Aplicando o limite quando $n \to \infty$, temos

$$\begin{aligned}
\lim_{n\to\infty} \mathbb{P}(X = x) &= \lim_{n\to\infty} \frac{\lambda^x}{x!}\left[\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\cdots\left(1-\frac{x+1}{n}\right)\right]\left(1-\frac{\lambda}{n}\right)^{n}\left(1-\frac{\lambda}{n}\right)^{-x} \\
&= \frac{\lambda^x}{x!}\lim_{n\to\infty}\left(1-\frac{\lambda}{n}\right)^{n},
\end{aligned}$$

pois, $\lim_{n\to\infty}\left(1-\frac{\lambda}{n}\right)^{-x} = 1$ e $\lim_{n\to\infty}\left(1-\frac{j}{n}\right) = 1$, para $j = 1, 2, \cdots, x+1$. Como

$$\lim_{n\to\infty}\left(1-\frac{\lambda}{n}\right)^{n} = e^{-\lambda},$$

temos que

$$\lim_{n\to\infty} \mathbb{P}(X = x) = \frac{\lambda^x}{x!}\lim_{n\to\infty}\left(1-\frac{\lambda}{n}\right)^{n} = \frac{e^{-\lambda}\lambda^x}{x!}.$$

□

Exemplo 5.26. Seja $X \sim \mathrm{B}(200, 0.01)$. Calcular $\mathbb{P}(X = 10)$ utilizando a distribuição Binomial e a aproximação pela Poisson.

Solução: Seja $X \sim \mathrm{B}(200, 0.01)$. Então, sob as condições

$$\begin{cases} n \to \infty \\ p \to 0 \\ 0 < np = 200 \times 0.01 = 2 \leqslant 10, \end{cases}$$

temos

i) Probabilidade Exata:

$$\begin{aligned}
\mathbb{P}(X = 10) &= \binom{200}{10} 0.01^{10}(1-0.01)^{200-10} \\
&= \text{dbinom(x=10, size=200, prob=0.01, log = FALSE)} \\
&= 0,000033.
\end{aligned}$$

ii) Probabilidade aproximada pela Distribuição Poisson:

$$\begin{aligned}
\mathbb{P}(X = 10) &= \frac{e^{-2}2^{10}}{10!} \\
&= \text{dpois(x=10, lambda=2, log = FALSE)} \\
&= 0,000038.
\end{aligned}$$

□

Exemplo 5.27. Seja $X \sim B(200, 0.01)$. Calcular $\mathbb{P}(X \leqslant 2)$ utilizando a distribuição Binomial e a aproximação pela Poisson.

Solução: Seja $X \sim B(200, 0.01)$. Então, sob as condições

$$\begin{cases} n \to \infty \\ p \to 0 \\ 0 < np = 200 \times 0.01 = 2 \leqslant 10, \end{cases}$$

temos

(i) Probabilidade Exata:

$$\begin{aligned} \mathbb{P}(X \leqslant 2) &= \mathbb{P}(X=0) + \mathbb{P}(X=1) + \mathbb{P}(X=2) \\ &= \binom{200}{0} 0.01^0 (1-0.01)^{200-0} + \binom{200}{1} 0.01^1 (1-0.01)^{200-1} + \binom{200}{2} 0.01^2 (1-0.01)^{200-2} \\ &= \text{pbinom(q=2, size=200, prob=0.01, lower.tail = TRUE, log.p = FALSE)} \\ &= 0,6766787. \end{aligned}$$

(ii) Probabilidade aproximada pela Distribuição Poisson:

$$\begin{aligned} \mathbb{P}(X \leqslant 2) &= \mathbb{P}(X=0) + \mathbb{P}(X=1) + \mathbb{P}(X=2) \\ &= \frac{e^{-2}2^0}{0!} + \frac{e^{-2}2^1}{1!} + \frac{e^{-2}2^2}{2!} \\ &= \text{ppois(q=2, lambda=2, lower.tail = TRUE, log.p = FALSE)} \\ &= 0,6766764. \end{aligned}$$

□

5.8 Binomial versus Hipergeométrica

Considere, por exemplo, uma urna com r bolas verdes e $N-r$ bolas brancas. Desta urna extrai-se uma amostra de n bolas e variável aleatória de interesse é X= número de bolas verdes extraídas na amostra.

$\rightarrow$ Se a extração é feita *com reposição*, então X segue distribuição Binomial: $X \sim \mathrm{B}\left(n, \frac{r}{N}\right)$.

$\rightarrow$ Se a extração é feita *sem reposição*, então X segue distribuição Hipergeométrica: $X \sim \mathrm{H}\left(N, r, n\right)$.

Exemplo 5.28. Assim, se $N = 16$, $r = 6$ e $n = 5$, obtenha a função massa de probabilidade de X considerando:

(i) amostragem com reposição;

(ii) amostragem sem reposição;

Solução:

(i) As retiradas são independentes. Sendo $n = 5$ retiradas. Considerando amostragem com reposição, a probabilidade de sucesso é $p = \frac{r}{N} = \frac{6}{16}$. Neste caso, $X \sim \mathrm{B}\left(n, \frac{6}{16}\right)$, ou seja,

$$\mathbb{P}(X = x) = \binom{5}{x}\left(\frac{6}{16}\right)^x\left(1 - \frac{6}{16}\right)^{5-x}, \quad x = 0, 1, \cdots, 5,$$

ou seja,

x	0	1	2	3	4	5
$p_X(x)$	0.095367432	0.286102295	0.343322754	0.205993652	0.061798096	0.007415771

(ii) As retiradas não são independentes. Sendo $n = 5$ retiradas. Com a amostragem sem reposição a relação $\frac{r}{N}$ se altera a cada retirada. Assim, temos a distribuição Hipergeométrica e as probabilidades são dadas por

$$\mathbb{P}(X = x) = \frac{\binom{r}{x}\binom{N-r}{n-x}}{\binom{N}{n}} = \frac{\binom{6}{x}\binom{16-6}{5-x}}{\binom{16}{5}}, \quad x = 0, 1, \cdots, 5,$$

ou seja,

x	0	1	2	3	4	5
$p_X(x)$	0.057692308	0.288461538	0.412087912	0.206043956	0.034340659	0.001373626

□

5.9 Relação entre Binomial e Hipergeométrica

Se a população for muito grande ($N \to \infty$), podemos aproximar a Distribuição Hipergeométricas utilizando a distribuição Binomial.

Assim para a uma variável aleatória $X \sim \mathrm{H}(N, r, n)$, tem-se que

$$\mathbb{P}(X = x) = \frac{\binom{r}{x}\binom{N-r}{n-x}}{\binom{N}{n}} \simeq \binom{n}{x} p^x (1-p)^{n-x}, \quad \text{onde} \quad p = \frac{r}{N}.$$

Desta forma, a distribuição $\mathrm{H}(N, r, n)$ possui a distribuição $\mathrm{B}\left(n, \frac{r}{N}\right)$ como limite, isto é, quando $N \to \infty$. Em geral, a aproximação é boa para $\frac{n}{N} \leqslant 0.1$.

A ideia da aproximação da distribuição Hipergeométrica pela distribuição Binomial pode ser vista de duas formas:

→ Pela facilidade de cálculo da distribuição Binomial.

→ Para populações muito grandes ($N \to \infty$), em amostragem, apesar de obter amostras sem reposição, é possível desenvolver o raciocínio supondo amostras independentes.

Exemplo 5.29. Uma firma compra lâmpadas por centenas. Examina sempre uma amostra de 10 lâmpadas para verificar se estão em boas condições. Se uma centena inclui 12 lâmpadas com defeito, qual a probabilidade de se escolher uma amostra com pelo menos uma lâmpada com defeito?

Solução: Temos que $N = 100$, $n = 10$ e $r = 12$. Como X é o número de lâmpadas com defeito na amostra, então $x \in \{0, 1, 2, \cdots, 10\}$.

Usa-se a distribuição Hipergeométrica pois as retiradas/amostragens não são independentes. Contudo é possível aproximar as probabilidades pela distribuição Binomial, pois N é bastante grande em relação ao tamanho da amostra extraída, ou seja, $\frac{n}{N} \leqslant 0.1$

$\rightarrow$ Probabilidade Exata: $X \sim \mathrm{H}\,(N = 100, r = 12, n = 10)$

$$\mathbb{P}(X \geqslant 1) = 1 - \mathbb{P}(X = 0) = 1 - \left[\frac{\binom{12}{0}\binom{100-12}{10-0}}{\binom{100}{10}}\right] = 0,7392497.$$

$\rightarrow$ Probabilidade Aproximada: $X \sim \mathrm{B}\,(n, p)$, com $n = 10$ e $\frac{r}{N} = \frac{12}{100} = 0,12$

$$\mathbb{P}(X \geqslant 1) = 1 - \mathbb{P}(X = 0) = 1 - \left[\binom{10}{0}(0,12)^0(1 - 0.12)^{10-0}\right] \simeq 0,721499.$$

$\square$

Capítulo 6

Principais Modelos Contínuas

Neste capítulo, apresentamos os principais modelos para variáveis aleatórias contínuas. Os diversos modelos serão caracterizados pela sua função densidade de probabilidade e, em vários casos, apresentamos a função distribuição, valor esperado e variância.

Neste capítulo são apresentadas algumas distribuições contínuas. Mais distribuições contínuas podem ser encontradas em https://bit.ly/3kfvEYK.

Iniciamos com o modelo Uniforme Contínuo.

6.1 Modelo Uniforme

Definição 6.1. Diremos que uma variável aleatória X segue o *modelo Uniforme*, no intervalo $[a, b] \in \mathbb{R}$, $a < b$, finito, se sua função densidade for dada por

$$f_X(x) = \frac{1}{b-a} I_{[a,b]}(x). \tag{6.1}$$

Notação: $X \sim \mathrm{U}[a, b]$, $a < b$.

Corolário 6.1. *Seja $X \sim U[a, b]$, $a < b$. Então $f_X(\cdot)$ dada pela equação (6.1) satisfaz as propriedades de uma função densidade de probabilidade, ou seja, é uma função densidade de probabilidade.*

Demonstração. Seja $f_X(\cdot)$ dada pela equação (6.1). Então $f_X(\cdot)$ satisfaz as condições para ser densidade. Temos que $f_X(\cdot)$ é não negativa e

$$\int_{-\infty}^{\infty} f_X(x)dx = \int_a^b \frac{1}{b-a}dx = 1.$$

Portanto. $f_X(\cdot)$ é uma função densidade de probabilidade.

$\square$

Resultado 6.1. Seja $X \sim \mathrm{U}[a,b]$, $a < b$. Então a função de distribuição do modelo uniforme em $[a,b]$ é dada por

$$F_X(x) = \begin{cases} 0, & \text{se } x < a; \\ \frac{x-a}{b-a}, & \text{se } a \leqslant x < b; \\ 1, & \text{se } x \geqslant b. \end{cases}$$

Demonstração. Para $x < a$, temos

$$F_X(x) = \int_{-\infty}^{x} f_X(y)dy = \int_{-\infty}^{x} 0dy = 0.$$

Para $a \leqslant x < b$, temos

$$F_X(x) = \int_{-\infty}^{x} f_X(y)dy = \int_{a}^{x} \frac{1}{b-a}dy = \frac{1}{b-a}\, y\Big|_a^x = \frac{x-a}{b-a}.$$

Para $x \geqslant b$, temos

$$F_X(x) = \int_{-\infty}^{x} f_X(y)dy = \int_{a}^{b} \frac{1}{b-a}dy = 1.$$

Portanto a Função de distribuição acumulada é dada por

$$F_X(x) = \begin{cases} 0, & \text{se } x < a; \\ \frac{x-a}{b-a}, & \text{se } a \leqslant x < b; \\ 1, & \text{se } x \geqslant b. \end{cases}$$

□

Proposição 6.1. *Seja* $X \sim \mathrm{U}[a,b]$, *então*

(i) $\mathbb{E}(X) = \frac{a+b}{2}$.

(ii) $\mathbb{E}(X^2) = \frac{b^2+ab+a^2}{3}$.

(iii) $\mathrm{Var}(X) = \frac{(b-a)^2}{12}$.

Demonstração. Seja $X \sim U[a,b]$.

(i) Cálculo de $\mathbb{E}(X)$.

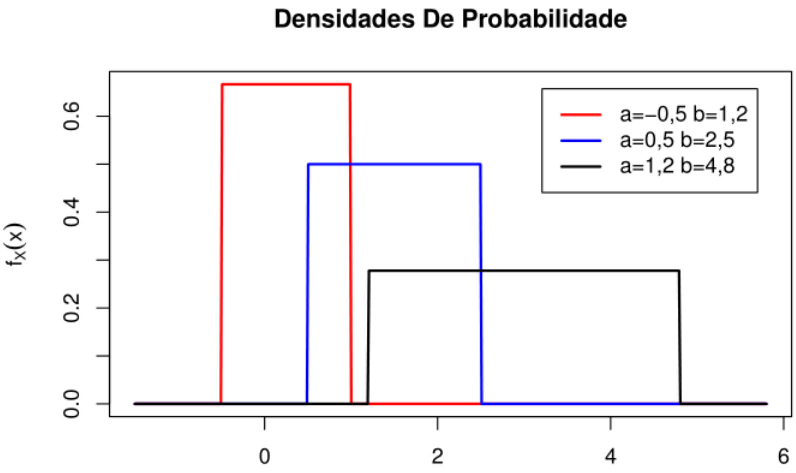


Figura 6.1: Distribuição U(a, b): função densidade de probabilidade.

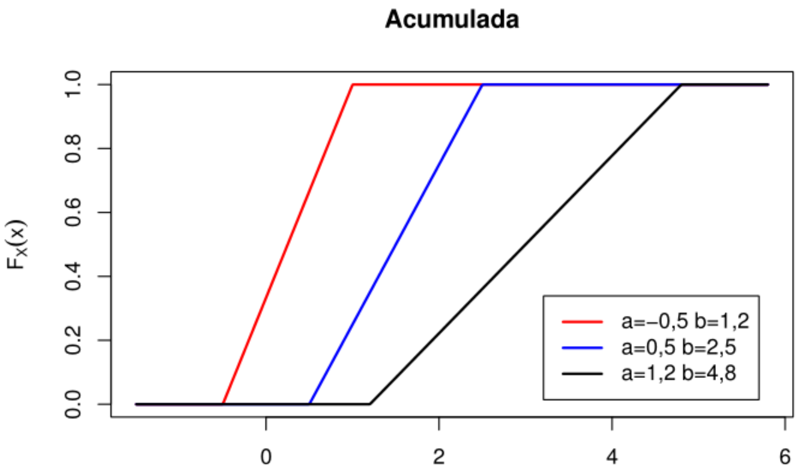


Figura 6.2: Distribuição U(a, b): função de distribuição acumulada.

$$
\begin{aligned}
\mathbb{E}(X) &= \int_{-\infty}^{\infty} x \times f_X(x)dx = \int_a^b x \times \frac{1}{b-a}dx = \frac{1}{b-a}\frac{x^2}{2}\Bigg|_a^b \\
&= \frac{1}{b-a}\left[\frac{x^2}{2}\right|_a^b = \frac{1}{b-a} \times \frac{b^2-a^2}{2} = \frac{1}{b-a} \times \frac{b^2-a^2}{2} \\
&= \frac{1}{b-a} \times \frac{(b-a)\times(b+a)}{2} = \frac{b+a}{2}
\end{aligned}
$$

(ii) Cálculo de $\mathbb{E}(X^2)$.

$$\begin{aligned}
\mathbb{E}(X^2) &= \int_{-\infty}^{\infty} x^2 \times f_X(x)dx = \int_a^b x^2 \times \frac{1}{b-a}dx = \frac{1}{b-a}\left[\frac{x^3}{3}\right|_a^b \\
&= \frac{1}{b-a} \times \frac{b^3-a^3}{3} = \frac{1}{b-a} \times \frac{(b-a)\times(b^2+ab+a^2)}{3} \\
&= \frac{b^2+ab+a^2}{3}
\end{aligned}$$

(iii) Cálculo de $\mathrm{Var}(X)$.

$$\begin{aligned}
\mathrm{Var}(X) &= \mathbb{E}(X^2) - [\mathbb{E}(X)]^2 = \frac{b^2+ab+a^2}{3} - \left(\frac{b+a}{2}\right)^2 \\
&= \frac{b^2+ab+a^2}{3} - \frac{b^2+2ab+a^2}{4} = \frac{4b^2+4ab+4a^2}{12} - \frac{3b^2+6ab+3a^2}{12} \\
&= \frac{b^2-2ab+a^2}{12} = \frac{(b-a)^2}{12}.
\end{aligned}$$

□

Exemplo 6.1. Um ponto é escolhido ao acaso no intervalo $[0, 2]$. Qual a probabilidade de que esteja ente 1 e 1.5?

Solução:

$$f_X(x) = \begin{cases} \frac{1}{2}, & \text{se } 0 \leqslant x \leqslant 2; \\ 0, & \text{c.c.} \end{cases}$$

$$\begin{aligned}
\mathbb{P}(1 \leqslant X \leqslant 1.5) &= \int_1^{1.5} \frac{1}{2}dx = \frac{x}{2}\Big|_1^{1.5} \\
&= \text{punif(q=1.5, min = 0, max = 2) − punif(q=1.0, min = 0, max = 2)} \\
&= \frac{1}{4}.
\end{aligned}$$

□

Exemplo 6.2. A dureza H de uma peça de aço pode ser pensada como uma variável aleatória com distribuição uniforme no intervalo $[50, 70]$ da escala de Rockwell. Calcular a probabilidade de que uma peça tenha dureza entre 55 e 60.

Solução:

$$f_X(x) = \begin{cases} \frac{1}{20}, & \text{se } 50 \leqslant x \leqslant 70; \\ 0, & \text{c.c.} \end{cases}$$

$$\begin{aligned}
\mathbb{P}(55 \leqslant X \leqslant 60) &= \int_{55}^{60} \frac{1}{20}dx = \frac{x}{20}\Big|_{55}^{60} \\
&= \text{punif(q=60, min = 50, max = 70) − punif(q=55, min = 50, max = 70))} \\
&= \frac{1}{4}.
\end{aligned}$$

□

Exemplo 6.3. A ocorrência de panes em qualquer ponto de uma rede telefônica de 7 km foi modelada por uma distribuição Uniforme no intervalo $[0,7]$. Qual é a probabilidade de que uma pane venha a ocorrer nos primeiros 800 metros? E qual a probabilidade de que ocorra nos 3 km centrais da rede?

Solução: A função densidade da distribuição Uniforme é dada por

$$f_X(x) = \frac{1}{7} I_{[0,7]}(x).$$

Assim, a probabilidade de ocorrer pane nos primeiros 800 metros é dada por

$$\mathbb{P}\left(X \leqslant 0,8\right) = \int_0^{0,8} \frac{1}{7} dx = \frac{0,8-0}{7} = \text{punif(q=0.8, min = 0, max = 7))} = 0,1142857.$$

e a probabilidade de ocorrer pane nos 3 km centrais da rede é

$$\begin{aligned}
\mathbb{P}\left(2 \leqslant X \leqslant 5\right) &= \int_2^5 \frac{1}{7} dx = \mathbb{P}\left(X \leqslant 5\right) - \mathbb{P}\left(X \leqslant 2\right) \\
&= \text{punif(q=5, min = 0, max = 7)} - \text{punif(q=2, min = 0, max = 7))} \\
&= 5/7 - 2/7 \approx 0,4285714.
\end{aligned}$$

□

Exemplo 6.4. Suponha que $Y \sim \mathrm{U}[0,5]$. Qual a probabilidade que a equação $4x^2 + 4Yx + 4 = 0$, tenha ambas as raízes reais?

Solução: Primeiramente observemos que para que uma equação de segundo grau tenha raízes reais é necessário que o discriminante da equação de segundo grau seja maior ou igual a zero, ou seja, que a fórmula abaixo seja maior ou igual a zero.

$$\Delta = 16Y^2 - 64 \geqslant 0 \Rightarrow Y^2 \geqslant 4.$$

Assim queremos encontrar $\mathbb{P}(Y^2 \geqslant 4)$. Então,

$$\begin{aligned}
\mathbb{P}(Y^2 \geqslant 4) &= 1 - \mathbb{P}(Y^2 \leqslant 4) = 1 - \mathbb{P}(-2 \leqslant Y \leqslant 2) = 1 - \left[\int_{-2}^0 0dy + \int_0^2 \frac{1}{5} dy\right] \\
&= 1 - \left[\frac{x}{5}\Big|_0^2\right. = 1 - \frac{2}{5} = \frac{3}{5}.
\end{aligned}$$

Assim a probabilidade de que ambas as raízes sejam reais é de $\frac{3}{5}$.

□

6.2 Modelo Exponencial

Esta é uma distribuição que se caracteriza por ter uma função de taxa de falha constante. A distribuição exponencial é a única com esta propriedade. Ela é considerada uma das mais simples em termos matemáticos. Esta distribuição tem sido usada extensivamente como um modelo para o tempo de vida de certos produtos, materiais, equipamento, tempo que se leva para completar uma tarefa. Ela descreve adequadamente o tempo de vida de óleos isolantes e dielétricos, entre outros.

Outros exemplo são: Tempo para realizar uma prova; Tempo de chegadas de pacotes em um roteador; Tempo de vida de aparelhos; Tempo de espera em restaurantes, caixas de banco, postos de saúde.

Além disso, a distribuição exponencial permite caracterizar o tempo/distância entre as ocorrências oriundas de um processo de Poisson.

Imagine que estejamos analisando um jogo de futebol e temos interesse em caracterizar o número de gols por partida, essa variável aleatória é uma Poisson. Podemos ainda caracterizar o tempo entre essas ocorrências, e essa variável aleatória é uma Exponencial.

A seguir definimos uma variável aleatória com distribuição Exponencial.

Definição 6.2. Diremos que uma variável aleatória X segue o *modelo Exponencial*, de parâmetro λ, $\lambda > 0$, se sua função densidade for dada por

$$f_X(x) = \lambda e^{-x\lambda} I_{(0,\infty)}(x). \tag{6.2}$$

Notação: $X \sim \text{Exp}(\lambda)$, para $\lambda > 0$. O parâmetro λ indica a taxa de ocorrência por unidade de medida, que pode ser tempo, distância ou volume, entre outras e $1/\lambda$ é o tempo médio entre chegadas.

Corolário 6.2. *Seja $X \sim Exp(\lambda)$, para $\lambda > 0$. Então $f_X(\cdot)$ dada pela equação (6.2) satisfaz as propriedades de uma função densidade de probabilidade, ou seja, é uma função densidade de probabilidade.*

Demonstração. Vamos verificar se $f_X(\cdot)$, dada pela equação (6.2) é função densidade de probabilidade. Temos que $f_X(\cdot)$ é não negativa e

$$\int_{-\infty}^{\infty} f_X(x)dx = \int_0^{\infty} \lambda e^{-x\lambda}dx = -e^{-x\lambda}\Big|_0^{\infty} = 1.$$

□

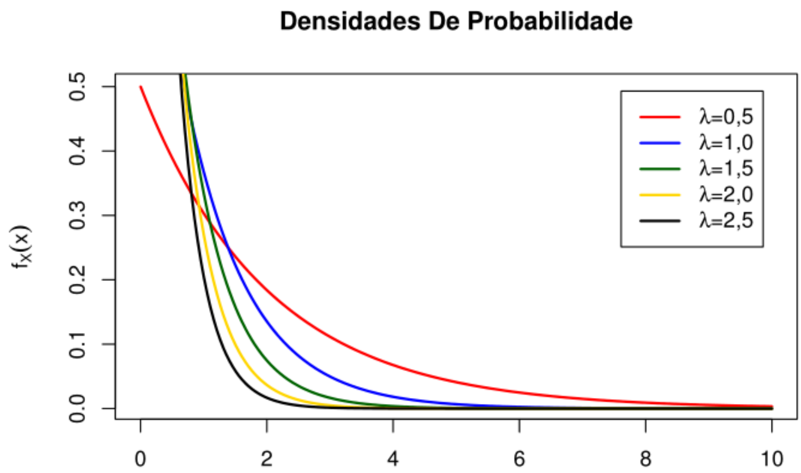


Figura 6.3: Distribuição Exp(λ): função densidade de probabilidade.

Resultado 6.2. Seja $X \sim \text{Exp}(\lambda)$, $\lambda > 0$. Então sua função de distribuição é dada por

$$F_X(x) = (1 - e^{-x\lambda}) I_{(0,\infty)}(x).$$

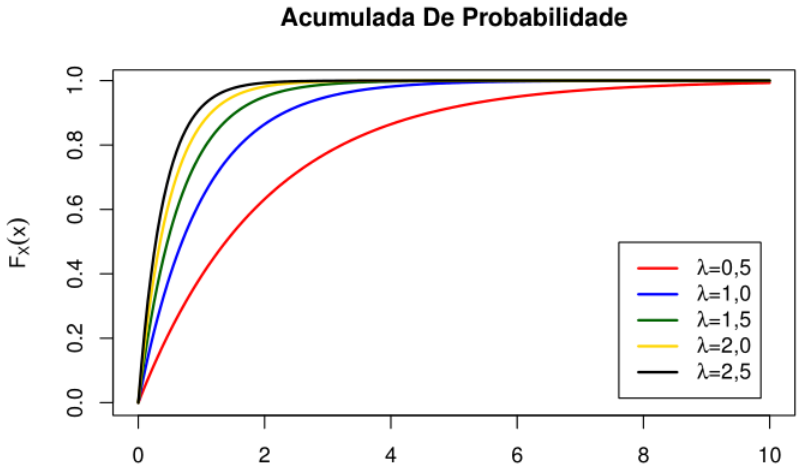


Figura 6.4: Distribuição Exp(λ): função de distribuição acumulada.

Demonstração. Para $x < 0$, temos

$$F_X(x) = \int_{-\infty}^{x} f_X(y)dy = \int_{-\infty}^{x} 0dy = 0.$$

Para $x \geqslant 0$, temos

$$F_X(x) = \int_{-\infty}^{x} f_X(y)dy = \int_{0}^{x} \lambda e^{-y\lambda}dy = -e^{-\lambda y}\Big|_0^x = 1 - e^{-\lambda x}.$$

Portanto a Função de distribuição acumulada é dada por

$$F_X(x) = (1 - e^{-x\lambda})I_{[0,\infty)}(x).$$

□

Proposição 6.2. *Seja* $X \sim \text{Exp}(\lambda)$, *então*

(i) $\mathbb{E}(X) = \frac{1}{\lambda}$.

(ii) $\mathbb{E}(X^2) = \frac{2}{\lambda^2}$.

(iii) $\text{Var}(X) = \frac{1}{\lambda^2}$.

Demonstração. Seja $X \sim \text{Exp}(\lambda)$.

(i) Cálculo de $\mathbb{E}(X)$.

$$\mathbb{E}(X) = \int_{-\infty}^{\infty} x f_X(x)dx = \int_{0}^{\infty} x\lambda e^{-x\lambda}dx = -\frac{(\lambda x + 1)\,\mathrm{e}^{-\lambda x}}{\lambda}\Bigg|_0^{\infty} = \frac{1}{\lambda}.$$

(ii) Cálculo de $\mathbb{E}(X^2)$.

$$\mathbb{E}(X^2) = \int_{-\infty}^{\infty} x^2 f_X(x)dx = \int_{0}^{\infty} x^2 \lambda e^{-x\lambda} dx = -\underbrace{\frac{\left(a^2x^2+2ax+2\right)\mathrm{e}^{-ax}}{a^2}}_{\text{lambda=a}}\Bigg|_0^{\infty} = \frac{2}{\lambda^2}.$$

(iii) Para o cálculo de $\text{Var}(X)$ utilizamos a sua definição, ou seja,

$$\text{Var}(X) = \mathbb{E}(X^2) - [\mathbb{E}(X)]^2 = \frac{2}{\lambda^2} - \left[\frac{1}{\lambda}\right]^2 = \frac{1}{\lambda^2}.$$

□

Exemplo 6.5. Suponha que o tempo de resposta X em um terminal de computador *on-line* específico (o tempo entre o final de uma consulta de um usuário e o começo da resposta do sistema para essa consulta) tenha distribuição exponencial com tempo de resposta esperado de 5 segundos. Então, $\frac{1}{\lambda} = 5$, ou seja, $\lambda = 0.2$. Calcule:

(i) a probabilidade de o tempo de resposta ser no máximo 10 segundos;

(ii) a probabilidade de o tempo de resposta estar entre 5 e 10 segundos.

Solução: Neste caso $X \sim \text{Exp}(0.2)$, então

(i) Cálculo de $\mathbb{P}(X \leqslant 10)$.

$$\mathbb{P}(X \leqslant 10) = F_X(10) = 1 - e^{-10(0.2)} = 1 - 0.135 = \text{pexp(q=10,rate = 0.2)} = 0,8646647.$$

(ii) Cálculo de $\mathbb{P}(5 \leqslant X \leqslant 10)$.

$$\begin{aligned} \mathbb{P}(5 \leqslant X \leqslant 10) &= F_X(10) - F_X(5) = (1 - e^{-10(0.2)}) - (1 - e^{-5(0.2)}) \\ &= \text{pexp(q=10,rate = 0.2)} - \text{pexp(q=5,rate = 0.2)} \\ &= 0.8646647 - 0,6321206 = 0,2325442. \end{aligned}$$

□

Exemplo 6.6. O tempo de vida (em horas) de um transistor pode ser considerado uma variável aleatória com distribuição exponencial. Seque-se que o a vida média do transistor é de 500 horas. Qual a probabilidade que ele dure mais do que a média?

Solução: Neste caso $\frac{1}{\lambda} = 500$, então $\lambda = 0.002$ e $X \sim \text{Exp}(0.002)$. Logo,

$$\begin{aligned} \mathbb{P}(X > 500) &= 1 - \mathbb{P}(X \leqslant 500) = 1 - F_X(500) = 1 - (1 - e^{-500(0.002)}) \\ &= 1 - \text{pexp(q=500,rate = 0.002)} \\ &= \text{pexp(q=500,rate = 0.002,lower.tail = FALSE)} \\ &= 0,3678794. \end{aligned}$$

□

Exemplo 6.7. Uma fábrica utiliza dois métodos para a produção de lâmpadas. 70% das lâmpadas são produzidas pelo método A e as demais pelo método B. A duração da lâmpada depende do método pelo qual ela foi produzida, sendo que as produzidas pelo método A seguem uma distribuição exponencial com parâmetro 1/80 e as do método B seguem uma exponencial de parâmetro 1/100. Qual a probabilidade de que, se escolhermos uma lâmpada ao acaso, ela dure mais de 100 horas?

Resolução: Sejam $X_A \sim$ Exp(1/80) e $X_B \sim$ Exp(1/100) e considere os evento C=Uma lâmpada durar mais de 100 horas, A=A lâmpada ter sido fabricada pelo método A e B=A lâmpada ter sido fabricada pelo método B. Assim usando o Teorema da Probabilidade Total obtemos que

$$\mathbb{P}(C) = \mathbb{P}(C|A)\mathbb{P}(A) + \mathbb{P}(C|B)\mathbb{P}(B) = \mathbb{P}(X_A \geqslant 100)0,7 + \mathbb{P}(X_B \geqslant 100)0,3$$

e, portanto,

$$\begin{aligned}
\mathbb{P}(C) &= 0,7\int_{100}^{\infty}\frac{1}{80}e^{-\frac{x}{80}}dx + 0,3\int_{100}^{\infty}\frac{1}{100}e^{-\frac{x}{100}}dx \\
&= \text{pexp(q=100,rate=1/80,lower.tail=FALSE)} \times 0,7 \\
&\quad +\text{pexp(q=100,rate=1/100,lower.tail=FALSE)} \times 0,3 \\
&= 0,2865048 \times 0,7 + 0,3678794 \times 0,3 = 0,3109172.
\end{aligned}$$

Portanto a probabilidade de que uma lâmpada escolhida ao acaso dure mais de 100 horas é de 0,3109172.

□

6.3 Modelo Normal

A distribuição normal é a mais importante de todas as distribuições. Muitas populações numéricas possuem distribuição que podem ser ajustadas por uma curva normal apropriada. Os exemplos incluem alturas, pesos e outras características físicas, erros em medidas em experimentos científicos, medidas antropométricas em fósseis, tempos de reação em experimentos psicológicos, medidas de inteligência e aptidão, pontuação em testes variados, e numerosas medidas e indicadores econômicos. Mesmo quando a distribuição é discreta, a curva normal frequentemente fornece aproximação excelente.Além disso, ainda que as próprias variáveis individuais não seja normalmente distribuídas, as somas e as médias das variáveis terão uma distribuição aproximadamente normal sob condições adequadas.

Definição 6.3. Uma variável aleatória X segue o *modelo Normal* se a sua densidade é dada por:

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{(x-\mu)^2}{2\sigma^2}}I_{(-\infty,\infty)}(x), \tag{6.3}$$

com μ, $\sigma \in \mathbb{R}$, $\sigma > 0$.

Notação: $X \sim \mathrm{N}(\mu,\sigma^2)$.

Corolário 6.3. *Seja $X \sim N(\mu,\sigma^2)$. Então $f_X(\cdot)$ dada pela equação (6.3) satisfaz as propriedades de uma função densidade de probabilidade, ou seja, é uma função densidade de probabilidade.*

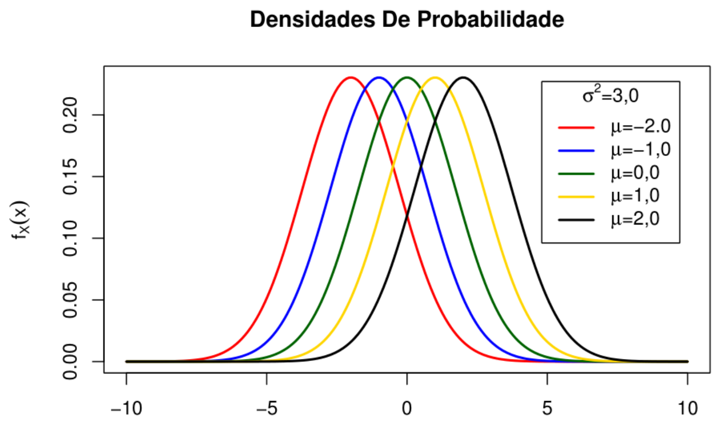


Figura 6.5: Gráficos da função densidade de probabilidade da distribuição normal $\mathrm{N}(\mu, \sigma^2)$: $\mu \in \{-2.0, -1.0, 0.0, 1.0, 2.0\}$ e $\sigma^2 = 3.0$.

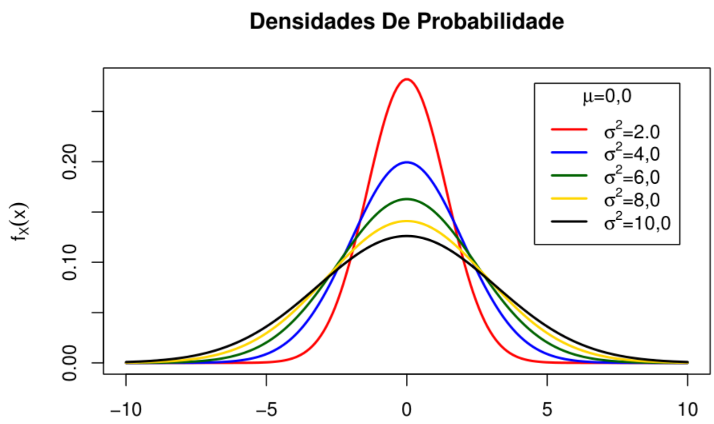


Figura 6.6: Gráficos da função densidade de probabilidade da distribuição normal $\mathrm{N}(\mu, \sigma^2)$: $\mu = 0,0$ e $\sigma^2 \in \{2.0, 4.0, 6.0, 8.0, 10.0\}$.

Demonstração. Seja $X \sim \mathrm{N}(\mu, \sigma^2)$. Temos que verificar se $f_X(\cdot)$, dada pela equação (6.3), satisfaz as propriedades de função densidade de probabilidade.

