Héliton Ribeiro Tavares
Maria Regina Madruga Tavares
José Gracildo de Carvalho Junior

INTRODUÇÃO À PROBABILIDADE com aspectos computacionais

Notas de Aula

UFPA

www.ufpa.br/heliton/arquivos/probabilidade

2012

O problema dos aniversários

Se você selecionar 23 pessoas ao acaso, a "probabilidade" de que pelo menos duas delas façam aniversário na mesma data já é superior a 50%; com 30 pessoas é cerca de 70%; com 40, 90%; com 50, 97% e com 60, 99%.

Ver também...

Problema dos encontros

Problema das 3 portas

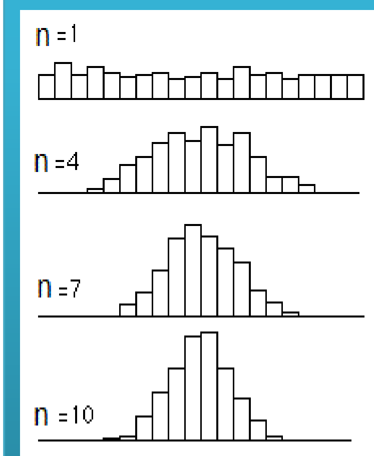
Tamanho da amostra

Em uma grande população, para estimar um determinado percentual P, o resultado obtido de uma amostra aleatória de cerca de 400 observações deve diferir de P por até 5% [*]; com 800, 3,5%; com 1100, 3%; com 1500, 2,5%; e com 2400, 2%.

[*] *Em apenas 5% das pesquisa a diferença deve ser superior à citada.*

Teorema Central do Limite

Se tomarmos amostras aleatórias de tamanho *n* de uma variável "qualquer" com média μ e desvio-padrão *σ*, então para valores grandes de *n* teremos que distribuição da média amostral se aproxima de uma distribuição normal com média μ e desvio padrão $\sigma/\sqrt{n}$. Ou seja, cada vez mais concentrada em torno de μ.



FACULDADE DE ESTATÍSTICA / UNIVERSIDADE FEDERAL DO PARÁ

Introdução à Probabilidade

---

com aspectos computacionais

Héliton Ribeiro Tavares

Maria Regina Madruga Tavares

José Gracildo de Carvalho Junior

Faculdade de Estatística,
Universidade Federal do Pará (UFPA)

Prefácio

A idéia de escrever um texto introdutório sobre a Teoria da Probabilidade com aspectos computacionais surgiu da necessidade de criar uma proximidade maior dos alunos de graduação com rotinas de programação, usando processos de simulação para entender vários conceitoss de Probabilidade e Estatística.

Embora a vontade seja de começar com linguagens mais avançadas, como em Ox, R, SPSS, SAS, por exemplo, nada mais natural do que usar algo que praticamente todos os computadores com sistema windows já têm instalado, o Excel, e sua linguagem subjacente, o Visual Basic for Applications (VBA). Assim trabalhamos o uso da Planilha e a Programação conjuntamente.

A maior preocupação foi a de escrever um texto que pudesse ser utilizado não só pelos estatísticos, mas também por interessados de outras áreas. Por conta disso o material é realmente introdutório, de forma que há muitos outros materiais disponíveis com maior formalismo matemático. Procuramos detalhar apenas os pontos que achamos mais interessantes para um texto introdutório e fornecer o maior número possível de referências bibliográficas que cobrissem os outros pontos.

De forma a suprir de material necessário para o desenvolvimento da teoria e exercícios, o primeiro capítulo apresenta diversos resultados de matemática básica. Sugestões e reclamações serão sempre bem-vindas de forma a melhorar continuamente estas notas de aula. Este material, bem como as macros desenvolvidas, estão disponibilizados em *www.helitontavares.com.* Sua atualização está sendo contínua, e ainda está em fase inicial de preparação, certamente com muitas imperfeições. E que fique registrada a colaboração de vários alunos de Graduação e Pós-Graduação.

Fevereiro 2012

Heliton Ribeiro Tavares
Maria Regina Madruga Tavares
José Gracildo de Carvalho Junior

Conteúdo

Lista de Figuras

Lista de Tabelas

Capítulo **1**

Matemática Básica

Neste capítulo são apresentadas algumas expressões que poderão ser usadas nos capítulos subsequentes. Talvez alguns ainda não sejam de domínio de estudantes em início de graduação, mas são de relativa facilidade, de forma que podem ser rapidamente absorvídos. O capítulo tem por objetivo ser de consulta rápida, podendo ser dispensado em uma primeira leitura, mas sugere-se um passar de olhos para um breve conhecimento do seu conteúdo.

1.1 Constantes importantes

Abaixo as principais constantes usadas a área de Estatística.

$$\pi = 3,14159265358979323846264338327950288419716939937510 58... \quad (1.1)$$

$$e = 2,718281828459045235360287... = \lim_{n\to\infty}\left(1+\frac{1}{n}\right)^n \quad (1.2)$$

$$\gamma = 0,577215664901532860606512... = \text{Constante de Euler} \quad (1.3)$$

$$= \lim_{n\to\infty}\left(1+\frac{1}{2}+\frac{1}{3}+...+\frac{1}{n}-\ln n\right) \quad (1.4)$$

$$\sqrt{\pi} = \Gamma\left(\frac{1}{2}\right) = 1,772453850905516027298167...(\text{onde } \Gamma \text{ é a função gamma}) \quad (1.5)$$

1.2 Fração e potências

(a) $a^{n+m} = a^n a^m$

(b) $(a^n)^m = a^{nm}$

(c) $a^{-n} = 1/a^n$

(d) $a^{1/n} = \sqrt[n]{a}$

1.3 Função Exponencial (exp)

(a) $x = e^{\ln(x)} = \ln(e^x)$

(b) $(e^x)^c = e^{cx}$, onde c é uma constante.

(c) $e^{a+b} = e^a \times e^b$, onde a e b são constantes quaisquer.

(d) $e^{a-b} = \frac{e^a}{e^b}$, onde a e b são constantes quaisquer.

1.4 Função Logaritmo (ln ou log)

a) $\log(x^c) = c\log x \neq (\log x)^c$

b) $\log\frac{a}{b} = \log a - \log b$

c) $\log(a \times b) = \log a + \log b \neq \log(a+b)$

1.5 Somatórios

Frequentemente precisamos representar séries e sequências em uma notação mais sintética. Por exemplo, considerando $x_i, i = 1, \cdots, n$, os termos de uma progressão aritmética ou geométrica, então a soma destes termos será representada pelo símbolo $\sum$, em que o primeiro e último termos serão colocados abaixo e acima do símbolo, resultando em:

$$x_1 + x_2 + \cdots + x_n = \sum_{i=1}^{n} x_i.$$

Costuma-se adotar as letras i, j, k e l nos somatórios, mas outras letras também podem ser encontradas com menor frequência. Ainda, quando temos uma soma dupla (em i e j, por exemplo) costuma-se *poupar* notação adotando-se apenas um símbolo de somatório, ou seja, $\sum\limits_i\sum\limits_j x_ix_j = \sum\limits_{i,j} x_ix_j$.

Obs: se $\{x_n\}_{n\geq 1}$ é uma sequência, então a soma infinita $x_1 + x_2 + \cdots + x_n + \cdots = \sum_{i=1}^{\infty} x_i$ é chamada de **Série**.

Propriedades: seja c uma constante qualquer

(a) $\sum\limits_{i=1}^{n} c = nc$. Caso especial, $n = \sum\limits_{i=1}^{n} 1$.

(b) $\sum\limits_{i=1}^{n} cx_i = c\sum\limits_{i=1}^{n} x_k$

(c) $\sum\limits_{i=1}^{n} (c + x_i) = nc + \sum\limits_{i=1}^{n} x_i$

(d) $\sum\limits_{i=1}^{n} (x_i + y_i) = \sum\limits_{i=1}^{n} x_i + \sum\limits_{i=1}^{n} y_i$

(e) $\frac{\sum\limits_{i=1}^{n} x_i}{\sum\limits_{i=1}^{n} y_i} = \frac{x_1+x_2+\cdots+x_n}{y_1+y_2+\cdots+y_n}$

(f) $\sum\limits_{i=1}^{n} \frac{x_i}{y_i} = \frac{x_1}{y_1} + \frac{x_2}{y_2} + \cdots + \frac{x_n}{y_n}$

(g) $\sum\limits_{i=1}^{n} x_iy_i = x_1y_1 + x_2y_2 + \cdots + x_ny_n$

(h) $(\sum_{i=1}^{n} x_i)\,(\sum_{i=1}^{n} y_i) = (x_1 + x_2 + \cdots + x_n) \times (y_1 + y_2 + \cdots + y_n) = \sum_{i=1}^{n}\sum_{j=1}^{n} x_i y_j = \sum_{i,j} x_i x_j.$

1.6 Produtórios

De forma similar ao somatório, o produto dos termos de uma sequência $x_1, x_2, \cdots, x_n$ será representada pelo símbolo $\prod$, resultando em

$$x_1 \times x_2 \times \cdots \times x_n = \prod_{i=1}^{n} x_i.$$

Propriedades:

(a) $\prod_{i=1}^{n} c = c^n$

(b) $\prod_{i=1}^{n} c\, x_i = c^n \prod_{i=1}^{n} x_i$

(c) $\prod_{i=1}^{n} x_i^c = (\prod_{i=1}^{n} x_i)^c$

(d) $\prod_{i=1}^{n} (x_i + y_i) = (x_1 + y_1) \times (x_2 + y_2) \times \cdots \times (x_n + y_n)$

(e) $\prod_{i=1}^{n} x_i \; + \; \prod_{i=1}^{n} y_i = (x_1 \times x_2 \times \cdots \times x_n) + (y_1 \times y_2 \times \cdots \times y_n)$

(f) $\frac{\prod_{i=1}^{n} x_i}{\prod_{i=1}^{n} y_i} = \frac{x_1 \times x_2 \times \cdots \times x_n}{y_1 \times y_2 \times \cdots \times y_n} = \prod_{i=1}^{n} \frac{x_i}{y_i}$

1.7 Miscelânea

(a) $\prod_{i=1}^{n} e^{x_i} = e^{x_1} \times e^{x_2} \times \cdots \times e^{x_k n} = e^{\sum_{i=1}^{n} x_i}$

(b) $\prod_{i=1}^{n} e^{c\, x_i} = e^{c \sum_{i=1}^{n} x_i}$

(c) $\prod_{i=1}^{n} c\, e^{x_i} = c^n e^{\sum_{i=1}^{n} x_i}$

(d) $\prod_{i=1}^{n} e^{c + x_i} = e^{nc + \sum_{i=1}^{n} x_i}$

(e) $\prod_{i=1}^{n} e^{x_i + y_i} = e^{\sum_{i=1}^{n} x_i + \sum_{i=1}^{n} y_i}$

(f) $\prod_{i=1}^{n} c\log(x_i) = c^n \prod_{i=1}^{n} \log(x_i) = \prod_{i=1}^{n} \log(x_i^c)$

(g) $\log(\prod_{i=1}^{n} x_i) = \sum_{i=1}^{n} \log(x_i)$

Exercício 1.7.1. *Escreva a soma* $\sum_{i=1}^{\infty}\sum_{j=i}^{\infty} a_{ij}$ *com os símbolos invertidos, primeiro o* j *e depois o* i *(note que se* $j \geq i$*, então na nova ordem teremos* $i \leq j$*).*

1.8 Contagem

Seja $n \in I\!N$. Define-se o *fatorial* de n por

$$n! = 1 \times 2 \times 3 \times \cdots \times n$$

I) Permutação: o número de maneiras de ordenar n objetos distintos é $n!$.

II) Arranjo: o número de maneiras de escolher k objetos distintos dentre n possíveis, importando-se a ordem, é $A(n,k) = \frac{n!}{k!}$.

III) Combinação de n, k a k: $C_{n,k} = \frac{n!}{k!(n-k)!}$, e representa-se por $\binom{n}{k}$.

IV) Fórmula Binomial para n Positivo Inteiro: Esta é a chamada Fórmula Binomial. Ela pode ser estendida para outros valores de n sendo, então, uma série infinita.

$(x+y)^n = x^n + nx^{n-1}y + \frac{n(n-1)}{2!}x^{n-2}y^2 + \frac{n(n-1)(n-2)}{3!}n^{n-3}y^3 + ... + y^n$

Que também pode ser escrita da seguinte forma:
$(x+y)^n = x^n + \binom{n}{1}x^{n-1}y + \binom{n}{2}x^{n-2}y^2 + \binom{n}{3}x^{n-3}y^3 + ... + \binom{n}{n}y^n$

onde os coeficientes, chamados coeficientes binomiais, são dados por
$\binom{n}{k} = \frac{n(n-1)(n-2)...(n-k+1)}{k!} = \frac{n!}{k!(n-k)!} = \binom{n}{n-k}$, $\quad n \geq k \geq 0$ *naturais*

Propriedades:

(a) $0! = 1$

(b) $\binom{n}{n} = \binom{n}{1} = \binom{n}{0} = 1$

(c) $C_{n,k} = 0$ se $n < k$

(d) $C_{n,k} = C_{n,n-k}$, ou seja, $\binom{n}{k} = \binom{n}{n-k}$

(e) $\binom{n}{k} + \binom{n}{k+1} = \binom{n+1}{k+1}$

(f) $\binom{n}{0}+\binom{n}{1}+\binom{n}{2}+...+\binom{n}{n}=2^n$

(g) $\binom{n}{n}+\binom{n+1}{n}+\binom{n+2}{n}+...+\binom{n+k}{n}=\binom{n+k+1}{n+1}$

(h) $\binom{n}{0}+\binom{n+1}{1}+\binom{n+2}{2}+...+\binom{n+k}{k}=\binom{n+k+1}{k}$

(i) $\binom{n}{0}^2+\binom{n}{1}^2+\binom{n}{2}^2+...+\binom{n}{n}^2=\binom{2n}{n}$

(j) $\binom{n}{0}-\binom{n}{1}+\binom{n}{2}+...+(-1)^n\binom{n}{n}=0$

(k) $1\cdot\binom{n}{1}+2\cdot\binom{n}{2}+3\cdot\binom{n}{3}+...+n\cdot\binom{n}{n}=n\cdot 2^{n-1}$

(l) $1\cdot\binom{n}{1}-2\cdot\binom{n}{2}+3\cdot\binom{n}{3}+...+(-1)^{n+1}\cdot n\binom{n}{n}=0$

(m) $\binom{n}{0}+\binom{n}{2}+\binom{n}{4}+...=\binom{n}{1}+\binom{n}{3}+\binom{n}{5}+...=2^{n-1}$

(n) $\binom{m}{0}\binom{n}{p}+\binom{m}{1}\binom{n}{p-1}+...+\binom{m}{p}\binom{n}{p-p}=\binom{m+n}{p}$

Fórmula Multinomial

(5) $(x_1+x_2+...+x_p)^n=\sum\frac{n!}{n_1!\,n_2!\,...\,n_p!}\cdot x_1^{n_1}\cdot x_2^{n_2}\cdot...\cdot x_p^{n_p}$
onde $n_1,n_2,...,n_p$ são inteiros não negativos tais que $n_1+n_2+...+n_p=n$ e a $\sum$ é sobre todos os n_i

(6) $(a+b+c)^2=a^2+b^2+c^2+2ab+2ac+2bc$

1.9 Triângulo de Pascal

$$
\begin{array}{c|ccccccc}
0 & \binom{0}{0} \\
1 & \binom{1}{0} & \binom{1}{1} \\
2 & \binom{2}{0} & \binom{2}{1} & \binom{2}{2} \\
3 & \binom{3}{0} & \binom{3}{1} & \binom{3}{2} & \binom{3}{3} \\
4 & \binom{4}{0} & \binom{4}{1} & \binom{4}{2} & \binom{4}{3} & \binom{4}{4} \\
5 & \binom{5}{0} & \binom{5}{1} & \binom{5}{2} & \binom{5}{3} & \binom{5}{4} & \binom{5}{5} \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots
\end{array}
\quad\Longleftrightarrow\quad
\begin{array}{c|cccccccc}
0 & 1 \\
1 & 1 & 1 \\
2 & 1 & 2 & & 1 \\
3 & 1 & 3 & & 3 & 1 \\
4 & 1 & \boxed{4} & + & \boxed{6} & 4 & 1 \\
5 & 1 & 5 & = & \boxed{10} & 10 & 5 & 1 \\
\vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & \vdots & \ddots
\end{array}
\qquad
\binom{n}{k}+\binom{n}{k+1}=\binom{n+1}{k+1}
$$

1.10 Funções Polinomiais

As expressões abaixo, do tipo $(x+y)^n$, podem ser genericamente escritas como

$$(x+y)^n=\sum_{k=0}^{n}\binom{n}{k}x^k y^{n-k}. \tag{1.6}$$

Alguns casos particulares são:

i) $(x+y)^2=x^2+2xy+y^2$

ii) $(x-y)^2 = x^2 - 2xy + y^2$

iii) $(x+y)^3 = x^3 + 3x^2y + 3xy^2 + y^3$

iv) $(x-y)^3 = x^3 - 3x^2y + 3xy^2 - y^3$

v) $(x+y)^4 = x^4 + 4x^3y + 6x^2y^2 + 4xy^3 + y^4$

vi) $(x-y)^4 = x^4 - 4x^3y + 6x^2y^2 - 4xy^3 + y^4$

Alguns casos especiais são:

i) $x^2 - y^2 = (x-y)(x+y)$

ii) $x^3 - y^3 = (x-y)(x^2+xy+y^2)$

iii) $x^3 + y^3 = (x+y)(x^2-xy+y^2)$

iv) $x^4 - y^4 = (x-y)(x+y)(x^2+y^2)$

v) $x^5 - y^5 = (x-y)(x^4+x^3y+x^2y^2+xy^3+y^4)$

Generalizações destas expressões são dadas por:

i) $x^{2n} - y^{2n} = (x-y)(x+y)(x^n + x^{n-1} + x^{n-2}y + x^{n-3}y^2 + ...)(x^{n-1}y + x^{n-3}y^2 - ...)$

$= (x-y)(x+y)\left(x^2 - 2xy\ cos\frac{\pi}{n} + y^2\right)\left(x^2 - 2xy\ cos\frac{2\pi}{n} + y^2\right) ... \left(x^2 - 2xy\ cos\frac{(n-1)\pi}{n} + y^2\right)$

ii) $x^{2n} + y^{2n} = \left(x^2 + 2xy\ cos\frac{\pi}{2n} + y^2\right)\left(x^2 + 2xy\ cos\frac{3\pi}{2n} + y^2\right) ... \left(x^2 + 23xy\ cos\frac{(2n-1)\pi}{2n} + y^2\right)$

iii) $x^{2n+1} - y^{2n+1} = (x-y)(x^{2n} + x^{2n-1}y + x^{2n-2}y^2 + ... + y2n) = (x-y)\left(x^2 - 2xy\ cos\frac{2\pi}{2n+1} + y^2\right)$
$\left(x^2 - 2xy\ cos\frac{4\pi}{2n+1} + y^2\right) ... \left(x^2 - 2xy\ cos\frac{2n\pi}{2n+1} + y^2\right)$

iv) $x^{2n+1} + y^{2n+1} = (x+y)(x^{2n} - x^{2n-1}y + x^{2n-2}y^2 - ... + y2n) = (x+y)\left(x^2 + 2xy\ cos\frac{2\pi}{2n+1} + y^2\right)$
$\left(x^2 + 2xy\ cos\frac{4\pi}{2n+1} + y^2\right) ... \left(x^2 + 2xy\ cos\frac{2n\pi}{2n+1} + y^2\right)$

Multinomial

Uma generalização da expressão (1.6) é dada por

$$(x_1 + x_2 + \cdots + x_p)^n = \sum_{n_1, n_2, \cdots, n_p} \frac{n!}{n_1! n_2! \cdots n_p!} x_1^{n_1} x_2^{n_2} \cdots x_p^{n_p},$$

onde n_1, $n_2, \cdots, n_p$ são inteiros não negativos tais que $\sum_{i=1}^p n_i = n$. Para $n = 2$, temos

$$\left(\sum_{i=1}^p x_i\right)^2 = \sum_{i,j} x_i x_j = \sum_{i=1}^p x_i^2 + 2\sum_{i=1}^n \sum_{i<j} x_i x_j. \tag{1.7}$$

1.11 Técnicas de Demonstração

Um teorema é uma proposição do tipo: "Se $\mathcal{H}$, então $\mathcal{T}$", a qual prova-se ser verdadeira sempre, ou seja, que: "$\mathcal{H}$ implica $\mathcal{T}$". As proposições '$\mathcal{H}$' e '$\mathcal{T}$' são denominadas *Hipótese* e *Tese*, respectivamente.

É usual denominar *corolário* a um teorema que é uma consequência quase direta de um outro já demonstrado (ou seja, cuja prova é trivial ou imediata). Adicionalmente, um teorema auxiliar que possui um resultado importante para a prova de um outro é usualmente denominado *lema*. Teoremas são fundamentais em Matemática, Estatística, dentre outras. Por exemplo, um teorema permite verificar se uma determinada afirmação é correta. Sob este ponto de vista, um teorema pode ser visto como um algorítmo que, prova-se, sempre funciona. Antes de iniciar uma demonstração, é fundamental identificar claramente a hipótese $\mathcal{H}$ e a tese $\mathcal{T}$. Por exemplo, considere o seguinte teorema: 0 é o único elemento neutro da adição no conjunto dos números naturais. Uma reescrita, identificando claramente a hipótese e a tese, é como segue: se 0 é elemento neutro da adição no conjunto dos números naturais, então 0 é o único elemento neutro da adição no conjunto dos números naturais. É importante observar que, na demonstração de que, de fato, $\mathcal{H}$ implica $\mathcal{T}$, a hipótese é suposta verdadeira. Consequêntemente, a hipótese não deve ser demonstrada. De fato, todas as teorias possuem um conjunto de premissas (hipóteses) que são supostas verdadeiras e sobre as quais todo o raciocínio é construído. A teoria dos conjuntos é baseada em uma premissa que é a noção de elemento, a qual é simplesmente suposta. A própria Estatística é construída sobre várias premissas, a partir das quais demonstra-se todos os resultados adotados na prática.

1.11.1 Demonstração Exaustiva

A conjectura pode ser provada verificando-se que ela é verdadeira para todos os elementos da coleção. Para provar a falsidade da conjectura, basta achar um contra-exemplo.

Exemplo 3: Prove a conjectura: Para todo inteiro positivo n, $n! \leq n^2$. Solução: A conjectura é falsa pois não é verdade para todo n: é falsa para $n = 4$.

1.11.2 Demontração Direta

A demontração direta de uma sentença $\mathcal{H} \Longrightarrow \mathcal{T}$ funciona da seguinte forma: assuma que o antecedente $\mathcal{H}$ é verdade e deduza a conclusão (ou consequente) $\mathcal{T}$.

1.11.3 Demonstração por Contraposição

A sentença condicional $\mathcal{H} \Longrightarrow \mathcal{T}$ pode ser provada mostrando-se que a sua contrapositiva é verdadeira: $\overline{\mathcal{T}} \Longrightarrow \overline{\mathcal{H}}$. Em suma, se $\mathcal{H}$ implica $\mathcal{T}$, então 'Não $\mathcal{T}$' implica 'Não $\mathcal{H}$'.

1.11.4 Demonstração por Absurdo

Para demonstrar $\mathcal{T}$, assumimos $\overline{\mathcal{T}}$ e mostramos que isso leva a uma contradição. Como $H \cap !F$ é verdadeira, concluimos que $\overline{\mathcal{H}}$ é falsa e portanto que p é verdadeira. De outra forma, para provar $\mathcal{H} \Longrightarrow \mathcal{T}$, basta mostrar ..., pois $(p\dot{} q!F)!(p!q)$ é uma tautologia.

1.11.5 Demonstração por Indução (Finita)

Também conhecida por Indução Matemática ou princípio da Indução, essa é uma técnica largamente aplicada. Por exemplo, algumas vezes nos defrontamos com afirmações envolvendo os números naturais e a questão que surge é: será tal afirmação verdadeira sempre, ou seja, vale para qualquer número natural $n = 1, 2, 3, \cdots$? Exemplos de afirmações são validades de fórmulas, tais como:

i) A soma dos n primeiros números naturais é $n(n+1)/2$, ou seja:

$$1 + 2 + 3 + \cdots + n = \frac{n(n+1)}{2}$$

ii) A soma dos n primeiros números ímpares é n^2:

$$1 + 3 + 5 + \cdots + (2n - 1) = n^2$$

iii) O trinômio $P(n) = n^2 + n + 41$, produz como resultado um número primo.

É claro que para responder às questões colocadas pelos problemas não basta "testar" a veracidade das "fórmulas" substituindo valores específicos para n. Por mais que as igualdades ganhem credibilidade, nunca poderemos garantir sua validade para algum valor de n que não tenha sido testado!

Quando uma proposição é enunciada em termos de números naturais, o Princípio de indução finita constitui um eficiente instrumento para demonstrar a proposição no caso geral. Na prática, o método pode ser entendido por um artifício muito simples. Vamos supor que temos uma série de peças do dominó colocados em fila, que começa por um deles e prossegue indefinidamente. Nosso objetivo é: empurrando apenas uma peça do dominó, garantir que todas caiam. Como derrubar todas as peças?

Para isso, basta nos assegurarmos de que: 1) A primeira peça cai; 2) As peças estão dispostas de tal modo que qualquer uma delas, toda vez que cai, automaticamente, empurra a peça seguinte e a faz cair também. Assim, mesmo que a fila se estenda indefinidamente, podemos afirmar que todas as peças cairão.

O sucesso desse método depende da demonstração de duas etapas:

Etapa 1) A proposição é válida para o primeiro termo ($n = 1$).

Etapa 2) Se a proposição é válida para o n-ésimo termo ($n = k$), então ela também é válida para o caso seguinte ($n = k + 1$).

Se ambas as partes forem verdadeiras, podemos afirmar que a proposição é válida para todo inteiro positivo n.

Exemplo 1.1. *Provar que se* $S_n = 1 + 3 + 5 + ... + (2n - 1)$*, então,* $S_n = n^2$.

Etapa 1: A fórmula vale para $n = 1$. De fato, $S_1 = 1^2 = 1$.
Etapa 2: Como $S_{n+1} = a_1 + a_2 + \cdots + a_{n+1} = S_n + a_{n+1}$, temos que

$$S_{n+1} = n^2 + [2(n+1) - 1] = n^2 + 2n + 1 = (n+1)^2.$$

Exercício 1.11.1. *Verificar a validades das proposições:*

i) $S_n = \sum_{k=1}^{n} k = \frac{n(n+1)}{2}$

ii) $S_n = \sum_{k=1}^{n} k^2 = \frac{n(n+1)(2n+1)}{6}$

iii) $S_n = \sum_{k=1}^{n} k^3 = \left[\frac{n(n+1)}{2}\right]^2$

iv) $S_n = \sum_{k=1}^{n} k^4 = \frac{n(n+1)(2n+1)(n^2-...)}{30}$

v) $a_n = \frac{4^{n+1}-1}{3}$ *é inteiro e ímpar.*

vi) $\ln a^n = n \ln a$ *para todo* $a > 0$ *e* $n \in I\!N$

vii) $(a+b)^n = \sum_{k=1}^{n} \binom{n}{k} a^k b^{n-k}$

viii) $S_n = \sum_{k=0}^{n} q^n = \frac{1-q^{n+1}}{1-q}$

ix) $(1+r)^n \geq 1 + rn, \forall r > 0.$

Exercício 1.11.2. *Encontrar a expressão* S_n *para as seguintes somas:*

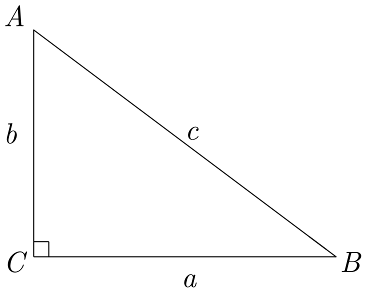
i) $2 + 4 + 6 + \cdots + 2n$

ii) $3 + 8 + 13 + \cdots + (5n-2)$

1.12 Trigonometria

Definição de Funções Trigonométricas para um triângulo reto

$$\begin{aligned}
\sin A &= \frac{a}{c} = \frac{\text{lado oposto}}{\text{hipotenusa}} && (1.8)\\
\cos A &= \frac{b}{c} = \frac{\text{lado adjacente}}{hipotenusa} && (1.9)\\
\tan A &= \frac{a}{c} = \frac{\text{lado oposto}}{lado\ adjacente} && (1.10)\\
\text{cotg}\, A &= \frac{b}{a} = \frac{\text{lado adjcente}}{\text{lado oposto}} && (1.11)\\
\sec A &= \frac{c}{b} = \frac{\text{hipotenusa}}{\text{lado adjacente}} && (1.12)\\
\text{cosec} A &= \frac{c}{a} = \frac{\text{hipotenusa}}{\text{lado oposto}} && (1.13)
\end{aligned}$$



Relações entre as Funções Trigonométricas

$i)$ $\tan A = \frac{\sin A}{\cos A}$

$ii)$ $\operatorname{cotg} A = \frac{1}{\tan A} = \frac{\cos A}{\sin A}$

$iii)$ $\sec A = \frac{1}{\cos A}$

$iv)$ $\operatorname{cosec} A = \frac{1}{\sin A}$

$v)$ $\sin^2 A + \cos^2 A = 1$

$vi)$ $\sec^2 A - \tan^2 A = 1$

$vii)$ $\operatorname{cosec}^2 A - \operatorname{cotg}^2 A = 1$

Funções de Ângulos Negativos

$i)$ $\sec(-A) = -\sin(A)$

$ii)$ $\cos(-A) = \cos(A)$

$iii)$ $\tan(-A) = -\tan(A)$

$iv)$ $\operatorname{cosec}(-A) = -\operatorname{cosec}(A)$

$v)$ $\sec(-A) = \sec(A)$

$vi)$ $\operatorname{cotg}(-A) = -\operatorname{cotg}(A)$

Fórmulas de Adição

$i)$ $\sin(A \pm B) = \sin A \cos B \pm \cos A \sin B$

$ii)$ $\cos(A \pm B) = \cos A \cos B \pm \sin A \sin B$

$iii)$ $\tan(A \pm B) = \frac{\tan A \pm \tan B}{1 \pm \tan A \tan B}$

Fórmulas de Ângulos Duplos

$i)$ $\sin\ 2A = 2 \cdot \sin A \cdot \cos A$

$ii)$ $\cos\ 2A = \cos^2 A - sen^2 A = 1 - 2\ \sin^2 A = 2 \cdot \cos^2 A - 1$

$iii)$ $\tan\ 2A = \frac{2 \tan A}{1 - \tan^2 A}$

Soma, Diferença e Produto de Funções Trigonométricas

$i)$ $\sin\ A + \sin B = 2\ \sin\ \frac{1}{2}(A+B) \cos \frac{1}{2}(A-B)$

$ii)$ $\sin\ A - \sin B = 2 \cos \frac{1}{2}(A+B)\ \sin\ \frac{1}{2}(A-B)$

$iii)$ $\cos A + \cos B = 2 \cos \frac{1}{2}(A+B) \cos \frac{1}{2}(A-B)$

$iv)$ $\cos A - \cos B = 2\ \sin\ \frac{1}{2}(A+B)\ \sin\ \frac{1}{2}(B-A)$

$v)$ $\sin\ A\ \sin B = \frac{1}{2}[\cos(A-B) - \cos(A+B)]$

$vi)$ $\cos A \cos B = \frac{1}{2}[\cos(A-B) + \cos(A+B)]$

$vii)$ $\sin\ A \cos B = \frac{1}{2}[\sin(A-B) + \sin\ (A+B)]$

1.13 Limites

1.14 Derivadas

Regras Gerais de Diferenciação

Nas expressões seguintes, u, v, w são funções de x; a, c, n são constantes.

$i)$ $\frac{d}{dx}(c) = 0$

$ii)$ $\frac{d}{dx}(cx) = c$

$iii)$ $\frac{d}{dx}(cx^n) = ncx^{n-1}$

$iv)$ $\frac{d}{dx}(u \pm v \pm w \pm ...) = \frac{du}{dx} \pm \frac{dv}{dx} \pm \frac{dw}{dx} \pm ...$

$v)$ $\frac{d}{dx}(uv) = u\frac{dv}{dx} + v\frac{du}{dx}$

$vi)$ $\frac{d}{dx}(\frac{u}{v}) = \frac{v(du/dx) - u(dv/dx)}{v^2}$

$vii)$ $\frac{d}{dx}(u^n) = nu^{n-1}\frac{du}{dx}$

viii) $\frac{d}{dx}(sen\, u) = cos\, u \frac{du}{dx}$

ix) $\frac{d}{dx}(cos\, u) = -sen\, u \frac{du}{dx}$

x) $\frac{d}{dx}(tg\, u) = sec^2\, u \frac{du}{dx}$

xi) $\frac{d}{dx} \ln u = \frac{1}{u} \frac{du}{dx}$

xii) $\frac{d}{dx}\, a^u = a^u \ln a \frac{du}{dx}$

xiii) $\frac{d}{dx}\, e^u = e^u \frac{du}{dx}$

Derivadas Superiores

A segunda, terceira e as derivadas superiores são definidas da seguinte forma:

i) Segunda Derivada $= \frac{d}{dx}\frac{dy}{dx} = \frac{d^2y}{dx^2} = f''(x) = y''$

ii) Terceira Derivada $= \frac{d}{dx}\frac{d^2y}{dx^2} = \frac{d^3y}{dx^3} = f'''(x) = y'''$

iii) Enésima Derivada $= \frac{d}{dx}\frac{d^{n-1}y}{dx^{n-1}} = \frac{d^ny}{dx^n} = f^{(n)}(x) = y^{(n)}$

Exemplo: Dado $y = 2x^3$ temos:

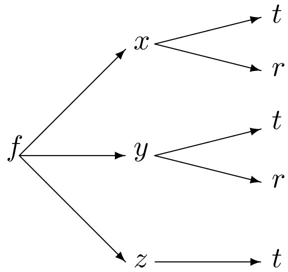
$\frac{dy}{dx} = y' = 6x^2$
$\frac{d^2y}{dx^2} = y'' = 12x$
$\frac{d^3y}{dx^3} = y''' = 12$
$\frac{d^4y}{dx^4} = y'''' = 0$

Derivadas de Funções Compostas (Várias Variáveis)

Seja $f(x, y, z)$ uma função tal que:

$x = f_1(r, t)$
$y = f_2(r, t)$
$z = f_3(t)$

De modo, que via indireta, f é uma função de r,t.



1.15 Integrais

Regras Gerais de Integração

Nas expressões seguintes, u, v, w são funções de x; a, n são constantes.

1) $\int a\, dx = ax$

2) $\int af(x)\, dx = a \int f(x)$

3) $\int (u \pm v \pm w \pm ...)dx = \int u\, dx \pm \int v\, dx \pm \int w\, dx \pm ...$

4) $\int u\, dv = uv - \int v\, du$ [Integração por partes]

5) $\int f(x)dx = \frac{1}{a} \int f(u)du$

6) $\int Ff(x)dx = \int F(u)\frac{dx}{du}\, du = \int \frac{F(u)}{f(x)}du$ onde $u = f(x)$

7) $\int u^n du = \frac{u^{n+1}}{n+1}, \;\; n \neq -1$ [Para n = -1, ver proximo item]

8) $\int \frac{du}{u} = \ln u$ [para u>0 ou u<0]

9) $\int e^u du = e^u$

10) $\int a^u du = \int e^{u \;\; \ln a} du = \frac{e^{u \; \ln a}}{\ln a} = \frac{a^u}{\ln a},$ $a > 0, \quad a \neq 1$

11) $\int \sin\; u\, du = -\cos\; u$

12) $\int \cos\; u\, du = \sin\; u$

13) $\int \frac{du}{u^2+a^2} = \frac{1}{a}\; \arctan \frac{u}{a}$

14) $\int \frac{du}{u^2-a^2} = \frac{1}{2a}\; \ln\left(\frac{u-a}{u+a}\right) = -\frac{1}{a}\; arc\; cotgh\; \frac{u}{a}$ $u^2 > a^2$

15) $\int \frac{du}{a^2-u^2} = \frac{1}{2a}\; \ln\left(\frac{a+u}{a-u}\right) = \frac{1}{a}\; arc\; tgh\; \frac{u}{a}$ $u^2 < a^2$

16) $\int \frac{du}{\sqrt{a^2-u^2}} = arc\; \sin\; \frac{u}{a}$

17) $\int \frac{du}{\sqrt{u^2+a^2}} = \ln\,(u + \sqrt{u^2+a^2})$

18) $\int \frac{du}{\sqrt{u^2-a^2}} = \ln\,(u + \sqrt{u^2-a^2})$

19) $\int \frac{du}{u\, \sqrt{u^2-a^2}} = \frac{1}{a}\; arc\; sec\, |\frac{u}{a}|$

20) $\int F(ax+b)\; dx = \frac{1}{a} \int F(u)\; du$ onde $u = ax + b$

21) $\int F(\sqrt{ax+b})\; dx = \frac{2}{a} \int u\; F(u)\; du$ onde $u = \sqrt{ax+b}$

22) $\int F(\sqrt[n]{ax+b})\; dx = \frac{n}{a} \int u^{n-1}\; F(u)\; du$ onde $u = \sqrt[n]{ax+b}$

23) $\int F(\sqrt{a^2 - x^2})\ dx = a \int\ F(a\ \cos u)\ cos u\ du$ onde $x = a\ \sin u$

24) $\int F(\sqrt{x^2 + a^2})\ dx = a\ \int F(a\ \sec u)\ \sec^2 u\ du$ onde $x = a \tan u$

25) $\int F(\sqrt{x^2 - a^2})\ dx = a\ \int F(a\ \tan u)\ \sec u\ \tan u\ du$ onde $x = a\ \sec u$

26) $\int F(e^{ax})\ dx = \frac{1}{a} \int \frac{F(u)}{u}\ du$ onde $u = e^{ax}$

27) $\int F(\ln x)\ dx = \int F(u)\ e^u du$ onde $u = \ln x$

28) $\int e^{ax}\ dx = \frac{e^{ax}}{a}$

29) $\int xe^{ax}\ dx = \frac{e^{ax}}{a} \left(x - \frac{1}{a}\right)$

30) $\int x^2 e^{ax}\ dx = \frac{e^{ax}}{a} \left(x^2 - \frac{2x}{a} + \frac{2}{a^2}\right)$

31) $\int x^n e^{ax}\ dx = \frac{x^n e^{ax}}{a} - \frac{n}{a} \int x^{n-1} e^{ax}\ dx$
$= \frac{e^{ax}}{a} \left(x^n - \frac{nx^{n-1}}{a} + \frac{n(n-1)x^{n-2}}{a} - ... \frac{(-1)^n n!}{a^n}\right)$ Se $n =$ número positivo inteiro.

32) $\int \frac{e^{ax}}{a}\ dx = \ln x + \frac{ax}{1 \cdot 1!} + \frac{(ax)^2}{2 \cdot 2!} + \frac{(ax)^3}{3 \cdot 3!} + ...$

33) $\int \frac{e^{ax}}{x^n}\ dx = \frac{-e^{ax}}{(n-1)x^{n-1}} + \frac{a}{n-1} \int \frac{e^{ax}}{x^{n-1}}\ dx$

Integrais Definidas

Integral definida de $f(x)$ em [a, b] é:

$$\int_a^b f(x)\ dx = \lim_{n\to\infty} S = \lim_{n\to\infty} \left(\sum_{i=0}^{n-1} f(a + i \cdot \Delta x) \cdot \Delta x\right) \tag{1.14}$$

e esse limite existirá se f(x) for contínua ponto a ponto.

Teorema Fundamental do Cálculo Integral:

O cálculo da integral anterior pode ser feito pela regra

$$\int_a^b f(x)\ dx = g(x)|_a^b = g(b) - g(a) \tag{1.15}$$

quando a condição $g'(x) = f(x)$ for satisfeita;

Propriedades das Integrais Definidas

1) $\int_a^b f(x) \pm g(x) \pm h(x) \pm ...dx = \int_a^b f(x)\, dx \pm \int_a^b g(x)\, dx \pm \int_a^b h(x)\, dx \pm ...$

2) $\int_a^b cf(x)\, dx = c\int_a^b f(x)\, dx,$ onde c é qualquer constante

3) $\int_a^a f(x)\, dx = 0$

4) $\int_a^b f(x)\, dx = -\int_b^a f(x)dx$

5) $\int_a^b f(x)\, dx = \int_a^c f(x)\, dx + \int_c^b f(x)\, dx$

1.16 Função Gamma (Γ)

Definição da Função Gama $\Gamma(n)$ para $n > 0$

$$\Gamma(n) = \int_0^\infty t^{n-1}e^{-t}\, dt \qquad n > 0 \tag{1.16}$$

Fórmula de Recorrência

1) $\Gamma(n+1) = n\Gamma(n)$

2) $\Gamma(n+1) = n!$ $\quad$ *se* $n = 0, 1, 2, ...$ onde $0! = 1$

1.17 Função Beta

Definição da Função Beta $\mathbf{B}(m,n)$

$$B(m,n) = \int_0^1 t^{m-1}(1-t)^{n-1}\, dt \qquad m > 0, n > 0 \tag{1.17}$$

Relação Entre a Função Beta e a Função Gamma

$$B(m,n) = \frac{\Gamma(m)\Gamma(n)}{\Gamma(m+n)}$$

Obs: $B(m,n) = B(n,m)$

1.18 PA, PG e Séries

Definição: Sejam $n =$ número natural e $a_n = f(n)$ uma função que associa o número a_n a cada n. Assim, $a_n = f(n)$ é definido como uma seqüência a_n de constantes.

A soma dos termos de uma P.A. é uma "Série Aritmética"

$$a_1 + a_2 + a_3 + ... + a_n = \frac{(a_1 + a_n) \cdot n}{2} \tag{1.18}$$

A série é harmônica (S.H.) quando for constituída por termos recíprocos aos termos de uma P.A.

$$a_1, a_2, a_3, ..., a_n \text{ é P.A.} \rightarrow \frac{1}{a_1} + \frac{1}{a_2} + \frac{1}{a_3} + \frac{1}{a_4} + \frac{1}{a_n} \text{ é S.H.} \tag{1.19}$$

Casos úteis (P.A. e P.G.)

i) $1 + 2 + 3 + ... + n = \frac{n(n+1)}{2}$

ii) $1 + 3 + 5 + ... + (2n - 1) = n^2$

iii) $2 + 4 + 6 + 8 + ... + (2n) = n(n + 1)$

iv) $1 \cdot 2 \cdot 3 \cdot 4 \cdot ... \cdot n = n!$

v) $1 \cdot 3 \cdot 5 \cdot 7 \cdot ... \cdot (2n - 1) = \frac{(2n)!}{2^n \cdot n!}$

vi) $2 \cdot 4 \cdot 6 \cdot 8 \cdot ... \cdot (2n) = 2^n \cdot n!$

vii) $a + ar + ar^2 + ... = \frac{a}{1-r} \quad se \quad |r| < 1.$

Soma de Potências de Números Inteiros Positivos

i) $1^2 + 2^2 + 3^2 + ... + n^2 = \frac{n(n+1)(2n+1)}{6}$

ii) $1^3 + 2^3 + 3^3 + ... + n^3 = \frac{n^2(n+1)^2}{4}$

iii) $1^4 + 2^4 + 3^4 + ... + n^4 = \frac{n(n+1)(2n+1)(3n^2+3n-1)}{30}$

Séries Envolvendo Recíprocas de potências de números positivos inteiros

i) $1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - ... = \ln 2$

ii) $1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \frac{1}{9} - ... = \frac{\pi}{4}$

1.19 Séries de Taylor

Série de Taylor para funções de uma variável

$f(x) = f(a) + f'(a)(x-a) + \frac{f'(a)(x-a)^2}{2!} + ... + \frac{f^{n-1}(a)(x-a)^{n-1}}{(n-1)!} + R_n$

Séries Binomiais

$i)$ $(a+x)^n = a^n + na^{n-1}x + \frac{n(n-1)}{2!}a^{n-2}x^2 + \frac{n(n-1)(n-2)}{3!}a^{n-3}x^3 + ...$ $\quad (a > x)$
$= a^n + \binom{n}{1}a^{n-1}x + \binom{n}{2}a^{n-2}x^2 + \binom{n}{3}a^{n-3}x^3 + ...$ $\quad$ (quando $n \in N$)

$ii)$ $(a+x)^2 = a^2 + 2ax + x^2$

$iii)$ $(a+x)^3 = a^3 + 3a^2x + 3ax^2 + x^3$

$iv)$ $(a+x)^4 = a^4 + 4a^3x + 6a^2x^2 + 4ax^3 + x^4$

$v)$ $(1+x)^{-1} = 1 - x + x^2 - x^3 + x^4 - ...$ $\quad -1 < x < 1$

$vi)$ $(1+x)^{-2} = 1 - 2x + 3x^2 - 4x^3 + 5x^4 - ...$ $\quad -1 < x < 1$

Séries para Funções Exponenciais e Logarítmicas

$i)$ $e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + ...$ $\quad -\infty < x < \infty$

$ii)$ $a^x = e^{x\ \ln a} = 1 + x\ \ln a + \frac{(x\ \ln a)^2}{2!} + \frac{(x\ \ln a)^3}{3!} + ...$ $\quad -\infty < x < \infty$

Séries para Funções Trigonométricas

$i)$ $\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + ...$ $\quad -\infty < x < \infty$

$ii)$ $\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + ...$ $\quad -\infty < x < \infty$

Convergência

$$S = \sum_{n=0}^{\infty} a_n \text{ é convergente quando sua soma tende a um número finito e bem definido.} \tag{1.20}$$

$$\text{Nota: Quando} \sum_{n=0}^{\infty} |a_n| \text{ é convergente, a série} \sum_{n=0}^{\infty} a_n \text{ é absolutamente convergente} \tag{1.21}$$

1.20 Alfabeto Grego

Nome Grego	Símbolos Gregos		Nome Grego	Símbolos Gregos	
	Minúsculas	Maiúsculas		Minúsculas	Maiúsculas
Alfa	α	A	Nu	ν	N
Beta	β	B	Csi	ξ	Ξ
Gama	γ	Γ	Ômicron	o	O
Delta	δ	Δ	Pi	π	Π
Épsilon	ϵ	E	Ro	ρ	P
Zeta	ζ	Z	Sigma	σ	Σ
Eta	η	H	Tau	τ	T
Teta	θ	Θ	Upsilon	υ	Υ
Iota	ι	I	Fi	ϕ	Φ
Capa	κ	K	Chi	χ	Ξ
Lambda	λ	Λ	Psi	ψ	Ψ
Mu	μ	M	Ômega	ω	Ω

Capítulo **2**

Informática Básica

Neste capítulo será apresentado um breve resumo de programação, com ênfase no *Visual Basil for Applications* através do Excel, disponível em *www.helitontavares.com*, desenvolvendo macros básicas para geração de números aleatórios, solução de problemas de probabilidades e outras necessidades inerentes à área. Um breve tópico sobre o a linguagem e ambiente `R` fechará o capítulo com seus principais comandos.

2.1 Uso de funções na planilha do Mocrosoft Excel

Nas células da planilha podem ser usadas diversas funções, digitando-se diretamente a função em qualquer célula com o símbolo de igualdade na frente. A seguir apresetamos algumas funções básicas.

Função	Descrição
=2*3	Produto de dois números
=2^3	Esponenciação 2^3
=ln(x)	Logaritmo de x na base neperiana
=exp(x)	e^x
=$aleatorio()$	Um número no intervalo [0,1)
=$aleatorio() * b$	Um número no intervalo [0,b)
=$(b - a) * aleatorio() + a$	Um número real no intervalo [a,b)

As planilhas trabalham com a Referência (linha, coluna), de tipo A1, indicando que é a célula da intesecção da coluna A e linha 1. Frequentemente usamos o símbolo $ para fixar a linha ou coluna, ou ambas, tal como A1. A tecla $F4$ também tem esse papel, pressionando várias vezes. Abaixo algumas funções que usam referências.

Função	Descrição
=soma(A1:A10)	Soma de uma sequência de valores no intervalo
=media(A1:A10)	Média dos valores no intervalo
=desvpad(A1:A10)	Desvio-padrão dos valores no intervalo
=somaquad(A1:A10)	Soma dos quadrados
=somarproduto(A1:A10;B1:B10)	
=med(A1:A10)	
=matriz.determ(A1:C3)	

Algumas funções fazem uso de uma estrutura especial chamada *MATRIZ*. O determinante de uma matriz quadrada é um escalar, mas a inversa, assim como o produto de duas matrizes, também é matriz. Para exemplificar, digite os números 1, 2, 3 e 4 nas células A1, B1, A2 e B2, respectivamente. Digite =MATRIZ.INVERSO(A1:B2) na célula D1, esta conterá a matriz inversa, mas só veremos seu primeiro elemento (-2). Para vermos os demais, marque as células D1:E2 (dimensão da matriz inversa), depois pressione a tecla F2, e depois CTRL+SHIFT+ENTER. O mesmo procedimento vale para quando fazemos o produto de duas matrizes, através do comando =MATRIZ.MULT(A1:B2;D1:E2), por exemplo. Preencha os intervalos citados e faça o teste.

2.2 Suplementos no Excel

Uma parcela do Excel não é instala inicialmente, mas apenas quando o usuário solicita. Estes são chamados de *suplementos* e servem, por exemplo, para montar um Histograma, para fazer uma Análise de Regessão, achar máximos e mínimos, ou raízes, de funções. A forma de instalação depende da versão do Excel, procure *Opções do Excel* e depois *Suplementos*, e será apresentado um quadro de suplementos. Clique em *Ir*, e será apresentada uma caixa para você marcar o suplemento desejado. Por exemplo, clique em *Ferramentas de Análise* e depois em *Ok* e a instalação será realizada. Pronto, agora você já pode testar várias ferramentas deste suplemento. Vá na aba *Dados* e clique em *Análise de Dados* para ver as opções de análise (Figura 2.2). Para exemplificar, baixe o arquivo denominado `Dados1.xlsx` do site *www.helitontavares.com/probabilidade*. Para montar o histograma de um conjunto de dados na planilha `Exemplo1`, que contém $n = 1000$ observações na coluna A (A1:A1000) e os intervalos de classe (bloco) para que sejam contadas as frequencias (B1:B10). Preencha as informações nas caixas conforme a Figura2.2.

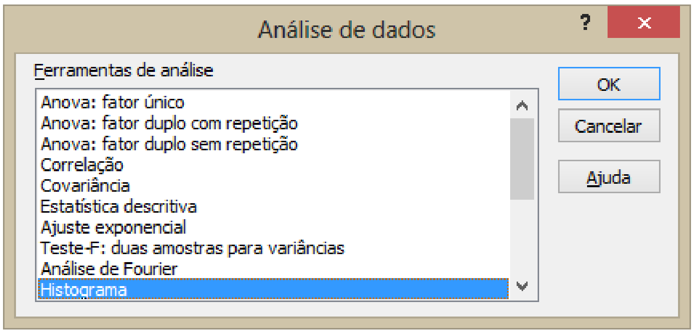


Figura 2.1 *Opções de Análise de Dados*

Exemplo 2.1. *No arquivo* `Dados1.xlsx`*, monte o histograma para as demais planilhas. A fórmula que gerará os dados está apenas na célula A1, mas você poderá copiá-las até a célula A1000.*

Instale também o suplemento *Solver* para trabalhar com funções e o *Análise de Dados - VBA* para a construção de rotinas (também chamadas de códigos, scripts ou macros) de programação.

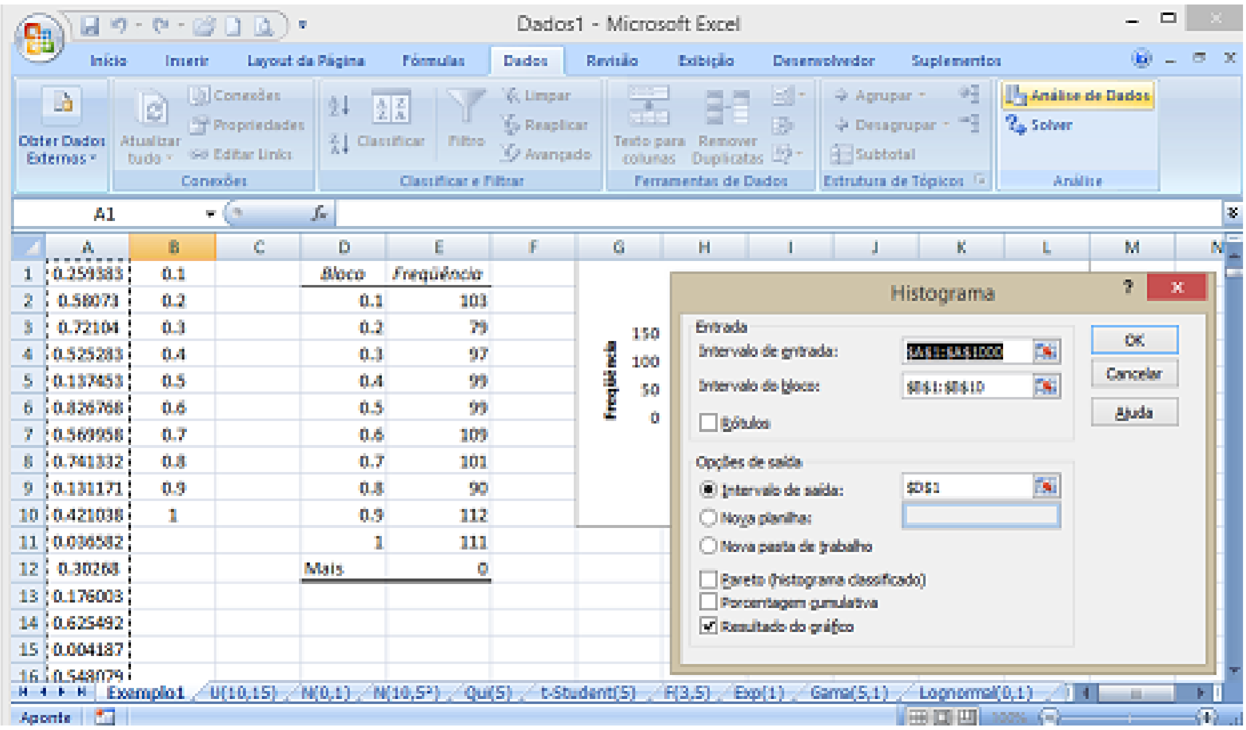


Figura 2.2 *Passos para montagem de um histograma*

2.3 Programação em VBA

2.3.1 Orientações gerais

Em muitas situações é mais conveniente elaborar macros no ambiente do Visual Basic for Applications (VBA), dentro do Excel. Pra abrir o ambiente deve-se inicialmente pressionar simultaneamente as teclas **ALT** e **F11** (esta ação é geralmente representada por **ALT+F11**). Com o ambiente aberto, simplemente aperte **F7** para iniciar o ambiente de digitação, ou clique sobre umas das planilhas (Plan1, por exemplo) e escolha a opção "exibir código". A Figura 2.3.1 apresenta o layout do editor em que você digitará sua macro.

A macro já pode ser digitada, começando sempre com `Sub NomeMacro()`, (onde `NomeMacro` é o nome da macro, a critério) e terminando com `End Sub`. Para executá-la, digite F5 (se quiser executar passo a passo, pressione F8 repetidamente).

Há muitas funções prontas para serem incluídas nas macros. Uma relação bem ampla pode ser encontrada no arquivo `Apostila_Excel_VBA.pdf` no site *www.helitontavares.com/prog*. No material que segue, vamos evitar inicialmente o uso de funções prontas, de forma a entender como construí-las.

Você pode preferir, pode configurar o Excel para apresentar os ícones do VBA em uma de suas abas. No Excel 2007 basta clicar no ícone redondo do `Office`, depois em `Opções do Excel > Personalizar`, e marcar a caixa `Mostrar guia Desenvolvedor na Faixa de Opções`. No Excel 2010 basta clicar em `Arquivos > Opções > Personalizar Faixa de Opções`, e marcar a caixa `Desenvolvedor` nas guias do lado direito.

Observação 2.1. *Algumas sugestões ou cuidados na elaboração de uma macro são:*

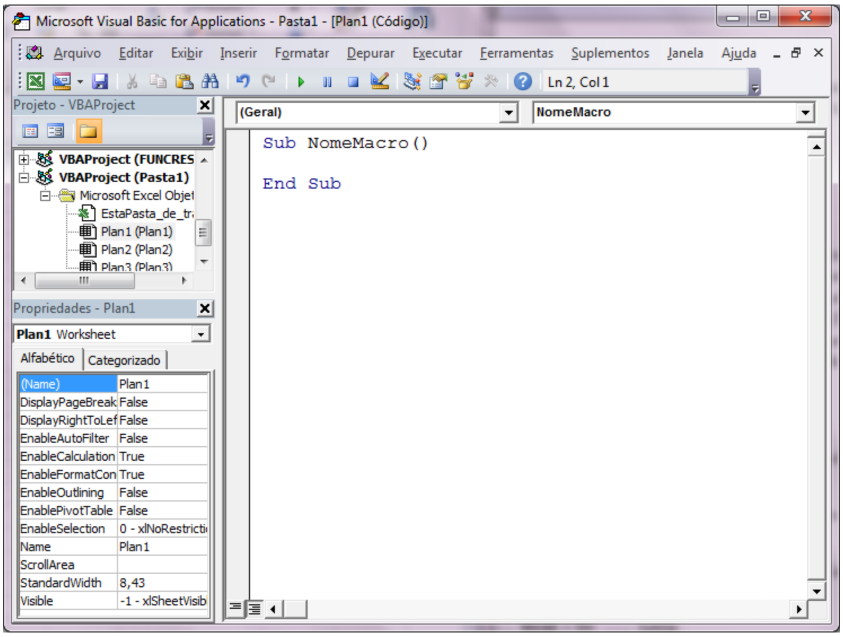


Figura 2.3 *Editor de VBA*

(a) *Quando quisermos colocar observações na macro, devemos usar aspas simples (') no início da observação.*

(b) *Normalmente devemos definir o tipo de variável antes de iniciar, mas vamos deixar esta etapa para quando necessário.*

(c) *É frequente em probabilidade a repetição de um experimento n vezes, por isso o uso de Loops é corriqueiro, particularmente do* `For ... Next`.

(d) *Estruturas alternativas de loops são* `While ... Wend` *e* `Do ... Loop`.

(e) *Também é frequente a verificação de uma condição, com o uso de IF ... Then. A estrutura IF ... Then ... Else é um pouco mais complexa, necessitando ser concluída com um* `End If`

(f) *Toda vez que formos acumular ou contar valores em uma variável, esta deve ser zerada antes de iniciar a operação. Por exemplo, com o comando Soma = 0 ou Cont = 0.*

(g) *Para solicitar a digitação de uma informação, use* `C= InputBox("Digite um caracter")`, *onde C é a variável que receberá o conteúdo digitado.*

(h) *Para apresentarmos na tela várias parcelas, o símbolo* `&` *deve separá-los.*

(i) *Podemos gerar valores dentro das células da planilha Excel. Para isso, usamos o comando* `=cells(i,j)`, *onde i e j são, respectivamente a linha e a coluna da planilha. Por exemplo, B3 equivale a cells(3,2) (ver Figura2.3.1).*

(j) *Se quisermos obter ou colocar um conteúdo em uma célula da planilha* `Plan1`, *basta usarmos* `Sheets("Plan1").Cells(i,j)`

(k) *Se desejar colocar vários comandos na mesma linha, basta separá-los por dois pontos (:). Por exemplo,* `a = 1 : b = 5 : c = 10`

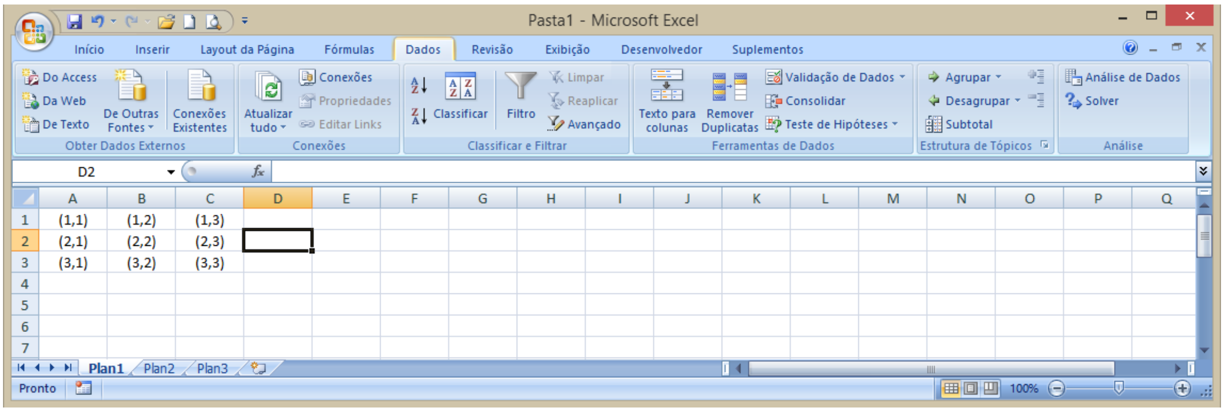


Figura 2.4 *Visual matricial da planilha do Excel*

A seguir serão apresentados alguns exemplos de macros. Evitaremos colocar acentos nas palavras, pois ao copiarmos uma macro de um arquivo pdf, os caracteres acentuados (ou o ç) podem dar origem a outros caracteres.

Exemplo 2.2. *Elabore uma macro para solicitar um número e imprimir na tela uma mensagem.*

```
Sub Informacao()
 C = InputBox("Digite um caracter")
 MsgBox "Voce digitou o caracter " & C
End Sub
```

Exemplo 2.3. *Elabore uma macro para solicitar um número e verificar se ele é igual a 1, imprimindo na tela uma mensagem.*

```
Sub Teste1()
 C = InputBox("Digite um numero")
 If C = 1 Then MsgBox "Voce digitou o caracter 1!!!"
End Sub
```

Nesta macro nenhuma mensagem será apresentada se for digitado um número diferente de 1. Podemos alterá-la para apresentar uma mensagem em qualquer caso.

```
Sub Teste2()
 C = InputBox("Digite um numero")
 If C = 1 Then
    MsgBox "Voce digitou o caracter 1!!!"
 Else
    MsgBox "Voce nao digitou o caracter 1!!!" & " Voce digitou o " & C
 End If
End Sub
```

Exemplo 2.4. *Elabore uma macro para solicitar dois numeros e indicar qual é o menor deles.*

```
Sub Menor1()
 Dim x, y As Integer
 x = InputBox("Digite o primeiro numero")
 y = InputBox("Digite o segundo numero")
 If x < y Then
    MsgBox "O menor numero e' o " & x
 Else
    MsgBox "O menor numero e' o " & y
 End If
End Sub
```

Vale ressaltar que em alguns casos é importante declarar o tipo de variável (`Dim x, y As Integer`), pois se o VBA identificar como texto (string), o resultado da condição `x < y` pode se inverter. Ilustraremos essa situação quando necessária.

Podemos alterar a macro anterior criando uma variável para receber o menor valor. Este procedimento é essencial quando temos mais de dois valores.

```
Sub Menor2()
 Dim x, y As Integer
 x = InputBox("Digite o primeiro numero")
 y = InputBox("Digite o segundo numero")

 If x < y Then
    Menor = x
 Else
    Menor = y
 End If
 MsgBox "O menor número e' o " & Menor
End Sub
```

Agora consideremos três números. Podemos ainda ter uma solução bem simples, fazendo uso do **AND**.

Exemplo 2.5. *Elabore uma macro para solicitar três números distintos e indicar qual é o menor deles.*

```
Sub Menor3()
 Dim x, y As Integer
 x = InputBox("Digite o primeiro número")
 y = InputBox("Digite o segundo número")
 z = InputBox("Digite o terceiro número")

 If x < y AND x < z Then Menor = x
 If y < x AND y < z Then Menor = y
 If z < x AND z < y Then Menor = z
```

```
 MsgBox "O menor número é o " & Menor
End Sub
```

Poderíamos ter uma saída alternativa, considerando o x inicialmente como o menor, e testando os demais com relação a x. Se y já for menor que x, então faremos a variável "Menor"receber o valor de y. Vejamos um exemplo, que pode ser estendido para um número maior de valores, substituindo os passos 2 e 3 abaixo por um **LOOP**.

```
Sub Menor4()
 Dim x, y As Integer
 x = InputBox("Digite o primeiro numero")
 y = InputBox("Digite o segundo numero")
 z = InputBox("Digite o terceiro numero")

 Menor = x                        ' Passo 1
 If y < Menor Then Menor = y      ' Passo 2
 If z < Menor Then Menor = z      ' Passo 3

 MsgBox "O menor número é o " & Menor
End Sub
```

Em muitas situações temos que ler não apenas três, mas um conjunto bem maior de valores. O Excel funciona como uma matriz denominada **CELLS** (células), em que podemos obter seus valores sabendo-se a linha e coluna. Para uso do CELLS (ou cells), as colunas de referência $(A, B, C, \cdots)$ terão suas letras substituidas por números $(A \equiv 1, B \equiv 2, \cdots)$. Por exemplo, o elemento $B3$ equivale a $cells(3, 2)$.

Exemplo 2.6. *Elabore uma macro para ler as células* $A1 : A10$ *e indicar qual é o menor deles.*

Na execução do exemplo abaixo, é didático acompanhar o resultado a cada loop. Tecle `F8` para executar uma linha de cada vez e coloque o cursor sobre a variável *Menor* para visualizar seu valor.

```
Sub Menor5()
 Menor = Cells(1, 1)   ' leu o elemento A1
 For i = 2 To 10
    If Cells(i, 1) < Menor Then Menor = Cells(i, 1)  ' lê mais um elemento
 Next
 MsgBox "O menor numero e' o " & Menor
End Sub
```

Naturalmente esse programa pode ser facilmente adaptado para indicar qual é o máximo de um conjunto de dados.

Exemplo 2.7. *Elabore uma macro para ler as células* $A1 : A10$ *e indicar qual é o maior deles.*

```
Sub Maior()
 Maior = Cells(1, 1)   ' leu o elemento A1
 For i = 2 To 10
    If Cells(i, 1) > Maior Then Maior = Cells(i, 1)  ' lê mais um elemento
```

```
 Next
 MsgBox "O maior numero e' o " & Maior
End Sub
```

Ao invés de ler o elemento da célula, você pode colocar o elemento na célula. Vejamos um exemplo:

Exemplo 2.8. *Elabore uma macro para colocar os números 1 a 10 nas células A1 a A10 da planilha.*

```
Sub Gera10()
 For i = 1 To 10
    Cells(i, 1) = i
 Next
End Sub
```

Também podemos mesclar números e string (texto), usando o operador `&` para juntá-los. Vejamos um exemplo:

```
Sub Gera10()
 For i = 1 To 10
    Cells(i, 1) = "Eu sou nota " & i
 Next
End Sub
```

2.4 Exercícios

Exercício 2.4.1. *Elabore uma macro para solicitar um número n e construir as seguintes quantidades:*

$$S1 = \sum_{k=1}^{n} k, \quad S2 = \sum_{k=1}^{n} k^2, \quad S3 = \sum_{k=1}^{n} k^3 \quad S4 = \sum_{k=1}^{n} k^4,$$

lembrando que o `"^"` *serve para exponenciação no VBA, e apresentando as respectivas mensagens. Verifique se as fórmulas I e II de PA e PG (pág. 17) estão corretas.*

Exercício 2.4.2. *Elabore uma macro para colocar nas células:*
i) A1:A10: números 1 a 10
ii) B1:B10: o dobro dos números na coluna A
iii) C1:C10: o quadrado dos números na coluna A
iv) D1:D10: valores da coluna A acumulados
v) E1:E10: valores da coluna C acumulados
vi) F1: soma dos valores da coluna A ($\sum_{i=1}^{n} X_i,\ n = 10$)
vii) F2: soma dos valores da coluna C ($\sum_{i=1}^{n} X_i^2$)
viii) F3: média dos valores da coluna A, por $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$
ix) F4: variância dos valores da coluna A, por $Var(X) = \frac{1}{n}\sum_{i=1}^{n} X_i^2 - \overline{X}^2$
x) F5: desvio-padrão dos valores da coluna A, por $DP(X) = \sqrt{Var(X)}$

Observação 2.2. *É frequente usarmos o conceito de* ***Momentos Amostrais de Ordem k da variável X****, definidos por*

$$M_{X,k} = \frac{1}{n}\sum_{i=1}^{n} X^k.$$

Neste caso temos que $E(X) = M_{X,1}$ *e* $Var(X) = M_{X,2} - M^2_{X,1}$. *Algumas vezes usa-se o* ***Momentos Amostrais de Ordem k em torno de*** $\overline{X}$, *definido por* $M'_{X,k} = \frac{1}{n}\sum_{i=1}^{n}(X - \overline{X})^k$ *e neste caso temos que* $Var(X) = M'_{X,2}$. *Se não houver risco de confusão, a denominação da variável será retirada da notação, ficando* M_k *e* M'_k *respectivamente.*

Exercício 2.4.3. *Elabore uma macro para, com os valores quaisquer de A1:A10:*
i) Obter o mínimo, colocar em D1
ii) Obter o máximo, colocar em D2
iii) Ordenar a coluna A em ordem crescente, colocar em B1:B10.
iv) Ordenar a coluna A em ordem decrescente, colocar em C1:C10.

Exercício 2.4.4. *Elabore uma macro para, com os valores quaisquer de A1:A1000:*
i) Construir a Amplitude Total: $X_{max} - X_{min}$.
ii) Construir o Desvio Médio: $DM = \frac{1}{n}\sum_{i=1}^{n}|X_i - \overline{X}|$.
iii) Construir o Desvio Médio Absoluto: $DM = \frac{1}{\overline{X}}\sum_{i=1}^{n}|X_i - \overline{X}|$.
iv) Construir o Primeiro Quartil (Q_1), *ou seja, 25% estão abaixo dele.*
v) Construir o Segundo Quartil (Q_2), *que é a Mediana* (Md), *com 50% abaixo dela.*
vi) Construir o Terceiro Quartil (Q_3), *ou seja, 75% estão abaixo dele.*
vii) Construir a Amplitude Semi-Interquartílica (ou Desvio-Quartílico): $DQ = \dfrac{Q_3 - Q_1}{2}$.
viii) Construir o Percentil de Ordem $k = 10, 25, 50, 75, 90$ *(Obs.:* $P_{25} = Q_1, \quad P_{50} = Q_2, \quad P_{75} = Q_3$*)*
ix) Construir a Moda (valor mais frequente, se houver).
x) Construir o Coeficiente de Assimetria de Pearson: $CAP = (\overline{X} - Md)/S$
xi) Construir o Coeficiente de Assimetria de Bowley: $CAB = (Q_1 + Q_3 - 2Md)/S$
xii) Construir o Coeficiente de Curtose: $K = \dfrac{Q_3 - Q_1}{2(P_{90} - P_{10})}$.
xiii) Construir o Coeficiente de Variação (ou Dispersão Relativa): $CV = S/\overline{X}$

Exercício 2.4.5. *Elabore uma macro para, com* $n = 10$ *valores quaisquer em A1:A10 e B1:B10, representando duas variáveis* X *e* Y, *obter:*

i) os primeiros Momentos Amostrais de X *e* Y, *dados por* $M_{X,1}$ *e* $M_{Y,1}$ *(simplificando,* M_X *e* M_Y*), e os desvios-padrão de* X *e* Y.

ii) a soma dos produtos cruzados: $S_{XY} = \sum\limits_{i=1}^{n} x_i y_i$

iii) a média dos produtos: $M_{XY} = S_{XY}/n$

iv) o Coeficiente de Covariância, dado por $Cov(X,Y) = M_{XY} - M_X M_Y$

v) o Coeficiente de Correlação, dado por $Corr(X,Y) = \dfrac{Cov(X,Y)}{DP(X)DP(Y)}$

2.5 Geração de números aleatórios

Na prática da Estatística acredita-se que os dados vieram de alguma distribuição de probabilidade (ou são realização de alguma variável aleatória), e por isso há necessidade de apresentar algumas das principais distribuições de probabilidades, para dados discretos e contínuos. A destas procuramos

	Função	Notação
1	=ALEATÓRIO()	$U(0,1)$
2	=ALEATÓRIO()*5+10	$U(5,10)$
3	=INV.NORMP(ALEATÓRIO())	$N(0,1)$
4	=INV.NORM(ALEATÓRIO();10;5)	$N(10,5^2)$
5	=INV.QUI(ALEATÓRIO();5)	$\chi(5)$
6	=INVT(ALEATÓRIO();5)	$t(5)$
7	=INVF(ALEATÓRIO();3;5)	$F(3,5)$
8	=INVGAMA(ALEATÓRIO();1;1)	$Exp(1)$
9	=INVGAMA(ALEATÓRIO();5;1)	$Gama(5,1)$
10	=INVLOG(ALEATÓRIO();0;1)	$LN(0,1)$

Tabela 2.1 *Exemplos das principais funções de geração de dados no Excel*

conhecer algumas características importantes, tais como o valor esperado (média), desvio-padrão (ou a variância), coeficientes de assimetria e curtose. Alguns desses conceitos serão vistos detalhadamente mais à frente, mas será importante já antecipar alguns resultados de forma a obter valores aproximados nesta seção.

Assim, nesta seção vamos apresentar algumas macros de geração de valores de variáveis (observações ou números aleatórios), em geral partindo de uma variável X com distribuição *Uniforme* no intervalo (0,1) [frequentemente representada por $X \sim U(0,1)$], ou seja, serão feitas funções ou transformações de variáveis de forma a obter a variável desejada. O procedimento geral será gerar n observações e com estes construir alguma função, tipo *soma*, média ou variancia, ou transformá-las, ou ainda verificar se alguma condição está satisfeita. O valor n, poderá, em alguns casos, ser chamado de tamanho da amostra. Na área de Estatística dizemos que quando o tamanho da amostra aumenta, o resultado da amostra (estimativa) converge para o verdadeiro valor (parâmetro populacional), e boa parte dos exemplos a seguir são exemplos desse conceito.

No VBA a função **RND()** simula uma observação de uma $U(0,1)$, com resultado restrito a [0,1), ou seja, ele pode até ser nulo, mas nunca será igual a 1. Na planilha do Excel a fórmula equivalente ao RND() é o `=ALEATORIO()`. A Tabela 2.1 apresenta um conjunto de funções para gerar valores das principais distribuições usadas na Estatística para modelagem de dados ou testes estatísticos.

Exemplo 2.9. *Elabore uma macro para acumular (somar) n observações de uma $X \sim U(0,1)$ e depois apresentar na tela o valor médio. Use $n = 10000$.*

```
Sub MacroMedia()
n = 10000
Soma = 0
For i = 1 To n
    x = Rnd()
    Soma = Soma + x
Next
MsgBox "Media = " & Soma / n
End Sub
```

Para a variável $U(0,1)$ temos que sua média teórica é 1/2, e representamos por $E(X) = 1/2$. Você poderá perceber que quando aumentamos o n o valor obtido será cada vez mais próximo do valor teórico.

Exemplo 2.10. *Complemente a macro anterior para calcular a Variância dos valores gerados, usando a fórmula já apresentada:* $Var(X) = \frac{1}{n}\sum_{i=1}^{n} X_i^2 - \overline{X}^2 = M_2 - M_1^2$. *(Lembre-se que* $Var(X) = E(X^2) - [E(X)]^2$*.)*

```
Sub MacroMediaVar()
n = 1000000
Soma = 0
Soma2 = 0
For i = 1 To n
   x = Rnd()
   Soma = Soma + x
   Soma2 = Soma2 + x ^ 2
Next

Media = Soma / n
Variancia = Soma2/n -  Media ^ 2

MsgBox "Media = " & Media
MsgBox "Variancia = " & Variancia

End Sub
```

Exemplo 2.11. *Elabore uma macro para gerar observações de uma* $X \sim U(0,1)$ *e transformá-las em Bernoullis* Y *de parâmetro* $p = 0.3$*, apresentando ao final o valor médio destas Bernoullis. Use* $n = 10000$.

Este procedimento consiste simplesmente em gerar variáveis $X \in [0,1]$ e usar a regra: se $X \leq p$, então $Y = 1$; se $X > p$, adotamos $Y = 0$. Dessa forma os eventos $\{Y = 1\}$ e $\{X \leq p\}$ são equivalentes, e daí $P(Y = 1) = P(X \leq p) = p$. Vale lembrar que se $X \sim U(a,b)$, então $F_X(x) = (x-a)/(b-a)$, e no caso particular $a = 0$ e $b = 1$, temos que $F_X(x) = x$ para $x \in (0,1)$.

```
Sub Macrobern()
n = 10000
p = 0.3
Soma = 0
For i = 1 To n
    If Rnd() < p Then
        x = 1
    Else
        x = 0
    End If
    Soma = Soma + x
```

```
Next
MsgBox "Média = " & Soma / n
End Sub
```

Exemplo 2.12. *Elabore uma macro para gerar observações de uma variável $X \sim U(0,1)$, transformá-la em uma $Y \sim U(a,b)$, com a e b, através da fórmula $Y = (b-a)X + a$ e depois apresentar na tela o valor médio de Y. Use $n = 1000$, $a = 10$ e $b = 30$. [Lembre-se que $E(Y) = (a+b)/2$ e $Var(X) = (b-a)^2/12$]*

```
Sub MacroUab()
n = 10000
a = 10
b = 30
Soma = 0
For i = 1 To n
    x = Rnd()
    y = (b - a) * x + a
    Soma = Soma + y
Next
MsgBox "Média = " & Soma / n
End Sub
```

Exemplo 2.13. *Elabore uma macro para gerar observações de uma variável $X \sim U(0,1)$, transformá-la em uma Exponencial com parâmetro λ através da fórmula $Y = -\frac{1}{\lambda}\ln(1-X)$ e depois apresentar na tela o valor médio de Y. Use $n = 1000$ e $\lambda = 0.1$.*

```
Sub Macro3()
n = 10000
lambda = 0.1
Soma = 0
For i = 1 To n
    x = Rnd()
    y = -1 / lambda * Log(1 - x)
    Soma = Soma + y
Next
MsgBox "Média = " & Soma / n
End Sub
```

Para gerar observações de uma distribuição normal o processo é um pouco mais complicado. Um procedimento é seguir um conhecido problema de probabilidade em que geramos aleatoriamente duas uniformes(0,1) e usamos uma transformação, mostrada a seguir.

Exemplo 2.14. *Elabore uma macro para gerar observações de $X \sim U(0,1)$ e $Y \sim U(0,1)$, e transformá-las em $N(0,1)$ através das fórmulas $Z = \sqrt{-2\ln(X)}\cos(2\pi Y)$ e $W = \sqrt{-2\ln(X)}\texttt{sen}(2\pi Y)$. Apresente na tela os valores médios de Z e W. Use $n = 10000$.*

```
Sub SomaNormais()
n = 10000
Pi = 4 * Atn(1)
SomaZ = 0: SomaW=0
For i = 1 To n
   x = Rnd() :  y = Rnd()
   Z = Sqr(-2 * Log(x)) * Cos(2 * Pi * y)
   w = Sqr(-2 * Log(x)) * Sin(2 * Pi * y)
   SomaZ = SomaZ + Z
   SomaW = SomaW + W
Next
MsgBox "Média de Z = " & SomaZ / n & "; Média de W = " & SomaW / n
End Sub
```

Um procedimento muito frequente é pegar um número e associar a uma determinada categoria ou intervalo. Na prática estamos discretizando essa variável. Vejamos abaixo.

Exemplo 2.15. *Elabore uma macro para gerar uma observação de uma $X \sim U(0,1)$ e transformá-las em um Uniforme discreta em $\{1, 2, \cdots, 10\}$. Em suma, estamos discretizando uma variável.*

A função $\mathsf{int}(x)$ trunca (arredonda pra baixo) o valor x. Com isso, $10 * x$ transformará inicialmente os valores da $U(0,1)$ em $U(0,10)$, e a função $\mathsf{int}(10 * x)$ transformará em $0, 1, \cdots, 9$, de forma que precisamos somar 1 a estes números para obter $1, 2, \cdots, 10$.

```
Sub Discretiza()
x = Rnd()
y = Int(10 * x) + 1
MsgBox "Valores gerados e transformados: " & x & "  " & y
End Sub
```

Outras situações podem ser construídas, mas tudo se resume a construir uma função geral, envolvendo a função $\mathsf{int}(x)$.

Exemplo 2.16. *Elabore uma macro para gerar uma observações de uma $X \sim U(0, 1000)$ e transformá-las em um Uniforme discreta em $\{10, 11, 12, \cdots, 30\}$.*

Exemplo 2.17. *Elabore uma macro para gerar observações de uma $X \sim U(50, 100)$ e transformá-las em um Uniforme discreta em $\{10, 13, 16, \cdots, 31\}$.*

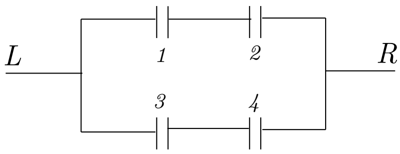
Exercício 2.5.1. *Elabore uma macro para responder à seguinte pergunta: se selecionarmos $r = 23$ pessoas ao acaso, qual a probabilidade de pelo menos duas fazerem aniversário no mesmo dia do ano? Considere que há 365 dias. Use n = 10000.*

Esta situação consiste apenas em gerar o nascimento de um indivíduo, ou seja, um número aleatório inteiro representando os dias do ano: $1, 2, \cdots, 365$. Da mesma forma gere o nascimento dos outros 22 indivíduos. Agora verifique se algum número se repetiu (coincidência). Para estimar a probabilidade, esse processo deve ser repetido um número grande de vezes (n), e verificando a proporção de ocorrências de coincidêcias.

```
Sub Aniversarios()
 Dim F(365) As Integer
 n = 10000
 dias = 365
 r = 23
 C = 0

 For i = 1 to n
    Coincidiu = 0
    For k = 1 to r
        x = Rnd()
        y = Int(dias * x) + 1 'Gera um número inteiro entre 1 e 365 (dias)
        F(y) = F(y) + 1       'Gera a frequência ao dia
    Next

    For y = 1 to dias
        If F(y) >= 2 Then Coincidiu = 1
        F(y) = 0
    Next

    If Coincidiu = 1 Then C = C + 1
 Next

 MsgBox "Probabilidade aproximada: " & C / n

 End Sub
```

Exercício 2.5.2. *Complemente o exercício anterior de forma a termos o resultado para valores de* $r = 10, 11, \cdots, 60$. *Plot um gráfico da probabilidade (y) como função de r (x).*

Exercício 2.5.3. *No circuito abaixo a probabilidade de que cada um dos quatro relés esteja fechado é p. Se todos funcionares independentemente, qual será a probabilidade de que a energia passe da esquerda (L) para a direita (R)?*



Denotemos esta probabilidade por $\theta(p)$. Com um cálculo simples podemos obter $\theta(p) = 2p^2 - p^4$. Este caso é extremamente simples, em sistemas complexos temos funções extremamente complexas.

Vamos tentar confirmar a função $\theta(p)$ via programa. Para isso, vamos criar $Bernoullis(p)$ para cada relé indicando se está aberto ou fechado. Para que haja energia entre L e R devemos ter $\{R1 = 1\} \bigcap \{R2 = 1\}$, que equivale a $\{R1 * R2 = 1\}$. Da mesma forma, podemos ter $\{R3 * R4 = 1\}$, ou ambas. Podemos sintetizar dizendo que que $L \leftrightarrow R$ se $\{R1 * R2 + R3 * R4 \geq 1\}$. É esta condição que deverá ser verificada no programa.

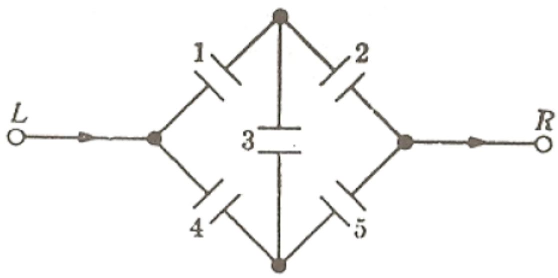
```
 Sub circuito()
```

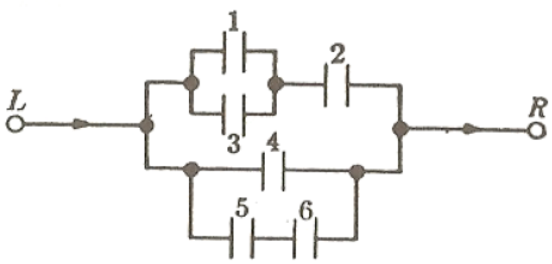
```
 n = 1000000 :  p = 0.2 :  conta = 0
 For i = 1 To n
    R1 = (Rnd() > p) + 1
    R2 = (Rnd() > p) + 1
    R3 = (Rnd() > p) + 1
    R4 = (Rnd() > p) + 1
    If (R1 * R2 + R3 * R4 >= 1) Then conta = conta + 1
Next
    MsgBox "Estimativa da Probabilidade: " & conta / n
End Sub
```

Observação 2.3. *Perceba que outras construções ou condições podem ser adotadas. Por exemplo,* $\{R1 + R2 + R3 + R4 \geq 3\}$ *também poderia ser usada para indicar o fluxo de corrente entre L e R.*

Exercício 2.5.4. *Complemente o exercício anterior de forma a termos o resultado para valores de* $p \in (0, 1)$ *com incremento (step)* $p = 0.01$. *Faça um gráfico da probabilidade (y) como função de p (x).*

Exercício 2.5.5. *Nos circuito abaixos a probabilidade de que cada um dos relés esteja fechado é p. Considerando que todos funcionam independentemente, obtenha a probabilidade de que a energia passe da esquerda (L) para a direita (R). Construa um programa para estimar essas funções de probabilidade.*





Os exercícios anteriores podem ser vistos da seguinte forma: dividimos o intervalo (0,1) em vários intervalos de mesmo tamanho, onde cada um está associado a um rótulo, tipo 1,2,3,$\cdots$,10. Depois geramos uma $U(0, 1)$ e verificamos em que intervalo caiu, mostrando seu rótulo. Mas podemos pensar em intervalos de tamanhos (probabilidades) diferentes. Com apenas dois intervalos estaremos gerando Bernoullis, e com mais interrvalos estaremos gerando Multinomiais.

Exemplo 2.18. *Elabore uma macro para gerar uma observação de uma Multinomial, assumindo o valor 1 com probabilidade 0,5, o valor 2 com probabilidade 0,3 e o valor 2 com probabilidade 0.2.*

```
Sub MacroMult()
n = 10000
p1 = 0.5: p2 = 0.3: p3 = 0.2
Cont1 = 0: Cont2 = 0: Cont3 = 0

For i = 1 To n
```

```
    p = Rnd()
    If p < p1 Then
        x = 1
        Cont1 = Cont1 + 1
    ElseIf p < p1 + p2 Then
        x = 2
        Cont2 = Cont2 + 1
    Else
        x = 3
        Cont3 = Cont3 + 1
    End If
Next
MsgBox "Media1 = " & Cont1 / n & ", Media2 = " & Cont2 / n
End Sub
```

O próximo passo é tentar identificar a distribuição de um conjunto de dados. Como já vimos, o Excel já tem algumas ferramentas para isso situado em *Análise de Dados* no Menu *Dados*, caso este suplemento tenha sido instalado (ver Seção 2.2).

Alternativamente, poderíamos mandar o Excel contar quantas observações estão abaixo de um determinado valor com a função **CONT.SE**, ou ainda **CONT.SES**. Após isso, podemos gerar o gráfico manualmente. Bastaria usar uma das funções abaixo para cada intervalo:

`=CONT.SE(A1:A20;"<0.1")` ou `=CONT.SES(A1:A20;">0";A1:A20;"<.1")`

No entanto, se o objetivo é gerar observações sem a necessidade de colocá-las na planilha, o que pode ser extremamente útil, temos que montar macros para gerar valores e logo identificar em que intervalo o valor se encontra, incrementando a frequência desse intervalo (ou seja, somando 1). Esse será o próximo conteúdo.

2.5.1 Montando uma distribuição de frequências

Exemplo 2.19. *Elabore uma macro para gerar n observações de uma $X \sim U(0,1)$, definir $K = 10$ intervalos (por exemplo [0;0,1], (0,1; 0,2], ..., (0,9;1]) e contar quantas observações caem em cada intervalo. Em suma, elaborar uma Tabela de Frequência de X. Use $n = 1000$.*

Uma forma de elaborar esta macro é discretizando as observações no intervalo $\{1, 2, \cdots, 10\}$, que são índices do vetor F.

```
Sub Freq()
Dim F(10) As Integer
n = 1000
For i = 1 To n
    x = Rnd()
    y = Int(10 * x) + 1
    F(y) = F(y) + 1
Next
For i = 1 To 10
```

```
    MsgBox "Frequencia na classe " & i & ": " & F(i)
Next
End Sub
```

Exemplo 2.20. *Elaborar uma macro para construir uma distribuição de frequência criando uma matriz que conterá os limites superiores dos intervalos, e use $X \sim U(-2,2)$. Leia os limites de uma planilha (*`B2:BK`*, onde* `K` *é o número de classes) e coloque as frequências relativas obtidas na mesma planilha com o comando* cells(linha,coluna).

```
Sub FreqInt()
Dim F(10, 2) As Single
K = InputBox("Introduza o Número de classes:")
n = 10000
For j = 1 To K
    F(j, 1) = Cells(j + 1, 2)
Next
For i = 1 To n
    y = Rnd() * 4 - 2
    For j = 1 To K
        If (y < F(j, 1)) Then F(j, 2) = F(j, 2) + 1
    Next
Next
For j = 1 To K
    Cells(j + 1, 3) = F(j, 2) / n
Next
MsgBox "Concluído!"
End Sub
```

Exemplo 2.21. *Elaborar uma macro para distribuição de frequência criando uma matriz que conterá também os limites dos intervalos, e use $X \sim N(0,1)$. Leia os limites de uma planilha e coloque as frequências relativas obtidas na mesma planilha com o comando* cells(linha,coluna)*, e plotar o um gráfico de dispersão com as frequências obtidas.*

```
Sub FreqInt()
Dim F(20, 2) As Single

K = 20
n = 10000
Pi = 4 * Atn(1)

For j = 1 To K
    F(j, 1) = Cells(j + 1, 2)
Next

For i = 1 To n
    R1 = Rnd()
    R2 = Rnd()
```

```
    y = Sin(2 * Pi * R1) * Sqr(-2 * Log(R2)) 'Gerando N(0,1)

    For j = 1 To K
        If (y < F(j, 1)) Then F(j, 2) = F(j, 2) + 1
    Next
Next

For j = 1 To K
    Cells(j + 1, 3) = (F(j, 2) - F(j - 1, 2)) / n
Next

' Geração de Gráfico
    ActiveSheet.Shapes.AddChart.Select
    ActiveChart.ChartType = xlXYScatter
    ActiveChart.SetSourceData Source:=Range("B2:C21")
    ActiveChart.SeriesCollection(1).Select
End Sub
```

Observação 2.4. *Vale notar que o VBA tem uma função própria para gerar observações N(0,1):* `y = Application.NormSInv(Rnd())`*, assim como o valor de* π*,* `Application.Pi()`*. Além disso, há outras formas de gerar observações normais. As linhas abaixo foram escritas em outra linguagem, mas mostra um algoritmo eficiente para geração de normais padrão.*

```
float x1, x2, w, y1, y2;

do {
        x1 = 2.0 * ranf() - 1.0;
        x2 = 2.0 * ranf() - 1.0;
        w = x1 * x1 + x2 * x2;
} while ( w >= 1.0 );

w = sqrt( (-2.0 * ln( w ) ) / w );
y1 = x1 * w;
y2 = x2 * w;
```

2.5.2 Gerando distribuições amostrais

Na etapa anterior geramos observações de algumas distribuições em particular, atribuindo valores aos parâmetros dessas distribuições. Também calculamos a média amostral e a variância amostral, verificando que quando o n (tamanho da amostra) cresce, o valor produzido por estas quantidades amostrais se aproximam dos valores teóricos. A Estatística funciona assim, ou seja, alguma distribuição, com um determinado parâmetro (ou mais de um) gerou os dados, e um dos principais objetivos é usar esses dados para tentar "descobrir"(ou *estimar*) esse parâmetro que gerou os dados.

As fórmulas adotadas, tal como a *média amostral* ou *variância amostral*, serão chamados de *estimadores* e os valores numéricos gerados por eles para cada amostra serão chamados de *estimativas*. E poderemos ter vários estimadores para um mesmo para um mesmo parâmetro. Naturalmente, se tivermos amostras diferentes (ainda que de mesmo tamanho n) poderemos ter estimativas diferentes, de forma que esses tais estimadores são também variáveis aleatórias, e, portanto, têm que ter uma distribuição associada. Para que tenhamos alguma informação sobre a distribuição dos estimadores, geraremos muitas dessas estimativas e estudaremos suas características (histograma, média, variância etc.). Esse será o procedimento geral nas próximas etapas.

Exemplo 2.22. *Elaborar uma macro para gerar amostras de tamanho $n = 5$ de uma $U(0, 1)$ e salvar na planilha a média amostral e a variância amostral. A macro deverá repetir esse processo $r = 10.000$ vezes, colocando em linhas distintas essas estimativas. Ao final, apresente a distribuição de frequência dessas estimativas. As estimativas médias estão próxims dos populacionais (teóricos)?*

Vamos inicialmente obter as estimativas médias:

```
Sub EstimativasU01()
n = 5
r = 10000

For k = 1 To r
   Soma = 0:   Soma2 = 0
   For i = 1 To n
      x = Rnd()
      Soma = Soma + x
      Soma2 = Soma2 + x ^ 2
   Next

   M1 = Soma / n   ' Momento Amostral de ordem 1
   M2 = Soma2 / n  ' Momento Amostral de ordem 2
   Cells(k, 1) = M1
   Cells(k, 2) = M2 - M1 ^ 2
Next
End Sub
```

Note que a estimativa média deve estar bem próxima do verdadeiro valor (teórico) $E(X) = \frac{1}{2}$. No entanto, a estimativa média para a variância deve estar um tanto abaixo do valor teórico $Var(X) = \frac{1}{12}$. Multiplique o valor obtido por $n/(n-1)$ e você terá uma proximidade bem melhor, sem viés. Isso ocorreu porque na fórmula da variância adotada dividimos por n, e para construir um estimador melhor devemos dividir por $n-1$. Neste caso dizemos que o estimador é *não viesado* (ou não viciado). Naturalmente, para valores grandes de n não fará diferença relevante dividir por n ou por $n-1$, por isso nesse caso a variância amostral viesada (dividindo-se por n) deverá ser um bom estimador da verdadeira variância populacional.

Agora faça a distribuição de frequência e veja o comportamento (distribuição) dos dois estimadores: média amostral e variância amostral não-viesada, dados pelas fórmuas abaixo:

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \quad e \quad S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2.$$

2.5.3 Organizando macros em sub-macros ou funções

A subdivisão de uma macro em várias menores pode facilitar a compreensão da mesma, embora exija maiores cuidados. Se o resultado da submacro for um valor, podemos chamá-la de *função*, com algumas particularidades. Uma função recebe um ou mais argumentos e calcula um valor. Vamos refazer a macro anterior usando uma sub-macro que será repetida para cada réplica dentro da macro mãe.

```
Sub Amostra(k)
n = 5
Soma = 0:   Soma2 = 0
For i = 1 To n
   x = Rnd()
   Soma = Soma + x
   Soma2 = Soma2 + x ^ 2
Next

M1 = Soma / n   ' Momento Amostral de ordem 1
M2 = Soma2 / n  ' Momento Amostral de ordem 2
Cells(k, 1) = M1
Cells(k, 2) = M2 - M1 ^ 2
End Sub
```

```
Sub EstimativasU01()
r = 10000
For k = 1 To r
   Amostra(k)
Next
End Sub
```

Um cuidado muito importante é com o *escopo* da variável. A variável r definida na rotina principal recebe o valor $r = 10000$ nesta rotina, mas na rotina *Amostra* a variável r estará sem valor algum.

Exemplo 2.23. *Construir uma rotina com o uso de funções para calcular* $C(n, k) = n!/(k!(n - k)!)$

```
Function Fatorial(k)
Fatorial = 1
For i = 1 To k
   Fatorial = Fatorial * i
Next
End Function

Function combinacao(n1, k1)
combinacao = Fatorial(n1) /(Fatorial(k1) * Fatorial(n1 - k1))
End Function

Sub calculacomb()
n = InputBox("Insira um número n")
k = InputBox("Insira um número k")
MsgBox "C(" & n & "," & k & ") = " & combinacao(n, k)
End Sub
```

2.5.4 Obtenção de Probabilidades sob a visão frequentista

O conceito freqüentista de probabilidade envolve basicamente a elaboração de uma seqüencia de repetições para um determinado evento. Essas repetições são normalmente denominadas de **Réplicas**. A idéia de repetições justifica a denominação “teoria freqüentista”, ou seja, baseada em frequências. A teoria ampara-se na regularidade estatística das freqüencias relativas e sustenta que a probabilidade de um dado acontecimento pode ser medida observando a freqüencia relativa desse acontecimento, em uma sucessão numerosa de experiências idênticas e independentes, cada uma delas resultando em *sucesso* ou *fracasso.*
Nesta linha de raciocínio, quando queremos determinar a probabilidade de ocorrência de um evento A, repetimos o experimento um número grande, digamos n, de vezes, de onde observamos n_A ocorrências do evento A (sucesso). A probabilidade de ocorrência de A, representada por $P(A)$, será aproximada por

$$P(A) \simeq \frac{n_A}{n}. \tag{2.1}$$

Um exemplo simples pode ser sobre a probabilidade de sair cara no lançamento de uma determinada moeda. O evento está bem claro, *A: sair cara no lançamento da moeda*, e se a moeda for regular (ou honesta), teremos que $P(A) = \frac{1}{2}$, mas nunca sabemos se é realmente regular. Para verificar podemos fazer n lançamentos da moeda e anotar os resultados. Se a moeda for honesta, devemos ter aproximadamente metade ($n_A = n/2$) lançamentos em que o resultado será cara, e teremos

$$P(A) \simeq \frac{n/2}{n} = \frac{1}{2}. \tag{2.2}$$

Exercício 2.5.6. *Considere um quadrado com $x \in [0,1]$ e $y \in [0,1]$. Sorteando um ponto (x,y) ao acaso, qual a probabilidade deste ponto cair no triângulo inferior do quadrado (área hachureada)?*

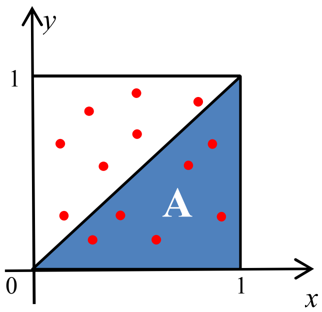


Figura 2.5 *Área do evento de interesse*

Primeiro devemos identificar a área de interesse (evento A), que neste caso é o conjunto dos pontos em que $x > y$, ou seja, $\boldsymbol{A} = \{(x,y) : x > y\}$. Feito isso, em termos de frequência relativa, este problema poderia ter a seguinte reformulação: *Elabore uma macro para gerar n observações de uma $X \sim U(0,1)$ e $Y \sim U(0,1)$. Verifique se elas estão dentro do triângulo inferior (ou seja, se $x > y$), contando os pontos em que essa condição é verdadeira e, ao final, obtendo a proporção destes pontos por n_A/n.*

```
Sub Triangulo()
```

```
n = 100000
Conta = 0

For i = 1 To n
   x = Rnd()
   y = Rnd()
   If x > y Then Conta = Conta + 1
Next
MsgBox "Probabilidade = " & Conta / n

End Sub
```

Observação 2.5. *Quando necessário, usaremos a seguinte notação de região:* $x \in [a, b]$ *e* $y \in [c, d]$*, então teremos o retângulo* $\boldsymbol{R} = [a, b] \times [c, d]$*. Ainda, quando a região for quadrada, tal como* $\boldsymbol{Q} = [a, b] \times [a, b]$*, usaremos* $\boldsymbol{Q} = [a, b]^2$*.*

Observação 2.6. *Como já visto em diversas situações, se gerarmos um valor x no intervalo (0,1), então a transformação* $(b - a)x + a$ *passará a ter valores no intervalo* (a, b)*.*

Exercício 2.5.7. *Considere um retângulo* $\boldsymbol{R} = [0, 1] \times [0, 2]$*. Sorteando um ponto* (x, y) *ao acaso, qual a probabilidade deste ponto cair na área em que* $x < y$*?*

Exercício 2.5.8. *Considere um quadrado* $\boldsymbol{Q} = [-1, 1]^2$*. Sorteando um ponto* (x, y) *ao acaso em* $\boldsymbol{Q}$*, qual a probabilidade deste ponto cair dentro do círculo unitário?*

Esta probabilidade pode ser calculada diretamente pela relação entre as áreas do círculo de raio 1, que é π, e a do quadrado $[-1, 1]^2$, que é 4. Portanto, a probabilidade desejada será $\pi/4 = 0,785$, aproximadamente.

Primeiro devemos identificar a área de interesse, que neste caso é o conjunto dos pontos em que $x^2 + y^2 < 1$, ou seja, $\boldsymbol{A} = \{(x, y) \in [-1, 1]^2 : x^2 + y^2 < 1\}$. Feito isso, em termos de frequência relativa, este problema consiste em sortear um número grande de pontos (x, y). Para cada ponto, o evento A (sucesso) ocorrerá se o ponto (x, y) estiver dentro do círculo. Em termos de implementação ele poderia ter a seguinte reformulação: *Elabore uma macro para gerar n observações de uma* $X \sim U(0, 1)$ *e* $Y \sim U(0, 1)$ *e depois transforme-as para o intervalo [-1;1]. Verifique se elas estão dentro do círculo de raio 1 (ou seja, se* $x^2 + y^2 \leq 1$*, contando qual a proporção destes pontos.*

```
Sub Circulo()
n = 100000
Conta = 0

For i = 1 To n
    x = Rnd()
    y = Rnd()
    x = 2 * x - 1
    y = 2 * y - 1
    r = Sqr(x ^ 2 + y ^ 2)
    If (r < 1) Then Conta = Conta + 1
Next
```

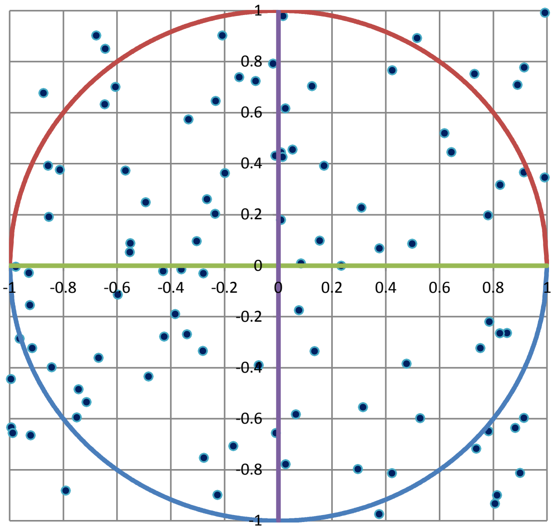


Figura 2.6 *Contagem do número de eventos favoráveis*

```
MsgBox "Probabilidade = " & Conta / n

End Sub
```

Exemplo 2.24. *Repita o exemplo anterior considerando uma esfera de raio 1 dentro de um cubo de lado 2, e conte a proporção de pontos que cairá na esfera.*

Este cálculo também pode ser feito exatamente pela relação entre o volume da esfera ($\frac{4}{3}\pi$) e o volume do cubo de lado 2 ($2^3 = 8$), resultando em $\pi/6 \equiv 0,5236$, aproximadamente.

```
 Sub Esfera()
 n = 10 ^ 7
 Conta = 0

For i = 1 To n
    x = Rnd()
    y = Rnd()
    Z = Rnd()

    x = 2 * x - 1
    y = 2 * y - 1
    Z = 2 * Z - 1
    r = Sqr(x ^ 2 + y ^ 2 + Z ^ 2)
    If (r < 1) Then Conta = Conta + 1
  Next
 MsgBox "Probabilidade = " & Conta / n
 End Sub
```

2.5.5 Aproximando uma integral

Podemos usar este mesmo princípio para aproximar o valor de uma integral, que em muitas situações equivale a calcular uma probabilidade. Por exemplo, desejamos obter a área da região A, contida em um retângulo R. Essa integral (probabilidade) pode ser calculada diretamente pela relação entre as áreas da região A e a de R. Considerando a região A definida $x \in (0,1)$ e $y \in (0, f(x))$ (ver Figura 2.8), claramente a área de R é 1, e a área de A é dada por

$$\int_0^1 f(x)dx = \int_0^1 x^2 dx = \left[\frac{x^3}{3}\right]_0^1 = \frac{1}{3}.$$

Portanto, a área desejada será $\frac{1/3}{1} = \frac{1}{3}$. Para fins de simulação, devemos gerar um ponto x e calcular $f(x)$. Este valor será usado para determinar a probabilidade de sucesso do evento A desejado. Portanto, deverá ser gerado um outro ponto y cuja probabilidade de sucesso (estar sob a curva) será dada por $f(x)$, ou seja, $y < f(x)$ será a condição de sucesso. Enfim, contaremos o número de sucessos.

Exercício 2.5.9. *Elabore uma macro para gerar n observações de uma $X \sim U(0,1)$ e $Y \sim U(0,1)$. Verifique se elas estão dentro da região A definida pelas funções $f(x) = 0$ e $f(x) = x^2$, $x \in (0,1)$, contando qual a proporção de pontos que caem em A.*

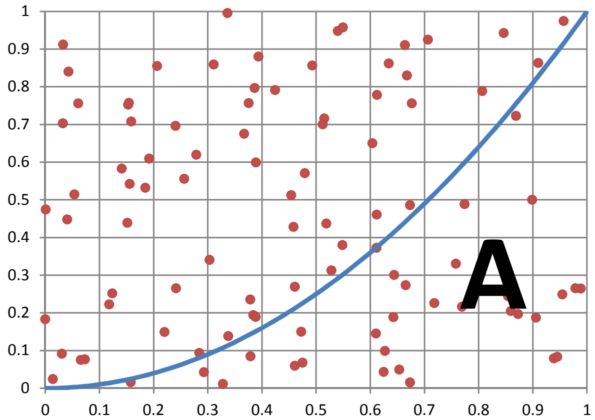


Figura 2.7 *Contagem do número de eventos favoráveis*

```
Sub Integral()
 n = 100000
 Conta = 0
 For i = 1 To n
    x = Rnd()
    f = x ^ 2
    y = Rnd()
    If y < f Then Conta = Conta + 1
 Next
 MsgBox "Probabilidade = " & Conta / n
End Sub
```

Caso a variação de x ou f não fosse o intervalo $(0, 1)$, deveríamos transformar x e y linearmente por $x^* = (b_x - a_x)x + a_x$ e $y^* = (b_y - a_y)x + a_y$, de forma a cobrir a área desejada.

Exemplo 2.25. *Elabore uma macro para obter uma aproximação para a integral de* $f(x) = x^2$, $x \in (0, 2)$.

Claramente essa integral resulta em 8/3. Notemos que $x \in (0, 2)$ e $f(x) \in (0, 4)$, de forma que devemos aumentar a área do quadrado que engloba a função f no intervalo desejado.

```
 Sub Integral()
 n = 100000
 Q = 8   ' área do quadrado [0,2]x[0,4]
 Conta = 0
 For i = 1 To n
    x = Rnd() * 2  'Intervalo [0,2]
    f = x ^ 2
    y = Rnd() * 4  'Intervalo [0,4]
    If y < f Then Conta = Conta + 1
 Next
 MsgBox "Integral = " & Conta / n * Q
End Sub
```

Vamos considerar agora o caso bidimensional. Seja $f(x, y) = \exp\{-(x^2 + y^2)/2\}, (x, y) \in \mathbb{R}^2$. Esta função é simétrica em torno da origem, com seu máximo no ponto $(0, 0)$, igual a $\exp\{0\} = 1$, com variação predominante no intervalo $[-4; 4]$. O resultado dessa integral é $2\pi \simeq 6,283$.

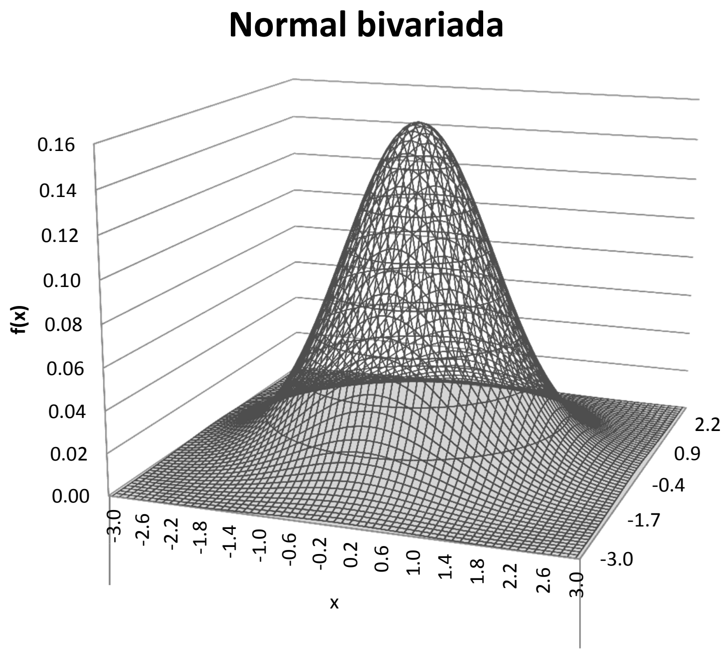


Figura 2.8 *Integral de uma normal bidimensional*

```
Sub Gaussiana2()
n = 10 ^ 7
a = 4    ' intervalo [-a , a]
Q = 1 * (2 * a) ^ 2 ' volume do paralelepípedo
For i = 1 To n
   x = Rnd() * (2 * a) - a ' [-a , a]
   y = Rnd() * (2 * a) - a ' [-a , a]
   f = Exp(-(x ^ 2 + y ^ 2) / 2)

   Z = Rnd()
   If Z < f Then Conta = Conta + 1
Next
Integ = Conta / n * Q
MsgBox "Integral: " & Integ
End Sub
```

2.6 Alguns problemas especiais de Probabilidade

Vamos agora considerar a seguinte situação: se $X_i \sim Exp(\lambda),\ i = 1, \cdots, m$, e $Y = \sum_{i=1}^{m} X_i$, vamos obter a $P(Y > K)$.