Temos que $f_X(\cdot)$ é uma função não negativa. Além disso,

$$\int_{-\infty}^{\infty} f_X(x)dx = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx$$

Fazendo $z = \frac{x-\mu}{\sigma}$, então $\frac{dz}{dx} = \frac{1}{\sigma}$, ou seja, $dx = \sigma dz$. Assim,

$$\int_{-\infty}^{\infty} f_X(x)dx = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \frac{1}{\sigma} e^{-\frac{1}{2}(z)^2} \sigma dz = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} dz = \mathrm{I}$$

Temos que calcular a integral I. Vamos então calcular I^2. Assim,

$$
\begin{aligned}
\mathrm{I}^2 &= \left[\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} dz\right]^2 = \left[\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} dz\right] \times \left[\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}s^2} ds\right] \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{1}{2\pi} e^{-\frac{1}{2}(z^2+s^2)} dz ds.
\end{aligned}
$$

Fazendo nova mudança de variável $z = r\cos(\theta)$ e $s = r\,\mathrm{sen}(\theta)$, então $drds = rdrd\theta$. Além disso, como $-\infty < z, s < \infty$, tenos que $r \geqslant 0$ e $0 \leqslant \theta \leqslant 2\pi$. Portanto, $r^2 + s^2 = r^2$ de modo que

$$
\mathrm{I}^2 = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{1}{2\pi} e^{-\frac{1}{2}(z^2+s^2)} dz ds = \int_{0}^{2\pi}\int_{0}^{\infty} \frac{1}{2\pi} e^{-\frac{1}{2}r^2} r dr d\theta = \int_{0}^{2\pi} \frac{1}{2\pi} \int_{0}^{\infty} r e^{-\frac{1}{2}r^2} dr d\theta
$$

Fazendo uma terceira mudança de variável $u = -\frac{r^2}{2}$, $\frac{du}{dr} = -r$, $du = -rdr$. Assim,

$$
\begin{aligned}
\mathrm{I}^2 &= \int_{0}^{2\pi} \frac{1}{2\pi} \int_{0}^{\infty} r e^{-\frac{1}{2}r^2} dr d\theta = \int_{0}^{2\pi} \frac{1}{2\pi} \int_{0}^{\infty} -e^{u} du d\theta \\
&= \int_{0}^{2\pi} \frac{1}{2\pi} \left[-e^{u} \Big|_{0}^{\infty} \right] d\theta = \int_{0}^{2\pi} \frac{1}{2\pi} d\theta = 1.
\end{aligned}
$$

Aplicando a raiz quadrada em I^2 para obtermos I. Assim,

$$
\int_{-\infty}^{\infty} f_X(x) dx = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = I = 1.
$$

Portanto, $f_X(\cdot)$, dada pela equação (6.3), satisfaz as propriedades de função densidade de probabilidade.

□

Definição 6.4. Distribuição Normal Padrão: Quando, na Definição 6.3, temos $\mu = 0$ e $\sigma^2 = 1$. Uma variável aleatória Z segue o modelo Normal Padrão se a sua densidade é dada por:

$$
f_Z(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} I_{(-\infty,\infty)}(z).
$$

Notação: $Z \sim \mathrm{N}(0, 1)$.

Proposição 6.3. *Seja* $X \sim \mathrm{N}(\mu, \sigma^2)$, *então*

(i) $\mathbb{E}(X) = \mu$.

(ii) $\mathbb{E}(X^2) = \sigma^2 + \mu^2$.

(iii) $\mathrm{Var}(X) = \sigma^2$.

Demonstração. Seja $X \sim \mathrm{N}(\mu, \sigma^2)$.

(i) Cálculo de $\mathbb{E}(X)$.

$$\mathbb{E}(X) = \int_{-\infty}^{\infty} x f_X(x) dx = \int_{-\infty}^{\infty} x \left[\frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}\right] dx = \int_{-\infty}^{\infty} x \left[\frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}\right] dx$$

Fazendo $z = \frac{x-\mu}{\sigma}$ $(x = \sigma z + \mu)$, então $\frac{dz}{dx} = \frac{1}{\sigma}$, ou seja, $dx = \sigma dz$. Assim,

$$\begin{aligned}
\mathbb{E}(X) &= \int_{-\infty}^{\infty} x \left[\frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}\right] dx = \int_{-\infty}^{\infty} (\sigma z + \mu) \left[\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}(z)^2}\right] \sigma dz \\
&= \int_{-\infty}^{\infty} \mu \left[\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2}\right] dz + \int_{-\infty}^{\infty} \sigma z \left[\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2}\right] dz \\
&= \mu \underbrace{\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} dz}_{\mathrm{I}} + \frac{\sigma}{\sqrt{2\pi}} \underbrace{\int_{-\infty}^{\infty} z e^{-\frac{1}{2}z^2} dz}_{\mathrm{II}}.
\end{aligned}$$

Temos que I $= 1$, pois $f_Z(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} I_{(-\infty,\infty)}(z)$ é a função densidade de probabilidade de uma variável aleatória $\mathrm{N}(0,1)$ (ver Definição 6.4).

Para encontrarmos o valor de II fazemos a mudança de variável $u = -\frac{r^2}{2}$, $\frac{du}{dr} = -r$, $du = -rdr$. Assim,

$$\mathrm{II} = \int_{-\infty}^{\infty} z e^{-\frac{1}{2}z^2} dz = 0.$$

Portanto,

$$\begin{aligned}
\mathbb{E}(X) &= \mu \underbrace{\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} dz}_{\mathrm{I}} + \frac{\sigma}{\sqrt{2\pi}} \underbrace{\int_{-\infty}^{\infty} z e^{-\frac{1}{2}z^2} dz}_{\mathrm{II}}. \\
&= \mu \times 1 + \frac{\sigma}{\sqrt{2\pi}} \times 0 = \mu.
\end{aligned}$$

(ii) Cálculo de $\mathbb{E}(X^2)$.

$$\mathbb{E}(X^2) = \int_{-\infty}^{\infty} x^2 f_X(x) dx = \int_{-\infty}^{\infty} x^2 \left[\frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}\right] dx = \int_{-\infty}^{\infty} x^2 \left[\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}\right] dx$$

Fazendo $z = \frac{x-\mu}{\sigma}$ $(x = \sigma z + \mu)$, então $\frac{dz}{dx} = \frac{1}{\sigma}$, ou seja, $dx = \sigma dz$. Assim,

$$
\begin{aligned}
\mathbb{E}(X^2) &= \int_{-\infty}^{\infty} x^2 \left[\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}\right] dx = \int_{-\infty}^{\infty} (\sigma z+\mu)^2 \left[\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}(z)^2}\right] \sigma dz \\
&= \int_{-\infty}^{\infty} (\sigma^2 z^2 + 2\mu\sigma z + \mu^2) \left[\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}z^2}\right] \sigma dz \\
&= \sigma^2 \int_{-\infty}^{\infty} z^2 \left[\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2}\right] dz + 2\mu\sigma \int_{-\infty}^{\infty} z \left[\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2}\right] dz + \mu^2 \int_{-\infty}^{\infty} \left[\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2}\right] dz \\
&= \sigma^2 \underbrace{\int_{-\infty}^{\infty} z^2 \frac{\sigma^2}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} dz}_{\mathrm{I}} + \frac{2\mu\sigma}{\sqrt{2\pi}} \underbrace{\int_{-\infty}^{\infty} z e^{-\frac{1}{2}z^2} dz}_{\mathrm{II}} + \mu^2 \underbrace{\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} dz}_{\mathrm{III}}
\end{aligned}
$$

Analisando cada uma das integrais acima temos

I $\displaystyle\int_{-\infty}^{\infty} z^2 \frac{\sigma^2}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} dz = 1$, fazendo derivação por parte com $u = z$ e $dv = ze^{-\frac{1}{2}z^2}$.

II $\displaystyle\int_{-\infty}^{\infty} z e^{-\frac{1}{2}z^2} dz = \mathbb{E}(Z) = 0$, onde $Z \sim \mathrm{N}(0,1)$.

III $\displaystyle\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} dz = 1$, pois $\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2}$ é a função densidade de probabilidade de uma variável aleatória $Z \sim \mathrm{N}(0,1)$.

Assim,

$$\mathbb{E}(X^2) = \sigma^2 \times 1 + 2\mu\sigma \times 0 + \mu^2 \times 1 = \sigma^2 + \mu^2.$$

(iii) Pela definição de $\mathrm{Var}(X)$, temos que

$$\mathrm{Var}(X) = \mathbb{E}(X^2) - [\mathbb{E}(X)]^2 = \sigma^2 + \mu^2 - \mu^2 = \sigma^2.$$

□

Resultado 6.3. Seja $X \sim \mathrm{N}(\mu, \sigma^2)$. Então a função de distribuição (acumulada) da variável aleatória X é dada por

$$F_X(x) = \int_{-\infty}^{x} f_X(y) dy = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(y-\mu)^2}{2\sigma^2}} dy$$

Não existe uma fórmula fechada para esta distribuição.

Proposição 6.4. *Sendo $X \sim N(\mu, \sigma^2)$, então $Z = \frac{X-\mu}{\sigma}$ possui distribuição $N(0,1)$.*

Demonstração. Seja $X \sim \mathrm{N}(\mu, \sigma^2)$, e defina $Z = g(X) = \frac{X-\mu}{\sigma}$. Então,

$$F_Z(z) = \mathbb{P}(Z \leqslant z) = \mathbb{P}\left(\frac{X-\mu}{\sigma} \leqslant z\right) = \mathbb{P}(X \leqslant \sigma z + \mu) = F_X(\sigma z + \mu).$$

Derivando a equação acima em relação a z temos

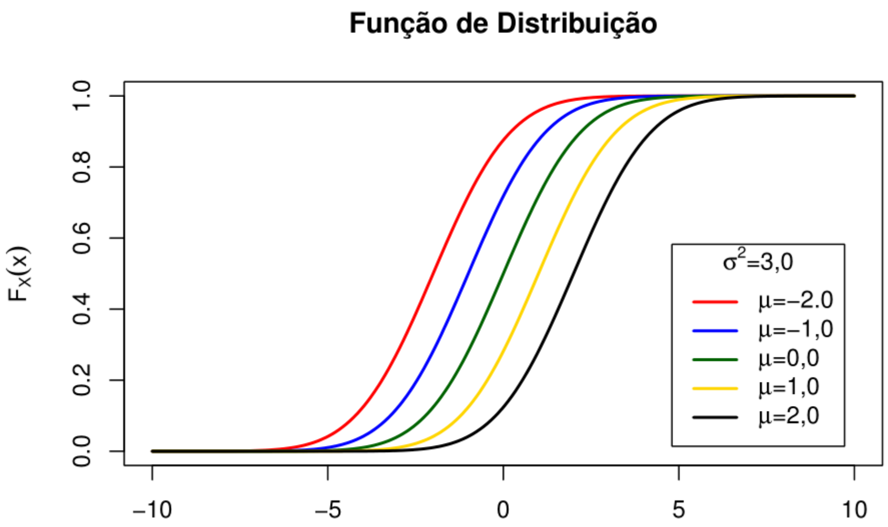


Figura 6.7: Função de distribuição (acumulada) da distribuição $N(\mu, \sigma^2)$: $\mu \in \{-2.0, -1.0, 0.0, 1.0, 2.0\}$ e $\sigma^2 = 3.0$.

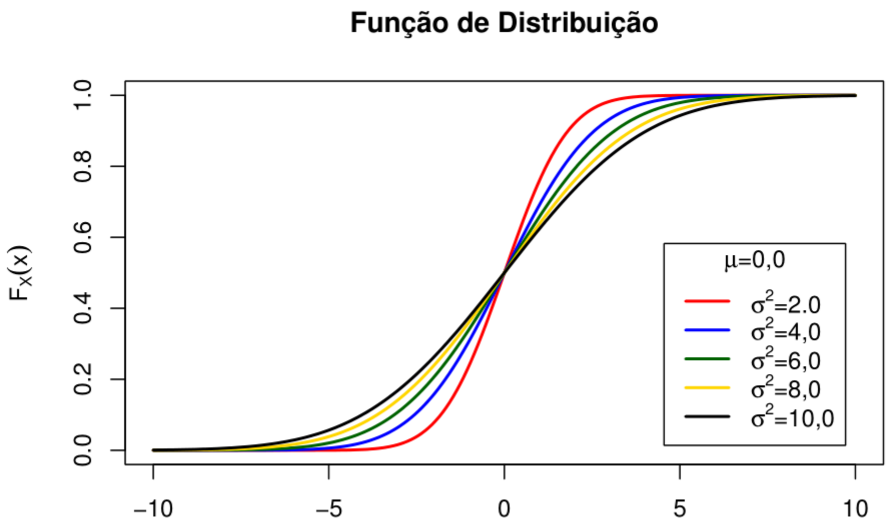


Figura 6.8: Função de distribuição (acumulada) da distribuição $N(\mu, \sigma^2)$: $\mu = 0,0$ e $\sigma^2 \in \{2.0, 4.0, 6.0, 8.0, 10.0\}$.

$$f_Z(z) = \frac{d}{dz} F_Z(z) = \frac{d}{dz} F_X(\sigma z + \mu) = f_X(\sigma z + \mu) \times \sigma.$$

Assim,

$$f_Z(z) = f_X(\sigma z + \mu) \times \sigma = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}.$$

Portanto, $Z = \frac{X-\mu}{\sigma}$ possui distribuição $N(0, 1)$.

$\square$

Proposição 6.5. *Sendo $Z \sim N(0,1)$, então $X = \sigma Z + \mu$ possui distribuição $N(\mu, \sigma^2)$.*

Demonstração. Seja $Z \sim \mathrm{N}(0,1)$ e defina $X = g(Z) = \sigma Z + \mu$, onde $\sigma, \mu \in \mathbb{R}$ e $\sigma \neq 0$, então

$$\begin{aligned} F_X(x) &= \mathbb{P}(X \leqslant x) = \mathbb{P}(g(Z) \leqslant x) = \mathbb{P}(\sigma Z + \mu \leqslant x) \\ &= \mathbb{P}(\sigma Z \leqslant x - \mu) = \begin{cases} \mathbb{P}\left(Z \leqslant \frac{x-\mu}{\sigma}\right), & \text{se } \sigma > 0, \\ \mathbb{P}\left(Z \geqslant \frac{x-\mu}{\sigma}\right), & \text{se } \sigma < 0. \end{cases} \end{aligned}$$

(i) $\sigma > 0$

$$F_X(x) = \mathbb{P}\left(Z \leqslant \frac{x-\mu}{\sigma}\right) = F_Z\left(\frac{x-\mu}{\sigma}\right)$$

Derivando temos

$$f_X(x) = \frac{d}{dx} F_Z\left(\frac{x-\mu}{\sigma}\right) = \frac{1}{\sigma} f_Z\left(\frac{x-\mu}{\sigma}\right) \tag{6.4}$$

(ii) $\sigma < 0$

$$\begin{aligned} F_X(x) &= \mathbb{P}\left(Z \geqslant \frac{x-\mu}{\sigma}\right) \overset{1-}{=} \mathbb{P}\left(Z < \frac{x-\mu}{\sigma}\right) \\ &\overset{1-}{=} \mathbb{P}\left(Z < \frac{x-\mu}{\sigma}\right) + \mathbb{P}\left(Z = \frac{x-\mu}{\sigma}\right) - \mathbb{P}\left(Z = \frac{x-\mu}{\sigma}\right) \\ &\overset{1-}{=} \mathbb{P}\left(Z \leqslant \frac{x-\mu}{\sigma}\right) \pm \mathbb{P}\left(Z = \frac{y-\mu}{\sigma}\right) \\ &= 1 - F_Z\left(\frac{x-\mu}{\sigma}\right), \end{aligned}$$

onde $\mathbb{P}\left(Z = \frac{y-\mu}{\sigma}\right) = 0$ pois Z é uma variável aleatória contínua.

Derivando temos

$$f_X(x) = \frac{d}{dx}\left[1 - F_Z\left(\frac{x-\mu}{\sigma}\right)\right] = -\frac{1}{\sigma} f_Z\left(\frac{x-\mu}{\sigma}\right) \tag{6.5}$$

As equações (7.3) e (6.5) são exatamente iguais. Na equação (6.5) temos o sinal negativo, mas $\sigma < 0$ e assim temos a função positiva.

Como Z é uma variável aleatória com distribuição $\mathrm{N}(0,1)$, sua função densidade de probabilidade é dada por

$$f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, \quad -\infty < x < \infty. \tag{6.6}$$

Substituindo a equação (7.1) na equação (7.3), temos

$$\begin{aligned} f_X(x) &= \frac{1}{\sigma} f_Z\left(\frac{x-\mu}{\sigma}\right) \\ &= \frac{1}{\sigma}\frac{1}{\sqrt{2\pi}} e^{-\frac{\left(\frac{x-\mu}{\sigma}\right)^2}{2}} \\ &= \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, x \in \mathbb{R}, \end{aligned}$$

pois $\mathbb{P}(X = x) = 0$ (caso $\sigma = 0$).

Portanto $X = \sigma Z + \mu$ possui distribuição $\mathrm{N}(\mu, \sigma^2)$.

□

Proposição 6.6. *Da Proposição 6.4 temos o seguinte resultado: seja $X \sim N(\mu, \sigma^2)$, então $Z = \frac{Z-\mu}{\sigma}$ possui distribuição $N(0, 1)$. Dessa forma,*

$$\mathbb{P}(a \leqslant X \leqslant b) = \mathbb{P}\left(\frac{a-\mu}{\sigma} \leqslant \frac{X-\mu}{\sigma} \leqslant \frac{b-\mu}{\sigma}\right) = \mathbb{P}\left(\frac{a-\mu}{\sigma} \leqslant Z \leqslant \frac{b-\mu}{\sigma}\right)$$

ou seja,

$$\mathbb{P}(a \leqslant X \leqslant b) = F_Z\left(\frac{b-\mu}{\sigma}\right) - F_Z\left(\frac{a-\mu}{\sigma}\right)$$

onde

$$\mathbb{P}(X \leqslant a) = F_Z\left(\frac{a-\mu}{\sigma}\right) \quad e \quad \mathbb{P}(X \geqslant b) = 1 - F_Z\left(\frac{a-\mu}{\sigma}\right)$$

Observação 6.1. Seja $Z \sim N(0, 1)$. Então,

(i) $f_Z(z)$ é simétrica em relação à origem (ver Figura 6.9). Então,

$$F_Z(z) = 1 - F_Z(-z).$$

Figura 6.9: Simetria da Distribuição Normal.

(ii) $f_Z(z)$ tem um único ponto crítico em $z = 0$ e $f_Z(0) = \frac{1}{2\pi}$ é o único máximo da função.

(iii) $z = 1$ e $z = -1$ são pontos de inflexão.

(iv) $\lim_{z \to \infty} f_Z(z) = 0 = \lim_{z \to -\infty} f_Z(z)$.

(v) $\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz = 1$.

(vi) $\mathbb{P}(Z \in A) = \int_A f_Z(z) dz$. Caso $A = [a, b]$, com $a < b$, temos que

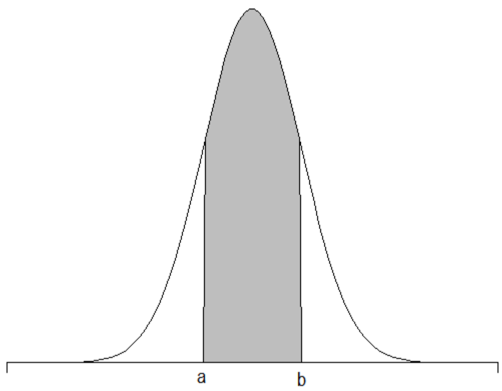


Figura 6.10: $\mathbb{P}(a \leqslant Z \leqslant b) = \int_a^b f_Z(z) dz$.

Exemplo 6.8. [Devore, 2016] Seja $Z \sim N(0, 1)$, isto é, a variável aleatória Z tem distribuição Normal padrão. Calcule as seguintes probabilidades:

(i) $\mathbb{P}(Z \leqslant 1.25)$;

(ii) $\mathbb{P}(Z > 1.25)$;

(iii) $\mathbb{P}(Z \leqslant -1.25)$;

(iv) $\mathbb{P}(-0.38 \leqslant Z \leqslant 1.25)$.

Solução: Será utilizada a Tabela da Distribuição Normal Padrão.

(i) $\mathbb{P}(Z \leqslant 1.25) = F_Z(1.25)$ é uma probabilidade que é tabulada. Pela Tabela da Distribuição Normal Padrão na intersecção na linha marcada com 1.2 e da coluna marcada com 5 (que significa 0.05). O número existente é 0.8944, portanto, $\mathbb{P}(Z \leqslant 1.25) = 0.8944$

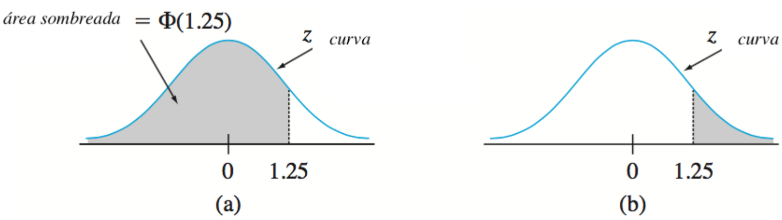


Figura 6.11: Área da curva Normal Padrão: (a) $\mathbb{P}(Z \leqslant 1.25)$; (b) $\mathbb{P}(Z > 1.25)$.

Comando R: $\mathbb{P}(Z \leqslant 1.25) = F_Z(1.25) =$ pnorm(1.25) $= 0,8943502$

(ii) $\mathbb{P}(Z > 1.25) = 1 - \mathbb{P}(Z \leqslant 1.25) = 1 - 0.8944 = 0.1056$;

Comando R: $\mathbb{P}(Z > 1.25) =$ pnorm(1.25,lower.tail=FALSE) $= 0,1056498$

(iii) $\mathbb{P}(Z \leqslant -1.25) = F_Z(-1.25)$. Pela Tabela da Distribuição Normal Padrão na intersecção na linha marcada com -1.2 e da coluna marcada com 5 (que significa 0.05). O número existente é 0.1056, portanto, $\mathbb{P}(Z \leqslant -1.25) = 0.1056$. Pela simetria da curva da distribuição Normal Padrão, $\mathbb{P}(Z \leqslant -1.25) = \mathbb{P}(Z > 1.25)$.

Comando R: $\mathbb{P}(Z \leqslant -1.25) = F_Z(-1.25) =$ pnorm(-1.25) $= 0,1056498$

(iv) $\mathbb{P}(-0.38 \leqslant Z \leqslant 1.25) = F_Z(1.25) - F_Z(-0.38) = 0.8944 - 0.3520 = 0.5424$.

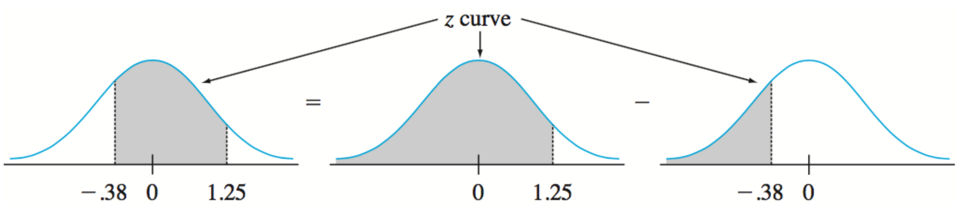


Figura 6.12: $\mathbb{P}(-0.38 \leqslant Z \leqslant 1.25)$ como a diferença entre duas áreas da função de distribuição (acumulada).

$$\begin{aligned}
\mathbb{P}(-0.38 \leqslant Z \leqslant 1.25) &= F_Z(1.25) - F_Z(-0.38) \\
&= \text{pnorm(1.25)} - \text{pnorm(-0.38)} \\
&= 0,8943502 - 0,3519727 = 0,5423775.
\end{aligned}$$

□

Exemplo 6.9 (Devore, 2016). Seja $X \sim N(\mu, \sigma^2)$, com $\mu = 1.25$ e $\sigma = 0.46$. Calcule $\mathbb{P}(1 \leqslant X \leqslant 1.75)$.

Solução: Temos que padronizar utilizando a distribuição Normal Padrão, ou seja,

$$\begin{aligned}
\mathbb{P}(1 \leqslant X \leqslant 1.75) &= \mathbb{P}\left(\frac{1 - 1.25}{0.46} \leqslant \frac{X - 1.25}{0.46} \leqslant \frac{1.75 - 1.25}{0.46}\right) = \mathbb{P}\left(-0.54 \leqslant Z \leqslant 1.09\right) \\
&= \mathbb{P}(Z \leqslant 1.09) - \mathbb{P}(Z \leqslant -0.54) = F_Z(1.09) - F_Z(-0.54) \\
&= \text{pnorm(1.09)} - \text{pnorm(-0.54)} \\
&= 0,8621434 - 0,2945985 = 0,5675449.
\end{aligned}$$

Veja a Figura 6.13 para ver graficamente a relação entre a distribuição $\mathrm{N}(\mu, \sigma^2)$ e a distribuição $\mathrm{N}(0, 1)$.

Comando R:

$$\begin{aligned}
\mathbb{P}(1 \leqslant X \leqslant 1.75) &= F_X(1,75) - F_X(1,0) \\
&= \text{pnorm(q=1.75, mean=1.25, sd=0.46)} - \text{pnorm(q=1.0, mean=1.25, sd=0.46)} \\
&= 0,861472 - 0,2934003 = 0,5680717.
\end{aligned}$$

□

Exemplo 6.10. Seja $X \sim N(\mu, \sigma^2)$, com $\mu = 100$ e $\sigma^2 = 25$. Calcule

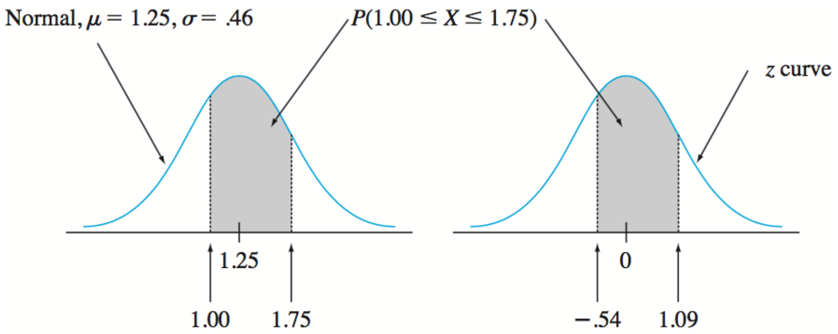


Figura 6.13: $\mathbb{P}(1 \leqslant X \leqslant 1.75) = \mathbb{P}(-0.54 \leqslant Z \leqslant 1.09)$.

(i) $\mathbb{P}(100 \leqslant X \leqslant 106)$.

(ii) $\mathbb{P}(87 \leqslant X \leqslant 107)$.

(iii) $\mathbb{P}(112 \leqslant X \leqslant 116)$.

(iv) $\mathbb{P}(X \geqslant 95)$.

Solução: Seja $X \sim N(\mu, \sigma^2)$, com $\mu = 100$ e $\sigma^2 = 25$. A parametrização a seguir pode ser utilizada para os 3 primeiros itens:

$$\begin{aligned}
\mathbb{P}(a \leqslant X \leqslant b) &= \mathbb{P}\left(\frac{a-\mu}{\sigma} \leqslant \frac{X-\mu}{\sigma} \leqslant \frac{b-\mu}{\sigma}\right) = \mathbb{P}\left(\frac{a-\mu}{\sigma} \leqslant Z \leqslant \frac{b-\mu}{\sigma}\right) \\
&= F_Z\left(\frac{b-\mu}{\sigma}\right) - F_Z\left(\frac{a-\mu}{\sigma}\right) = F_Z\left(\frac{b-100}{5}\right) - F_Z\left(\frac{a-100}{5}\right),
\end{aligned}$$

onde $\frac{X-\mu}{\sigma} = Z \sim N(0,1)$. Assim,

(i) Cálculo de $\mathbb{P}(100 \leqslant X \leqslant 106)$.

$$\begin{aligned}
\mathbb{P}(100 \leqslant X \leqslant 106) &= F_Z\left(\frac{b-100}{5}\right) - F_Z\left(\frac{a-100}{5}\right) \\
&= F_Z\left(\frac{106-100}{5}\right) - F_Z\left(\frac{100-100}{5}\right) \\
&= \text{pnorm(q=(106-100)/5)} - \text{pnorm(q=0)} \\
&= 0.8849303 - 0.5 = 0.3849303.
\end{aligned}$$

Comando R:

$$\begin{aligned}
\mathbb{P}(100 \leqslant X \leqslant 106) &= \text{pnorm(q=106, mean = 100, sd = 5)} - \text{pnorm(q=100, mean = 100, sd = 5)} \\
&= 0.8849303 - 0.5 = 0.3849303
\end{aligned}$$

(ii) Cálculo de $\mathbb{P}(87 \leqslant X \leqslant 107)$.

$$
\begin{aligned}
\mathbb{P}(87 \leqslant X \leqslant 107) &= F_Z\left(\frac{b-100}{5}\right) - F_Z\left(\frac{a-100}{5}\right) \\
&= F_Z\left(\frac{107-100}{5}\right) - F_Z\left(\frac{87-100}{5}\right) \\
&= \text{pnorm(q=(107-100)/5)} - \text{pnorm(q=(87-100)/5)} \\
&= 0.9192433 - 0.004661188 = 0.9145822.
\end{aligned}
$$

Comando R:

$$
\begin{aligned}
\mathbb{P}(87 \leqslant X \leqslant 107) &= \text{pnorm(q=107, mean = 100, sd = 5)} - \text{pnorm(q=87, mean = 100, sd = 5)} \\
&= 0.9192433 - 0.004661188 = 0.9145822
\end{aligned}
$$

(iii) Cálculo de $\mathbb{P}(112 \leqslant X \leqslant 116)$.

$$
\begin{aligned}
\mathbb{P}(112 \leqslant X \leqslant 116) &= F_Z\left(\frac{b-100}{5}\right) - F_Z\left(\frac{a-100}{5}\right) \\
&= F_Z\left(\frac{116-100}{5}\right) - F_Z\left(\frac{112-100}{5}\right) \\
&= \text{pnorm(q=(116-100)/5)} - \text{pnorm(q=(112-100)/5)} \\
&= 0.9993129 - 0.9918025 = 0.007510398.
\end{aligned}
$$

Comando R:

$$
\begin{aligned}
\mathbb{P}(112 \leqslant X \leqslant 116) &= \text{pnorm(q=116, mean = 100, sd = 5)} - \text{pnorm(q=112, mean = 100, sd = 5)} \\
&= 0.9993129 - 0.9918025 = 0.007510398.
\end{aligned}
$$

Os cálculos a seguir serão utilizados para resolver os próximo item.

$$
\mathbb{P}(X \geqslant b) = 1 - \mathbb{P}(X < b) = 1 - \mathbb{P}\left(\frac{X-\mu}{\sigma} < \frac{b-\mu}{\sigma}\right) = 1 - \mathbb{P}\left(Z < \frac{b-\mu}{\sigma}\right) = 1 - F_Z\left(\frac{b-\mu}{\sigma}\right),
$$

onde $\frac{X-\mu}{\sigma} = Z \sim N(0,1)$. Assim

(iv) Cálculo de $\mathbb{P}(X \geqslant 95)$.

$$
\begin{aligned}
\mathbb{P}(X \geqslant 95) &= 1 - F_Z\left(\frac{95-\mu}{\sigma}\right) 1 - F_Z\left(\frac{95-100}{5}\right) \\
&= 1 - \text{pnorm(q=(95-100)/5)} = 1 - 0.1586553 = 0.8413447 \\
&= \text{pnorm(q=(95-100)/5,lower.tail = FALSE)} = 0.8413447
\end{aligned}
$$

Comando R:

$$
\mathbb{P}(X \geqslant 95) = \text{pnorm(q=95, mean=100, sd=5,lower.tail=FALSE)} = 0.8413447.
$$

□

Observação 6.2. Seja $X \sim \mathrm{N}(\mu, \sigma^2)$. Então,

(i) $f_X(x)$ é simétrica com relação à μ. Portanto,

$$\mathbb{P}(X \leqslant \mu) = 1/2 = \mathbb{P}(X \geqslant \mu)$$

e

$$\mathbb{P}(X \leqslant \mu - a) = \mathbb{P}(X \geqslant \mu + a).$$

(ii) $\mu + \sigma$ e $\mu - \sigma$ são pontos de inflexão.

(iii) Para $X \sim \mathrm{N}(\mu, \sigma^2)$, temos que

$$\begin{aligned}
\mathbb{P}(|X-\mu| < k\sigma) &= \mathbb{P}(-k\sigma < X - \mu < k\sigma) = \mathbb{P}(\mu - k\sigma < X < \mu + k\sigma) \\
&= \mathbb{P}\left(\frac{-k\sigma}{\sigma} < \frac{X-\mu}{\sigma} < \frac{k\sigma}{\sigma}\right) = \mathbb{P}\left(-k < \frac{X-\mu}{\sigma} < k\right) \\
&= \mathbb{P}(-k < Z < k) = \mathbb{P}(|Z| < k) = F_Z(k) - F_Z(-k) = 2F_Z(k) - 1,
\end{aligned}$$

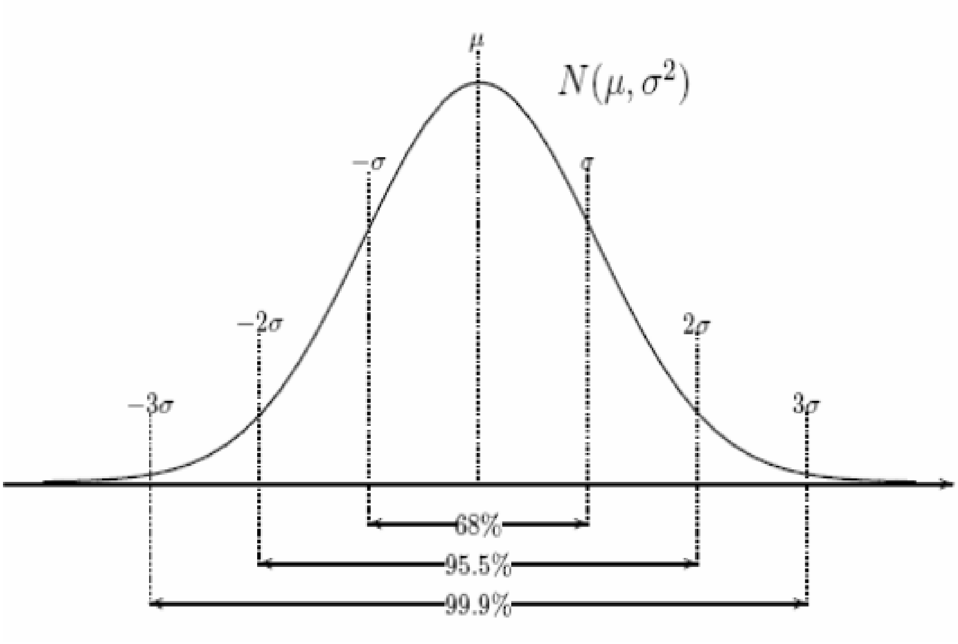
onde $Z \sim \mathrm{N}(0,1)$ e $Z = \frac{X-\mu}{\sigma}$.

Assim,

(a) Para $k = 1$, $\mathbb{P}(|X-\mu| < \sigma) = 2F_Z(1) - 1 = 2 \times 0.8413 - 1 = 0.6826$.

(b) Para $k = 2$, $\mathbb{P}(|X-\mu| < 2\sigma) = 2F_Z(2) - 1 = 2 \times 0.9772 - 1 = 0.9544$.

(c) Para $k = 3$, $\mathbb{P}(|X-\mu| < 3\sigma) = 2F_Z(3) - 1 = 2 \times 0.9987 - 1 = 0.9974$.



Resultado 6.4. Notação z_α para os valores críticos de z:

Em inferência estatística, precisaremos dos valores do eixo z horizontal das medidas que encerram pequenas áreas da cauda abaixo da curva normal padrão.

> **Notação:** A quantidade z_α representará o valor no eixo z para o qual uma área α abaixo da curva da distribuição Normal Padrão fica à direita de z_α. Ou seja,
>
> $$\mathbb{P}(Z \geqslant z_\alpha) = \alpha \text{ ou } \mathbb{P}(Z \leqslant z_\alpha) = 1 - \alpha.$$

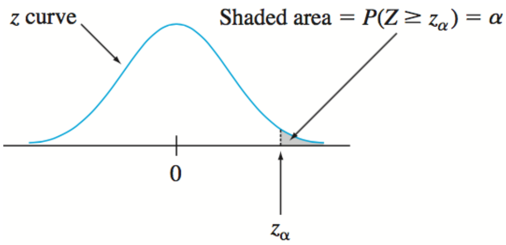


Figura 6.14: Ilustração da notação z_α.

Uma vez que α é a área abaixo da curva da distribuição Z que encontra-se à direita de z_α, $1-\alpha$ é a área que encontra-se a esquerda. Assim, z_α é o $100(1-\alpha)$-ésimo percentil da distribuição normal padrão. Os valores z_α usualmente são considerados os valores críticos de z.

Tabela 6.1: Percentis Normal Padrão e Valores Críticos.