Exercício 2.6.1. *Elabore uma macro para gerar m observações de uma $X \sim Exp(\lambda)$ e somá-las. Verifique se esta soma é superior a K. Repita o processo n vezes, calculando a proporção de pontos que satisfaz a condição. Use $m = 15$, $\lambda = 20$, $K = 365$ e $n = 10000$.*

```
Sub Exponencial()

n = 10000
m = 15
lambda = 20
K = 365

Conta = 0
For i = 1 To n
    Soma = 0
    For j = 1 To m
        x = -lambda * Log(1 - Rnd())
        Soma = Soma + x
        Next
    If (Soma > K) Then Conta = Conta + 1
Next
Prob = Conta / n
MsgBox "Probabilidade: " & Prob

End Sub
```

2.7 O Método da Transformação Inversa

Uma forma mais geral de geração de variáveis aleatórias é através de sua Função de Distribuição $F_X(x) = P(X \leq x) \in [0, 1]$, chamado *Método da Transformação Inversa*. Gerando $Y \sim U(0, 1)$, e resolvendo $Y = F(X)$, chegamos a $X = F^{-1}(Y)$ com a distribuição desejada.

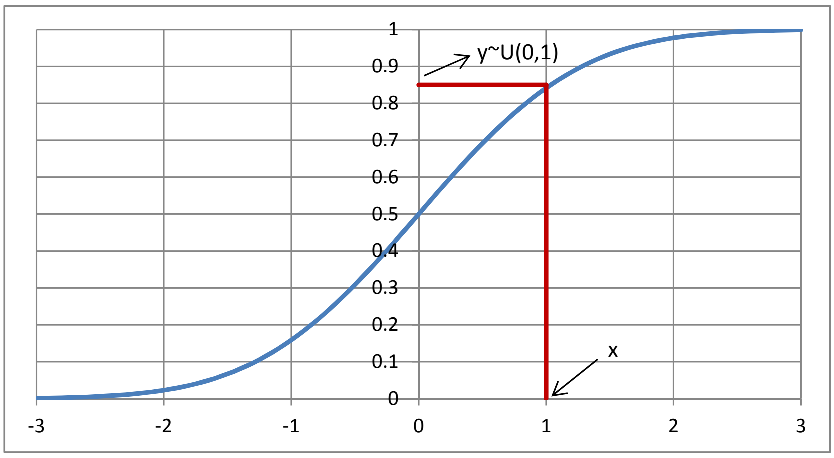


Figura 2.9 *Transformação Inversa*

Exemplo 2.26. *Considerando* $X \sim U(a, b)$, *temos que* $F_X(x) = (x - a)/(b - a)$. *Resolvendo* $y = (x - a)/(b - a)$ *obtemos* $x = (b - a)y + a$. *Basta agora gerar* $Y \sim U(0, 1)$ *e fazer a transformação. Podemos notar que* $x \in (a, b)$.

Exemplo 2.27. *Considerando* $X \sim Exp(\lambda)$, *temos que* $F_X(x) = 1 - e^{-\lambda x}$. *Resolvendo* $y = 1 - e^{-\lambda x}$ *obtemos* $x = -\frac{1}{\lambda} \ln(1 - y)$. *Basta agora gerar* $Y \sim U(0, 1)$ *e fazer a transformação. Podemos notar que* $x \in (0, \infty)$.

2.8 O Ambiente R

R é uma plataforma livre para computação estatística que inclue um excelente ambiente gráfico. O **R** começou a ser desenvolvido por Robert Gentleman e Ross Ihaka do Departamento de Estatística da Universidade de Auckland na Nova Zelândia, mais conhecidos por "R & R", apelido do qual se originou o nome **R** do programa. Com o incentivo de um dos primeiros usuários deste programa, Martin Mächler do ETH Zürich (Instituto Federal de Tecnologia Zurique da Suíça), "R & R"lançaram o código fonte do **R** em 1995, disponível por ftp (uma forma de se transferir dados pela internet), sobre os termos de Free Software Foundations GNU general license, que seria um tipo de "licença para softwares livres". Desde então o **R** vem ganhando cada vez mais adeptos em todo o mundo, em parte devido ao fato de ser totalmente gratuito e também por ser um programa que exige do usuário o conhecimento das análises que está fazendo, diminuindo assim as chances de uma interpretação errada dos resultados. Outro fator importante para a difusão do **R** é a sua compatibilidade com quase todos os sistemas operacionais.

Apresentaremos a seguir um breve resumo de suas funções, deixando os detalhes para os muitos manuais disponíveis na Internet; na página *http://www.helitontavares/R/* há vários deles. Iniciamos com algumas observações gerais antes de apresentar as funções mais usuais. Naquela página também serão disponibilizados os executáveis para instalação.

Caso sensitivo Todas as funções em R são caso-sensitivas, ou seja, variáveis com letras maiúsculas são diferentes de variáveis com letras minúsculas. Por exemplo, *x=2* e *X=3* serão duas variáveis distintas com valores diferentes.

Help Para obter ajuda sobre um comando *plot*, use *help(plot)* ou *?plot*. O mesmo vale para qualquer outro comando.

Script O conjunto de comandos será chamado de script. às vezes encontrase com outras denominações (macro, sintaxe, programa etc). Podemos usar a combinação de teclas **CONTROL+R** para rodar as linhas selecionadas de script.

RStudio O RStudio é uma interface que organiza de forma bem eficiente várias janelas no R. Recomenda-se fortemente a sua instalação, após a instalação do R. A partir disso, basta abrir o RStudio que ele já ativará o R.

package De forma geral, os comandos do R estão dentro depacotes (packages), de forma que antes de executar comandos específicos é necessário carregar um ou mais pacotes.

Abaixo apresentamos o visual do RStudio. Ele está dividido em quatro janelas, em que a primeira (*superior-esquerda*) contém o conjunto de comandos que queremos executar (Script), podendo conter vários scripts, em abas distintas, um ao lado do outro. A janela *superior-direita* contém o workspace (todas as variáveis criadas), bem como o histórico. Na janela *inferior-esquerda*, chamada **console**, são apresentados os resultados dos processamentos dos scripts. Na janela *inferior-direita*, podemos visualizar todos os gráficos gerados, além de help, pacotes (packages) e arquivos. Um breve resumo dos principais comandos do R será apresentado a aseguir.

2.8.1 Rodando scripts

É sempre aconselhável escrever os scripts e salvá-los em um arquivo, que terão a extensão `.R`. Também é imporante que os scritps contenham explicações sobre o que faz cada operação; usa-se o caracter ♯ para iniciar um comentário. Na página www.helitontavares,com/R com o arquivo *script1.R* que contém todos os comandos aqui apresentados.

Para rodar um scritp podemos ir até a janela de scripts, marcar as linhas a serem executadas e clicar no ícone de Run, ao topo desta janela. Também podemos teclar CONTROL+R. Alternativamente, podemos copiar ou digitar diretamente no *console* e teclar *enter*. Veremos que o R enumera as linhas com [1], [2] ...

2.8.2 Obtendo ajuda sobre algum comando

Na janela de ajuda, clique o nome do comando comando desejado. Alternativamente, podemos escrever no script o comando help(plot) ou ?plot e executá-lo.

2.8.3 Fomas de atribuições de valores

i) `a<-2`

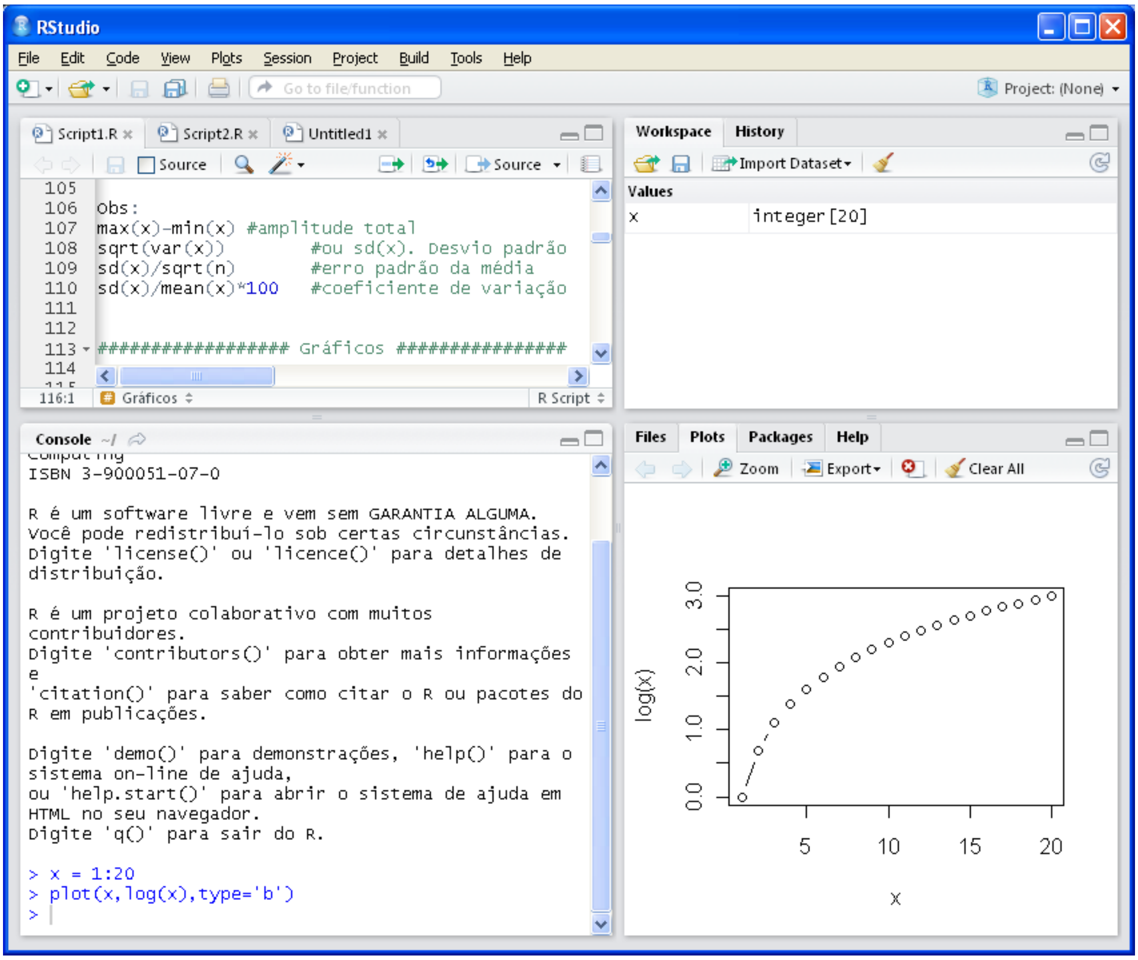


Figura 2.10 *Apresentação do RStudio*

ii) `2->a`

iii) `a=2`

2.8.4 Construindo sequências

i) `A=1:10` produz 1,2,3,...,9,10

ii) `C = seq(from=1, to=10)` mesmo que 1:10

iii) `D = seq(length=51, from=-5, by=.2)`

iv) `C = seq(from=1, to=4, length=6)`

v) `C = seq(from=1, to=4, by=0.5)`

vi) `C = seq(2,50,2)` produz 2,4,6,8, ... ,50

vii) `E = rep(2, times=5)`

viii) `rep(5, times=3)` produz 5 5 5

ix) `rep(1:5, 3)` produz [1] 1 2 3 4 5 1 2 3 4 5 1 2 3 4 5

2.8.5 Montando vetores e matrizes

A função **c()** é usada para juntar (concatenar) diferentes elementos para formar um arranjo maior. Vejamos alguns exemplos:

i) `B=c(1,2,3)` produz 1,2,3

ii) `F = c(B, 0, A)`

iii) `a = c(3.4, pi, exp(-1))`

iv)

v)

Uma primeira aplicação: imc

O índice de massa corpórea (imc), muito utilizado em diversas áreas médicas, bem como em Educação Física, é obtido pela fórmula: $imc = peso/altura^2$. Nesta aplicação foram obtidos os pesos e alturas de seis pessoas e desejamos calcular seus imc's.

```
peso = c(60, 72, 57, 90, 95, 72)
altura = c(1.75, 1.80, 1.65, 1.90, 1.74, 1.91)
imc = peso/(altura^2)
imc
```

2.8.6 Principais funções matemáticas

i) `log(x)       Log de base e de x`

ii) `exp(x)       Antilog de x (e^x)`

iii) `log(x,n)     Log de base n de x`

iv) `log10(x)     Log de base 10 de x`

v) `sqrt(x)      Raiz quadrada de x`

vi) `choose(n,x)  n!/(x!(n-x)!)`

vii) `cos(x), sin(x), tan(x) Funções trigonométricas de x em radianos`

viii) `acos(x), asin(x), atan(x) Funções trig. inversas de x em radianos`

ix) `abs(x) Valor absoluto de x`

x) `round(x, dig = 1)`

xi) `sort(x)    # Organizando os dados`

xii) `sort(x,decreasing=T)`

Principais funções Estatísticas

i)	`min(x)`	`Mínimo do vetor x`
ii)	`max(x))`	`Máximo do vetor x`
iii)	`sum(x)`	`Soma dos elementos de x`
iv)	`length(x)`	`Num. elementos de x`
v)	`mean(x)`	`Média amostral`
vi)	`var(x)`	`Variância amostral`
vii)	`sd(x)`	`Desvio padrão amostral`
viii)	`median(x)`	`Mediana amostral`
ix)	`quantile(x,p)`	`Quantis p=(p1,p2...) dos elementos de x. Ex. p=c(.25,.5,.75)`
x)	`cor(x,y)`	`Correlação amostral entre X e Y`
xi)	`range(x)`	`Equivalente a c(min(x),max(x))`
xii)	`table()`	`Frequências`
xiii)	`summary(x)`	`Resumo de estatísticas descritivas`
xiv)	`cumsum(x)`	`Acumulada prod(x)`
xv)	`cumprod(x)`	`Acumula os produtos`
xvi)	`diff(x)`	`x(i)-x(i-1)`
xvii)	`which.min(x)`	`Posição do mínimo`
xviii)	`which.max(x)`	`Posição do máximo`

Observação 2.7. *: Algumas funções podem ser construídas a partir das funções.*

```
max(x)-min(x)             amplitude total
colMeans(H)               Media para cada coluna
sd(x)/mean(x)*100         coeficiente de variação
sd(x) = sqrt(var(x))
diag(var(H))              variancias por coluna
```

2.8.7 Gráficos bidimensionais

A principal função para plotar uma função $f(x)$ é a *plot*, que tem várias opções. Basicamente, devemos criar um vetor x com os pontos em que serão calculados os valores da função $y = f(x)$. A cada comando plot será gerado um novo gráfico, e todos poderão ser acessados através das setas (esquerda e direita) da janela plots. Abaixo segue um conjunto de exemplos e opções desta função. O help com as opções pode ser obtido com *help*(*plot*).

i) `plot(x,y,type='b')`

ii) `plot(x,log(x),type='b')`

iii) `plot(x, y, type = "l")`

iv) `plot(x, y, type = "p")`

v) `plot(x, y, type = "o")`

vi) `plot(x, y, type = "b")`

vii) `plot(x, y, type = "h")`

viii) `plot(x, y, type = "S")`

ix) `plot(x, y, type = "s")`

x) `plot(x, y, type = "n")`

xi) `points(rev(x),y)#adiciona pontos ao gráfico ativo`

xii) `lines(x,100-y)  #adiciona linhas ao gráfico ativo`

Alterando alguns padrões dos gráficos

i) `points(x,8000-y,pch="*") #símbolo asterisco`

ii) `points(rev(x),y,pch=3)   #adiciona cruzes`

iii) `plot(x,y,pch="@")`

iv) `plot(x,y,pch=1:3)`

v) `plot(0:20,0:20,pch=0:20)`

vi) `curve(100*(x^3-x^2)+15, from=0, to=1)`

Gráficos bidimensionais com subgráficos

i) `par(mfrow = c(1,1)) #padrão`

ii) `par(mfrow = c(3,2)) #gráficos múltiplos`

iii) `par(mfrow = c(4, 2), mar = c(2, 2, 0.3, 0.3), mgp = c(1.5, 0.6,0))`

Gráficos tridimensionais

O principal comando para gráficos em $3d$ é o `open3d()`.

i) `x = sort(rnorm(1000))`

ii) `y = rnorm(1000)`

iii) `z = rnorm(1000) + atan2(x,y)`

iv) `plot3d(x, y, z, col=rainbow(1000))`

2.8.8 Operações matriciais

i) `A = 1:9`

ii) `A = matrix(A, nrow = 3, ncol = 3, byrow = TRUE)`

iii) `A[3,3]=10 # Muda o elemento (3,3) da matriz A`

iv) `A[,1]`

v) `B[1,3]; B[c(1,2),3];B[c(1,3),c(2,4)] #selecionando partes de uma matriz`

vi) `B = t(A)`

vii) `rowSums(A[,2:3])`

viii) `diag(c(1, 2, 3))`

ix) `A * B   # Produto ponto a ponto`

x) `A %*% B  # Produto matricial`

xi) `A.inv = solve(A)) # Inversa de A`

xii) `A %*% A.inv  #`

xiii) `det(A)   determinante da matriz A`

xiv) `pmin, pmax`

2.8.9 Distribuições de probabilidade e funções associadas

Há muitas distribuições de probabilidade na área de Estatística. Para cada distribuição podemos estar interessados em quatro tipos de operações, representadas por d, p, q, r, descritas abaixo.

d : função (densidade) de probabilidade avaliada em um ponto;

p : probabilidade associada a algum evento,

q : valores para construir Intervalos de Confiança (IC),

r : observações pseudo-aleatórias.

Estas letras precedem as funções das distribuições. Os principais casos são:

i) `Normal:  dnorm(), pnorm(), qnorm(), rnorm()`

ii) `t-Student: dt(), pt(), qt(), rt()`

iii) `Qui-quadrado:  dchisq(), pchisq(), qchisq(), rchisq()`

iv) `Poisson: dpois(), ppois(), qpois(), rpois()`

v) `Uniforme:`

vi) `Gama:`

vii) `Cauchy:`

viii) `Geométrica:`

ix) `Hipergeométrica:`

x) `F:`

xi) `Exponencial:`

xii) `Beta:`

xiii) `Weibull:`

Exemplo 2.28. *Obter o valor $f(x)$ da densidade da $N(0,1)$ avaliada no ponto $x = 1$.*

```
dnorm(x=1,mean=0,sd=1)
```

Resultado, $f(1) = 0.2419707$

Exemplo 2.29. *Plotar o gráfico da fdp $N(0,1)$.*

```
x = seq(-3,3,0.01)
plot(x,dnorm(x,0,1),type="l",xlab="x",ylab="f(x)")
```

Exemplo 2.30. *Calcular a probabilidade $P(Z > -1)$, onde $Z \sim N(0,1)$.*

```
pnorm(-1)
```

Exemplo 2.31. *Obter $P(Z < -1) = 0.1586553$ (duas opções)*

```
pnorm(q=-1,mean=0,sd=1,lower.tail=T)
1-pnorm(q=1,mean=0,sd=1,lower.tail=F)
```

Exemplo 2.32. *Ober o valor de z tal que $P(Z > z) = 0.95$. Resultado: $z = 1.644854$.*

```
qnorm(p=0.95,mean=0,sd=1)
```

2.8.10 Geração de valores de variáveis aleatórias no R

Uma das necessidades mais importantes é a geração de valores de cada ditribuição, simulando o que ocorre na prática. A letra **r** que determina essas funções, provém da palavra *random*, que significa aleatório. A cada vez que executamos um comando, são gerados diferentes valores aleatórios, mas se quisermos repetição dos valores podemos resetar a semente (ou raiz) com o comando `set.seed(1)`. A seguir apresentamos funções definidas no R para geração de valores de variáveis aleatórias para várias distribuições.

Exemplo 2.33. *Gerar 20 observações de uma N(0,1).*
`rnorm(n=20)`

Exemplo 2.34. *Gerar 20 observações de uma* $N(10, 9)$.
`rnorm(n=20,mean=10,sd=3)`

Descritivas

Para gerar estatísticas descritivas de um conunto de dados podemos usar a função `summary`.

Exemplo 2.35. *Gerar 1000 valores de uma* $N(0, 1)$ *e apresentar um sumário.*

```
nrv = rnorm(mean=0,sd=1,n=10000)
print(summary(nrv))
```

2.8.11 Construindo um histograma

Para plotar um histograma de um conunto de dados, usamos a função `hist`.

Tabela 2.2 *Funções para geração de observações no R*

Função	Distribuições
runif(x,min,max)	Uniforme
rnorm(x,mean,sd)	Normal
rlnorm(x,mean,sd)	Lognormal
rt(x,df)	t-Student
rf(x,df1,df2)	F
rchisq(x,df)	Qui-quadrado
rexp(x)	Exponencial
rgamma(x,shape,scale)	Gama
rbeta(x,a,b)	Beta
rcauchy(x,location,scale)	Cauchy
rbinom(x,n,p)	Binomial
rgeom(x,p)	Geométrica
rpois(x,lambda)	Poisson
rhyper(x,m,n,k)	Hipergeométrica
rweibull(x,m)	Weibull

Exemplo 2.36. *Gerar 1000 valores de uma $N(0,1)$ e apresentar um histograma.*

```
nrv = rnorm(mean=0,sd=1,n=10000)
hist(nrv)
```

Exemplo 2.37. *Gerar 1000 valores de uma $N(0,1)$ e apresentar um histograma.*

```
x = rnorm(1000)                      #Geração das observações
par(mfrow=c(2,1))                    #Serão dois gráficos
hist(x,main=1)                       #histograma padrão
hist(x,breaks=seq(-5,5,.1),main=2)   #histograma com opções
```

Perceba que foram apresentados dois gráficos na mesma figura com o `par(mfrow=c(2,1))` .

Exemplo 2.38. *Gerar 1000 valores de uma $\chi^2(10)$ e apresentar um histograma.*

```
x=rchisq(1000,10);              #Geração das 1000 observações
hist(x)                         #histograma padrãao de x
hist(x,                         #histograma de x com opções:
main="Histograma Qui-quadrado", #título
xlab="Valores",                 #texto do eixo das abscissas
ylab="Prob",                    #texto do eixo das ordenadas
br=c(c(0,5),c(5,15),5*3:7),     #int das classes
xlim=c(0,30),                   #limites do eixo de x
ylim=c(0,0.1),                  #limites do eixo y
col="lightblue",                #cor das colunas
border="white",                 #cor das bordas das colunas
prob=T,                         #mostrar  probabilidades.
right=T,                        #int fechados à direita
adj=0,                          #alinhamento dos textos
col.axis="red")                 #cor do texto nos eixos
```

Perceba que foram apresentados dois gráficos em janelas disintas.

Exemplo 2.39. *Entrar com um determinado conjunto de dados e apresentar um histograma.*

```
dados<-c(25,27,18,16,21,22,21,20,18,23,27,21,19,20,21,16)
hist(dados,                     #histograma de dados
nc=6,                           #número de classes igual a 6
right=F,                        #int fechado à esquerda
main="Histograma",              #título do histograma
xlab="tempo (em minutos)",      #texto do eixo x
ylab="frequencia",              #texto do eixo y
col=2)                          #usa a cor cinza nas barras
```

2.8.12 Construindo um gráfico de barras

Para plotar um gráfico de barras de um conunto de dados, usamos a função `barplot`.

Exemplo 2.40. *Entrar com um determinado conjunto de dados e apresentar um gráfico de barras.*

```
barplot(table(c("a","a","a","a","a","b","b","b","c","c","v","v")))
```

Exemplo 2.41. *Entrar com as frequências obtidas em determinado conjunto de dados e apresentar um gráfico de barras.*

```
dados<-c("a"=4,"b"=7)
barplot(dados)
```

Exemplo 2.42. *Entrar com dados e apresentar um gráfico de barras horizontais.*

```
barplot(table(c("a","a","a","a","a","b","b","b","c","c","v","v")), hor=T)
```

Exemplo 2.43. *Gerar $n = 100$ valores $X \sim B(10, 5^2)$ e de $\chi(10)$ e plotar um gráfico de barras*

```
x = rnorm(100,10,4)
y = rchisq(100,10)
boxplot(x,y)
```

2.8.13 Construindo um gráfico de setores (pizza)

Para plotar um gráfico de pizza de um conunto de dados, usamos a função `pie`.

Exemplo 2.44. *Geração de gráfico de pizza*

```
x=c(1,1,2,2,2,2,3,3,3)
pie(x)

pie(x, labels = paste(round(tp3, dig=2), "%"), col = c(2,4))
legend("topleft", legend=names(t3), fill = c(2,4))
```

2.8.14 Integração numérica simples

Há vários métodos para obtenção de aproximações para integrais. Apresentamos, para fins de ilustração, uma ideia bem simples de como proceder.

Exemplo 2.45. *Obter valor aproximado para a integral da função $f(t) = \cos(t)$no intervalo $[0, \pi/6]$.*

```
dt = 0.005
t = seq(0, pi/6, by = dt)
ft = cos(t)
plot(t,ft)
I = sum(ft) * dt
```

2.8.15 Programando com o R

Expressõe condicionais

```
if (logical_expression) { expression_1 ... }

if (logical_expression) { expression_1 ... } else { expression_2 ... }
```

Loops

```
for (i in 1:n) {}


while (logical_expression) { expression_1 ... }
```

Capítulo **3**

Conceitos Básicos

3.1 Introdução

Muitas das decisões que ocorrem na prática, sejam por grandes empresas, bolsas de valores, planos de saúde, liberação de remédios para comercialização, dentre outras, são baseadas em algo *altamente favorável*. Ressalta-se o *altamente* porque as pessoas são diferentes, os momentos podem ser diferentes, as condições são fiferentes, de forma que raramente obtemos sucesso em 100% dos casos.

Todas as vezes que se estuda algum fenômeno observável, objetiva-se distinguir o próprio fenômeno e o modelo matemático (determinístico ou probabilístico), que melhor o explique. Os fenômenos estudados pela estatística são fenômenos cujo resultado, mesmo em condições normais de experimentação varia de uma observação para outra, dificultando dessa forma a previsão de um resultado futuro. Para a explicação desses fenômenos (fenômenos aleatórios), adotaremos um modelo matemático probabilístico. Que neste caso será o Cálculo das Probabilidades.

3.2 Experimentos Aleatórios

Hoje em dia grande quantidade de jogos é oferecida, entre os quais citamos, por exemplo: A Loteria Federal, A Sena e a Loteria Esportiva. É natural que se pense nas chances de ganhar um prêmio antes de se decidir em qual deles jogar. Um torcedor procura avaliar as chances de vitória de seu clube antes de cada jogo do qual o mesmo participará. A loteria esportiva foi criada em função do interesse do brasileiro pelo futebol e de sua paixão por jogos. Na loteria esportiva a cada rodada é escolhido um determinado número de jogos e a aposta consiste da escolha em cada jogo de um dos possíveis resultados, ou seja, vitória de um dos dois clubes ou empate.

Muitas vezes ao acordar nos perguntamos: Será que vai chover? De um modo ou de outro atribuímos um valor à chance de chover e então decidimos o tipo de roupa que usaremos e se levaremos ou não um guarda chuva conosco. Pode-se imaginar uma série de outras situações na qual nos deparamos, com a incerteza quanto à ocorrência de uma das possíveis alternativas dentro de um contexto no qual se está inserido. Por exemplo, ao chegar a uma bifurcação na qual há duas opções de tráfego para se dirigir a um local desejado, procura-se avaliar as condições de trânsito nos dois caminhos para então decidir-se por um deles.

Um analista de sistemas atribui chances aos possíveis números de usuários que estarão ligados, a uma rede durante certo período. Um engenheiro industrial avalia as chances de um determinado processo encontrar-se em equilíbrio, ou atribui chances para as possíveis proporções de peças defeituosas por ele produzidas. Um médico defronta-se com a incerteza em relação ao efeito provocado pela administração de um novo remédio a um determinado paciente.

Há uma grande classe de experimentos que, ao serem repetidos nas mesmas condições, produzem resultados diferentes. Em outras palavras, experimentos que, quando realizados, não apresentam resultados previsíveis de antemão.

Definição 3.2.1. *Experimentos que ao serem repetidos nas mesmas condições não produzem o mesmo resultado são denominados experimentos aleatórios.*
Apresenta-se a seguir alguns exemplos de experimentos aleatórios:

Exemplo 3.2.1. Quando retiramos um lote de peças em um processo de produção, observamos que o número de peças defeituosas varia de lote para lote.

Exemplo 3.2.2. O número de chamados telefônicas que chegam a uma central em um determinado intervalo de tempo não pode ser determinado de antemão.

Exemplo 3.2.3. Se escolhermos uma lâmpada do processo de fabricação e observarmos o seu tempo de duração, verificaremos que este tempo varia de lâmpada para lâmpada.

Exemplo 3.2.4. Se lançarmos um dado sobre uma superfície plana e observarmos o número que aparece na face superior, não poderemos determinar "a priori"qual será esse número.

Exemplo 3.2.5. Se selecionarmos um casal dentre vários outros casais e observarmos o sexo do primogênito, não será possível determiná-lo "a priori", e o mesmo irá variar de casal para casal.

Definição 3.2.2. *Os experimentos que ao serem repetidos nas mesmas condições conduzem ao mesmo resultado são denominados determinísticos.*
Eis alguns exemplos de experimentos determinísticos

Exemplo 3.2.6. Se deixarmos uma pedra cair de certa altura pode-se determinar sua posição e velocidade para qualquer instante de tempo posterior à queda.

Exemplo 3.2.7. Se aquecermos a água a 100 graus centígrados, a mesma entrará em ebulição.
Objetiva-se, portanto, construir um modelo matemático para representar experimentos aleatórios. Isso será feito em duas etapas: Na primeira descreveremos para cada experimento aleatório o conjunto de seus resultados possíveis, e na segunda procuraremos atribuir pesos a cada resultado que reflitam a sua maior ou menor chance de ocorrer, quando o experimento é realizado.

3.3 Espaço Amostral e Eventos

Definição 3.3.1. *Denominaremos espaço amostral associado a um experimento o conjunto de seus resultados possíveis.*
O espaço amostral será representado por um conjunto S, cujos elementos serão denominados eventos simples ou pontos amostrais. Sempre que o experimento for realizado, irá se supor que ocorrerá um e apenas um evento simples.

Exemplo 3.3.1. No Exemplo 4.2.4, que corresponde ao lançamento de um dado, o espaço amostral é o conjunto:

$$S = \{1, 2, 3, 4, 5, 6\}$$

Exemplo 3.3.2. Uma moeda é lançada duas vezes sobre uma superfície plana. Em cada um dos dois lançamentos pode ocorrer cara (K) ou coroa (C). O espaço amostral é o conjunto:

$$S = \{KK, KC, CK, CC\}$$

Exemplo 3.3.3. Três peças são retiradas de uma linha de produção. Cada peça é classificada como boa (B) ou defeituosa (D). O espaço amostral associado a esse experimento é:

$$\{BBB, BBD, BDB, DBB, BDD, DBD, DDB, DDD\}$$

Nos Exemplos 4.3.1; 4.3.2; 4.3.3; O espaço amostral é finito. Apresentaremos a seguir exemplos de experimentos aleatórios, cujos espaços amostrais não são finitos.

Exemplo 3.3.4. Uma moeda é lançada sucessivamente até que apareça cara pela primeira vez. Se ocorrer cara no primeiro lançamento o experimento termina. Se ocorrer coroa no primeiro lançamento , faz-se um segundo lançamento e se então ocorrer cara o experimento termina. Se não ocorrer cara nos dois primeiros lançamentos, faz-se um terceiro lançamento e caso não ocorra cara, faz-se um quarto lançamento e assim por diante até se verificar a ocorrência da primeira cara, que é quando o experimento termina. O espaço amostral é o conjunto:

$$\{K, CK, CCK, CCCK, CCCCK, ...\}$$

Note que os pontos desse espaço amostral, podem ser postos em correspondência biunívuca com o conjunto dos números naturais e portanto, ele é infinito, porém enumerável.

Exemplo 3.3.5. Considere o Exemplo 4.2.2, a onde observamos o número de chamadas telefônicas que chegam a uma central durante um intervalo de tempo. O espaço amostral é o conjunto:

$$S = \{0, 1, 2, 3, 4, ...\}$$

Exemplo 3.3.6. Considere a situação do Exemplo 4.2.3, a onde observamos o tempo de vida de uma lâmpada. O espaço amostral é o conjunto dos números reais não negativos. Ou seja:

$$\{x : x \ real, x \geq 0\}$$

Exemplo 3.3.7. A umidade do ar pode ser registrada por meio de higrômetro. Um higrômetro pode ser acoplado a um dispositivo que possui um ponteiro, que desliza sobre papel milimetrado e registra em cada instante a umidade do ar. Se as leituras são feitas no intervalo de tempo $[0, T]$. Então o resultado é uma curva que a cada $t \in [0, T]$, associa-se $x(t)$, que designa a umidade do ar no instante t. É razoável supor-se que $x(t)$, é uma função contínua de t. no intervalo $[0, T]$. O espaço amostral nesse caso é o conjunto:

$$S = \{x : x \text{ é uma função contínua em } [0, T]\}.$$

Nos Exemplos 4.3.4 e 4.3.5, o espaço amostral é infinito, porém enumerável, isto é, pode ser posto em correspondência biunívoca com o conjunto dos naturais.

Nos Exemplos 4.3.6 e 4.3.7, o espaço amostral é infinito e não enumerável.

Seja S o espaço amostral associado a um experimento aleatório.

Definição 3.3.2. *Denominaremos de evento a todo resultado ou subconjunto de resultados de um experimento.*

Os eventos serão representados por subconjuntos do espaço amostral. Os eventos representados por um conjunto unitário, isto é, contendo apenas um ponto do espaço amostral são denominados eventos simples. Diremos que o evento "A"ocorre quando o resultado do experimento é um evento simples pertencente à "A".

Exemplo 3.3.8. No Exemplo 4.3.3, consideremos o evento “A”: duas peças são boas. Tem-se:

$$A = \{BBD, BDB, DBB\}$$

Então “A”ocorre se ocorrer um dos três eventos simples BBD, BDB, ou DBB.

Exemplo 3.3.9. No Exemplo 4.3.4, consideramos o evento “A”: A primeira cara ocorre em um lançamento que é um múltiplo de 3. Temos então:

$$A = \{CCK, CCCCCK, CCCCCCCCK, ...\}$$

Os eventos simples de “A”têm 3n-1, coroas que precedem a ocorrência da primeira cara na posição 3n, para n = 1, 2, 3,$\cdots$

Exemplo 3.3.10. No Exemplo 4.3.2, considere o evento “B”: O número de caras é igual ao número de coroas,

$$B = \{KC, CK\}$$

3.4 Operações entre Eventos

Definição 3.4.1. *A reunião de dois eventos “A”e “B”, denotada $A \cup B$, é o evento que ocorre se pelo menos um deles ocorre.*

Definição 3.4.2. *A interseção de dois eventos “A” e “B”, denotada $A \cap B$, é o evento que ocorre se ambos ocorrem.*

Definição 3.4.3. *O complementar do evento “A”, denotado A^c é o evento que ocorre quando “A”não ocorre.*

Como os eventos são subconjuntos do espaço amostral, podemos representar a reunião de dois eventos, a interseção de dois eventos e o complementar de um evento, pelos diagramas utilizados para representar subconjuntos de um determinado conjunto.

Dado dois conjuntos A e B, as seguintes operações podem ser realizadas: União, Interseção e Complementar de eventos.

Exemplo 3.4.1. Uma urna contém bolas numeradas de 1 à 15. Uma bola é retirada da urna e seu número anotado. Sejam "A" e "B", os seguintes eventos,
A: O número da bola retirada é par;
B: O número da bola retirada é múltiplo de 3.
Determinaremos os eventos $A \cup B$, $A \cap B$ e A^c.

O espaço amostral S associado a esse experimento é o conjunto:

$$S = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15\}$$

Para "A" , "B"; $A \cup B$; $A \cap B$ e A^c; Temos:

$$A = \{2, 4, 6, 8, 10, 12, 14\}; \quad (3.1)$$
$$B = \{3, 6, 9, 12, 15\}; \quad (3.2)$$
$$A \cup B = \{2, 3, 4, 6, 8, 9, 10, 12, 14, 15\}; \quad (3.3)$$
$$A \cap B = \{6, 12\} \quad (3.4)$$
$$A^c = \{1, 3, 5, 7, 9, 11, 13, 15\} \quad (3.5)$$

Dizemos que o evento "A", implica o evento "B", que denotamos por $A \subset B$, se para todo $w \in$ "A", tivermos $w \in$ "B". Isto corresponde à situação em que a ocorrência de "A", garante inevitavelmente a ocorrência de "B".

Os eventos "A" e "B", são iguais se $A \subset B$ e $B \subset A$. Os eventos "A" e "B", são ditos mutuamente exclusivos, se eles não podem ocorrer simultaneamente. Isto é equivalente à $A \cap B = \phi$.

Apresenta-se a seguir um lema com algumas propriedades dessas operações entre eventos.

Lema 3.4.1. *Sejam "A", "B" e "C", eventos do espaço amostral S, temos:*
a) $(A \cup B) \cap C = (A \cap C) \cup (B \cap C)$;
b) $(A \cap B) \cup C = (A \cup C) \cap (B \cup C)$;
c) $(A \cup B)^c = A^c \cap B^c$;
d) $(A \cap B)^c = A^c \cup B^c$.

Demonstração:
Vamos demonstrar a) e d).
Para demonstrar a igualdade em a), precisamos mostrar que todo elemento pertencente ao lado esquerdo da igualdade, pertence ao lado direito da igualdade e vise-versa.

Se $w \in (A \cup B) \cap C$, então $w \in (A \cup B)$ e $w \in C$. Dai decorre que ($w \in A$ ou $w \in B$) e $w \in C$ e portanto, ($w \in A$ e $w \in C$) ou ($w \in B$ e $w \in C$), ou seja, ($w \in (A \cap C)$) ou ($w \in (B \cap C)$), que implica em $w \in (A \cap C) \cup (B \cap C)$. Podemos percorrer estas implicações de trás para frente e verificar que são verdadeiras, de onde decorre a igualdade dos conjuntos.

Para demonstrar a igualdade d), precisamos mostrar que;
Seja $w \in (A \cap B)^c$, então $w \notin (A \cap B)$, o que implica em $w \notin A$ ou $w \notin B$, que por sua vez implica em $w \in A^c$ ou $w \in B^c$, isto é, $w \in (A^c \cup B^c)$. Partindo de $w \in (A^c \cup B^c)$ e fazendo o percurso inverso nós obtemos a igualdade.

A seguir apresenta-se definições de operações de uma reunião e de uma interseção enumerável de eventos.

Definição 3.4.4. *O evento* $\bigcup_{i=1}^{\infty} A_i$*, é o evento que ocorre quando pelo menos um dos eventos* A_i*, para* $i = 1, 2, 3, ...,$ *ocorre.*

Definição 3.4.5. *O evento* $\bigcap_{i=1}^{\infty} A_i$ *é o evento que ocorre quando todos os eventos* A_i*,* $i = 1, 2, 3, ...,$ *ocorrem.*

3.5 Definições: Clássica, Freqüentista e Subjetiva de Probabilidade

A definição dita clássica baseia-se no conceito primitivo de eventos igualmente possíveis. Consideremos um experimento com número "Finito"de eventos simples. Vamos supor que podemos, por alguma razão, uma razão de simetria, por exemplo, atribuir a mesma chance de ocorrência a cada um dos eventos simples desse experimento. Nessas condições adotaremos a seguinte definição de probabilidade.

Definição 3.5.1. *Consideremos um espaço amostral S com "N"eventos simples, onde supõe-se que os mesmos são igualmente possíveis. Seja "A"um evento de S composto de "m"eventos simples. A probabilidade de "A", que denotaremos por P(A), é definida por:*

$$P(A) = \frac{m}{N} \tag{3.6}$$

Observamos que assim caracterizada, a probabilidade é uma função a qual defini-se na classe dos eventos ou, o que é equivalente, na classe dos subconjuntos do espaço amostral e satisfaz as propriedades estabelecidas no lema a seguir:

Lema 3.5.1. *Seja S um espaço amostral finito satisfazendo as condições da Definição 4.5.1, a probabilidade definida pela Equação* (1.1)*, satisfaz:*

(i) $0 \leq P(A) \leq 1$;

(ii) $P(A) \geq 0$, *para todo* $A \subset S$;

(iii) *P(S) = 1;*

(iv) *Se "A" e "B", são eventos mutuamente exclusivos, então:* $P(A \cup B) = P(A) + P(B)$

Demonstração (ii): Como $N > 0$ e $m \geq 0$, segue que $P(A) \geq 0$.
Suponha que "A"tem "m_1" eventos simples e que "B"tem "m_2" eventos simples. Como "A"e "B"são mutuamente exclusivos, segue que os mesmos não possuem eventos simples em comum, logo o número de eventos simples de $A \cup B$ é $m_1 + m_2$. Usando a definição obtemos (iv).

Como o número de eventos simples de S é N, segue baseado na definição que $P(S) = 1$.

Exemplo 3.5.1. No experimento que consiste em se lançar um dado, supondo-se que o mesmo é balanceado (honesto), pode-se atribuir probabilidade 1/6, a cada um dos eventos simples: 1, 2, 3, 4, 5 e 6. O evento "O número obtido quando se lança o dado é par"tem probabilidade 0,5, ou seja, 1/2.

Nas situações em que a definição clássica se aplica, para calcular a probabilidade de um evento "A", precisamos contar o número de eventos simples do espaço amostral e de "A". Para facilitar essa tarefa será expresso alguns métodos de contagem.

3.5.1 Definição Freqüentista de Probabilidade

Ao concluirmos que um evento é aleatório, desejamos poder atribuir ao mesmo um número que reflita, suas chances de ocorrência quando o experimento é realizado. Vimos anteriormente que em determinadas circunstâncias podemos atribuir, a mesma chance a todos os eventos simples associados ao experimento. Quando o número de eventos simples do espaço amostral não for finito, esta possibilidade fica descartada.

Uma outra maneira de determinar a probabilidade de um evento consiste em repetir-se o experimento aleatório, digamos "n"vezes, e verificar quantas vezes o evento "A"associado a esse experimento ocorre. Seja $n(A)$, o número de vezes em que o evento "A"ocorreu nas "n"repetições do experimento. A razão

$$f_{n,A} = \frac{n(A)}{n} \tag{3.7}$$

é denotada de freqüência relativa de "A"nas "n", repetições do experimento.

Repetindo-se um experimento um grande número de vezes, nas mesmas condições, e de modo que as repetições sucessivas não dependam dos resultados anteriores, observa-se que a freqüência relativa de ocorrências do evento "A"tende a uma constante "p".

A estabilidade da freqüência relativa, para um grande número de observações, foi inicialmente notada em dados demográficos e em resultados de lançamentos de dados.

Buffon, no século XVIII, realizou 4040 lançamentos de uma moeda e observou a ocorrência de 2048 caras. A freqüência relativa observada desse experimento foi 0,5069. Karl Pearson fez 24000 lançamentos de uma moeda, tendo obtido freqüência relativa de 0,5005 para a face cara.

Os dados seguintes referem-se ao número de nascimentos durante um ano de classificação quanto ao sexo.

Tabela 3.1 *Nascimentos Durante Um Ano*

Meses	Jan.	Fev.	Mar.	Abr.	Mai.	Jun.	Jul.	Ago.	Set.	Out.	Nov.	Dez.
Masc.	3743	3550	4017	4173	4117	3944	3964	3797	3712	3512	3392	3761
Femin.	3537	3407	3866	3711	3775	3665	3621	3596	3491	3391	3160	3371
Total	7280	6957	7883	7884	7892	7609	7585	7393	7203	6903	6552	7132

Para os dados do ano todo, a freqüência relativa de nascimentos do sexo masculino foi:

$$\frac{45682}{88273} = 0,5175.$$

Seja S o espaço amostral associado a um experimento aleatório. Considerando-se "n"repetições desse experimento nas mesmas condições, observemos que a freqüência relativa está definida na classe dos eventos de S e suas propriedades são dadas no lema a seguir:

Lema 3.5.2. *A "Freqüência Relativa $f_{n,\ A}$, definida na classe dos eventos do espaço amostral S satisfaz as seguintes condições:"*

(a) *Para todo evento "A",* $0 \leq f_{n,\ A} \leq 1$;

(b) *Se "A"e "B"são dois eventos de S mutuamente exclusivos, temos:* $f_{n,\ A\ \cup\ B} = f_{n,\ A} + f_{n,\ B}$;

(c) $f_{n,\ S} = 1$.

Demonstração:
A parte a) decorre do fato de que $n(A) \geq 0$. Como os eventos "A"e "B", são mutuamente exclusivos, toda vez que um deles ocorre o outro não ocorre e portanto, o número de ocorrências de $A \cup B$, é igual a soma do número de ocorrências de "A"com o número de ocorrências de "B", isto é: $n(A \ \cup \ B) = n(A) + n(B)$. Dividindo-se por "n"obtemos b). Como em toda realização do experimento algum ponto de S ocorre, segue-se que c) é verdadeira.

Houveram tentativas de se definir a probabilidade como limite da freqüência relativa, já que se observa a maneira como foi mensionada, a freqüência relativa $f_{n,\ A}$, se aproxima de uma constante quando "n"tende a infinito. A definição clássica satisfaz o Lema 1.5.1. São também satisfeitas as condições pela freqüência relativa baseado no Lema 1.5.2 e servem de sustentação intuitiva para a definição axiomática de probabilidade.

3.5.2 Definição Subjetiva de Probabilidade

Toda fundamentação freqüentista de probabilidade está baseada na hipótese de que existe uma realidade física e que as probabilidades descrevem aspectos dessa realidade, de modo análogo aos que as leis da mecânica fazem no caso de um modelo determinístico. A probabilidade de um evento associado a um experimento independe, portanto, do observador, sendo obtida como o valor do qual se aproxima a freqüência relativa de ocorrências desse evento, em um grande número de repetições desse experimento. Há, no entanto, situações em que a repetição do experimento não pode ser realizada e outras em que não pode ser realizada em idênticas condições.

Apresenta-se a seguir alguns exemplos dessas situações:

(a) Deseja-se saber quem vencerá o próximo jogo entre Flamengo e Fluminense;

(b) Um paciente é submetido a um novo tipo de cirurgia e deseja-se saber se ele ficará curado;

(c) Deseja-se saber se haverá um tremor de terra no Rio Grande do Norte no próximo ano.

No primeiro exemplo que se relaciona a um jogo entre dois clubes, sabemos que há estatísticas de um elevado número de jogos entre Flamengo e Fluminense, mas que as condições entre um jogo e outro variam bastante. No exemplo seguinte, não se pode falar em repetição do experimento, pois, trata-se de uma nova técnica cirúrgica que estará sendo aplicada. Com relação ao terceiro exemplo, temos notícias de raras ocorrências de tremores de terra no Rio Grande do Norte.

Uma corrente de probabilistas considera a probabilidade de um evento, como sendo a medida da crença que o observador possui na ocorrência do evento. Dessa forma, a probabilidade será em geral diferente para distintas pessoas, em decorrência das diferentes opiniões que elas têm sobre a ocorrência do evento. Em uma outra descrição equivalente, a probabilidade de um evento é o valor que cada observador estaria propenso a apostar na realização do evento.

3.5.3 Definição Axiomática de Probabilidade

A probabilidade será definida em uma classe de eventos do espaço amostral que satisfaz certas propriedades. Todas as operações a serem definidas entre os eventos conduzem a novos eventos que pertencem a essa classe.

Definição 3.5.2. *Seja uma medida de probabilidade P, uma função definida em uma classe f de eventos de S a qual satisfaz as seguintes condições:*

1. $P(S) = 1$;
2. $P(A) \geq 0$ para todo $A \in f$;
3. Se $A_n, n \geq 1$, são seqüências mutuamente disjuntas em f, então $A_1, A_2, A_3, \cdots, A_n \in f$, logo, chega-se a:
$$P(\bigcup_{n\geq 1} A_n) = \sum_{n\geq 1} P(A_n) \tag{3.8}$$

Definição 3.5.3. *Um espaço de probabilidade é uma terna (S, f, P), onde*

1. *S é um conjunto não vazio;*
2. *f é um subconjunto de S, e*
3. *P é uma medida de probabilidade em f.*

3.5.4 Propriedades da Probabilidade

- (P1) $P(\phi) = 0$;
- (P2) $P(A) = 1 - P(A^c)$;
- (P3) Sendo A e B, dois eventos quaisquer, garante-se que
$$P(B) = P(B \cap A) + P(B \cap A^c);$$
- Se $A \subset B$, então $P(A) \leq P(B)$;
- Regra da Adição de Probabilidade (Generalizável para qualquer "n"):
$$P(A \cup B) = P(A) + P(B) - P(A \cap B);$$
- (P6) Para eventos quaisquer $A_1, A_2, A_3, \cdots, A_n$
$$P(\bigcup_{i=1}^{\infty} A_i) \leq \sum_{i=1}^{\infty} P(A_i);$$
- (P7) Se $A_n \uparrow A$, então $P(A_n) \uparrow P(A)$. De maneira similar, Se $A_n \downarrow A$, então $P(A_n) \downarrow P(A)$.

Demonstração:
(P1) Denotemos por ϕ o evento impossível. Seja A um evento de S (espaço amostral), de probabilidade positiva; Seja ϕ o evento impossível (improvável), podemos exprimir o evento A da seguinte maneira:
$$A = A \bigcup_{i=1}^{\infty} \phi_i,$$
onde para todo $i \geq 1$, $\phi_i = \phi$.

Então baseado na Equação (1.3), segue que:

$$P(A) = P(A) + \sum_{i=1}^{\infty} P(\phi).$$

Subtraindo $P(A)$, em Ambos os membros, segue que a igualdade anterior só faz sentido se $P(\phi) = 0$, logo, a demonstração está completa.

(P2) Os eventos A e A^c formam uma partição de S e, portanto, baseado na Equação (1.3), temos que

$$P(S) = P(A \cup A^c) = P(A) + P(A^c);$$

devido ao fato que $P(S) = 1$, logo

$$1 = P(A) + P(A^c) \Rightarrow P(A) = 1 - P(A^c).$$

(P3) Para dois eventos A e B quaisquer, é sempre possível escrever o evento B da seguinte maneira: $B = (B \cap A) \cup (B \cap A^c)$. Note que essa é uma união disjunta e, portanto, baseado na Equação (1.3), chega-se que o resultado segue de forma imediata, ou seja,

$$P(B) = P(B \cap A) + P(B \cap A^c) - P((B \cap A) \cap (B \cap A^c))$$

$$P(B) = P(B \cap A) + P(B \cap A^c) - 0 \implies P(B) = P(B \cap A) + P(B \cap A^c).$$

(P4) Se $A \subset B$ então o evento B pode ser particionado nos moldes usados em (P3). Assim,

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

Então, como a união é disjunta, vem que

$$P(B) = P(A) + P(B \cap A^c) \geq P(A),$$

uma vez que $P(A) \geq 0, P(B \cap A^c) \geq 0$. Portanto, $\quad P(A) \leq P(B)$.

(P5) Vamos escrever $A \cup B$ como a seguinte união disjunta:

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

Portanto, baseado na Equação (1.3), segue que:

$$P(A \cup B) = P(A \cap B^c) + P(B \cap A^c) + P(A \cap B);$$

Aplicando-se (P3), nos eventos A e B, eles podem ser escritos da seguinte forma:

$$P(A) = P(A \cap B) + P(A \cap B^c)$$

e

$$P(B) = P(B \cap A) + P(B \cap A^c);$$

De modo que obtemos:

$$P(A \cup B) = P(A) - P(A \cap B) + P(B) - P(B \cap A) + P(A \cap B),$$

e o resultado segue, uma vez que a intersecção é comutativa.

(P6) Vamos escrever $\cup_{i=1}^{\infty} A_i$, como a união de uma seqüência disjunta de eventos. Temos

$$\bigcup_{i=1}^{\infty} A_i = A_1 \cup (A_1^c \cap A_2) \cup (A_1^c \cap A_2^c \cap A_3) \cup \cdots$$

Portanto, pela Equação (4.3), podemos escrever

$$P(\bigcup_{i=1}^{\infty} A_i) \leq P(A_1) + P(A_2) + P(A_3) + \cdots$$

visto que,

$$P(\bigcup_{i=1}^{\infty} A_i) = P(A_1) + P(A_1^c \cap A_2) + P(A_1^c \cap A_2^c \cap A_3) + ...;$$

Como para qualquer j,

$$A_1^c \cap A_2^c \cap A_3^c \cap \cdots \cap A_{j-1}^c \cap A_j \subset A_j.$$

Baseado na Equação (1.3), logo

$$P(\bigcup_{i=1}^{\infty} A_i) \leq \sum_{i=1}^{\infty} P(A_i).$$

(P7) Lembramos que a notação $A_n \uparrow A$, indica que temos uma seqüência monótona não decrescente de eventos $A_1, A_2, A_3, \cdots$, tais que, pode-se garantir que

$$\lim_{i \to \infty} A_i = A = \bigcup_{i=1}^{\infty} A_i.$$

Uma vez que $A_1 \subset A_2 \subset A_3 \subset \cdots$, segue que $P(A_i)$ é não decrescente em i pela propriedade (P4). Como também é limitada então $\lim_{n \to \infty} P(A_n)$

existe. Escrevendo o evento A da mesma maneira como foi feito na demonstração de (P6), vem que

$$P(A) = P(\bigcup_{i=1}^{\infty} A_i) = P(A_1) + P(A_1^c \cap A_2) + P(A_1^c \cap A_2^c \cap A_3) + \cdots ,$$

Neste caso, vale $P(A_1^c \cap A_2^c \cap A_3^c \cap \cdots \cap A_{j-1}^c \cap A_j) = P(A_j) - P(A_{j-1}),$

para qualquer j. Assim pela definição de convergência de séries infinitas,

$$P(A) = \lim_{n \longrightarrow \infty} (P(A_1) + [P(A_2) - P(A_1)] + \cdots + [P(A_n) - P(A_{n-1})]),$$

e o resultado segue após simplificação.