Percentil	90.00	95.00	97.50	99.00	99.50	99.90	99.95
α	0.100	0.050	0.025	0.010	0.005	0.001	0.001
Z_α	1.280	1.645	1.960	2.330	2.580	3.080	3.270

Comando R:

(i) 90º percentil:

$$\mathbb{P}(Z \geqslant z_{0.1}) = 0.1 \quad \rightarrow \quad \text{qnorm(p=0.1,lower.tail = FALSE)} = 1.281552$$

ou

$$\mathbb{P}(Z \leqslant z_{0.1}) = 0.9 \quad \rightarrow \quad \text{qnorm(p=0.9)} = 1.281552$$

(ii) 95º percentil:

$$\mathbb{P}(Z \geqslant z_{0.05}) = 0.05 \quad \rightarrow \quad \text{qnorm(p=0.05,lower.tail = FALSE)} = 1.644854$$

ou

$$\mathbb{P}(Z \leqslant z_{0.05}) = 0.95 \quad \rightarrow \quad \text{qnorm(p=0.95)} = 1.644854$$

(iii) 97.5º percentil:

$$\mathbb{P}(Z \geqslant z_{0.025}) = 0.025 \quad \rightarrow \quad \text{qnorm(p=0.025,lower.tail = FALSE)} = 1.959964$$

ou

$$\mathbb{P}(Z \leqslant z_{0.025}) = 0.975 \quad \rightarrow \quad \text{qnorm(p=0.975)} = 1.959964$$

(iv) 97.5º percentil:

$$\mathbb{P}(Z \geqslant z_{0.01}) = 0.01 \quad \rightarrow \quad \text{qnorm(p=0.01,lower.tail = FALSE)} = 2.326348$$

ou

$$\mathbb{P}(Z \leqslant z_{0.01}) = 0.99 \quad \rightarrow \quad \text{qnorm(p=0.99)} = 2.326348$$

□

Exemplo 6.11. A quantidade $z_{0.05}$ é o $100(1-0.05)$-ésimo $= 95^0$ percentil da distribuição normal padrão, de modo que $z_{0.05} = 1.645$. A área abaixo da curva normal padrão à esquerda de $-z_{0.05}$ também é 0.05.

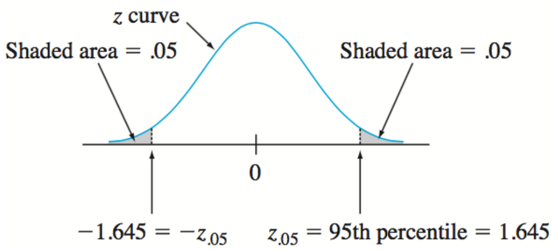


Figura 6.15: Ilustração da notação $z_{0.05}$ e $-z_{0.05}$.

□

O $100(1-\alpha)$-ésimo percentil da distribuição normal com média μ e variância σ^2 pode ser facilmente relacionado com o $100(1-\alpha)$-ésimo percentil da distribuição normal padrão.

> **Proposição 6.7.** *Relação entre os percentis:*
>
> $$100(1-\alpha)\text{-}\textit{ésimo percentil da } N(\mu, \sigma^2) = \mu + 100(1-\alpha)\text{-}\textit{ésimo percentil da } N(0,1) \times \sigma.$$

Exemplo 6.12. Os autores de *"Assessment of lifetime of railway axle" (Intl. J. of Fatigue, 2013: 40-46)* utilizaram dados coletados de um experimento com o comprimento inicial de fissura específico e diversos ciclos de carregamento para propor uma distribuição normal com valor médio 5,496 mm e desvio padrão 0,067 mm para a variável X={profundidade final da fissura}. para este modelo, qual valor da profundidade final da fissura seria excedido por apenas 0,5% de todas as fissuras nestas circunstâncias?

Solução: Denotamos por c o valor solicitado. Então, a condição desejada é que $\mathbb{P}(X > c) = 0,005$, ou de maneira equivalente, que $\mathbb{P}(X \leqslant c) = 0,995$. Assim, c é o 99,5º percentil da distribuição normal com $\mu = 5,496$ e $\sigma = 0,067$. Temos que o 99,5º da distribuição normal padrão é $z_{0,005} =$ qnorm(p=0.995) = 2.575829, então

$$c = \mu + z_{0,005} \times \sigma = 5,496 + 2.575829 \times 0,067 = 5.668581.$$

□

6.4 Aproximação da Binomial pela Normal

Proposição 6.8. [Aproximação da Binomial pela Normal] *Seja $X \sim \mathrm{B}(n,p)$. A medida que n cresce a distribuição $\mathrm{B}(n,p)$ se aproxima da distribuição $\mathrm{N}(\mu,\sigma^2)$, onde $\mu = np$ e $\sigma^2 = npq$. As condições mais importantes são: $np > 5$ e $n(1-p) > 5$.*

A demonstração da validade desta aproximação é feita utilizando-se o Teorema do Limite Central, que será estudado em tempo. Ver Figuras 6.16 e 6.17.

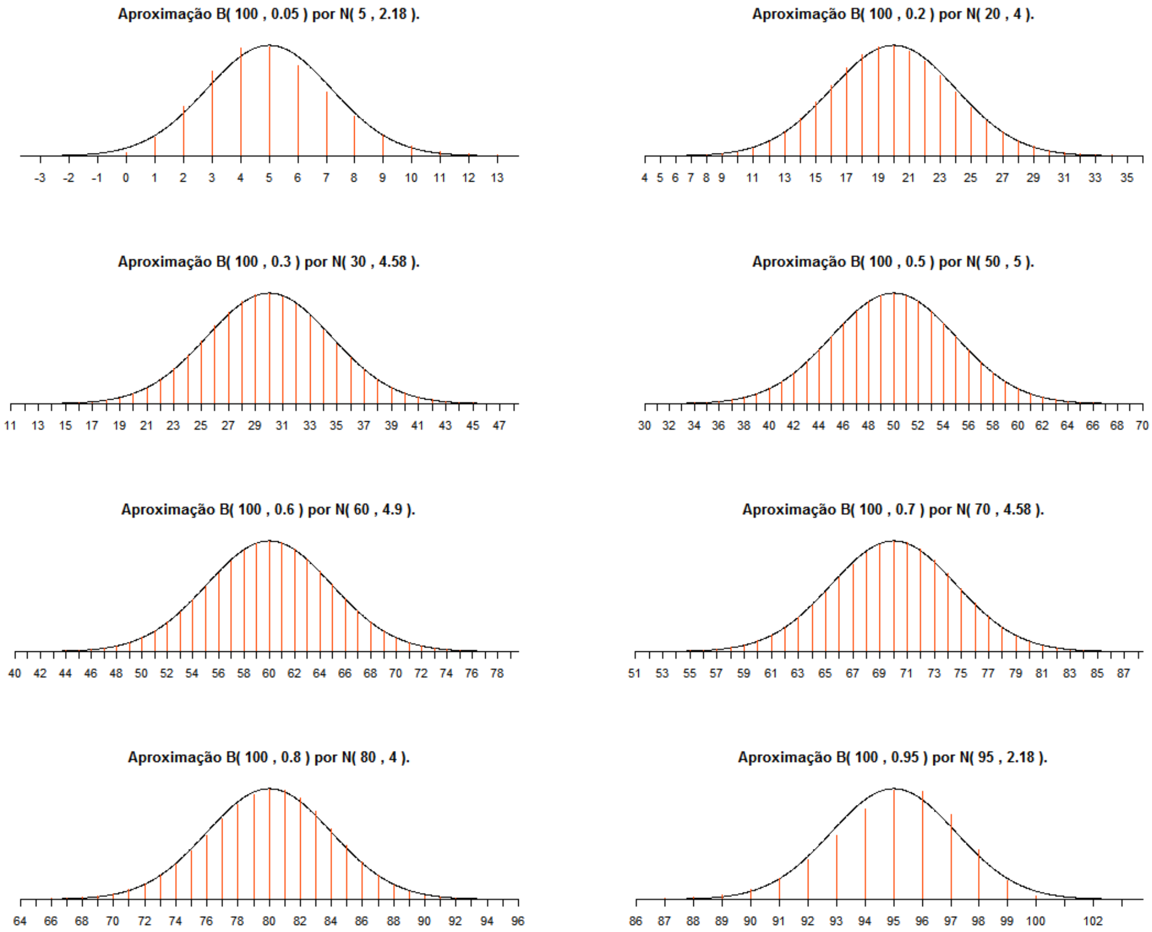


Figura 6.16: Aproximação da distribuição $\mathrm{B}(n,p)$ pela distribuição $\mathrm{N}(\mu,\sigma^2)$, onde $n = 100$ e $p \in \{0.05, 0.2, 0.3, 0.5, 0.6, 0.7, 0.8, 0.95\}$.

Cálculo da Probabilidade

Seja $X \sim \mathrm{B}(n,p)$, tal que $np > 5$. Calcule $\mathbb{P}(a \leqslant X \leqslant b)$, para $a, b \in \mathbb{N}$ utilizando a Proposição 6.8, ou seja utilizando uma variável aleatória $Y \sim \mathrm{N}(\mu,\sigma^2)$. Então,

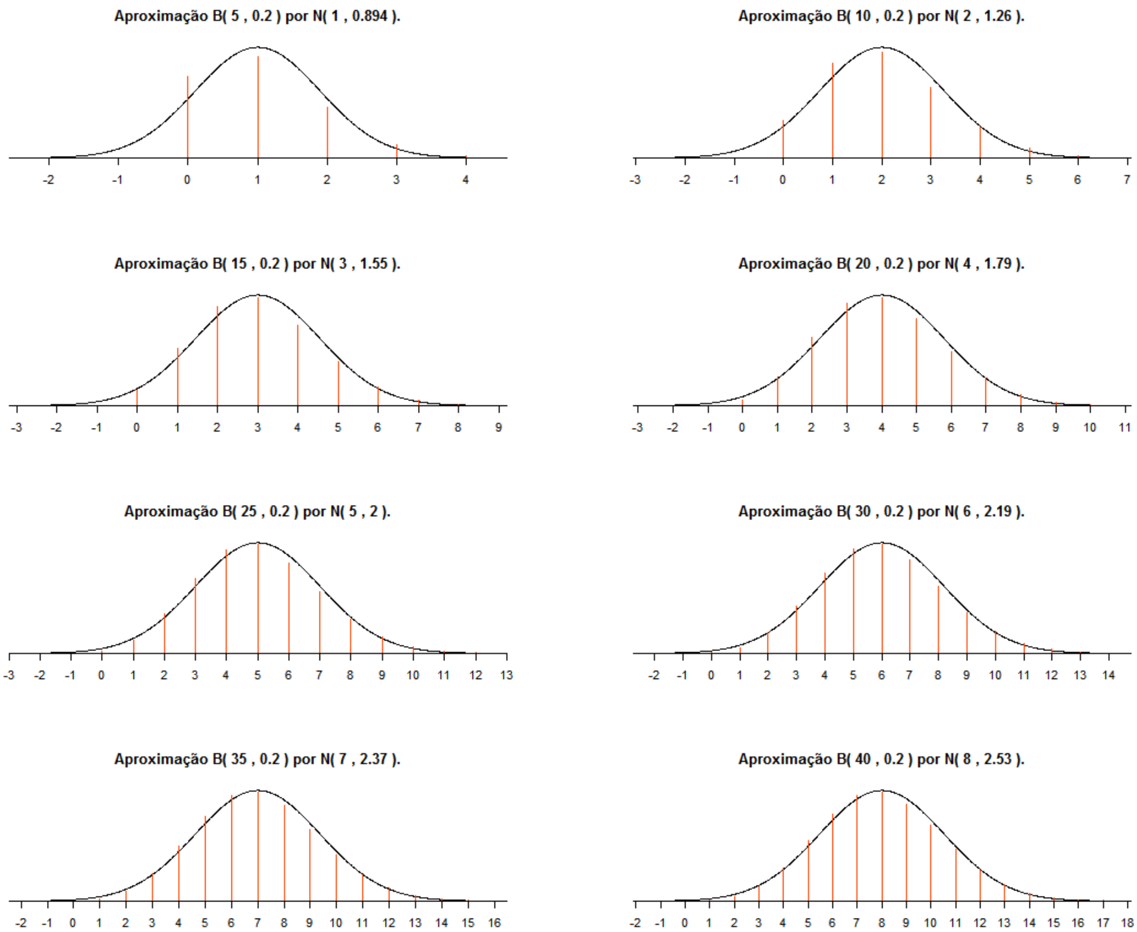


Figura 6.17: Aproximação da distribuição $\mathrm{B}(n, p)$ pela distribuição $\mathrm{N}(\mu, \sigma^2)$, onde $p = 0.25$ e $n \in \{5, 10, 15, 20, 25, 30, 35, 40\}$.

$$\begin{aligned}
\mathbb{P}(a \leqslant X \leqslant b) &\simeq \mathbb{P}(a \leqslant Y \leqslant b) \\
&= \mathbb{P}\left[\frac{a - np}{\sqrt{np(1-p)}} \leqslant \frac{Y - np}{\sqrt{np(1-p)}} \leqslant \frac{b - np}{\sqrt{np(1-p)}}\right] \\
&= \mathbb{P}\left[\frac{a - np}{\sqrt{np(1-p)}} \leqslant Z \leqslant \frac{b - np}{\sqrt{np(1-p)}}\right],
\end{aligned}$$

onde $Z \sim \mathrm{N}(0, 1)$.

A aproximação pode ser melhorada através do uso da Correção de Continuidade.

Correção de Continuidade

A correção de continuidade é um procedimento que pode ser aplicado para melhorar a aproximação de distribuições discretas através de distribuições contínuas. Em particular, na aproximação da distribuição binomial pela normal temos o seguinte:

$$\begin{aligned}\mathbb{P}(a \leqslant X \leqslant b) &\simeq \mathbb{P}\left(a - \frac{1}{2} \leqslant Y \leqslant b + \frac{1}{2}\right)\\ &= \mathbb{P}\left[\frac{a - \frac{1}{2} - np}{\sqrt{np(1-p)}} \leqslant \frac{Y - np}{\sqrt{np(1-p)}} \leqslant \frac{b + \frac{1}{2} - np}{\sqrt{np(1-p)}}\right]\\ &= \mathbb{P}\left[\frac{a - \frac{1}{2} - np}{\sqrt{np(1-p)}} \leqslant Z \leqslant \frac{b + \frac{1}{2} - np}{\sqrt{np(1-p)}}\right]\end{aligned}$$

onde $Z \sim \mathrm{N}(0,1)$.

Caso particular:

$$\begin{aligned}\mathbb{P}(X = a) &\simeq \mathbb{P}\left(a - \frac{1}{2} \leqslant Y \leqslant a + \frac{1}{2}\right)\\ &= \mathbb{P}\left[\frac{a - \frac{1}{2} - np}{\sqrt{np(1-p)}} \leqslant \frac{Y - np}{\sqrt{np(1-p)}} \leqslant \frac{a + \frac{1}{2} - np}{\sqrt{np(1-p)}}\right]\\ &= \mathbb{P}\left[\frac{a - \frac{1}{2} - np}{\sqrt{np(1-p)}} \leqslant Z \leqslant \frac{a + \frac{1}{2} - np}{\sqrt{np(1-p)}}\right]\end{aligned}$$

Exemplo 6.13. Seja $X \sim \mathrm{B}(225, 0.2)$. Calcule $\mathbb{P}(39 \leqslant X \leqslant 48)$.

Solução: Temos que $\mathbb{E}(X) = np = 225 \times 0,2 = 45$ e $\mathrm{Var}(X) = np(1-p) = 225 \times 0,2 \times 0,8 = 36$. Logo a distribuição da variável aleatória X pode ser aproximada pelo distribuição $\mathrm{N}(\mu, \sigma^2)$, com $\mu = 45$ e $\sigma^2 = 36$. Ver Figura 6.18.

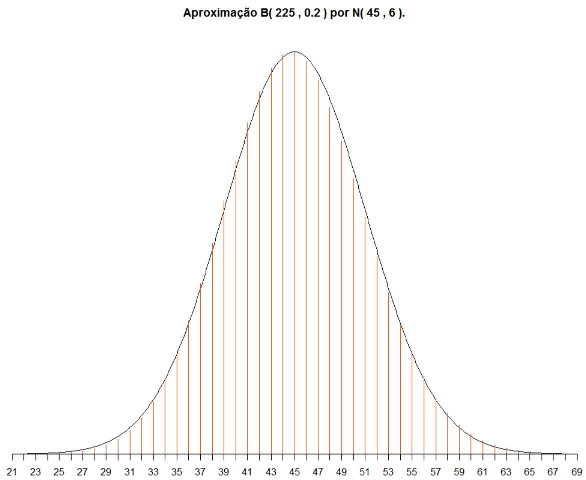


Figura 6.18: Aproximação da distribuição $\mathrm{B}(225, 0, 2)$ pela distribuição $\mathrm{N}(\mu, \sigma^2)$, onde $\mu = 45$ e $\sigma^2 = 36$.

Valor Exato:

$$\mathbb{P}(39 \leqslant X \leqslant 48) = 0,5853.$$

Valor sem o fator de Correção:

$$\begin{aligned}\mathbb{P}(39 \leqslant X \leqslant 48) &\simeq \mathbb{P}(39 \leqslant Y \leqslant 48)\\ &= \mathbb{P}\left(\frac{39-45}{6} \leqslant \frac{Y-45}{6} \leqslant \frac{48-45}{6}\right)\\ &= \mathbb{P}(-1,0 \leqslant Z \leqslant 0,5) = 0,5328.\end{aligned}$$

Valor com o fator de Correção:

$$\begin{aligned}\mathbb{P}(39 \leqslant X \leqslant 48) &\simeq \mathbb{P}(39-0,5 \leqslant Y \leqslant 48+0,5)\\ &= \mathbb{P}\left(\frac{39-0,5-45}{6} \leqslant \frac{Y-45}{6} \leqslant \frac{48+0,5-45}{6}\right)\\ &= \mathbb{P}(-1,08 \leqslant Z \leqslant 0,58) = 0,5808.\end{aligned}$$

□

Cuidado na hora de calcular probabilidade com o fator de correção:

$X \sim \mathrm{B}(n,p)$	$Y \sim \mathrm{N}(\mu,\sigma^2)$
$\mathbb{P}(a \leqslant X \leqslant b)$	$\mathbb{P}(a-0,5 \leqslant Y \leqslant b+0,5)$
$\mathbb{P}(X = k)$	$\mathbb{P}(k-0,5 \leqslant Y \leqslant k+0,5)$
$\mathbb{P}(X < k)$	$\mathbb{P}(Y < k-0,5)$
$\mathbb{P}(X \leqslant k)$	$\mathbb{P}(Y \leqslant k+0,5)$
$\mathbb{P}(X > k)$	$\mathbb{P}(Y > k+0,5)$
$\mathbb{P}(X \geqslant k)$	$\mathbb{P}(Y \geqslant k-0,5)$

Exemplo 6.14. Um sistema é formado por 100 componentes, cada um dos quais com confiabilidade (probabilidade de funcionar adequadamente num certo período) igual a $0,9$. Se esses componentes funcionarem de forma independente um do outro e se o sistema funcionar adequadamente enquanto pelo menos 87 componentes estiverem funcionando, qual é a confiabilidade do sistema?

Solução: Seja X: número de componentes que funcionam adequadamente. Então $X \sim B(100, 0,9)$. Logo $\mathbb{E}(X) = np = 100 \times 0,9 = 90$ e $\mathrm{Var}(X) = np(1-p) = 100 \times 0,9 \times 0,1 = 9$. Logo a distribuição da variável aleatória X pode ser aproximada pelo distribuição $N(\mu, \sigma^2)$, com $\mu = 90$ e $\sigma^2 = 9$

O sistema é confiável quando $\mathbb{P}(X \geqslant 87)$.

Valor Exato:

$$\mathbb{P}(X \geqslant 87) = 0,8761.$$

Valor sem o fator de Correção:

$$\begin{aligned}\mathbb{P}(X \geqslant 87) &\simeq \mathbb{P}(Y \geqslant 87)\\ &= \mathbb{P}\left(\frac{Y-90}{3} \geqslant \frac{87-90}{3}\right)\\ &= \mathbb{P}(Z \geqslant -1) = 0,8413.\end{aligned}$$

Valor com o fator de Correção:

$$\begin{aligned}\mathbb{P}(X \geqslant 87) &\simeq \mathbb{P}(Y \geqslant 87-0.5)\\ &= \mathbb{P}\left(\frac{Y-90}{3} \geqslant \frac{87-0.5-90}{3}\right)\\ &= \mathbb{P}(Z \geqslant -1.16) = 0,8769.\end{aligned}$$

□

Exemplo 6.15. Uma moeda honesta é lançada 100 vezes.

(a) Calcular a probabilidade do número de caras estar entre 40% e 70% dos lançamentos, inclusive.

(b) Determinar um intervalo simétrico em torno do número médio de caras, tal que a probabilidade de observar um valor de X nesse intervalo é 80%.

Solução:

(a) Calcular a probabilidade do número de caras estar entre 40% e 70% dos lançamentos, inclusive.

Seja X:{número de caras em 100 lançamentos}. Então $X \sim B(100, 0.5)$. Logo $\mathbb{E}(X) = np = 100 \times 0,5 = 50$ e $\text{Var}(X) = np(1-p) = 100 \times 0,5 \times 0,5 = 25$. Logo a distribuição da variável aleatória X pode ser aproximada pelo distribuição $N(\mu, \sigma^2)$, com $\mu = 50$ e $\sigma^2 = 25$. Queremos calcular $\mathbb{P}(40 \leqslant X \leqslant 70)$. Vamos utilizar o fator de correção. O Valor exato é 0.9824.

$$\begin{aligned}
\mathbb{P}(40 \leqslant X \leqslant 70) &\simeq \mathbb{P}(40 - 0.5 \leqslant Y \leqslant 70 + 0.5) \\
&= \mathbb{P}\left(\frac{40 - 0.5 - 50}{5} \leqslant \frac{Y - 50}{5} \leqslant \frac{70 + 0.5 - 50}{5}\right) \\
&= \mathbb{P}(-2.1 \leqslant Z \leqslant 4.1) = 0,9821.
\end{aligned}$$

(b) Determinar um intervalo simétrico em torno do número médio de caras, tal que a probabilidade de observar um valor de X nesse intervalo é 80%.

Intervalo simétrico em torno da média: $(50 - a, 50 + a)$. Assim temos que encontrar o valor de a tal que

$$\mathbb{P}(50 - a \leqslant X \leqslant 50 + a) = 0,8.$$

Para tanto, vamos utilizar a aproximação da Binomial pela Normal, com fator de correção.

$$\begin{aligned}
\mathbb{P}(50 - a \leqslant X \leqslant 50 + a) &\simeq \mathbb{P}(50 - 0.5 - a \leqslant Y \leqslant 50 + 0.5 + a) \\
&= \mathbb{P}\left(\frac{50 - 0.5 - a - 50}{5} \leqslant \frac{Y - 50}{5} \leqslant \frac{50 + 0.5 + a - 50}{5}\right) \\
&= \mathbb{P}\left(\frac{-0.5 - a}{5} \leqslant Z \leqslant \frac{0.5 + a}{5}\right) = 0,8.
\end{aligned}$$

Precismos encontrar o valor de a tal que

$$\begin{aligned}
\mathbb{P}\left(-\frac{0.5 + a}{5} \leqslant Z \leqslant \frac{0.5 + a}{5}\right) &= \mathbb{P}\left(Z \leqslant \frac{0.5 + a}{5}\right) - \mathbb{P}\left(Z \leqslant -\frac{0.5 + a}{5}\right) \\
&= F_Z\left(\frac{0.5 + a}{5}\right) - F_Z\left(-\frac{0.5 + a}{5}\right) \\
&= 1 - 2F_Z\left(-\frac{0.5 + a}{5}\right) = 0,8.
\end{aligned}$$

Assim,

$$1 - 2F_Z\left(-\frac{0.5 + a}{5}\right) = 0,8 \quad \longrightarrow \quad F_Z\left(-\frac{0.5 + a}{5}\right) = -\frac{0.8 - 1}{2} = 0.1$$

Pela tabela da distribuição Normal Padrão, temos que $F_Z(-1.281) \simeq 0.1000968$. Logo,

$$-\frac{0.5 + a}{5} = -1.281 \quad \longrightarrow \quad a = 5 \times 1.281 - 0.5 \quad \longrightarrow \quad a = 5.905$$

Intervalo Procurado: $[50 - 5.905; 50 + 5.905] = [44.095; 55.905]$.

Interpretação: A probabilidade de em 100 lançamentos termos entre 44 e 56 caras é aproximadamente 80%.

Probabilidade Exata: $\mathbb{P}(44 \leqslant X \leqslant 56)$, onde $X \sim B(100, 0.5)$ é de 0.8066521.

□

6.5 Aproximação da Poisson pela Normal

Proposição 6.9. [Aproximação da Poisson pela Normal] *Seja $X \sim P(\lambda)$. A medida que λ cresce a distribuição $P(\lambda)$ se aproxima da distribuição $N(\mu, \sigma^2)$, onde $\mu = \lambda$ e $\sigma^2 = \lambda$. A condição mais importante é: $\lambda > 5$.*

A demonstração da validade desta aproximação é feita utilizando-se o Teorema do Limite Central, que será estudado em tempo. Ver Figura 6.19.

Exemplo 6.16. Um processo de produção produz 10 itens defeituosos por hora. Encontre a probabilidade de que entre 8 e 15, inclusive,sejam defeituosos numa retirada aleatória por hora.

Solução: Sabemos que $X \sim P(10)$. Então podemos calcular a probabilidade de maneira exata da seguinte forma

$$\mathbb{P}(8 \leqslant X \leqslant 15) = \sum_{x=8}^{15} \frac{e^{-10}10^x}{x!} = 0.731039.$$

Difícil de ser calculado a mão. Calculamos computacionalmente.

Vamos utilizar a aproximação Normal. Seja $Y \sim N(\mu, \sigma^2)$, com $\mu = \lambda$ e $\sigma^2 = \lambda$.

Valor sem o fator de Correção:

$$\begin{aligned}
\mathbb{P}(8 \leqslant X \leqslant 15) &\simeq \mathbb{P}(8 \leqslant Y \leqslant 15) \\
&= \mathbb{P}\left(\frac{8 - 10}{\sqrt{10}} \leqslant \frac{Y - 10}{\sqrt{10}} \leqslant \frac{15 - 10}{\sqrt{10}}\right) \\
&= \mathbb{P}(-0.63 \leqslant Z \leqslant 1.58) = 0.6785.
\end{aligned}$$

Valor com o fator de Correção:

$$\begin{aligned}
\mathbb{P}(8 \leqslant X \leqslant 15) &\simeq \mathbb{P}(8 - 0.5 \leqslant Y \leqslant 15 + 0.5) \\
&= \mathbb{P}\left(\frac{8 - 0.5 - 10}{\sqrt{10}} \leqslant \frac{Y - 10}{\sqrt{10}} \leqslant \frac{15 + 0.5 - 10}{\sqrt{10}}\right) \\
&= \mathbb{P}(-0.79 \leqslant Z \leqslant 1.74) = 0.7443066.
\end{aligned}$$

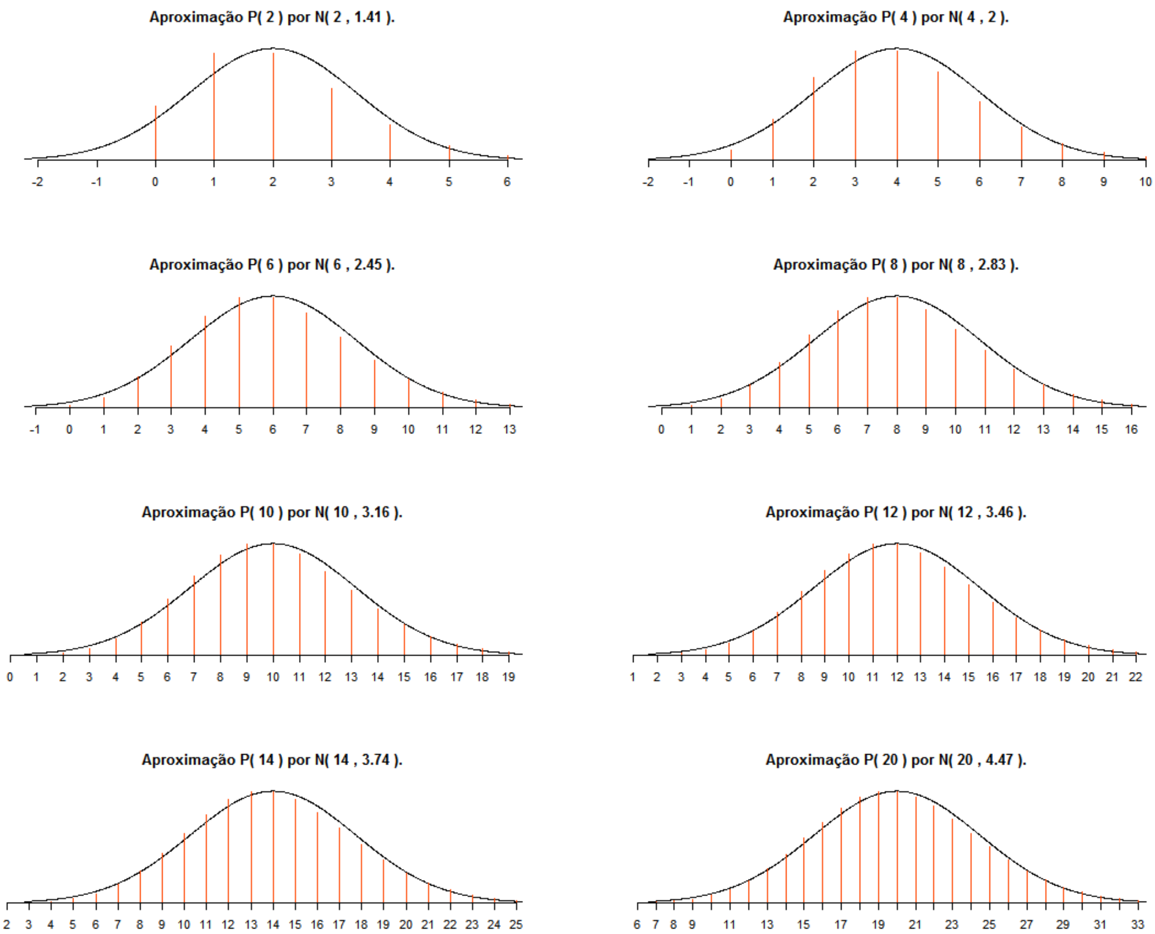


Figura 6.19: Aproximação da distribuição $P(\lambda)$ pela distribuição $N(\mu, \sigma^2)$, onde $\lambda \in \{2, 4, 6, 8, 10, 12, 14, 20\}$.

□

Exemplo 6.17. Considere que o número de partículas em uma superfície segue uma distribuição Poisson. Suponha que esperamos observar 1000 partículas por m^2. Analisamos um metro quadrado da superfície. Qual a probabilidade de observarmos entre 850 e 1050 partículas, inclusive?

Solução: Sabemos que $X \sim P(1000)$. Então podemos calcular a probabilidade de maneira exata da seguinte forma

$$\mathbb{P}(850 \leqslant X \leqslant 1050) = \sum_{x=850}^{1050} \frac{e^{-1000} 1000^x}{x!} = 0.9440.$$

Impossível de ser calculado a mão. Somente computacionalmente.

Vamos utilizar a aproximação Normal. Seja $Y \sim N(\mu, \sigma^2)$, com $\mu = \lambda$ e $\sigma^2 = \lambda$. **Valor sem o fator de Correção:**

$$\begin{aligned}
\mathbb{P}(850 \leqslant X \leqslant 1050) &\simeq \mathbb{P}(850 \leqslant Y \leqslant 1050) \\
&= \mathbb{P}\left(\frac{850 - 1000}{\sqrt{1000}} \leqslant \frac{Y - 1000}{\sqrt{1000}} \leqslant \frac{1050 - 1000}{\sqrt{1000}}\right) \\
&= \mathbb{P}(-4.74 \leqslant Z \leqslant 1.58) = 0.9431.
\end{aligned}$$

Valor com o fator de Correção:

$$\begin{aligned}
\mathbb{P}(850 \leqslant X \leqslant 1050) &\simeq \mathbb{P}(850 - 0.5 \leqslant Y \leqslant 1050 + 0.5)\\
&= \mathbb{P}\left(\frac{850 - 0.5 - 1000}{\sqrt{1000}} \leqslant \frac{Y - 1000}{\sqrt{1000}} \leqslant \frac{1050 + 0.5 - 1000}{\sqrt{1000}}\right)\\
&= \mathbb{P}(-4.76 \leqslant Z \leqslant 1.60) = 0.9449.
\end{aligned}$$

□

6.6 Modelo Gama

A distribuição gama é uma generalização da distribuição exponencial, que utiliza a função Gama, cuja definição apresentamos a seguir.

6.6.1. Função Gama

A função gama é definida pela seguinte integral:

$$\Gamma(\alpha) = \int_0^\infty x^{\alpha-1}e^{-x}dx, \quad \alpha \geqslant 1.$$

Note que o argumento da função é α que aparece no expoente da variável de integração x.

A função gama tem a seguinte propriedade recursiva: $\Gamma(\alpha+1) = \alpha\Gamma(\alpha)$. Para demonstrar esse resultado, iremos usar integração por partes.

$$\Gamma(\alpha+1) = \int_0^\infty x^{\alpha}e^{-x}dx.$$

Fazendo: $u = x^\alpha \to du = \alpha x^{\alpha-1}$ e $dv = e^{-x}dx \to v = -e^{-x}$. Portanto,

$$\begin{aligned}
\Gamma(\alpha+1) &= \int_0^\infty x^{\alpha}e^{-x}dx\\
&= -x^{\alpha-1}e^{-x}\Big|_0^\infty - \alpha\int_0^\infty -x^{\alpha}e^{-x}dx\\
&= 0 + \alpha\int_0^\infty x^{\alpha}e^{-x}dx\\
&= \alpha\Gamma(\alpha).
\end{aligned}$$

Assumindo agora, que $\alpha = n$. Então

$$\Gamma(1) = \int_0^\infty x^{1-1}e^{-x}dx = \int_0^\infty e^{-x}dx = 1$$

$$\Gamma(2) = 1 \times \Gamma(1) = 1 = 1!$$

$$\Gamma(3) = 2 \times \Gamma(2) = 2 \times 1 = 2!$$

$$\Gamma(4) = 3 \times \Gamma(3) = 3 \times 2 \times 1 = 3!$$

$$\Gamma(5) = 4 \times \Gamma(4) = 4 \times 3 \times 2 \times 1 = 4!$$

Generalizando, temos que

$$\Gamma(n) = (n-1)!$$

A seguir apresentamos o resumo de algumas propriedades da Distribuição Gama que podem ser encontradas em Magalhães (2006) e Hoel, Port e Stone (1971).

(i) $\Gamma(\alpha) = (\alpha - 1)\Gamma(\alpha - 1)$, para $\alpha \geqslant 1$.

(ii) $\Gamma(n) = (n-1)!$, onde n é um inteiro positivo.

(iii) $\Gamma(1/2) = \sqrt{\pi}$.

6.6.2. Distribuição Gama

A distribuição Gama é uma das mais gerais distribuições, pois diversas distribuições são caso particular dela como por exemplo a exponencial, a qui-quadrado, entre outras. Essa distribuição tem como suas principais aplicações à análise de tempo de vida de produtos.

A distribuição Gama também é usada para modelar valores de dados positivos que são assimétricos à direita e maiores que 0. Comumente usada em estudos de sobrevivência de confiabilidade.

A seguir definimos a função densidade de probabilidade de uma variável aleatória contínua com distribuição Gama.

Definição 6.5. Diremos que uma variável aleatória X segue o *modelo Gama*, se sua função densidade de probabilidade for dada por

$$f_X(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x} I_{(0,\infty)}(x), \tag{6.7}$$

sendo α e β dois parâmetros positivos e com $\Gamma(\alpha)$ sendo a função gama dada por

$$\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} e^{-x} dx, \quad \alpha > 0.$$

Notação: $X \sim \Gamma(\alpha, \beta)$.

Corolário 6.4. *Seja $X \sim \Gamma(\alpha, \beta)$. Então $f_X(\cdot)$ dada pela equação (6.7) satisfaz as propriedades de uma função densidade de probabilidade, ou seja, é uma função densidade de probabilidade.*

Demonstração. Seja $X \sim \Gamma(\alpha, \beta)$. Temos que verificar se $f_X(\cdot)$, dada pela equação (6.7), satisfaz as propriedades de função densidade de probabilidade.

Temos que $f_X(\cdot)$ é uma função não negativa. Além disso,

$$\int_{-\infty}^\infty f_X(x)dx = \int_0^\infty \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x} dx = \frac{\beta^\alpha}{\Gamma(\alpha)} \int_0^\infty x^{\alpha-1} e^{-\beta x} dx$$

Fazendo $y = \beta x$, então $\frac{dy}{dx} = \beta$, ou seja, $dx = \frac{dy}{\beta}$. Assim,

$$\begin{aligned}\int_{-\infty}^{\infty} f_X(x)dx &= \frac{\beta^\alpha}{\Gamma(\alpha)}\int_0^\infty x^{\alpha-1}e^{-\beta x}dx \\ &= \frac{\beta^\alpha}{\Gamma(\alpha)}\int_0^\infty \left(\frac{y}{\beta}\right)^{\alpha-1} e^{-y}\frac{dy}{\beta} \\ &= \frac{1}{\Gamma(\alpha)}\int_0^\infty y^{\alpha-1}e^{-y}dy = \frac{1}{\Gamma(\alpha)}\Gamma(\alpha) = 1\end{aligned}$$

Portanto, $f_X(\cdot)$, dada pela equação (6.7), é uma função densidade de probabilidade.

□

Resultado 6.5. Seja $X \sim \Gamma(\alpha, \beta)$. Então a função de distribuição (acumulada) da variável aleatória X é dada por:
Para $x < 0$, $F_X(x) = 0$. Para $x \geqslant 0$,

$$F_X(x) = \int_{-\infty}^{x} f_X(y)dy = \int_0^x \frac{\beta^\alpha}{\Gamma(\alpha)} y^{\alpha-1}e^{-\beta y}dy$$

Não existe uma fórmula fechada para esta distribuição.

Proposição 6.10. *Seja $X \sim \Gamma(\alpha, \beta)$, então*

(i) $\mathbb{E}(X) = \frac{\alpha}{\beta}$.

(ii) $\mathbb{E}(X^2) = \frac{(\alpha+1)\alpha}{\beta^2}$

(iii) $Var(X) = \frac{\alpha}{\beta^2}$.

Demonstração. Seja $X \sim \Gamma(\alpha, \beta)$.

(i) Cálculo de $\mathbb{E}(X)$.

$$\mathbb{E}(X) = \int_{-\infty}^{\infty} x f_X(x)dx = \int_0^\infty x\left[\frac{\beta^\alpha}{\Gamma(\alpha)}x^{\alpha-1}e^{-\beta x}\right]dx = \frac{\beta^\alpha}{\Gamma(\alpha)}\int_0^\infty x^\alpha e^{-\beta x}dx$$

Fazendo $y = \beta x$ ($x = \frac{y}{\beta}$), então $\frac{dy}{dx} = \beta$, ou seja, $dx = \frac{dy}{\beta}$. Assim,

$$\begin{aligned}\mathbb{E}(X) &= \frac{\beta^\alpha}{\Gamma(\alpha)}\int_0^\infty x^\alpha e^{-\beta x}dx = \frac{\beta^\alpha}{\Gamma(\alpha)}\int_0^\infty \left(\frac{y}{\beta}\right)^\alpha e^{-y}\frac{dy}{\beta} \\ &= \frac{\beta^\alpha}{\Gamma(\alpha)}\int_0^\infty \frac{y^\alpha}{\beta^{\alpha+1}}e^{-y}dy = \frac{1}{\beta\Gamma(\alpha)}\int_0^\infty y^\alpha e^{-y}dy \\ &= \frac{1}{\beta\Gamma(\alpha)}\int_0^\infty y^{(\alpha+1)-1}e^{-y}dy = \frac{1}{\beta\Gamma(\alpha)}\Gamma(\alpha+1) = \frac{1}{\beta\Gamma(\alpha)}\alpha\Gamma(\alpha) = \frac{\alpha}{\beta},\end{aligned}$$

pois

$$\Gamma(\alpha) = \int_0^\infty x^{\alpha-1}e^{-x}dx, \quad \alpha > 0.$$

(ii) Cálculo de $\mathbb{E}(X^2)$.

$$\mathbb{E}(X^2) = \int_{-\infty}^{\infty} x^2 f_X(x)dx = \int_0^{\infty} x^2 \left[\frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}\right] dx = \frac{\beta^\alpha}{\Gamma(\alpha)} \int_0^{\infty} x^{\alpha+1} e^{-\beta x} dx$$

Fazendo $y = \beta x$ ($x = \frac{y}{\beta}$), então $\frac{dy}{dx} = \beta$, ou seja, $dx = \frac{dy}{\beta}$. Assim,

$$\begin{aligned}
\mathbb{E}(X) &= \frac{\beta^\alpha}{\Gamma(\alpha)} \int_0^{\infty} x^{\alpha+1} e^{-\beta x} dx = \frac{\beta^\alpha}{\Gamma(\alpha)} \int_0^{\infty} \left(\frac{y}{\beta}\right)^{\alpha+1} e^{-y} \frac{dy}{\beta} \\
&= \frac{\beta^\alpha}{\Gamma(\alpha)} \int_0^{\infty} \frac{y^{\alpha+1}}{\beta^{\alpha+2}} e^{-y} dy = \frac{1}{\beta^2\Gamma(\alpha)} \int_0^{\infty} y^{\alpha+1} e^{-y} dy \\
&= \frac{1}{\beta^2\Gamma(\alpha)} \int_0^{\infty} y^{(\alpha+2)-1} e^{-y} dy = \frac{1}{\beta^2\Gamma(\alpha)} \Gamma(\alpha+2) \\
&= \frac{1}{\beta^2\Gamma(\alpha)} (\alpha+1)\Gamma(\alpha+1) = \frac{1}{\beta^2\Gamma(\alpha)} (\alpha+1)\alpha\Gamma(\alpha) = \frac{(\alpha+1)\alpha}{\beta^2},
\end{aligned}$$

pois

$$\Gamma(\alpha) = \int_0^{\infty} x^{\alpha-1} e^{-x} dx, \quad \alpha > 0.$$

(iii) Pela definição de $\text{Var}(X)$, temos que

$$\text{Var}(X) = \mathbb{E}(X^2) - [\mathbb{E}(X)]^2 = \frac{(\alpha+1)\alpha}{\beta^2} - \left[\frac{\alpha}{\beta}\right]^2 = \frac{\alpha}{\beta^2}.$$

$\square$

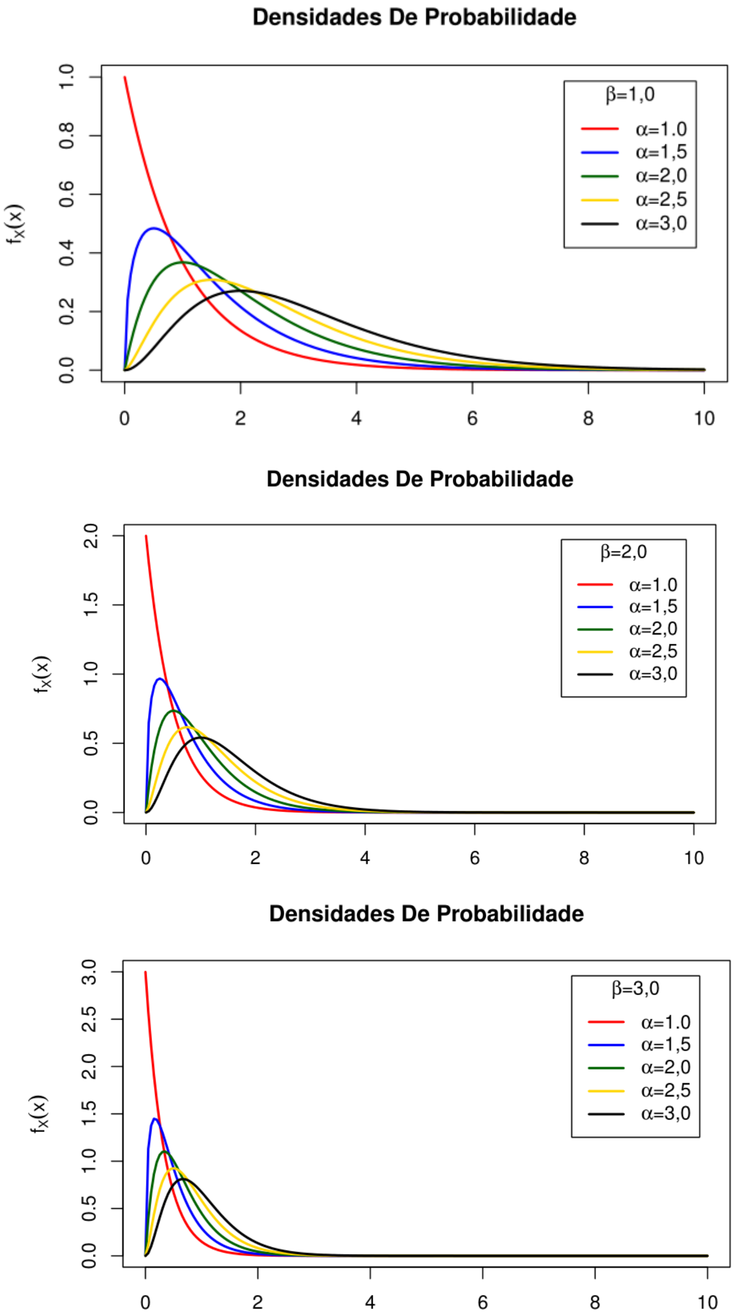


Figura 6.20: Distribuição $\Gamma(\alpha, \beta)$: Função densidade de probabilidade. Parâmetros $\alpha \in \{1.0, 1.5, 2.0, 2.5, 3.0\}$ e $\beta \in \{1.0, 2.0, 3.0\}$.

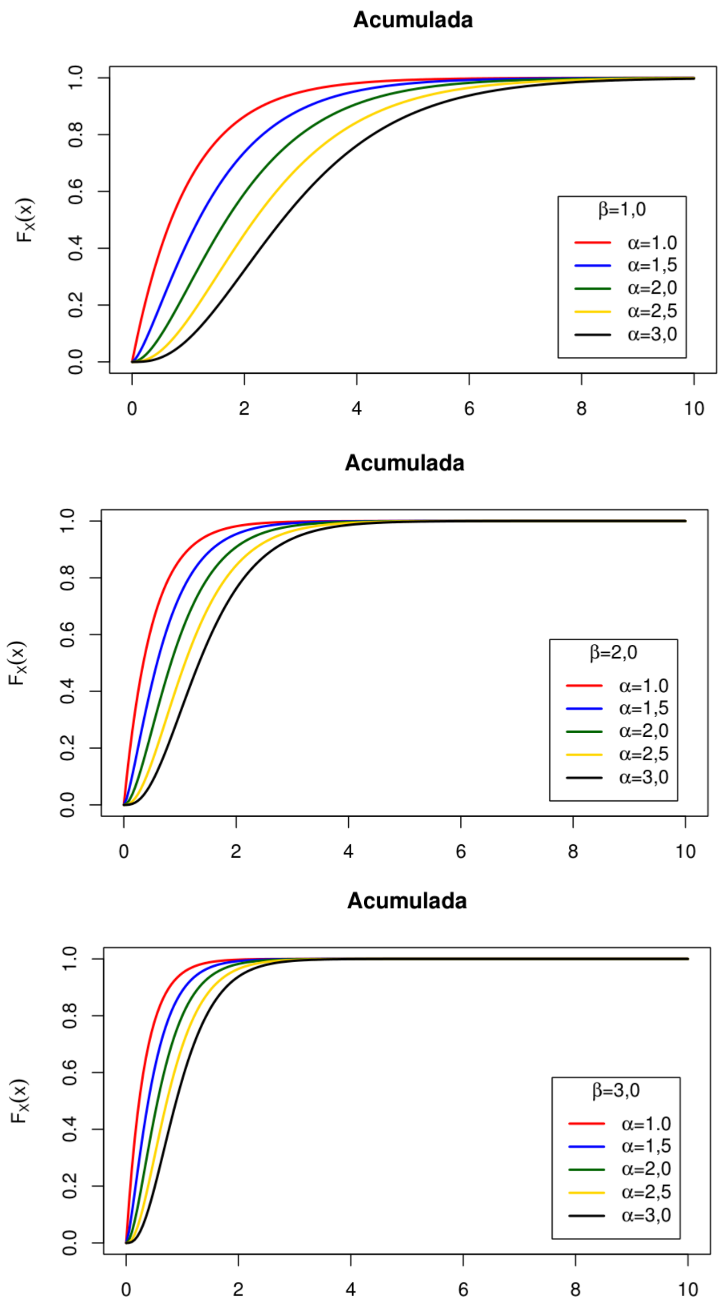


Figura 6.21: Distribuição $\Gamma(\alpha, \beta)$: Função Distribuição de probabilidade. Parâmetros $\alpha \in \{1.0, 1.5, 2.0, 2.5, 3.0\}$ e $\beta \in \{1.0, 2.0, 3.0\}$.

Exemplo 6.18. A renda doméstica mensal num certo bairro do Rio de Janeiro é uma variável aleatória com distribuição Gama com média R$ 2000 e desvio padrão R$ 400.

(i) Ache os parâmetros α e β desta densidade.

(ii) Calcule a probabilidade da renda média mensal de um domicílio exceder R$2500.

(iii) Qual a mediana desta distribuição Gama?

Resolução:

(i) Ache os parâmetros α e β desta densidade.

Temos que $\mathbb{E}(X) = \frac{\alpha}{\beta} = 2000 \quad \to \quad \alpha = 2000\beta$

Por outro lado, temos que

$$\sqrt{\text{Var}(X)} = \sqrt{\frac{\alpha}{\beta^2}} = 400 \quad \to \quad \beta = \frac{1}{80}.$$

Assim, $\alpha = 2000\beta = \frac{2000}{80} = 25$.

Portanto $X \sim \Gamma(25, \frac{1}{80})$.

(ii) Calcule a probabilidade da renda média mensal de um domicílio exceder R$2500.

Queremos $\mathbb{P}(X > 2500)$.

$$\mathbb{P}(X > 2500) = 1 - \mathbb{P}(X \leqslant 2500) = 1 - \int_0^{2500} \frac{\frac{1}{80}^{25}}{\Gamma(25)} x^{25-1} e^{-\frac{1}{80}x} dx.$$

Vamos utilizar a função *pgamma* do software R, ou seja,

$$\mathbb{P}(X > 2500) = 1 - \mathbb{P}(X \leqslant 2500) = 1 - \text{pgamma(q=2500, shape=25, rate} = 1/80) = 0.1104299,$$

noindent onde shape=α, rate=β

(iii) Qual os quartis desta distribuição Gama?

Para responder a está pergunta vamos utilizar a função *qgamma* do software R Core Team (2019).

$$\mathbb{P}(X \leqslant \gamma_{\frac{1}{4}}) = \frac{1}{4} \quad \to \quad \text{qgamma(p=0.25, shape=25, rate} = 1/80) = 1717.683,$$

$$\mathbb{P}(X \leqslant \gamma_{\frac{1}{2}}) = \frac{1}{2} \quad \to \quad \text{qgamma(p=0.5, shape=25, rate} = 1/80) = 1973.397,$$

$$\mathbb{P}(X \leqslant \gamma_{\frac{3}{4}}) = \frac{3}{4} \quad \to \quad \text{qgamma(p=0.75, shape=25, rate} = 1/80) = 2253.344.$$

□

Observação 6.3. Principais casos da Distribuição Gama.

(i) Chamaremos de distribuição Gama padrão quando $\beta = 1$, na distribuição Gama. Dessa forma, a função densidade de probabilidade de uma variável aleatória X com distribuição padrão é dada por

$$f_X(x) = \frac{x^{\alpha-1}e^{-x}}{\Gamma(\alpha)} I_{(0,\infty)}(x), \quad \alpha > 0.$$

(ii) Quando $\alpha = 1$ e $\beta > 0$ na distribuição Gama, temos a distribuição Exponencial ($X \sim \text{Exp}(\lambda)$).

(iii) Quando $\alpha = \frac{n}{2}$, $n > 0$ inteiro e $\beta = \frac{1}{2}$ na distribuição Gama, temos a distribuição Qui-quadrado com n graus de liberdade ($X \sim \chi^2_n$).

6.7 Modelo Chi-Quadrado χ^2

Um caso importante da distribuição $\Gamma(\alpha, \beta)$, $\alpha, \beta > 0$ é obtido, se tomarmos $\alpha = n/2$ e $\beta = 1/2$, onde n é um inteiro positivo. Obteremos uma família de distribuições de um parâmetro.
Se $X \sim \Gamma(\alpha, \beta)$,

$$f_X(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x} I_{(0,\infty)}(x).$$

Tomando $\alpha = n/2$ e $\beta = 1/2$, temos para $x > 0$,

$$\begin{aligned} f_Z(z) &= \frac{(1/2)^{n/2}}{\Gamma(n/2)} z^{\frac{n}{2}-1} e^{-\frac{z}{2}} I_{(0,\infty)}(z) \\ &= \frac{1}{2^{n/2}\Gamma(n/2)} z^{\frac{n}{2}-1} e^{-\frac{z}{2}} I_{(0,\infty)}(z). \end{aligned}$$

A variável aleatória Z, que tem função densidade de probabilidade $f_Z(z)$, é chamada de chi-quadrado, com n graus de liberdade.

Notação: $Z \sim \chi^2_n$, ou $Z \sim \chi^2_{(n)}$.

Por ser um caso particular da distribuição $\Gamma(\alpha, \beta)$, a distribuição χ^2_n, tem esperança e variância dados por

$$\begin{cases} X \sim \Gamma(\alpha, \beta) \\ \mathbb{E}(X) = \frac{\alpha}{\beta} \\ \text{Var}(X) = \frac{\alpha}{\beta^2} \end{cases} \longrightarrow \alpha = \frac{n}{2}, \beta = \frac{1}{2} \longrightarrow \begin{cases} Z \sim \chi^2_n \\ \mathbb{E}(Z) = n \\ \text{Var}(Z) = 2n. \end{cases}$$

Para $n = 1$, temos

$$f_Z(z) = \frac{1}{2^{1/2}\Gamma(1/2)} z^{\frac{1}{2}-1} e^{-\frac{z}{2}} I_{(0,\infty)}(z) = \frac{1}{\sqrt{2\pi}} z^{\frac{-1}{2}} e^{-\frac{z}{2}} I_{(0,\infty)}(z).$$

Para $n = 2$, temos

$$f_Z(z) = \frac{1}{2\Gamma(1)} z^0 e^{-\frac{z}{2}} I_{(0,\infty)}(z) = \frac{1}{2} e^{-\frac{z}{2}} I_{(0,\infty)}(z),$$

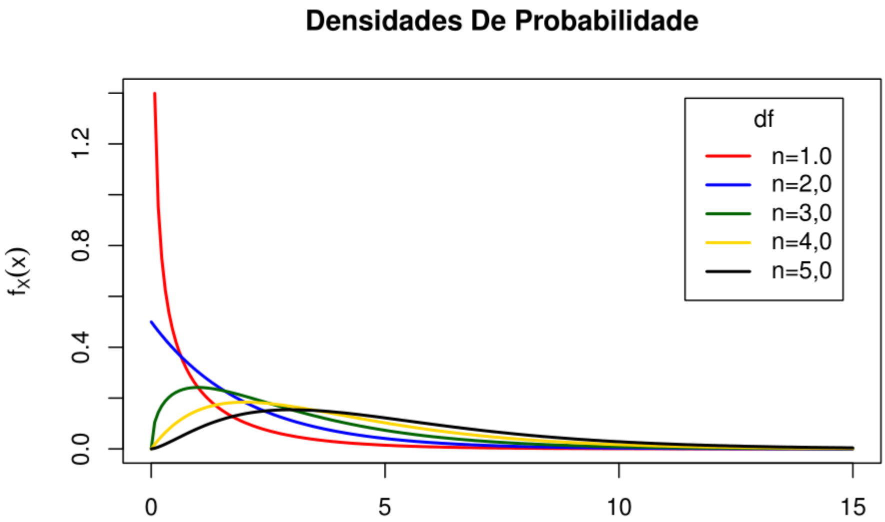


Figura 6.22: Distribuição χ_n^2: função densidade de probabilidade.

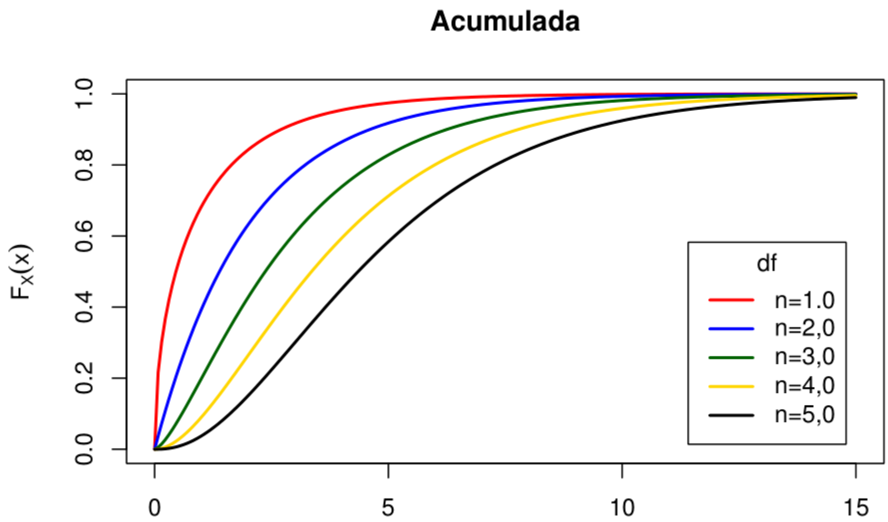


Figura 6.23: Distribuição χ_n^2: função de distribuição acumulada.

a qual é a f.d.p da distribuição exponencial.

Para $n > 30$, podemos utilizar uma aproximação normal à distribuição chi-quadrado. Especificamente, temos o seguinte resultado: Se $Z \sim \chi_n^2$, com n graus de liberdade, então a variável aleatória

$$Y = \sqrt{2Z} - \sqrt{2n-1} \sim N(0,1).$$

Exemplo 6.19. Consultando a tabela temos que, para $n = 30$, $\mathbb{P}(Z > 40.25) = 0.1$.

Utilizando a relação acima, temos que $z = \sqrt{2 \times 40,256} - \sqrt{2 \times 30 - 1} = 1,291$.

Portanto, $\mathbb{P}(Y > 1,291) = 0.099$, onde $Y \sim N(0,1)$, que resulta em uma boa aproximação.

$\square$

Resultado 6.6. Considere $Z \sim N(0,1)$ e a variável aleatória $Y = Z^2$. Qual a distribuição de Y.

$$F_Y(y) = \mathbb{P}(Y = y) = \mathbb{P}(Z^2 < y) = \mathbb{P}(-\sqrt{y} < Z < \sqrt{y}) = F_Z(\sqrt{y}) - F_Z(-\sqrt{y}).$$

Derivando a expressão acima temos

$$\begin{aligned} f_Y(y) &= F'_Y(y) = \frac{1}{2\sqrt{y}}\left[F'_Z(\sqrt{y}) - F'_Z(-\sqrt{y})\right] = \frac{1}{2\sqrt{y}}\left[f_Z(\sqrt{y}) - f_Z(-\sqrt{y})\right] \\ &= \frac{1}{2\sqrt{y}}\left[\frac{1}{\sqrt{2\pi}}e^{-y/2} - \frac{1}{\sqrt{2\pi}}e^{-y/2})\right] = \frac{1}{\sqrt{2\pi}}y^{-1/2}e^{-y/2} \end{aligned}$$

Logo, $Y \sim \chi^2_1$.

□

Exemplo 6.20. Seja $X \sim \chi^2_{15}$. Encontre o valor de $\chi^2_{0.1}$ tal que $\mathbb{P}(\chi^2_{15} \geqslant \chi^2_{0.1}) = 0.10$

Solução: Iremos utilizar a função qchisq para encontrar tal valor.

$$\mathbb{P}(\chi^2_{15} \geqslant \chi^2_{0.1}) = 0.10 \quad \rightarrow \quad \chi^2_{0.1} = \text{qchisq(p=0.10, df=15, lower.tail = FALSE)} = 22.30713$$

ou

$$\mathbb{P}(\chi^2_{15} \leqslant \chi^2_{0.1}) = 0.90 \quad \rightarrow \quad \chi^2_{0.1} = \text{qchisq(p=0.90, df=15, lower.tail = TRUE)} = 22.30713$$

□

Exemplo 6.21. Seja $X \sim \chi^2_{15}$. Encontre o valor de χ^2_α tal que $\mathbb{P}(\chi^2_{15} \geqslant \chi^2_\alpha) = \alpha$

(i) $\alpha = 0.10$

(ii) $\alpha = 0.95$

Solução: Iremos utilizar a função qchisq para encontrar tal valor. Veja a Figura 6.24 para ilustração.

(i) Considerando $\alpha = 0.10$.

$$\mathbb{P}(\chi^2_{15} \geqslant \chi^2_{0.1}) = 0.10 \quad \rightarrow \quad \chi^2_{0.1} = \text{qchisq(p=0.10, df=15, lower.tail = FALSE)} = 22.30713$$

(ii) Considerando $\alpha = 0.95$

$$\mathbb{P}(\chi^2_{15} \geqslant \chi^2_{0.95}) = 0.95 \quad \rightarrow \quad \chi^2_{0.95} = \text{qchisq(p=0.95, df=15, lower.tail = FALSE)} = 7.260944$$

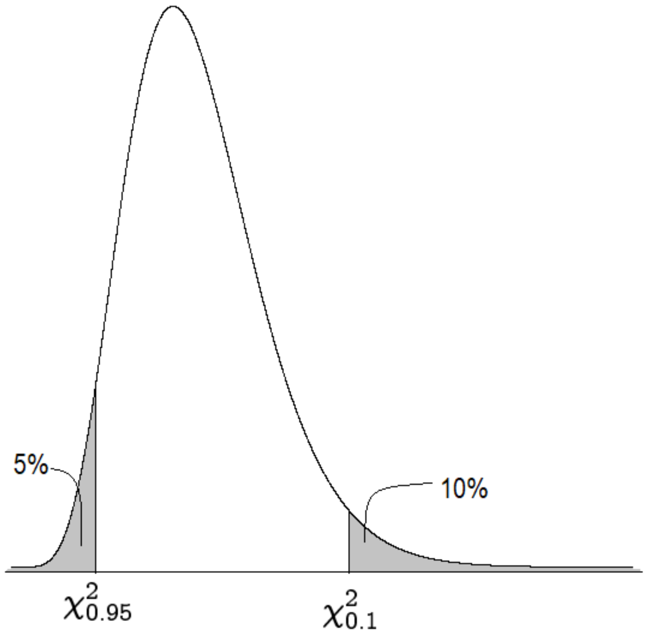


Figura 6.24: Distribuição χ^2_{15}: valores críticos $\chi^2_{0.95}$ e $\chi^2_{0.10}$

$\square$

6.8 Distribuição t de student

Definição 6.6. Sejam $X \sim N(0,1)$ e $Y \sim \chi^2_n$, X e Y são variáveis aleatórias independentes. Então, a variável aleatória

$$T = \frac{X}{\sqrt{Y/n}} = \frac{\sqrt{n}X}{\sqrt{Y}},$$

é dita ter distribuição t de student com n graus de liberdade.
A função densidade de probabilidade da variável aleatória T é dada por

$$f_T(t) = \frac{\Gamma(\frac{n+1}{2})}{\Gamma(n/2)\sqrt{n\pi}}\left(1 - \frac{t^2}{n}\right)^{-(n+1)/2}, \quad t \in \mathbb{R}.$$

Notação: $T \sim t_n$ ou $T \sim t(n)$.

Observação 6.4. 1) Caso particular: se $n = 1$, temos a distribuição Cauchy Padrão com $\alpha = 0$ e $\beta = 1$, onde

$$f_T(t) = \frac{1}{\pi}(1 + t^2)^{-1}, \quad t \in \mathbb{R}.$$

2) Para n grande, a distribuição t-student se aproxima da distribuição normal.

3) A função densidade de probabilidade da distribuição t-student é simétrica em $t = 0$ e $\lim_{t\to\infty} f_T(t) = 0 = \lim_{t\to-\infty} f_T(t)$

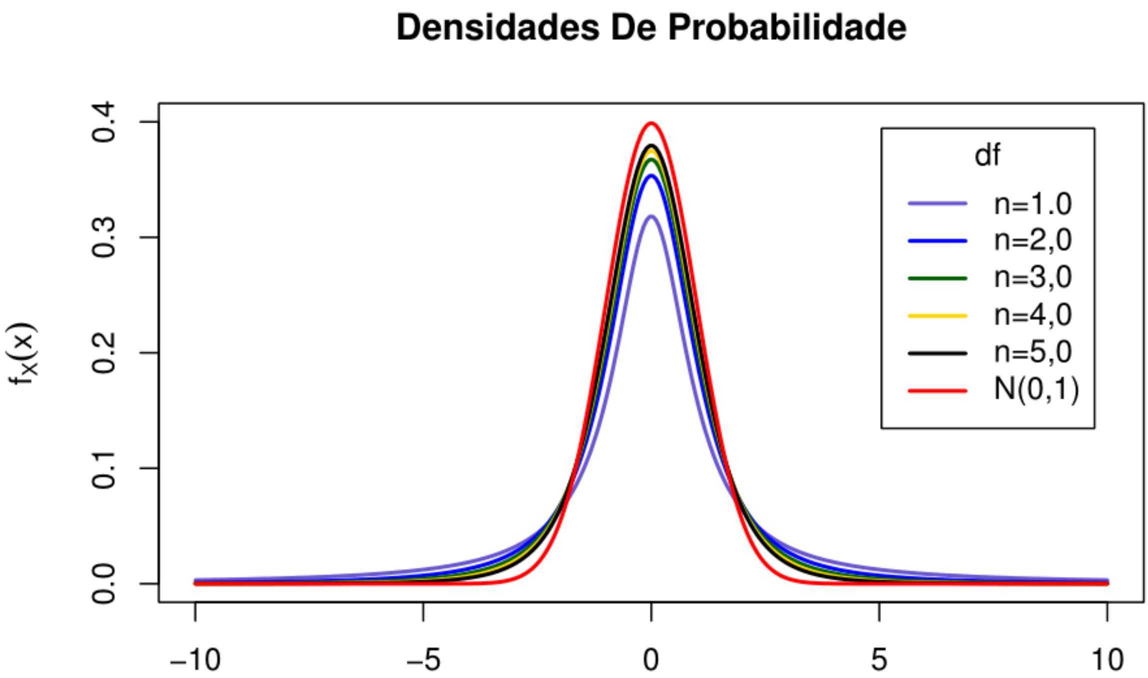


Figura 6.25: Distribuição t_n: Função densidade de probabilidade.

Proposição 6.11. *Seja $X \sim t_n$, então*

(i) $\mathbb{E}(X) = 0$, *se* $n > 1$.

(ii) $\text{Var}(X) = \frac{n}{n-2}$, *se* $n > 2$.

Exemplo 6.22. Seja $X \sim t_5$. Encontre o valor de t_c tal que $\mathbb{P}(-t_{n,c} \leqslant T \leqslant t_{n,c}) = 0.9$

Solução: Iremos utilizar a função qt para encontrar tal valor.

$$\mathbb{P}(T \geqslant t_{5,c}) = 0.05 \quad \to \quad t_{5,c} = \text{qt(p=0.05, df=5, lower.tail = FALSE)} = 2.015048$$

Como a distribuição t_n é simétrica em relação a zero, $-t_{5,c} = -2.015048$.

Sabemos que quando $n \to \infty$, a distribuição t_n se aproxima da distribuição normal padrão. Considere $n = 200$, então

$$\mathbb{P}(T \geqslant t_{200,c}) = 0.05 \quad \to \quad t_{200,c} = \text{qt(p=0.05, df=200, lower.tail = FALSE)} = 1.652508$$

e

$$\mathbb{P}(Z \geqslant z_{0.05}) = 0.05 \quad \to \quad z_{0.05} = \text{qnorm(p=0.05,lower.tail = FALSE)} = 1.644854.$$

$$\mathbb{P}(T \geqslant t_{10000,c}) = 0.05 \quad \to \quad t_{10000,c} = \text{qt(p=0.05, df=10000, lower.tail = FALSE)} = 1.645006$$

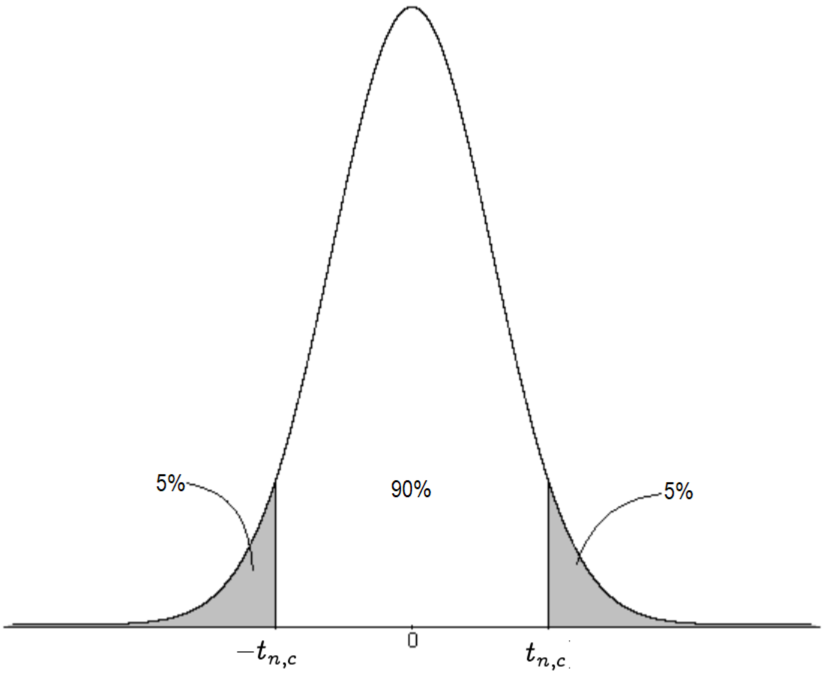


Figura 6.26: Distribuição t_n: $t_{n,c}$ - valor crítico.

□

6.9 Modelo F-Snedecor

Definição 6.7. Sejam X e Y variáveis aleatórias independentes com distribuição χ^2_m e χ^2_n, respectivamente. A variável aleatória

$$F = \frac{X/m}{Y/n} = \frac{n}{m}\frac{X}{Y},$$

é dita ter distribuição F-Snedecor com (m, n) graus de liberdade.
A função densidade de probabilidade da variável aleatória F é dada por

$$f_F(x) = \frac{\Gamma(\frac{m+n}{2})}{\Gamma(\frac{m}{2})\Gamma(\frac{n}{2})}\left(\frac{m}{n}\right)\left(\frac{m}{n}x\right)^{\frac{m}{2}-1}\left(1+\frac{m}{n}x\right)^{-(\frac{m+n}{2})} I_{(0,\infty)}(x).$$

Notação: $F \sim \mathrm{F}_{m,n}$ ou $F \sim \mathrm{F}(m, n)$.

Observação 6.5. Propriedades:

(i) Se $X \sim \mathrm{F}(m, n)$ então, $\frac{1}{X} \sim \mathrm{F}(n, m)$.

(ii) Se $Z \sim \mathcal{C}(0, 1) = t_1$, então $Z^2 \sim F(1, 1)$.

(iii) $\mathbb{E}(\mathrm{F}) = \frac{n}{n-2}$, para $n > 2$.

(iv) $\mathrm{Var}(\mathrm{F}) = \frac{n^2(2m+2n-4)}{n(n-2)^2(n-4)}$, para $n > 4$.

(v) Se $X \sim t_n$, então $X^2 \sim \mathrm{F}(1, n)$.

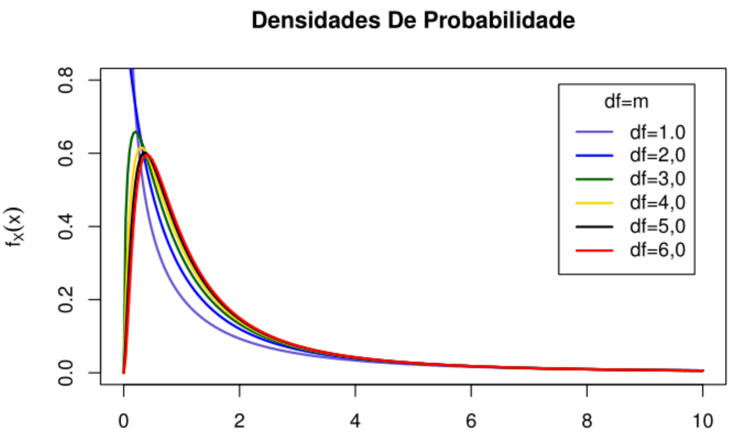


Figura 6.27: Função Densidade de Probabilidade da distribuição $\mathrm{F}(m, n)$: $m \in \{1.0, \cdots, 6.0\}$ e $n = 3, 0$.

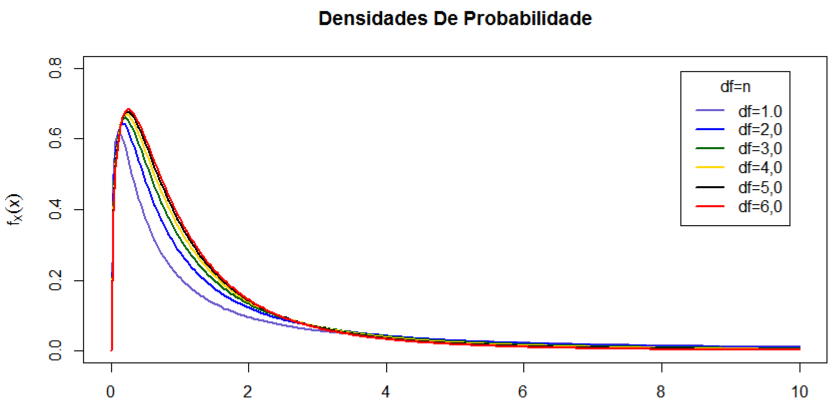


Figura 6.28: Função Densidade de Probabilidade da distribuição $\mathrm{F}(m, n)$: $m = 3, 0$ e $n \in \{1.0, \cdots, 6.0\}$.