Considere agora o caso em que $A_n \downarrow A$. Temos agora uma seqüência monótona não crescente com: $A_1 \supset A_2 \supset A_3 \supset \cdots$, de modo que

$$\lim_{i \longrightarrow \infty} A_i = A = \bigcap_{i=1}^{\infty} A_i,$$

e também $P(A_i)$ é não crescente em i por (P4). Tomando os complementares dos A'_{is}, recaímos no caso anterior de seqüências monótonas não decrescentes e o resultado segue sem dificuldades.

Exemplo 3.5.2. Um dado equilibrado é lançado duas vezes e as faces resultantes são observadas. Um espaço amostral natural seria:

$$S = \{1, 2, 3, 4, 5, 6\} \times \{1, 2, 3, 4, 5, 6\}.$$

A face f pode ser o conjunto das partes e P é a probabilidade uniforme em todos os pontos de S, isto é, $P(\{S\}) = 1/36$. Note que fica mais simples considerar que $S = \{S_1, S_2\}$. Dessa maneira o espaço amostral é constituído de pares de valores, representando os resultados do primeiro e segundo lançamentos, respectivamente. Assim, o espaço amostral completo será:

$$\begin{gathered} S = \{(1,1); (1,2); (1,3); (1,4); (1,5); (1,6); (2,1); (2,2); (2,3); (2,4); (2,5); (2,6); \\ (3,1); (3,2); (3,3); (3,4); (3,5); (3,6); (4,1); (4,2); (4,3); (4,4); (4,5); (4,6); \\ (5,1); (5,2); (5,3); (5,4); (5,5); (5,6); (6,1); (6,2); (6,3); (6,4); (6,5); (6,6)\}. \end{gathered}$$

Considere os eventos:
A: A soma dos resultados é ímpar;
B: O resultado do primeiro lançamento é ímpar;
C: O produto dos resultados é ímpar.
Portanto, as probabilidades associadas a cada um dos três eventos será:

$$P(A) = 18/36; \; P(B) = 18/36 \;\; e \;\; P(C) = 9/36.$$

Para a união de A e B, temos:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) = 18/36 + 18/36 - 9/36 \; = \; 27/36 \; = \; 3/4 \; \Longrightarrow$$

$$P(A \cup B) = 0,75.$$

O cálculo da probabilidade da união de A, B e C, pode ser feito com aplicações sucessivas da regra de adição de probabilidades, que é a propriedade (P5), apresentada anteriormente. Logo,

$$P(A \cup B \cup C) = P[(A \cup B) \cup C] = P(A \cup B) + P(C) - P[(A \cup B) \cap C]$$

$$P(A \cup B \cup C) = P(A) + P(B) - P(A \cap B) + P(C) - [P((A \cap C) \cup (B \cap C))]$$

$$P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - [P(A \cap C) + P(B \cap C) - P(A \cap B \cap C)]$$

$$P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) - P(B \cap C) + P(A \cap B \cap C);$$

Assim, atribuindo o valor das probabilidades aos seus respectivos eventos chega-se a:

$$P(A \cup B \cup C) = 18/36 + 18/36 + 9/36 - 9/36 - 0 - 9/36 + 0 \; = \; 27/36 \; = \; 3/4,$$

assim

$$P(A \cup B \cup C) = 0,75.$$

Teorema 3.5.1. *Sejam os eventos* $A_n,\ n \geq 1$, *se*

$$A_1 \subset A_2 \subset A_3 \subset ...... \subset A_n \qquad e \qquad A = \bigcup_{n=1}^{\infty} A_n.$$

Seja $B_1 = A_1$ *e, para todo* $n \geq 2$;
Seja $B_n = A_n \cap A_{n-1}^c$. *Então, temos que*

$$\lim_{n \longrightarrow \infty} P(A_n) = P(A).$$

Prova: Baseado nas afirmações acima conclui-se que:

(a) Os $B_{ns}^{'}$ são eventos disjuntos (exclusivos);

(b) $A_n = \bigcup_{i=1}^{n} B_i$,;

(c) $A = \bigcup_{n=1}^{\infty} A_n = \bigcup_{n=1}^{\infty} \bigcup_{i=1}^{n} B_i = \bigcup_{i=1}^{\infty} B_i$

logo,

$$P(A_n) = \sum_{i=1}^{n} P(B_i) \ \ e \ \ P(A) = \sum_{i=1}^{\infty} P(B_i).$$

Mas, como

$$\lim_{n \longrightarrow \infty} \sum_{i=1}^{n} P(B_i) = \sum_{i=1}^{\infty} P(B_i),$$

segue que

$$\lim_{n \longrightarrow \infty} P(A_n) = P(A).$$

3.6 Métodos de Enumeração (Análise Combinatória)

3.6.1 Arranjo (Amostras Ordenadas)

Princípio Fundamental da Contagem

Suponha que uma tarefa pode ser executada em duas etapas. Se a primeira etapa pode ser realizada de "n"maneiras e a segunda etapa de "m"maneiras, então a tarefa completa pode ser desenvolvida de $n \times m$ maneiras.

Definição 3.6.1. *Uma amostra de tamanho "n"de um conjunto C que tem N elementos é uma coleção de "n"elementos de C.*

Definição 3.6.2. *Uma amostra é dita ordenada se os seus elementos forem ordenados, isto é, se duas amostras com os mesmos elementos, porém em ordens distintas, forem consideradas diferentes.*

Lema 3.6.1. *O número de amostras ordenadas (sem reposição) de tamanho "n", de um conjunto com N elementos, que será denotado por* $(N)_n$, *é dado por*

$$(N)_n = \frac{N!}{(N-n)!} = N \times (N-1) \times (N-2) \times ...... \times (N-n+1).$$

Exemplo 3.6.1. Considere o conjunto das quatro primeiras letras do alfabeto $\{a, b, c, d\}$. O número de amostras ordenadas sem reposição de tamanho três é igual à $(4)_3 = 4 \times 3 \times 2 = 24$, a seguir lista-se essas amostras.

$$\begin{array}{cccccc}
(a,b,c) & (a,b,d) & (a,c,b) & (a,c,d) & (a,d,c) & (a,d,b) \\
(b,a,c) & (b,a,d) & (b,c,a) & (b,c,d) & (b,d,a) & (b,d,c) \\
(c,a,b) & (c,a,d) & (c,b,a) & (c,b,d) & (c,d,a) & (c,d,b) \\
(d,a,b) & (d,a,c) & (d,b,a) & (d,b,c) & (d,c,a) & (d,c,b)
\end{array}$$

Lema 3.6.2. *O número de amostras ordenadas (com reposição) de tamanho "n", de um conjunto de N elementos é igual a N^n.*

Exemplo 3.6.2. Lança-se "n" vezes uma moeda equilibrada. Determinar a probabilidade de obter pelo menos uma cara.

Lançar uma moeda "n" vezes equivale a selecionar uma amostra com reposição de tamanho "n" de uma população de dois elementos, $\{K, C\}$. Logo, existem 2^n resultados possíveis e igualmente prováveis. Sejam
$A = \{$Obter pelo menos uma Cara$\}$ e
$A_i = \{$Obter Cara no i-ésimo lançamento$\}$.

Então,

$$A = \bigcup_{i=1}^{n} A_i$$

e

$$P(A) = 1 - P(A^c) = 1 - P\left(\left(\bigcup_{i=1}^{n} A_i\right)^c\right) = 1 - P\left(\bigcap_{i=1}^{n} A_i^c\right);$$

Visto que, com base na **Lei de Morgan** garante-se que:

(i) $\left(\bigcup_{i=1}^{n} A_i\right)^c = \bigcap_{i=1}^{n} A_i^c$

(ii) $\left(\bigcap_{i=1}^{n} A_i\right)^c = \bigcup_{i=1}^{n} A_i^c$.

Agora, $\bigcap_{i=1}^{n} A_i^c$ ocorre se todos os "n" lançamentos produzirem coroas. Logo, conclui-se que:

$$P\left(\bigcap_{i=1}^{n} A_i^c\right) = (1/2)^n = \frac{1}{2^n};$$

De modo que

$$P(A) = 1 - \frac{1}{2^n}.$$

Supondo $n = 1$, chega-se que, lançando a moeda uma única vez, a probabilidade de se obter a face cara é:

$$P(A) = 1 - \frac{1}{2^1} = 1 - \frac{1}{2},$$

ou seja,

$$P(A) = \frac{1}{2} = 0,5 = 50\%.$$

Supondo $n = 5$,

$$P(A) = 1 - \frac{1}{2^5} = 1 - \frac{1}{32} = 1 - 0,03 = 0,97.$$

Exemplo 3.6.3. (Problema do Aniversário) Suponha que os aniversários das pessoas ocorram com igual probabilidade entre os 365 dias do ano. Determinar a probabilidade de que em um grupo de "n"pessoas não existam duas com aniversários comuns, ou seja, todas tenham nascido em dias diferentes.

Seja $A = \{$as "n"pessoas nasceram em dias diferentes$\}$. Temos que o número de possíveis datas para as "n"pessoas será 365^n (Amostra ordenada com reposição de tamanho "n"de um conjunto com 365 elementos).

Agora, datas distintas de nascimento das "n"pessoas correspondam a amostras ordenadas sem reposição de tamanho "n"de um conjunto com 365 elementos, cujo número é $(365)_n$. Assim

$$\frac{(365)_n}{(365)^n} = \frac{(365) \times (365-1) \times (365-2) \times (365-3) \times ...... \times (365-n+1)}{365^n}$$
$$= \left(1 - \tfrac{1}{365}\right) \times \left(1 - \tfrac{2}{365}\right) \times \left(1 - \tfrac{3}{365}\right) \times ...... \times \left(1 - \tfrac{n-1}{365}\right)$$

Portanto, caso se queira a probabilidade de pelo menos duas pessoas com aniversários no mesmo dia, deve-se proceder de maneira que:

$$P(A) = 1 - P(A^c) \Longrightarrow P(A^c) = 1 - P(A).$$

A tabela a seguir apresenta alguns valores para a $P(A^c)$ em função de "n".

Tabela 3.2 Número de pessoas com aniversários coincidentes.

Número de Pessoas	$P(A^c)$
10	0,1169
20	0,4114
30	0,7063
40	0,8912
50	0,9704
60	0,9941

3.7 Permutações

Definição 3.7.1. *Uma amostra ordenada sem reposição de tamanho "n", de um conjunto com "n"elementos será denominada uma permutação dos "n"elementos.*

Lema 3.7.1. *O número de permutações de "n"elementos denotado* P_n *é dado por* $P_n = n!$.

Exemplo 3.7.1. Suponha que temos "n"caixas distintas e "n"bolas distintas. Essas bolas são distribuídas ao acaso nas "n"caixas, de modo que, cada caixa contenha exatamente uma bola. Qual a probabilidade de que uma bola específica, digamos uma bola i, esteja em uma caixa específica, digamos caixa j?

O número de maneiras de distribuir as "n"bolas nas "n"caixas, da forma descrita acima, é dado por $n!$. Agora, se a bola i está na caixa j, restam $(n-1)$ bolas para serem distribuídas em $(n-1)$ caixas. Isto pode ser feito de $(n-1)!$ maneiras. Portanto, a probabilidade solicitada é dada por

$$p = \frac{(n-1)!}{n!} = \frac{1}{n}$$

Exemplo 3.7.2. Considere o conjunto dos inteiros de 1 a 3. O número de permutações desse conjunto é $P_3 = 3 \times 2 \times 1 = 6$ e as permutações são as seguintes:

$$123; 132; 213; 231; 312 \ e \ 321$$

Observações:

(i) Se permutamos "n"objetos entre si, a probabilidade de que um objeto específico esteja em uma posição específica é $1/n$;

(ii) Se permutamos "n"objetos entre si, a probabilidade de que k objetos específicos estejam em k posições específicas é $(n-k)!/n!$.

3.8 Combinações (Amostras Não Ordenadas)

Lema 3.8.1. *O número de amostras não ordenadas sem reposição, de tamanho "n", de um conjunto com N elementos, é dado por*

$$C_{N,\ n} = \binom{N}{n} = \frac{(N)_n}{P_n} = \frac{N!}{n!(N-n)!}$$

Exemplo 3.8.1. Suponha que se distribui "n"bolas em "n"caixas. determine a probabilidade de que somente a caixa 1 fique vazia. O número total de maneiras de distribuir é n_n. Seja $A = \{$Apenas a caixa 1 fica vazia $\}$. Queremos obter $P(A)$. Agora, o evento A ocorre $\Longleftrightarrow$ a caixa 1 fica vazia, uma das $(n-1)$ caixas restantes fica com duas bolas e as $(n-2)$ caixas restantes tem exatamente 1 bola. Para $j = 2, 3, 4, ..., n$, seja $B_j = \{$A caixa j tem 2 bolas, a caixa 1 está vazia e as $(n-2)$ caixas restantes tem exatamente uma bola$\}$. Então, os B_j são eventos disjuntos e $A = \cup_{j=2}^{n} B_j$. Portanto, $P(A) = \Sigma_{j=2}^{n} P(B_j)$. Para obter $P(B_j)$, observe que:

$\binom{n}{2}$: Número de maneiras de escolher as duas bolas para a caixa j.

$(n-2)!$: Número de maneiras de distribuir as $(n-2)$ bolas nas $(n-2)$ caixas restantes com uma bola por caixa.

Logo,

$$P(B_j) = \frac{\binom{n}{2}(n-2)!}{n^n}$$

e conseqüentemente,

$$P(A) = \frac{(n-1)\binom{n}{2}(n-2)!}{n^n} = \frac{\binom{n}{2}(n-1)!}{n^n}$$

Exemplo 3.8.2. Uma comissão formada por três estudantes deve ser escolhida, em uma classe de vinte estudantes para organizar os jogos interclasses. De quantas maneiras essa comissão pode ser escolhida?

Como a comissão deve ter três membros distintos, as amostras devem ser selecionadas sem reposição, e como a ordem da escolha dos participantes é irrelevante, trata-se de amostras não ordenadas.

Portanto,

$$C_{N,n} = \frac{(N)_n}{n!} = \binom{N}{n} = \frac{N!}{n!(N-n)!} = \frac{20!}{3!(20-3)!} = \frac{20 \times 19 \times 18 \times 17!}{3! \times 17!} = \frac{6840}{6}$$

$$C_{N,n} = 1.140$$

Assim, essa comissão pode ser escolhida de 1.140 maneiras diferentes.

3.9 Partições

Lema 3.9.1. *O número de partições de um conjunto de N elementos em k subconjuntos, onde verifica-se $n_1, n_2, n_3, n_4, ..., n_k$ elementos, respectivamente, é igual a*

$$\frac{N!}{n_1! \times n_2! \times n_3! \times n_4! \times ...... \times n_k!}$$

Lema 3.9.2. *Suponha que temos um conjunto de "n"objetos, tais que* n_1 *são do tipo 1,* n_2 *são do tipo 2,......,*n_k *são do tipo "k", com* $n_1 + n_2 + n_3 + ... + n_m = m$*. O número de maneiras de se selecionar uma amostra sem reposição de "r"objetos, contendo* k_1 *objetos do tipo 1,* k_2 *objetos do tipo 2,* k_3 *objetos do tipo 3,......,* k_m *objetos do tipo m, com* $k_1 + k_2 + k_3 + k_4 + ... + k_m = r$*, é*

$$\binom{n_1}{k_1}\binom{n_2}{k_2}\binom{n_3}{k_3}......\binom{n_m}{k_m}$$

Exemplo 3.9.1. Em uma mão de 13 cartas de um baralho comum, obtenha a probabilidade de ocorrer exatamente 3 paus, 4 ouros, 4 copas e 2 espadas.
A probabilidade solicitada é:

$$\frac{\binom{13}{3}\binom{13}{4}\binom{13}{4}\binom{13}{2}}{\binom{52}{13}} = \frac{1,1404 \times 10^{10}}{6,3501 \times 10^{11}} = 0,018.$$

3.10 União de eventos

Lema 3.10.1. *Sejam* $A_1, A_2, A_3, ..., A_n$ *eventos de um espaço amostral S. Temos que*

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{n} P(A_i) - \sum_{i=1}^{n}\sum_{j>i}^{n} P(A_i \cap A_j) + \sum_{i=1}^{n}\sum_{j>i}^{n}\sum_{k>j}^{n} P(A_i \cap A_j \cap A_k) + ...$$

$$... + (-1)^{n-1} P(A_1 \cap A_2 \cap ...... \cap A_n).$$

Exemplo 3.10.1. O código morse consiste de uma seqüência de pontos e traços em que repetições são permitidas.

(i) Quantas letras se pode codificar usando exatamente "n"símbolos?
O número de letras será 2^n.

(ii) Qual é o número de letras que se pode codificar usando "n"ou menos símbolos?
$\sum_{i=1}^{n} 2^i = 2^1 + 2^2 + 2^3 + 2^4 + 2^5 + ... + 2^n = \dfrac{2^{n+1} - 2}{2 - 1} = 2\ (2^n - 1)$
Logo, o número de letras que se pode codificar é igual a $2\ (2^n - 1)$.

3.11 Probabilidade condicional

Os conceitos de probabilidade condicional e de independência de eventos são conceitos típicos da teoria das probabilidades, e que servem para distingui-la de outros ramos da matemática. Para a introdução do conceito de probabilidade condicional considera-se uma situação especial, em que o espaço amostral tem eventos equiprováveis.

Considerando o experimento que consiste em lançar um dado duas vezes em uma superfície plana e observar o número de pontos na face superior do dado em cada um dos lançamentos. Supõe-se que não se presencie os lançamentos do dado, mas se receba a seguinte informação: "em cada um dos lançamentos, o número de pontos observados é menor ou igual a dois."Denota-se por "A"esse evento.

Nestas condições, pergunta-se: Qual é a probabilidade de que a soma dos pontos nos dois lançamentos seja igual a quatro? Ou seja, designando-se por “B”o evento “soma dos pontos nos dois lançamentos igual a quatro”, queremos saber qual é a probabilidade de ocorrer o evento “B”, sabendo-se que o evento “A”ocorreu? Para o espaço amostral associado aos dois lançamentos e para os eventos “A”e “B”temos:

$$S = \{(i, j) : 1 \leq i \leq 6;\ 1 \leq j \leq 6\}, \quad \text{onde} : i, j \text{ são inteiros.}$$

$$A = \{(1, 1); (1, 2); (2, 1); (2, 2)\} \quad \text{e} \quad B = \{(1, 3); (2, 2); (3, 1)\}$$

Dizer que o evento “A”ocorreu é equivalente a dizer que pode não se levar em conta, qualquer ponto do espaço amostral que não pertença a “A”, ou seja, pode considerar-se o evento “A”como o novo espaço amostral para o experimento. Dessa maneira a probabilidade de “B”ocorrer dado “A”é igual a $1/4$, pois dos quatro pontos de “A”apenas o ponto $(2, 2) \in$ “B”e os quatro pontos são equiprováveis. Para espaços amostrais com eventos equiprováveis pode-se adotar este procedimento como definição de probabilidade condicional do evento “B”dado o evento “A”. Ele serve, no entanto, de motivação para a seguinte definição:

Definição 3.11.1. *Sejam “A”e “B”dois eventos de um espaço amostral e supondo que $P(A) > 0$, a probabilidade condicional de “B”dado “A”é definida por:*

$$P(B \mid A) = \frac{P(B \cap A)}{P(A)} \tag{3.9}$$

Retornando ao exemplo anterior. Temos $A \cap B = \{(2, 2)\}$ e portanto, $P(A \cap B) = 1/36$, $P(A) = 4/36$. Aplicando-se a Equação (1.4) obtém-se o valor $1/4$, já encontrado anteriormente.

Exemplo 3.11.1. Considere uma caixa contendo "r" bolas vermelhas numeradas de 1 a "r" e "b" bolas pretas numeradas de 1 a "b". Suponha que a probabilidade de extrair uma bola qualquer da caixa é $(b+r)^{-1}$. Sabendo-se que a bola extraída da caixa foi vermelha, qual a probabilidade de que seu número seja 1?

Sejam "A"$=$ {a bola extraída é vermelha} e
"B"$=$ {a bola extraída é a de número 1}.
Assim,

$$P(A) = \frac{r}{b+r} \quad e \quad P(B \cap A) = \frac{1}{b+r}.$$

Logo,

$$P(B \mid A) = \frac{\frac{1}{(b+r)}}{\frac{r}{(b+r)}} = \frac{1}{r}.$$

Exemplo 3.11.2. Considere o lançamento de duas moedas idênticas e "honestas". Calcule a probabilidade condicional de:

(a) Obter duas caras, dado que se obteve cara na primeira moeda;
Sejam "A"$=$ {obteve-se cara na primeira moeda } e "B"$=$ {obteve-se cara na segunda moeda}. Assim queremos calcular:

$$P(A \cap B \mid A) = \frac{P(A \cap B)}{P(A)} = \frac{\frac{1}{4}}{\frac{1}{2}} = \frac{1}{2}.$$

(b) Obter duas caras, dado que se obteve pelo menos uma cara; Neste caso, queremos calcular:

$$P(A \cap B \mid A \cup B) = \frac{P(A \cap B)}{P(A \cup B)} = \frac{\frac{1}{4}}{\frac{3}{4}} = \frac{1}{3}.$$

Baseado na Equação (1.4) que define a probabilidade condicional do evento "B" dado o evento "A", obtemos a seguinte expressão:

$$P(A \cap B) = P(A) \times P(B \mid A) = P(B) \times P(A \mid B) \tag{3.10}$$

Esta expressão e sua generalização para uma interseção de "n" eventos, permitem construir probabilidades em espaços amostrais que representam experimentos realizados em seqüência, em que a ocorrência de um evento na k-ésima etapa depende das ocorrências nas (k-1) etapas anteriores. A Equação (1.5) pode ser generalizada de modo a apresentar a probabilidade da interseção de "n" eventos, $A_1, A_2, A_3, ..., A_n$, por meio das probabilidades condicionais sucessivas.

Exemplo 3.11.3. Considere uma urna com três bolas brancas e sete vermelhas. Duas bolas são retiradas da urna, uma após a outra sem reposição. Determinar o espaço amostral e as probabilidades associadas a cada ponto amostral.

O espaço amostral é o conjunto $\{B_1B_2; B_1V_2; V_1B_2; V_1V_2\}$.

O evento (B_1B_2) é o evento que corresponde a ocorrer branca na primeira retirada e branca na segunda retirada. Para os outros pontos do espaço amostral a interpretação é análoga.

Utilizando a Equação (1.5) temos:

$$P(B_1 \cap B_2) = P(B_2 \mid B_1)P(B_1) = \frac{2}{9} \times \frac{3}{10} = \frac{2}{30}$$

$$P(B_1 \cap V_2) = P(V_2 \mid B_1)P(B_1) = \frac{7}{9} \times \frac{3}{10} = \frac{7}{30}$$

$$P(V_1 \cap B_2) = P(B_2 \mid V_1)P(V_1) = \frac{3}{9} \times \frac{7}{10} = \frac{7}{30}$$

$$P(V_1 \cap V_2) = P(V_2 \mid V_1)P(V_1) = \frac{6}{9} \times \frac{7}{10} = \frac{14}{30}$$

3.11.1 Proposição: Regra do Produto de Probabilidades

Lema 3.11.1. *Sejam $A_1, A_2, A_3, \cdots, A_n$, eventos do espaço amostral S, onde está definida a probabilidade P, tem-se:*

$$P(A_1 \cap A_2 \cap A_3 \cap \cdots \cap A_n) = P(A_1) \times P(A_2 \mid A_1) \times P(A_3 \mid A_1 \cap A_2) \times \cdots$$
$$\cdots \times P(A_n \mid A_1 \cap A_2 \cap A_3 \cap \cdots \cap A_{n-1}) \quad (3.11)$$

Demonstração: Faz-se a demonstração por indução. Para n=2 esta equação reduz-se a Equação (1.5). Suponha que a Equação (1.6) vale para n-1 eventos, isto é,

$$P(A_1 \cap A_2 \cap A_3 \cap \cdots \cap A_{n-1}) = P(A_1) \times P(A_2 \mid A_1) \times P(A_3 \mid A_1 \cap A_2) \times \cdots$$
$$\cdots \times P(A_{n-1} \mid A_1 \cap A_2 \cap A_3 \cap \cdots \cap A_{n-2}) \quad (3.12)$$

Aplicando-se a Equação (1.4) aos eventos $A_1 \cap A_2 \cap A_3 \cap \cdots \cap A_{n-1}$ e A_n, temos:

$$P(A_1 \cap A_2 \cap A_3 \cap \cdots \cap A_{n-1} \cap A_n) = P(A_1 \cap A_2 \cap A_3 \cap \cdots \cap A_{n-1}) \times P(A_n \mid A_1 \cap A_2 \cap A_3 \cap \cdots \cap A_{n-1})$$

Substituindo nesta igualdade a expressão de $P(A_1 \cap A_2 \cap A_3 \cap \cdots \cap A_{n-1})$ dada pela Equação (1.7) obtemos a Equação(1.6).

Exemplo 3.11.4. Tomando como referência o Exemplo 1.11.3 e calculando-se a probabilidade de obter o seguinte resultado: $(B_1B_2V_3V_4B_5)$ em cinco retiradas de bolas da urna, sem reposição. O índice representa o número da retirada. Pela Equação (1.6) temos:

$$P(B_1 \cap B_2 \cap V_3 \cap V_4 \cap B_5) = P(B_1) \times P(B_2 \mid B_1) \times P(V_3 \mid B_1 \cap B_2) \times P(V_4 \mid B_1 \cap B_2 \cap V_3) \times$$
$$\times P(B_5 \mid B_1 \cap B_2 \cap V_3 \cap V_4) = \frac{3}{10} \times \frac{2}{9} \times \frac{7}{8} \times \frac{6}{7} \times \frac{1}{6} = \frac{1}{120}$$

As probabilidades condicionais sucessivas foram calculadas levando-se em conta, as mudanças na composição da urna após cada retirada. Assim, após a primeira retirada, na qual saiu uma bola branca, a urna fica com duas bolas brancas e sete vermelhas; Após a segunda retirada na qual saiu novamente uma bola branca, a urna fica com uma bola branca e sete vermelhas; Após a terceira retirada em que saiu uma bola vermelha, a urna apresenta-se com uma bola branca e seis vermelhas e assim por diante.

3.11.2 Fórmula das Probabilidades Totais e Fórmula de Bayes.

Lema 3.11.2. *Seja $B_1, B_2, B_3, ..., B_n$ uma partição do espaço amostral S, isto é, esses eventos são mutuamente exclusivos e sua reunião é S; Seja "A"uma evento e P uma medida de probabilidade definida nos eventos de S, temos:*

$$P(A) = \sum_{i=1}^{n} P(A \mid B_i) \times P(B_i) \tag{3.13}$$

Demonstração:
Como $S = \bigcup_{i=1}^{n} B_i$, temos que:

$$A = A \cap S = A \cap (\cup_{i=1}^{n} B_i) = \bigcup_{i=1}^{n} A \cap B_i$$

Calculando-se a probabilidade de "A"obtemos:

$$P(A) = \sum_{i=1}^{n} P(A \cap B_i) = \sum_{i=1}^{n} P(A \mid B_i) \times P(B_i)$$

Sendo que, esta última igualdade foi obtida através da $P(A \cap B_i)$, pela Equação (1.5).

A Equação (1.8) é conhecida como a fórmula das probabilidades totais. Note que, ela permite calcular a probabilidade de um evento "A"quando se conhece as probabilidades de um conjunto de eventos disjuntos, cuja reunião é o espaço amostral (uma partição de S) e as probabilidades condicionais de "A"dado cada um deles.

Exemplo 3.11.5. São dadas três Urnas com as seguintes composições: A Urna 1 tem três bolas brancas e cinco vermelhas; A Urna 2 tem quatro bolas brancas e duas vermelhas e a Urna 3 tem uma bola branca e três vermelhas. Escolhe-se uma das três Urnas de acordo com as seguintes probabilidades: Urna 1 com probabilidade 2/6; Urna 2 com probabilidade 3/6 e Urna 3 com probabilidade 1/6. Uma bola é retirada da Urna selecionada. Calcule a probabilidade da bola escolhida ser branca.

Designemos por A = {Retirar uma bola branca da Urna selecionada}, e por B_i, para $1 \leq i \leq 3$, o evento "A Urna i é selecionada". Temos, para n=3,

$$P(A) = P(A \mid B_1) \times P(B_1) + P(A \mid B_2) \times P(B_2) + P(A \mid B_3) \times P(B_3).$$

Os dados do exemplo nos fornecem:

$$P(B_1) = \frac{2}{6}, \quad P(B_2) = \frac{3}{6}, \quad P(B_3) = \frac{1}{6}, \quad P(A \mid B_1) = \frac{3}{8}, \quad P(A \mid B_2) = \frac{4}{6} \quad e \quad P(A \mid B_3) = \frac{1}{4}$$

substituindo estes valores na expressão a seguir chega-se que:

$$P(A) = \frac{3}{8} \times \frac{2}{6} + \frac{4}{6} \times \frac{3}{6} + \frac{1}{4} \times \frac{1}{6} \quad \Longrightarrow \quad P(A) = \frac{1}{2}.$$

A probabilidade condicional, definida na classe dos eventos do espaço amostral, satisfaz as propriedades estabelecidas no seguinte lema.

Lema 3.11.3. *Seja "A"um evento tal que $P(A) > 0$. A probabilidade condicional satisfaz:*

(i) *Para todo evento "B", temos* $P(B \mid A) \geq 0$

(ii) *Se* $B_1, B_2, B_3, ..., B_n$ *são eventos mutuamente exclusivos, então*

$$P(\cup_{i=1}^{n} B_i \mid A) = \sum_{i=1}^{n} P(B_i \mid A)$$

(iii) *Se* S *denota o espaço amostral, então* $P(S \mid S) = 1$.

Demonstração:

(i) Decorre imediatamente da definição de probabilidade condicional e do fato da probabilidade de um evento ser sempre não negativa.

(ii) Decorre da definição de probabilidade condicional e da aditividade da probabilidade. De fato, sejam $B_1, B_2, B_3, ..., B_n$ eventos mutuamente exclusivos,

$$P(\cup_{i=1}^{n} B_i \mid A) = \frac{P((\cup_{i=1}^{n} B_i) \cap A)}{P(A)} = \frac{P(\cup_{i=1}^{n} (B_i \cap A))}{P(A)} = \frac{\sum_{i=1}^{n} P(B_i \cap A)}{P(A)} =$$

$$= \frac{\sum_{i=1}^{n} P(B_i \mid A) \times P(A)}{P(A)}$$

(iii) A demonstração é imediata, pois, $P(S \mid S) = \frac{P(S)}{P(S)} = 1$

Lema 3.11.4. *Seja "B"um evento e* A_1 *e* A_2 *uma partição do espaço amostral S, isto é,* $A_1 \cap A_2 = \phi$ *e* $A_1 \cup A_2 = S$. *Seja P uma probabilidade definida nos eventos de S. Temos para* $i = 1, 2, 3, ..., n$

$$P(A_i \mid B) = \frac{P(B \mid A_i) \times P(A_i)}{\sum_{k=1}^{n} P(B \mid A_k) \times P(A_k)} =$$

$$= \frac{P(B \mid A_i) \times P(A_i)}{P(B \mid A_1) \times P(A_1) + P(B \mid A_2) \times P(A_2) + ... + P(B \mid A_k) \times P(A_k)} \tag{3.14}$$

Fórmula de Bayes

Supondo $i = 1, 2$; Chega-se que:

$$P(A_i \mid B) = \frac{P(B \mid A_i) \times P(A_i)}{P(B \mid A_1) \times P(A_1) + P(B \mid A_2) \times P(A_2)}.$$

Demonstração:
Com base na definição de probabilidade condicional garante-se que:

$$P(A_i \mid B) = \frac{P(A_i \cap B)}{P(B)},$$

Porém, $P(A_i \cap B) = P(B \mid A_i) \times P(A_i)$ e expressando $P(B)$ através da fórmula das probabilidades totais dada pela Equação (1.8) obtém-se a Equação (1.9). Ressaltando algumas observações sobre a fórmula de Bayes chega-se que:

(i) Ela permanece válida quando se considera uma partição finita do espaço amostral S, ou seja, se $A_1, A_2, A_3, ..., A_n$ é uma partição de S, e "B"é um evento de S, então, para todo $i = 1, 2, 3, ..., n$, tem-se:

$$P(A_i \mid B) = \frac{P(A_i) \times P(B \mid A_i)}{\sum_{k=1}^{n} P(A_k) \times P(B \mid A_k)} \tag{3.15}$$

(ii) Como $A_1, A_2, A_3, ..., A_n$ é uma partição do espaço amostral S, isto é, $\cup_{i=1}^{n} A_i = S$ e ainda $A_i \cap A_j = \phi$, segue-se que $P(A_1) + P(A_2) + P(A_3) + ... + P(A_n) = 1$. Portanto, pode-se verificar facilmente que $\Sigma_{i=1}^{n} P(A_i \mid B) = 1$. De fato, isto decorre das propriedades (ii) e (iii) do lema 1.11.3.

Exemplo 3.11.6. Uma companhia monta rádios cujas peças são produzidas em fábricas denominadas $A_1,\ A_2\ e A_3$. Elas produzem, respectivamente, 15%, 35% *e* 50% do total. As probabilidades das fábricas $A_1,\ A_2\ e\ A_3$ produzirem peças defeituosas são respectivamente $0,01;\ 0,05\ e\ 0,02$; Respectivamente. Uma peça é escolhida ao acaso do conjunto das peças produzidas. Essa peça é testada e verifica-se que é defeituosa. Qual é a probabilidade que tenha sido produzida pela *i-ésima* fábrica, para $i = 1, 2, 3$?

Sejam $A_i = \{$ A peça selecionada foi produzida pela fábrica A_i $\}$, $\quad i = 1, 2, 3$ e

D = { A peça selecionada é defeituosa }.

Sabemos que $P(A_1) = 0,15;\ P(A_2) = 0,35;\ P(A_3) = 0,50;\ P(D \mid A_1) = 0,01;\ P(D \mid A_2) = 0,05\ e\ P(D \mid A_3) = 0,02$. Desejamos calcular $P(A_i \mid D)$, para $i = 1, 2, 3$.

Baseado na fórmula de Bayes temos, para $i = 1, 2, 3$,

$$P(A_i \mid D) = \frac{P(D \mid A_i) \times P(A_i)}{P(D \mid A_1) \times P(A_1) + P(D \mid A_2) \times P(A_2) + P(D \mid A_3) \times P(A_3)}$$

com:

$P(D \mid A_1) \times P(A_1) + P(D \mid A_2) \times P(A_2) + P(D \mid A_3) \times P(A_3) =$

$= 0,01 \times 0,15 + 0,05 \times 0,35 + 0,02 \times 0,5 = 0,0015 + 0,0175 + 0,01 = 0,029$. Logo,

$$P(A_i \mid D) = \frac{P(D \mid A_i) \times P(A_i)}{0,029};$$

Para $i = 1$,

$$P(A_1 \mid D) = \frac{P(D \mid A_1) \times P(A_1)}{0,029} = \frac{0,01 \times 0,15}{0,029} = \frac{0,0015}{0,029} \cong 0,052;$$

Para $i = 2$,

$$P(A_2 \mid D) = \frac{P(D \mid A_2) \times P(A_2)}{0,029} = \frac{0,05 \times 0,35}{0,029} = \frac{0,0175}{0,029} \cong 0,6;$$

Para $i = 3$,

$$P(A_3 \mid D) = \frac{P(D \mid A_3) \times P(A_3)}{0,029} = \frac{0,02 \times 0,05}{0,029} = \frac{0,01}{0,029} \cong 0,345$$

Exemplo 3.11.7. O Doente Sadio e o Sadio Doente (Degroot e Schervich [02]).
Uma das formas de avaliar a eficiência de um teste para detectar uma doença é quantificar a probabilidade de erro. Em geral, testes sofisticados envolvem vários procedimentos laboratoriais e diversos equipamentos. Denominamos "*falso positivo*" ao erro em que o teste indica positivo para um paciente que não tem a doença. Por outro lado, teremos um erro "*falso negativo*" se o teste não acusar a doença em um paciente doente. Não é difícil imaginar os inconvenientes dessas ocorrências. Os erros originam "*Doentes*" sadios e "*Sadios*" doentes, isto é, pessoas sadias indicadas como doentes pelo teste e pessoas doentes apontadas como sadias. As probabilidades dos erros são calculadas condicionalmente à situação do paciente. Seus complementares fornecem as probabilidades de acerto do teste, ou seja, a probabilidade de indicar doença para os doentes e não doentes para os sadios.

Considere que um determinado teste resulta positivo para não doentes, com probabilidades 0,1. Também com probabilidade 0,1, o teste será negativo para um paciente doente. As informações fornecidas se referem aos erros que podem ser cometidos ao realizar o teste. Se a incidência da doença na população é de 1 para cada 10 mil habitantes, qual é a probabilidade de uma pessoa estar realmente doente se o teste deu positivo? Observe que, estando ou não doente, existe uma probabilidade não nula de o teste indicar a presença da doença.

A situação acima ocorre de forma similar em várias áreas e é típica para a aplicação do teorema de Bayes. Defini-se os eventos por:

D = {A pessoa está doente};

A = {O teste é positivo}.

Assim, as informações disponíveis são as seguintes:

$P(D) = 0,0001; \quad P(A \mid D^c) = 0,1 \ \ e \ \ P(A^c \mid D) = 0,1.$

Note que, pela propriedade do complementar, temos
$P(D^c = 0,9999)$ e também $P(A \mid D) = 0,9$. Com a notação utilizada, a probabilidade desejada é $P(D \mid A)$ e será calculada através do teorema de Bayes. Portanto,

$$P(D \mid A) = \frac{P(A \mid D) \times P(D)}{P(A \mid D) \times P(D) + P(A \mid D^c) \times P(D^c)} = \frac{0,9 \times 0,0001}{0,9 \times 0,0001 + 0,1 \times 0,9999} = 0,0009.$$

Essa probabilidade é de aproximadamente 1 em 1000. Ela é bastante pequena apesar de ser dez vezes maior que a probabilidade da doença na população. É interessante notar que a probabilidade calculada depende fortemente da eficiência do aparelho.

Passando agora a considerar a eficiência do teste igual à 99%, ou seja, cada um dos erros é igual a 0,01. Assim, as informações disponíveis são as seguintes:

$P(D) = 0,0001; \quad P(A \mid D^c) = 0,01 \ \ e \ \ P(A^c \mid D) = 0,01.$

Note que, pela propriedade do complementar, temos $P(D^c = 0,9999)$ e também $P(A \mid D) = 0,99$. Com a notação utilizada, a probabilidade desejada é $P(D \mid A)$ e será calculada através do teorema de Bayes. Portanto,

$$P(D \mid A) = \frac{P(A \mid D) \times P(D)}{P(A \mid D) \times P(D) + P(A \mid D^c) \times P(D^c)} = \frac{0,99 \times 0,0001}{0,99 \times 0,0001 + 0,01 \times 0,9999} = 0,0098.$$

Assim, percebe-se que essa probabilidade é de aproximadamente 1 em 100. Ela é pequena apesar de ser cem vezes maior que a probabilidade da doença na população e dez vezes maior que a probabilidade obtida com erros de 0,1. Novamente pode-se observar que a probabilidade calculada depende fortemente da eficiência do aparelho, que nesse caso aumentou de 90% para 99%.

3.12 Independência de eventos

Inicia-se a noção de independência para dois eventos e posteriormente estende-se a definição para um número qualquer de eventos.

Definição 3.12.1. *Sejam "A" e "B"dois eventos e suponha que* $P(A) > 0$*. O evento "B"é dito independente do evento "A" se*
$$P(B \mid A) = P(B) \tag{3.16}$$

Esta definição corresponde à noção intuitiva da independência do evento "B"em relação ao evento "A", pois diz que a probabilidade de "B"não se altera com a informação de que o evento "A"ocorreu. Se o evento "B"é independente do evento "A"decorre da Equação (1.5) que:

$$P(A \cap B) = P(A) \times P(B) \tag{3.17}$$

Se o evento "B"é independente do evento "A"então espera-se que "A"também seja independente de "B". De fato isto ocorre, como pode ser verificado a seguir:

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} = \frac{P(A) \times P(B)}{P(B)} = P(A).$$

A primeira igualdade é a definição de probabilidade condicional, e a segunda decorre da Equação (1.12).

Se a Equação (1.12) é válida, então essas igualdades mostram que "A"é independente de "B". Logo, "A"e "B"são independentes se somente se a Equação (1.12) valer. Então podemos adotar essa equação como a definição de independência. Note que esta equação, apesar de não ser tão intuitiva como a expressão da Definição (1.12.1), é simétrica em relação aos dois eventos, e além disso ela é adequada para a generalização da definição de independência para mais de dois eventos.

Exemplo 3.12.1. Seja S o quadrado no plano $0 \leq X \leq 1$ e $0 \leq Y \leq 1$. Considere o espaço uniforme de probabilidade sobre o quadrado, e sejam $A = \{(X, Y) \in S : 0 \leq X \leq 1/2, 0 \leq Y \leq 1\}$ e $B = \{(X, Y) \in S : 0 \leq X \leq 1, 0 \leq Y \leq 1/4\}$. Mostre que "A"e "B"são independentes.

Precisamos mostrar que $P(A \cap B) = P(A) \times P(B)$. temos que
$P(A) = 1/2 \times 1 = 1/2$; $P(B) = 1 \times 1/4 = 1/4$ e $A \cap B = \{(X, Y) \in S : 0 \leq X \leq 1/2,\ 0 \leq Y \leq 1/4\}$.
Logo, $P(A \cap B) = 1/2 \times 1/4 = 1/8 = P(A) \times P(B)$ Para definir a independência para "n"eventos, sendo "n"um inteiro positivo, nós devemos exigir a validade da forma produto para todo subconjunto de "k"dos "n"eventos, para $2 \leq k \leq n$.

Exemplo 3.12.2. Lança-se uma moeda três vezes. O espaço amostral é o conjunto das seqüências de caras e coroas de comprimento 3, ou seja,

$$S = \{KKK, KKC, KCK, KCC, CKK, CKC, CCK, CCC\}.$$

Consideremos os eventos:
$$A = \{CCC, CCK, CKC, CKK\}$$
$$B = \{CKC, CKK, KCC, KCK\}$$
$$C = \{CKK, KCC, KCK, KKC\}.$$

Decorre então que: $P(A) = P(B) = P(C) = 4/8 = 2/4 = 1/2$. Temos

$$P(A \cap B) = \frac{2}{8} = \frac{1}{4} = P(A) \times P(B)$$

$$P(A \cap C) = \frac{1}{8} \neq \frac{1}{4} = P(A) \times P(C)$$

$$P(B \cap C) = \frac{3}{8} \neq \frac{1}{4} = P(B) \times P(C)$$

e

$$P(A \cap B \cap C) = \frac{1}{8} = P(A) \times P(B) \times P(C).$$

As igualdades acima mostram que os eventos "A"e "B"são independentes, e que os eventos "B"e "C"e os eventos "A"e "C", não são independentes e que "A", "B"e "C"satisfazem a forma produto.

Exemplo 3.12.3. Duas moedas são lançadas. Considere os seguintes eventos:

A = {Cara no primeiro lançamento }
B = {Cara no segundo lançamento }
C = {Em dois lançamentos ocorre a mesma face }

logo,

$$A = \{KK, KC\}; \quad B = \{KK, CK\} \quad C = \{KK, CC\}$$

O espaço amostral é o conjunto $S = \{KK, KC, CK, CC\}$, onde K designa cara e C coroa. Supondo que a moeda é balanceada e portanto que os pontos do espaço amostral têm probabilidade 1/4. Portanto,

$$A \cap B = \{KK\}; \quad A \cap C = \{KK\}; \quad B \cap C = \{KK\}; \quad A \cap B \cap C = \{KK\}$$

Calculando-se as probabilidades desses eventos, obtemos:

$$P(A) = P(B) = P(C) = 1/2$$
$$P(A \cap B) = P(A \cap C) = P(B \cap C) = P(A \cap B \cap C) = 1/4.$$

De onde decorre que vale a Equação (1.12) para os eventos tomados dois a dois, mas

$$P(A \cap B \cap C) \neq P(A) \times P(B) \times P(C).$$

Considerando o exemplo onde três moedas equilibradas são lançadas, os eventos:

A = {Cara no primeiro lançamento }
B = {Cara no segundo lançamento }
C = {Cara no terceiro lançamento }

logo,

A = {KKK, KKC, KCK, KCC }
B = {KKK, KKC, CKK, CKC }
C = {KKK, KCK, CKK, CCK }

Verifica-se que para esses eventos, todas as quatro expressões relacionadas as probabilidades: $P(A \cap B)$; $P(A \cap C)$; $P(B \cap C)$ *e* $P(A \cap B \cap C)$, que para esse exemplo se transformam em igualdades, isto é,

1. $P(A \cap B) = P(A) \times P(B)$;

2. $P(A \cap C) = P(A) \times P(C)$;

3. $P(B \cap C) = P(B) \times P(C)$;

4. $P(A \cap B \cap C) = P(A) \times P(B) \times P(C)$

Senão vejamos,

$$
\begin{aligned}
P(A \cap B) &= \frac{1}{4} = P(A) \times P(B) = \frac{1}{2} \times \frac{1}{2} \\
P(A \cap C) &= \frac{1}{4} = P(A) \times P(C) = \frac{1}{2} \times \frac{1}{2} \\
P(B \cap C) &= \frac{1}{4} = P(B) \times P(C) = \frac{1}{2} \times \frac{1}{2} \\
P(A \cap B \cap C) &= \frac{1}{8} = P(A) \times P(B) \times P(C) = \frac{1}{2} \times \frac{1}{2} \times \frac{1}{2};
\end{aligned}
$$

Portanto,

$$P(A \cap B \cap C) = P(A) \times P(B) \times P(C) \tag{3.18}$$

Com base nas afirmações anteriores, concluí-se que, para definirmos independência de três eventos, devemos impor que essas igualdades (1 à 4) sejam verdadeiras. Esses exemplos nos sugerem que para definirmos a independência para "n"eventos, sendo "n"um inteiro positivo, devemos exigir a validade da forma produto para todo subconjunto de "k"dos "n"eventos, para $2 \leq k \leq n$.

Definição 3.12.2. *Os eventos $A_1, A_2, A_3, ..., A_n$, são independentes se*

$$P(A_{i1} \cap A_{i2} \cap A_{i3} \cap ...... \cap A_{ik}) = P(A_{i1}) \times P(A_{i2}) \times P(A_{i3}) \times ...... \times P(A_{ik}),$$

para todo $k = 2, 3, 4, ..., n$ e todo $\{i_1, i_2, i_3, ..., i_k\} \subset \{1, 2, 3, ..., n\}$, tal que,
$i_1 < i_2 < i_3 < ...... < i_k$.

O número de equações que devem ser satisfeitas nesta definição é $2^n - n - 1$. De fato, para todo "k", $2 \leq k \leq n$, têm-se uma equação para cada subconjunto de tamanho "k"dos "n"eventos e como o número desses subconjuntos é $\binom{n}{k}$, temos para cada "k"esse número de equações. O total de equações é, portanto,

$$\sum_{k=2}^{n} \binom{n}{k} = \sum_{k=0}^{n} \binom{n}{k} - \binom{n}{1} - \binom{n}{0} = 2^n - n - 1.$$

3.13 Lista de Exercícios

1) Determine o espaço amostral dos seguintes experimentos aleatórios:

   (a) Lançamento de uma moeda honesta;

   (b) Lançamento de um dado;

   (c) Lançamento de duas moedas;

   (d) Retirada de uma carta de um baralho completo de 52 cartas;

   (e) Determinação da vida útil de um componente eletrônico.

2) Em uma caixa existem cinco bolas brancas e oito azuis. Duas bolas são retiradas simultaneamente da caixa, de forma aleatória. Qual a probabilidade de serem brancas?

3) Escolhem-se ao acaso dois números naturais distintos, de 1 a 20. Qual a probabilidade de que o produto dos números escolhidos seja ímpar?

4) Em quatro lançamentos de uma moedas, qual a probabilidade de ocorrer pelo menos uma cara?

5) Em uma gaveta contém seis pares de meias verdes e dois pares de meias azuis. Tiram-se duas meias ao acaso. Qual a probabilidade de se formar:

   (a) Um par verde?

   (b) Um par de meias da mesma cor?

   (c) um par com meias de cores diferentes?

6) As probabilidades de um aluno da Escola Naval, ser selecionado para a turma A, B e C são respectivamente: 0,2; 0,4 e 0,4. Escolhe-se ao acaso três alunos desta Instituição de ensino. Qual a probabilidade de que pelo menos um seja selecionado para a turma B ou C?

7) Uma sala possui dois soquetes para lâmpadas de caixa com dez lâmpadas, das quais oito estão boas, retiram-se duas lâmpadas ao acaso e colocam-se as mesmas nos soquetes. Qual a probabilidade de que:

   (a) Todas acendam?

   (b) Pelo menos uma lâmpada acenda?

8) Uma caixa contém 10 bolas numeradas de 1 a 10. Extrai-se ao acaso uma bola da caixa. Determine a probabilidade de que o número da bola seja 5, 7 ou 9.

9) Suponha que se lance um dado duas vezes e que os 36 resultados possíveis são igualmente prováveis. Determine a probabilidade de que a soma dos números observados seja par.

10) Qual a probabilidade da soma dos pontos de 2 dados ser 8, se sabemos que os dados tem que mostrar pelo menos 2 pontos cada um?

11) Suponha que no depósito de uma empresa exista uma caixa contendo p fusíveis em perfeito estado de funcionamento e q fusíveis defeituosos. Extrai-se ao acaso um fusível da caixa e a seguir extrai-se, também ao acaso, um segundo fusível dentre os que ficaram na caixa. Determine a probabilidade de que:

    (a) Ambos fusíveis sejam perfeitos para utilização (não apresentam defeito);

    (b) O primeiro fusível seja perfeito para utilização e o segundo seja defeituoso;

    (c) O primeiro fusível seja defeituoso e o segundo seja perfeito para utilização;

    (d) Ambos fusíveis sejam defeituosos.

12) Suponha que num lote com 20 peças existam cinco defeituosas. Escolhe-se quatro peças do lote ao acaso, ou seja, uma amostra de quatro elementos, de modo que a ordem dos elementos seja irrelevante. Qual a probabilidade de duas peças defeituosas estarem contidas nesta amostra?

13) Uma moeda é viciada, isto é, a probabilidade de se obter cara nesta moeda é 0,8. Sorteamos ao acaso uma moeda e a lançamos cinco vezes. Obtemos três caras e duas coroas. Qual a probabilidade de termos escolhido a moeda viciada?

14) Dados os algarismos 1, 2, 3, 4, 5, 6 e 7. Tomando-se dois números dentre os sete possíveis. Qual a probabilidade do produto entre eles ser ímpar?

15) Numa empresa trabalham 10 homens e 5 mulheres. Para se formar uma comissão composta de quatro pessoas é realizado um sorteio. Qual a probabilidade de a comissão ter 2 homens e 2 mulheres?

16) Um colégio é composto de 70% de homens e 30% de mulheres. Sabe-se que 40% dos homens e 60% das mulheres são fumantes. Qual é a probabilidade de que um estudante que foi visto fumando seja homem?

17) Num prédio de 5 andares há 4 apartamentos por andar. Apenas 5 apartamentos estão ocupados. Qual a probabilidade de que cada um dos 5 andares tenha um apartamento ocupado?

18) Duas pessoas A e B praticam tiro ao alvo, a probabilidade do atirador A acertar o alvo é $\frac{1}{3}$ e a probabilidade do atirador B acertar o alvo é $\frac{2}{3}$. Admitindo que A e B são independentes, se os dois atiram, qual a probabilidade de:

    (a) Ambos atingirem o alvo;

    (b) Pelo menos um atingir o alvo.

19) Uma moeda é viciada, de modo que $P(cara) = \dfrac{3}{4}$, a mesma é lançada três vezes. qual a probabilidade de se obter pelo menos duas caras?

20) Suponha que A, B e C sejam eventos tais que: $P(A) = P(B) = P(C) = \dfrac{1}{4}$, $P(A \cap B) = P(C \cap B)$ $= P(A \cap C) = \dfrac{1}{8}$. Calcular a probabilidade de que pelo menos um dos eventos ocorra.

21) Em uma urna há dez bolas, numeradas de 1 a 10. Retira-se uma bola ao acaso. Pede-se determinar a probabilidade de seu número ser par ou maior que 4.

22) Caixa I contém 3 bolas brancas e 2 pretas, caixa II contém 2 bolas brancas e 1 preta e caixa III contém 1 bola branca e 3 pretas.

   (a) Extrai-se uma bola de cada caixa. Determine a probabilidade de que todas as bolas sejam brancas.

   (b) Seleciona-se uma caixa e dela extrai-se uma bola. Determine a probabilidade de que a bola extraída seja branca.

   (c) Calcule em (b) a probabilidade de que a primeira caixa foi selecionada, dado que a bola extraída é branca.

23) Considere dois eventos: A e B, mutuamente exclusivos. Com $P(A) = 0,3$ e $P(B) = 0,5$. Calcule:

   (a) $P(A \cap B)$;

   (b) $P(A \cup B)$;

   (c) $P(A \mid B)$;

   (d) $P(A^c)$;

   (e) $P((A \cup B)^c)$.

24) Um restaurante popular apresenta dois tipos de refeições: salada completa e um prato a base de carne. 20% dos fregueses do sexo masculino preferem salada, e 30% das mulheres preferem carne. 75% dos fregueses são homens. Considere os seguintes eventos:
H: O freguês é homem, M: O freguês é mulher, A: O freguês prefere salada,
B: O freguês prefere carne. Calcule:

   (a) $P(A \mid H)$ e $P(B \mid H)$;

   (b) $P(A \cup H)$ e $P(A \cap H)$;

   (c) $P(M \mid A)$.

Capítulo **4**

Variáveis Aleatórias

4.1 Introdução

Quando se joga uma moeda certo número de vezes estamos fazendo um *Experimento* (representemos por $\mathcal{E}_1$). Pode se estar interessado não na particular sucessão obtida, mas somente no numero de caras. Do mesmo modo, quando se joga uma moeda até obter uma cara ($\mathcal{E}_2$), pode-se estar unicamente interessado no número de jogadas necessárias para obtê-la. Estas são características numéricas associadas aos experimentos.

Se S é um espaço amostral associado ao esperimento $\mathcal{E}_1$, podemos definir várias quantidades de interesse baseadas nos elementos de S. Essas quantidades são essenciais na prática, e facilitam a construção de eventos de interesse.

Definição 4.1.1. *Chamaremos de* ***variável aleatória*** *(v.a.) a toda função real definida sobre S*

As variáveis aleatórias serão definidas usualmente com as letras maiúsculas $X, Y, Z, \cdots$, e os valores que elas assumem são representados por letras minúsculas. Por exemplo, se X é uma variável aleatória com resultados enumeráveis, e $X(S) \in \{x_1, x_2, x_3, \cdots\}$. Os números x_i são os valores da variável aleatória.

Suponha que se atire duas moedas e considera-se o espaço amostral associado a esse experimento. Isto é,

$$S = \{KK, KC, CK, CC\}.$$

Podemos definir uma variável aleatória da seguinte maneira: X é o número de caras (K), obtidas nas duas moedas. Daí,

$$X(KK) = 2, \quad X(KC) = X(CK) = 1, \quad e \quad X(CC) = 0.$$

Portanto, $R_X =$ Valores possíveis de $X = \{0, 1, 2\}$.

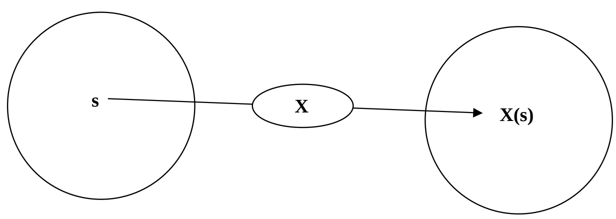


Figura 4.1 Função de Variável Aleatória.

É muito importante compreender uma exigência fundamental de uma função (unívoca): A cada

$s \in S$ corresponderá exatamente um valor $X(s)$. Isto pode ser verificado na Figura 1. No entanto, diferentes valores de s podem levar ao mesmo valor de X. Por exemplo, na Figura 4.1, verifica-se que $X(KC) = X(CK) = 1$.

O espaço R_X, conjunto de todos os valores possíveis de X, é algumas vezes denominado de contra-domínio. De certo modo, pode-se considerar R_X como um outro espaço amostral. O espaço amostral (original) S corresponde ao resultado (possivelmente não numérico) do experimento, enquanto R_X é o espaço amostral associado à variável aleatória X, representando a característica numérica que nos poderá interessar. Se for $X(s) = s$, teremos $S = R_X$. Vale ressaltar que pode-se pensar em uma variável aleatória X, de duas maneiras

(a) Realiza-se um experimento ε que dá um resultado $s \in S$; A seguir calcula-se o valor de $X(s)$.

(b) Realiza-se um experimento ε, obtêm-se o resultado s, e (imediatamente) calcula-se $X(s)$. Neste caso, o número $X(s)$ é pensado como o próprio resultado do experimento e R_X se torna o espaço amostral do experimento.

Exemplo 4.1.1. Três moedas são atiradas sobre a mesa. A partir do momento que as moedas repousem, a fase “aleatória” do experimento terminou. Um resultado simples s poderia consistir na descrição detalhada de como e onde as moedas repousaram. Conseqüentemente se estará interessado somente em certas características numéricas associadas a este experimento. Por exemplo, pode-se avaliar

$$\begin{cases} X(s)\text{: Número de caras que apareceram;} \\ Y(s)\text{: Distância máxima entre duas moedas quaisquer;} \\ Z(s)\text{: Distância mínima das moedas a uma borda qualquer da mesa.} \end{cases}$$

Caso seja a variável X que interesse, pode-se incluir a avaliação de $X(s)$ na descrição deste experimento e, depois. simplesmente afirma-se que o espaço amostral associado ao experimento é $\{0, 1, 2, 3\}$, correspondendo aos valores de X. No entanto, muito freqüêntemente vem-se adotar esta interpretação, compreendendo que a contagem do número de caras é feita “depois” que os aspectos aleatórios do experimento tenham terminado.

Quando se estiver interessado nos eventos associados a um espaço amostral S, verifica-se a necessidade de examinar os eventos relativamente à variável aleatória X, isto é, subespaços do contradominio R_X. Freqüentemente, certos eventos associados a s são “relacionados” (em um sentido a ser explicado) a eventos associados com R_X, na seguinte forma

Definição 4.1.2. *Sejam um experimento $\mathcal{E}$ e seu espaço S. Seja X uma variável aleatória definida em S e seja R_X seu contradomínio. Seja B um evento definido em relação a R_X, isto é, $B \subset R_X$.*
Então, A será definido como

$$A = \{s \in S \mid X(s) \in B\}$$

A será constituído por todos os resultados em S, para os quais $X(s) \in B$ (como pode ser visto na Figura 4.2). Neste caso, pode-se dizer que A e B são eventos equivalentes.

Pode-se dizer que os conjuntos A e B serão equivalentes sempre que ocorram juntos. Isto é, quando A ocorre, B ocorre e vice-versa. Porque se A tiver ocorrido, então um resultado s terá ocorrido, para o qual $X(s) \in B$ e, portanto, B ocorreu. De maneira recíproca, se B ocorreu, um valor $X(s)$ terá sido observado, para o qual $s \in A$ e, portanto, A ocorreu.

Figura 4.2 *Função de Variável Aleatória para Eventos Equivalentes.*

É importante compreender que, em definição de eventos equivalentes, A e B são associados a espaços amostrais diferentes.

Exemplo 4.1.2. Considere a jogada de duas moedas. Daí, $S = \{KK, KC, CK, CC\}$. Seja X o número de caras obtido. Portanto, $R_X = \{0, 1, 2\}$. Seja $B = \{1\}$. Visto que, $X(KC) = X(CK) = 1$ se somente se, $X(s) = 1$, temos que $A = \{KC, CK\}$ é equivalente a B.

Uma importante definição apresenta-se a seguir

Definição 4.1.3. *Seja B um evento no contradomínio Rx. Nesse caso, defini-se P(B) da seguinte maneira* $P(B) = P(A)$, *onde* $A = \{s \in S \mid X(s) \in B\}$. *Defini-se P(B) igual à probabilidade do evento* $A \subset S$, *o qual é equivalente a B, no sentido de* $A = \{s \in S \mid X(s) \in B\}$.

- Admite-se que as probabilidades possam ser associadas a eventos em S. Portanto, a definição anterior torna possível atribuir probabilidades a eventos associados a R_X, em termos de probabilidades definidas sobre S;

- É realmente possível demonstrar que $\mathrm{P}(B)$ deve ser definida tal como foi procedido. Contudo, isto envolveria algumas dificuldades teóricas que deseja-se evitar e, por isso, procede-se como a forma acima;

- Desde que na formulação de $\mathrm{P}(B) = \mathrm{P}(A)$, onde $A = \{s \in S | X(s) \in B\}$, os eventos A e B se referem a espaços amostrais diferentes, se deveria empregar realmente notação diferente, quando se fizesse referência a probabilidades definidas sobre S e aquelas definidas sobre *Rx*, por exemplo, alguma coisa tal como $\mathrm{P}(A)$ e $P_x(B)$. No entanto, não irá se fazer isso, mas continuar-se simplesmente a escrever $\mathrm{P}(A)$ e $\mathrm{P}(B)$.

- Quando se adiciona uma variável aleatória X e seu contradomínio R_X, se adicionando probabilidades nos eventos associados a R_X, as quais serão estritamente determinadas se as probabilidades associadas a eventos em S forem especificadas.

Exemplo 4.1.3. Se as moedas consideradas no Exemplo 2 forem “equilibradas”, teremos $P(KC) = P(CK) = \dfrac{1}{4}$. Portanto, $P(KC, CK) = \dfrac{1}{4} + \dfrac{1}{4} = \dfrac{1}{2}$. (os cálculos anteriores são uma conseqüência direta da suposição fundamental referente à propriedade de equilibrio ou simetria das moedas). Visto que, o evento $\{X = 1\}$ é equivalente ao evento $\{KC, CK\}$, empregado a $A = \{s \in S | X(s) \in B\}$, teremos que $P(X = 1) = P(KC, CK) = \dfrac{1}{2}$. [na verdade não existe escolha para o valor de $P(X = 1)$ coerente com a $\mathrm{P}(B) = \mathrm{P}(A)$, onde $A = \{s \in S \mid X(s) \in B\}$, uma vez que $P(KC, CK)$ tenha sido determinada. É nesse sentido que probabilidades associadas a eventos de R_X são induzidos.]

Uma vez que as probabilidades associadas aos vários resultados (eventos) no contra domínio R_X tenham sido determinadas (mais precisamente, induzidas), ignora-se freqüentemente o espaço amostral original S, que deu origem a essas probabilidades. Assim, no Exemplo 2.1.6, estaremos interessados em $R_X = \{0, 1, 2\}$ e as probabilidades associadas $(\frac{1}{4}, \frac{1}{2}, \frac{1}{4})$. O fato, de que essas probabilidades sejam "determinadas"por uma função de probabilidade definida sobre o espaço amostral original S, não interessa, quando se está apenas interessado em estudar os valores da variável aleatória X.

4.2 Variáveis Aleatórias Discretas

Definição 4.2.1. Seja S o espaço amostral e seja X uma variável aleatória. Dado que o número de valores dos quais a variável aleatória X poderá assumir, ou seja, R_X, o contradomínio, sendo finito ou infinito enumerável de valores $X_1, X_2, X_3, \cdots$, denomina-se X de variável aleatória discreta. Portanto, a variável aleatória X, poderá assumir valores que, podem ser postos em lista como $X_1, X_2, X_3, \cdots, X_n$. Para a situação onde X é finito, e no caso de X ser infinito enumerável, essa seqüência continua de forma indefinida. Logo, dado $\{X_1, X_2, X_3, \cdots\}$ um subconjunto dos números reais $\mathbb{R}$, tal que, $\{S : X(s) = x_i\}$ é um evento para todo i. Então $\{S : X(s) = x_i\}$ é por definição um evento e portanto se pode falar sobre sua probabilidade. Abreviadamente representa-se em geral o evento $\{S : X(s) = x_i\}$ por $\{X = x_i\}$, e a probabilidade desse evento é dada por $P(X = x_i)$, ao invés de $P(\{S : X(s) = x_i\})$.

Seja X uma variável aleatória discreta real. Realmente, se $X_1, X_2, X_3, ...$ são os valores para X assumir, $\{S : X(s) = x_i\}$ é um evento de acordo com a definição de uma variável aleatória discreta real. Se x não é um desses valores, então $\{S : X(s) = x\} = \emptyset$, que também é um evento.

Se os valores possíveis de uma variável aleatória discreta X consistem apenas de números inteiros ou números inteiros não negativos, dizemos que X é uma variável aleatória inteira ou uma variável aleatória inteira não negativa, respectivamente. A maioria das variáveis aleatórias discretas, que ocorrem nas aplicações são inteiras não negativas.

Definição 4.2.2. Seja S o espaço amostral e seja X uma variável aleatória discreta. Todavia, vemos que R_X (o contra-domínio de X) será contituído no máximo por um número infinito enumerável de valores $x_1, x_2, x_3, ...$ para cada possível resultado de x_i associa-se um número $P(x_i) = P(X = x_i)$, denominado de probabilidade de x_i. Os números $P(x_i), i = 1, 2, 3, ...$ devem satisfazer às seguintes condições:

(i) $P(x_i) \geq 0$, *para todo i;*

(ii) $\sum\limits_{i=1}^{\infty} P(x_i) = 1.$

Concluí-se portanto que, chama-se função discreta de densidade X a uma função real $P(x_i)$ definida por $P(x_i) = P(X = x_i)$. Diz-se que um número real x é um valor possível de X se $P(x_i) > 0$.

Exemplo 4.2.1. Seja X a variável aleatória obtida pelo experimento de lançar uma moeda três vezes, onde a probabilidade de obter a face cara em um lançamento individual é p. Logo, o espaço amostral será: $X(s) = \{3, 2, 2, 2, 1, 1, 1, 0\}$, então, para cada um dos $s \in S$, $X(s)$ assume os valores: 3, 2, 1, 0. Dado que,

$$P\{s\} = \{p^3, p^2(1-p), p^2(1-p), p^2(1-p), p(1-p)^2, p(1-p)^2, p(1-p)^2, (1-p)^3\}.$$

Sendo $p = 0,4$. Então, X tem a função de probabilidade discreta $p(x_i)$ dada por

$$p(0) = (1-p)^3 \Rightarrow p(0) = (1-0,4)^3 = 0,216;$$

$$p(1) = 3 \times [p(1-p)^2] = 3 \times 0,4 \times (1-0,4)^2 = 0,432;$$

$$p(2) = 3 \times [p^2(1-p)] = 3 \times (0,4)^2 \times 0,6 = 0,288;$$

$$p^3 = (0,4)^3 = 0,064;$$

Logo,

$$p(0) = 0,216; \quad p(1) = 0,432; \quad p(2) = 0,288; \quad p(3) = 0,064.$$

Dado a condição (ii), ou seja, $\sum_{i=1}^{\infty} p(x_i) = 1$, chega-se que, $\sum_{i=1}^{4} p(x_i) = p(0)+p(1)+p(2)+p(3) = 0,216 + 0,432 + 0,288 + 0,064 = 1$. Portanto, a condição (ii) está satisfeita. Pode-se representar esta função por meio de um diagrama como ilustra a Figura 4.2.1.

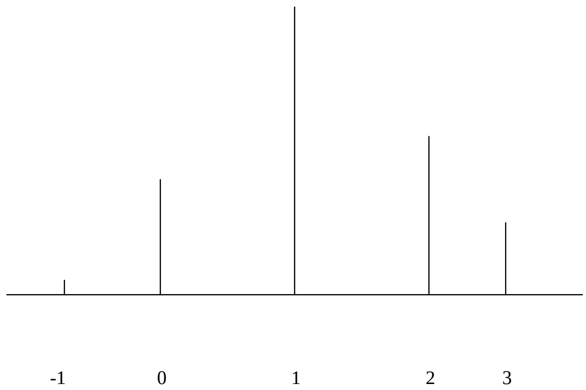


Figura 4.3 Função da Distribuição de $p(x_i)$, onde x_i assume os valores -1, 0, 1, 2 e 3.

4.2.1 Variável Aleatória Constante

Seja "c"um número real. Então a função X definida através de $X(s) = c$, para todo s é uma variável aleatória discreta, visto que, $\{s : X(s) = c\}$ é todo conjunto S e S é um evento. Claramente $P(X = c) = 1$, de modo que a F.P. f de X é simplesmente $f(c) = 1$ e $f(x) = 0$, $x \neq c$. Esta variável aleatória chama-se variável constante. É sob este ponto de vista que uma constante numérica é considerada uma variável aleatória.

4.2.2 Valor Médio de Uma Variável Aleatória Discreta

Definição 4.2.3. Dada a variável aleatória X discreta, assumindo os valores $x_1, x_2, x_3, ..., x_n$, chama-se valor médio ou **esperança** matemática de X ao valor

$$E(X) = \sum_{i=1}^{n} x_i \ P(X = x_i) = \sum_{i=1}^{n} x_i \ P_i \tag{4.1}$$

Se X tem apenas um número finito de valores possíveis $x_1, x_2, x_3, ..., x_n$, então a Equação (4.1) é simplesmente definido como

$$E(X) = \sum_{i=1}^{n} x_i \ P(X = x_i)$$

No caso discreto geral, esta definição é válida desde que a soma $\sum_{i=1}^{n} |x_i| \; P(X = x_i) < \infty$, é exigido, então $\sum_{i=1}^{n} x_i \; P(X = x_i)$ será bem definida. Isso pode levar a conclusão que:

Seja X uma variável aleatória discreta com f.p. $f(x \mid \theta)$. Se $\sum_{i=1}^{n} x_i \; P(X = x_i) < \infty$, dizemos que X tem esperança finita e definimos sua esperança através de

$$E(X) = \sum_{i=1}^{n} x_i \; P(X = x_i)$$

Por outro lado, se $\sum_{i=1}^{n} |x_i| \; P(X = x_i) = \infty$, dizemos que X não tem esperança finita e E(X) é indefinida. Se X é uma variável aleatória não-negativa, geralmente indica-se por $E(X) < \infty$ o fato de que ela tem esperança finita.

Exemplo 4.2.2. Uma questão que se pode ter o interesse em resolver, se refere ao lucro médio por um determinado conjunto fornecido pela Tabela 2.1. Observa-se que 56% das montagens devem produzir um lucro (X) de 15 reais, 23% um lucro (X) de 10 reais, etc.

Tabela 4.1 *Distribuição da V. A. X*

X	P(X)
15	0,56
10	0,23
5	0,02
-5	0,19
Total	1,00

Logo, $E(X) = \sum_{i=1}^{4} x_i \; P(X = x_i) = 15 \times 0,56 + 10 \times 0,23 + 5 \times 0,02 + (-5) \times 0,19 = 9,85$.

Isto é, caso sejam verdadeiras as suposições feitas para determinar a distribuição da Variável Aleatória (v.a.), o empresário espera ter um lucro de 9,85 reais por conjunto montado.

Exemplo 4.2.3. Seja X a V.A. introduzida sob a forma de considerá-la como o resultado do experimento de lançar uma moeda três vezes, no qual a probabilidade de se obter cara em um lançamento indidual é $p = 0,4$. De posse dos dados fornecidos pela Tabela 2.2, pode-se calcular a média de obtenções da face cara no experimento.
Então, $E(X) = \sum_{i=1}^{3} x_i \; P(X = x_i) = 0 \times 0,216 + 1 \times 0,432 + 2 \times 0,288 + 3 \times 0,064 = 1,2$.

Isto é, o número médio de obtenções da face cara, no experimento de três lançamentos dessa moeda (não honesta), sob probabilidade de se obter a face cara igual à 0,4 será 1,2, para experimentos realizados sob essas mesmas condições.

4.2.3 Variância de Uma Variável Aleatória Discreta

A variância de uma variável aleatória é expressa em função do valor médio desta mesma variável, ou seja, de sua esperança matemática, tal como, se observa na Equação (4.2) a seguir,

$$\begin{aligned} Var(X) &= E(X^2) - [E(X)]^2 \\ &= (X - E(X))^2 \end{aligned} \tag{4.2}$$

Tabela 4.2 *Distribuição da Variável Aleatória X*

X	P(X)
0	0,216
1	0,432
2	0,288
3	0,064
Total	1,00

Exemplo 4.2.4. Para o Exemplo 2.2.3 se tem que, a variância da variável aleatória X é dada por:

Tabela 4.3 *Calculo da Variância Variável Aleatória X.*

X	P(X)	X^2	P(X^2)
0	0,216	0	0,216
1	0,432	1	0,432
2	0,288	4	0,288
3	0,064	9	0,064
Total	1,00		1,00

Portanto,
$E(X^2) = \sum_{i=1}^{3} x_i^2 \; P(X^2 = x_i^2) = 0 \times 0,216 + 1 \times 0,432 + 4 \times 0,288 + 9 \times 0,064 = 2,16.$
Consequêntemente, a partir da Equação (4.2) a variância da variável aleatória X é igual à:

$$\begin{aligned} Var(X) &= 2,16 - [1,2]^2 \\ &= 2,16 - 1,44 \\ &= 0,72. \end{aligned}$$

Assim, se pode concluir que a variabilidade associada ao experimento de obtenções da face Cara em três lançamentos de uma moeda é igual à 0,72.

4.2.4 Função de Distribuição Acumulada Discreta

Definição 4.2.4. *Seja X uma variável aleatória, discreta. Defini-se a função "F"como a função de distribuição acumulada da variável aleatória X.*

Teorema 4.2.1. *Se X for uma variável aleatória discreta*

$$F(x) = P(X \leq x) = \sum_j P(x_j),$$

onde o somatório é estendido a todos os índices j que satisfazem à condição $x_j \leq x$.

Observa-se que o domínio de "F"é tido conjunto dos números reais, ao passo que o contradomínio é o intervalo $[0, 1]$.

Exemplo 4.2.5. Voltando ao problema do empresário mostrado anteriormente e usando a função de distribuição acumulada de X definida na tabela a seguir, a função de distribuição aculumada (f.d.a) de X será dada por

$$F(x) = \begin{cases} 0 & se \;\; x \; < \; -5 \\ 0,19 & se \;\; -5 \; \leq \; x \; < \; 5 \\ 0,21 & se \;\; 5 \; \leq \; x \; < \; 10 \\ 0,44 & se \;\; 10 \; \leq \; x \; < \; 15 \\ 1 & se \;\; x \; \geq \; 15. \end{cases}$$

Comentário: Observe que $P(X = x_i)$ é igual ao salto que a função F dá no ponto x_i; por exemplo, $P(X = 10) = 0,23 = F(10) - F(10^-)$. De modo geral, $P(X = x_i) = F(x_i) - F(x_i^-)$, ou seja, $P(X = x_i) = F(x_i) - F(x_i^-)$, onde lembramos que

$$F(a^-) = \lim_{x \to a^-} F(x)$$

Exemplo 4.2.6. Suponha que a variável aleatória X tome os valores 0, 1 e 2, com probabilidades $\frac{1}{3}$, $\frac{1}{6}$ e $\frac{1}{2}$, respectivamente. Então,

$$F(x) = \begin{cases} 0 \;\; se \; x \; < \; 0 \\ \\ \frac{1}{3} \;\; se \; 0 \; \leq \; x \; < \; 1 \\ \\ \frac{1}{2} \;\; se \; 1 \; \leq \; x \; < \; 2 \\ \\ 1 \;\; se \; x \; \geq \; 2 \end{cases}$$

Observe que é de extrema importância o fato de se indicar a inclusão ou a exclusão dos limites, na descrição dos diversos intervalos. Pode se mostrar graficamente a F para o exemplo 5.2.5.

Os gráficos para as funções de distribuição acumulada são, bastante típicos, no seguinte sentido:

(i) Se X for uma variável aleatória discreta, com um número finito de valores possíveis, o gráfico da função de distribuição acumulada será constituído por segmentos de reta horizontais (nesse caso, a função de distribuição acumulada se denomina "Função em Degraus").

(ii) A função de distribuição acumulada de F é definida para todos os valores de x, o que é um motivo importante para considerá-la.

Existem duas outras importantes propriedades da função de distribuição acumulada, que será sintetizada no teorema seguinte

Teorema 4.2.2. *Garante-se que*

(i) *A função F é não-decrescente. Isto é, se $x_1 \leq x_2$, teremos $F(x_1) \leq F(x_2)$.*

(ii) $\lim_{x \to -\infty} F(x) = 0$ *e* $\lim_{x \to \infty} F(x) = 1$. [Freqüentemente, escreve-se isto como $F(-\infty) = 0$ e $F(\infty) = 1$.]

4.3 Variáveis Aleatórias Contínuas

Considerou-se anteriormente as variáveis aleatórias discretas e suas funções de probabilidade como, por exemplo, Bernoulli, Binomial, Geométrica, Hipergeométrica e Poisson. Nas aplicações, estas variáveis aleatórias representam tipicamente o número de objetos de uma certa característica, como o número de bolas vermelhas em uma amostra aleatória de tamanho "n", com ou sem reposição, ou número de chamadas que chegam a uma central telefônica em um minuto.

Existem muitas situações, tanto teóricas quanto aplicadas, em que as variáveis aleatórias naturais a considerar são "contínuas" ao invés de discretas. como aproximação inicial, podemos definir uma variável aleatória contínua X em um espaço de probabilidade S como uma função $X(s)$, $s \in S$, tal que, $P(\{s \mid X(s) = x\}) = 0$, $-\infty < x < +\infty$, isto é, tal que, X assume qualquer valor específico x com probabilidade zero.

É facil pensar em exemplos de variáveis aleatórias contínuas. Considera-se inicialmente, um modelo probabilístico para os tempos de desintegração de um número finito de particulas radioativas. Seja T o tempo que decorre até a desintegração da primeira partícula. Então, T é uma variável aleatória contínua, pois a probabilidade de que a primeira desintegração ocorra em um tempo específico (por exemplo, $T = 2.0000...$ segundos) é zero. Como segundo exemplo, considera-se o experimento de escolher ao acaso um ponto de um subconjunto S do espaço euclidiano *"n-dimensional"* finito não nulo. Seja X a variável aleatória que representa a primeira coordenada do ponto escolhido. Supõem-se por exemplo, que $n = 2$ e que S seja um disco de raio unitário no plano, centrado na origem. Então, o conjunto de pontos em S que tem a primeira coordenada zero é um segmento de reta no plano. Qualquer segmento como este tem área zero e portanto probabilidade zero.

De forma geral, as variáveis aleatórias que representam medidas de grandezas físicas como coordenadas espaciais, peso, tempo, temperatura e voltagem são descritas mais adequadamente como variáveis aleatórias contínuas. Variáveis aleatórias associadas às contagens de objetos ou eventos são exemplos típicos de variáveis aleatórias discretas.

Entretanto, existem casos em que tanto a formulação discreta como à contínua poderiam ser apropriadas. Assim, embora normalmente a medida de comprimento seja considerada como uma variável aleatória contínua, pode-se considerar a medida arredondada para um certo número de casas decimais, e portanto, como sendo uma variável aleatória discreta.

4.3.1 Variáveis Aleatórias e suas Funções de Distribuição

Nas aplicações, uma variável aleatória representa uma quantidade numérica definida em termos do resultado de um experimento aleatório. Matematicamente, entretanto, uma variável aleatória X é uma função real definida em um espaço de probabilidade. Naturalmente deseja-se que $P(X \leq x)$ seja definida para todo número real x. Em outras palavras, se S for um espaço de probabilidade sobre o qual se define X, deseja-se que

$$\{s \mid X(s) \leq x\}.$$

Seja um evento (isto é, um elemento do espaço amostral). Isto implica nas definições a seguir

Definição 4.3.1. *Uma variável aleatória X em um espaço de probabilidade S é uma função real $X(s), s \in S$, tal que $\{s \mid X(s) \leq x\}$ é um evento para $-\infty < x < +\infty$.*

Definição 4.3.2. *A função de distribuição F de uma variável aleatória X é uma função*

$$F(X) = P(X \leq x), \ -\infty < x < +\infty$$

A função de distribuição torna-se útil na determinação de diferentes probabilidades associadas com a variável aleatória X. Um exemplo é:

$$P(a < x \leq b) = F(b) - F(a), \quad a \leq b$$

4.3.2 Propriedades de Funções de Distribuição Contínuas

Nem todas as funções ocorrem como funções de distribuição, pois as funções de distribuição devem satisfazer certas condições.

Seja X uma variável aleatória e seja F uma função de distribuição. Então

- (i) $0 \leq F(x) \leq 1$ para todo x.
- (ii) F é uma função não decrescente de x.

  A propriedade (i) decorre imediatamente da propriedade de definição $F(x) = P(X \leq x)$.

  Para se verifivar a validade de (ii), precisa-se simplesmente observar que se $x < y$, então

  $$F(y) - F(x) = P(x < X \leq y) \geq 0$$

  Diz-se que uma função f tem um limite L à direita (à esquerda) no ponto x se $f(x+h) \to L$ para $h \to 0$, através de valores positivos (negativos). Quando existem, representam-se os limites à direita e à esquerda por $f(x^+)$ e $f(x^-)$ existem para todo x. Sob as mesmas condições, f tem os limites $f(-\infty)$ para $x \to -\infty$ e $f(+\infty)$ para $x \to +\infty$.

  Das propriedades (i) e (ii), segue-se que a função de distribuição F tem os limites $F(-\infty)$ e $F(+\infty)$.

- (iii) $F(-\infty) = 0$ e $F(+\infty) = 1$
- (iv) $F(x^+) = F(x)$ para todo x.

Para avaliar $F(-\infty)$ e $F(+\infty)$ precisa-se apenas obter os limites de $F(n)$ para $n \to -\infty$. (Isso decorre do fato de que F é não-decrescente).

Definição 4.3.3. *Chama-se à variável aleatória X de variável aleatória contínua se*

$$P(X = x) = 0, \quad -\infty < x < \infty.$$

Observa-se que X é uma variável aleatória contínua se, sua função de distribuição F for contínua para todo x, isto é, se F for uma função contínua. Portanto, caso X seja uma variável aleatória contínua, então além da Definição 2.10.1 temos que

$$P(a < X < b) = P(a \leq X \leq b) = P(a \leq X < b) = F(b) - F(a),$$

de modo que $<$ e $\leq$ podem ser usados indiscriminadamente neste contexto.

Definição 4.3.4. *Uma função de distribuição é qualquer função F que satisfaz as propriedades* (i)-(iv)*, isto é,*

(i) $0 \leq F(x) \leq 1$ *para todo x,*

(ii) *F é uma função não-decrescente de x,*

(iii) $F(-\infty) = 0$ *e* $F(+\infty) = 1$

(iv) $F(X^+) = F(x)$ *para todo x.*

Em textos mais avançados demonstra-se que se F for uma função de distribuição, necessariamente existe um espaço de probabilidade e uma variável aleatória X definida neste espaço tal que F é a função de distribuição de X.

4.3.3 Densidades de Variáveis Aleatórias Contínuas

Na prática defini-se geralmente as funções de distribuição em termos de funções de densidade.

Definição 4.3.5. Uma função de densidade (em relação a integração) é uma função não-negativa f tal que

$$\int_{-\infty}^{+\infty} f(x)\ dx = 1 \tag{4.3}$$

observe que se f for uma função de densidade, então a função F definida por

$$F(x) = \int_{-\infty}^{x} f(y)\ dy, \quad -\infty < x < \infty, \tag{4.4}$$

é uma função contínua que satisfaz as propriedades (i)-(iv). Assim, a Equação (4.3) define uma função de distribuição contínua. Dizemos que esta função de distribuição tem densidade f. É possível, embora difícil, construir exemplos de funções de distribuições contínuas que não possuem densidades. As funções de distribuições contínuas que realmente possuem densidades são chamadas de funções de distribuição "absolutamente contínuas".

4.3.4 Valor Médio de Uma Variável Aleatória Contínua

Definição 4.3.6. *Como a função $f(x)$ é senpre não-negativa, pode-se escrever a esperança como:*

$$E(X) = \int_{-\infty}^{+\infty} X\ f(x)\ dx. \tag{4.5}$$

4.3.5 Variância de Uma Variável Aleatória Contínua

A extensão do conceito de variância para variáveis aleatórias contínuas é feita de maneira semelhante e equivalente ao caso de uma variável aleatória discreta, ou seja,

$$\begin{aligned} Var(X) &= E(X^2) - [E(X)]^2 \\ &= (X - E(X))^2 \\ &= \int_{-\infty}^{+\infty} (X - E(X))^2 \, f(x) \, dx. \end{aligned} \tag{4.6}$$

Exemplo 4.3.1. O ponteiro dos segundos de um relógio mecânico pode parar a qualquer momento (instante), devido a algum defeito técnico, ou término da bateria, e indica-se por X o ângulo que esse ponteiro forma com o eixo imaginário passando pelo centro do mostrador e pelo número 12, como é apresentado na Tabela 2.10.

Tabela 4.4 *Deslocamento dos Ponteiros de um Relógio*

X	0^o	6^o	12^o	18^o	$\cdots$	342^o	348^o	354^o	360^o
$P(x)$	$\frac{1}{60}$	$\frac{1}{60}$	$\frac{1}{60}$	$\frac{1}{60}$	$\cdots$	$\frac{1}{60}$	$\frac{1}{60}$	$\frac{1}{60}$	$\frac{1}{60}$

Pergunta-se, qual o ângulo médio formado por esse ponteiro, em relação ao eixo imaginário que passa pelo centro do mostrador e pelo número 12.
Sabe-se que

$$E(X) = \int_{-\infty}^{+\infty} X \, f(x) \, dx,$$

então

$$\begin{aligned} E(X) &= \int_0^{360} X \, \frac{1}{360} \, dx \;=\; \frac{1}{360} \int_0^{360} X \, dx \;=\; \frac{1}{360} \;\times\; \frac{X^2}{2}\Big|_0^{360} \\ &= \frac{1}{360} \times [(360)^2 - (0)^2] \times \frac{1}{2} \;=\; \frac{1}{360} \times [(360)^2] \times \frac{1}{2} \\ &= \frac{1}{360} \times \frac{(360)^2}{2} \;=\; \frac{360}{2} \\ &= 180. \end{aligned}$$

Calcula-se conseqüentemente a variância para a variável aleatória X, ou seja,

$$Var(X) = \int_{-\infty}^{+\infty} (X - E(X))^2 \, f(x) \, dx$$

logo, chega-se que

$$\begin{aligned} Var(X) &= \int_0^{360} (X - 180)^2 \times \frac{1}{360} \, dx \;=\; \frac{1}{360} \times \left[\int_0^{360} (X^2 - 2 \times X \times 180 + (180)^2) \, dx\right] \\ &= \frac{1}{360} \times \left[\int_0^{360} X^2 \, dx - \int_0^{360} 2 \times X \times 180 \, dx + \int_0^{360} (180)^2 \, dx\right] \end{aligned}$$

$$
\begin{aligned}
&= \frac{1}{360} \times \left[\frac{X^3}{3}\Big|_0^{360} - 360 \times \frac{X^2}{2}\Big|_0^{360} + (180)^2 \times X\Big|_0^{360} \right] \\
&= \frac{1}{360} \times \left[\left(\frac{(360)^3}{3} - \frac{0^3}{3} \right) - (180 \times (360^2 - 0^2)) + \big((180)^2 \times (360 - 0)\big) \right] \\
&= \frac{1}{360} \times \big[15.552 \;-\; 23.328 \;+\; 11.664\big] \times 10^3 \;=\; \frac{1}{360} \times (3880) \times 10^3 \\
&= 10.800
\end{aligned}
$$

4.4 Suporte de um modelo de probabilidade

Definição 4.4.1. *O conjunto* $A(x) = \{x : f(x \mid \theta) > 0\}$ *é denominado o* **suporte** *da variável aleatória* X.

4.5 Espaço paramétrico de um modelo de probabilidade

Definição 4.5.1. *O conjunto* Θ *dos valores possíveis de* θ *é denominado* **Espaço Paramétrico**.

4.6 Parâmetro

Definição 4.6.1. *Um parâmetro é uma medida usada para descrever uma característica da população.*

Tabela 4.5 Símbolos Mais Comuns Para Medidas Populacionais e Amostrais

	Estatística	Parâmetro
Média	$\overline{X}$	μ
Variância	S^2	σ^2
N^o de elementos	n	N
Proporção	$\widehat{p}$	p

Definição 4.6.2. - *Modelos de Probabilidade:* Seja X uma variável aleatória discreta ou contínua, com função de probabilidade (f.p.) ou função densidade de probabilidade (f.d.p.) respectivamente, dadas por $f(x \mid \theta)$. Portanto, tem-se que θ é o parâmetro associado ao modelo de probabilidade da variável aleatória X.

4.7 Modelos de Probabilidade Discretos

4.7.1 Modelo de Probabilidade Uniforme Discreta

Experimentos que tenham n resultados possíveis, todos com a mesma probabilidade $1/n$, *é dito seguir um modelo uniforme discreta. Assim,*

$$P(X = x_i) = \frac{1}{n}, \quad i = 1, 2, \cdots, n.$$

- *O lançamento de um dado honesto tem como resultado as faces 1,2,3,4,5,6, todos com probabilidade* $1/6$*;*

- *O resultado da megasena tem como resultado uma das* $\binom{60}{6}$ *combinações possíveis, todas com a mesma probabilidade.*

- *O ângulo inteiro em uma roleta de um cassino honesto.*

Para esta variável temos que

$$E(X) = \frac{1}{n}\sum_{i=1}^{n} x_i, \quad V(X) = \frac{1}{n}\sum_{i=1}^{n}(x_i - E(X))^2.$$

O principal caso é quando $x_i = i$*, ou seja,* $x_i \in \{1, 2, ..., n\}$*, onde temos*

$$E(X) = \frac{n+1}{n}, \quad V(X) = \frac{n^2-1}{12}.$$

4.7.2 Modelo de Probabilidade Bernoulli

Muitos experimentos são tais que os resultados apresentam ou não uma determinada característica. Por exemplo:

- Uma moeda é lançada: o resultado é cara ou coroa;

- Um dado é lançado: ou ocorre face 5 ou não (ocorrendo então uma das demais faces 1, 2, 3, 4 ou 6);

- Uma peça é escolhida ao acaso dentre 1000: é ou não defeituosa;

- Uma pessoa é escolhida dentre 500: é ou não do sexo masculino;

- Uma pessoa é escolhida ao acaso dentre os moradores de uma cidade e verifica-se quanto a ela ser ou não favorável a um projeto municipal.

Em todos esses casos, estamos interessados na ocorrência de sucesso (cara; face 5; a peça é defeituosa; a pessoa é do sexo masculino e a pessoa é favorável ao projeto municipal) ou fracasso (coroa; uma face diferente de 5; a peça não é defeituosa; a pessoa é do sexo feminino e a pessoa não é favorável a um projeto municipal). Essa terminologia (sucesso e fracasso) será usada freqüentemente.

Para cada experimento citado anteriormente, podemos definir uma variável aleatória X, que assume apenas dois valores: 1, se ocorrer sucesso, e 0, se ocorrer fracasso. Indica-se por θ a probabilidade de sucesso, isto é, $P\ (sucesso) = P\ (S) = \theta, \ \ 0 < \theta < 1.$

$$f(x) = \begin{cases} \theta^x(1-\theta)^{1-x}, & x = 0, 1 \\ \\ 0, & para\ outros\ valores\ de\ x. \end{cases}$$

A variável aleatória X, que assume apenas os valores 0 e 1, é representada por $X \sim Bernoulli(\theta)$ e com função de probabilidade dada por

$$f(x \mid \theta) \ = \ P(X = x \mid \theta) \ = \ \theta^x(1-\theta)^{1-x}, \quad x = 0, 1; \quad 0 < \theta < 1. \tag{4.7}$$

$$E(X) = \theta \qquad Var(X) = \theta(1-\theta).$$

Tal que,

$$P(X = 0) = \theta^0(1-\theta)^{1-0} = (1-\theta)$$
$$P(X = 1) = \theta^1(1-\theta)^{1-1} = \theta,$$

portanto, X é chamada variável aleatória de *Bernoulli(θ)*.

Exemplo 4.7.1. Vamos supor o experimento onde um dado é lançado e a variável X representa a obtenção da face 5, logo os possíveis resultados desse experimento serão: ocorre a face 5 ou ocorre qualquer uma das faces diferentes de 5; supondo o dado perfeito (equilibrado) tem-se

$$P(X = 0) = \theta^0(1-\theta)^{1-0} = (1-\theta) = 1 - \frac{1}{6} = \frac{5}{6},$$

$$P(X = 1) = \theta^1(1-\theta)^{1-1} = \theta = \frac{1}{6},$$

conseqüentemente,

$$\mu = E(X) = \frac{1}{6}$$

visto que, $\mu = E(X) = \theta$,

$$\begin{aligned} Var(X) &= \theta(1-\theta) \\ &= \frac{1}{6} \times \frac{5}{6} = \frac{5}{36} \end{aligned}$$

4.7.3 Modelo de Probabilidade Binomial

Considere "n"repetições independentes de um experimento do tipo sucesso e fracasso (Bernoulli). Seja S_n o número de sucessos em "n"repetições. Então S_n é uma variável aleatória que pode assumir somente os valores 0, 1, 2, 3,..., n. Sabe-se que, para um número inteiro x, $0 \leq x \leq n$, resulta em

$$P(S_n = x) = \binom{n}{x}\theta^x(1-\theta)^{n-x};$$

Portanto, a função de probabilidade f de S_n é dada por

$$f(x) = \begin{cases} \binom{n}{x}\theta^x(1-\theta)^{n-x}, & x = 0, 1, 2, ..., n, \\ \\ 0, & \textit{para outros valores de } x. \end{cases}$$

Definição 4.7.1. *A variável aleatória X, que assume os valores 0, 1, 2, 3,..., n é representada por: $X \sim$ Binomial(n, θ) e com função de probabilidade*

$$f(x \mid \theta) = P(X = x \mid \theta) = \frac{n!}{x!(n-x)!}\,\theta^x(1-\theta)^{n-x}, \quad x = 0, 1, 2, 3, ..., n; \quad 0 < \theta < 1. \tag{4.8}$$

$$E(X) = n\theta \qquad Var(X) = n\theta(1-\theta).$$

Esta função está entre as mais importantes que ocorrem na teoria de probabilidades, função de probabilidade Binomial de parâmetros n *e* θ.

Exemplo 4.7.2. Seja X a variável aleatória introduzida sob a forma à considerá-la como o experimento de lançar uma moeda três vezes, onde a probabilidade de obter a face cara em um lançamento individual é 0,4. Logo, o espaço amostral será: $X(s) = \{3, 2, 2, 2, 1, 1, 1, 0\}$, então, $X = 3, 2, 1, 0$. Este experimento é uma função de distribuição Binomial de parâmetros $n = 3$ e $\theta = 0,4$. Portanto, a probabilidade de se obter uma, duas ou três caras respectivamente nos três lançamentos é

$$P(X = 1) = \frac{3!}{1!(3-1)!}\,(0,4)^1(1-0,4)^{3-1} = \frac{3!}{2!}\,(0,4)^1(0,6)^2 = 0,432$$

$$P(X = 2) = \frac{3!}{2!(3-2)!}\,(0,4)^2(1-0,4)^{3-2} = \frac{3!}{2!}\,(0,4)^2(0,6)^1 = 0,288$$

$$P(X = 3) = \frac{3!}{3!(3-3)!}\,(0,4)^3(1-0,4)^{3-3} = \frac{3!}{3!}\,(0,4)^3(0,6)^0 = 0,064$$

conseqüentemente,

$$\mu = E(X) = 3 \times 0,4 = 1,2$$

visto que, $\mu = E(X) = n\theta$,

$$Var(X) = n\theta(1-\theta) = 3 \times 0,4 \times (1-0,4) = 3 \times 0,4 \times 0,6 = 0,72$$

Refere-se freqüentemente a uma variável aleatória X, com f.p. Binomial, dizendo que X tem uma distribuição Binomial (com parâmetros n e θ, quando se deseja ser mais preciso). Usa-se também denominação semelhante para outras variáveis aleatórias que tem designação específica.

A partir do estudo de análise combinatória, verifica-se que a distribuição Binomial ocorre no processo de amostragem aleatória com reposição. No caso de um processo de amostragem aleatória ocorrido sem reposição, caracteriza-se a distribuição Hipergeométrica.

4.7.4 Modelo de Probabilidade Geométrica

Considere "n"repetições independentes de um experimento do tipo sucesso e fracasso. Porém, o único evento que nos interessa é o que representa o sucesso no experimento, logo faz-se "n"ensaios de *Bernoulli* independentes e a distribuição geométrica se caracterizará a partir do momento em que se observar o primeiro sucesso no experimento, quando então o mesmo será finalizado.

$$f(x) = \begin{cases} \theta(1-\theta)^{x-1}, & x = 1, 2, 3, ..., \\ \\ 0, & para\ outros\ valores\ de\ x. \end{cases}$$

Definição 4.7.2. *A variável aleatória X, que assume os valores 1, 2, 3,... é representada por:* $X \sim$ *Geométrica(θ) e com função de probabilidade*

$$f(x \mid \theta) = P(X = x \mid \theta) = \theta(1-\theta)^{x-1}, \quad x = 1, 2, 3, ...; \quad 0 < \theta < 1. \tag{4.9}$$

$$E(X) = \frac{1}{\theta} \qquad Var(X) = \frac{1-\theta}{\theta^2}$$

Exemplo 4.7.3. Supondo um experimento onde cinco moedas são lançadas e observa-se a ocorrência da face cara nas moedas. Admitindo-se que a probabilidade de ocorrência da face cara é igual a 0,4, ou seja, $\theta = 0,4$, logo, para se calcular a probabilidade de obter-se no $1^o, 2^o, 3^o, 4^o$ e 5^o lançamentos respectivamente a face cara deve-se proceder de maneira que

$$P(X = 1) = 0,4(1 - 0,4)^{1-1} = 0,4 \times (0,6)^0 = 0,400$$
$$P(X = 2) = 0,4(1 - 0,4)^{2-1} = 0,4 \times (0,6)^1 = 0,240$$
$$P(X = 3) = 0,4(1 - 0,4)^{3-1} = 0,4 \times (0,6)^2 = 0,144$$
$$P(X = 4) = 0,4(1 - 0,4)^{4-1} = 0,4 \times (0,6)^3 = 0,086$$
$$P(X = 5) = 0,4(1 - 0,4)^{5-1} = 0,4 \times (0,6)^4 = 0,052$$

conseqüentemente,

$$\mu = E(X) = \frac{1}{0,4} = 2,5$$

visto que, $\mu = E(X) = \dfrac{1}{\theta}$,

$$Var(X) = \frac{1-\theta}{\theta^2} = \frac{1-0,4}{(0,4)^2} = \frac{0,6}{0,16} = 3,75$$

4.7.5 Modelo de Probabilidade Hipergeométrica

Considera-se uma população de "r"objetos dos quais "r_1", são de um tipo e "$r_2 = r - r_1$", são de um segundo tipo. Suponha que se extraia desta população uma amostra aleatória sem reposição de tamanho $n \leq r$. Seja X o número de objetos do primeiro tipo na amostra. Então X é uma variável aleatória cujos valores possíveis são: 0, 1, 2,..., n.

Definição 4.7.3. *A variável aleatória X, que assume os valores 0, 1, 2, 3,..., n é representada por: $X \sim$ Hipergeométrica(n, r) e a partir dos resultados de partições do espaço amostral, chega-se a função de probabilidade*

$$f(x \mid r) = P(X = x \mid r) = \frac{\binom{r_1}{x}\binom{r-r_1}{n-x}}{\binom{r}{n}}; \tag{4.10}$$

onde $x = 0, 1, 2, ..., n;\quad r_1 = 0, 1, 2, ..., r;\quad r = 1, 2, ...;\quad n = 1, 2, ..., r.$

$$E(X) = n\,\frac{r_1}{r} \qquad Var(X) = n\,\frac{r_1}{r}\,\frac{r-r_1}{r}\,\frac{r-n}{r-1}.$$

A Equação (4.10) pode ser escrita como

$$\frac{\binom{r_1}{x}\binom{r-r_1}{n-x}}{\binom{r}{n}} = \frac{(r_1)_x(r-r_1)_{n-x}}{x!(n-x)!}\,\frac{n!}{(r)_n} = \binom{n}{x}\frac{(r_1)_x(r-r_1)_{n-x}}{(r)_n}$$

Assim pode-se escrever a f.p. f de X de duas formas

$$f(x) = \begin{cases} \dfrac{\dbinom{r_1}{x}\dbinom{r-r_1}{n-x}}{\dbinom{r}{n}}; & x = 0, 1, 2, ..., n. \\ \\ 0, & para\ outros\ valores\ de\ x. \end{cases}$$

ou ainda,

$$f(x) = \begin{cases} \dbinom{n}{x}\dfrac{(r_1)_x(r-r_1)_{n-x}}{(r)_n}; & x = 0, 1, 2, ..., n. \\ \\ 0, & para\ outros\ valores\ de\ x. \end{cases}$$

Esta função de probabilidade chama-se função de probabilidade Hipergeométrica.

Exemplo 4.7.4. Em problemas de controle da qualidade, lotes com "N" ítens são examinados. O número de ítens com defeito (atributo A), r, é desconhecido. Colhemos uma amostra de "n"ítens e determinamos "k". Suponha que um lote de $N = 100$ peças, $r = 10$ sejam defeituosas; sabe-se que a função de probabilidade Hipergeométrica pode ser escrita como

$$P(X = k) = P_k = \frac{\dbinom{r}{k}\dbinom{N-r}{n-k}}{\dbinom{N}{n}}; \quad onde\ 0 \ \leq \ k \ \leq \ min(r, n)$$

Escolhendo-se $n = 5$, ou seja, retira-se cinco peças sem reposição, a probabilidade de não se obter peças defeituosas é

$$P(X = 0) = P_0 = \frac{\dbinom{10}{0}\dbinom{100-10}{5-0}}{\dbinom{100}{5}} = \frac{\dbinom{90}{5}}{\dbinom{100}{5}} \cong 0,584.$$

Enquanto a probabilidade de se obter pelo menos uma defeituosa é

$$P(X = 1) + P(X = 2) + P(X = 3) + P(X = 4) + P(X = 5) = P_1 + P_2 + P_3 + P_4 + P_5$$

$$\frac{\dbinom{10}{1}\dbinom{90}{4}}{\dbinom{100}{5}} + \frac{\dbinom{10}{2}\dbinom{90}{3}}{\dbinom{100}{5}} + \frac{\dbinom{10}{3}\dbinom{90}{2}}{\dbinom{100}{5}} + \frac{\dbinom{10}{4}\dbinom{90}{1}}{\dbinom{100}{5}} + \frac{\dbinom{10}{5}\dbinom{90}{0}}{\dbinom{100}{5}} \cong 0,416$$

Ou equivalentemente, para obter-se a probabilidade de pelo menos uma peça dentre as cinco escolhidas apresentar defeito calcula-se

$$P(X \geq 1) = 1 - P(X = 0) = 1 - 0,584 = 0,416$$

4.7.6 Modelo de Probabilidade Poisson

Sabe-se que em muitos fenômenos aleatórios que envolvem um processo de contagem, seguem aproximadamente uma distribuição de *Poisson*. Alguns exemplos de tais fenômenos são o número de átomos de uma substância radioativa que se desintegram na unidade de tempo; o número de chamadas que chegam a uma central telefônica em um determinado período de tempo; o número de erros de impressão por página de um livro e o número de colônias de bactérias em um recipiente de *petri untato* com uma suspensão de bactérias.

Definição 4.7.4. *A variável aleatória X, que assume os valores 0, 1, 2, 3,... é representada por:* $X \sim$ *Poisson(*θ*), e com função de probabilidade*

$$f(x \mid \theta) = P(X = x \mid \theta) = \frac{e^{-\theta}\theta^x}{x!}; \quad x = 0, 1, 2, ...; \quad \theta > 0 \tag{4.11}$$

$$E(X) = \theta \qquad Var(X) = \theta$$

Portanto, seja θ um número positivo. A f.p. de *poisson* de parâmetro θ é expressa por

$$f(x) = \begin{cases} \dfrac{\theta^x e^{-\theta}}{x!}; & x = 0, 1, 2, ..., n. \\ \\ 0, & para\ outros\ valores\ de\ x. \end{cases}$$

É evidente que esta função satisfaz as seguintes propriedades: $f(x) \geq 0$, $X \in \mathbb{R}$ e $\{x : f(x) \neq 0\}$ da definição de Variável Aleatória Discreta. A propriedade $\sum_{i=1}^{n} P(x_i) = 1$, segue-se imediatamente da expansão em série de *Taylor* da função exponencial que

$$e^{\theta} = \sum_{x=0}^{\infty} \frac{\theta^x}{x!}$$

Exemplo 4.7.5. Um PBX recebe, em média, cinco chamadas por minuto. Supondo que a distribuição de *poisson* seja adequada nessa situação, para obter-se a probabilidade de que o PBX não receba chamadas durante um intervalo de um minuto. Faz-se $\theta = 5$ e

$$P(X = 0) = \frac{5^0 e^{-5}}{0!} = e^{-5} = 0,0067$$

Por outro lado, se o objetivo for calcular a probabilidade de obter-se no máximo duas chamadas em quatro minutos, tem-se $\theta = 20$ chamadas em quatro minutos, visto que, em média o PBX recebe cinco chamadas por minuto, logo

$$\begin{aligned} P(X \leq 2) &= P(X = 0) + P(X = 1) + P(X = 2) \\ &= \frac{20^0 e^{-20}}{0!} + \frac{20^1 e^{-20}}{1!} + \frac{20^2 e^{-20}}{2!} \\ &= e^{-20}(1 + 20 + 200) = 221e^{-20}, \end{aligned}$$

que é um número muito próximo de zero. Esse exemplo mostra que a probabilidade de "*k*"ocorrências em um intervalo fixo de comprimento "*t*"pode ser escrita como

$$P(X = k) = \frac{e^{-\theta t}(\theta t)^k}{k!}; \quad k = 0, 1, 2, \cdots$$

onde θ representa o número médio de ocorrências naquele intervalo. Denota-se uma variável aleatória X com distribuição de *poisson* de parâmetro θ *por* $X \sim poisson(\theta)$.

4.8 Modelos de Probabilidade Contínuos

4.8.1 Modelo de Probabilidade Uniforme Contínua

Definição 4.8.1. *A variável aleatória* X *tem distribuição uniforme no intervalo* $[a,\ b]$ *se sua função densidade de probabilidade* $(f.d.p.)$ *é dada por*

$$f(x) = \begin{cases} \dfrac{1}{b-a}; & se \ \ a \leq x \leq b, \\ \\ 0, & para\ outros\ valores\ de\ x. \end{cases}$$

O gráfico da f.d.p. é dado pela Figura 5.5 e o Gráfico da f.d.a. é dado pela Figura 5.6. Portanto,

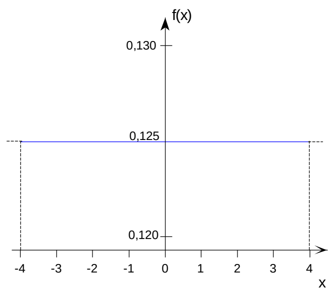


Figura 4.4 *Gráfico da função densidade de probabilidade de uma variável aleatória com distribuição uniforme no intervalo [- 4; 4].*

$$f(x; a, b) = \frac{1}{b-a}\ I_{[a,b]}\ (x),\ \ a \leq x \leq b;\ -\infty < a < b < +\infty. \tag{4.12}$$

$$E(X) = \frac{a+b}{2} \qquad e \qquad Var(X) = \frac{(b-a)^2}{12}.$$

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(x)\ dx = \begin{cases} 0, & se\ x\ <\ a \\ \\ \dfrac{x-a}{b-a}, & se\ \ a\ \leq\ x\ <\ b \\ \\ 1, & se\ x\ \geq\ b \end{cases}$$

Assim, para dois valores quaisquer c e d, onde $c < d$, tem-se portanto

$$P(c < x \leq d) = F(d) - F(c),$$

Notação: usa-se a notação $X \sim U(a, b)$ para indicar que a variável aleatória X tem distribuição uniforme de parâmetros [a, b].

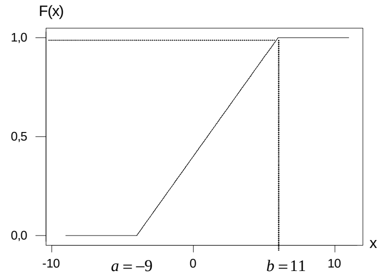


Figura 4.5 *Gráfico da função de distribuição acumulada de uma variável aleatória uniforme no intervalo [- 9; 11].*

Exemplo 4.8.1. Dada uma variável aleatória X, a qual tem distribuição uniforme tem-se que, um caso particular bastante interessante quanto a esta distribuição é quando os parâmetros são iguais à [a = -1/2, b = 1/2]. Portanto, tem-se que

$$f(x) = \begin{cases} 1, \ se \ -1/2 \leq x \leq 1/2 \\ \\ 0, \ caso \ contrario. \end{cases}$$

Verifica-se que,

$$E(X) = \frac{a+b}{2} = \frac{-1/2+1/2}{2} = \frac{0}{2} = 0 \quad e \quad Var(X) = \frac{(b-a)^2}{12} = \frac{(1/2+1/2)^2}{12} = \frac{1}{12}.$$

A f.d.a. será dada por

$$F_X(x) = \begin{cases} 0, \ se \ x < -1/2 \\ \\ x + 1/2, \ se \ -1/2 \leq x < 1/2 \\ \\ 1, \ se \ x > 1/2. \end{cases}$$

Por exemplo, $P(-1/4 \leq X \leq 1/4) = F_X(1/4) - F_X(-1/4) = 1/2$.