Exemplo 6.23. Seja $X \sim \mathrm{F}(m, n)$. Encontre o valor de $F_{m,n,c}$ tal que $\mathbb{P}(X \geqslant F_{m,n,c}) = 0.05$. Considerem $m = 2$ e $n = 5$.

Solução: Iremos utilizar a função qt para encontrar tal valor.

$$\mathbb{P}(X \geqslant F_{2,5,c}) = 0.05 \quad \rightarrow \quad F_{2,5,c} = \text{qf(p=0.05, df1=2, df2=5, lower.tail=FALSE)} = 5.786135.$$

$\square$

6.10 Modelo Weibull

A distribuição Weibull foi proposta originalmente por W. Weibull (1954) em estudos relacionados ao tempo de falha devido a fadiga de metais. Ela é frequentemente usada para descrever o tempo de vida de produtos industriais. A sua popularidade em aplicações práticas deve-se ao fato dela apresentar uma grande variedade de formas, todas com uma propriedade básica: a sua função de taxa de falha é monótona. Isto é, ou ela é crescente ou decrescente ou constante. Ela descreve adequadamente a vida de mananciais, componentes eletrônicos, cerâmicas, capacitores e dielétricos.

Definição 6.8. Uma variável aleatória X tem distribuição Weibull se tiver função densidade de probabilidade dada por:

$$f_X(x) = \frac{\alpha}{\beta^\alpha} x^{\alpha-1} \exp\left[-\left(\frac{x}{\beta}\right)^\alpha\right] I_{[0,\infty)}(x). \tag{6.8}$$

Notação: $X \sim \mathrm{W}(\alpha, \beta)$.

Corolário 6.5. *Seja $X \sim W(\alpha,\beta)$. Então $f_X(\cdot)$ dada pela equação (6.8) satisfaz as propriedades de uma função densidade de probabilidade, ou seja, é uma função densidade de probabilidade.*

Demonstração. Seja $X \sim \mathrm{W}(\alpha, \beta)$. Temos que verificar se $f_X(\cdot)$, dada pela equação (6.8), satisfaz as propriedades de função densidade de probabilidade.

Temos que $f_X(\cdot)$ é uma função não negativa. Além disso,

$$\int_{-\infty}^{\infty} f_X(x)dx = \int_0^\infty \frac{\alpha}{\beta^\alpha} x^{\alpha-1} \exp\left[-\left(\frac{x}{\beta}\right)^\alpha\right] dx = \frac{\alpha}{\beta^\alpha} \int_0^\infty x^{\alpha-1} \exp\left[-\left(\frac{x}{\beta}\right)^\alpha\right] dx$$

Fazendo $y = \left(\frac{x}{\beta}\right)^\alpha$ $(x = \beta y^{\frac{1}{\alpha}})$, então $\frac{dy}{dx} = \frac{\alpha}{\beta^\alpha} x^{\alpha-1}$, ou seja, $dx = \frac{\beta^\alpha}{\alpha x^{\alpha-1}} dy$. Assim,

$$\begin{aligned}
\int_{-\infty}^{\infty} f_X(x)dx &= \frac{\alpha}{\beta^\alpha} \int_0^\infty x^{\alpha-1} \exp\left[-\left(\frac{x}{\beta}\right)^\alpha\right] dx \\
&= \frac{\alpha}{\beta^\alpha} \int_0^\infty x^{\alpha-1} \exp(-y) \frac{\beta^\alpha}{\alpha x^{\alpha-1}} dy \\
&= \int_0^\infty e^{-y} dy = 1
\end{aligned}$$

Portanto, $f_X(\cdot)$, dada pela equação (6.8), é uma função densidade de probabilidade.

□

Resultado 6.7. Seja $X \sim \mathrm{W}(\alpha, \beta)$. Então sua função de distribuição (ou acumulada) é dada por

$$F(x) = \begin{cases} 0, & \text{se } x < 0 \\ 1 - \exp\left[-\left(\frac{x}{\beta}\right)^\alpha\right], & \text{se } x \geqslant 0 \end{cases}$$

Demonstração. Seja $X \sim \mathrm{W}(\alpha, \beta)$. Considerando $x < 0$, temos que

$$F_X(x) = \int_{-\infty}^{x} f_X(y)dy = 0$$

Considerando $x \leqslant 0$, temos que

$$F_X(x) = \int_{-\infty}^{x} f_X(y)dy = \int_{0}^{x} \frac{\alpha}{\beta^\alpha} y^{\alpha-1} \exp\left[-\left(\frac{y}{\beta}\right)^\alpha\right] dy$$

Fazendo $z = \left(\frac{y}{\beta}\right)^\alpha$ $(y = \beta z^{\frac{1}{\alpha}})$, então $\frac{dz}{dy} = \frac{\alpha}{\beta^\alpha} y^{\alpha-1}$, ou seja, $dy = \frac{\beta^\alpha}{\alpha y^{\alpha-1}} dz$. Assim,

$$\begin{aligned} y = 0 \quad &\rightarrow \quad z = 0 \\ y = x \quad &\rightarrow \quad z = \left(\frac{x}{\beta}\right)^\alpha \end{aligned}$$

e

$$\begin{aligned} F_X(x) &= \int_{0}^{x} \frac{\alpha}{\beta^\alpha} y^{\alpha-1} \exp\left[-\left(\frac{y}{\beta}\right)^\alpha\right] dy \\ &= \int_{0}^{\left(\frac{x}{\beta}\right)^\alpha} \frac{\alpha}{\beta^\alpha} y^{\alpha-1} \exp(-z) \frac{\beta^\alpha}{\alpha y^{\alpha-1}} dz = \int_{0}^{\left(\frac{x}{\beta}\right)^\alpha} \exp(-z) dz \\ &= -e^{-z}\Big|_{0}^{\left(\frac{x}{\beta}\right)^\alpha} = 1 - \exp\left[-\left(\frac{x}{\beta}\right)^\alpha\right] \end{aligned}$$

Portanto,

$$F(x) = \begin{cases} 0, & \text{se } x < 0 \\ 1 - \exp\left[-\left(\frac{x}{\beta}\right)^\alpha\right], & \text{se } x \geqslant 0 \end{cases}$$

□

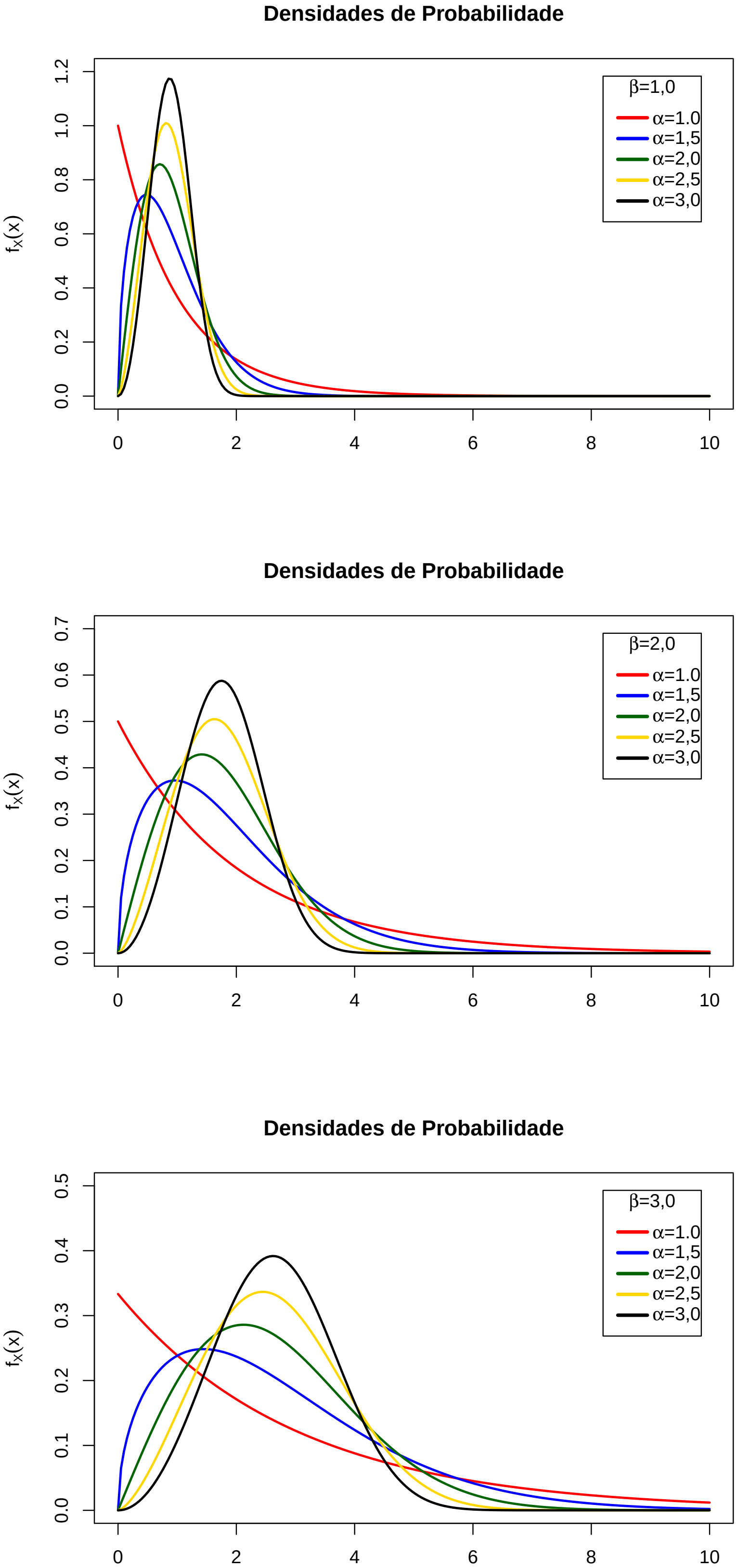


Figura 6.29: Distribuição $W(\alpha, \beta)$: Função densidade de probabilidade. Parâmetros $\alpha \in \{1.0, 1.5, 2.0, 2.5, 3.0\}$ e $\beta \in \{1.0, 2.0, 3.0\}$.

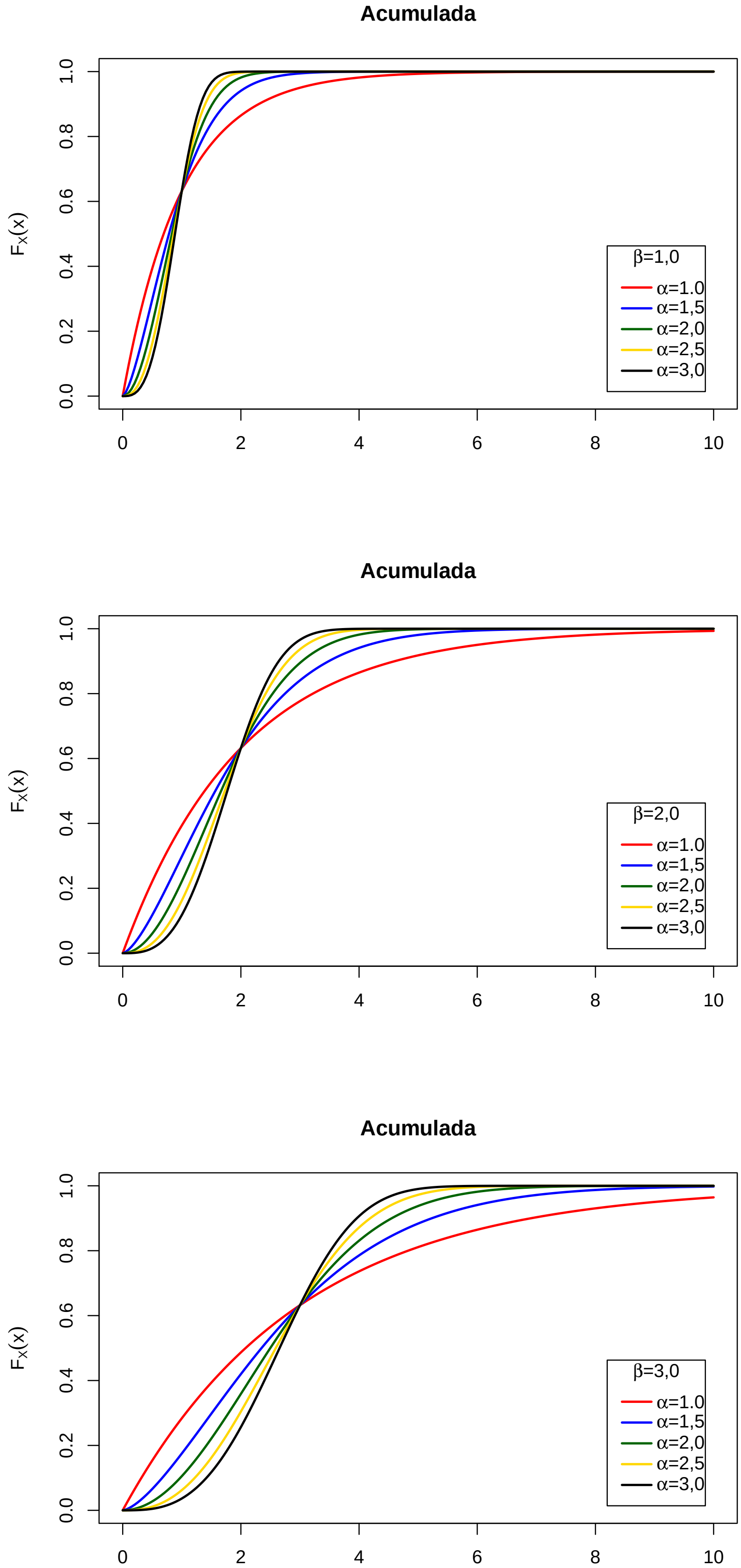


Figura 6.30: Distribuição W(α, β): Função Distribuição de probabilidade. Parâmetros $\alpha \in \{1.0, 1.5, 2.0, 2.5, 3.0\}$ e $\beta \in \{1.0, 2.0, 3.0\}$.

Proposição 6.12. *Seja* $X \sim \mathrm{W}(\alpha, \beta)$. *Então*

(i) $\mathbb{E}(X^k) = \beta^k \Gamma\left(\frac{k+\alpha}{\alpha}\right)$.

(ii) $\mathrm{Var}(X) = \beta^2 \left\{\Gamma\left(\frac{2+\alpha}{\alpha}\right) - \left[\Gamma\left(\frac{1+\alpha}{\alpha}\right)\right]^2\right\}$.

Demonstração. Seja $X \sim \mathrm{W}(\alpha, \beta)$.

i) Iniciamos calculando $\mathbb{E}(X^k)$. Assim,

$$\mathbb{E}(X^k) = \int_0^\infty x^k \left[\frac{\alpha}{\beta^\alpha} x^{\alpha-1} \exp\left\{-\left(\frac{x}{\beta}\right)^\alpha\right\}\right] dx.$$

Fazendo $y = \frac{x}{\beta}$ ($x = \beta y$), então $\frac{dy}{dx} = \frac{1}{\beta}$, ou seja, $dx = \beta dy$. Logo,

$$\begin{aligned}
\mathbb{E}(X^k) &= \int_0^\infty x^k \left[\frac{\alpha}{\beta^\alpha} x^{\alpha-1} \exp\left\{-\left(\frac{x}{\beta}\right)^\alpha\right\}\right] dx \\
&= \int_0^\infty (\beta y)^k \left[\frac{\alpha}{\beta^\alpha} (\beta y)^{\alpha-1} \exp\left\{-\left(\frac{\beta y}{\beta}\right)^\alpha\right\}\right] \beta dy. \\
&= \beta^k \int_0^\infty y^k \left[\alpha y^{\alpha-1} \exp\{-y^\alpha\}\right] dy \\
&= \beta^k \int_0^\infty \alpha y^{k+\alpha-1} e^{-y^\alpha} dy
\end{aligned}$$

Fazendo uma segunda mudança de variável $u = y^\alpha$ ($y = u^{\frac{1}{\alpha}}$), $\frac{du}{dy} = \alpha y^{\alpha-1}$, ou seja, $du = \alpha y^{\alpha-1} dy$. Logo,

$$\begin{aligned}
\mathbb{E}(X^k) &= \beta^k \int_0^\infty \alpha y^k y^{\alpha-1} e^{-y^\alpha} dy \\
&= \beta^k \int_0^\infty \left(u^{\frac{1}{\alpha}}\right)^k e^{-y} du = \beta^k \int_0^\infty u^{\frac{k}{\alpha}} e^{-y} du \\
&= \beta^k \int_0^\infty u^{\frac{k}{\alpha}+1-1} e^{-y} du = \beta^k \int_0^\infty u^{\frac{k+\alpha}{\alpha}-1} e^{-y} du = \beta^k \Gamma\left(\frac{k+\alpha}{\alpha}\right),
\end{aligned}$$

pois,

$$\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} e^{-x} dx, \quad \alpha > 0.$$

ii) Pela definição de $\mathrm{Var}(X) = \mathbb{E}(X^2) - [\mathbb{E}(X)]^2$, temos que calcular

$$\mathbb{E}(X) = \beta \Gamma\left(\frac{1+\alpha}{\alpha}\right) \quad \text{e} \quad \mathbb{E}(X^2) = \beta^2 \Gamma\left(\frac{2+\alpha}{\alpha}\right)$$

Portanto,

$$\begin{aligned}\text{Var}(X) &= \mathbb{E}(X^2) - [\mathbb{E}(X)]^2 = \beta^2\Gamma\left(\frac{2+\alpha}{\alpha}\right) - \left[\beta\Gamma\left(\frac{1+\alpha}{\alpha}\right)\right]^2 \\ &= \beta^2\left\{\Gamma\left(\frac{2+\alpha}{\alpha}\right) - \left[\Gamma\left(\frac{1+\alpha}{\alpha}\right)\right]^2\right\}.\end{aligned}$$

□

Exemplo 6.24. Nos últimos anos, a distribuição Weibull tem sido usada para modelar emissões de poluentes de vários motores. Assuma X como o valor da emissão de NO_x (g/gal) a partir de certo tipo de motor de quatro tempos selecionado aleatoriamente e suponha que X possua distribuição Weibull com $alpha = 2$ e $\beta = 10$ (sugerido pelas informações do artigo "*Qualification of variability and uncertainty in lawn and garden equipament NO_x and total hydrocarbon emission factor", (J. of the Air and Waste Management Assoc., 2002: 435-448).*

$$\mathbb{P}(X \leqslant 10) = 1 - \exp\left\{-\left(\frac{10}{10}\right)^2\right\} = 1 - e^{-1} = 0,632.$$

Da mesma forma, $\mathbb{P}(X \leqslant 25) = 0,998$, portanto a distribuição é concentrada quase que totalmente nos valores entre 0 e 25. O valor de c que separa 5% de todos os motores que têm as maiores quantidades de emissões de NO_x dos outros 95% satisfaz

$$0,95 = \mathbb{P}(X \leqslant c) = 1 - \exp\left\{-\left(\frac{c}{10}\right)^2\right\}$$

Isolando o termo exponencial de um lado, utilizando a função logarítmica e solucionando a equação resultante tem-se $c \simeq 17,3$, como o 95º percentil da distribuição da emissão.

□

6.11 Modelo Log-Normal

Assim como a distribuição Weibull, a distribuição Log-Normal é muito usada para caracterizar tempo de vida de produtos e materiais. Isto inclui fadiga de metal, semicondutores, diodos e isolação elétrica.

Definição 6.9. Uma variável aleatória X tem distribuição Log-Normal se a variável aleatória $Y = \ln(X)$ possui distribuição Normal. A função densidade de probabilidade de uma variável aleatória da variável aleatória $\ln(X)$ é dada por

$$f_X(x) = \frac{1}{x\sigma\sqrt{2\pi}} \exp\left[\frac{-(\ln(x) - \mu)^2}{2\sigma^2}\right] I_{[0,\infty)}(x). \tag{6.9}$$

Cuidado: os parâmetros μ e σ não são a média e o desvio padrão de X , mas de $\ln(X)$.

Notação: $X \sim \text{LN}(\mu, \sigma^2)$.

Corolário 6.6. *Seja $X \sim \text{LN}(\mu, \sigma^2)$. Então $f_X(\cdot)$ dada pela equação (6.9) satisfaz as propriedades de uma função densidade de probabilidade, ou seja, é uma função densidade de probabilidade.*

Demonstração. Seja $X \sim \mathrm{LN}(\mu, \sigma^2)$. Temos que verificar se $f_X(\cdot)$, dada pela equação (6.9), satisfaz as propriedades de função densidade de probabilidade.

Temos que $f_X(\cdot)$ é uma função não negativa. Além disso,

$$\int_{-\infty}^{\infty} f_X(x)dx = \int_{0}^{\infty} \frac{1}{x\sigma\sqrt{2\pi}} \exp\left[\frac{-(\ln(x)-\mu)^2}{2\sigma^2}\right] dx.$$

Fazendo $y = \ln(x)$ $(x = e^y)$, então $\frac{dy}{dx} = \frac{1}{x}$, ou seja, $dy = \frac{1}{x}dx$. Assim,

$$\begin{aligned}\int_{-\infty}^{\infty} f_X(x)dx &= \int_{0}^{\infty} \frac{1}{x\sigma\sqrt{2\pi}} \exp\left[\frac{-(\ln(x)-\mu)^2}{2\sigma^2}\right] dx \\ &= \int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} \exp\left[\frac{-(y-\mu)^2}{2\sigma^2}\right] dy = 1\end{aligned}$$

pois

$$f_Y(y) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[\frac{-(y-\mu)^2}{2\sigma^2}\right],$$

é a função densidade de probabilidade de uma variável aleatória $\mathrm{N}(\mu, \sigma^2)$.

Portanto, $f_X(\cdot)$, dada pela equação (6.9), é uma função densidade de probabilidade.

□

Proposição 6.13. *Se X segue uma distribuição $N(\mu, \sigma^2)$ então, $Y = e^X$ tem distribuição log-normal com parâmetros μ e σ^2.*

Demonstração. De fato se X segue $\mathrm{N}(\mu, \sigma^2)$, então temos que

$$F_Y(y) = \mathbb{P}(Y \leqslant y) = \mathbb{P}(e^X \leqslant y) = \mathbb{P}(X \leqslant \ln y) = F_X(\ln(y))$$

Derivando em relação a y temos

$$f_Y(y) = \frac{d}{dy}F_Y(y) = \frac{d}{dy}F_X(\ln(y)) = f_X(\ln(y))\frac{1}{y}$$

Substituindo a função densidade de probabilidade temos

$$f_Y(y) = f_X(\ln(y))\frac{1}{y} = \frac{1}{y\sigma\sqrt{2\pi}} \exp\left\{-\frac{[\ln(y)-\mu]^2}{2\sigma^2}\right\}.$$

Portanto, a função densidade de probabilidade de Y é uma log-normal com os parâmetros μ e σ^2.

□

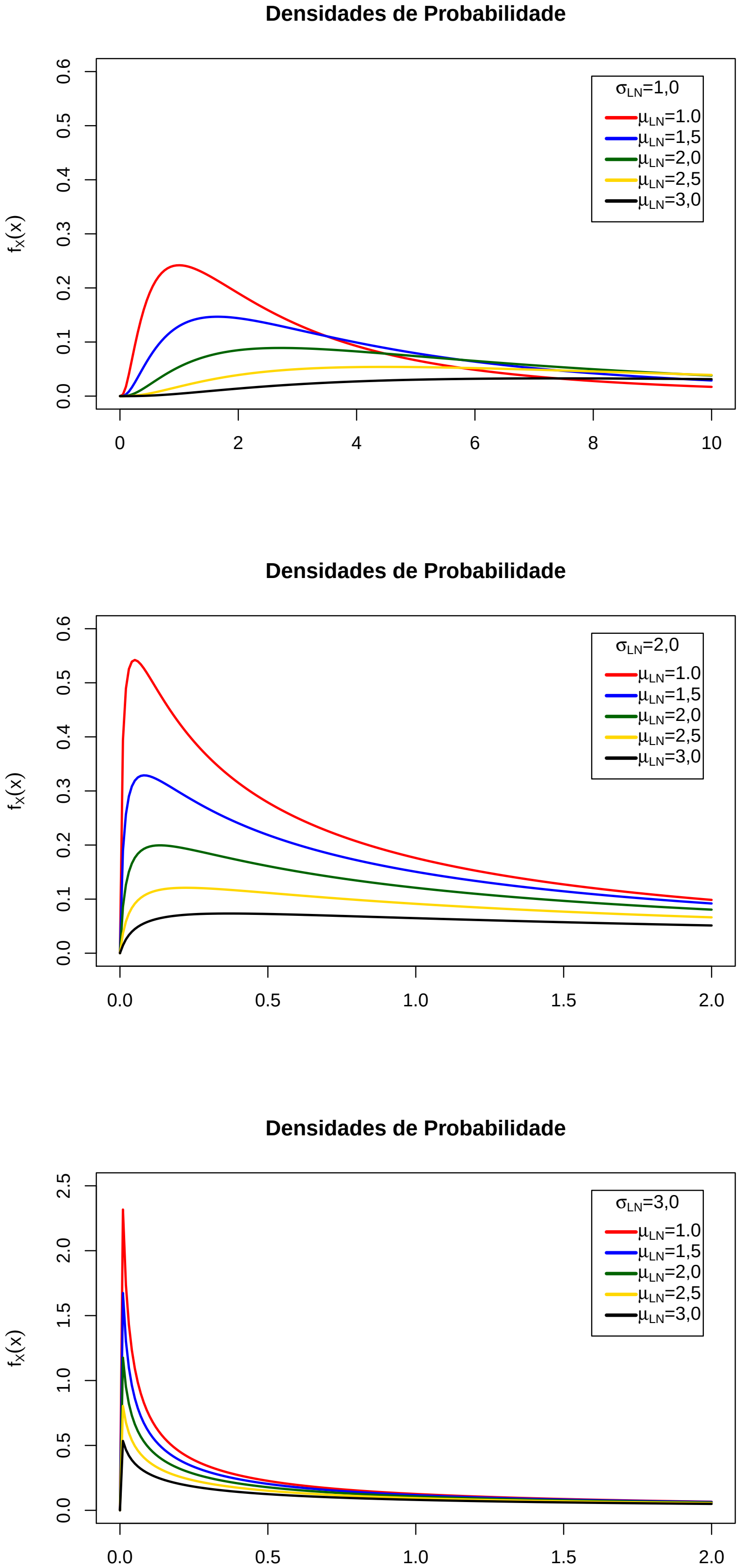


Figura 6.31: Distribuição $\mathrm{LN}(\mu, \sigma^2)$: Função densidade de probabilidade. Parâmetros $\mathbb{E}(X) = \mu_{\mathrm{LN}} \in \{1.0, 1.5, 2.0, 2.5, 3.0\}$ e $\sqrt{\mathrm{Var}(X)} = \sigma_{\mathrm{LN}} = \in \{1.0, 2.0, 3.0\}$.

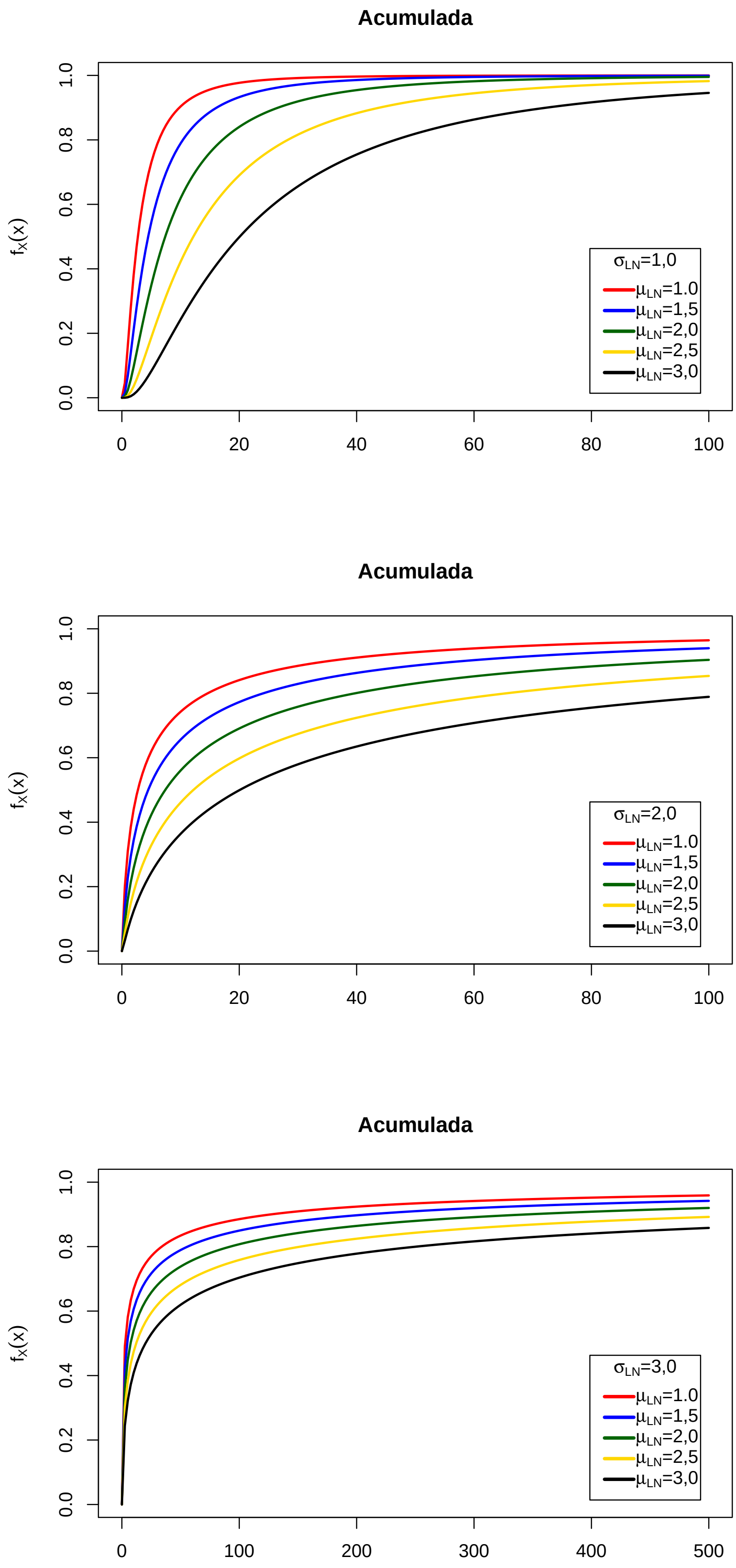


Figura 6.32: Distribuição $\mathrm{W}(\alpha, \beta)$: Função Distribuição de probabilidade. Parâmetros $\mathbb{E}(X) = \mu_{\mathrm{LN}} \in \{1.0, 1.5, 2.0, 2.5, 3.0\}$ e $\sqrt{\mathrm{Var}(X)} = \sigma_{\mathrm{LN}} =\in \{1.0, 2.0, 3.0\}$.

Observação 6.6. Relação da distribuição Lognormal.

(i) Se $X \sim \mathrm{N}(\mu, \sigma^2)$, então $\exp(X) \sim \mathrm{LN}(\mu, \sigma^2)$.

(ii) Se $X \sim \mathrm{LN}(\mu, \sigma^2)$, então $\ln(X) \sim \mathrm{N}(\mu, \sigma^2)$.

(iii) Se $X \sim \mathrm{LN}(\mu, \sigma^2)$, então $Y = aX \sim \mathrm{LN}(\mu + \ln(a), \sigma^2)$.

(iv) Se $X \sim \mathrm{LN}(\mu, \sigma^2)$, então $Y = \frac{1}{X} \sim \mathrm{LN}(-\mu, \sigma^2)$.

(v) Se $X \sim \mathrm{LN}(\mu, \sigma^2)$, então $Y = X^a \sim \mathrm{LN}(a\mu, a^2\sigma^2)$.

Proposição 6.14. *Seja* $Y \sim \mathrm{LN}(\mu, \sigma^2)$. *Então*

(i) $\mathbb{E}(Y) = \exp\left\{\mu + \frac{\sigma^2}{2}\right\}$.

(ii) $\mathbb{E}(Y^n) = \exp\left\{n\mu + \frac{n^2\sigma^2}{2}\right\}$, *para* $n \in \mathbb{N}$.

(iii) $\mathrm{Var}(Y) = \exp(2\mu + \sigma^2)[\exp(\sigma^2) - 1]$.

Demonstração. Seja $X \sim \mathrm{LN}(\mu, \sigma^2)$.

i) Iremos utilizar o resultado apresentado na Proposição 6.13. Se X segue uma distribuição $\mathrm{N}(\mu, \sigma^2)$ então, $Y = e^X$ tem distribuição log-normal com parâmetros μ e σ. Vamos calcular

$$\mathbb{E}(Y) = \mathbb{E}(e^X) = \int_{-\infty}^{\infty} e^x f_X(x) dx,$$

onde X segue uma distribuição $\mathrm{N}(\mu, \sigma^2)$. O cálculo será feito em duas etapas.

Etapa (1) - Seja X segue uma distribuição $\mathrm{N}(0, \sigma^2)$ então,

$$\begin{aligned}
\mathbb{E}(Y) = \mathbb{E}(e^X) &= \int_{-\infty}^{\infty} e^x f_X(x) dx = \int_{-\infty}^{\infty} e^x \left[\frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{x^2}{2\sigma^2}\right\}\right] dx \\
&= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{x - \frac{x^2}{2\sigma^2}\right\} dx = \int_{0}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{(x^2 - 2x\sigma^2)}{2\sigma^2}\right\} dx \\
&= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{(x^2 - 2x\sigma^2 + (\sigma^2)^2 - (\sigma^2)^2)}{2\sigma^2}\right\} dx \\
&= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{(x - \sigma^2)^2 - \sigma^4}{2\sigma^2}\right\} dx \\
&= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{(x - \sigma^2)^2}{2\sigma^2} + \frac{\sigma^4}{2\sigma^2}\right\} dx \\
&= e^{\frac{\sigma^2}{2}} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{(x - \sigma^2)^2}{2\sigma^2}\right\} dx.
\end{aligned}$$

Como

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{(x - \sigma^2)^2}{2\sigma^2}\right\}$$

é a função densidade de probabilidade de uma variável aleatória com distribuição $\mathrm{N}(\sigma^2, \sigma^2)$, integral na equação anterior é igual a 1, logo $Y \sim \mathrm{LN}(0, \sigma^2)$, possui valor esperado

$$\mathbb{E}(Y) = \mathbb{E}(e^X) = e^{\frac{\sigma^2}{2}}.$$

Etapa (2) - Seja X segue uma distribuição $\mathrm{N}(\mu, \sigma^2)$ então,

$$\mathbb{E}(Y) \;=\; \mathbb{E}(e^X) = \int_{-\infty}^{\infty} e^x f_X(x)dx = \int_{-\infty}^{\infty} e^x \left[\frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\}\right] dx.$$

Fazendo a mudança de variável $z = x - \mu$ $(x = z + \mu)$, $\frac{dz}{dx} = 1$, ou seja $dx = dz$, temos

$$\begin{aligned}
\mathbb{E}(Y) = \mathbb{E}(e^X) &= \int_{-\infty}^{\infty} e^x \left[\frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\}\right] dx \\
&= \int_{-\infty}^{\infty} e^{(z+\mu)} \left[\frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{z^2}{2\sigma^2}\right\}\right] dx \\
&= e^{\mu} \int_{-\infty}^{\infty} e^z \left[\frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{z^2}{2\sigma^2}\right\}\right] dz.
\end{aligned}$$

Agora,

$$\int_{-\infty}^{\infty} e^z \left[\frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{z^2}{2\sigma^2}\right\}\right] dz = \mathbb{E}(e^Z) = e^{\frac{\sigma^2}{2}},$$

pois $Z \sim \mathrm{N}(0, \sigma^2)$. Portanto

$$\mathbb{E}(Y) = \mathbb{E}(e^X) = e^{\mu} \times e^{\frac{\sigma^2}{2}} = \exp\left\{\mu + \frac{\sigma^2}{2}\right\}.$$

ii) Iremos generalizar o resultado do item anterior. Se X segue uma distribuição $\mathrm{N}(\mu, \sigma^2)$ então, $Y = e^X$ tem distribuição log-normal com parâmetros μ e σ^2. Vamos calcular

$$\mathbb{E}(Y^n) = \mathbb{E}(e^{nX}) = \int_{-\infty}^{\infty} e^{nx} f_X(x)dx,$$

onde X segue uma distribuição $\mathrm{N}(\mu, \sigma^2)$. O cálculo será feito em duas etapas.