4.8.2 Modelo de Probabilidade Normal

Apresenta-se agora um modelo fundamental em Probabilidade e Inferência Estatística. Suas origens remetem-se a Gauss em seus trabalhos sobre erros de observação astronômica, por volta de 1810, de onde se dá a origem do nome distribuição Gaussiana para tal modelo.

Definição 4.8.2. *Diz-se que a variável aleatória X tem "Distribuição Normal" com parâmetros μ e σ^2, onde $-\infty < \mu < +\infty$ e $\sigma > 0$, se sua função densidade de probabilidade é dada por*

$$f(x) = \begin{cases} \dfrac{1}{\sqrt{2\pi\sigma^2}} \, e^{-\dfrac{(x-\mu)^2}{2\sigma^2}}; \quad se \quad -\infty < x < +\infty, \\ \\ 0, \quad caso \ contrario. \end{cases}$$

Portanto,

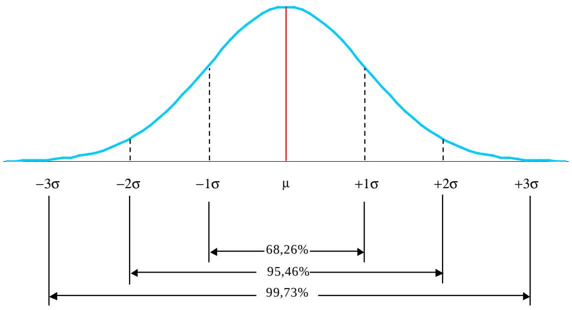


Figura 4.6 *Gráfico da função densidade de probabilidade de uma variável aleatória normal com média e desvio padrão.*

$$f(x;\mu,\sigma^2)=\frac{1}{\sqrt{2\pi\sigma^2}}\,e^{-\frac{(x-\mu)^2}{2\sigma^2}};\qquad -\infty\leq x\leq+\infty,\ -\infty<\mu<+\infty,\ \sigma>0. \tag{4.13}$$

$$E(X)=\mu \quad e \quad Var(X)=\sigma^2.$$

Verifica-se que, $f(x;\mu,\sigma^2)\ \rightarrow\ 0$, quando $x\ \rightarrow\ \pm\infty$, $\mu-\sigma$ e $\mu+\sigma$ são pontos de inflexão de $f(x;\mu,\sigma^2)$, $x=\mu$ é ponto de máximo de $f(x;\mu,\sigma^2)$, e o valor máximo é $\dfrac{1}{\sqrt{2\pi\sigma^2}}$.
A densidade $f(x;\mu,\sigma^2)$, é simétrica em relação à reta $x=\mu$, isto é,

$$f(\mu+x;\mu,\sigma^2),f(\mu-x;\mu,\sigma^2),$$

para todo x real. Com o objetivo de simplificar a notação, denota-se a função densidade de probabilidade normal simplesmente por $f(x)$ e escreve-se simbolicamente,

$$X\sim N(\mu,\sigma^2).$$

Quando $\mu=0$ e $\sigma^2=1$, tem-se uma distribuição normal padrão ou normal reduzida, ou simplesmente $N(0,1)$. Portanto, a função densidade de probabilidade reduz-se a

$$\phi(z)=\frac{1}{\sqrt{2\pi}}\ e^{-\frac{z^2}{2}};\qquad -\infty\leq z\leq+\infty. \tag{4.14}$$

O gráfico da Normal padrão é dado por Se $X\sim N(\mu,\sigma^2)$, então a variável aleatória definida por $Z=\frac{x-\mu}{\sigma}$ terá média 0 e variância 1.
A função de distribuição acumulada $F(y)$ de uma variável aleatória normal X, com média μ e variância σ^2 é obtida integrando-se $f(x;\mu,\sigma^2)=\dfrac{1}{\sqrt{2\pi\sigma^2}}\ e^{-\frac{(x-\mu)^2}{2\sigma^2}}$, de $-\infty$ até y, ou seja,

$$F(y)=\int_{-\infty}^{y}f(x;\mu,\sigma^2)\ dx, y\in\mathbb{R}. \tag{4.15}$$

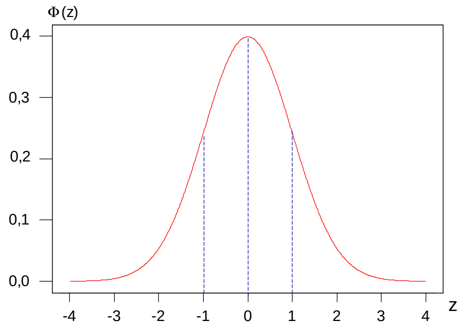


Figura 4.7 *Função de distribuição acumulada da variável aleatória X com distribuição Normal Padrão: $Z \sim N(0,1)$.*

A integral de $F(y)$ corresponde a área, sob $f(x)$ desde $-\infty$ até y. No caso específico da normal padrão, utiliza-se a seguinte notação, a qual é universal

$$\phi(y) = \int_{-\infty}^{y} \phi(z)\, dz = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{y} e^{-\frac{z^2}{2}}\, dz. \tag{4.16}$$

Exemplo 4.8.2. Os depósitos efetuados em um determinado banco durante o mês de janeiro são distribuídos normalmente, com média R\$ 10,00 e variância igual à 2,25. Um depósito é selecionado ao acaso dentre todos os referentes ao mês de janeiro. Encontrar a probabilidade de que o depósito seja

(a) R\$ 10,00 ou menos;

(b) Pelo menos R\$ 10,00;

(c) Um valor entre R\$ 12,00 e R\$ 15,00;

(d) Maior do que R\$ 20,00.

Sabe-se que $\mu = 10$ e $\sigma^2 = 2,25$, conseqüentemente, $\sigma = 1,5$. Sendo X a variável aleatória que representa o referido depósito, então para

(a)

$$P(X \leq 10) = P\left(\frac{x-\mu}{\sigma} \leq \frac{10-10}{1,5}\right) = P(Z \leq 0) = 0,5.$$

(b)

$$P(X \geq 10) = P\left(\frac{x-\mu}{\sigma} \geq \frac{10-10}{1,5}\right) = P(Z \geq 0) = 0,5.$$

(c)

$$\begin{aligned} P(12 < X < 15) &= P\left(\frac{12-10}{1,5} < \frac{x-\mu}{\sigma} < \frac{15-10}{1,5}\right) \\ &= P(1,33 < Z < 3,33) = 0,09133. \end{aligned}$$

(d)

$$P(X > 20) = P\left(\frac{x-\mu}{\sigma} \geq \frac{20-10}{1,5}\right) = P(Z > 6,67) \approx 0.$$

(e)

$$P(0 < X < 5) = P\left(\frac{0-10}{1,5} < \frac{x-\mu}{\sigma} < \frac{5-10}{1,5}\right)$$
$$= P(-6,67 < Z < -3,33) \approx 0.$$

(f)

$$P(5 < X < 10) = P\left(\frac{5-10}{1,5} < \frac{x-\mu}{\sigma} < \frac{10-10}{1,5}\right)$$
$$= P(-3,33 < Z < 0) = 0,5.$$

4.8.3 Modelo de Probabilidade Exponencial

Uma distribuição de probabilidade importantíssima e que tem aplicações em confiabilidade de sistemas, assim como é capaz de mensurar o tempo de vida útil de equipamentos eletrônicos é a distribuição exponencial.

Definição 4.8.3. *A variável aleatória X tem distribuição exponencial com parâmetro $\theta > 0$, onde a função densidade de probabilidade exponencial apresenta-se da seguinte forma*

$$f(x) = \begin{cases} \frac{1}{\theta}\, e^{-\frac{x}{\theta}}; & se \quad x > 0, \\ 0, & x \leq 0. \end{cases}$$

O gráfico da f.d.p. é dad

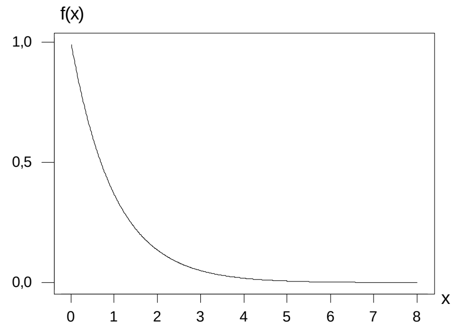


Figura 4.8 *Gráfico da função densidade de probabilidade de uma variável aleatória com distribuição exponencial de parâmetro $\theta = 1$.*

Portanto,

$$f(x;\theta) = \frac{1}{\theta}\, e^{-\frac{x}{\theta}}, \; x > 0; \; \theta > 0. \tag{4.17}$$

$$E(X) = \theta \quad e \quad Var(X) = \theta^2.$$

$$F(x) = P(X \leq x) = \int_0^s f(x)\ dt = \begin{cases} 0, & se\ x\ \leq\ 0 \\ 1 - e^{-\dfrac{x}{\theta}}, & se\ \ 0\ <\ x\ <\ s \\ 1, & se\ x\ \geq\ s. \end{cases}$$

O gráfico da f.d.a. é dado pela Figura (2.10).

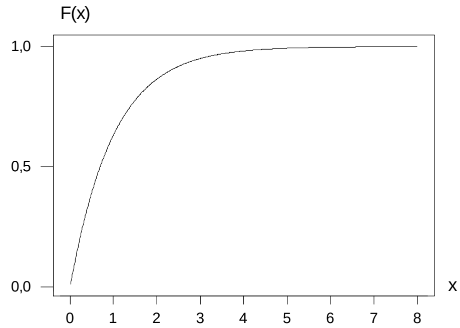


Figura 4.9 *Gráfico da função de distribuição acumulada de uma variável aleatória com distribuição exponencial de parâmetro* $\theta = 1$.

Notação: usa-se a notação $X\ \sim\ Exp\ (\theta)$ para indicar que a variável aleatória X tem distribuição exponencial de parâmetro θ.

Exemplo 4.8.3. O tempo de vida (em horas) de um fusível pode ser pode ser considerado uma variável aleatória com distribuição exponencial de parâmetro $\theta = 500$. Sabe-se que a vida média de um fusível é 500 horas, ou $E(T) = 500$ Horas e conseqüentemente a probabilidade de que este fusível dure mais do que a média é

$$\begin{aligned} P(T > 500) &= \int_{500}^{+\infty} f(t)\ dt\ =\ \int_{500}^{+\infty} \frac{1}{500}\ e^{-\dfrac{t}{500}}\ dt\ =\ \frac{1}{500}\int_{500}^{+\infty} e^{-\dfrac{t}{500}} \\ &= \frac{1}{500}\ \left[(-500) \times e^{-\dfrac{t}{500}}\right]_{500}^{+\infty}\ =\ \left[-\ e^{-\dfrac{t}{500}}\right]_{500}^{+\infty} \\ &= \ -\ e^{-\dfrac{\infty}{500}}\ +\ e^{-\dfrac{500}{500}}\ =\ e^{-1}\ =\ 0,3678. \end{aligned}$$

Conseqüentemente, para $t = 500$ e $\theta = 500$, ou seja, para os dados deste exemplo tem-se que,

$$F(t)\ =\ P(T \leq 500) =\ 1\ -\ e^{-1}\ = 1\ -\ 0,3678\ = 0,6322$$

4.8.4 Modelo de Probabilidade Gama

Definição 4.8.4. *Seja X a variável aleatória que tem distribuição Gama com parâmetros* α *e* β*, onde a função densidade de probabilidade dessa distribuição apresenta-se da seguinte forma*

$$f(x) = \begin{cases} \dfrac{\beta^\alpha}{\Gamma(\alpha)}\, x^{\alpha-1}e^{-\beta x}; \quad se \quad x > 0, \\ \\ 0, \quad x \leq 0 \end{cases}$$

Portanto,

$$f(x \mid \alpha, \beta) = \frac{\beta^\alpha}{\Gamma(\alpha)}\, x^{\alpha-1}e^{-\beta x},\ x > 0,\ \alpha > 0\ e\ \beta > 0 \tag{4.18}$$

$$E(X) = \frac{\alpha}{\beta} \quad e \quad Var(X) = \frac{\alpha}{\beta^2}$$

Notação: usa-se a notação $X \sim Gama\,(\alpha, \beta)$ para indicar que a variável aleatória X tem distribuição gama com parâmetros α *e* β.

4.8.5 Modelo de Probabilidade Qui-quadrado

Definição 4.8.5. *Seja X a variável aleatória que tem distribuição Qui-quadrado com parâmetro n, onde a função densidade de probabilidade dessa distribuição apresenta-se da seguinte forma*

$$f(x) = \begin{cases} \dfrac{(1/2)^{n/2}}{\Gamma(n/2)}\, x^{\frac{n}{2}-1}e^{-\frac{x}{2}}; \quad se \quad x > 0, \\ \\ 0, \quad x \leq 0 \end{cases}$$

Portanto,

$$f(x \mid n) = \frac{(1/2)^{n/2}}{\Gamma(n/2)}\, x^{\frac{n}{2}-1}e^{-\frac{x}{2}},\ x > 0,\ e\ n \geq 1\ (inteiro) \tag{4.19}$$

$$E(X) = n \quad e \quad Var(X) = 2n$$

Notação: usa-se a notação $X \sim \chi^2_n$ para indicar que a variável aleatória X tem distribuição de Qui-quadrado com n graus de liberdade.

4.8.6 Modelo de Probabilidade F-Snedecor

Definição 4.8.6. *Seja X a variável aleatória que tem distribuição F-Snedecor com parâmetro m e n, onde a função densidade de probabilidade dessa distribuição apresenta-se da seguinte forma*

$$f(x) = \begin{cases} \dfrac{\Gamma\left(\dfrac{m+n}{2}\right)}{\Gamma(m/2)\Gamma(n/2)} \left(\dfrac{m}{n}\right)^{m/2} x^{\frac{m}{2}-1}\left(1 + \dfrac{m}{n}\,x\right)^{-\frac{m+n}{2}}; \quad se \quad x > 0, \\ \\ 0, \quad x \leq 0 \end{cases}$$

Portanto,

$$f(x \mid n) = \frac{\Gamma\left(\dfrac{m+n}{2}\right)}{\Gamma(m/2)\Gamma(n/2)} \left(\frac{m}{n}\right)^{m/2} x^{\frac{m}{2}-1}\left(1 + \frac{m}{n}\,x\right)^{-\frac{m+n}{2}},\ x > 0,\ e\ m,n \geq 1\ (inteiros) \tag{4.20}$$

$$E(X) = \frac{n}{n-2} \quad e \quad Var(X) = \frac{2n^2(m+n-2)}{m(n-2)^2(n-4)}$$

Notação: usa-se a notação $X \sim F_{m,n}$ para indicar que a variável aleatória X tem distribuição de F-Snedecor com m e n graus de liberdade.

4.8.7 Modelo de Probabilidade t-Student

Definição 4.8.7. *Seja X a variável aleatória que tem distribuição t-Student com parâmetro n, onde a função densidade de probabilidade dessa distribuição apresenta-se da seguinte forma*

$$f(x) = \begin{cases} \dfrac{\Gamma\left(\dfrac{n+1}{2}\right)}{\Gamma(n/2)} \dfrac{1}{\sqrt{n\pi}} \left(1 + \dfrac{x^2}{n}\right)^{-\frac{n+1}{2}} & ; \ para \ -\infty < x < +\infty \end{cases}$$

Portanto,

$$f(x \mid n) = \frac{\Gamma\left(\dfrac{n+1}{2}\right)}{\Gamma(n/2)} \frac{1}{\sqrt{n\pi}} \left(1 + \frac{x^2}{n}\right)^{-\frac{n+1}{2}}, \ -\infty < x < +\infty, \ e \ n \geq 1 \ (inteiro) \tag{4.21}$$

$$E(X) = 0 \quad e \quad Var(X) = \frac{n}{n-2}$$

Notação: usa-se a notação $X \sim t_n$ para indicar que a variável aleatória X tem distribuição de t-Student com n graus de liberdade.

4.8.8 Modelo de Probabilidade Beta

Definição 4.8.8. *Seja X a variável aleatória que tem distribuição Beta com parâmetro a e b, onde a função densidade de probabilidade dessa distribuição apresenta-se da seguinte forma*

$$f(x) = \begin{cases} \dfrac{\Gamma\left(a+b\right)}{\Gamma(a)\ \Gamma(b)} \ x^{a-1} \left(1-x\right)^{b-1}; & para \ \ 0 < x < 1 \\ 0 & \end{cases}$$

Portanto,

$$f(x \mid a, b) = \frac{\Gamma\left(a+b\right)}{\Gamma(a)\ \Gamma(b)} \ x^{a-1} \left(1-x\right)^{b-1}, \ 0 < x < 1, \ a > 0 \ e \ b > 0 \tag{4.22}$$

$$E(X) = \frac{a}{a+b} \quad e \quad Var(X) = \frac{ab}{(a+b)^2\ (a+b+1)}$$

Notação: usa-se a notação $X \sim Beta(a,b)$ para indicar que a variável aleatória X tem distribuição de Beta com parâmetros a e b.

4.9 Transformações de Variáveis

Em muitas situações não estamos interessados no comportamento de uma variavel propriamente, mas sim em uma transformação dela, $Y = T(X)$, tais como $Y = |X|$, $Y = aX + b$ ou $Y = X^2$. Na teoria dos testes estatísticos essa é uma situação extremamente frequente.

Uma das formas de conhecer o comportamento de $Y = T(X)$ e obter sua função densidade $f_Y(y)$ ou função função de distribuição $F_Y(y)$. Obviamente podemos usar a propriedade $F'(y) = f(y)$. De forma geral, com o uso da regra da cadeia, obtemos,

$$f(y) = \frac{\partial F(y)}{\partial y} = f(x)\left|\frac{dx}{dy}\right|. \tag{4.23}$$

Exemplo 4.1. *Seja $X \sim U(0,1)$. Obtenha a distribuição de $Y = 1 - X$ através da Função de Distribuição de X*

Temos que $f_X(x) = 1, x \in (0,1)$ e que $F_X(x) = x$, com $y \in (0,1)$. Com base nisso,

$$F_Y(y) = P(Y \leq y) = P(1-X \leq y) = P(X \geq 1-y) = 1-P(X < 1-y) = 1-F_X(1-y) = 1-(1-y) = y.$$

Como $F_Y(y) = y$, podemos concluir diretamente que $Y \sim U(0,1)$. Podemos ainda obter sua função densidade, dada por $f_Y(y) = F'_Y(y) = 1$.

Poderíamos aplicar diretamente a expressão 4.23. Notemos que $f_Y(1-x) = 1$ e $\frac{dx}{dy} = -1$, pois $x = 1-y$, obtendo $f_Y(y) = 1 \times |-1| = 1$.

Exemplo 4.2. *Seja $X \sim U(0,1)$. Obtenha a distribuição de $Y = -\frac{1}{\lambda}\ln(1-X)$, através da Função de Distribuição de X*

Temos que $F_X(x) = 1 - e^{-\lambda x}$ e $y \in (0,\infty)$. Com base nisso,

$$\begin{aligned} F_Y(y) &= P(Y \leq y) = P\left(-\frac{1}{\lambda}\ln(1-X) \leq y\right) = P(-\ln(1-X) \leq \lambda y) = P(\ln(1-X) \geq -\lambda y) \\ &= P(1-X \geq e^{-\lambda y}) = P(X \leq 1-e^{-\lambda y}) = F_X(1-e^{-\lambda y}) = 1-e^{-\lambda y}. \end{aligned}$$

Como $F_Y(y) = 1 - e^{-\lambda y}$, podemos concluir diretamente que $Y \sim Exp(\lambda)$. Podemos ainda obter sua função densidade, dada por $f_Y(y) = F'_Y(y) = \lambda e^{-\lambda x}$.

Para aplicar diretamente a expressão 4.23, notemos que $x = 1 - e^{-\lambda y}$, de onde obtemos $f_Y(y) = 1$ e $\frac{dx}{dy} = \lambda e^{-\lambda y}$, obtendo $f_Y(y) = 1 \times |\lambda e^{-\lambda y}| = \lambda e^{-\lambda y}$.

4.10 Exercícios

1) Suponha que uma variável aleatória discreta X tenha uma distibuição com as probabilidades dadas pela Tabela (4.6). a seguir

Tabela 4.6 *Distribuição de Probabilidade da Variável Aleatória Z.*

Z	0	1	2	3	4	5
P(z)	0	p^2	p^2	p	p	p^2

(a) Obtenha o valor de p;

(b) A distribuição acumulada;

(c) $P(Z \geq 4)$;

(d) $P(Z < 3)$.

2) Uma urna contém 4 bolas brancas e 6 pretas. 3 bolas são retiradas com reposição. Seja X a variável aleatória que representa o número de bolas brancas, calcule:

(a) E(X);

(b) Var(X);

(c) P(X = 3).

3) Sabe-se que uma moeda mostra a face cara o quádruplo de vezes do que a face coroa, quando lançada. Esta moeda é lançada 4 vezes. Seja X o número de vezes que se obtem a face cara. determine:

(a) E(X);

(b) Var(X);

(c) $P(X \geq 2)$;

(d) $P(1 \leq X \leq 3)$.

4) A função de probabilidade da variável aleatória X é:

Tabela 4.7 *Distribuição de Probabilidade da Variável Aleatória X.*

X	1	2	3	4	5
P(x)	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{5}$

Calcule:

(a) $E(X)$;

(b) $E(X^2)$;

(c) $E(X+3)^2$;

(d) Var(3X-2).

5) Em um experimento de lançar uma moeda honesta sobre a superfície de uma mesa plana e verificar a face da moeda. Observa-se 20 lançamentos dessa moeda. Qual a probabilidade de saírem 10 caras?

6) Um "sinal"consiste de uma séries de vibrações de magnitude X. Um "ruído" consiste de uma série de vibrações de magnitude Y, apresentando os valores 2, 0 e -2, com probabilidades $\frac{1}{6}$, $\frac{2}{3}$ e $\frac{1}{6}$ respectivamente, se ruídos e sinais são combinados de vibrações sincronizadas, a soma dos mesmos consiste de vibrações de magnitude $Z = X+Y$. Construir a função de probabilidade de Z e calcular: $E(Z)$ e $Var(Z)$, admitindo independência entre ruído e sinal. a variável X assume os valores 1, 0 e -1, cada um com probabilidade $\frac{1}{3}$.

7) As probabilidades de que haja em cada carro que vá a Santos no sábado com 1, 2, 3, 4, 5 ou 6 pessoas, são respectivamente: 0,05; 0,20; 0,40; 0,15; 0,12 e 0,08. Qual o número médio de pessoas por carro? Se chagarem a Santos 4.000 carros por hora, qual o número esperado de pessoas na cidade, em 10 horas de contagem?

8) Seja X: o número de caras e Y: o número de coroas, quando são lançadas 3 moedas. Calcular a média e a variância de $Z = 2X + Y$.

9) As variáveis aleatórias W e S são independentes e tem as distribuições de probabilidade expressas por:

Tabela 4.8 *Distribuição de Probabilidade das Variáveis Aleatórias W e S*

W	1	2	S	3	4
P(w)	0,4	0,6	P(s)	0,2	0,8

Considerando a variável aleatória $Z = W + S$, construir a tabela de distribuição de Z e de posse da mesma. Calcular: E(Z) e Var(Z).

10) Dada a distribuição conjunta de probabilidade da variável (X, Y), apresentada na Tabela (4.9). Calcule:

Tabela 4.9 *Distribuição de Probabilidades Marginais das Variáveis Aleatórias X e Y.*

Y	0	2	4	$\mathbf{P}(X = x)$
X				
0	0,5	0	0	0,5
2	0	0,2	0,05	0,25
4	0	0	0,25	0,25
$\mathbf{P}(Y = y)$	0,5	0,2	0,3	**1**

(a) E(X); E(X^2) e Var(X);

(b) E(Y); E(Y^2) e Var(Y).

11) O tempo T em minutos, necessário para um operário processar certa peça é uma variável aleatória com a distribuição de probabilidade dada pela Tabela (4.10).

Calcule o tempo médio de processamento para cada peça, assim como a variabilidade do processo. O operário ganha um fixo de R$ 2,00, mas se ele processa a peça em menos de 6 minutos, ganha R$ 0,50, em cada minuto poupado de trabalho, ou seja, se ele processa a peça em 4 minutos, recebe a quantia adicional de R$ 1,00.

Para o caso do operário processar certa peça em um máximo de tempo igual a 5 minutos, Calcule a probabilidade deste processo de produção ocorrer e o ganho adicional do operário.

Tabela 4.10 *Distribuição de Probabilidade da Variável Aleatória T*

T	2	3	4	5	6	7
P(t)	0,1	0,1	0,3	0,2	0,2	0,1

12) Considere a variável aleatória X tendo distribuição de Poisson. Se a probabilidade de um indivíduo acusar certa reação negativa à injeção de determinado soro é 0,001. Determinar a probabilidade de que, em 2000 indivíduos,

(a) Pelo menos 5 acusem a reação negativa;

(b) Exatamente 3 indivíduos acusem reação negativa;

(c) Nenhum indivíduo acuse reação negativa;

(d) No máximo 2 indivíduos acusem reação negativa;

(e) Pelo menos um indivíduo acuse reação negativa.

13) De acordo com a Estatística vital do Departamento de Saúde dos Estados Unidos, a média anual de afogamentos acidentais nos EUA é 3,0 por 100.000 indivíduos, considera-se esse fenômeno apresentando uma distribuição de Poisson. Determine a probabilidade de que, em uma cidade com 200.000 habitantes, se verifique:

(a) Nenhum afogamento acidental em um ano;

(b) Dois afogamentos acidentais em um ano;

(c) Entre quatro e oito afogamentos em um ano;

(d) Menos de três afogamentos acidentais em um ano;

(e) Pelo menos um afogamento acidental em um ano.

14) Uma companhia de seguros acredita que 0,005% de uma população falece de um certo tipo de acidente. Admite-se que este processo tem uma distribuição de Poisson. Qual é a probabilidade que a companhia tenha que pagar mais do que três pessoas das 10.000 asseguradas contra este tipo de acidentes em um dado ano?

15) Durante um concurso realizado, verificou-se que o tempo gasto no exame de vestibular de uma universidade tem distribuição normal, com média igual à 120 minutos e variância igual à 225 minutos. De posse dessas informações, e sabendo-se que T: tempo gasto no exame de vestibular. $T \sim N(120, 225)$.pergunta-se:

(a) Sorteando-se um aluno ao acaso, qual a probabilidade dele terminar o exame antes de 100 minutos?

(b) Sorteando-se novamente um aluno ao acaso, qual é a probabilidade dele terminar no máximo em duas horas e meia (150 minutos)?

16) Dado uma variável aleatória X a qual apresenta distribuição normal com média 20 e variância 25, ou seja, $X \sim N(20, 25)$. Calcular:

(a) $\mathrm{P}(X \leq 30)$;

(b) $\mathrm{P}(5 \leq X \leq 20)$;

(c) $\mathrm{P}(X \leq 10)$.

17) Dado uma variável aleatória X a qual apresenta distribuição normal com média 10 e variância 16, ou seja, $X \sim N(10, 16)$. Calcular:

(a) $\mathrm{P}(X \leq 5)$;

(b) $\mathrm{P}(15 \leq X \leq 35)$;

(c) $\mathrm{P}(X \leq 20)$.

18) A variável aleatória X tem f.d.p. dada por:

$$f(x) = \begin{cases} \dfrac{1}{40}\left(\dfrac{x}{10} + 1\right); & para \;\; 0 \leq x \leq 20, \\ \\ 0, & caso \;\; contrario. \end{cases}$$

Calcule,

(a) $\mathrm{E}(X)$;

(b) $\mathrm{Var}(X)$.

19) A variável aleatória contínua bidimensional (X, Y) tem a distribuição conjunta dada por:

$$f(x) = \begin{cases} \dfrac{3xy^2}{4}; & para \;\; 0 \leq x \leq 1, \;\; 0 \leq y \leq 2 \\ \\ 0, & caso \;\; contrario. \end{cases}$$

Determine,

(a) As distribuições marginais de X e Y;

(b) Se X e Y são independentes;

(c) A covariância de X e Y.

20) Dada a distribuição conjunta de probabilidade da variável (X, Y), apresentada na Tabela (4.11) a seguir. Calcule o coeficiente de correlação.

21) As alturas dos alunos de uma escola do ensino médio de um determinado município do Estado do Pará são normalmente distribuidos com média 1,60 metros e desvio padrão 0,30 metros. Encontre a probabilidade de um aluno medir:

(a) Entre 1,50 e 1,80 metros;

(b) Menos de 1,50 metros;

Tabela 4.11 *Distribuição de Probabilidade Conjunta da Variável Aleatória de X e Y.*

$X \setminus Y$	0	2	4
0	0,5	0	0
2	0	0,2	0,05
4	0	0	0,25

(c) Entre 1,65 e 1,80 metros;

(d) Mais de 1,90 metros.

22) O tempo necessário para um medicamento contra a dor fazer efeito foi modelado segundo uma função densidade de probabilidade Uniforme no intervalo de 5 a 15 minutos, tendo por base experimentos conduzidos em animais. Um paciente, que esteja sofrendo dor, recebe o remédio e, supondo válido o modelo uniforme, qual a probabilidade da dor:

(a) Cessar em até 10 minutos;

(b) Demorar pelo menos 12 minutos.

23) A função densidade de probabilidade dada abaixo

$$f(x) = \begin{cases} 2e^{-2x}; & para \;\; x \geq \; 0 \\ & \\ 0; & para \;\; x \; \leq \; 0 \end{cases}$$

Representa a distribuição do Índice de Acidez (X) de um determinado produto alimentício. O produto é consumível se este índice for menor que 2. O setor de fiscalização do I.A.L., apreendeu 30 unidades do mesmo. Qual a probabilidade de que pelo menos 10% da amostra seja imprópria para o consumo.

24) O tempo de vida de um dispositivo é dado pela f.d.p. dada abaixo:

$$f(x) = \begin{cases} k(10-x); & para \;\; 5 \leq x \leq \; 10 \\ & \\ kx; & para \;\; se \;\; 0 \; \leq \; x \; < \; 5 \\ & \\ 0; & para \;\; x \; < \; 0 \; ou \; x \; > \; 10 \\ & \end{cases}$$

25) Suponha que uma variável aleatória X tenha f.d.p. dada por:

$$f(x) = \begin{cases} \dfrac{2x}{5} + k; & para \;\;\; 0 \; < \; x < \; 4 \\ & \\ 0; & Caso \;\; contrario \end{cases}$$

Calcule:

(a) Qual é o valor de k?

(b) Quanto vale b, tal que $P(X > b) = \dfrac{1}{4}$?

(c) Calcule a $\mathrm{E}(X)$.

26) A função densidade de probabilidade dada abaixo

$$f(x) = \begin{cases} kx; & para\ 0 \ < \ x \leq \ 1 \\ \\ 0; & para \ \ x \ \leq \ 0 \ ou \ x \ > \ 1 \end{cases}$$

Calcule:

(a) O valor de k para que $f(x)$ seja uma f.d.p.;

(b) $P(0 \ \leq \ X \ \leq \dfrac{1}{2})$;

(c) $\mathrm{E}(X)$;

(d) $\mathrm{Var}(X)$.

27) Suponha uma variável aleatória discreta X tendo a seguinte distribuição de probabilidades: Obtenha

Tabela 4.12 *Distribuição de Probabilidade da Variável Aleatória X*

X	1	2	3	4	5
P(x)	0,1	0,2	0,4	0,2	0,1

a função de distribuição acumulada.

28) Uma moeda é lançada 20 vezes. Qual a probabilidade de saírem 10 caras?

29) Dado um certo experimento, o qual caracteriza-se por apresentar uma distribuição de *Poisson* com parâmetro θ, ou seja, $X \sim Poisson(\theta)$, onde a $P(X = 0) = 0,2$; Obtenha o valor de θ e calcule, $P(X < 2)$.

30) Um jogador lança um dado. Se aparecer os números 1, 2 ou 3, recebe R$ 10,00. Se no entanto, aparecer 4 ou 5, recebe R$ 5,00. Se aparecer 6 ganha R$20,00. Qual o ganho médio do jogador.

Tabela 4.13 *Tabela da Distribuição* $Normal(0;1)$ *ou Normal Padrão Reduzida.*

Valores de p **tais que** $P(0 \leq Z \leq z) = p$

Segunda Decimal de z

Parte Inteira e Primeira Decimal de z	z	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
	0,0	0,0000	0,0040	0,0080	0,0120	0,0160	0,0199	0,0239	0,0279	0,0319	0,0359
	0,1	0,0398	0,0438	0,0478	0,0517	0,0557	0,0596	0,0636	0,0675	0,0714	0,0753
	0,2	0,0793	0,0832	0,0871	0,0910	0,0948	0,0987	0,1026	0,1064	0,1103	0,1141
	0,3	0,1179	0,1217	0,1255	0,1293	0,1331	0,1368	0,1406	0,1443	0,1480	0,1517
	0,4	0,1154	0,1591	0,1628	0,1664	0,1700	0,1736	0,1772	0,1808	0,1844	0,1879
	0,5	0,1915	0,1950	0,1985	0,2019	0,2054	0,2088	0,2123	0,2157	0,2190	0,2224
	0,6	0,2257	0,2291	0,2324	0,2357	0,2389	0,2422	0,2454	0,2486	0,2517	0,2549
	0,7	0,2580	0,2611	0,2642	0,2673	0,2704	0,2734	0,2764	0,2794	0,2823	0,2852
	0,8	0,2881	0,2910	0,2939	0,2967	0,2995	0,3023	0,3051	0,3078	0,3106	0,3133
	0,9	0,3159	0,3186	0,3212	0,3238	0,3264	0,3289	0,3315	0,3340	0,3365	0,3389
	1,0	0,3413	0,3438	0,3461	0,3485	0,3508	0,3531	0,3554	0,3577	0,3599	0,3621
	1,1	0,3643	0,3665	0,3686	0,3708	0,3729	0,3749	0,3770	0,3790	0,3810	0,3830
	1,2	0,3849	0,3869	0,3888	0,3907	0,3925	0,3944	0,3962	0,3980	0,3997	0,4015
	1,3	0,4032	0,4049	0,4066	0,4082	0,4099	0,4115	0,4131	0,4147	0,4162	0,4177
	1,4	0,4192	0,4207	0,4222	0,4236	0,4251	0,4265	0,4279	0,4292	0,4306	0,4319
	1,5	0,4332	0,4345	0,4357	0,4370	0,4382	0,4394	0,4406	0,4418	0,4429	0,4441
	1,6	0,4452	0,4463	0,4474	0,4484	0,4495	0,4505	0,4515	0,4525	0,4535	0,4545
	1,7	0,4554	0,4564	0,4573	0,4582	0,4591	0,4599	0,4608	0,4616	0,4625	0,4633
	1,8	0,4641	0,4649	0,4656	0,4664	0,4671	0,4678	0,4686	0,4693	0,4699	0,4706
	1,9	0,4713	0,4719	0,4726	0,4732	0,4738	0,4744	0,4750	0,4756	0,4761	0,4767
	2,0	0,4772	0,4778	0,4783	0,4788	0,4793	0,4798	0,4803	0,4808	0,4812	0,4817
	2,1	0,4821	0,4826	0,4830	0,4834	0,4838	0,4842	0,4846	0,4850	0,4854	0,4857
	2,2	0,4861	0,4864	0,4868	0,4871	0,4875	0,4878	0,4881	0,4884	0,4887	0,4890
	2,3	0,4893	0,4896	0,4898	0,4901	0,4904	0,4906	0,4909	0,4911	0,4913	0,4916
	2,4	0,4918	0,4920	0,4922	0,4925	0,4927	0,4929	0,4931	0,4932	0,4934	0,4936
	2,5	0,4938	0,4940	0,4841	0,4943	0,4945	0,4946	0,4948	0,4949	0,4951	0,4952
	2,6	0,4953	0,4955	0,4956	0,4957	0,4959	0,4960	0,4961	0,4962	0,4963	0,4964
	2,7	0,4965	0,4966	0,4967	0,4968	0,4969	0,4970	0,4971	0,4972	0,4973	0,4974
	2,8	0,4974	0,4975	0,4976	0,4977	0,4977	0,4978	0,4979	0,4979	0,4980	0,4981
	2,9	0,4981	0,4982	0,4982	0,4983	0,4984	0,4984	0,4985	0,4985	0,4986	0,4986
	3,0	0,4987	0,4987	0,4987	0,4988	0,4988	0,4989	0,4989	0,4989	0,4990	0,4990
	3,1	0,4990	0,4991	0,4991	0,4991	0,4992	0,4992	0,4992	0,4992	0,4993	0,4993
	3,2	0,4993	0,4993	0,4994	0,4994	0,4994	0,4994	0,4994	0,4995	0,4995	0,4995
	3,3	0,4995	0,4995	0,4995	0,4996	0,4996	0,4996	0,4996	0,4996	0,4996	0,4997
	3,4	0,4997	0,4997	0,4997	0,4997	0,4997	0,4997	0,4997	0,4997	0,4997	0,4998
	3,5	0,4998	0,4998	0,4998	0,4998	0,4998	0,4998	0,4998	0,4998	0,4998	0,4998
	3,6	0,4998	0,4998	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999
	3,7	0,4999	0,4999	0,4999	0.4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999
	3,8	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999
	3,9	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000

Capítulo **5**

Variáveis Aleatórias Bidimensionais e Multidimensionais

Embora o desenvolvimento visto até aqui refere-se à apresentação das principais variáveis aleatórias, é comum na prática a observação de várias variáveis ao mesmo tempo, considerando que elas estão, de certa forma, relacionadas. Por exemplo, a dureza X e a tensão de ruptura Y de uma peça manufaturada de aço poderão interessar, e considera-se (X, Y) como um único resultado experimental. Pode-se estudar a estatura H e o peso P de alguma pessoa escolhida, o que forneceria o resultado (h, p). Finalmente, pode-se observar a altura total da chuva R e a temperatura T em uma certa localidade, durante um mês especificado, dando origem ao resultado do (r, t). Em uma situação mais abrangente, teremos o caso em que várias observações são realizadas ao mesmo tempo, dando origem a um vetor aleatório multidimensional $\boldsymbol{X} = (X_1, X_2, \cdots, X_n)$. Vamos começar apresentado toda a teoria para o caso bidimensional, fazendo a extensão ao cado multidimensional quando necessário.

5.1 Definição

*Sejam E um **Experimento** e $\Omega = (\omega_i)$ ou $S = (s_i)$ o **Espaço Amostral**, ou seja, o conjunto dos resultados possíveis do experimento*

$$S = \{s_1, s_2, s_3, \ldots\ldots\}$$

***Variável Aleatória** X: É uma função que associa a cada elemento do espaço amostral um n^o real.*

$$(X, Y) : S \longrightarrow \mathbb{R}^2$$

$$s_i \longrightarrow X(s_i) = x_i, \quad Y(s_i) = y_i,$$

Definição 5.1. *Sejam E um experimento e S um espaço amostral associado a E. Sejam $X = X(s)$ e $Y = Y(s)$ duas funções, cada uma ssociada a um número real a cada resultado $s \in S$. Denominaremos (X, Y) uma variável aleatória bidimensional (ou vetor aleatório).*

OBS:

1. *Se $X_1 = X_1(s), X_2 = X_2(s), \ldots X_n = X_n(s)$ forem n funções cada uma associando um n^o real a cada resultado $s \in S$, denominaremos $(X_1, \ldots, X_n)$ uma variável aleatória n-dimensional (ou vetor aleatório n-dimensional).*

2. *Não estamos interessados na natureza funcional de $X(s)$ e $Y(s)$, mas sim nos valores possíveis que X e Y tomam.*

   R_{XY} : *contradomínio de $(X, Y) \longrightarrow$ conjunto dos valores possíveis de (X, Y).*

OBS: *No caso bidimensional, R_{XY} será um subconjunto do plano euclidiano $I\!R^n$.*

Exemplo 5.1. *Retira-se uma amostra de 20 alunos do curso de estatística da UFPA e anota-se a idade (X) e ano de ingresso no curso (Y). O vetor (X, Y) será bidimensional. Outras variáveis poderiam ser medidas.*

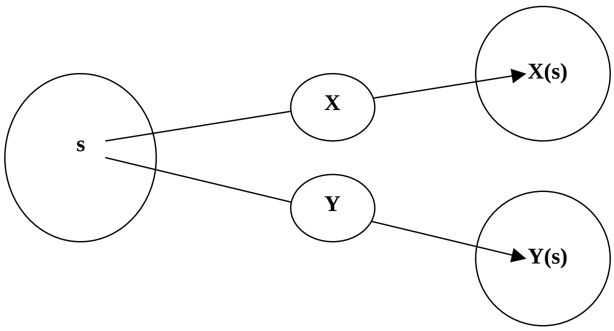


Figura 5.1 Função de Variável Aleatória Bidimensional.

Definição 5.2. *O vetor (X, Y) será uma variável aleatória **Discreta** Bidimensional se o conjunto dos valores possíveis de (X, Y) for finito ou infinito enumerável. Isto é, os valores possóveis de (X, Y) possam ser representados por (x_i, y_j), $\quad i = 1, 2, \ldots; \quad j = 1, 2, \ldots$*

Definição 5.3. *O vetor (X, Y) será uma variável aleatória **Contínua** Bidimensional se (X, Y) puder tomar todos os valores em algum conjunto não-enumerável do plano auclidiano.*

Exemplo 5.2.

$$\{(x, y) : a \leq x \leq b, \quad c \leq y \leq d\} \quad - \quad \text{Retângulo}$$
$$\{(X, Y) : x^2 + y^2 \leq 1\} \quad - \quad \text{circulo}$$

5.1.1 Variáveis discretas

Seja (X, Y) uma variável aleatória discreta bidimensioal. A cada resultado possível (x_i, y_j) associaremos um valor $p(x_i, y_j)$ representando $P(X = x_i, Y = y_i)$ e satisfazendo ás seguinte condições:

1) $p(x_i, y_j) \geq 0, \forall (x, y)$.

2) $\sum\limits_{j=1}^{\infty} \sum\limits_{i=1}^{\infty} p(x_i, y_j) = 1$

*A função **p** definida para todo (x_i, y_j) no contra domínio de (X, Y) é denominada a Função de Probabilidade Conjunta de (X, Y) . O conjunto $[x_i, y_j, \quad p(x_i, y_j)]$, $i, \; j = 1, 2, \ldots$ é, algumas vezes, denominado Distribuição de Probabilidade Conjunta de (X, Y) . Na Tabela 5.1.1 temos a representação usual da distribuição conjunta.*

Exemplo 5.3. *Lançam-se dois dados perfeitos. X indica o n^o obtido no primeiro dado, e Y o maior ou o n^o comum nos dois dados. Encontre a distribuição de probabilidade conjunta de (X, Y) .*

Exemplo 5.4. *Lançam-se dois dados perfeitos. X indica o n^o obtido no primeiro dado, e Y a soma dos dois dados. Encontre a distribuição de probabilidade conjunta de (X, Y) .*

Exemplo 5.5. *Lançam-se dois dados perfeitos. X indica o máximo dos dois resultados e Y a soma dos dois dados. Encontre a distribuição de probabilidade conjunta de (X, Y) .*

Tabela 5.1 Distribuição Conjunta de Probabilidade das Variáveis X e Y

X/Y	y_1	y_2	y_3	...	y_m
x_1	p(x_1, y_1)	p(x_1, y_2)	p(x_1, y_3)	...	p(x_1, y_m)
x_2	p(x_2, y_1)	p(x_2, y_2)	p(x_2, y_3)	...	p(x_2, y_m)
x_3	p(x_3, y_1)	p(x_3, y_2)	p(x_3, y_3)	...	p(x_3, y_m)
.	.	.	.	...	.
.	.	.	.	...	.
.	.	.	.	...	.
x_n	p(x_n, y_1)	p(x_n, y_2)	p(x_n, y_3)	...	p(x_n, y_m)
Soma	p(x_i, y_1)	p(x_i, y_2)	p(x_i, y_3)	...	1

$$S = \begin{bmatrix} (1,1) & (1,2) & (1,3) & (1,4) & (1,5) & (1,6) \\ (2,1) & (2,2) & (2,3) & (2,4) & (2,5) & (2,6) \\ (3,1) & (3,2) & (3,3) & (3,4) & (3,5) & (3,6) \\ (4,1) & (4,2) & (4,3) & (4,4) & (4,5) & (4,6) \\ (5,1) & (5,2) & (5,3) & (5,4) & (5,5) & (5,6) \\ (6,1) & (6,2) & (6,3) & (6,4) & (6,5) & (6,6) \end{bmatrix}$$

Exemplo 5.6. *Uma moeda perfeita é lançada 3 vezes. Para as variáveis (X, Y, Z) definidas a seguir, encontrar a Distribuição Conjunta.*
X: n^o de caras obtidas nos dois primeiros lançamentos.
Y: n^o de caras obtidas no último lançamento.
Z: n^o Total de caras.

Exercício 5.1.1. *Montar uma macro para obter uma aproximação para a distribuição conjunta de (X, Y, Z).*

Exemplo 5.1.1. Supondo que se esteja interessado em estudar a composição de famílias com três crianças, quanto ao sexo. Defini-se então as seguintes variáveis X = Número de nascimentos de homens

$$Y = \begin{cases} 1 \text{ se no primeiro nascimento for um homem} \\ \\ 0, \text{ se no primeiro nascimento for uma mulher} \end{cases}$$

Exemplo 5.1.2. Supondo que se esteja interessado em estudar a composição de famílias com três crianças, quanto ao sexo. Defini-se então as seguintes variáveis

X = Número de meninos
Z = Número de vezes em que há variação do sexo entre um nascimento e outro, dentro de uma mesma família.

5.1.2 Variáveis contínuas

Seja (X, Y) uma variável aleatória bidimensional contínua tomando todos os valores em alguma região R_{XY} do plano euclidiano. Uma função f que satisfaça ás seguintes condições:

Tabela 5.2 *Resultados do lançamento de dois dados*

Y \ X	1	2	3	4	5	6	Total
1	1/36	0	0	0	0	0	1/36
2	1/36	2/36	0	0	0	0	3/36
3	1/36	1/36	3/36	0	0	0	5/36
4	1/36	1/36	1/36	4/36	0	0	7/36
5	1/36	1/36	1/36	1/36	5/36	0	9/36
6	1/36	1/36	1/36	1/36	1/36	6/36	11/36
Total	6/36	6/36	6/36	6/36	6/36	6/36	1

Tabela 5.3 *Espaço amostral*

Resultados	Probab	X	Y	Z
ccc	1/8	2	1	3
cc$\bar{c}$	1/8	2	0	2
c$\bar{c}$c	1/8	1	1	2
c$\bar{c}\bar{c}$	1/8	1	0	1
$\bar{c}$cc	1/8	1	1	2
$\bar{c}$c$\bar{c}$	1/8	1	0	1
$\bar{c}\bar{c}$c	1/8	0	1	1
$\bar{c}\bar{c}\bar{c}$	1/8	0	0	0
Total	1	8	4	12

i) $f(x,y) \geq 0 \quad \forall (x,y) \in \quad R_{XY}$

ii) $\displaystyle\iint_{R_{XY}} f(x,y)dxdy = 1$

é denominada ***Função Densidade de Probabilidade Conjunta*** *de* (X,Y) *. Na Figura 5.1.2 temos a ilustração de uma v.a. contínua bidimensional.*

Observação 5.1.

1) $f(x,y)$ *não representa propriamente a probabilidade de coisa alguma. Esse valor pode, inclusive, ser maior que 1. Contudo, para* Δx *e* Δy *positivos e suficientemente pequenos,* $f(x,y)\Delta x \Delta y$ *é aproximadamente igual* $P(x - \Delta x \leq X \leq x + \Delta x, y - \Delta y \leq Y \leq y + \Delta y)$

2) Assim como no caso unidimensional, adotaremos a conveção de que $f(x,y) = 0$ *se* $(x,y) \notin R_{XY}$*, de forma que*

$$\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x,y)dxdy = 1$$

Tabela 5.4 *Distribuição Conjunta de* (X, Y, Z)

(x, y, z)	Probabilidade
(2,1,3)	1/8
(2,0,2)	1/8
(1,1,2)	2/8
(1,0,1)	2/8
(0,1,1)	1/8
(0,0,0)	1/8
Total	1

Tabela 5.5 *Primeira composição familiar com três crianças.*

X/Y	0	1	$\mathrm{P}(X = x)$
0	1/8	0	1/8
1	2/8	1/8	3/8
2	1/8	2/8	3/8
3	0	1/8	1/8
$\mathrm{P}(Y = y)$	4/8	4/8	1

Tabela 5.6 *Segunda composição familiar com três crianças.*

X/Z	0	1	2	$\mathrm{P}(X = x)$
0	1/8	0	0	1/8
1	0	2/8	1/8	3/8
2	0	2/8	1/8	3/8
3	1/8	0	0	1/8
$\mathrm{P}(Z = z)$	2/8	4/8	2/8	1

3) Se B *for um evento no contradomínio de* (X, Y) *, teremos:*

$$P(B) = \sum_B \sum p(x_i, y_j)$$

se (X, Y) *for Discreta, onde a soma é feita para os índices* (i, j) *tais que* $(x_i, y_j) \in B$*. E, se* (X, Y) *for contínua:*

$$P(B) = \iint_B f(x, y) dx dy$$

Exemplo 5.7. *Suponha que a variável aleatória* (X, Y) *tenha f.d.p conjunta dada por:*

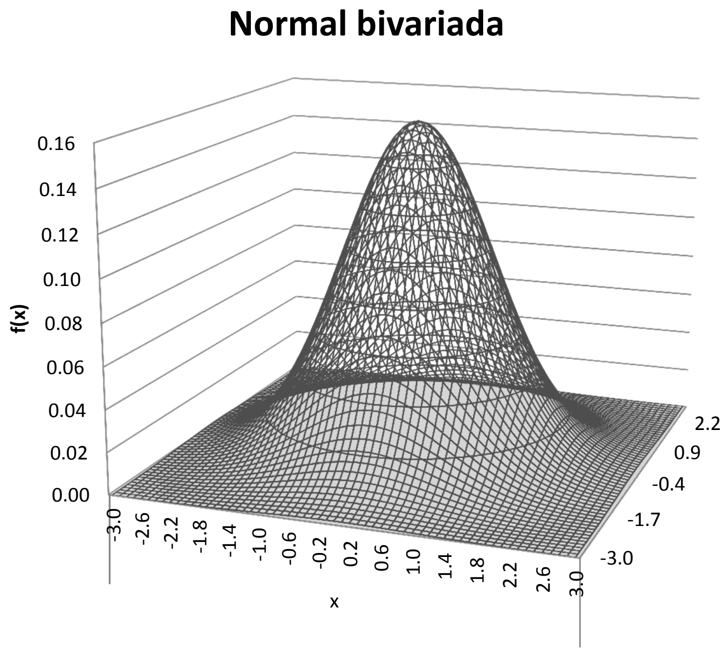


Figura 5.2 *Ilustração de uma v.a. contínua bidimensional*

$$f(x,y) = \begin{cases} e^{-(x+y)} & x > 0, y > 0 \\ 0, & c.c. \end{cases}$$

a) Calcule $P(0 < X < 1, \quad 1 < Y < 2)$

b) Desenhe a região $B = \{X > Y\} = \{(x,y) : x > y\}$

c) Calcule $P(X > Y)$

Solução de a):

$$\begin{aligned} P(0 < X < 1, 1 < Y < 2) &= \int_1^2 \int_0^1 f(x,y)dxdy \\ &= \int_1^2 \int_0^1 e^{-(x+y)}dxdy = \int_1^2 e^{-y}\left[-e^{-x}\right]_0^1 dy \\ &= \int_1^2 e^{-y} \quad [-e^{-1}+1]dy = (1-e^{-1}) \ \left[-e^{-y}\right]_1^2 \\ &= (1-e^{-1})[-e^{-2}+e^{-1}] \simeq 0,147 \end{aligned}$$

Solução de b):

$$\begin{aligned}
P(X>Y) &= \int_0^{+\infty}\int_y^{+\infty} f(x,y)dxdy \\
&= \int_0^{+\infty}\int_y^{+\infty} e^{-(x+y)}dxdy \\
&= \int_0^{+\infty} e^{-y}\left[-e^{-x}\right]_y^{+\infty} dy = \int_0^{+\infty} e^{-y}(0+e^{-y})dy \\
&= \int_0^{+\infty} e^{-2y}dy = \frac{-e^{-2y}}{2} \quad |_0^{+\infty} = 0+\frac{1}{2} = \frac{1}{2}.
\end{aligned}$$

```
Sub Expbi()
 n = 10 ^ 6
 For i = 1 To n
    x = Rnd(): y = Rnd() + 1
    f = Exp(-(x + y))

    Z = Rnd()
    If Z < f Then conta = conta + 1
 Next
 Prob = conta / n
 MsgBox "Probabilidade: " & Prob
End Sub
```

Para obter uma aproximação para $P(X>Y)$ podemos usar o mesmo procedimento, com alguns ajustes. A integração deverá ser feita em um quadrado maior, $[0,a]^2$, com $a=10$, onde a altura máxima da função é $h=f(0,0)=1$. Assim, estaremos inserindo a função desejada no parlelepídedo de volume $V=a^2\times 1$.

```
 Sub Expbi2()
 n = 10 ^ 7
 a = 10: h = 1
 V = (a ^ 2) * h  'Volume do cubo
 Conta = 0

 For i = 1 To n
    x = a * Rnd()   ' x in (0,a)
    y = a * Rnd()   ' y in (0,a)

    F = Exp(-(x + y))
    Z = Rnd() * h
    If (Z < F And x < y) Then Conta = Conta + 1
 Next
 Prob = Conta / n
 MsgBox "Probabilidade: " & Prob * V
 End Sub
```

Exemplo 5.8. *Suponha que a variável aleatória (X, Y) tenha f.d.p conjunta dada por:*

$$f(x,y) = \begin{cases} x^2 + \frac{xy}{3} & 0 \leq x \leq 1, 0 \leq y \leq 2 \\ 0, & c.c. \end{cases}$$

a) Verifique que integra 1

b) Desenhe a região $B = \{X + Y \geq 1\}$

c) Calcule $P(X + Y \geq 1)$

```
Sub Ex68()
n = 10 ^ 7
h = 2
V = 1 * 2 * h      ' Volume do paralelepípedo que vai cobrir a densidade
Conta = 0

For i = 1 To n
   x = 1 * Rnd()   ' x in (0,1)
   y = 2 * Rnd()   ' y in (0,2)

   F = x ^ 2 + x * y / 3
   Z = h * Rnd()
   If (Z < F And x + y > 1) Then Conta = Conta + 1
Next
Prob = Conta / n   ' fração de pontos que cai na região de interesse
MsgBox "Probabilidade: " & Prob * V
End Sub
```

Exemplo 5.9 (Caso discreto)**.** *No exemplo dos dois dados:*

$$P(X \geq 1, Y = 3) = \frac{5}{36}$$

Definição 5.4. *Seja (X, Y) uma variável aleatória bidimensional. A função de distribuição acumulada (*f.d.a*) F de (X, Y) é definida por*

$$F(x, y) = P(X \leq x, Y \leq y)$$

Se F for a fda de uma v.a bidimensional contínua com fdp *f, então:*

$$\frac{\partial^2 F(x,y)}{\partial x \partial y} = f(x,y)$$

sempre que F for derivável.

Exemplo 5.10. *Suponha que a v.a (X,Y) tenha fdp dada por $f(x,y) = e^{-(x+y)}, \quad x > 0, y > 0$. Encontre $F(x,y)$.*

$$\begin{aligned}
F(x,y) &= P(X \leq x, Y \leq y) = \int_0^y \int_0^x f(u,t)dudt \\
&= \int_0^y \int_0^x e^{-(u+t)}dudt = \int_0^y (-e^{-u+t}|_0^x)dt \\
&= \int_0^y [-e^{-(x+t)} + e^{-t}]dt = e^{-(x+t)} - e^{-t}|_0^y \\
&= e^{-(x+y)} - e^{-y} - e^{-x} + 1
\end{aligned}$$

Portanto,

$$\frac{\partial^2 F(x,y)}{\partial x \partial y} = \frac{\partial}{\partial x}[-e^{-(x+y)} + e^{-y}] = e^{-(x+y)} = f(x,y)$$

5.2 Distribuiçao de Probabilidade Marginal

Seja (X,Y) uma variável aleatória bidimensional. Desejamos obter as distribuição individuais (Marginais) de probabilidade de X e de Y, a partir da distribuição conjunta.

Da Teoria de Conjuntos *temos que se A é um conjunto de S e $B_i, i = 1, \cdots, n$, formam uma partição disjunta (mutuamente exclusiva) de S, ou seja, $S = \bigcup_{i=1}^n B_i$, então $C_i = A \bigcap B_i$ são disjuntos. E, ainda,*

$$A = A \bigcap S = A \bigcap (\bigcup_{i=1}^n B_i) = \bigcup_{i=1}^n (A \bigcap B_i)$$

Com isso,

$$P(A) = P(\bigcup_{i=1}^n C_i) = \sum_{i=1}^n P(C_i) = \sum_{i=1}^n P(A \bigcap B_i).$$

Assim, se $A = \{X = x\}$ e $B_i = \{Y = y_i\}$ então

$$P(X = x) = \sum_{i=1}^n P(X = x, Y = y_i). \tag{5.1}$$

Exemplo 5.11. *Um dado perfeito é lançado 2 vezes. Sejam*

X: n^o obtido no primeiro dado
Y: Y o maior ou o n^o comum nos dois dados

$$P(X = 1) = P(X = 1, Y = 1) + P(X = 1, Y = 2) + P(X = 1, Y = 3) + \ldots + P(X = 1, Y = 6) = \frac{6}{36}$$

Tabela 5.7 *Resultados do lançamento de dois dados*

Y \ X	1	2	3	4	5	6	Total
1	1/36	0	0	0	0	0	1/36
2	1/36	2/36	0	0	0	0	3/36
3	1/36	1/36	3/36	0	0	0	5/36
4	1/36	1/36	1/36	4/36	0	0	7/36
5	1/36	1/36	1/36	1/36	5/36	0	9/36
6	1/36	1/36	1/36	1/36	1/36	6/36	11/36
Total	6/36	6/36	6/36	6/36	6/36	6/36	1

Distribuição Marginal de X

x_i	*1*	*2*	*3*	*4*	*5*	*6*	*Total*
$p(x_i)$	*6/36*	*6/36*	*6/36*	*6/36*	*6/36*	*6/36*	*1*

Distrbuição Marginal de Y

y_i	*1*	*2*	*3*	*4*	*5*	*6*	*Total*
$p(y_i)$	*1/36*	*3/36*	*5/36*	*7/36*	*9/36*	*11/36*	*1*

a) Caso Discreto:

$$\begin{aligned} p(x_i) &= P(X = x_i) = P(X = x_i \ , \ Y = y_1 \ ou \ X = x_i \ , \ Y = y_2 \ ou \ \ldots) \\ &= \sum_{j=1}^{\infty} P(X = x_i, Y = y_j) = \sum_{j=1}^{\infty} p(x_i \, y_j). \\ q(y_i) &= P(Y = y_j) = \sum_{i=1}^{\infty} p(x_i, y_j) \end{aligned}$$

A função p, definida para $x_1, x_2, \ldots$, representa a distribuição de probabilidade marginal de X, e a função q, representa distribuição de probabilidade marginal de Y.

b) Caso Contínuo:

$$g(x) = \int_{-\infty}^{+\infty} f(x, y) dy \quad f.d.p \ marginal \ de \ X$$

$$h(y) = \int_{-\infty}^{+\infty} f(x, y) dx \quad f.d.p \ marginal \ de \ Y$$

$$\begin{aligned} P(c \leq X \leq d) &= P(c \leq X \leq d \ , -\infty < Y < +\infty) \\ &= \int_{c}^{d} \int_{-\infty}^{+\infty} f(x, y) dy dx = \int_{c}^{d} g(x) dx \end{aligned}$$

Exemplo 5.12. *Dois característicos do desenpenho do motor de um foguete são o empuxo X e a taxa* *de mistura Y. Suponha que (X, Y) seja uma variável aleatória com f.d.p conjunta dada por:*

$$f(x, y) = 2(x + y - 2xy), \quad 0 \leq x \leq 1 \;\; 0 \leq y \leq 1$$

Encontrar as f.d.p's marginais de X e Y.

$$\begin{aligned} g(x) &= \int_{-\infty}^{+\infty} f(x, y)dy = \int_0^1 2(x + y - 2xy)dy = 2(xy + \frac{y^2}{2} - xy^2)|_0^1 \\ &= 2(x + \frac{1}{2} - x) = 1. \end{aligned}$$

Portanto, $g(x) = 1, \quad 0 \leq x \leq 1, \Rightarrow \quad X$ *tem distribuição* $U(0, 1)$.

$$\begin{aligned} h(y) &= \int_{-\infty}^{+\infty} f(x, y)dx = \int_0^1 2(x + y - 2xy)dx \\ &= 2(\frac{x^2}{2} + yx - x^2y)|_0^1 = 2(\frac{1}{2} + y - y) = 1. \end{aligned}$$

$$h(y) = 1 \quad ; \quad 0 \leq y \leq 1.$$

Portanto, $h(x) = 1, \quad 0 \leq x \leq 1, \Rightarrow \quad Y$ *tem distribuição* $U(0, 1)$.

Definição 5.5. *Dizemos que a variável aleatória contínua bidimensional é uniformemente distribuída sobre a região R do plano euclidiano quando:*

$$f(x, y) = \begin{cases} \frac{1}{Area(R)} & (x, y) \in R \\ 0 & c.c. \end{cases}$$

Exemplo 5.13. *Suponha que a variável aleatória (X, Y) seja uniformemente distribuída sobre a região R definida por $R = \{(x, y) : 0 < x < 1, 0 < y < 2\}$. Encontre a fdp conjunta de (X, Y) e as fdp's marginais de X e Y.*

$$f(x, y) = \frac{1}{Area(R)}$$

$$Area\ (R) = \int_0^2 \int_0^1 dxdy = 2$$

$$Logo: \quad f(x, y) = \frac{1}{2} \quad , \quad 0 < x < 1, \quad 0 < y < 2$$

$$g(x) = \int_{-\infty}^{+\infty} f(x, y)dy = \int_0^2 \frac{1}{2}dy = 1, \quad 0 < x < 1$$

$$h(y) = \int_{-\infty}^{+\infty} f(x, y)dx = \int_0^1 \frac{1}{2}dx = \frac{1}{2}, \quad 0 < y < 2$$

Exemplo 5.14. *Suponha que a variável aleatória* (X, Y) *seja uniformemente distribuída sobre a região* R *definida por* $R = \{(x, y) : 0 < x < 1, x^2 < y < x\}$. *Encontre a f.d.p conjunta de* (X, Y) *e as fdp's marginais de* X *e* Y.

$$f(x, y) = \frac{1}{Area(R)}$$

$$Area\ (R) = \int_0^1 \int_{x^2}^{x} dydx = \int_0^1 (x - x^2)dx = \left[\frac{x^2}{2} - \frac{x^3}{3}\right]_0^1 = \frac{1}{2} - \frac{1}{3} = \frac{1}{6}$$

$$Logo: \quad f(x, y) = 6 \quad , \quad 0 < x < 1 \quad , \quad x^2 < y < x$$

$$g(x) = \int_{-\infty}^{+\infty} f(x, y)dy = \int_{x^2}^{x} 6dy = 6(x - x^2)$$

$$h(y) = \int_{-\infty}^{+\infty} f(x, y)dx = \int_{y}^{\sqrt{y}} 6dx = 6(\sqrt{y} - y)$$

$$g(x) = 6(x^2 - x) \quad ; \quad 0 \leq x \leq 1$$

$$h(y) = 6(\sqrt{y} - y) \quad ; \quad 0 \leq y \leq 1$$

Para obter a constante $1/6$ *via simulação, basta usarmos o código abaixo.*

```
Sub Constante6()
n = 1000000
Soma1 = 0
Soma2 = 0
For i = 1 To n
    x = Rnd()
    y = Rnd()
    fxy = 1
    Soma1 = Soma1 + fxy
    If (y > x ˆ 2 And y < x) Then Soma2 = Soma2 + fxy
Next
MsgBox "Constante = " & Soma2 / Soma1
End Sub
```

5.3 Distribuição de Probabilidade Condicional

a) Caso Discreto:

Função de probabilidade condicional de X *dado* $Y = y_j$.

$$p(x_i \mid y_j) = P(X = x_i \mid Y = y_j) = \frac{P(X = x_i \;\; Y = y_j)}{P(Y = y_j)} = \frac{p(x_i \;\; y_j)}{q(y_j)} \quad se \quad q(y_j) > 0.$$

Função de probabilidade condicional de Y *dado* $X = x_i$.

$$q(y_j \mid x_i) = P(Y = y_j \mid X = x_i) = \frac{P(X = x_i \;\; Y = y_j)}{P(X = x_i)} = \frac{p(x_i \;\; y_j)}{p(x_i)} \quad se \quad p(x_i) > 0.$$

OBS: *Para um dado* j, $p(x_i \mid y_j)$ *satisfaz todas as condições de uma distribuição de probabilidade:*

i) $p(x_i \mid y_j) \geq 0$

ii) $\sum_{i=1}^{\infty} p(x_i \mid y_j) = \sum_{i=1}^{\infty} \frac{p(x_i, y_j)}{q(y_j)} = \frac{1}{q(y_j)} \sum_{i=1}^{\infty} p(x_i \;\; y_j) = \frac{1}{q(y_j)} q(y_j) = 1$

b) Caso contínuo: *Seja (X, Y) uma variável aleatória contínua bidimensional com f.d.p marginais de X e Y representadas por $h(x)$ e $g(y)$, respectivamente.*

• *A f.d.p condicional de X, dado $Y = y$, é definida por:*

$$g(x \mid y) = \frac{f(x,y)}{h(y)} \quad , \quad h(y) > 0$$

A f.d.p condisional de Y, dado $X = x$, é definida por:

$$h(x \mid Y) = \frac{f(x,y)}{g(x)} \quad , \quad g(x) > 0$$

Observação 5.2. *As f.d.p's condicionais satisfazem todas as condições para uma f.d.p unidimensional. Para Y fixado, temos:*

i) $g(x \mid y) \geq 0$

ii) $\int_{-\infty}^{+\infty} g(x \mid y)dx = \int_{-\infty}^{+\infty} \frac{f(x,y)}{h(y)} dx = \frac{1}{h(y)} \int_{-\infty}^{+\infty} f(x,y)dx = \frac{1}{h(y)} h(y) = 1$

Exemplo 5.15. *Considerando a densidade conjunta abaixo, obter as densidades marginais e condicionais.*

$$f(x,y) = 6 \quad , \quad 0 < x < 1, \quad x^2 < y < x$$

$$\begin{aligned}
g(x) &= 6(x^2 - x) \quad ; \quad 0 \leq x \leq 1 \\
h(y) &= 6(\sqrt{y} - y) \quad ; \quad 0 \leq y \leq 1 \\
g(x \mid y) &= \frac{f(x,y)}{h(y)} = \frac{6}{6(\sqrt{y} - y)} = \frac{1}{\sqrt{y} - y} \quad ; \quad y < x < \sqrt{y} \\
h(y \mid x) &= \frac{f(x,y)}{g(x)} = \frac{6}{6(x - x^2)} = \frac{1}{x - x^2} \quad ; \quad x^2 < y < x, \quad 0 < x < 1.
\end{aligned}$$

Exercício 5.3.1. *Considerando a densidade conjunta abaixo, obter as densidades marginais e condicionais.*

$$f(x,y) = \begin{cases} x^2 + \frac{xy}{3} & 0 \leq x \leq 1, 0 \leq y \leq 2 \\ 0, & c.c. \end{cases}$$

Exemplo 5.16. *Retiram-se duas cartas de um baralho. Sejam $X = n^o$ de azes obtidos e $Y = n^o$ de damas obtidas.*

a) *Distribuição Conjunta de (X, Y)*
b) *Distribuição Marginal de X*

Y/X	0	1	2	Total
0	0,714	0,133	0,004	0,851
1	0,133	0,012	0	0,145
2	0,004	0	0	0,004
Total	0,851	0,145	0,004	1,00

x_i	0	1	2	Total
$p(x_i)$	0,851	0,145	0,004	1,00

c) *Distribuição Marginal de Y*

$P(x \mid y) =?$

$$y = 0 \quad \Longrightarrow p(x \mid y = 0) =?$$

$$P(x = 0 \mid y = 0) = \frac{p(0,0)}{q(0)} = \frac{0,714}{0,851} = 0,839$$
$$P(X = 1 \mid y = 0) = \frac{p(1,0)}{q(0)} = \frac{1,33}{0,851} = 0,156$$
$$P(X = 2 \mid y = 0) = \frac{p(2,0)}{q(0)} = \frac{0,004}{0,851} = 0,005$$

Distribuição Condicional de X, dado $Y = 0$:
Distribuição Condicional de X, dado $Y = 1$:
Distribuição Condicional de X, dado $Y = 2$:

$$P(X = 0 \mid Y = 2) = 1$$

Distribuição condicional de Y, dado $X = 0$:
Distribuição condicional de Y, dado $X = 1$:
Distribuição condicional de Y, dado $X = 2$:

$$P(Y = 0 \mid X = 2) = \frac{0,004}{0,004} = 1$$

5.4 Independência

Definição 5.6 (Variáveis aleatórias independentes)**.**

a) Seja (X, Y) uma v.a ***discreta*** *bidimensional. Diremos que X e Y são v.a's independentes se, e somente se, $P(x_i, y_i) = p(x_i)q(y_j)$ para quaisquer $i\, e j$. Isto é, $P(X = x_i, Y = y_j) = P(X = x_i)P(Y = y_j)$ para todo $i\, e j$.*

y_i	0	1	2	Total
$p(y_i)$	0,851	0,145	0,004	1,00

$p(x \mid y=0)$	0	1	2	Total
$p(x \mid y=0)$	0,839	0,156	0,005	1,00

b) Seja (X, Y) uma v.a ***contínua*** *bidimensional. Diremos que X e Y são v.a's independentes se, e somente se, $f(x, y) = g(x)h(y)$ para todo (x, y), onde f é a f.d.p conjunta, e g e h são as f.d.p marginais de X e Y, respectivamente.*

Exemplo 5.17. *1. Suponha que uma máquina seja utilizada para determinada tarefa durante a manhã e para uma tarefa diferente durante a tarde. Sejam X e Y, respectivamente, o n^o de vezes que a máquina pára por defeito de manhã e à tarde. A tabela a seguir dá a distribuição de probabilidade conjunta de (X, Y). Verifique se X e Y são inpendentes.*

$Y \backslash X$	*0*	*1*	*2*	*Total*
0	*0,1*	*0,2*	*0,2*	*0,5*
1	*0,04*	*0,08*	*0,08*	*0,2*
2	*0,06*	*0,12*	*0,12*	*0,3*
Total	*0,20*	*0,40*	*0,40*	*1,00*

$$p(0,0) = 0,1 = p(0) \times q(0) = 0,2 \times 0,5$$

$$p(0,1) = 0,04 = p(0) \times q(1) = 0,2 \times 0,2$$

$$p(0,2) = 0,06 = p(0) \times q(2) = 0,2 \times 0,3$$

$$p(1,0) = 0,2 = p(1) \times q(0) = 0,4 \times 0,5$$

$$p(1,1) = 0,08 = p(1) \times q(1) = 0,4 \times 0,2$$

$$p(1,2) = 0,12 = p(1) \times q(2) = 0,4 \times 0,3$$

Logo, $p(x_i, y_j) = p(x_i) \times q(y_j), \forall\ i$ e $j \Rightarrow X$ e Y são independentes

Exemplo 5.18. *Sejam X e Y a duração da vida de dois dispositivos eletrônicos. Suponha-se que sua f.d.p conjunta seja dada pela função abaixo. Verifique se Xe Y são independentes.*

$$f(x, y) = e^{-(x+y)} \quad ;\ x \geq 0,\ y \geq 0.$$

$$g(x) = \int_0^{+\infty} e^{-(x+y)} dy = e^{-x} \int_0^{+\infty} e^{-y} dy = e^{-x} \left[-e^{-y} \Big|_0^{\infty} \right] = e^{-x} [0 + 1] = e^{-x}$$

$$\therefore g(x) = e^{-x} \quad ;\ x \geq 0$$

x_i	0	1	Total
$p(x_i \mid y = 1)$	0,917	0,083	1,00

y_i	0	1	2	Total
$q(y_i \mid 0)$	0,839	0,156	0,005	1,00

Da mesma forma,

$$\therefore h(y) = \int_0^{+\infty} e^{-(x+y)}dx = e^{-y}$$

$$h(y) = e^{-y} \ ; \ y \geq 0$$

Logo,

$$f(x,y) = e^{-(x+y)} = e^{-x} \cdot e^{-y} = g(x)h(y) \Rightarrow \quad X \text{ e } Y \text{ são independentes} .$$

Teorema 1 (Definição alternativa de independência)**.**

*a) Seja (X, Y) uma variável aleatória **discreta** bidimensional. Nesse caso, X e Y serão independentes se, e somente se, $p(x_i|y_j) = p(x_i)$ para todo i e j [ou, o que é equivalente se, e somente se, $q(y_j|x_i) = q(y_j)$ para todo i e j].*

*b) Seja (X, Y) uma variável aleatória **contínua** bidimensional. Nesse caso, X e Y serão independentes se, e somente se, $g(x|y) = g(x)$, ou equivalente, se e somente se, $h(y|x) = h(y)$, para todo (x, y)*

Teorema 2. *Seja (X, Y) uma variável aleatória bidimensional. Sejam A e B eventos cuja ocorrência (ou não ocorrência) dependa apenas de X e Y, respectivamente. (Isto é, A é um subconjunto de R_X, o contradomínio de X, enquanto B é um subconjunto de R_Y, o contradomínio de Y). Então, se X e Y forem variáveis aleatórias independentes, teremos $P(A \cap B) = P(A)P(B)$.*

Demonstração (apenas para o caso contínuo):

$$\begin{aligned} P(A \cap B) &= \int\int_{A \cap B} f(x,y)dxdy = \int\int_{A \cap B} g(x)h(y)dxdy \\ &= \int_A g(x)dx \int_B h(y)dy = P(A)P(B). \end{aligned}$$

Teorema 3. *Sejam X e Y duas v.a. independentes. Quaisquer funções isoladas de X e Y também serão independentes. Ou seja, sendo $Z = g(X)$ e $W = h(Y)$, então Z e W serão independentes.*

Exemplo 5.19. *Sejam X e Y duas v.a. independentes. Serão independentes, por exemplo,*

i) $Z = X^a$ e $W = (Y - b)/c$, onde a, b e c são constantes;

y_i	0	1	Total
$q(y_i \mid 1)$	0,917	0,083	1,00

ii) $Z = e^{t_1 X}$ *e* $W = e^{t_2 Y}$ *, onde* t_1 *e* t_2 *são constantes.*

Observação 5.3. *Este teorema é extremamente importante e pode ser generalizado uma uma número maior de variáveis aleatórias. Por exemplo, se temos variáveis* $X_1, X_2, \cdots, X_n$*, então funções de subconjuntos disjuntos de* $X_1, X_2, \cdots, X_n$ *serão independentes, tais como* $Y_1 = g(X_1, X_2)$ *e* $Y_2 = h(X_3, \cdots, X_n)$*,* $n \geq 3$*.*

Uma outra propriedade importante é a Independência Condicional *(também chama da de* Independência Local*), necessária em várias situações para fins de desenvolvimento de métodos de estimação.*

Definição 5.7 (Variáveis aleatórias condicionalmente independentes)**.**

a) Seja (X, Y) *uma v.a.* ***discreta*** *e* Z *uma v.a. Diremos que* X *e* Y *são v.a.'s condicionlmente independentes em* Z *se, e somente se,* $P(x_i, y_j|z_k) = p(x_i|z_k)q(y_j|z_k)$ *para quaisquer* i *e* j*, e cada* k*. Isto é,* $P(X = x_i, Y = y_j, Z = z_k) = P(X = x_i, Z = z_k)P(Y = y_j, Z = z_k)$*.*

b) Seja (X, Y) *uma v.a* ***contínua*** *bidimensional. Diremos que* X *e* Y *são v.a's condicionlmente independentes se, e somente se,* $f(x, y|z) = g(x|z)h(y|z)$ *para todo* (x, y)*, e cada* z*.*

c) A definição é similar para o caso em que X *e* Y *são v.a.'s discretas e* Z *é uma v.a. contínua:* $P(x_i, y_j|z) = p(x_i|z)q(y_j|z)$*.*

d) A generalização é natural para um número maior de v.a.'s: $P(x_1, x_2, \cdots, x_n|z) = \prod_{i=1}^{n} P(x_i|z)$*.*

5.5 Algumas funções de variáveis aleatórias

Seja (X, Y) *uma v.a. e* $Z = H(X, Y)$ *uma função de* $\mathbb{R}^2 \to \mathbb{R}$*. Desejamos obter a distribuição de probabilidade de* Z*. Primeiro temos que observar que* Z *é uma v.a.*

Exemplo 5.20. *Exemplos de funções de v.a.*

$$\begin{aligned} Z &= X + Y, & Z &= X - Y, \\ Z &= XY, & Z &= \frac{X}{Y}, \\ Z &= \min(X, Y), & Z &= \max(X, Y). \end{aligned}$$

Exemplo 5.21. *Duas linhas de produção fabricam um certo tipo de peça. Suponha que a capacidade (em qualquer dia) seja 5 peças na linha I e 3 peças na linha II. Admita que o número de peças realmente produzidas em qualquer linha seja uma variável aleatória, e que* (X, Y) *represente a variável aleatória bidimensional que fornece o número de peças produzidas pela linha I e a linha II, respectivamente.*

A tabela a seguir dá a distribuição de probabilidade conjunta de (X,Y)*. Cada casa representa*

$$p(x_i, y_j) = P(X = x_i, Y = y_j)$$

Assim, $p(2,3) = P(X = 2, Y = 3) = 0,04$ *etc. Portanto, se B for definido como*

B= {Mais peças são produzidas pela linha I que pela linha II} encontraremos que

$P(B) = \boxed{0{,}01+\ 0{,}03+\ 0{,}05+\ 0{,}07+\ 0{,}09} + \boxed{0{,}04+\ 0{,}05+\ 0{,}06+\ 0{,}08} + 0,05 + 0,05 + 0,06 + 0,06 + 0,05 = 0,75.$

Y\X	0	1	2	3	4	5	Total
0	0	$\boxed{0{,}01}$	$\boxed{0{,}03}$	$\boxed{0{,}05}$	$\boxed{0{,}07}$	$\boxed{0{,}09}$	0,25
1	0,01	0,02	0,04	0,05	0,06	0,08	0,26
2	0,01	0,03	0,05	0,05	0,05	0,06	0,25
3	0,01	0,02	0,04	0,06	0,06	0,05	0,24
Total	0,03	0,08	0,16	0,21	0,24	0,28	1,00

Encontre a distribuição de probabilidade das seguintes v.a's:
U = *mín* (X,Y) = *menor* n^o *de peças produzidas pelas duas linhas.*
V = *máx*(X,Y) = *maior* n^o *de peças produziadas pelas duas linhas.*
$W = X + Y = n^o$ *total de peças produzidas pelas duas linhas.*

DISTRIB. DE U = Mín(X,Y)

u_i	0	1	2	3	Total
$p(u_i)$	0,28	0,30	0,25	0,17	1,00

DISTRIB. DE V = Máx(X,Y)

v_i	1	2	3	4	5	Total
$p(v_i)$	0,04	0,16	0,28	0,24	0,28	1,00

DISTRIB. DE W = X + Y

w_i	1	2	3	4	5	6	7	8	Total
$P(w_i)$	0,02	0,06	0,13	0,19	0,24	0,19	0,12	0,05	1,00

5.6 Soma de Variáveis Aleatórias

A soma de variáveis é um dos casos mais importantes para a Estatística, pois dela decorrerão as principais propriedades da média amostral (que é a soma dividida pelo tamanho da amostra), variância amostral e muitos outros estimadores. Mas, por simplicidade, tratemos separadamente os casos discretos e contínuo.

5.6.1 Caso discreto

Seja (X, Y) *uma v.a discreta. Para obtermos a distribuição da variável* $Z = X+Y$ *devemos considerar todas as possibilidades em que a soma dê* z*, ou seja*

$$P(Z = z) = \sum_k P(X = k, Y = z - k) = \sum_k P(X = k)P(Y = z - k). \tag{5.2}$$

Esta expressão é decorrente do fato que os eventos $\{X = k, Y = z - k\}, \forall\ k$*, são mutuamente exclusivos, de forma que vale o apresentado em (5.1).*

Exemplo 5.22. *Sejam* X *e* Y *v.a.i. com distribuições Bin(*n_1, p*) e Bin(*n_2, p*), respectivamente. Determinar a distribuição de* $Z = X + Y$*.*

Vale notar que z *é um valor inteiro e que a v.a.* X *só pode assumir valores de 0 a* n_1*. Assim, para* $z = 0, 1, 2, \cdots, n_1 + n_2$*, temos*

$$\begin{aligned}
P(Z = z) &= \sum_{k=0}^{z} P(X = k)P(Y = z - k) \\
&= \sum_{k=0}^{z} \binom{n_1}{k} p^k (1-p)^{n_1-k} \times \binom{n_2}{z-k} p^{z-k}(1-p)^{n_2-(z-k)} \\
&= p^z (1-p)^{n_1+n_2-z} \sum_{k=0}^{z} \binom{n_1}{k}\binom{n_2}{z-k} \\
&= \binom{n_1+n_2}{z} p^z (1-p)^{n_1+n_2-z}.
\end{aligned}$$

Portanto, a soma de v.a. binomiais de mesmo parâmetro p também é Binomial. Esse fato é bastante intuitivo, pois como X *representa o número de sucessos em* n_1 *ensaios independentes de Bernoulli (tipo lançamentos de uma moeda), todos com mesma probabilidade de sucesso p, e sendo* Y *o número de sucessos em* n_2 *ensaios independentes de Bernoulli com o mesmo parâmetro de sucesso p, então* Z *será o número de sucessos em* $n_1 + n_2$ *ensaios independentes de Bernouli de mesmo parâmetro p, que é Binomial de parâmetros* $n_1 + n_2$ *e p.*

Exemplo 5.23. *Sejam* X *e* Y *v.a.i. com distribuições Poisson(*λ_1*) e Poisson(*λ_2*), respectivamente. Determinar a distribuição de* $Z = X + Y$*.*

Inicialmente devemos notar que z *é um valor inteiro e que a v.a.* X *só pode assumir valores de 0 a* z*. Assim, para* $k = 0, 1, 2, \cdots, z$*, temos*

$$
\begin{aligned}
P(Z=z) &= \sum_{k=0}^{z} P(X=k)P(Y=z-k) \\
&= \sum_{k=0}^{z} \frac{e^{-\lambda_1}\lambda_1^k}{k!} \times \frac{e^{-\lambda_2}\lambda_2^{z-k}}{(z-k)!} \\
&= e^{-(\lambda_1+\lambda_2)} \sum_{k=0}^{z} \frac{1}{k!(n-k)!}\lambda_1^k\lambda_2^{z-k} \\
&= \frac{e^{-(\lambda_1+\lambda_2)}}{z!} \sum_{k=0}^{z} \binom{z}{k}\lambda_1^k\lambda_2^{z-k} \\
&= \frac{e^{-(\lambda_1+\lambda_2)}(\lambda_1+\lambda_2)^z}{z!}
\end{aligned}
$$

Portanto, a soma de v.a. Poissons também é Poisson, com parâmetro igual à soma dos parâmetros. Esse fato também é bastente intuitivo, mas por quê?

5.6.2 Caso contínuo

De forma similar ao caso de variáveis discretas, a soma de v.a. pode ser obtida "somando-se"(que neste caso é a integral) para todos os valores x de X, ou seja,

$$g(z) = \int_{-\infty}^{\infty} f_X(x) f_Y(z-y)dx \tag{5.3}$$

Exemplo 5.24. *Sejam $X \sim Exp(\lambda_1)$ e $Y \sim Exp(\lambda_2)$, independentes, qual a distribuição de $Z = X+Y$?*

Exemplo 5.25. *Sejam $X \sim N(0,1)$ e $Y \sim N(0,1)$, independentes, qual a distribuição de $Z = X+Y$?*

Exercício 5.6.1. *Sejam $X \sim N(\mu_1, \sigma_1^2)$ e $Y \sim N(\mu_2, \sigma_2^2)$, independentes, qual a distribuição de $Z = X+Y$?*

No geral, se (X,Y) for uma v.a contínua e se $Z = H_1(X,Y)$ for uma função contínua de (X,Y), então Z será uma v.a contínua. Qual a f.d.p de Z? Em alguns casos (soma, diferença, por exemplo) podemos obtê-la diretamente por (5.3), mas em outros tem que haver um procedimento mais geral, que também vale nos casos citados.

Observação 5.4. *A solução é a versão bidimensional da transformação de variáveis do caso unidimensional, em que temos uma v.a. X com fdp $f(x)$ e uma transformada $Y = H(X)$. Então, a fdp g de Y é dada por $g(y) = f(x)\left|\frac{dx}{dy}\right|$.*

Procedimento para obter a fdp de $Z = H_1(X,Y)$:

1º) Introduzir uma segunda v.a: $W = H_2(X,Y)$;

2º) Obter a f.d.p conjunta de Z e W (vamos denominar esta conjunta de $K(z,w)$)

3º) Integra-se a f.d.p conjunta $K(z,w)$ com relação a W e obtém-se a f.d.p de Z:

$$g(z) = \int_{-\infty}^{+\infty} K(z,w)dw$$

Problemas:

1. *Como escolher a v.a W apropriada?*
2. *Como encontrar $K(z,w)$?*

- *Deve-se fazer a escolha mais simples para W;*

- *O Teorema a seguir indica como encontrar $K(z,w)$.*

Teorema 4. *Suponha que (X,Y) seja uma variável aleatória contínua bidimensional com f.d.p conjunta f. Sejam $Z = H_1(X,Y)$ e $W = H_2(X,Y)$, e admitamos que as funções H_1 e H_2 satisfaçam às seguintes condições:*

a) As equações $z = H_1(x,y)$ e $w = H_2(x,y)$ podem ser univocamente resolvidas para x e y, em termos de z e w, isto é, $x = G_1(z,w)$ e $y = G_2(z,w)$.

b) As derivadas parciais $\partial x/\partial z$, $\partial x/\partial w$, $\partial y/\partial z$ e $\partial y/\partial w$ existem e são contínuas.

Nessas circunstâncias, a f.d.p conjunta de (Z,W), isto é, $k(z,w)$ é dada pela seguinte expressão: $k(z,w) = f(x,y)\ |J(z,w)|$, onde $J(z,w)$ é o determinante 2×2:

$$J(z,w) = \begin{vmatrix} \frac{\partial x}{\partial z} & \frac{\partial x}{\partial w} \\ \frac{\partial y}{\partial z} & \frac{\partial y}{\partial w} \end{vmatrix}$$

Este determinante é denominado oJacobiano da transformação $(x,y) \to (z,w)$ e, algumas vezes, é denotado por $\partial(x,y)/\partial(z,w)$. Salientamos que $k(z,w)$ será não-nula para aqueles valores de (z,w) correspondentes a valores de (x,y) para os quais $f(x,y)$ não seja nula.

Observação 5.5.

a) Muito embora não demonstremos este teorema, indicaremos ao menos o que se deseja e onde residem as dificuldades. Consideremos a f.d. conjunta da variável aleatória bidimensional (Z,W), isto é,

$$K(z,w) = P(Z \leq z, W \leq w) = \int_{-\infty}^{w} \int_{-\infty}^{z} k(s,t)dsdt,$$

na qual k é a f.d.p procurada. Como se supõe que a transformação $(x,y) \to (z,w)$ seja biunívoca [veja a hipótese (a), acima], poderemos achar o evento, equivalente a $\{z \leq z, W \leq w\}$, em termos de X e Y. Suponhamos que este evento seja denotado por C. Sendo assim, $\{(X,Y) \in C\}$se, e *somente se,* $\{Z \leq z, W \leq w\}$. *Consequentemente,*

$$\int_{-\infty}^{w}\int_{-\infty}^{z} k(s,t)dsdt = \iint_{C} f(x,y)dxdy$$

Como se admite f conhecida, a integral do segundo membro pode ser calculada. O cálculo de suas derivadas em relação a z e w fornecerá a fdp pedida. Na maior parte dos manuais de cálculo avançado, mostra-se que essas técnicas conduzem ao resultado, tal como foi enunciado no teorema acima.

b) *Observe-se a acentuada semelhança entre o resultado acima e o resultado obtido no caso unidimensional, explicado no capítulo anterior. A exigência de monotonicidade para a função $y = H(x)$ é substituída pela suposição de que a correspondência entre (x, y) e (z, w) seja biunívoca. A condição de derivabilidade é substituída por algumas hipóteses sobre as derivadas parciais consideradas. A solução final obtida é, também, muito semelhante aquela obtida no caso unidimensional: as variáveis x e y são simplesmente substituídas por suas expressões equivalentes em termos de z e w, e o valor absoluto de dx/dy é substituído pelo valor absoluto do jacobiano.*

Para o caso da função soma, $Z = X + Y$, geralmente escolhemos $W = X$. Neste caso teremos

$$\begin{cases} Z = X+Y \\ W = X \end{cases} \Rightarrow \begin{cases} x = w \\ y = z - w \end{cases} \qquad \begin{cases} x = G_1(z,w) \\ y = G_2(z,w) \end{cases}$$

Portanto, $J = \begin{vmatrix} \frac{\partial x}{\partial z} & \frac{\partial x}{\partial w} \\ \frac{\partial y}{\partial z} & \frac{\partial y}{\partial w} \end{vmatrix} = \begin{vmatrix} 0 & 1 \\ 1 & -1 \end{vmatrix} = -1$

$K(z,w) = f(G_1(z,w), G_2(z,w)) \times |J| = f(w,\ z-w) \times |-1| = f(w,\ z-w)$
Com isso, a densidade de Z é dada por

$$g(z) = \int_{-\infty}^{+\infty} K(z,w)dw = \int_{-\infty}^{+\infty} f(x, z-x)dx.$$

que é a mesma expressão apresentada em (5.3).

Exemplo 5.26. *Sejam X e Y v.a's independentes, cada uma tendo distribuição uniforme no intervalo $(0,1)$. Seja $Z = X + Y$. Encontre a f.d.p de Z.*

$$X \sim U(0,1) \quad \Rightarrow \quad f_X(x) = \begin{cases} 1, & x \in (0,1) \\ 0, & c.c. \end{cases}$$

$$Y \sim U(0,1) \quad \Rightarrow \quad f_Y(y) = \begin{cases} 1, & y \in (0,1) \\ 0, & c.c. \end{cases}$$

$$X \text{ e } Y \text{ independentes} \quad \Rightarrow \quad f(x,y) = \begin{cases} 1; & 0 \leq x \leq 1,\ 0 \leq y \leq 1 \\ 0; & c.c. \end{cases}$$

$$\begin{cases} Z = X+Y \\ W = X \end{cases} \Rightarrow \begin{cases} x = w \\ y = z - w \end{cases} \qquad \begin{cases} x = G_1(z,w) \\ y = G_2(z,w) \end{cases}$$

$$J = \begin{vmatrix} \frac{\partial x}{\partial z} & \frac{\partial x}{\partial w} \\ \frac{\partial y}{\partial z} & \frac{\partial y}{\partial w} \end{vmatrix} = \begin{vmatrix} 0 & 1 \\ 1 & -1 \end{vmatrix} = -1$$

$$K(z,w) = f(G_1(z,w), G_2(z,w)) \times |J|$$

$$= f(w,\ z-w) \times |-1|$$

$$K(z,w) = \begin{cases} 1; & 0 \le w \le 1\,,\ 0 \le z-w \le 1 \\ & ,\quad w \le z \le 1+w \\ 0; & c.c \end{cases}$$

$$g(z) = \int_{-\infty}^{+\infty} K(z,w)dw$$
$$= \begin{cases} \int_0^z K(z,w)dw & ; \quad 0 \le z \le 1 \\ \int_{z-1}^1 K(z,w)dw & ; \quad 1 < z \le 2 \end{cases}$$

$$= \begin{cases} \int_0^z dw & ; \quad 0 \le z \le 1 \\ \int_{z-1}^1 dw & ; \quad 1 < z \le 2 \end{cases}$$

$$= \begin{cases} z & ; \quad 0 \le z \le 1 \\ 2-z & ; \quad 1 < z \le 2 \end{cases}$$

Exemplo 5.27. *Suponha-se que estejamos fazendo mira em um alvo circular, de raio unitário, que tenha sido colocado de modo que seu centro se situe na origem de um sistema de coordenadas regulares conforme a figura. Admita-se que as coordenadas (X,Y) do ponto de impacto estejam uniformemente distribuídas sobre o circulo. Isto é, $f(x,y) = 1/\pi$, se (x,y) estiver dentro (ou na circunferência) do círculo, $f(x,y) = 0$, se em qualquer outra parte. Obter a densidade da variável aleatória R, que representa a distância da origem, ou seja, $R = \sqrt{X^2+Y^2}$.*

Encontraremos a f.d.p de R, digamos g, assim:

Seja $\Phi = tg^{-1}(Y/X)$. Portanto, $X = H_1(R,\Phi)$ e $Y = H_2(R,\Phi)$, onde $x = H_1(r,\phi) = r\ cos\ \phi$ e $y = H_2(r,\phi) = r\ sen\ \phi$. (Estamos apenas introduzindo coordenadas polares).

$$f(x,y) = \begin{cases} \frac{1}{\pi}, & se \quad (x,y) \text{ é círculo;} \\ 0, & se \quad c.c. \end{cases}$$

$R = \sqrt{X^2+Y^2}$ *(Distância da Origem)*

$g(r) = ?$

$$\begin{cases} \Phi = & arctg(Y/X); \\ R = & \sqrt{X^2+Y^2}. \end{cases}$$

Encontrar a f.d.p conjunta de ϕ e R.

$$\begin{cases} \frac{y}{x} = tg\,\phi & \qquad x = G_1(\phi, r) \\ X^2+Y^2 = r^2 & \qquad y = G_2(\phi, r) \end{cases}$$

$y = x\,tg\,\phi$

$$\begin{aligned} x^2 + (x\,tg\,\phi)^2 = r^2 &\Rightarrow x^2 + x^2 tg^2\,\phi = r^2 \\ &\Rightarrow x^2(1 + tg^2\,\phi) = r^2 \\ &\Rightarrow x^2 = \frac{r^2}{1 + tg^2\,\phi} \end{aligned}$$

No entanto,

$$\begin{aligned} &1 + tg^2\,\phi = 1 + \frac{sen^2\,\phi}{cos^2\,\phi} = \frac{1}{cos^2\,\phi} \\ \Rightarrow\ & x^2 = r^2 cos^2\,\phi \\ \Rightarrow\ & x = r\,cos\phi \end{aligned}$$

$$\begin{aligned} y = x\,tg\,\phi \quad &\Rightarrow \quad y = r\,cos\,\phi\,tg\,\phi \\ &\Rightarrow \quad r\,cos\,\phi \tfrac{sen\,\phi}{cos\,\phi} \\ &\Rightarrow \quad y = r\,sen\,\phi \end{aligned}$$

$$\begin{cases} x = & r\,cos\,\phi \\ y = & r\,sen\,\phi \end{cases} \qquad \textit{coordenadas polares}$$

Jacobianos:

$$J = \begin{vmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \phi} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \phi} \end{vmatrix} = \begin{vmatrix} cos\phi & -rsen\phi \\ sen\phi & rcos\phi \end{vmatrix} = rcos^2\phi + rsen^2\phi = r(cos^2\phi + sen^2) = r$$

$$K(\phi, r) = f(G_1(\phi, r), G_2(\phi, r)) \cdot |r| = f(rcos\phi, rsen\,\phi) \cdot r = \tfrac{1}{\pi} \cdot r$$

Portanto,
$K(\phi, r) = \frac{r}{\pi} \qquad 0 \le r \le 1 \quad , 0 \le \phi \le 2\pi$

$$g(r) = \int_{-\infty}^{+\infty} K(\phi, r) d\phi = \int_0^{2\pi} \frac{r}{\pi}\, d\phi \ = \ \frac{r}{\pi}\,\phi\,\Big|_0^{2\pi} \ = \ \frac{r}{\pi}\,2\pi - 0 = 2r$$

Concluindo que

$$g(r) = \begin{cases} 2r, & 0 \le r \le 1 \\ 0, & c.c. \end{cases}$$

5.7 Produto de Variáveis Aleatórias Independentes

Teorema 5. *Seja (X, Y) uma v.a contínua bidimensional e admita-se que X e Y sejam independentes. Seja $V = XY$. Neste caso, a f.d.p de V, digamos p, é dada por:*

$$p\,(v) = \int_{-\infty}^{+\infty} g(w)\ h\left(\frac{v}{w}\right)\ \left|\frac{1}{w}\right|\ dw.$$

Demonstração:

$$\left\{\begin{array}{l} v = xy \\ w = x \end{array}\right. \quad \Rightarrow \quad \left\{\begin{array}{l} x = w \\ y = \frac{v}{w} \end{array}\right.$$

Logo, o Jacobiano é

$$J = \left|\begin{array}{cc} \frac{\partial x}{\partial w} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial w} & \frac{\partial y}{\partial v} \end{array}\right| \quad = \quad \left|\begin{array}{cc} 1 & 0 \\ -\frac{v}{w^2} & \frac{1}{w} \end{array}\right| = \frac{1}{w}$$

Portanto, a f.d.p conjunta de $V = XY$ e $W = X$ é dada por:

$$K(v,w) = f\left(w\,,\,\frac{v}{w}\right)\cdot|J\,| = g(w)\cdot h\left(\frac{v}{w}\right)\cdot\left|\frac{1}{w}\right|$$

A f.d.p marginal de V, será

$$p(v) = \int_{-\infty}^{+\infty} K(v,w)dw = \int_{-\infty}^{+\infty} g(w)h\left(\frac{v}{w}\right)\cdot\left|\frac{1}{w}\right|\,dw$$

Observação 5.6.

1. Os valores de v para os quais $p(v) > 0$ dependeram dos valores de (x,y) para os quais $f(x,y) > 0$

2. Temos que

$$\int_{-\infty}^{+\infty} g(w)h\left(\frac{v}{w}\right)\cdot\left|\frac{1}{w}\right|\,dw = \int_{0}^{+\infty} g(w)h\left(\frac{v}{w}\right)\cdot\frac{1}{w}\,dw - \int_{-\infty}^{0} g(w)h\left(\frac{v}{w}\right)\cdot\frac{1}{w}\,dw$$

Exemplo 5.28. *Suponha que temos um circuito no qual tanto à corrente I como a resistência R variem de algum modo aleatório. Particularmente, suponhamos que I e R sejam variáveis aleatórias contínuas independentes com as com as fdp's abaixo. Obter a distribuição de $E = IR$.*

$$g(i) = \left\{\begin{array}{ll} 2i\ ; & 0 \leq i \leq 1 \\ 0\ \ ; & c.c \end{array}\right.$$

$$h(r) = \left\{\begin{array}{ll} \frac{r^2}{9}\ ; & 0 \leq r \leq 3 \\ 0\ \ ; & c.c \end{array}\right.$$

A densidade de E será dada por

$$\begin{aligned} p\,(e) &= \int_{-\infty}^{+\infty} g(w)\cdot h\left(\frac{e}{w}\right)\cdot\left|\frac{1}{w}\right| dw \\ &= \int_{-\infty}^{+\infty} g(i)\cdot h\left(\frac{e}{i}\right)\cdot\left|\frac{1}{i}\right| di \\ &= \int_{\frac{e}{3}}^{1} 2i\cdot\frac{e^2}{9\,i^2}\cdot\frac{1}{i}\, di \end{aligned}$$

$$\left.\begin{array}{ll} 0\leq i\leq 1 & \\ 0\leq \frac{e}{i}\leq 3 & \Rightarrow \quad 0\leq e\leq 3i \end{array}\right\} \quad \Rightarrow \quad \tfrac{e}{3}\leq i\leq 1$$

$$p\,(e) = \int_{\frac{e}{3}}^{1}\frac{2e^2}{9\,i^2}\,di = \frac{2e^2}{9}\left[-\frac{1}{i}\Big|_{\frac{e}{3}}^{1}\right] = \frac{2e^2}{9}\cdot\left[-1+\frac{3}{e}\right] = \frac{2e}{9}\,(3-e)$$

Portanto,

$$p\,(e) = \begin{cases} \frac{2e}{9}\,(3-e)\;; & 0\leq e\leq 3 \\ 0 \qquad\qquad\;\; ; & c.c \end{cases}$$

5.8 Quociente de Variáveis Aleatórias Independentes

Teorema 6. *Seja (X,Y) uma variável aleatória bidimensional contínua e suponhamos que X e Y sejam independentes. [Portanto, a fdp de (X,Y) pode ser escrita como $f(x,y) = g(x)h(y)$]. Seja $Z = \frac{X}{Y}$. Deste modo, a fdp de Z, digamos q, será dada por*

$$q(z) = \int_{-\infty}^{+\infty} g(vz)h(v)|v|dv$$

Demonstração:

Sejam $z = x/y$ e $v = y$. Portanto, $x = vz$ e $y = v$. O jacobiano é

$$J = \begin{vmatrix} v & z \\ 0 & 1 \end{vmatrix} = v$$

Daí a fdp conjunta de $Z = X/Y$ e $V = Y$ ser igual a

$$t(z,v) = g(vz)h(v)|v|$$

Integrando esta fdp conjunta em relação a v obtém-se a fdp marginal de Z procurada.

Exemplo 5.29. *Admita-se que X e Y representem a duração da vida de duas lâmpadas fabricadas por processos diferentes. Suponha-se que X e Y sejam variáveis aleatórias independentes, com fdp respectivamente f e g dadas a seguir. Obter a fdp de X/Y.*

$f(x) = e^{-x}$, $x \geq 0$ *e 0 para outros quaisquer valores;*
$g(y) = 2e^{-2y}$, $y \geq 0$ *e 0 para outros valores;*

A variável aleatória X/Y representa o quociente das duas durações de vida. Seja q a fdp de Z. Pelo teorema temos que $q(z) = \int_{-\infty}^{+\infty} g(vz)h(v)|v|dv$. Como X e Y podem tomar somente valores não-negativos, a integração precisa ser feita apenas sobre os valores positivos da variável de integração. Além disso, o integrando será positivo somente quando ambas as fdp que aparecem sejam positivas. Isto significa que deveremos ter $v \geq 0$ e $vz \geq 0$. Visto que $z > 0$, essas desigualdades determinam que $v \geq 0$. Portanto, a expressão acima se torna

$$q(z) = \int_0^{\infty} e^{-vz}\ 2e^{-2v}\, v\, dv =$$
$$= 2\int_0^{\infty} ve^{-v(2+z)}\ dv$$

Integrando por partes, obtemos:

$$q(z) = \frac{2}{(z+2)^2}\ , \qquad z \geq 0$$

5.9 Distribuição do Mínimo e do Máximo de duas v.a's contínuas Independentes

Em algumas situações práticas temos interesse em estudar um sistema formado por vários componentes, em série ou paralelo. Se estiverem em série, o sistema falhará quando o primeiro falhar (mínimo), mas se estiver em paralelo, falhará quando o último falhar (máximo). Estas distribuições são muitas vezes denominadas de Estatísticas de Ordem*.*

5.9.1 Máximo

Sejam X e Y v.a's independentes, com f.d.p's dadas, respectivamente, por $f_1(x)$ e $f_2(y)$. Nosso objetivo é encontrar Encontre a f.d.p de $M = \max(X,Y)$, digamos $g(m)$.

Devemos inicialmente observar que se o máximo entre um conjunto de números é menor que m, então todos os números serão menores que m. Estes argumentos levam ao uso da Função de Distribuição, e posteriormente a função de probabilidade ou densidade. Com base nisso, a f.d de M será:

$$\begin{aligned} G(m) &= P(M \leq m) = P(Max(X,Y) \leq m) = P(X \leq m\ \ Y \leq m) \\ &= P(X \leq m)P(Y \leq m) = F_1(m)F_2(m) \end{aligned}$$

Logo a f.d.p de M será obtida derivando-se $G(m)$ com relação a m:

$$g(m) = \frac{\partial G(m)}{\partial m} = F_1^{'}F_2(m) + F_1(m)F_2^{'}(m) = f_1(m)F_2(m) + f_2(m)F_1(m)$$

Portanto,

$$g(m) = f_1(m)F_2(m) + f_2(m)F_1(m)$$

Observação 5.7. *Se X e Y, além de independentes tiverama amesma distribuição de probabilidade, a f.d.p de M tornam-se:*

$$g(m) = 2f(m)F(m),$$

com $f(m) = f_1(m) = f_2(m)$ e $F(m) = F_1(m) = F_2(m)$.

5.9.2 Mínimo

De forma similar á distribuição do máximo, consideraremos X e Y v.a's independentes, com f.d.p's dadas, respectivamente, por $f_1(x)$ e $f_2(y)$. Nosso objetivo é encontrar a f.d.p de $Z = \min(X, Y)$, digamos $h(z)$.

Devemos inicialmente observar que se o mínimo entre um conjunto de números é maior que z, então todos os números serão maiores que z. Estes argumentos levam ao uso da Função de Distribuição, e posteriormente a função de probabilidade ou densidade. Com base nisso, a f.d de Z será:

$$\begin{aligned} H(z) &= P(Z \leq z) = P(Min(X,Y) \leq z) = 1 - P(Min(X,Y) > z) \\ &= 1 - P(X > z \ \ Y > z) = 1 - P(X > z)P(Y > z) = 1 - [1 - F_1(z)][1 - F_2(z)] \\ &= 1 - [1 - F_2(z) - F_1(z) + F_1(z)F_2(z)] = F_1(z) + F_2(z) - F_1(z)F_2(z). \end{aligned}$$

Com isso,

$$\begin{aligned} h(z) &= \frac{\partial H(z)}{\partial z} = F_1^{'}(z) + F_2^{'}(z) - [F_1^{'}(z)F_2(z) + F_1(z)F_2^{'}(z)] \\ &= f_1(z) + f_2(z) - f_1(z)F_2(z) - F_1(z)f_2(z) \\ &= f_1(z)[1 - F_2(z)] + f_2(z)[1 - F_1(z)] \end{aligned}$$

Observação 5.8. *Se X e Y tiverem a mesma distribuição de probabilidade, então :*

$$h(z) = 2f(z)[1 - F(z)],$$

onde $f(z) = f_1(z) = f_2(z)$ e $F(z) = F_1(z) = F_2(z)$.

5.10 Distribuições Condicionais

Em certas situações desejamos obter a distribuição de uma função de variáveis, dada uma determinada condição, ou no condicionamento pode estar uma função. Vejamos alguns casos:

Exemplo 5.30. *Sejam $X \sim Bin(n, p)$ e $Y \sim Bin(n, p)$, independentes. Então a distribuição de X, dado que $X + Y = m$ é Hipergeométrica, com*

$$P(X = k | X + Y = m) = \frac{\binom{n}{k}\binom{n}{m-n}}{\binom{2n}{m}}.$$

Pelo Exemplo 5.22 já temos que $X+Y \sim Bin(2n,p)$. Portanto, para $0 \leq k \leq min(n,n)$,

$$\begin{aligned} P(X=k|X+Y=n) &= \frac{P(X=k, X+Y=n)}{P(X+Y=n)} = \frac{P(X=k, Y=n-k)}{P(X+Y=n)} \\ &= \frac{P(X=k)P(Y=n-k)}{P(X+Y=n)} \\ &= \frac{\binom{n}{k}p^k(1-p)^{n-k} \times \binom{n}{m-k}p^{m-k}(1-p)^{n-(m-k)}}{\binom{2n}{m}p^m(1-p)^{2n-m}} \\ &= \frac{\binom{n}{k}\binom{n}{m-n}}{\binom{2n}{m}}. \end{aligned}$$

Exemplo 5.31. *Sejam $X \sim Poisson(\lambda_1)$ e $Y \sim Poisson(\lambda_2)$, independentes. Então a distribuição de X, dado que $X+Y=n$ é $Bin(n, \frac{\lambda_1}{\lambda_1+\lambda_2})$.*

Já sabemos do Exemplo 5.23 que $X+Y \sim Poisson(\lambda_1+\lambda_2)$. Portanto,

$$\begin{aligned} P(X=k|X+Y=n) &= \frac{P(X=k, X+Y=n)}{P(X+Y=n)} = \frac{P(X=k, Y=n-k)}{P(X+Y=n)} \\ &= \frac{P(X=k)P(Y=n-k)}{P(X+Y=n)} = \frac{e^{-\lambda_1}\lambda_1^k/k! \times e^{-\lambda_2}\lambda_1^{n-k}/(n-k)!}{e^{-(\lambda_1+\lambda_2)}(\lambda_1+\lambda_2)^n/n!} \\ &= \binom{n}{k}\left(\frac{\lambda_1}{\lambda_1+\lambda_2}\right)^k\left(\frac{\lambda_2}{\lambda_1+\lambda_2}\right)^{n-k} \end{aligned}$$

Exercício 5.10.1. *Sejam $X \sim Poisson(\lambda)$ e $Y|(X=x) \sim Bin(x,p)$. Então,*
a) A distribuição de Y é $Poisson(\lambda p)$.
b) A distribuição condicional de $X|(Y=y)$ é $\mathcal{P}oisson(\lambda(1-p))$.