Etapa (1) - Seja X segue uma distribuição $\mathrm{N}(0, \sigma^2)$ então,

$$
\begin{aligned}
\mathbb{E}(Y^n) = \mathbb{E}(e^{nX}) &= \int_{-\infty}^{\infty} e^{nx} f_X(x)dx = \int_{-\infty}^{\infty} e^{nx}\left[\frac{1}{\sqrt{2\pi\sigma^2}}\exp\left\{-\frac{x^2}{2\sigma^2}\right\}\right]dx \\
&= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}}\exp\left\{nx - \frac{x^2}{2\sigma^2}\right\}dx = \int_{0}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}}\exp\left\{-\frac{(x^2 - 2nx\sigma^2)}{2\sigma^2}\right\}dx \\
&= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}}\exp\left\{-\frac{(x^2 - 2nx\sigma^2 + (n\sigma^2)^2 - (n\sigma^2)^2)}{2\sigma^2}\right\}dx \\
&= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}}\exp\left\{-\frac{(x - n\sigma^2)^2 - n^2\sigma^4}{2\sigma^2}\right\}dx \\
&= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}}\exp\left\{-\frac{(x - n\sigma^2)^2}{2\sigma^2} + \frac{n^2\sigma^4}{2\sigma^2}\right\}dx \\
&= e^{\frac{n^2\sigma^2}{2}} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}}\exp\left\{-\frac{(x - n\sigma^2)^2}{2\sigma^2}\right\}dx.
\end{aligned}
$$

Como

$$
f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}}\exp\left\{-\frac{(x - n\sigma^2)^2}{2\sigma^2}\right\}
$$

é a função densidade de probabilidade de uma variável aleatória com distribuição $\mathrm{N}(n\sigma^2, \sigma^2)$, integral na equação anterior é igual a 1, logo $Y \sim \mathrm{LN}(0, \sigma^2)$, possui n-ésimo momento

$$
\mathbb{E}(Y^n) = \mathbb{E}(e^{nX}) = e^{\frac{n^2\sigma^2}{2}}.
$$

Etapa (2) - Seja X segue uma distribuição $\mathrm{N}(\mu, \sigma^2)$ então,

$$
\mathbb{E}(Y^n) = \mathbb{E}(e^{nX}) = \int_{-\infty}^{\infty} e^{nx} f_X(x)dx = \int_{-\infty}^{\infty} e^{nx}\left[\frac{1}{\sqrt{2\pi\sigma^2}}\exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\}\right]dx.
$$

Fazendo a mudança de variável $z = x - \mu$ $(x = z + \mu)$, $\frac{dz}{dx} = 1$, ou seja $dx = dz$, temos

$$
\begin{aligned}
\mathbb{E}(Y^n) = \mathbb{E}(e^{nX}) &= \int_{-\infty}^{\infty} e^{nx}\left[\frac{1}{\sqrt{2\pi\sigma^2}}\exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\}\right]dx \\
&= \int_{-\infty}^{\infty} e^{n(z+\mu)}\left[\frac{1}{\sqrt{2\pi\sigma^2}}\exp\left\{-\frac{z^2}{2\sigma^2}\right\}\right]dx \\
&= e^{n\mu}\int_{-\infty}^{\infty} e^{nz}\left[\frac{1}{\sqrt{2\pi\sigma^2}}\exp\left\{-\frac{z^2}{2\sigma^2}\right\}\right]dz.
\end{aligned}
$$

Agora,

$$
\int_{-\infty}^{\infty} e^{nz}\left[\frac{1}{\sqrt{2\pi\sigma^2}}\exp\left\{-\frac{z^2}{2\sigma^2}\right\}\right]dz = \mathbb{E}(e^{nZ}) = e^{\frac{n^2\sigma^2}{2}},
$$

pois $Z \sim \mathrm{N}(0, \sigma^2)$. Portanto

$$
\mathbb{E}(Y^n) = \mathbb{E}(e^{nX}) = e^{m\mu} \times e^{n^2\frac{\sigma^2}{2}} = \exp\left\{n\mu + \frac{n^2\sigma^2}{2}\right\}.
$$

iii) Por definição temos que $\text{Var}(Y) = \mathbb{E}(Y^2) - [\mathbb{E}(Y)]^2$. Pelo item (i), $\mathbb{E}(Y) = \exp\left\{\mu + \frac{\sigma^2}{2}\right\}$, e pelo item (ii) com $n = 2$, $\mathbb{E}(Y^2) = \exp\left\{2\mu + \frac{2^2\sigma^2}{2}\right\}$, logo

$$
\begin{aligned}
\text{Var}(Y) = \mathbb{E}(Y^2) - [\mathbb{E}(Y)]^2 &= \exp\left\{2\mu + \frac{2^2\sigma^2}{2}\right\} - \left[\exp\left\{\mu + \frac{\sigma^2}{2}\right\}\right]^2 \\
&= \exp\left\{2\mu + 2\sigma^2\right\} - \exp\left\{2\mu + \sigma^2\right\} \\
&= \exp\left\{2\mu\right\}\exp\left\{2\sigma^2\right\} - \exp\left\{2\mu\right\}\exp\left\{\sigma^2\right\} \\
&= \exp\left\{2\mu\right\}\exp\left\{\sigma^2\right\}\left[\exp\left\{\sigma^2\right\} - 1\right] \\
&= \exp\left\{2\mu + \sigma^2\right\}\left[\exp\left\{\sigma^2\right\} - 1\right]
\end{aligned}
$$

□

Proposição 6.15. *Seja* $Y \sim \text{LN}(\mu, \sigma^2)$. *Então*

$$F_Y(y) = F_Z\left(\frac{\ln(y) - \mu}{\sigma}\right),$$

onde $Z \sim \text{N}(0, 1)$.

Demonstração. Seja $Y \sim LN(\mu, \sigma^2)$. Sabemos que $\ln(Y)$ possui distribuição $\text{N}(\mu, \sigma^2)$. Portanto,

$$
\begin{aligned}
F_Y(y) &= \mathbb{P}(Y \leqslant y) = \mathbb{P}(\ln(Y) \leqslant \ln(y)) = \mathbb{P}\left(\frac{\ln(Y) - \mu}{\sigma} \leqslant \frac{\ln(y) - \mu}{\sigma}\right) \\
&= \mathbb{P}\left(Z \leqslant \frac{\ln(y) - \mu}{\sigma}\right) = F_Z\left(\frac{\ln(y) - \mu}{\sigma}\right),
\end{aligned}
$$

onde $Z \sim N(0, 1)$.

□

Exemplo 6.25. [Devore (2018)] De acordo com o artigo "*Predictive model for pitting corrosion in buried oil and gas pipelines." (Corrosion, 2009: 332-342).* a distribuição lognormal foi relatada como a melhor opção para descrever os dados da distribuição da profundidade máxima do ponto de corrosão das tubulações de ferro fundido no solo. Os autores sugerem que a distribuição lognormal com $\mu = 0,353$ e $\sigma = 0,754$ é apropriada para a profundidade máxima do ponto de corrosão (mm) dos canos enterrados. Para esta distribuição, o valor médio e a variância da profundidade do ponto de corrosão são

$$\mathbb{E}(X) = \exp\left\{\mu + \frac{\sigma^2}{2}\right\} = \exp\left\{0,353 + \frac{0,754^2}{2}\right\} = 1,891$$

$$\text{Var}(X) = \exp\left\{2\mu + \sigma^2\right\}\left[\exp\left\{\sigma^2\right\} - 1\right] = \exp\left\{2(0,353) + 0,754^2\right\}\left[\exp\left\{0,754^2\right\} - 1\right]$$

$$\text{Var}(X) = (3,57697)(0,765645) = 2,7387.$$

A probabilidade do de a profundidade máxima do ponto de corrosão estar entre 1 e 2 mm é

$$\begin{aligned}
\mathbb{P}(1 \leqslant X \leqslant 2) &= \mathbb{P}(\ln(1) \leqslant \ln(X) \leqslant \ln(2)) = \mathbb{P}(0 \leqslant \ln(X) \leqslant \ln(2)) \\
&= \mathbb{P}\left(\frac{0-0,353}{0,754} \leqslant \frac{\ln(X)-0,353}{0,754} \leqslant \frac{\ln(2)-0,353}{0,754}\right) \\
&= \mathbb{P}\left(\frac{0-0,353}{0,754} \leqslant Z \leqslant \frac{\ln(2)-0,353}{0,754}\right) \\
&= F_Z(0,47) - F_Z(-0,45) \\
&= \texttt{pnorm(0.47)} - \texttt{pnorm(-0.47)} \\
&= 0,361645.
\end{aligned}$$

A Figura 6.33 representa a probabilidade calculada.

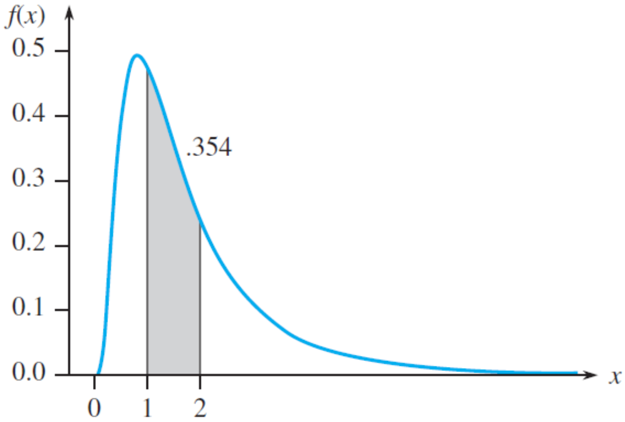


Figura 6.33: Curva de densidade lognormal com $\mu = 0,353$ e $\sigma = 0,754$.
Fonte: Devore (2018)

Qual o valor de c faz com que apenas 1% das amostras tenham uma profundidade máxima do ponto de corrosão excedente a c?

Este valor de c satisfaz

$$0,99 = \mathbb{P}(X \leqslant c) = \mathbb{P}(\ln(X) \leqslant \ln(c)) = \mathbb{P}\left(\frac{\ln(X)-0,353}{0,754} \leqslant \frac{\ln(c)-0,353}{0,754}\right) = \mathbb{P}\left(Z \leqslant \frac{\ln(c)-0,353}{0,754}\right).$$

Temos que $\mathbb{P}(Z \leqslant z_{0,99}) = \texttt{qnorm(0.99)} = 2,33$. Isso implica que

$$\frac{\ln(c)-0,353}{0,754} = 2,33$$

do qual temos que $\ln(c) = 2,1098$, ou seja, $c = 8,247$. Dessa forma, $8,247$ é o 99º percentil da distribuição da profundidade máxima do ponto de corrosão. □

Exemplo 6.26. Seja $X \sim \text{LN}(10,4)$. Calcule

(i) $\mathbb{P}(X \leqslant 100)$.

(ii) $\mathbb{P}(800 \leqslant X \leqslant 1150)$.

Solução:

(i) Cálculo de $\mathbb{P}(X \leqslant 100)$.

$$\mathbb{P}(X \leqslant 100) = \mathbb{P}(\ln(X) \leqslant \ln(100)) = \mathbb{P}\left(\frac{\ln(X) - 10}{2} \leqslant \frac{\ln(100) - 10}{2}\right) = \mathbb{P}\left(Z \leqslant \frac{\ln(100) - 10}{2}\right),$$

pois $\frac{\ln(X)-10}{2} \sim N(0,1)$. Portanto,

$$\mathbb{P}(X \leqslant 100) = \mathbb{P}\left(Z \leqslant \frac{\ln(100) - 10}{2}\right) = \mathbb{P}(Z \leqslant -1,55) = \text{pnorm(-1.55)} = 0.06057076.$$

(ii) Cálculo de $\mathbb{P}(800 \leqslant X \leqslant 1150)$.

$$\begin{aligned}
\mathbb{P}(800 \leqslant X \leqslant 1150) &= \mathbb{P}(\ln(800) \leqslant \ln(X) \leqslant \ln(1150)) \\
&= \mathbb{P}\left(\frac{\ln(800) - 10}{2} \leqslant \frac{\ln(X) - 10}{2} \leqslant \frac{\ln(1150) - 10}{2}\right) \\
&= \mathbb{P}\left(\frac{\ln(800) - 10}{2} \leqslant Z \leqslant \frac{\ln(1150) - 10}{2}\right) \\
&= \mathbb{P}(-1.657694 \leqslant Z \leqslant -1.476241) \\
&= \mathbb{P}(Z \leqslant -1.476241) - \mathbb{P}(Z < -1.657694) \\
&= \text{pnorm(-1.476241)} - \text{pnorm(-1.657694)} \\
&= 0,0699396 - 0,04868962 = 0,02124998,
\end{aligned}$$

pois $\frac{\ln(X)-10}{2} \sim N(0,1)$

Capítulo 7

Função Geradora de Momentos e Função Característica

Neste capítulo, apresentamos as *funções geradora de momentos e característica*. Como outras transformadas, elas facilitam o cálculo de probabilidades e de outras quantidades relacionadas.

7.1 Função Geradora de Momentos

A seguir, apresentamos a definição de função geradora de momentos.

Definição 7.1. A *função geradora de momentos* da variável aleatória X é definida por

$$\mathrm{M}_X(t) = \mathbb{E}(e^{tX}),$$

desde que a esperança seja finita para todo t real em algum intervalo $-t_0 < t < t_0$, com $t_0 > 0$.

A função geradora de momentos é o valor esperado de uma função positiva da variável aleatória. Segundo nossa definição de esperança, o valor infinito era possível e, portanto, a expressão acima também poderia produzir esse valor para alguns intervalos de t. Entretanto, a definição de função geradora de momentos requer que a integral seja finita para valores de t numa vizinhança de zero. Isso garante algumas propriedades importantes dessa função. Assim, ao apresentar a função geradora de momentos, convém sempre mencionar o intervalo de valores de t (sempre ao redor de zero) em que ela existe.

Resultado 7.1. Definição de Função Geradora de Momentos, segundo do tipo de variável aleatória.

(i) Se X for uma variável aleatória discreta, com função massa de probabilidade $f_X(x) = \mathbb{P}(X = x)$, a função $\mathrm{M}_X(t)$, denominada função geradora de momentos de X, é definida por

$$\mathrm{M}_X(t) = \sum_{x\in\Omega_X} e^{tx} f_X(x).$$

(ii) Se X for uma variável aleatória contínua, função densidade de probabilidade $f_X(\cdot)$, definimos a função geradora de momentos por

$$\mathrm{M}_X(t) = \int_{-\infty}^{\infty} e^{tx} f_X(x)dx.$$

(iii) $\mathrm{M}_X(0) = \mathrm{M}_X(t)|_{t=0} = \mathbb{E}(e^{0X}) = 1.$

Exemplo 7.1. Seja $X \sim \mathrm{B}(n, p)$. Encontre a função geradora de momentos de X.

Solução: De fato,

$$\begin{aligned} M_X(t) = \mathbb{E}(e^{tX}) &= \sum_{x=0}^{n} e^{tx} f_X(x) = \sum_{x=0}^{n} e^{tx} \binom{n}{x} p^x (1-p)^{n-x} \\ &= \sum_{x=0}^{n} \binom{n}{x} (e^t p)^x (1-p)^{n-x} \\ &= (e^t p + (1-p))^n, \text{ para todo } t \in \mathbb{R}, \end{aligned}$$

pois

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

□

Exemplo 7.2. Seja $X \sim \mathrm{P}(\lambda)$, $\lambda > 0$. Encontre a função geradora de momentos de X.

Solução: De fato,

$$f_X(x) = \mathbb{P}(X = x) = \frac{e^{-\lambda}\lambda^x}{x!}, \quad x = 0, 1, 2, \cdots.$$

$$\mathrm{M}_X(t) = \mathbb{E}(e^{tX}) = \sum_{x\geqslant 0} e^{tx} f_X(x) = \sum_{x\geqslant 0} e^{tx}\frac{e^{-\lambda}\lambda^x}{x!} = e^{-\lambda}\sum_{x\geqslant 0}\frac{(e^t\lambda)^x}{x!}.$$

Como $e^y = \sum_{j\geqslant 0}\frac{y^j}{j!}$, temos

$$\mathrm{M}_X(t) = e^{-\lambda} e^{e^t\lambda} = e^{-\lambda(1-e^t)}, \text{ para todo } t \in \mathbb{R}.$$

□

Exemplo 7.3. Seja $X \sim \text{Exp}(\alpha)$. Encontre a função geradora de momentos de X.

Solução: De fato,

$$\mathrm{M}_X(t) = \mathbb{E}(e^{tX}) = \int_0^\infty e^{tx} \alpha e^{-\alpha x} dx = \int_0^\infty \alpha e^{-x(\alpha - t)} dx.$$

Considerando $\alpha - t > 0$, ou seja $t < \alpha$, temos

$$\mathrm{M}_X(t) = -\alpha \frac{e^{-x(\alpha-t)}}{\alpha - t}\Bigg|_{x=0}^{\infty} = \frac{-\alpha}{\alpha - t}\left(\lim_{x\to\infty} e^{-x(\alpha-t)} - 1\right) = \frac{-\alpha}{\alpha - t}(0-1) = \frac{\alpha}{\alpha - t}.$$

Logo,

$$\mathrm{M}_X(t) = \frac{\alpha}{\alpha - t}, \quad \text{para} \quad t < \alpha.$$

□

Exemplo 7.4. Seja $X \sim \Gamma(\alpha, \beta)$. Encontre a função geradora de momentos de X.

Solução: Lembre que

$$f_x(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} e^{-\beta x} x^{\alpha - 1} I_{(0,\infty)}(x).$$

Logo,

$$\mathrm{M}_X(t) = \mathbb{E}(e^{tX}) \quad = \quad \int_0^\infty e^{tx} \frac{\beta^\alpha}{\Gamma(\alpha)} e^{-\beta x} x^{\alpha-1} dx = \frac{\beta^\alpha}{\Gamma(\alpha)} \int_0^\infty e^{-x(\beta - t)} x^{\alpha - 1} dx.$$

A integral converge desde que $\beta > t$ pois, caso contrário,

$$\lim_{x\to\infty} e^{-x(\beta-t)} = \infty.$$

Fazendo $x(\beta - t) = u$, temos

$$\mathrm{M}_X(t) = \frac{\beta^\alpha}{\Gamma(\alpha)(\beta - t)^\alpha} \int_0^\infty u^{\alpha-1} e^{-u} du.$$

Lembre que, $\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} e^{-x} dx$. Logo

$$\mathrm{M}_X(t) = \frac{\beta^\alpha}{\Gamma(\alpha)(\beta - t)^\alpha} \Gamma(\alpha) = \left(\frac{\beta}{\beta - t}\right)^\alpha.$$

□

Observação 7.1. A distribuição χ^2_n é um caso particular da distribuição $\Gamma(\alpha, \beta)$. A relação entre as duas distribuições é dada por

$$\begin{cases} X \sim \Gamma(\alpha,\beta) \\ \mathbb{E}(X) = \frac{\alpha}{\beta} \\ \mathrm{Var}(X) = \frac{\alpha}{\beta^2} \end{cases} \longrightarrow \alpha = \frac{n}{2}, \beta = \frac{1}{2} \longrightarrow \begin{cases} Z \sim \chi^2_n \\ \mathbb{E}(Z) = n \\ \mathrm{Var}(Z) = 2n. \end{cases}$$

A função geradora de momentos de $Z \sim \chi^2_n$, é dada por

$$\mathrm{M}_Z(t) = \left(\frac{1/2}{1/2 - t}\right)^{\frac{n}{2}} = \left(\frac{1}{1-2t}\right)^{\frac{n}{2}}, \quad \text{para} \quad t < \frac{1}{2}.$$

Exemplo 7.5. Seja $X \sim \mathrm{N}(0,1)$. Encontre a função geradora de momentos de X.

Solução: De fato,

$$\begin{aligned} \mathrm{M}_X(t) = \mathbb{E}(e^{tX}) &= \int_{-\infty}^{\infty} e^{tx}\frac{1}{\sqrt{2\pi}}e^{\frac{-x^2}{2}}dx = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}}e^{\frac{-(x^2-2tx)}{2}}dx \\ &= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}}e^{\frac{-(x-t)^2}{2}+\frac{t^2}{2}}dx = e^{\frac{t^2}{2}}\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}}e^{\frac{-(x-t)^2}{2}}dx. \end{aligned}$$

Temos que $f_X(x) = \frac{1}{\sqrt{2\pi}}e^{\frac{-(x-t)^2}{2}}$ é a função densidade de probabilidade de uma variável aleatória com distribuição $\mathrm{N}(t,1)$, logo

$$\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}}e^{\frac{-(x-t)^2}{2}}dx = 1.$$

Portanto,

$$\mathrm{M}_X(t) = e^{\frac{t^2}{2}}, \text{ para } t \in \mathbb{R}.$$

□

Teorema 7.1. *Suponha que uma variável aleatória X tenha função geradora de momentos $\mathrm{M}_X(t)$. Seja $Y = \alpha X + \beta$. Então $\mathrm{M}_Y(t)$, a função geradora de momentos da variável aleatória Y, será dada por*

$$\mathrm{M}_Y(t) = e^{\beta t}\mathrm{M}_X(\alpha t).$$

Demonstração. Seja $Y = \alpha X + \beta$ uma variável aleatória Então, sua f.g.m é dada por

$$\begin{aligned} \mathrm{M}_Y(t) &= \mathbb{E}(e^{tY}) = \mathbb{E}(e^{t(\alpha X+\beta)}) = \mathbb{E}(e^{t\alpha X}e^{t\beta}) \\ &= e^{t\beta}\mathbb{E}(e^{t\alpha X}) = e^{\beta t}\mathrm{M}_X(\alpha t). \end{aligned}$$

□

Exemplo 7.6. Seja X uma variável aleatória com distribuição $\mathrm{N}(\mu, \sigma^2)$. Encontre a função geradora de momentos de X.

Solução: Lembre que $Z = \frac{X-\mu}{\sigma}$ possui distribuição $\mathrm{N}(0,1)$. Então, $X = \sigma Z + \mu$, logo

$$\begin{aligned} \mathrm{M}_X(t) &= \mathbb{E}(e^{tX}) = \mathbb{E}(e^{t(\sigma Z+\mu)}) = \mathbb{E}(e^{t\sigma Z}e^{t\mu}) \\ &= e^{t\mu}\mathbb{E}(e^{t\sigma Z}) = e^{\mu t}\mathrm{M}_Z(\sigma t) \\ &= e^{\mu t}e^{\frac{(\sigma t)^2}{2}} = e^{\mu t+\frac{(\sigma t)^2}{2}} \text{ para } t \in \mathbb{R}. \end{aligned}$$

□

Teorema 7.2. *Suponha que a função geradora de momentos da variável aleatória X exista para $|t| < t_0$, $t_0 > 0$. Então, $\mathbb{E}(X^n)$ existe para $n = 1, 2, 3, \cdots$ e temos que*

$$\mathbb{E}(X^n) = \left.\frac{\partial^n M_X(t)}{\partial t^n}\right|_{t=0}.$$

Demonstração. Temos que, para $x \in \mathbb{R}$, a função e^x, pode ser escrita na forma de uma série de Taylor como

$$e^x = 1 + x + \frac{(x)^2}{2!} + \frac{(x)^3}{3!} + \frac{(x)^4}{4!}\cdots$$

Dessa forma

$$e^{tx} = 1 + tx + \frac{(tx)^2}{2!} + \frac{(tx)^3}{3!} + \frac{(tx)^4}{4!}\cdots$$

Seja $t_0 > 0$, tal que $\mathbb{E}(e^{tX})$ seja finita em $-t_0 < t < t_0$. Então,

$$\mathrm{M}_X(t) = \mathbb{E}(e^{tX}) = 1 + t\mathbb{E}(X) + \frac{t^2\mathbb{E}(X^2)}{2!} + \frac{t^3\mathbb{E}(X^3)}{3!} + \cdots + \frac{t^n\mathbb{E}(X^n)}{n!}\cdots$$

Derivando a expressão acima em relação a t, temos

$$\frac{d\mathrm{M}_X(t)}{dt} = \mathrm{M}'_X(t) = \mathbb{E}(X) + 2t\frac{\mathbb{E}(X^2)}{2!} + 3t^2\frac{\mathbb{E}(X^3)}{3!} + \cdots + nt^{n-1}\frac{\mathbb{E}(X^n)}{n!}\cdots$$

Fazendo $t = 0$, verificamos que $\left.\mathrm{M}'_X(t)\right|_{t=0} = \mathbb{E}(X)$.

Derivando $\mathrm{M}'_X(t)$ em relação a t e aplicando $t = 0$, obtemos

$$\left.\mathrm{M}''_X(t)\right|_{t=0} = \mathbb{E}(X^2).$$

Prosseguindo desta forma, pela n-ésima derivada, obtemos

$$\left.\frac{d^n\mathrm{M}_X(t)}{dt^n}\right|_{t=0} = \left.\mathrm{M}_X^{(n)}(t)\right|_{t=0} = \mathbb{E}(X^n).$$

□

Observação 7.2.

$$\mathrm{Var}(X) = \mathbb{E}(X^2) - [\mathbb{E}(X)]^2 = \mathrm{M}''_X(t)\Big|_{t=0} - \left[\mathrm{M}'_X(t)\Big|_{t=0}\right]^2$$

Exemplo 7.7. Seja X uma variável aleatória com distribuição $\mathrm{B}(n,p)$. Obtenha $\mathbb{E}(X)$ e $\mathrm{Var}(X)$ através função geradora de momentos.

Solução: Lembre que

$$\mathrm{M}_X(t) = (e^t p + (1-p))^n = (pe^t + q)^n.$$

Logo,

$$\mathrm{M}'_X(t) = n(pe^t + q)^{n-1} + (pe^t)$$

e

$$\mathrm{M}''_X(t) = np[(n-1)(pe^t + q)^{n-2}(pe^t)e^t + (pe^t + q)^{n-1}e^t].$$

Fazendo $t = 0$, obtemos o primeiro e o segundo momentos

$$\mathbb{E}(X) = \mathrm{M}'_X(t)\Big|_{t=0} = np$$

e

$$\mathbb{E}(X^2) = \mathrm{M}''_X(t)\Big|_{t=0} = np^2(n-1) + np = n^2p^2 - np^2 + np = np(np - p + 1).$$

Portanto,

$$\mathrm{Var}(X) = \mathbb{E}(X^2) - [\mathbb{E}(X)]^2 = n^2p^2 - np^2 + np - n^2p^2 = np(1-p) = npq.$$

□

Exemplo 7.8. Seja X uma variável aleatória com distribuição $\mathrm{N}(0,1)$. Obtenha $\mathbb{E}(X)$ e $\mathrm{Var}(X)$ através função geradora de momentos.

Solução: Lembre que

$$\mathrm{M}_X(t) = e^{\frac{t^2}{2}}, \quad t \in \mathbb{R}.$$

Logo,

$$\mathrm{M}'_X(t) = te^{\frac{t^2}{2}}$$

e

$$\mathrm{M}''_X(t) = e^{\frac{t^2}{2}} + t(te^{\frac{t^2}{2}}).$$

Fazendo $t = 0$, obtemos o primeiro e o segundo momentos

$$\mathbb{E}(X) = \mathrm{M}'_X(t)\Big|_{t=0} = 0$$

e

$$\mathbb{E}(X^2) = \mathrm{M}''_X(t)\Big|_{t=0} = 1.$$

Portanto,

$$\mathrm{Var}(X) = \mathbb{E}(X^2) - [\mathbb{E}(X)]^2 = 1.$$

□

Exemplo 7.9. Seja X uma variável aleatória com distribuição $\mathrm{N}(\mu, \sigma^2)$. Obtenha $\mathbb{E}(X)$ e $\mathrm{Var}(X)$ através função geradora de momentos.

Solução: Lembre que

$$\mathrm{M}_X(t) = e^{\mu t + \frac{(\sigma t)^2}{2}} \text{ para } t \in \mathbb{R}.$$

Logo,

$$\mathrm{M}'_X(t) = e^{\mu t + \frac{(\sigma t)^2}{2}} \frac{d}{dt}\left[\mu t + \frac{(\sigma t)^2}{2}\right] = e^{\mu t + \frac{(\sigma t)^2}{2}} \left[\mu + \sigma^2 t\right]$$

e

$$\begin{aligned}
\mathrm{M}''_X(t) &= \frac{d^2}{dt^2}\mathrm{M}_X(t) = \frac{d}{dt}\mathrm{M}'_X(t) = \frac{d}{dt} e^{\mu t + \frac{(\sigma t)^2}{2}} \left[\mu + \sigma^2 t\right] \\
&= \frac{d}{dt}\mu e^{\mu t + \frac{(\sigma t)^2}{2}} + \frac{d}{dt}\sigma^2 t e^{\mu t + \frac{(\sigma t)^2}{2}} \\
&= \left\{\mu e^{\mu t + \frac{(\sigma t)^2}{2}} \left[\mu + \sigma^2 t\right]\right\} + \left\{\sigma^2 e^{\mu t + \frac{(\sigma t)^2}{2}} + \sigma^2 t e^{\mu t + \frac{(\sigma t)^2}{2}} \left[\mu + \sigma^2 t\right]\right\}
\end{aligned}$$

Fazendo $t = 0$, obtemos o primeiro e o segundo momentos

$$\mathbb{E}(X) = \mathrm{M}'_X(t)\Big|_{t=0} = \mu$$

e

$$\mathbb{E}(X^2) = \mathrm{M}''_X(t)\Big|_{t=0} = \mu^2 + \sigma^2.$$

Portanto,

$$\mathrm{Var}(X) = \mathbb{E}(X^2) - [\mathbb{E}(X)]^2 = \mu^2 + \sigma^2 - [\mu]^2 = \sigma^2.$$

□

Teorema 7.3. *Se duas variável aleatória's têm função geradora de momentos que existem e são iguais, então elas têm a mesma função de distribuição.*

Exemplo 7.10. A função geradora de momentos da variável aleatória X é dada por

$$\mathrm{M}_X(t) = \left(\frac{1}{3}e^t + \frac{2}{3}\right)^4.$$

Comparando a expressão acima com a função geradora de momentos de uma variável aleatória com distribuição $\mathrm{B}(n, p)$, podemos concluir que $X \sim \mathrm{B}(4, 1/3)$.

□

7.2 Função Característica

Na função geradora de momentos a $\mathbb{E}(X)$ pode nem sempre existir. Definiremos uma nova transformada que sempre existe.

Definição 7.2. A *função característica* de uma variável aleatória X é definida por

$$\begin{aligned}\varphi_X(t) &= \mathbb{E}(e^{itX}) = \mathbb{E}(\cos(tX) + i\,\mathrm{sen}(tX)) \\ &= \mathbb{E}(\cos(tX)) + i\mathbb{E}(\mathrm{sen}(tX)),\end{aligned}$$

para todo $t \in \mathbb{R}$ e $i = \sqrt{-1}$ (número imaginário).

Exemplo 7.11. Seja $X \sim \mathrm{U}[a, b]$, $a < b$. Encontre a função característica de X.

Solução: De fato,

$$\begin{aligned}\varphi_X(t) = &= \mathbb{E}(e^{itX}) = \mathbb{E}(\cos(tX) + i\,\mathrm{sen}(tX)) \\ &= \int_a^b \frac{\cos(tx)}{b-a}dx + i\int_a^b \frac{\mathrm{sen}(tx)}{b-a}dx \\ &= \frac{1}{b-a}\int_a^b \cos(tx)dx + \frac{i}{b-a}\int_a^b \mathrm{sen}(tx)dx \\ &= \frac{1}{b-a}\left[\frac{\mathrm{sen}(tx)}{t} - i\frac{\cos(tx)}{t}\Big|_a^b\right] \\ &= \frac{1}{t(b-a)}(\mathrm{sen}(tb) - i\cos(tb) - (\mathrm{sen}(ta) - i\cos(ta)))\frac{i}{i} \\ &= \frac{1}{it(b-a)}(\cos(tb) + i\,\mathrm{sen}(tb) - (\cos(ta) + i\,\mathrm{sen}(ta))) \\ &= \frac{e^{itb} - e^{ita}}{it(b-a)}, \quad t \in \mathbb{R}.\end{aligned}$$

□

Observação 7.3. Levando em consideração as proximidades das definições de $\varphi_X(t)$ e $\mathrm{M}_X(t)$, se $\mathrm{M}_X(t)$ existir e é conhecida, ela pode ser usada para calcular a função característica $\varphi_X(t)$, isto é,

$$\varphi_X(t) = \mathrm{M}_X(it).$$

Exemplo 7.12. Seja $X \sim \mathrm{B}(n, p)$. Encontre sua função característica.

Solução: Lembre que $\mathrm{M}_X(t) = (pe^t + q)^n$. Logo,

$$\varphi_X(t) = \mathrm{M}_X(it) = (pe^{it} + q)^n$$

□

Exemplo 7.13. Seja $X \sim \mathrm{N}(\mu, \sigma^2)$. Encontre sua função característica.

Solução: Lembre que $\mathrm{M}_X(t) = e^{\mu t + \frac{(\sigma t)^2}{2}}$. Logo,

$$\varphi_X(t) = \mathrm{M}_X(it) = e^{i\mu t + \frac{(i\sigma t)^2}{2}} = e^{i\mu t - \frac{(\sigma t)^2}{2}}.$$

□

Proposição 7.1. *Seja X uma variável aleatória qualquer e $\varphi_X(t)$ sua função característica. Então,*

(i) $\varphi_X(0) = 1$.

De fato:

$$\varphi_X(0) = \mathbb{E}(e^{i0X}) = \mathbb{E}(1) = 1.$$

(ii) $|\varphi_X(t)| \leqslant 1$.

De fato:

$$\begin{aligned}
|\varphi_X(t)| &= |\mathbb{E}(e^{itX})| = |\mathbb{E}(\cos(tX)) + i\mathbb{E}(\,\mathrm{sen}(tX))| \\
&= \sqrt{(\mathbb{E}(\cos(tX)))^2 + (\mathbb{E}(\,\mathrm{sen}(tX)))^2} \qquad (7.1) \\
&\leqslant \sqrt{\mathbb{E}(\cos^2(tX)) + \mathbb{E}(\,\mathrm{sen}^2(tX))} \qquad (7.2) \\
&= \sqrt{\mathbb{E}(\cos^2(tX) + \mathrm{sen}^2(tX))} \qquad (7.3) \\
&= \sqrt{\mathbb{E}(1)} = 1
\end{aligned}$$

A equação (7.1) seque pois, para um número real $Z = a + bi$, o seu módulo é dado por $|Z| = \sqrt{a^2 + b^2}$. A equação (7.2) seque pela desigualdade de Jensen, isto é, $\psi(\mathbb{E}(Y)) \leqslant \mathbb{E}(\psi(Y))$ para $\psi(\cdot)$ convexa. Neste caso $\psi(z) = z^2$. A equação (7.3) seque pois $\mathrm{sen}^2(a) + \cos^2(a) = 1$.

(iii) $\overline{\varphi_X(t)} = \varphi_X(-t)$, onde $\bar{c}$ é o complexo conjugado de c, isto é, $c = a + bi$, $\bar{c} = a - bi$.
De fato:

$$\begin{aligned}\varphi_X(-t) &= \mathbb{E}(e^{-itX}) = \mathbb{E}(\cos(tX)) - i\mathbb{E}(\operatorname{sen}(tX)) \\ &= \overline{\mathbb{E}(\cos(tX)) + i\mathbb{E}(\operatorname{sen}(tX))} \\ &= \overline{\mathbb{E}(e^{itX})} = \overline{\varphi_X(t)}\end{aligned}$$

□

Teorema 7.4. *Suponha que a função característica da variável aleatória X exista. Então, $\mathbb{E}(X^n)$ existe para $n = 1, 2, 3, \cdots$ e são gerados por $\varphi_X(t)$ através da relação*

$$\left.\frac{\partial^n \varphi_X(t)}{\partial t^n}\right|_{t=0} = i^n \mathbb{E}(X^n),$$

para $n \geqslant 1$ e se $\mathbb{E}(|X^n|) < \infty$.

Exemplo 7.14. Seja $X \sim \mathrm{N}(\mu, \sigma^2)$. Encontre $\mathbb{E}(X)$ e $\mathrm{Var}(X)$.

Solução: Lembre que $\varphi'_X(t) = e^{i\mu t - \frac{(\sigma t)^2}{2}}$.
Então,

$$\varphi'_X(t) = e^{i\mu t - \frac{(\sigma t)^2}{2}}(i\mu - \sigma^2 t).$$

Logo,

$$i\mathbb{E}(X) = \left.\varphi'_X(t)\right|_{t=0} = i\mu.$$

Portanto, $\mathbb{E}(X) = \mu$.
Da mesma forma, temos que

$$\varphi''_X(t) = e^{i\mu t - \frac{(\sigma t)^2}{2}}(i\mu - \sigma^2 t)^2 + e^{i\mu t - \frac{(\sigma t)^2}{2}}(-\sigma^2).$$

Logo,

$$i^2\mathbb{E}(X^2) = \left.\varphi''_X(t)\right|_{t=0} = i^2\mu^2 - \sigma^2.$$

Portanto,

$$\mathbb{E}(X^2) = \mu^2 - \frac{\sigma^2}{i^2} = \mu^2 + \sigma^2.$$

Assim,

$$\mathrm{Var}(X) = \mathbb{E}(X^2) - [\mathbb{E}(X)]^2 = \mu^2 + \sigma^2 - \mu^2 = \sigma^2.$$

□

Capítulo 8

Anexo

8.1 Integrais

(i) $\displaystyle\int x^k dx = \frac{x^{k+1}}{k+1}$, para $k \neq -1$;

(ii) $\displaystyle\int e^{\alpha x} dx = \frac{e^{\alpha x}}{\alpha}$, para α constante;

(iii) Caso particular: $\displaystyle\int e^x dx = e^x$;

(iv) $\displaystyle\int [f(x) + g(x)] dx = \int f(x) dx + \int g(x) dx$;

(v) $\displaystyle\int cf(x) dx = c \int f(x) dx$, para $c \neq 0$ constante;

8.2 Derivadas

(i) $\frac{d}{dx} k = 0$, para k constante;

(ii) $\frac{d}{dx} kx = k$, para k constante;

(iii) $\frac{d}{dx} x^n = nx^{n-1}$;

(iv) $\frac{d}{dx} e^x = e^x$;

(v) $\frac{d}{dx} e^{\alpha x} = \alpha e^{\alpha x}$, para α constante;

(vi) $\frac{d}{dx}[kf(x)] = k \frac{d}{dx} f(x) = f'(x)$;

(vii) $\frac{d}{dx}[f(x) + g(x)] = \frac{d}{dx} f(x) + \frac{d}{dx} g(x) = f'(x) + g'(x)$;

Capítulo 9

Bibliografia

1. [Peternelli, 2004] http://bit.do/ezN6S

2. [Prob. Estat. IM-UFRJ, 2017] http://www.im.ufrj.br/probest/

3. [Magalhães, 2006] Magalhães, M.N. (2006). *Probabilidade e Variáveis Aleatórias*, EDUsp, 2º edição, São Paulo.

4. [Devore, 2016] Devore, J.L. (2016). *Probability and Statistics for Engineering and the Sciences.* Cengage Learning, 9th Edition. Brasil.

5. [Rohatgi, 1976] Rohatgi, V.K. (1976). *An Introduction to Probability Theory and Mathematical Statistics.* John wiles & Sons, Inc. New York.

6. [Sindelar, et al. 2014] Sindelar, F.C.W., Conto, S.M. e L. Ahlert. (2014). *Teoria e prática em Estatística para cursos de graduação.* Univates, Lajeado.

7. [Blitzstein e Jessica, 2019] Blitzstein, Joseph K.; Hwang, Jessica. Introduction to probability. 2nd Edition. Chapman and Hall/CRC, 2019. http://bit.do/eL6bY

8. [Casella e Berger, 2001] Casella, G. e R.L. Berger. (2001). *Statistical Inference.* Austrália: Duxbury/Thomson Learning.

9. [Degroot, 1989] Degroot, M.T. (1989). *Probability and Statistics.*New York: Addison-Wesley.

10. [Meyer, 2006] Meyer, Paul L. Introductory probability and statistical applications. Oxford and IBH Publishing, 2nd Edition, 2006.

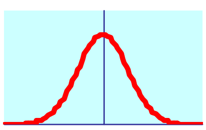
11. [Guidorizzi, 2008] Guidorizzi, Hamilton Luiz (2008). Um curso de cálculo, vol. 1,2,3 e 4 . Grupo Gen-LTC.

12. [Montgomery, 2003] Montgomery, D.C. e Runger, G.C. (2003). *Estatística Aplicada e Probabilidade Para Engenheiros*, LTC, Segunda Edição.

13. [Bussab e Morettin, 2017] Bussab W.O. e Morettin P.A., Estatística Básica, Saraiva, Sao Paulo, 9ed, 2017.Códigos R: https://www.ime.usp.br/~pam/scriptsR.html

Tabela 1 - Valores da distribuição normal padrão - $\Phi(z) = P(Z \le z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{\frac{-u^2}{2}} du$

	0	1	2	3	4	5	6	7	8	9
-3	0,0013	0,0010	0,0007	0,0005	0,0003	0,0002	0,0002	0,0001	0,0001	0,0000
-2,9	0,0019	0,0018	0,0017	0,0017	0,0016	0,0016	0,0015	0,0015	0,0014	0,0014
-2,8	0,0026	0,0025	0,0024	0,0023	0,0023	0,0022	0,0021	0,0021	0,0020	0,0019
-2,7	0,0035	0,0034	0,0033	0,0032	0,0031	0,0030	0,0029	0,0028	0,0027	0,0026
-2,6	0,0047	0,0045	0,0044	0,0043	0,0041	0,0040	0,0039	0,0038	0,0037	0,0036
-2,5	0,0062	0,0060	0,0059	0,0057	0,0055	0,0054	0,0052	0,0051	0,0049	0,0048
-2,4	0,0082	0,0080	0,0078	0,0075	0,0073	0,0071	0,0069	0,0068	0,0066	0,0064
-2,3	0,0107	0,0104	0,0102	0,0099	0,0096	0,0094	0,0091	0,0089	0,0087	0,0084
-2,2	0,0139	0,0136	0,0132	0,0129	0,0126	0,0122	0,0119	0,0116	0,0113	0,0110
-2,1	0,0179	0,0174	0,0170	0,0166	0,0162	0,0158	0,0154	0,0150	0,0146	0,0143
-2,0	0,0228	0,0222	0,0217	0,0212	0,0207	0,0202	0,0197	0,0192	0,0188	0,0183
-1,9	0,0287	0,0281	0,0274	0,0268	0,0262	0,0256	0,0250	0,0244	0,0238	0,0233
-1,8	0,0359	0,0352	0,0344	0,0336	0,0329	0,0322	0,0314	0,0307	0,0300	0,0294
-1,7	0,0446	0,0436	0,0427	0,0418	0,0409	0,0401	0,0392	0,0384	0,0375	0,0367
-1,6	0,0548	0,0537	0,0526	0,0516	0,0505	0,0495	0,0485	0,0475	0,0465	0,0455
-1,5	0,0668	0,0655	0,0643	0,0630	0,0618	0,0606	0,0594	0,0582	0,0570	0,0559
-1,4	0,0808	0,0793	0,0778	0,0764	0,0749	0,0735	0,0722	0,0708	0,0694	0,0681
-1,3	0,0968	0,0951	0,0934	0,0918	0,0901	0,0885	0,0869	0,0853	0,0838	0,0823
-1,2	0,1151	0,1131	0,1112	0,1093	0,1075	0,1056	0,1038	0,1020	0,1003	0,0985
-1,1	0,1357	0,1335	0,1314	0,1292	0,1271	0,1251	0,1230	0,1210	0,1190	0,1170
-1,0	0,1587	0,1562	0,1539	0,1515	0,1492	0,1469	0,1446	0,1423	0,1401	0,1379
-0,9	0,1841	0,1814	0,1788	0,1762	0,1736	0,1711	0,1685	0,1660	0,1635	0,1611
-0,8	0,2119	0,2090	0,2061	0,2033	0,2005	0,1977	0,1949	0,1922	0,1894	01867
-0,7	0,2420	0,2389	0,2358	0,2327	0,2297	0,2266	0,2236	0,2206	0,2177	0,2148
-0,6	0,2743	0,2709	0,2767	0,2643	0,2611	0,2578	0,2546	0,2514	0,2483	0,2451
-0,5	0,3085	0,3050	0,3015	0,2981	0,2946	0,2912	0,2877	0,2843	0,2810	0,2776
-0,4	0,3446	0,3409	0,3372	0,3336	0,3300	0,3264	0,3228	0,3192	0,3156	0,3121
-0,3	0,3821	0,3783	0,3745	0,3707	0,3669	0,3632	0,3594	0,3557	0,3520	0,3483
-0,2	0,4207	0,4168	0,4129	0,4090	0,4052	0,4013	0,3974	0,3936	0,3897	0,3859
-0,1	0,4602	0,4562	0,4522	0,4483	0,4443	0,4404	0,4364	0,4325	0,4286	0,4247
-0,0	0,5000	0,4960	0,4920	0,4880	0,4840	0,4801	0,4761	0,4721	0,4681	0,4641
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	0,7703	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	0,9278	0,9292	0,9306	0,9319
1,5	0,9332	0,9345	0,9357	0,9370	0,9382	0,9394	0,9406	0,9418	0,9430	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	0,9648	0,9656	0,9664	0,9671	0,9678	0,9686	0,9693	0,9700	0,9706
1,9	0,9713	0,9719	0,9726	0,9732	0,9738	0,9744	0,9750	0,9756	0,9762	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	0,9874	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	0,9987	0,9990	0,9993	0,9995	0,9997	0,9998	0,9998	0,9999	0,9999	1,000



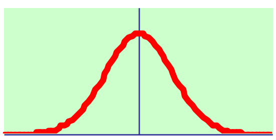
OBS.: **(1)** Se X não é padronizada, seus valores devem ser reduzidos por: $Z = (X - \mu)/\sigma$, i. é, $P(X \le x)$ $\Phi[(x - \mu)/\sigma]$.

(2) Para $z \ge 4$, $\Phi(z) = 1$ e para $z \le -4$, $\Phi(z) = 0$, com aproximação de 4 decimais.

(3) Os valores na linha do -3 e +3 tem aproximação decimal e não centesimal, como o resto da tabela.

Tabela 2 - Valores críticos da distribuição t de Student

	P(|t de Student| ≥ valor tabelado) = α ⇔ Valores bilaterais								
G. L.	**0.5000**	**0.2000**	**0.1000**	**0.0500**	**0.0400**	**0.0200**	**0.0100**	**0.0050**	**0.0010**
1	1.000	3.078	6.314	12.706	15.894	31.821	63.656	127.321	636.578
2	0.816	1.886	2.920	4.303	4.849	6.965	9.925	14.089	31.600
3	0.765	1.638	2.353	3.182	3.482	4.541	5.841	7.453	12.924
4	0.741	1.533	2.132	2.776	2.999	3.747	4.604	5.598	8.610
5	0.727	1.476	2.015	2.571	2.757	3.365	4.032	4.773	6.869
6	0.718	1.440	1.943	2.447	2.612	3.143	3.707	4.317	5.959
7	0.711	1.415	1.895	2.365	2.517	2.998	3.499	4.029	5.408
8	0.706	1.397	1.860	2.306	2.449	2.896	3.355	3.833	5.041
9	0.703	1.383	1.833	2.262	2.398	2.821	3.250	3.690	4.781
10	0.700	1.372	1.812	2.228	2.359	2.764	3.169	3.581	4.587
11	0.697	1.363	1.796	2.201	2.328	2.718	3.106	3.497	4.437
12	0.695	1.356	1.782	2.179	2.303	2.681	3.055	3.428	4.318
13	0.694	1.350	1.771	2.160	2.282	2.650	3.012	3.372	4.221
14	0.692	1.345	1.761	2.145	2.264	2.624	2.977	3.326	4.140
15	0.691	1.341	1.753	2.131	2.249	2.602	2.947	3.286	4.073
16	0.690	1.337	1.746	2.120	2.235	2.583	2.921	3.252	4.015
17	0.689	1.333	1.740	2.110	2.224	2.567	2.898	3.222	3.965
18	0.688	1.330	1.734	2.101	2.214	2.552	2.878	3.197	3.922
19	0.688	1.328	1.729	2.093	2.205	2.539	2.861	3.174	3.883
20	0.687	1.325	1.725	2.086	2.197	2.528	2.845	3.153	3.850
21	0.686	1.323	1.721	2.080	2.189	2.518	2.831	3.135	3.819
22	0.686	1.321	1.717	2.074	2.183	2.508	2.819	3.119	3.792
23	0.685	1.319	1.714	2.069	2.177	2.500	2.807	3.104	3.768
24	0.685	1.318	1.711	2.064	2.172	2.492	2.797	3.091	3.745
25	0.684	1.316	1.708	2.060	2.167	2.485	2.787	3.078	3.725
26	0.684	1.315	1.706	2.056	2.162	2.479	2.779	3.067	3.707
27	0.684	1.314	1.703	2.052	2.158	2.473	2.771	3.057	3.689
28	0.683	1.313	1.701	2.048	2.154	2.467	2.763	3.047	3.674
29	0.683	1.311	1.699	2.045	2.150	2.462	2.756	3.038	3.660
30	0.683	1.310	1.697	2.042	2.147	2.457	2.750	3.030	3.646
31	0.682	1.309	1.696	2.040	2.144	2.453	2.744	3.022	3.633
32	0.682	1.309	1.694	2.037	2.141	2.449	2.738	3.015	3.622
33	0.682	1.308	1.692	2.035	2.138	2.445	2.733	3.008	3.611
34	0.682	1.307	1.691	2.032	2.136	2.441	2.728	3.002	3.601
35	0.682	1.306	1.690	2.030	2.133	2.438	2.724	2.996	3.591
36	0.681	1.306	1.688	2.028	2.131	2.434	2.719	2.990	3.582
37	0.681	1.305	1.687	2.026	2.129	2.431	2.715	2.985	3.574
38	0.681	1.304	1.686	2.024	2.127	2.429	2.712	2.980	3.566
39	0.681	1.304	1.685	2.023	2.125	2.426	2.708	2.976	3.558
40	0.681	1.303	1.684	2.021	2.123	2.423	2.704	2.971	3.551
41	0.681	1.303	1.683	2.020	2.121	2.421	2.701	2.967	3.544
42	0.680	1.302	1.682	2.018	2.120	2.418	2.698	2.963	3.538
43	0.680	1.302	1.681	2.017	2.118	2.416	2.695	2.959	3.532
44	0.680	1.301	1.680	2.015	2.116	2.414	2.692	2.956	3.526
45	0.680	1.301	1.679	2.014	2.115	2.412	2.690	2.952	3.520
46	0.680	1.300	1.679	2.013	2.114	2.410	2.687	2.949	3.515
47	0.680	1.300	1.678	2.012	2.112	2.408	2.685	2.946	3.510
48	0.680	1.299	1.677	2.011	2.111	2.407	2.682	2.943	3.505
49	0.680	1.299	1.677	2.010	2.110	2.405	2.680	2.940	3.500
50	0.679	1.299	1.676	2.009	2.109	2.403	2.678	2.937	3.496
60	0.679	1.296	1.671	2.000	2.099	2.390	2.660	2.915	3.460
70	0.678	1.294	1.667	1.994	2.093	2.381	2.648	2.899	3.435
80	0.678	1.292	1.664	1.990	2.088	2.374	2.639	2.887	3.416
90	0.677	1.291	1.662	1.987	2.084	2.368	2.632	2.878	3.402
100	0.677	1.290	1.660	1.984	2.081	2.364	2.626	2.871	3.390
110	0.677	1.289	1.659	1.982	2.078	2.361	2.621	2.865	3.381
120	0.677	1.289	1.658	1.980	2.076	2.358	2.617	2.860	3.373
∞	0.674	1.282	1.645	1.960	2.054	2.326	2.576	2.807	3.290
	0,2500	**0,1000**	**0,0500**	**0,0250**	**0,0200**	**0,0100**	**0,0050**	**0,0025**	**0,0005**
	P(t de Student ≥ valor tabelado) = α ⇔ Valores unilaterais								

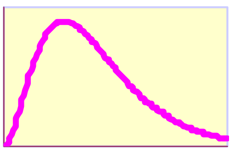


OBS.: **(1)** G. L. = Graus de Liberdade

(2) Para valores à esquerda, i. é, teste unilateral à esquerda (ou mesmo bilateral), basta trocar o sinal dos valores da tabela, pois a distribuição t é simétrica em torno de zero.

Tabela 3 - Valores críticos (unilaterais à esquerda) da distribuição Qui-Quadrado
P(χ^2 com n graus de liberdade ≥ valor tabelado) = α

	0,995	0,99	0,975	0,95	0,9	0,1	0,05	0,025	0,01	0,005
1	0,000	0,000	0,001	0,004	0,016	2,706	3,841	5,024	6,635	7,879
2	0,010	0,020	0,051	0,103	0,211	4,605	5,991	7,378	9,210	10,597
3	0,072	0,115	0,216	0,352	0,584	6,251	7,815	9,348	11,345	12,838
4	0,207	0,297	0,484	0,711	1,064	7,779	9,488	11,143	13,277	14,860
5	0,412	0,554	0,831	1,145	1,610	9,236	11,070	12,832	15,086	16,750
6	0,676	0,872	1,237	1,635	2,204	10,645	12,592	14,449	16,812	18,548
7	0,989	1,239	1,690	2,167	2,833	12,017	14,067	16,013	18,475	20,278
8	1,344	1,647	2,180	2,733	3,490	13,362	15,507	17,535	20,090	21,955
9	1,735	2,088	2,700	3,325	4,168	14,684	16,919	19,023	21,666	23,589
10	2,156	2,558	3,247	3,940	4,865	15,987	18,307	20,483	23,209	25,188
11	2,603	3,053	3,816	4,575	5,578	17,275	19,675	21,920	24,725	26,757
12	3,074	3,571	4,404	5,226	6,304	18,549	21,026	23,337	26,217	28,300
13	3,565	4,107	5,009	5,892	7,041	19,812	22,362	24,736	27,688	29,819
14	4,075	4,660	5,629	6,571	7,790	21,064	23,685	26,119	29,141	31,319
15	4,601	5,229	6,262	7,261	8,547	22,307	24,996	27,488	30,578	32,801
16	5,142	5,812	6,908	7,962	9,312	23,542	26,296	28,845	32,000	34,267
17	5,697	6,408	7,564	8,672	10,085	24,769	27,587	30,191	33,409	35,718
18	6,265	7,015	8,231	9,390	10,865	25,989	28,869	31,526	34,805	37,156
19	6,844	7,633	8,907	10,117	11,651	27,204	30,144	32,852	36,191	38,582
20	7,434	8,260	9,591	10,851	12,443	28,412	31,410	34,170	37,566	39,997
21	8,034	8,897	10,283	11,591	13,240	29,615	32,671	35,479	38,932	41,401
22	8,643	9,542	10,982	12,338	14,041	30,813	33,924	36,781	40,289	42,796
23	9,260	10,196	11,689	13,091	14,848	32,007	35,172	38,076	41,638	44,181
24	9,886	10,856	12,401	13,848	15,659	33,196	36,415	39,364	42,980	45,558
25	10,520	11,524	13,120	14,611	16,473	34,382	37,652	40,646	44,314	46,928
26	11,160	12,198	13,844	15,379	17,292	35,563	38,885	41,923	45,642	48,290
27	11,808	12,878	14,573	16,151	18,114	36,741	40,113	43,195	46,963	49,645
28	12,461	13,565	15,308	16,928	18,939	37,916	41,337	44,461	48,278	50,994
29	13,121	14,256	16,047	17,708	19,768	39,087	42,557	45,722	49,588	52,335
30	13,787	14,953	16,791	18,493	20,599	40,256	43,773	46,979	50,892	53,672
31	14,458	15,655	17,539	19,281	21,434	41,422	44,985	48,232	52,191	55,002
32	15,134	16,362	18,291	20,072	22,271	42,585	46,194	49,480	53,486	56,328
33	15,815	17,073	19,047	20,867	23,110	43,745	47,400	50,725	54,775	57,648
34	16,501	17,789	19,806	21,664	23,952	44,903	48,602	51,966	56,061	58,964
35	17,192	18,509	20,569	22,465	24,797	46,059	49,802	53,203	57,342	60,275
36	17,887	19,233	21,336	23,269	25,643	47,212	50,998	54,437	58,619	61,581
37	18,586	19,960	22,106	24,075	26,492	48,363	52,192	55,668	59,893	62,883
38	19,289	20,691	22,878	24,884	27,343	49,513	53,384	56,895	61,162	64,181
39	19,996	21,426	23,654	25,695	28,196	50,660	54,572	58,120	62,428	65,475
40	20,707	22,164	24,433	26,509	29,051	51,805	55,758	59,342	63,691	66,766
41	21,421	22,906	25,215	27,326	29,907	52,949	56,942	60,561	64,950	68,053
42	22,138	23,650	25,999	28,144	30,765	54,090	58,124	61,777	66,206	69,336
43	22,860	24,398	26,785	28,965	31,625	55,230	59,304	62,990	67,459	70,616
44	23,584	25,148	27,575	29,787	32,487	56,369	60,481	64,201	68,710	71,892
45	24,311	25,901	28,366	30,612	33,350	57,505	61,656	65,410	69,957	73,166
46	25,041	26,657	29,160	31,439	34,215	58,641	62,830	66,616	71,201	74,437
47	25,775	27,416	29,956	32,268	35,081	59,774	64,001	67,821	72,443	75,704
48	26,511	28,177	30,754	33,098	35,949	60,907	65,171	69,023	73,683	76,969
49	27,249	28,941	31,555	33,930	36,818	62,038	66,339	70,222	74,919	78,231
50	27,991	29,707	32,357	34,764	37,689	63,167	67,505	71,420	76,154	79,490



OBS.: **(1)** G.L. = Graus de Liberdade

(2) Para graus de liberdade que não estão na tabela, isto é acima de 50, use a aproximação:

$\chi_p^2 = \frac{1}{2}\left(z_p + \sqrt{2k-1}\right)^2$, onde z_p é o valor correspondente na normal padrão.

Tabela 4 - Valores críticos (unilaterais à direita) da distribuição F (significância de 5%)

	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**
1	161,45	199,50	215,71	224,58	230,16	233,99	236,77	238,88	240,54	241,88	242,98	243,90	244,69	245,36	245,95	246,47
2	18,51	19,00	19,16	19,25	19,30	19,33	19,35	19,37	19,38	19,40	19,40	19,41	19,42	19,42	19,43	19,43
3	10,13	9,55	9,28	9,12	9,01	8,94	8,89	8,85	8,81	8,79	8,76	8,74	8,73	8,71	8,70	8,69
4	7,71	6,94	6,59	6,39	6,26	6,16	6,09	6,04	6,00	5,96	5,94	5,91	5,89	5,87	5,86	5,84
5	6,61	5,79	5,41	5,19	5,05	4,95	4,88	4,82	4,77	4,74	4,70	4,68	4,66	4,64	4,62	4,60
6	5,99	5,14	4,76	4,53	4,39	4,28	4,21	4,15	4,10	4,06	4,03	4,00	3,98	3,96	3,94	3,92
7	5,59	4,74	4,35	4,12	3,97	3,87	3,79	3,73	3,68	3,64	3,60	3,57	3,55	3,53	3,51	3,49
8	5,32	4,46	4,07	3,84	3,69	3,58	3,50	3,44	3,39	3,35	3,31	3,28	3,26	3,24	3,22	3,20
9	5,12	4,26	3,86	3,63	3,48	3,37	3,29	3,23	3,18	3,14	3,10	3,07	3,05	3,03	3,01	2,99
10	4,96	4,10	3,71	3,48	3,33	3,22	3,14	3,07	3,02	2,98	2,94	2,91	2,89	2,86	2,85	2,83
11	4,84	3,98	3,59	3,36	3,20	3,09	3,01	2,95	2,90	2,85	2,82	2,79	2,76	2,74	2,72	2,70
12	4,75	3,89	3,49	3,26	3,11	3,00	2,91	2,85	2,80	2,75	2,72	2,69	2,66	2,64	2,62	2,60
13	4,67	3,81	3,41	3,18	3,03	2,92	2,83	2,77	2,71	2,67	2,63	2,60	2,58	2,55	2,53	2,51
14	4,60	3,74	3,34	3,11	2,96	2,85	2,76	2,70	2,65	2,60	2,57	2,53	2,51	2,48	2,46	2,44
15	4,54	3,68	3,29	3,06	2,90	2,79	2,71	2,64	2,59	2,54	2,51	2,48	2,45	2,42	2,40	2,38
16	4,49	3,63	3,24	3,01	2,85	2,74	2,66	2,59	2,54	2,49	2,46	2,42	2,40	2,37	2,35	2,33
17	4,45	3,59	3,20	2,96	2,81	2,70	2,61	2,55	2,49	2,45	2,41	2,38	2,35	2,33	2,31	2,29
18	4,41	3,55	3,16	2,93	2,77	2,66	2,58	2,51	2,46	2,41	2,37	2,34	2,31	2,29	2,27	2,25
19	4,38	3,52	3,13	2,90	2,74	2,63	2,54	2,48	2,42	2,38	2,34	2,31	2,28	2,26	2,23	2,21
20	4,35	3,49	3,10	2,87	2,71	2,60	2,51	2,45	2,39	2,35	2,31	2,28	2,25	2,22	2,20	2,18
21	4,32	3,47	3,07	2,84	2,68	2,57	2,49	2,42	2,37	2,32	2,28	2,25	2,22	2,20	2,18	2,16
22	4,30	3,44	3,05	2,82	2,66	2,55	2,46	2,40	2,34	2,30	2,26	2,23	2,20	2,17	2,15	2,13
23	4,28	3,42	3,03	2,80	2,64	2,53	2,44	2,37	2,32	2,27	2,24	2,20	2,18	2,15	2,13	2,11
24	4,26	3,40	3,01	2,78	2,62	2,51	2,42	2,36	2,30	2,25	2,22	2,18	2,15	2,13	2,11	2,09
25	4,24	3,39	2,99	2,76	2,60	2,49	2,40	2,34	2,28	2,24	2,20	2,16	2,14	2,11	2,09	2,07
26	4,23	3,37	2,98	2,74	2,59	2,47	2,39	2,32	2,27	2,22	2,18	2,15	2,12	2,09	2,07	2,05
27	4,21	3,35	2,96	2,73	2,57	2,46	2,37	2,31	2,25	2,20	2,17	2,13	2,10	2,08	2,06	2,04
28	4,20	3,34	2,95	2,71	2,56	2,45	2,36	2,29	2,24	2,19	2,15	2,12	2,09	2,06	2,04	2,02
29	4,18	3,33	2,93	2,70	2,55	2,43	2,35	2,28	2,22	2,18	2,14	2,10	2,08	2,05	2,03	2,01
30	4,17	3,32	2,92	2,69	2,53	2,42	2,33	2,27	2,21	2,16	2,13	2,09	2,06	2,04	2,01	1,99
31	4,16	3,30	2,91	2,68	2,52	2,41	2,32	2,25	2,20	2,15	2,11	2,08	2,05	2,03	2,00	1,98
32	4,15	3,29	2,90	2,67	2,51	2,40	2,31	2,24	2,19	2,14	2,10	2,07	2,04	2,01	1,99	1,97
33	4,14	3,28	2,89	2,66	2,50	2,39	2,30	2,23	2,18	2,13	2,09	2,06	2,03	2,00	1,98	1,96
34	4,13	3,28	2,88	2,65	2,49	2,38	2,29	2,23	2,17	2,12	2,08	2,05	2,02	1,99	1,97	1,95
35	4,12	3,27	2,87	2,64	2,49	2,37	2,29	2,22	2,16	2,11	2,07	2,04	2,01	1,99	1,96	1,94
36	4,11	3,26	2,87	2,63	2,48	2,36	2,28	2,21	2,15	2,11	2,07	2,03	2,00	1,98	1,95	1,93
37	4,11	3,25	2,86	2,63	2,47	2,36	2,27	2,20	2,14	2,10	2,06	2,02	2,00	1,97	1,95	1,93
38	4,10	3,24	2,85	2,62	2,46	2,35	2,26	2,19	2,14	2,09	2,05	2,02	1,99	1,96	1,94	1,92
39	4,09	3,24	2,85	2,61	2,46	2,34	2,26	2,19	2,13	2,08	2,04	2,01	1,98	1,95	1,93	1,91
40	4,08	3,23	2,84	2,61	2,45	2,34	2,25	2,18	2,12	2,08	2,04	2,00	1,97	1,95	1,92	1,90
41	4,08	3,23	2,83	2,60	2,44	2,33	2,24	2,17	2,12	2,07	2,03	2,00	1,97	1,94	1,92	1,90
42	4,07	3,22	2,83	2,59	2,44	2,32	2,24	2,17	2,11	2,06	2,03	1,99	1,96	1,94	1,91	1,89
43	4,07	3,21	2,82	2,59	2,43	2,32	2,23	2,16	2,11	2,06	2,02	1,99	1,96	1,93	1,91	1,89
44	4,06	3,21	2,82	2,58	2,43	2,31	2,23	2,16	2,10	2,05	2,01	1,98	1,95	1,92	1,90	1,88
45	4,06	3,20	2,81	2,58	2,42	2,31	2,22	2,15	2,10	2,05	2,01	1,97	1,94	1,92	1,89	1,87

Tabela 5 - Valores críticos (unilaterais à direita) da distribuição F (significância de 1%)

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	4052,18	4999,34	5403,53	5624,26	5763,96	5858,95	5928,33	5980,95	6022,40	6055,93	6083,40	6106,68	6125,77	6143,00	6156,97	6170,01
2	98,50	99,00	99,16	99,25	99,30	99,33	99,36	99,38	99,39	99,40	99,41	99,42	99,42	99,43	99,43	99,44
3	34,12	30,82	29,46	28,71	28,24	27,91	27,67	27,49	27,34	27,23	27,13	27,05	26,98	26,92	26,87	26,83
4	21,20	18,00	16,69	15,98	15,52	15,21	14,98	14,80	14,66	14,55	14,45	14,37	14,31	14,25	14,20	14,15
5	16,26	13,27	12,06	11,39	10,97	10,67	10,46	10,29	10,16	10,05	9,96	9,89	9,82	9,77	9,72	9,68
6	13,75	10,92	9,78	9,15	8,75	8,47	8,26	8,10	7,98	7,87	7,79	7,72	7,66	7,60	7,56	7,52
7	12,25	9,55	8,45	7,85	7,46	7,19	6,99	6,84	6,72	6,62	6,54	6,47	6,41	6,36	6,31	6,28
8	11,26	8,65	7,59	7,01	6,63	6,37	6,18	6,03	5,91	5,81	5,73	5,67	5,61	5,56	5,52	5,48
9	10,56	8,02	6,99	6,42	6,06	5,80	5,61	5,47	5,35	5,26	5,18	5,11	5,05	5,01	4,96	4,92
10	10,04	7,56	6,55	5,99	5,64	5,39	5,20	5,06	4,94	4,85	4,77	4,71	4,65	4,60	4,56	4,52
11	9,65	7,21	6,22	5,67	5,32	5,07	4,89	4,74	4,63	4,54	4,46	4,40	4,34	4,29	4,25	4,21
12	9,33	6,93	5,95	5,41	5,06	4,82	4,64	4,50	4,39	4,30	4,22	4,16	4,10	4,05	4,01	3,97
13	9,07	6,70	5,74	5,21	4,86	4,62	4,44	4,30	4,19	4,10	4,02	3,96	3,91	3,86	3,82	3,78
14	8,86	6,51	5,56	5,04	4,69	4,46	4,28	4,14	4,03	3,94	3,86	3,80	3,75	3,70	3,66	3,62
15	8,68	6,36	5,42	4,89	4,56	4,32	4,14	4,00	3,89	3,80	3,73	3,67	3,61	3,56	3,52	3,49
16	8,53	6,23	5,29	4,77	4,44	4,20	4,03	3,89	3,78	3,69	3,62	3,55	3,50	3,45	3,41	3,37
17	8,40	6,11	5,19	4,67	4,34	4,10	3,93	3,79	3,68	3,59	3,52	3,46	3,40	3,35	3,31	3,27
18	8,29	6,01	5,09	4,58	4,25	4,01	3,84	3,71	3,60	3,51	3,43	3,37	3,32	3,27	3,23	3,19
19	8,18	5,93	5,01	4,50	4,17	3,94	3,77	3,63	3,52	3,43	3,36	3,30	3,24	3,19	3,15	3,12
20	8,10	5,85	4,94	4,43	4,10	3,87	3,70	3,56	3,46	3,37	3,29	3,23	3,18	3,13	3,09	3,05
21	8,02	5,78	4,87	4,37	4,04	3,81	3,64	3,51	3,40	3,31	3,24	3,17	3,12	3,07	3,03	2,99
22	7,95	5,72	4,82	4,31	3,99	3,76	3,59	3,45	3,35	3,26	3,18	3,12	3,07	3,02	2,98	2,94
23	7,88	5,66	4,76	4,26	3,94	3,71	3,54	3,41	3,30	3,21	3,14	3,07	3,02	2,97	2,93	2,89
24	7,82	5,61	4,72	4,22	3,90	3,67	3,50	3,36	3,26	3,17	3,09	3,03	2,98	2,93	2,89	2,85
25	7,77	5,57	4,68	4,18	3,85	3,63	3,46	3,32	3,22	3,13	3,06	2,99	2,94	2,89	2,85	2,81
26	7,72	5,53	4,64	4,14	3,82	3,59	3,42	3,29	3,18	3,09	3,02	2,96	2,90	2,86	2,81	2,78
27	7,68	5,49	4,60	4,11	3,78	3,56	3,39	3,26	3,15	3,06	2,99	2,93	2,87	2,82	2,78	2,75
28	7,64	5,45	4,57	4,07	3,75	3,53	3,36	3,23	3,12	3,03	2,96	2,90	2,84	2,79	2,75	2,72
29	7,60	5,42	4,54	4,04	3,73	3,50	3,33	3,20	3,09	3,00	2,93	2,87	2,81	2,77	2,73	2,69
30	7,56	5,39	4,51	4,02	3,70	3,47	3,30	3,17	3,07	2,98	2,91	2,84	2,79	2,74	2,70	2,66
31	7,53	5,36	4,48	3,99	3,67	3,45	3,28	3,15	3,04	2,96	2,88	2,82	2,77	2,72	2,68	2,64
32	7,50	5,34	4,46	3,97	3,65	3,43	3,26	3,13	3,02	2,93	2,86	2,80	2,74	2,70	2,65	2,62
33	7,47	5,31	4,44	3,95	3,63	3,41	3,24	3,11	3,00	2,91	2,84	2,78	2,72	2,68	2,63	2,60
34	7,44	5,29	4,42	3,93	3,61	3,39	3,22	3,09	2,98	2,89	2,82	2,76	2,70	2,66	2,61	2,58
35	7,42	5,27	4,40	3,91	3,59	3,37	3,20	3,07	2,96	2,88	2,80	2,74	2,69	2,64	2,60	2,56
36	7,40	5,25	4,38	3,89	3,57	3,35	3,18	3,05	2,95	2,86	2,79	2,72	2,67	2,62	2,58	2,54
37	7,37	5,23	4,36	3,87	3,56	3,33	3,17	3,04	2,93	2,84	2,77	2,71	2,65	2,61	2,56	2,53
38	7,35	5,21	4,34	3,86	3,54	3,32	3,15	3,02	2,92	2,83	2,75	2,69	2,64	2,59	2,55	2,51
39	7,33	5,19	4,33	3,84	3,53	3,30	3,14	3,01	2,90	2,81	2,74	2,68	2,62	2,58	2,54	2,50
40	7,31	5,18	4,31	3,83	3,51	3,29	3,12	2,99	2,89	2,80	2,73	2,66	2,61	2,56	2,52	2,48
41	7,30	5,16	4,30	3,81	3,50	3,28	3,11	2,98	2,87	2,79	2,71	2,65	2,60	2,55	2,51	2,47
42	7,28	5,15	4,29	3,80	3,49	3,27	3,10	2,97	2,86	2,78	2,70	2,64	2,59	2,54	2,50	2,46
43	7,26	5,14	4,27	3,79	3,48	3,25	3,09	2,96	2,85	2,76	2,69	2,63	2,57	2,53	2,49	2,45
44	7,25	5,12	4,26	3,78	3,47	3,24	3,08	2,95	2,84	2,75	2,68	2,62	2,56	2,52	2,47	2,44
45	7,23	5,11	4,25	3,77	3,45	3,23	3,07	2,94	2,83	2,74	2,67	2,61	2,55	2,51	2,46	2,43

Bruno Monte de Castro – IME – USP

Exponencial $[Exp(\lambda),\ \lambda > 0]$

$f(x) = \lambda \exp\{-\lambda x\} \underset{(0,\infty)}{I}(x)$; $\mathbb{E}(X) = \frac{1}{\lambda}$; $\mathbb{E}(X^k) = \frac{k!}{\lambda^k}$; $K(X) = 6$;

$F(x) = 1 - \exp\{-\lambda x\} \underset{(0,\infty)}{I}(x)$; $x_p = \frac{-\ln(1-p)}{\lambda}$; $Var(X) = \frac{1}{\lambda^2}$;

$Mo(X) = 0$; $M_X(t) = \frac{\lambda}{\lambda - t},\ t < \lambda$; $A(X) = 2$; $Med(X) = \frac{\ln 2}{\lambda}$;

- ∘ $X \sim Exp(\lambda) \Rightarrow Y = \lambda X \sim Exp(1)$
- ∘ $X_1, \cdots, X_n$ i.i.d. com $X_i \sim Exp(\lambda) \Rightarrow Y = \min\{X_1, \cdots, X_n\} \sim Exp(n\lambda)$
- ∘ X_1, X_2 i.i.d. com $X_i \sim Exp(\lambda) \Rightarrow Y = X_1 - X_2 \sim La(0, \lambda)$
- ∘ $X \sim Exp(\lambda) \Leftrightarrow X \sim Gama(1, \lambda)$
- ∘ $X_1, \cdots, X_n$ i.i.d. com $X_i \sim Exp(\lambda) \Rightarrow Y = \sum_{i=1}^n X_i \sim Gama(n, \lambda)$

Weibull I $[W_1(a,b),\ a > 0, b > 0]$

$f(x) = abx^{b-1}\exp\{-ax^b\} \underset{(0,\infty)}{I}(x)$; $Med(X) = (\frac{\ln 2}{a})^{1/b}$;

$Mo(X) = (\frac{b-1}{ab})^{1/b}$; $F(x) = 1 - \exp\{-ax^b\} \underset{(0,\infty)}{I}(x)$;

$Var(X) = a^{\frac{-2}{b}}[\Gamma(1+\frac{2}{b}) - \Gamma^2(1+\frac{1}{b})]$;

$x_p = [-\frac{\ln(1-p)}{a}]^{\frac{1}{b}}$; $\mathbb{E}(X^k) = a^{\frac{-k}{b}}\Gamma(1+\frac{k}{b})$;

Weibull II $[W_2(\theta,\lambda),\ \theta > 0, \lambda > 0]$

$f(x) = (\frac{\theta}{\lambda})(\frac{x}{\lambda})^{\theta-1}\exp\{-(\frac{x}{\lambda})^\theta\} \underset{(0,\infty)}{I}(x)$;

$F(x) = 1 - \exp\{-(\frac{x}{\lambda})^\theta\} \underset{(0,\infty)}{I}(x)$;

$\mathbb{E}(X^k) = \lambda^k\Gamma(1+\frac{k}{\theta})$;

$Var(X) = \lambda^2[\Gamma(1+\frac{2}{\theta}) - \Gamma^2(1+\frac{1}{\theta})]$;

- ∘ $X \sim Rayleigh(\beta) \Leftrightarrow X \sim W_1(1/2\beta^2, 2)$
- ∘ $X \sim W_1(\lambda^{-1/\theta}, \theta) \Leftrightarrow X \sim W_2(\lambda, \theta)$
- ∘ $X \sim W_1(\lambda, 1) \Leftrightarrow X \sim Exp(\lambda)$
- ∘ $X \sim W_1(a,b)$ e $Y = -\ln X \Rightarrow$ $Y \sim Gumbel(b^{-1}\ln a, b^{-1})$

Normal $[N(\mu,\sigma^2),\ \mu \in \mathbb{R}, \sigma > 0]$

$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}}\exp\Big\{-\frac{(x-\mu)^2}{2\sigma^2}\Big\} \underset{(-\infty,\infty)}{I}(x)$; $\mathbb{E}[(X - \mathbb{E}(X))^k] = \begin{cases} 0, & \text{se } k \text{ é ímpar} \\ \frac{\sigma^k k!}{2^{k/2}\Gamma(1+\frac{k}{2})}, & \text{se } k \text{ é par} \end{cases}$

$A(X) = K(X) = 0$; $M_X(t) = e^{\mu t + \sigma^2\frac{t^2}{2}}$; $\mathbb{E}(X) = Med(X) = Mo(X) = \mu$; $Var(X) = \sigma^2$;

- ∘ $X_1, \cdots, X_n$ i.i.d. com $X_i \sim N(0,1) \Rightarrow$ $Y = \sum_{i=1}^n X_i^2 \sim \chi^2_{(n)}$
- ∘ $X \sim N(\mu, \sigma^2) \Rightarrow Y = (X-\mu)/\sigma \sim N(0,1)$
- ∘ $X_1, \cdots, X_n$ i.i.d. com $X_i \sim N(\mu, \sigma^2) \Rightarrow$ $Y = \sum_{i=1}^n a_i X_i \sim N(\mu\sum_{i=1}^n a_i, \sigma^2\sum_{i=1}^n a_i^2)$
- ∘ X_1, X_2 i.i.d. com $X_i \sim N(0,1) \Rightarrow Y = \frac{X_1}{X_2} \sim Ca(0,1)$

Uniforme $[U(a,b),\ -\infty < a < b < \infty]$

$F(x) = \begin{cases} 0, & \text{se } x < a \\ \frac{x-a}{b-a}, & \text{se } a \le x < b \\ 1, & \text{se } x \ge b \end{cases}$;

$f(x) = \frac{1}{(b-a)} \underset{(a,b)}{I}(x)$; $\mathbb{E}(X) = Med(X) = \frac{a+b}{2}$; $Var(X) = \frac{(b-a)^2}{12}$;

$\mathbb{E}(X^k) = \sum_{i=0}^k \frac{a^{k-i}b^i}{k+1}$; $\mathbb{E}[(X - \mathbb{E}(X))^k] = \begin{cases} 0, & \text{se } k \text{ é ímpar} \\ \frac{(b-a)^k}{2^k(k+1)}, & \text{se } k \text{ é par} \end{cases}$

$M_X(t) = \frac{e^{tb} - e^{ta}}{t(b-a)}$; $A(X) = 0$; $K(X) = -\frac{6}{5}$;

- ∘ $X \sim U(0,1) \Rightarrow Y = 1 - X^{\frac{1}{n}} \sim Beta(1,n)$
- ∘ $X \sim U(0,1) \Rightarrow Y = X^2 \sim Beta(1/2, 1)$
- ∘ $X \sim U(0,1) \Rightarrow Y = -\frac{\ln X}{\lambda} \sim Exp(\lambda)$
- ∘ $X_1, \cdots, X_n$ i.i.d. com $X_i \sim U(0,1) \Rightarrow$ $Y = \min\{X_1, \cdots, X_n\} \sim Beta(1,n)$ e $Z = \max\{X_1, \cdots, X_n\} \sim Beta(n,1)$
- ∘ $X \sim U(0,1) \Rightarrow Y = tg[\pi(x - \frac{1}{2})] \sim Ca(0,1)$

Rayleigh $[Rayleigh(\sigma),\ \sigma > 0]$

$f(x) = \frac{x}{\sigma^2}\exp\{-\frac{x^2}{2\sigma^2}\} \underset{(0,\infty)}{I}(x)$; $F(x) = 1 - \exp\{\frac{-x^2}{2\sigma^2}\} \underset{(0,\infty)}{I}(x)$; $A(X) = \frac{(\pi-3)\sqrt{\frac{\pi}{2}}}{(2-\frac{\pi}{2})^{\frac{3}{2}}}$; $Var(X) = \frac{\sigma^2(4-\pi)}{2}$;

$Med(X) = \sigma\sqrt{\ln 4}$; $\mathbb{E}(X^k) = 2^{\frac{k}{2}}\sigma^k\Gamma(\frac{k}{2}+1)$; $\mathbb{E}(X) = \sigma\sqrt{\frac{\pi}{2}}$; $K(X) = \frac{8 - \frac{3\pi^2}{4}}{(2-\frac{\pi}{2})^2} - 3$; $Mo(X) = \sigma$;

- ∘ $X \sim Rayleigh(1) \Leftrightarrow X \sim \chi^2_{(2)}$
- ∘ $X_1, \cdots, X_n$ i.i.d. com $X_i \sim Rayleigh(\sigma) \Rightarrow$ $Y = \sum_{i=1}^n X_i^2 \sim Gama(n, 2\sigma^2)$

Pareto $[Pareto(\alpha,\beta),\ \alpha > 0, \beta > 0]$

$f(x) = \frac{\alpha\beta^\alpha}{x^{\alpha+1}} \underset{(\beta,\infty)}{I}(x)$; $F(x) = 1 - \big(\frac{\beta}{x}\big)^\alpha \underset{(\beta,\infty)}{I}(x)$; $\mathbb{E}(X^k) = \frac{\alpha\beta^k}{\alpha - k}$; $Var(X) = \frac{\alpha\beta^2}{(\alpha-2)(\alpha-1)^2}$;

$x_p = \frac{\beta}{\sqrt[\alpha]{(1-p)}}$; $Med(X) = \beta\sqrt[\alpha]{2}$; $A(X) = \frac{2(\alpha+1)}{\alpha-3}\sqrt{\frac{\alpha-2}{\alpha}}$; $K(X) = \frac{6(\alpha^3+\alpha^2-6\alpha-2)}{\alpha(\alpha-3)(\alpha-4)}$; $Mo(X) = \beta$;

- ∘ $X_1, \cdots, X_n$ i.i.d. com $X_i \sim Pareto(\alpha, \beta) \Rightarrow$ $Y = \min\{X_1, \cdots, X_n\} \sim Pareto(n\alpha, \beta)$
- ∘ $X \sim Pareto(\alpha, \beta) \Rightarrow Y = \ln(\frac{X}{\beta}) \sim Exp(\alpha)$

Log – Normal $[LN(\mu,\sigma),\ \mu \in \mathbb{R},\ \sigma > 0]$

$f(x) = \frac{1}{x\sqrt{2\pi\sigma^2}}\exp\{\frac{-(\ln x - \mu)^2}{2\sigma^2}\} \underset{(0,\infty)}{I}(x)$; $\mathbb{E}(X^k) = e^{k\mu + \frac{k^2\sigma^2}{2}}$; $Var(X) = e^{2\mu+\sigma^2}(e^{\sigma^2} - 1)$;

$Med(X) = \mu$; $Mo(X) = e^{\mu - \sigma^2}$; $A(X) = (e^{\sigma^2}+2)\sqrt{e^{\sigma^2}-1}$; $K(X) = e^{4\sigma^2} + 2e^{3\sigma^2} + 3e^{2\sigma^2} - 6$;

- ∘ $X \sim LN(\mu, \sigma) \Rightarrow Y = \ln X \sim N(\mu, \sigma^2)$
- ∘ $X_1, \cdots, X_n$ i.i.d. com $X_i \sim LN(\mu, \sigma) \Rightarrow$ $Y = \prod_{i=1}^n X_i \sim LN(n\mu, \sqrt{n}\sigma)$

Cauchy $[Ca(a,b),\ a \in \mathbb{R},\ b > 0]$

$f(x) = \frac{1}{\pi b\big[1+\big(\frac{x-a}{b}\big)^2\big]} \underset{(-\infty,\infty)}{I}(x)$; $F(x) = \frac{1}{2} + tg^{-1}\big(\frac{x-a}{b}\big)$;

$Med(X) = a$; $C_X(t) = e^{iat - b|t|}$; $\mathbb{E}(X^k) = indefinido$; $Mo(X) = a$;

- ∘ $X \sim Ca(a,b) \Rightarrow Y = cX + d \sim Ca(ac+d, b|c|)$
- ∘ $X_1, \cdots, X_n$ i.i.d. com $X_i \sim Ca(a_i, b_i) \Rightarrow Y = \sum_{i=1}^n X_i \sim Ca(\sum_{i=1}^n a_i, \sum_{i=1}^n b_i)$
- ∘ $X \sim Ca(0,1) \Rightarrow Y = \frac{2X}{1-X^2} \sim Ca(0,1)$
- ∘ $X \sim Ca(0,b) \Rightarrow Y = 1/X \sim Ca(0, \frac{1}{b})$

Gama $[Gama(\alpha,\beta),\ \alpha > 0,\ \beta > 0]$

$f(x) = \frac{\beta^\alpha}{\Gamma(\alpha)}x^{\alpha-1}e^{-\beta x} \underset{(0,\infty)}{I}(x)$; $\mathbb{E}(X^k) = \frac{\Gamma(\alpha+k)}{\Gamma(\alpha)\beta^k}$; $\mathbb{E}(X) = \frac{\alpha}{\beta}$; $Var(X) = \frac{\alpha}{\beta^2}$;

$A(X) = \frac{2}{\sqrt{\alpha}}$; $K(X) = \frac{6}{\alpha}$; $M_X(t) = \big(\frac{\beta}{\beta - t}\big)^\alpha,\ t < \beta$; $Mo(X) = \frac{\alpha-1}{\beta}$;

- ∘ $X \sim Gama(\frac{n}{2}, \frac{1}{2}) \Leftrightarrow X \sim \chi^2_{(n)}$
- ∘ $X \sim Maxwell(a) \Rightarrow Y = X^2 \sim Gama(\frac{3}{2}, \frac{1}{2a^2})$
- ∘ $X_1,\ X_2$ ind. com $X_1 \sim Gama(\alpha, \beta)$ e $X_2 \sim Gama(\theta, \beta) \Rightarrow$ $Y = \frac{X_1}{X_1+X_2} \sim Beta(\alpha, \theta)$

Beta $[Beta(a,b),\ a > 0,\ b > 0]$

$f(x) = \frac{1}{B(a,b)}x^{a-1}(1-x)^{b-1} \underset{(0,1)}{I}(x)$; $\mathbb{E}(X^k) = \frac{\Gamma(a+b)\Gamma(a+k)}{\Gamma(a+b+k)\Gamma(a)}$; $Var(X) = \frac{ab}{(a+b)^2(a+b+1)}$;

$\mathbb{E}(X) = \frac{a}{a+b}$; $Mo(X) = \frac{a-1}{a+b+2}$; $A(X) = \frac{2(b-a)\sqrt{a+b+1}}{(a+b+2)\sqrt{ab}}$; $K(X) = \frac{3(a+b+1)[2(a+b)^2 + ab(a+b+6)]}{ab(a+b+2)(a+b+3)}$;

- ∘ $X \sim Beta(a,b) \Rightarrow Y = 1 - X \sim B(b,a)$
- ∘ $X \sim Beta(a,b) \Rightarrow Y = \frac{X}{1-X} \sim Beta\ Prime(a,b)$
- ∘ $X \sim Beta(\frac{n}{2}, \frac{m}{2}) \Rightarrow Y = \frac{mX}{n(1-X)} \sim F(n,m)$
- ∘ $X \sim Beta(a,1) \Rightarrow Y = -\ln X \sim Exp(a)$

Qui – Quadrado $[\chi^2_{(n)},\ n > 0]$

$f(x) = \frac{(\frac{1}{2})^{\frac{n}{2}}}{\Gamma(\frac{n}{2})}x^{\frac{n}{2}-1}e^{-\frac{x}{2}} \underset{(0,\infty)}{I}(x)$; $\mathbb{E}(X^k) = 2^k\frac{\Gamma(k+\frac{n}{2})}{\Gamma(\frac{n}{2})}$; $\mathbb{E}(X) = n$; $Var(X) = 2n$;

$A(X) = \sqrt{\frac{8}{n}}$; $K(X) = \frac{12}{n}$; $M_X(t) = (1-2t)^{-\frac{n}{2}}$; $Mo(X) = n - 2,\ n > 2$;

- ∘ $X \sim \chi^2_{(2)} \Leftrightarrow X \sim Exp(\frac{1}{2})$
- ∘ $X_1 \sim \chi^2_{(n)}$ e $X_2 \sim \chi^2_{(m)} \Rightarrow Y = \frac{X_1/n}{X_2/m} \sim F(n,m)$
- ∘ $X \sim \chi^2_{(n)}$ e $c > 0 \Rightarrow Y = \frac{X}{c} \sim Gama(\frac{n}{2}, \frac{c}{2})$
- ∘ $X_1, \cdots, X_n$ ind. com $X_i \sim \chi^2_{(n_i)} \Rightarrow Y = \sum_{i=1}^n X_i \sim \chi^2_{(\sum_{i=1}^n n_i)}$

t – Student $[t(n),\ n > 0]$

$f(x) = \frac{\Gamma(\frac{n+1}{2})}{\Gamma(\frac{n}{2})\sqrt{\pi n}}\left(1 + \frac{x^2}{n}\right)^{-\frac{(n+1)}{2}} \underset{(-\infty,\infty)}{I}(x)$; $\mathbb{E}(X^k) = \begin{cases} 0, & \text{se } k \text{ é ímpar} \\ \frac{\Gamma(\frac{k+1}{2})\Gamma(\frac{n-k}{2})n^{\frac{k}{2}}}{\sqrt{\pi}\Gamma(\frac{n}{2})}, & \text{se } k \text{ é par} \end{cases}$

$Mo(X) = Med(X) = 0$; $Var(X) = \frac{n}{(n-2)},\ n > 2$; $A(X) = 0,\ n > 3$; $K(X) = \frac{6}{n-4},\ n > 4$;

- ∘ $X \sim t(1) \Leftrightarrow X \sim Ca(0,1)$
- ∘ $X_1 \sim N(0,1)$ e $X_2 \sim \chi^2_{(n)}$, ind. $\Rightarrow Y = \frac{X_1}{\sqrt{X_2/n}}$
- ∘ $X \sim t(n) \Rightarrow Y = X^2 \sim F(1,n)$

F – Snedecor $[F(n,m),\ n > 0,\ m > 0]$

$f(x) = \frac{n^{\frac{n}{2}}m^{\frac{m}{2}}x^{\frac{n}{2}-1}}{B(\frac{n}{2},\frac{m}{2})(m+nx)^{\frac{n+m}{2}}} \underset{(0,\infty)}{I}(x)$; $\mathbb{E}(X^k) = \frac{(\frac{m}{n})^k\Gamma(\frac{n}{2}+k)\Gamma(\frac{m}{2}-k)}{\Gamma(\frac{n}{2})\Gamma(\frac{m}{2})}$; $Var(X) = \frac{2m^2(n+m-2)}{n(m-2)^2(m-4)},\ m > 4$;

$Mo(X) = \frac{m(n-2)}{n(m+2)}$; $A(X) = \frac{(2n+m-2)\sqrt{8(m-4)}}{(m-6)\sqrt{n(n+m-2)}},\ m > 6$; $K(X) = \frac{12[(m-2)^2(m-4) + n(n+m-2)(5m-22)]}{n(m-6)(m-8)(n+m-2)},\ m > 8$;

- ∘ $X \sim F(n,m) \Rightarrow Y = \frac{n\frac{X}{m}}{1+n\frac{X}{m}} \sim B(\frac{n}{2}, \frac{m}{2})$
- ∘ $X \sim F(n,m) \Rightarrow Y = \lim_{m\to\infty} nX \sim \chi^2_{(n)}$
- ∘ $X \sim F(n,m) \Rightarrow Y = X^{-1} \sim F(m,n)$ e $x_p = \frac{1}{y_{1-p}}$

Bruno Monte de Castro – IME – USP

Laplace $[La(a,b),\ a \in \mathbb{R},\ b>0]$

$f(x) = \frac{1}{2b} e^{-\frac{|x-a|}{b}} \underset{(-\infty,\infty)}{I}(x)$; $M_X(t) = \frac{e^{at}}{1-(bt)^2},\ |t| < \frac{1}{b}$; $F(x) = \begin{cases} \frac{1}{2} e^{\frac{x-a}{b}}, & \text{se } x < a \\ 1 - \frac{1}{2} e^{-\frac{x-a}{b}}, & \text{se } x \geq a \end{cases}$

$\mathbb{E}[(X-\mathbb{E}(X))^k] = \begin{cases} 0, & \text{se } k \text{ é ímpar} \\ k!b^k, & \text{se } k \text{ é par} \end{cases}$ $\mathbb{E}(X) = Med(X) = Mo(X) = a$; $Var(X) = 2b^2$; $A(X) = 0$; $K(X) = 3$;

- ∘ $X \sim La(a,b) \Rightarrow Y = cX + d \sim La(ca+d, cb)$
- ∘ $X \sim La(a,b) \Rightarrow Y = |X-a| \sim Exp(1/b)$
- ∘ X_1, X_2 i.i.d. com $X_i \sim La(a,b) \Rightarrow Y = \frac{|X_1-a|}{|X_2-a|} \sim F(2,2)$
- ∘ $X_1|X_2 \sim N(a, X_2)$ com $X_2 \sim Rayleigh(b) \Rightarrow X_1 \sim La(0,1)$
- ∘ X_1, X_2 i.i.d. com $X_i \sim U(0,1) \Rightarrow Y = \ln(X_1/X_2) \sim La(0,1)$

Logistica $[Logistica(a,b),\ a \in \mathbb{R},\ b>0]$

$f(x) = \frac{\exp\{-(x-a)/b\}}{b(1+\exp\{-(x-a)/b\})^2} \underset{(-\infty,\infty)}{I}(x)$; $F(X) = (1+\exp\{-(x-a)/b\})^{-1}$;

$\mathbb{E}(X) = Med(X) = Mo(X) = a$; $M_X(t) = e^{at}B(1-bt, 1+bt)$;

$A(X) = 0$; $K(X) = 1,2$; $Var(X) = \frac{(b\pi)^2}{3}$;

- ∘ $X \sim Logistica(a,b) \Rightarrow Y = cX + d \sim Logistica(ca+d, cb)$
- ∘ $X \sim U(0,1) \Rightarrow Y = a + b(\ln(X) - \ln(1-X)) \sim Logistica(a,b)$
- ∘ $X \sim Exp(1) \Rightarrow Y = a - b\ln(\frac{e^{-X}}{1-e^{-X}}) \sim Logistica(a,b)$
- ∘ X_1, X_2 i.i.d. com $X_i \sim Exp(1) \Rightarrow Y = a - b\ln\left(\frac{X_1}{X_2}\right) \sim Logistica(a,b)$

Triangular $[Tri(a,c,b),\ \infty < a \leq c \leq b < \infty,\ a < b]$

$f(x) = \frac{2(x-a)}{(b-a)(c-a)} \underset{(a,c)}{I}(x) + \frac{2(b-x)}{(b-a)(b-c)} \underset{(c,b)}{I}(x)$; $\mathbb{E}(X) = \frac{a+b+c}{3}$; $Med(X) = \begin{cases} a + \frac{\sqrt{(b-a)(c-a)}}{\sqrt{2}}, & \text{se } c \geq \frac{a+b}{2} \\ b - \frac{\sqrt{(b-a)(b-c)}}{\sqrt{2}}, & \text{se } c \leq \frac{a+b}{2} \end{cases}$

$F(x) = \frac{(x-a)^2}{(b-a)(c-a)} \underset{(a,c)}{I}(x) + \left(1 - \frac{(b-x)^2}{(b-a)(b-c)}\right) \underset{(c,b)}{I}(x) + \underset{(b,\infty)}{I}(x)$; $Var(X) = \frac{a^2+b^2+c^2-ac-ab-bc}{18}$; $Mo(X) = c$;

$M_X(t) = \frac{2(b-c)e^{at} - (b-a)e^{ct} + (c-a)e^{bt}}{t^2(b-a)(c-a)(b-c)}$; $A(X) = \frac{\sqrt{2}(a+b-2c)(2a-b-c)(a-2b+c)}{5(a^2+b^2+c^2-ab-ac-bc)^{3/2}}$; $K(X) = -\frac{3}{5}$;

- ∘ X_1, X_2 i.i.d. com $X_i \sim U(0,1) \Rightarrow$ $Y = \frac{X_1+X_2}{2} \sim Tri(0, \frac{1}{2}, 1)$

Kumaraswamy $[Kum(a,b),\ a>0,\ b>0]$

$f(x) = abx^{a-1}(1-x^a)^{b-1} \underset{(0,1)}{I}(x)$; $F(x) = 1 - (1-x^a)^b \underset{(0,1)}{I}(x) + \underset{(1,\infty)}{I}(x)$; $\mathbb{E}(X^k) = bB(1+\frac{k}{a}, b)$;

$Mo(X) = \left(\frac{a-1}{ab-1}\right)^{1/a}$; $Med(X) = (1-2^{-1/b})^{1/a}$; $Var(X) = bB(1+\frac{2}{a}, b) - b^2B(1+\frac{1}{a}, b)^2$;

- ∘ $X \sim U(0,1) \Rightarrow Y = (1-(1-X)^{\frac{1}{b}})^{\frac{1}{a}} \sim Kum(a,b)$
- ∘ $X \sim Kum(a,1) \Rightarrow Y = (1-X) \sim Kum(1,a)$
- ∘ $X \sim Kum(1,1) \Leftrightarrow X \sim U(0,1)$
- ∘ $X \sim Kum(a,1) \Rightarrow Y = -\ln(X) \sim Exp(a)$

Beta Prime $[Beta\ Prime(a,b),\ a>0,\ b>0]$

$f(x) = \frac{x^{a-1}}{B(a,b)(x+1)^{a+b}} \underset{(0,\infty)}{I}(x)$; $\mathbb{E}(X^k) = \frac{B(a+k, b-k)}{B(a,b)}$; $Var(X) = \frac{a(a+b-1)}{(b-2)(b-1)^2},\ b>2$;

$Mo(X) = \frac{a-1}{b+1},\ a \geq 1$; $A(X) = \frac{2(2a+b-1)}{b-3}\sqrt{\frac{b-2}{a(a+b-1)}},\ b>3$;

- ∘ $X \sim Beta\ Prime(a,b) \Rightarrow Y = \frac{1}{X} \sim B(b,a)$
- ∘ $X \sim F(a,b) \Rightarrow Y = \frac{a}{b}X \sim Beta\ Prime(\frac{a}{2}, \frac{b}{2})$
- ∘ X_1, X_2 ind. com $X_i \sim G(a_i, 1)$ $\Rightarrow Y = \frac{X_1}{X_2} \sim Beta\ Prime(a_1, a_2)$

Binomial $[Bin(n,p),\ n \in \mathbb{N}, 0 \leq p \leq 1]$

$\mathbb{P}(X=x) = \binom{n}{x}p^x(1-p)^{n-x} \underset{\{0,\cdots,n\}}{I}(x)$; $\mathbb{E}(X) = np$; $M_X(t) = (pe^t + 1 - p)^n$; $\mathbb{E}\left[\frac{X!(n-k)!}{(X-k)!n!}\right] = p^k$;

$Med(X) = \begin{matrix} \lfloor np \rfloor \\ \lceil np \rceil \end{matrix}$ ou; $Mo(X) = \begin{matrix} \lfloor (n+1)p \rfloor \\ \lceil (n+1)p \rceil \end{matrix}$ ou; $Var(X) = np(1-p)$; $A(X) = \frac{1-2p}{\sqrt{np(1-p)}}$; $K(X) = \frac{1-6p(1-p)}{np(1-p)}$;

- ∘ $X \sim Bin(1,p) \Leftrightarrow X \sim Ber(p)$
- ∘ $X_1, \cdots, X_n$ i.i.d com $X_i \sim Ber(p) \Rightarrow$ $Y = \sum_{i=1}^n X_i \sim Bin(n,p)$
- ∘ $X_1, \cdots, X_n$ ind com $X_i \sim Bin(n_i, p) \Rightarrow$ $Y = \sum_{i=1}^n X_i \sim Bin(\sum_{i=1}^n n_i, p)$

Geométrica I $[G_0(p),\ 0 \leq p \leq 1]$

$\mathbb{P}(X=x) = p(1-p)^x \underset{\{0,1,\cdots\}}{I}(x)$; $M_X(t) = \frac{p}{1-(1-p)e^t}$;

$Mo(X) = 0$; $F(x) = 1 - (1-p)^{x+1}$; $\mathbb{E}(X) = \frac{1-p}{p}$;

$A(X) = \frac{2-p}{\sqrt{1-p}}$; $K(X) = 6 + \frac{p^2}{1-p}$; $Var(X) = \frac{1-p}{p^2}$;

Geométrica II $[G_1(p),\ 0 \leq p \leq 1]$

$\mathbb{P}(X=x) = p(1-p)^{x-1} \underset{\{1,2,\cdots\}}{I}(x)$; $M_X(t) = \frac{pe^t}{1-(1-p)e^t}$;

$Mo(X) = 1$; $F(x) = 1 - (1-p)^x$; $\mathbb{E}(X) = \frac{1}{p}$;

$A(X) = \frac{2-p}{\sqrt{1-p}}$; $K(X) = \frac{6-6p+p^2}{(1-p)}$; $Var(X) = \frac{1-p}{p^2}$;

- ∘ $X_1, \cdots, X_n$ ind. com $X_i \sim G_1(p_i) \Rightarrow$ $Y = \min_{i=1,\cdots,n}\{X_i\} \sim G_1(1 - \prod_{i=1}^n (1-p_i))$
- ∘ $X \sim Exp(\lambda) \Rightarrow Y = \lfloor X \rfloor \sim G_0(1 - e^{-\lambda})$
- ∘ $X_1, \cdots, X_r$ i.i.d. com $X_i \sim G_0(p) \Rightarrow$ $Y = \sum_{i=1}^r X_i \sim BN(r,p)$

Binomial Negativa $[BN(r,p),\ r>0,\ 0 \leq p \leq 1]$

$\mathbb{P}(X=x) = \frac{\Gamma(r+x)}{\Gamma(x+1)\Gamma(r)}p^r(1-p)^x \underset{\{0,1,\cdots\}}{I}(x)$; $K(X) = \frac{p^2+6(1-p)}{r(1-p)}$;

$\mathbb{E}(X) = \frac{r(1-p)}{p}$; $Var(X) = \frac{r(1-p)}{p^2}$; $\mathbb{E}\left[\frac{X!}{(X-k)!}\right] = \frac{(r+k-1)!(1-p)^k}{(r-1)!p^k}$;

$A(X) = \frac{2-p}{\sqrt{r(1-p)}}$; $M_X(t) = \left(\frac{p}{1-(1-p)e^t}\right)^r$; $Mo(X) = \lfloor \frac{(r-1)(1-p)}{p} \rfloor$;

Pascal $[Pa(r,p),\ r \in \mathbb{N},\ 0 \leq p \leq 1]$

$\mathbb{P}(X=x) = \binom{x-1}{r-1}p^r(1-p)^{x-r} \underset{\{r,r+1,\cdots\}}{I}(x)$; $A(X) = \frac{2-p}{\sqrt{r(1-p)}}$;

$\mathbb{E}\left[\frac{(X+k)!}{(X-1)!}\right] = \frac{(k+r)!}{(r-1)!p^{k+1}}$; $K(X) = \frac{6-6p+p^2}{r(1-p)}$;

$\mathbb{E}(X) = \frac{r}{p}$; $M_X(t) = \left(\frac{pe^t}{1-(1-p)e^t}\right)^r$; $Var(X) = \frac{r(1-p)}{p^2}$;

- ∘ $X \sim Pa(1,p) \Leftrightarrow$ $X \sim G_1(p)$
- ∘ $X \sim BN(r,p)$ e $\lambda = \frac{rp}{1-p} \Rightarrow X \sim Po(\lambda)$, quando $r \to \infty$

Uniforme Discreta $[UD(n),\ n \in \mathbb{N}]$

$\mathbb{P}(X=x) = \frac{1}{n} \underset{\{1,\cdots,n\}}{I}(x)$; $\mathbb{E}(X) = \frac{n+1}{2}$; $Var(X) = \frac{n^2-1}{12}$; $M_X(t) = \frac{e^t(1-e^{nt})}{n(1-e^t)}$; $A(X) = 0$; $K(X) = -\frac{6(n^2+1)}{5(n^2-1)}$;

- ∘ $X \sim Beta\ Binomial(n-1, 1, 1)$ $\Leftrightarrow X \sim UD(n)$

Poisson $[Po(\lambda),\ \lambda > 0]$

$\mathbb{P}(X=x) = \frac{\lambda^x e^{-\lambda}}{x!} \underset{\{0,1,\cdots\}}{I}(x)$; $\mathbb{E}(X) = Var(X) = \lambda$; $Mo(X) = \begin{matrix} \lfloor \lambda \rfloor \\ \text{ou;} \\ \lceil \lambda \rceil - 1 \end{matrix}$ $A(X) = \lambda^{-1}$;

$\mathbb{E}[(1-a)^X] = e^{-a\lambda}$; $K(X) = \lambda^{-1/2}$; $M_X(t) = e^{\lambda(e^t-1)}$; $\mathbb{E}\left[\frac{X!}{(X-k)!}\right] = \lambda^k$;

- ∘ $X_1, X_2, \cdots$ ind. com $X_i \sim Po(\frac{r^i}{i}) \Rightarrow Y = \sum_{i=1}^\infty iX_i \sim G_0(\frac{1-r}{r})$
- ∘ $X_1, \cdots, X_n$ i.i.d. com $X_i \sim Po(\lambda) \Rightarrow Y = \sum_{i=1}^n X_i \sim Po(n\lambda)$
- ∘ $X_1, \cdots, X_n$ ind. com $X_i \sim Po(\lambda_i) \Rightarrow$ $Y = X_i|(\sum_{j=1}^n X_j = k) \sim Bin(k, \lambda_i/\sum_{j=1}^n \lambda_j)$
- ∘ $X \sim Bin(n,p)$ e $\lambda = np \Rightarrow X \xrightarrow{\mathcal{D}} Po(\lambda)$, quando $n \to \infty$

Hipergeométrica $[Hipergeo(N,k,n),\ N \in \mathbb{N},\ k \in \{0,\cdots,N\},\ n \in \{0,\cdots,N\}]$

$\mathbb{P}(X=x) = \frac{\binom{k}{x}\binom{N-k}{n-x}}{\binom{N}{n}} \underset{\{\max\{0,n-(N-k)\},\cdots,\min\{k,n\}\}}{I}(x)$; $Var(X) = n\frac{k}{N}(1-\frac{k}{N})(\frac{N-n}{N-1})$;

$\mathbb{E}(X) = n\frac{k}{N}$; $Mo(X) = \lfloor \frac{(n+1)(k+1)}{N+2} \rfloor$; $A(X) = \frac{(N-2k)(N-2n)\sqrt{N-1}}{(N-2)\sqrt{nk(N-k)(N-n)}}$;

- ∘ $X \sim Hipergeo(N,k,1) \Leftrightarrow X \sim Bernoulli(\frac{k}{N})$
- ∘ $X \sim Hipergeo(N,k,n)$ e $p = \frac{k}{N} \Rightarrow$ $X \xrightarrow{\mathcal{D}} Bin(n,p)$, quando $n \to \infty$

Apêndice

- ∘ $Var(X) = \mathbb{E}[(X-\mathbb{E}(X))^2]$;
- ∘ $M_X(t) = \mathbb{E}(e^{tX})$; $C_X(t) = \mathbb{E}(e^{itX})$
- ∘ $x_p = F^{-1}(p) \Rightarrow Med(X) = x_{1/2}$;
- ∘ $A(X) = \frac{\mathbb{E}[(X-\mathbb{E}(X))^3]}{Var(X)^{3/2}}$;
  - $A(X) < 0$ (cauda pesada à esquerda);
  - $A(X) > 0$ (cauda pesada à direita);
  - $A(X) = 0$ (distribuição simétrica);
- ∘ $K(X) = \frac{\mathbb{E}[(X-\mathbb{E}(X))^4]}{Var(X)^2} - 3$;
  - $K(X) < 0$ (platicúrtica [achatada]);
  - $K(X) > 0$ (leptocúrtica [afuniliada]);
  - $K(X) = 0$ (mesocúrtica);

- ∘ $e^x = \sum_{n=0}^\infty \frac{x^n}{n!}$;
- ∘ $\ln(x+1) = \sum_{n=0}^\infty \frac{(-1)^n x^{n+1}}{n+1},\ |x| < 1$;
- ∘ $sen\ x = \sum_{n=0}^\infty \frac{(-1)^n x^{2n+1}}{(2n+1)!}$;
- ∘ $cos\ x = \sum_{n=0}^\infty \frac{(-1)^n x^{2n}}{(2n)!}$;
- ∘ $\sum_{i=1}^n i = \frac{n(n+1)}{2}$;
- ∘ $\sum_{i=1}^n i^2 = \frac{n(n+1)(2n+1)}{6}$;
- ∘ $\sum_{i=1}^n i^3 = \left(\frac{n(n+1)}{2}\right)^2$;
- ∘ $\sum_{i=1}^n i^4 = \frac{n(n+1)(2n+1)(3n^2+3n-1)}{30}$;
- ∘ $\left(\sum_{i=1}^m a_i\right)^n = \sum_{x_1+\cdots+x_m=n} \frac{n!}{\prod_{j=1}^m x_j} \prod_{j=1}^m a_j^{x_j}$;

- ∘ $\binom{n}{x} = \frac{n!}{x!(n-x)!}$;
- ∘ $\binom{n}{x} + \binom{n}{x-1} = \binom{n+1}{x},\ n \in \mathbb{N}$ e $x \in \mathbb{Z}$;
- ∘ $\binom{-n}{x} = (-1)^x\binom{n+x-1}{x}$;
- ∘ $n! \sim (2\pi)^{1/2}e^{-n}n^{n+1/2}$;
- ∘ $(a+b)^n = \sum_{i=0}^n \binom{n}{i}a^ib^{n-i}$;
- ∘ $\binom{N_1+N_2}{n} = \sum_{i=0}^n \binom{N_1}{i}\binom{N_2}{n-i}$;
- ∘ $(1-x)^{-2} = \sum_{i=0}^\infty (i+1)x^i$;
- ∘ $\max(a,b) = \frac{a+b+|a-b|}{2}$;
- ∘ $\min(a,b) = \frac{a+b-|a-b|}{2}$;

- ∘ $\Gamma(a+1) = a\Gamma(a)$;
- ∘ $B(a,b) = \frac{\Gamma(a)(b)}{\Gamma(a+b)}$;
- ∘ $\Gamma(a)\Gamma(1-a) = \frac{\pi}{sen\ (\pi a)}$;
- ∘ $\Gamma(\frac{a}{2}) = \frac{\sqrt{\pi}(a-1)!}{2^{a-1}(\frac{a-1}{2})!}$;
- ∘ $a_1(x-b_1)^2 + a_2(x-b_2)^2$ $= (a_1+a_2)(x-c)^2$ $+ \frac{a_1a_2}{a_1+a_2}(b_1-b_2)^2$, sendo $c = \frac{a_1b_1+a_2b_2}{a_1+a_2}$;
- ∘ $e^{ix} = cos\ x + isen\ x$;
- ∘ $e^a = \lim_{n\to\infty}\left(1+\frac{a}{n}\right)^n$;