5.11 Variáveis Aleatórias n-dimensionais

Praticamente todos os conceitos introduzidos para o caso bidimensional podem ser facilmente estendidos para o caso em que temos várias variáveis em estudo, por isso faremos um breve resumo sobre este caso. Consideremoa a variável n-dimensional $(X_1, X_2, \cdots, X_n)$.

Seja $(X_1, X_2, \cdots, X_n)$ uma variável aleatória n-dimensional contínua tomando todos os valores em alguma região $I\!R^n$ do espaço auclidiano. Uma função f que satisfaça ás seguintes condições:

i) $f(x_1, \cdots, x_n) \geq 0 \quad \forall(x_1, \cdots, x_n) \in I\!R^n$

ii) $\displaystyle\int \cdots \int f(x_1, \cdots, x_n)dx_1 \cdots dx_n = 1$

é denominada ***Função Densidade de Probabilidade Conjunta*** *de $(X_1, X_2, \cdots, X_n)$.*
Sendo $C \in I\!R^n$, a probabilidade associada à C é dada por

$$P(X_1, X_2, \cdots, X_n) \in C) = \int_C \cdots \int f(x_1, \cdots, x_n)dx_1 \cdots dx_n.$$

As distribuições marginais ou de um subconjunto das variáveis em $(X_1, X_2, \cdots, X_n)$ *podem ser obtidas integrando-se com relação às demais. Por exemplo, para dimensão* $n = 3$*, a marginal de* X_1 *e a conjunta de* X_2 *e* X_3 *são dadas por*

$$f_1(x_1) = \int \int f(x_1, x_2, x_3) dx_2 dx_3$$

$$f_{23}(x_2, x_3) = \int f(x_1, x_2, x_3) dx_1$$

5.11.1 Método do jacobiano para o caso n-dimensional

Teorema 7. *Suponha que* $\boldsymbol{X} = (X_1, X_2, \cdots, X_n)$ *seja uma variável aleatória contínua n-dimensional com f.d.p conjunta* f*. Sejam* $Z_i = H_i(\boldsymbol{X})$*, e admitamos que as funções* H_i *satisfaçam às seguintes condições:*

a) As equações $z = H_i(\boldsymbol{x})$ *podem ser univocamente resolvidas para* $\boldsymbol{x}$ *em termos de* z*, isto é,* $x_i = G_i(\boldsymbol{z})$

b) As derivadas parciais $\partial x_i / \partial z_i$*, existem e são contínuas.*

Nessas circunstâncias, a f.d.p conjunta de $\boldsymbol{Z}$*, isto é,* $k(\boldsymbol{z})$ *é dada pela seguinte expressão:* $k(\boldsymbol{z}) = f(\boldsymbol{x}) \; |J(\boldsymbol{z})|$ *, onde* $J(.)$ *é o determinante* $n \times n$ *:*

$$J(\boldsymbol{z}) = \begin{vmatrix} \frac{\partial x_1}{\partial z_1} & \cdots & \frac{\partial x}{\partial w} \\ \frac{\partial y}{\partial z} & \cdots & \frac{\partial y}{\partial w} \\ \vdots & \ddots & \vdots \\ \frac{\partial y}{\partial z} & \cdots & \frac{\partial y}{\partial w} \end{vmatrix}$$

Capítulo **7**

Valor Esperado de uma função de uma v.a. bidimensional

7.1 Definição

Seja (X, Y) uma variável bidimensional e seja $Z = H(X, Y)$ uma função real de (X, Y). Normalmente, para obtermos o valor esperado de Z temos que encontrar a distribuição da variável para, então, calcular sua esperança. Veremos que esta etapa intermediária, da obtenção da distribuição de Z não é necessária.

a) Se Z for uma v.a discreta, então:

$$E(Z) = \sum_{i=1}^{\infty} z_i p(z_i), \qquad \textit{onde} \quad p(z_i) = P(Z = z_i)$$

b) Se Z for uma v.a Contínua, com f.d.p g, então:

$$E(Z) = \int_{-\infty}^{+\infty} z g(z) dz$$

Exemplo 7.1. *Seja (X, Y) uma v.a com a distribuição conjunta a seguir. Encontre a $E(Z)$, onde $Z = X^2 + Y$.*

Y/X	-1	0	1	Total
0	0,2	0	0,4	0,6
1	0,1	0,1	0,2	0,4
Total	0,3	0,1	0,6	1,0

Distribuição da v.a Z

z_i	1	2	Total
$p(z_i)$	0,7	0,3	1,0

$$E(Z) = 1 \times 0,7 + 2 \times 0,3 = 1,3$$

Exemplo 7.2. *Sejam X e Y v.a's independentes, cada uma tendo uma distribuição $U(0,1)$. Seja $Z = X + Y$, encontre a $E(Z)$.*

$$g(z) = \begin{cases} z; & 0 \leq z \leq 1 \\ 2 - z; & 1 \leq z \leq 2 \end{cases}$$

$$\begin{aligned} E(Z) &= \int_0^2 zg(z)dz = \int_0^1 z^2 dz + \int_1^2 (2z - z^2)dz \\ &= \frac{z^3}{3}|_0^1 + (z^2 - \frac{z^3}{3})|_1^2 = \frac{1}{3} + (4 - \frac{8}{3} - 1 + \frac{1}{3}) = 4 - 2 - 1 = 1 \end{aligned}$$

Teorema 8. *Seja (X,Y) uma v.a bidimensional e seja $Z = H(X,Y)$.*

a) Se (X,Y) for uma v.a discreta e se $p(x_i,\ y_j) = P(X = x_i,\, Y = y_j)\, j, i = 1, 2, \ldots,$ então:

$$E(Z) = \sum_{j=1}^{\infty} \sum_{i=1}^{\infty} H(x_i, y_j) p(x_i, y_j)$$

b) Se (X,Y) for uma v.a contínua com f.d.p conjunta f, entao:

$$E(Z) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} H(x,y) f(x,y) dxdy$$

Exemplo 7.3. *Refazer os Exemplos 7.1 e 7.2 através do Teorema 8.*

7.1: *Temos que $Z = X^2 + Y$. Portanto,*

$$\begin{aligned} E(Z) &= \sum_{j=1}^{\infty} \sum_{i=1}^{\infty} H(x_i\ \ y_j) p(x_i\ \ y_j) \\ &= H(-1,0)p(-1,0) + H(-1,1)p(-1,1) + H(0,1)p(0,1) + H(1,0)p(1,0) + H(1,1)p(1,1) \\ &= 1 \times 0,2 + 2 \times 0,1 + 1 \times 0,1 + 1 \times 0,4 + 2 \times 0,2 = 1,3 \end{aligned}$$

7.2: *Temos que $Z = X + Y$, logo,*

$$f(x,y) = \begin{cases} 1; & 0 \leq x \leq 1 \\ 0; & c.c \end{cases}$$

Portanto,

$$\begin{aligned} E(Z) &= \int_0^1 \int_0^1 H(x,y) f(x,y) dxdy = \int_0^1 \int_0^1 (x + y) dxdy = \int_0^1 [\frac{x^2}{2} + yx]|_0^1 dy \\ &= \int_0^1 (\frac{1}{2} + y) dy = \frac{y}{2} + \frac{y^2}{2}\,|_0^1 = \frac{1}{2} + \frac{1}{2} = 1. \end{aligned}$$

7.2 Algumas Propriedades envolvendo Esperança e Variância

Seja (X, Y) uma variável aleatória bidimensional com distribuição de probabilidade conjunta $f(x, y)$. Seja $Z = H_1(X, Y)$ e $W = H_2(X, Y)$. Então, $E(Z + W) = E(Z) + E(W)$.

Demostração:

$$\begin{aligned} E(Z+W) &= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}[H_1(x,y)+H_2(x,y)]f(x,y)dxdy \\ &= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}H_1(x,y)f(x,y)dxdy + \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}H_2(x,y)f(x,y)dxdy \\ &= E(Z)+E(W). \end{aligned} \tag{7.1}$$

Propriedade 7.4) *Seja X e Y duas variáveis aleatórias quaisquer. Então:*

$$E(X+Y) = E(X) + E(Y).$$

Demostração:

Isto decorre imediatamente de (7.1), ao se fazer $H_1(X, Y) = X$ e $H_2(X, Y) = Y$.

Propriedade 7.5) *Sejam n variáveis aleatórias $X_1, X_2, \ldots, X_n$. Então:*

$$E(X_1 + \ldots + X_n) = E(X_1) + \cdots + E(X_n).$$

Demostração:

Isto decorre imediatamente da Propriedade 7.4, pela aplicação da indução matemática.
Comentário: Conbinando-se esta propriedade com outra já conhecida, $E(aX + b) = aE(X) + b$, obteremos:

$$E\left(\sum_{i=1}^{n} a_i X_i\right) = \sum_{i=1}^{n} a_i E(X_i),$$

onde os a_i são constantes.

Propriedade 7.6) *Seja (X, Y) uma variável aleatória bidimensional e suponha-se que X e Y sejam independentes. Então,*

$$E(XY) = E(X)E(Y).$$

Demostração:

$$\begin{aligned} E(XY) &= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}xyf(x,y)dxdy = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}xyg(x)h(y)dxdy \\ &= \int_{-\infty}^{+\infty}xg(x)dx\int_{-\infty}^{+\infty}yh(y)dy = E(X)E(Y). \end{aligned}$$

Propriedade 7.7) *Se C for uma constante,*

$$Var(X+C) = Var(X)$$

Demostração:

$$Var(X+C) = E[(X+C) - E(X+C)]^2 = E[(X+C) - E(X) - C]^2$$

$$= E[X - E(X)^2] = Var(X).$$

Propriedade 7.8) *Se for uma constante,*

$$Var(CX) = C^2 Var(X)$$

Demostração:

$$\begin{aligned} Var(CX) &= E(CX)^2 - [E(CX)]^2 = C^2[E(X)]^2 \\ &= C^2[E(X^2) - [E(X)]^2] = C^2 Var(X) \end{aligned}$$

Propriedade 7.9) *Se (X, Y) for uma variável aleatória bidimensional, e se X e Y forem independentes, então*

$$Var(X+Y) = Var(X) + Var(Y).$$

Demostração:

$$\begin{aligned} Var(X+Y) &= E(X+Y)^2 - [E(X+Y)]^2 \\ &= E(X^2 + 2XY + Y^2) - [E(X)]^2 - 2E(X)E(Y) - [E(Y)]^2 \\ &= E(X^2) - [E(X)]^2 + E(Y^2) - [E(Y)]^2 = Var(X) + Var(Y). \end{aligned}$$

Propriedade 7.10) *Sejam $X_1, \ldots, X_n$, variáveis aleatórias independente. Então,*

$$Var(X_1 + \ldots + X_n) = Var(X_1) + \ldots + Var(X_n) = \sum_{i=1}^{n} Var(X_i).$$

Demostração:
Isto decorre da Propriedade 7.9, por indução matemática.
Propriedade 7.11) *Sejam $X_1, \ldots, X_n$, variáveis aleatórias quaisquer. Então,*

$$Var(X_1 + \ldots + X_n) = \sum_{i=1}^{n} Var(X_i) + 2\sum_{i=1}^{n}\sum_{j>i}^{n} Cov(X_i, X_j).$$

Prova: *decorre diretamente de (1.7):*

$$\left(\sum_{i=1}^{p} x_i\right)^2 = \sum_{i,j} x_i x_j = \sum_{i=1}^{p} x_i^2 + 2\sum_{i=1}^{n}\sum_{j>i}^{n} x_i x_j.$$

7.3 Expressões Aproximadas da Esperança e da Variância

Teorema 7.7) *Seja (X, Y) uma variável aleatória bidimensional. Suponha-se que $E(X) = \mu_X, E(Y) = \mu_Y, V(X) = \sigma_X^2$ e $V(Y) = \sigma_Y^2$. Seja $Z = H(X, Y)$, derivável em (μ_X, μ_Y). Então, se X e Y forem independentes, ter-se-á*

$$\begin{aligned} E(Z) &\simeq H(\mu_X, \mu_Y) + \frac{1}{2}\left[\frac{\partial^2 H}{\partial x^2}\sigma_X^2 + \frac{\partial H}{\partial y^2}\sigma_Y^2\right] \\ Var(Z) &\simeq \left[\frac{\partial H}{\partial x}\right]^2 \sigma_X^2 + \left[\frac{\partial H}{\partial y}\right]^2 \sigma_Y^2. \end{aligned}$$

Demostração: *A demostração envolve o desenvolvimento de H em série de Taylor, no ponto (μ_X, μ_Y), para um de dois termos, o abandono do resto, e, a seguir, o cálculo da esperança e da variância de ambos os membros, tal como foi feito na demostração do Teor. 7.6.*

Exemplo 7.4.

$$\begin{cases} X \sim U(0,1), & E(X)=0.5 \quad Var(X) = 1/12 \\ Y \sim U(0,1), & E(Y)=0.5 \quad Var(Y) = 1/12 \end{cases}$$

$$Z = X + Y, \quad E(Z) = 1 \quad e \quad Var(Z) = 1/6$$

Valores aproximados:

$$E(Z) \simeq H(0.5; 0.5) + \frac{1}{2}\left[\frac{\partial^2 H}{\partial x^2}\frac{1}{12} + \frac{\partial^2 H}{\partial y^2}\frac{1}{12}\right]$$

$$Var(Z) \simeq \left[\frac{\partial H}{\partial x}\right]^2 \frac{1}{12} + \left[\frac{\partial H}{\partial y}\right]^2 \frac{1}{12}$$

$$E(Z) \simeq 0,5 + 0,5 + \frac{1}{12} \times 0 = 1.0$$

$$Var(Z) \simeq 1^2 \times \frac{1}{12} + 1^2 \times \frac{1}{12} = \frac{2}{12} = \frac{1}{6}$$

7.4 Esperança Condicional

Em muitas situações estamos interessados em determinadas características (distribuição, média etc.) de uma variável, sujeita a determinada condição. Neste caso passamos a trabalhar com distribuições condicionais, fixada uma outra variável.

Definição 7.1.

a) *Se (X, Y) for uma variável aleatória* ***contínua*** *bidimensional, definiremos o valor esperado condicional de X, dado $Y = y$, como*

$$E(X|y) = \int_{-\infty}^{+\infty} xg(x|y)dx \qquad E(Y|x) = \int_{-\infty}^{+\infty} yh(y|x)dy$$

b) *Se (X, Y) for uma v.a* ***discreta*** *bidimensional, definiremos o valor esperado condicional de X, dado $Y = y_j$, como*

$$E(X|y_j) = \sum_{i=1}^{\infty} x_i p(x_i|y_j)$$

$$E(Y|x_i) = \sum_{j=1}^{\infty} y_j p(y_j|x_i)$$

OBS:

1. *$E(X|Y)$ é uma função de Y ; (v.a)*
2. *$E(Y|X)$ é uma função de X ; (v.a)*
3. *Como $E(X|Y)$ e $E(Y|X)$ são v.a's, terá sentido falarmos em seu valor esperado.*

Teorema 9. *Seja (X, Y) for uma variável aleatória uma variável aleatória qualquer. Temos que*

$$E(E(X|Y)) = E(X)$$

$$E(E(Y|X)) = E(Y)$$

Demostração: (p/ caso contínuo)

$$E(X|y) = \int_{-\infty}^{+\infty} xg(x,y)dx = \int_{-\infty}^{+\infty} x\frac{f(x,y)}{h(y)}dx = \frac{1}{h(y)}\int_{-\infty}^{+\infty} xf(x,y)dx$$

Portanto,

$$\begin{aligned} E[E(X|Y)] &= \int_{-\infty}^{+\infty} E(X|y)h(y)dy \int_{-\infty}^{+\infty} \left[\frac{1}{h(y)}\int_{-\infty}^{+\infty} xf(x,y)dx\right] h(y)dy \\ &= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} xf(x,y)dxdy = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} xf(x,y)dydx = \\ &= \int_{-\infty}^{+\infty} x[\int_{-\infty}^{+\infty} f(x,y)dy]dx = \int_{-\infty}^{+\infty} xg(x)dx = E(X) \end{aligned}$$

Teorema 10. *Se X e Y forem variáveis aleatórias independentes, então*

$$E(X|Y) = E(X)$$

$$E(Y|X) = E(Y)$$

7.5 Variância condicional

Teorema 11.

$$Var(X) = E[Var(X|Y)] + Var[E(X|Y)]$$

$$Var(Y) = E[Var(Y|X)] + Var[E(Y|X)]$$

Exemplo 7.5. *Adimita que um inseto ponha ovos segundo uma distribuição de Poisson de parâmetro 4 e que a probabilidade de que um ovo dê origem a um novo inseto seja 0,6. Admitimos que os ovos produzam novos insetos de maneira independente, encontre o número esperado de novos insetos gerados pelo inseto.*

$X = n^o$ de ovos produzidos pelo inseto.

$$X \sim P(4)$$

$Y = n^o$ de novos insetos gerados.

$$E(Y) = ?$$

$$(Y|X = x) \sim B(x; 0,6) \qquad E(Y|x) = 0,6x$$

Portanto,

$$E(Y) = E(E(Y|X)) = E(0,6X) = 0,6 \times E(X) = 0,6 \times 4 = 2,4.$$

Exemplo 7.6. *Suponha-se que a variável aleatória bidimensional (X, Y) seja uniformemente distribuída sobre a região triangular a seguir. Obter $E(X|Y)$ e $E(Y|X)$*

$$T = \{(x, y)|0 < x < y < 1\}.$$

Temos que

$$f(x, y) = \begin{cases} 2, & (x, y) \in T \\ 0, & c.c \end{cases}$$

Capítulo **8**

Coeficientes de Covariância e Correlação

A principal medida de o quanto duas variáveis estão relacionadas é o coeficiente de covariância. *No entanto, sua ordem de grandeza depende da unidade de medida das variáveis. Para contornar este problema, costuma-se adotar o* coeficiente de correlação *que é a normalização da covariância, restringindo-se ao intervalo [-1,1].*

8.1 Coeficiente de covariância

Definição 8.1. *Seja (X, Y) uma variável aleatória bidimensional com médias μ_X e μ_Y, respectivamente. Definiremos o coeficiente de correlação $Cov(X, Y)$ entre X e Y (às vezes representado por σ_{XY}) por*

$$Cov(X, Y) = E\left[(X - \mu_X)(Y - \mu_Y)\right]$$

Teorema 12.

$$Cov(X, Y) = E(XY) - E(X)E(Y).$$

Demostração:

$$\begin{aligned} Cov(X, Y) &= E\{[X - E(X)][Y - E(Y)]\} \\ &= E[XY - XE(Y) - YE(X) - E(X)E(Y)] \\ &= E(XY) - E(X)E(Y) - E(Y)E(X) + E(X)E(Y) \\ &= E(XY) - E(X)E(Y). \end{aligned}$$

Como consequência deste teorema temos que

$$Cov(X, X) = Var(X). \tag{8.1}$$

Também é fácil perceber que

$$Cov(X, Y) = Cov(Y, X).$$

8.2 Coeficiente de correlação

Definição 8.2. *Seja (X, Y) uma variável aleatória bidimensional com médias μ_X e μ_Y e desvios-padrão σ_X e σ_Y, respectivamente. Definiremos ρ_{XY}, o coeficiente de correlação entre X e Y, por*

$$\rho_{XY} = \frac{E\{[X - \mu_X][Y - \mu_Y]\}}{\sqrt{Var(X)Var(Y)}}$$

Propriedade 1.

$$\rho_{XY} = \frac{E(XY) - E(X)E(Y)}{\sqrt{Var(X)Var(Y)}} = \frac{Cov(X,Y)}{\sigma_X \sigma_Y}$$

Teorema 13. *Se X e Y forem independentes, então $\rho_{XY} = 0$.*

Demostração: *Isto decorre imediatamente do Teorema 12, pois sendo X e Y independentes, temos*

$$E(XY) = E(X)E(Y).$$

Observação 8.1.

1.) A recíproca não é verdadeira

2.) $\rho = 0$ significa que X e Y são não correlacionadas

Teorema 14.

$$-1 \leq \rho \leq 1.$$

Demostração: *Considere a seguinte função da variável real t:*

$$q(t) = E[v + tw]^2$$

onde $V = X - E(X)$ e $W = Y - E(Y)$.

$$Como \quad [v + tw]^2 \geq 0 \quad \Longrightarrow \quad q(t) \geq 0 \quad \forall t$$

Agora

$$q(t) = E[v^2 + 2tvw + t^2w^2] = E(v^2) + 2tE(vw) + E(vw) + t^2E(w^2)$$

Como $q(t)$ é uma expressão quadrática de t, que só assume valores ≥ 0, para todo t, então

$$\triangle = b^2 - 4ac \quad \leq 0$$

Não possui duas raízes distintas. Logo,

$$b^2 - 4ac = 4[E(vw)]^2 - 4E(v^2)E(w^2) \leq 0$$

$$\Longrightarrow \frac{[E(vw)^2]}{E(v^2)E(w^2)} \leq 1$$

$$\Longrightarrow \frac{[E(X - E(X))(y - E(Y))]^2}{E[(x - E(X))^2]E[(y - E(Y))^2]} \leq 1$$

$$\Longrightarrow \rho^2 \leq 1 \quad \Longrightarrow \quad -1 \leq \rho \leq 1$$

Teorema 15. *Se $\rho_{xy} = \pm 1$, então Y será uma função linear de X (com probabilidade 1).*

Teorema 16. *Suponha que X e Y sejam duas v.a's, para as quais $Y = AX + B$, onde A e B são constantes. Então $|\rho| = 1$. Se $A > 0$, $\rho = +1$; se $A < 0$, $\rho = -1$.*

Teorema 17. *Se ρ_{XY} for o coeficiente de correlação entre X e Y, e se $V = AX + B$ e $W = CY + D$, onde* A, B, C *e D são constantes, com $A \neq 0$ e $C \neq 0$, então*

$$\rho_{VW} = \frac{AC}{|AC|}\rho_{XY}. \tag{8.2}$$

Capítulo **9**

Construção de Algumas Distribuições Especiais

9.1 Distribuições Discretas

9.1.1 Distribuição Multinomial

Suponha que existam k resultados possíveis para um experimento: $A_1, A_2, \ldots, A_k$. Seja $p_i = P(A_i),\ i = 1, \cdots, k$. Assim, devemos ter $\sum_{i=1}^{k} p_i = 1$. Suponha que repetimos o experimentos n vezes independentemente. Seja X_i o n^o de vezes que ocorre o resultado $A_i,\ i = 1, 2, \cdots, k$. Então, a f.p conjunta das v.a's $X_1, \cdots, X_k$ é dada por:

$$P(X_1 = x_1,\ X_2 = x_2, \cdots, X_k = x_k) = \frac{n!}{x_1!x_2!\cdots x_k!} p_1^{x_1} p_2^{x_2} \cdots p_k^{x_k}$$

com $x_i = 0, 1, \ldots, n, \quad \sum_{i=1}^{k} x_i = n.$

Teorema 18.

$$E(X_i) = np_i \quad e \quad Var(X_i) = np_i(1 - p_i), \quad i = 1, \ldots, k.$$

Observação 9.1.

- *Cada v.a X_i tem distribuição binomial, ou seja, $X_i \sim Binomial(n, p_i)$*
- *As X_i's não são independentes.*

Exemplo 9.1. *Consideremoa o experimento de lançarmos um dado $n = 12$ vezes. Temos $k = 6$ resultados possíveis e $A_i = \{resultado\ é\ face\ i\}$, com $p_i = \frac{1}{6}$. A probabilidade de obtermos 2 resultados de cada face ée dada por*

$$\begin{aligned} P(X_1 = 2,\ X_2 = 2, X_3 = 2,\ X_4 = 2,\ X_5 = 2,\ X_6 = 2) &= \frac{12!}{2!2!2!2!2!2!} \left(\frac{1}{6}\right)^2 \left(\frac{1}{6}\right)^2 \left(\frac{1}{6}\right)^2 \left(\frac{1}{6}\right)^2 \left(\frac{1}{6}\right)^2 \left(\frac{1}{6}\right)^2 \\ &= \frac{12!}{(2!)^6} \left(\frac{1}{6}\right)^{12} \end{aligned}$$

Exemplo 9.2. *Uma barra de comprimento especificado é fabricada. Admita-se que o comprimento real X (polegada) seja uma variável aleatória uniformemente distribuída sobre [10,12]. Suponha-se que somente interesse saber se um dos três eventos seguinte terá ocorrido.*

$$A_1 = \{X < 10,5\}, \quad A_2 = \{10,5 \le X \le 11,8\} \quad e \quad A_3 = \{X > 11,8\}.$$

$$p_1 = P(A_1) = 0,25, \quad p_2 = P(A_2) = 0,65 \quad e \quad p_3 = P(A_3) = 0,1$$

9.2 Distribuições Contínuas

Distribuição Normal

A variável aleatória X é dita ter distribuíção Normal com parâmetros μ e σ^2 se sua f.d.p é dada por

$$f(x) = \frac{1}{\sqrt{2\pi\sigma}} e^{\frac{-1}{2\sigma^2}(x-\mu)^2}$$

$$-\infty < x < +\infty \quad , \quad -\infty < \mu < +\infty, \quad \sigma > 0$$

Notação: $X \sim N(\mu, \sigma^2)$

Teorema 19. *Se $X \sim N(\mu, \sigma^2)$, então*

$$E(X) = \mu \quad e \quad Var(X) = \sigma^2$$

Observação 9.2.

1. *f é simétrica em torno de μ*
2. *Não há solução explícita para a Função Distribuição, apenas numérica, com tabelas na maioria dos livros*
3. *Se $\mu = 0$ e $\sigma^2 = 1$, dizemos que X tem distribuição* normal padrão *ou* padronizada *(tabelada).*

Observação 9.3. *Há livros que apresentam tabelas com a Normal Acumulada (NA) de $(-\infty, d)$, outros (maioria) no intervalo $(0, d)$. Neste último caso devemos usar simetria. Por exemplo, que $NA(-\infty, 0) = 0,5$ e que $NA(-\infty, -d) = NA(d, \infty)$. Alguns autores preferem adotar a notação: $NA(-\infty, d) = \Phi(d)$.*

Teorema 20. *Se $X \sim N(\mu, \sigma^2)$, então:*

(i) $Z = \frac{X-\mu}{\sigma} \sim N(0, 1)$ *[Transformamos e usamos a tabela]*

(ii) $Y = aX + b \sim N(a\mu + b, a^2\sigma^2)$

Teorema 21. *Se $X_i \sim N(\mu_i, \sigma_i^2)$, $\quad i = 1, \ldots, n$ e as v.a's X_i são independentes, então*

$$S = \sum_{i=1}^{n} = X_i = X_1 + X_2 + \ldots + X_n \sim N(\mu, \sigma^2)$$

com

$$\mu = \sum_{i=1}^{n} \mu_i \quad \text{e} \quad \sigma^2 = \sum_{i=1}^{n} \sigma_i^2.$$

Exemplo 9.3. *Na distribuição $X \sim N(100, 100)$, encontre:*

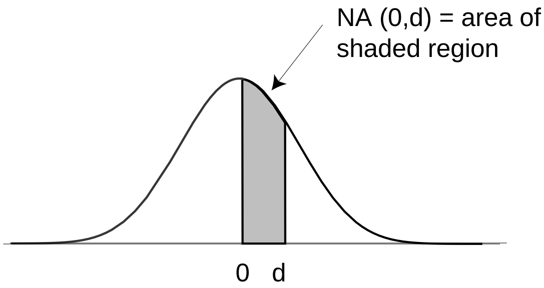
a) $P(X < 115)$

b) $P(X \geq 80)$

c) $P(|X - 100| \leq 100)$

Appendix A

Normal distribution table



	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.0000	.0040	.0080	.0120	.0160	.0199	.0239	.0279	.0319	.0359
0.1	.0398	.0438	.0478	.0517	.0557	.0596	.0636	.0675	.0714	.0753
0.2	.0793	.0832	.0871	.0910	.0948	.0987	.1026	.1064	.1103	.1141
0.3	.1179	.1217	.1255	.1293	.1331	.1368	.1406	.1443	.1480	.1517
0.4	.1554	.1591	.1628	.1664	.1700	.1736	.1772	.1808	.1844	.1879
0.5	.1915	.1950	.1985	.2019	.2054	.2088	.2123	.2157	.2190	.2224
0.6	.2257	.2291	.2324	.2357	.2389	.2422	.2454	.2486	.2517	.2549
0.7	.2580	.2611	.2642	.2673	.2704	.2734	.2764	.2794	.2823	.2852
0.8	.2881	.2910	.2939	.2967	.2995	.3023	.3051	.3078	.3106	.3133
0.9	.3159	.3186	.3212	.3238	.3264	.3289	.3315	.3340	.3365	.3389
1.0	.3413	.3438	.3461	.3485	.3508	.3531	.3554	.3577	.3599	.3621
1.1	.3643	.3665	.3686	.3708	.3729	.3749	.3770	.3790	.3810	.3830
1.2	.3849	.3869	.3888	.3907	.3925	.3944	.3962	.3980	.3997	.4015
1.3	.4032	.4049	.4066	.4082	.4099	.4115	.4131	.4147	.4162	.4177
1.4	.4192	.4207	.4222	.4236	.4251	.4265	.4279	.4292	.4306	.4319
1.5	.4332	.4345	.4357	.4370	.4382	.4394	.4406	.4418	.4429	.4441
1.6	.4452	.4463	.4474	.4484	.4495	.4505	.4515	.4525	.4535	.4545
1.7	.4554	.4564	.4573	.4582	.4591	.4599	.4608	.4616	.4625	.4633
1.8	.4641	.4649	.4656	.4664	.4671	.4678	.4686	.4693	.4699	.4706
1.9	.4713	.4719	.4726	.4732	.4738	.4744	.4750	.4756	.4761	.4767
2.0	.4772	.4778	.4783	.4788	.4793	.4798	.4803	.4808	.4812	.4817
2.1	.4821	.4826	.4830	.4834	.4838	.4842	.4846	.4850	.4854	.4857
2.2	.4861	.4864	.4868	.4871	.4875	.4878	.4881	.4884	.4887	.4890
2.3	.4893	.4896	.4898	.4901	.4904	.4906	.4909	.4911	.4913	.4916
2.4	.4918	.4920	.4922	.4925	.4927	.4929	.4931	.4932	.4934	.4936
2.5	.4938	.4940	.4941	.4943	.4945	.4946	.4948	.4949	.4951	.4952
2.6	.4953	.4955	.4956	.4957	.4959	.4960	.4961	.4962	.4963	.4964
2.7	.4965	.4966	.4967	.4968	.4969	.4970	.4971	.4972	.4973	.4974
2.8	.4974	.4975	.4976	.4977	.4977	.4978	.4979	.4979	.4980	.4981
2.9	.4981	.4982	.4982	.4983	.4984	.4984	.4985	.4985	.4986	.4986
3.0	.4987	.4987	.4987	.4988	.4988	.4989	.4989	.4989	.4990	.4990
3.1	.4990	.4991	.4991	.4991	.4992	.4992	.4992	.4992	.4993	.4993
3.2	.4993	.4993	.4994	.4994	.4994	.4994	.4994	.4995	.4995	.4995
3.3	.4995	.4995	.4995	.4996	.4996	.4996	.4996	.4996	.4996	.4997
3.4	.4997	.4997	.4997	.4997	.4997	.4997	.4997	.4997	.4997	.4998
3.5	.4998	.4998	.4998	.4998	.4998	.4998	.4998	.4998	.4998	.4998
3.6	.4998	.4998	.4999	.4999	.4999	.4999	.4999	.4999	.4999	.4999
3.7	.4999	.4999	.4999	.4999	.4999	.4999	.4999	.4999	.4999	.4999
3.8	.4999	.4999	.4999	.4999	.4999	.4999	.4999	.4999	.4999	.4999
3.9	.5000	.5000	.5000	.5000	.5000	.5000	.5000	.5000	.5000	.5000

d) o valor "a"tal que $P(100 - a \leq X \leq 100 + a) = 0,95$

a)

$$\begin{aligned} P(X < 115) &= P\left(\frac{X-100}{10} < \frac{115-100}{10}\right) = P(Z < 1,5) \\ &= 0,5 + 0,43319 = 0,93319 \end{aligned}$$

b)

$$\begin{aligned} P(X \geq 80) &= P\left(Z \geq \frac{80-100}{10}\right) = P(Z \geq -2) \\ &= = 0,5 + 0,47725 = 0,97725 \end{aligned}$$

C)

$$\begin{aligned} P(|X - 100| \leq 10) &= P(-10 \leq X - 100 \leq 10) \\ &= P\left[-\frac{10}{10} \leq Z \leq \frac{10}{10}\right] = P(-1 \leq Z \leq 1) \\ &= 2P(0 < Z < 1) = 2 \times 0,34134 = 0,68268 \end{aligned}$$

D)

$$\begin{aligned} & P(100 - a \leq X \leq 100 + a) = 0,95 \\ \Leftrightarrow\ & P\left(\frac{100-a-100}{10} \leq \frac{X-100}{10} \leq \frac{100+a-100}{10} = 0,95\right) \\ \Leftrightarrow\ & P\left(-\frac{a}{10} \leq Z \leq \frac{a}{10}\right) = 0,95 \\ \Leftrightarrow\ & P\left(0 < Z < \frac{a}{10}\right) = \frac{0,95}{2} = 0,475 = \frac{a}{10} = 1,96 \Longrightarrow a = 19,6 \end{aligned}$$

Exemplo 9.4. *Na distribuição* $X \sim N(\mu, \sigma^2)$*, encontre:*

a) $P(X \leq \mu + 2\sigma)$

b) $P(|X - \mu| \leq \sigma)$

c) o nº "a"tal que $P(\mu - a\sigma \leq X \leq \mu + a\sigma) = 0,99$

d) o nº "a"tal que $P(X > a) = 0,90$

a)

$$P(X \leq \mu + 2\sigma) = P\left(\frac{X-\mu}{\sigma} \leq \frac{\mu + 2\sigma - \mu}{\sigma}\right) = P(Z \leq 2) = 0,97725$$

b)

$$P(|X - \mu| \leq \sigma) = P(-\sigma \leq X - \mu < \sigma) = P(-1 \leq Z \leq 1) = 0,68268$$

c)

$$P(\mu - a\sigma \leq X \leq \mu + a\sigma) = 0,99$$

$$P\left(\frac{\mu - a\sigma - \mu}{\sigma} \leq Z \leq \frac{\mu + a\sigma - \mu}{\sigma}\right) = 0,99$$

$$P(-a \leq Z \leq a) = 0,99$$

$$P(0 \leq Z \leq a) = \frac{0,99}{2} = 0,495$$

$$a = 2,58$$

d)

$$P(X > a) = 0,90 \Longrightarrow P(Z > \frac{a-\mu}{\sigma}) = 0,90$$

$$\Longrightarrow \frac{a-\mu}{\sigma} = -1,28 \Longrightarrow a = \mu - 1,28\sigma$$

Distribuição Normal Bidimensional

Definição 9.1. *Seja (X, Y) uma variável aleatória contínua, bidimensional, tomando todos os valores no plano euclidiano. Diremos que (X, Y) tem uma distribuição normal bidimensional se sua fdp conjunta for dada pela seguinte expressão*

$$f(x,y) = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}} \exp\left\{-\frac{1}{2(1-\rho^2)}\left[\left(\frac{x-\mu_X}{\sigma_x}\right)^2 - 2\rho\frac{(x-\mu_X)(y-\mu_Y)}{\sigma_x\sigma_y} + \left(\frac{y-\mu_Y}{\sigma_y}\right)^2\right]\right\}$$

$$-\infty < x < \infty \quad , \quad -\infty < y < \infty$$

Teorema 22. *Suponha-se que (X, Y) tenha fdp como a dada pela equação acima. Então, nesse caso:*

a) As distribuições marginais de X e Y são $N(\mu_X, \sigma_X^2)$ e $N(\mu_X, \sigma_y^2)$, respectivamente.

b) O parâmetro ρ é o coeficiente de correlação entre X e Y.

c) As distribuições condicionais de X (dado que $Y = y$) e de Y (dado que $X = x$) serão, respectivamente:

$$N\left[\mu_X + \rho\frac{\sigma_x}{\sigma_y}(y-\mu_Y) \quad , \quad \sigma_x^2(1-\rho^2)\right],$$

$$N\left[\mu_Y + \rho\frac{\sigma_y}{\sigma_x}(x-\mu_Y) \quad , \quad \sigma_y^2(1-\rho^2)\right]$$

Observação 9.4.

1. *A recíproca de a) no Teorema não vale.*

2. *Se $\rho = 0$, então X e Y são independentes. Isto só vale na distribuição Normal bidimensional.*

Distribuição Normal n-dimensional

Definição 9.2. *Seja $\boldsymbol{X} = (X_1, X_2, \cdots, X_n)$ uma variável aleatória contínua, n-dimensional, tomando todos os valores no espaço euclidiano. Diremos que $\boldsymbol{X}$ tem uma distribuição normal n-dimensional ou multivariada se sua fdp conjunta for dada pela seguinte espressão*

$$f(\boldsymbol{x}) = \frac{1}{(2\pi|\boldsymbol{\Sigma}|)^{n/2}} \exp\left\{-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\right\}, \quad \boldsymbol{x} \in I\!R^n,$$

onde $\boldsymbol{\mu} = (\mu_1, \cdots, \mu_n)$ é o vetor das médias das X_i, $i = 1, \cdots, n$.

$$\boldsymbol{\Sigma} = \begin{pmatrix} \sigma_1^2 & \sigma_{12} & \sigma_{13} & \cdots & \sigma_{1n} \\ \sigma_{21} & \sigma_2^2 & \sigma_{23} & \cdots & \sigma_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \sigma_{n1} & \sigma_{n2} & \sigma_{n3} & \cdots & \sigma_n^2 \end{pmatrix},$$

onde $\sigma_i^2 = Var(X_i)$ e σ_{ij} é a $Cov(X_i, X_j) = \rho_{ij}\sigma_i\sigma_j$. Ainda, $|\boldsymbol{\Sigma}|$ representa o determinante de $\boldsymbol{\Sigma}$ e $\boldsymbol{\Sigma}^{-1}$ é a inversa de $\boldsymbol{\Sigma}$.

Teorema 23. *Suponha-se que $\boldsymbol{X}$ tenha fdp como a dada pela equação acima. Então,*

a) A distribuição marginal de X_i, $i = 1, \cdots, n$, é $N(\mu_i, \sigma_i^2)$.

b) O parâmetro ρ_{ij} é o coeficiente de correlação entre X_i e X_j.

c) A distribuição condicional de X_i dado que $X_j = x$ é dada por

9.2.1 Função Gama

Definição: *A função gama, denotada por Γ é definida para $p > 0$ por*

$$\Gamma(p) = \int_0^\infty x^{p-1}e^{-x}dx$$

Propriedades:

1. $\Gamma(p) = (p-1)\Gamma(p-1)$
2. $\Gamma(n) = (n-1)!$ *se n for um inteiro positivo.*
3. $\Gamma(\frac{1}{2}) = \sqrt{\pi}$

9.2.2 Distribuição Gama

$$f(x \mid \alpha, \beta) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1}e^{-\beta x} \quad , \quad x > 0 \;\; \alpha > 0 \;\; \beta > 0$$

$$X \sim Gama(\alpha, \beta)$$

$$E(X) = \frac{\alpha}{\beta}, \qquad Var(X) = \frac{\alpha}{\beta^2}$$

9.2.3 Distribuição Beta

$$f(x \mid a,b) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} x^{a-1}(1-x)^{b-1} \quad , \quad 0<x<1 \ , \ a>0 \ , \ b>0$$

$$E(X) = \frac{a}{a+b}$$

$$Var(X) = \frac{ab}{(a+b)^2(a+b+1)}$$

OBS:

$$B(a,b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)} = \int_0^1 u^{a-1}(1-u)^{b-1}du \quad , \quad a>0 \ \ b>0$$

$$f(x) = \frac{1}{B(a,b)} x^{a-1}(1-x)^{b-1}$$

9.2.4 Distribuição Qui-quadrado

Dizemos que a v.a X tem distribuição qui-quadrado com n graus de liberdade se sua f.d.p for dada por

$$f(x) = \frac{1}{2^{\frac{n}{2}}\Gamma(\frac{n}{2})} x^{\frac{n}{2}-1}e^{-\frac{x}{2}} \quad , \quad x>0 \quad n: inteiro \ \ positivo$$

Notação:

$$X \sim \chi^2_n$$

Observação 9.5. *A f.d.p acima é um caso particular da distribuição Gama(α,β), com $\alpha = \frac{n}{2}$ e $\beta = \frac{1}{2}$.*

Teorema 24. *Se $X \sim \chi^2_n$, então $E(X) = n$ e $Var(X) = 2n$*

Teorema 25. *Se $Z \sim N(0,1)$ então $X^2 \sim \chi^2_1$*

Teorema 26. *Sejam $X_1, \ldots, X_k$ v.a's independentes tais que $X_i \sim \chi^2_{n_i}$, $i = 1, \ldots, k$. Então,*

$$Y = \sum_{i=1}^{k} X_i = X_1 + \ldots + X_k \sim \chi^2_n$$

com

$$n = \sum_{i=1}^{k} n_i$$

Teorema 27. *Seja $Y \sim \chi^2_n$. Então, para n suficientemente grande, a v.a $\sqrt{2Y}$ tem aproximadamente a distribuição $N(\sqrt{2n-1}, 1)$.*

Uso da Tabela

Exemplo 9.5. *Seja $X \sim \chi^2_{15}$, obtenha o valor de "a"tal que:*

a) $P(X \leq a) = 0,10$

b) $P(X > a) = 0,95$

c) $P(X \leq a) = 0,99$

d) $P(X \geq a) = 0,025$

Exemplo 9.6. *Seja* $X \sim \chi^2_{23}$*, calcule:*

a) $P(X \leq 11,689)$

b) $P(X \leq 38,076)$

C) $P(X > 18,137)$

Exemplo 9.7. *Seja* $X \sim \chi^2_{20}$*, obtenha o valor de "a"tal que:*

$$P(X > a) = 0,80$$

$$P(X \leq a) = 0,20$$

Temos que

$$\begin{cases} P(X \leq 12,443) = 0,10 \\ P(X \leq 15,452) = 0,25 \end{cases}$$

Portanto, devemos fazer uma **Interpolação:**

$$0,15 \longrightarrow 3,009$$

$$0,10 \longrightarrow x$$

$$\Longrightarrow x = \frac{0,10 \times 3,009}{0,15} = 2,006$$

$$a = 12,443 + 2,006 = 14,449$$

Exemplo 9.8. *Seja* X_1, X_2 *e* X_3 *v.a's independentes, cada uma tendo distribuição normal com média 5 e variância 16. Encontre a distribuição das seguintes v.a's:*

a) $Y = \left(\frac{X_2-5}{4}\right)^2$

b) $S = \sum_{i=1}^{3}\left[\frac{X_i-5}{4}\right]^2$

Exemplo 9.9. *Seja* $Y \sim \chi^2_{50}$*, calcule:*

a) $P(Y > 50)$

b) $P(Y < 40,5)$

$$Y \sim \chi^2_{50} \Longrightarrow \sqrt{2Y} \sim N(\sqrt{99};1)$$

a)

$$P(Y > 50) = P(2Y > 100) = P(\sqrt{2Y} > 10) =$$

$$P\left(\frac{\sqrt{2Y}-\sqrt{99}}{1} > \frac{10-\sqrt{99}}{1}\right) = P(Z > 0,0501) = 1 - 0,5199 = 0,4801$$

b)

$$P(Y < 40,5) = P(2Y < 81) = P(\sqrt{2Y} < 9) =$$

$$= P(Z < 9 - \sqrt{99}) = P(Z < -0,94987) = 0,1711$$

9.2.5 Distribuição t-Student

Dizemos que a v.a contínua X tem distribuição t-student com n graus de liberdade se sua f.d.p for dada por

$$f_X(x) = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\Gamma(\frac{n}{2})\sqrt{\pi n}}\left(1+\frac{x^2}{n}\right)^{\frac{-(n+1)}{2}} \quad ; \quad -\infty < x < +\infty \quad n: inteiro\ \ positivo$$

Notação: $X \sim t_n$

Teorema 28. *Se $X \sim t_n$ então $E(X) = 0$ e $Var(X) = \frac{n}{n-2}$.*

Teorema 29. $\lim_{n\to\infty} f(x) = \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$*, ou seja, se $X \sim t_n$ então quando $n \longrightarrow \infty$, $X \longrightarrow N(0,1)$*

Teorema 30. *Sejam $Z \sim N(0,1)$ e $Y \sim \chi^2_k$ v.a's independentes, então a v.a $T = \frac{Z}{\sqrt{\frac{Y}{k}}} \sim t_k$*

Uso da Tabela t-Student

1. *Seja $X \sim t_{24}$, calcule:*

   a) $P(X \le 1,71)$ *= 0,95*

   b) $P(X > 0,68)$ *= 1-0,75 = 0,25*

   c) $P(X > -1,32)$ *= 0,90*

   d) $P(1,32 < X < 2,49)$ *= 0,99-0,90 = 0,009*

   e) $P(-0,68 \le X \le 1,32)$ *= 0,90-[1-0,75] = 0,90-0,25 = 0,65*

2. *Seja $X \sim t_{35}$, econtre o valor de "a"tal que*

   a) $P(X \le a) = 0,975$ *a=2,0301*

   b) $P(X > a) = 0,10$ *a=1,3062*

   c) $P(|X| \le a) = 0,90$ $\quad P[-a \le x \le a] = 0,90 \Longrightarrow a = 1,6896$

9.2.6 Distribuição F de Snedecor

Dizemos que a v.a contínua X tem distribuição F de Snedecor com n_1 graus de liberdade (g.l.) no numerador e n_2 g.l. no denominador se sua f.d.p for dada por

$$f(x) = \frac{\Gamma\left(\frac{n_1+n_2}{2}\right)}{\Gamma\left(\frac{n_1}{2}\right)\Gamma\left(\frac{n_2}{2}\right)}\left(\frac{n_1}{n_2}\right)^{\frac{n_1}{2}} x^{\frac{n_1}{2}-1}\left[1+\frac{n_1}{n_2}x\right]^{\frac{-(n_1+n_2)}{2}} \qquad ; \quad x > 0$$

Notação: $X \sim F(n_1, n_2)$

Teorema 31. *Se $X \sim F_{n_1,n_2}$, então*

$$E(X) = \frac{n_2}{n_2-2}, \qquad para\ n_2 \geq 2.$$

e

$$Var(X) = \frac{2n_2^2(n_1+n_2-2)}{n_1(n_2-2)^2(n_2-4)}, \qquad para\ n_2 \geq 4.$$

Teorema 32. *Sejam $X_1 \sim \chi^2_m$ e $X_2 \sim \chi^2_n$ v.a's independentes. Então, a v.a*

$$F = \frac{X_1/m}{X_2/n} \sim F_{m,n}$$

Teorema 3: *Se $X \sim F_{m,n}$, então a v.a $Y = \frac{1}{X} \sim F_{n,m}$*

Uso da Tabela

Exemplo 9.10. *Seja $X \sim F_{8,10}$, calcule*

a) $P(X > 3,07)$

b) $P(Z \leq 3,35)$, *com* $Z = \frac{1}{X}$

c) $P(X \leq 5,06)$

d) o valor "a"tal que $P(X > a) = 0,10$

Exercício 9.2.1. *Sejam $X \sim N(0,16)$ e $Y = \chi^2_{16}$ v.a's independentes. Encontre a distribuição da v.a $U = \frac{X}{\sqrt{Y}} \sim t_{16}$*

Exercício 9.2.2. *Sejam $X_k \sim \chi^2_k$, k=1,2,3, v.a's independentes.*

a) Obtenha a distribuição da v.a $W = \frac{5X_1}{X_2+X_3} \sim F_{1,5}$

b) Calcule $P[W \leq 10] = 0,975$

Capítulo **10**

Características especiais

10.1 A Desigualdade de Tchebycheff

Seja X uma v.a com $E(X)=\mu$ e seja c uma constante real qualquer. Então, se $E(X-c)^2$ for finita e ε for qualquer nº positivo, teremos

$$P[|X-c|\geq\varepsilon]\leq\frac{1}{\varepsilon^2}E(X-c)^2.$$

As seguintes formas são equivalentes:

a) *Considerando o evento complementar, teremos:*

$$P[|X-c|<\varepsilon]\geq 1-\frac{1}{\varepsilon^3}E(X-c)^2$$

b) *Escolhendo $c=\mu$, temos:*

$$P[|X-\mu|\varepsilon]\leq\frac{\sigma^2}{\varepsilon^2}$$

c) *Escolhendo $c=\mu$ e $\varepsilon=k\sigma$*

$$P[|X-\mu|\geq k\sigma]\leq\frac{1}{k^2}$$

$$P[|X-\mu|\leq k\sigma]\geq 1-\frac{1}{k^2}$$

Observação 10.1. *Esta última forma é particularmente informativa de como a variância mede o "grau de concentração" da probabilidade em torno de $E(X)=\mu$.*

Demostração: (Contínuo)

$$p=P[|X-c|\geq\varepsilon]=\int_A f(x)dx \qquad A=\{x:|x-c|\geq\varepsilon\}$$

Mas,

$$|x-c|\geq\varepsilon\Leftrightarrow(x-c)^2\geq\varepsilon^2\Leftrightarrow\frac{(x-c)^2}{\varepsilon^2}\geq 1$$

Logo,

$$\begin{aligned}
p &= \int_A f(x)dx\leq\int_A\frac{(x-c)^2}{\varepsilon^2}f(x)dx\\
&= \int_{-\infty}^{c-\varepsilon}\frac{(x-c)^2}{\varepsilon^2}f(x)dx+\int_{c+\varepsilon}^{+\infty}\frac{(x-c)^2}{\varepsilon^2}f(x)dx\leq\int_{-\infty}^{+\infty}\frac{(x-c)^2}{\varepsilon^2}f(x)dx\\
&= \frac{1}{\varepsilon^2}E[(X-c)^2]
\end{aligned}$$

10.2 A Desigualdade de Jensen

Capítulo **11**

Função Geradora de Momentos

A Função Geradora (ou Geratriz) de Momentos é um caso particular de função de variáveis aleatórias, onde

$$H(\boldsymbol{X}) = \exp\{\boldsymbol{tX}\} = \exp\{t_1X_1 + t_2X_2 + \cdots + t_nX_n\}, \quad n \geq 1.$$

Muitos resultados importantes são obtidos através dessa função. Até a obtenção de $E(X)$ e $Var(X)$ fica bastante simplificada.

Definição 11.1. *Seja X uma v.a discreta ou contínua. A função geradora de momentos (f.g.m) da v.a X é definida por*

$$M_X(t) = E(e^{tX}).$$

Portanto, se X for uma v.a discreta com f.p. $p(x_i) = P(X = x_i)$, $i = 1, 2, \ldots,$ a f.g.m será

$$M_X(t) = \sum_{i=1}^{+\infty} e^{tx_i} p(x_i)$$

E se X for contínua,

$$M_X(t) = \int_{-\infty}^{+\infty} e^{tx} f_X(x)$$

Observação 11.1.

1. *A f.g.m poderá não existir para todos os valores de t;*

2. *A f.g.m existe sempre para $t = 0$, e é igual a 1, ou seja, $M_X(0) = 1$ para qualquer v.a X.*

11.1 Principais distribuições

Exemplo 11.1. $X \sim U(a, b)$

$$\begin{aligned} M_X(t) &= \int_a^b \frac{e^{tx}}{b-a}dx = \frac{1}{b-a}\int_a^b e^{tx}dx = \frac{1}{b-a}\left[\frac{e^{tx}}{b-a}\Big|_a^b\right] \\ &= \frac{1}{b-a}\left[\frac{e^{tb}}{t} - \frac{e^{t^a}}{t}\right] = \frac{1}{(b-a)t}(e^{tb} - e^{ta}) \end{aligned}$$

$$M_X(t) = \begin{cases} 1, & t\text{=0;} \\ \frac{e^{tb}-e^{ta}}{(b-a)t}, & t \neq 0. \end{cases}$$

Exemplo 11.2. $X \sim B(n,p)$

$$\begin{aligned} M_X(t) &= \sum_{k=0}^{n} e^{tk} P(X=k) = \sum_{k=0}^{n} e^{tk} \binom{n}{k} p^k (1-p)^{n-k} \\ &= \sum_{k=0}^{n} \binom{n}{k} (pe^t)^k (1-p)^{n-k} = \left[pe^t + 1 - p\right]^n \end{aligned}$$

Exemplo 11.3. $X \sim P(\lambda)$

$$\begin{aligned} M_X(t) &= \sum_{k=0}^{\infty} e^{tk} P(X=k) = \sum_{k=o}^{\infty} e^{tk} \frac{e^{-\lambda}\lambda^k}{k!} = e^{-\lambda} \sum_{k=0}^{\infty} \frac{(e^t\lambda)^k}{k!} \\ &= e^{-\lambda} e^{e^t\lambda} = e^{\lambda(e^t-1)} \end{aligned}$$

Observação 11.2. *Série de Maclaurin:*

$$e^y = \sum_{n=0}^{\infty} \frac{y^n}{n!}$$

Exemplo 11.4. $X \sim Exp(\alpha)$

$$M_X(t) = \int_0^{\infty} e^{tx} \alpha e^{-\alpha x} dx = \alpha \int_0^{\infty} e^{(t-\alpha)x}$$

que só converge se $t - \alpha < 0$, ou seja, $t < \alpha$. Admitindo que esta condição seja satisfeita, teremos:

$$M_X(t) = \alpha \left[\frac{e^{(t-\alpha)x}}{t-\alpha} |_0^{\infty} \right] = \alpha \left[0 - \frac{1}{t-\alpha} \right] = \alpha \left(-\frac{1}{t-\alpha} \right) = \frac{\alpha}{\alpha - t}$$

$$M_X(t) = \frac{\alpha}{\alpha - t} \quad ; \quad t < \alpha$$

Exemplo 11.5. $X \sim N(\mu, \sigma^2)$

$$M_X(t) = \int_{-\infty}^{+\infty} e^{tx} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{+\infty} e^{tx} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx$$

Seja:

$$s = \frac{x-\mu}{\sigma} \Longrightarrow x = s\sigma + \mu \qquad dx = \sigma ds$$

Logo,

$$\begin{aligned}
M_X(t) &= \frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{+\infty} e^{t(s\sigma+\mu)}e^{-\frac{1}{2}s^2}\sigma ds = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} e^{t\mu}e^{\sigma ts-\frac{s^2}{2}}ds \\
&= \frac{e^{t\mu}}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} e^{-\frac{1}{2}(s^2-2\sigma ts)}ds \qquad \textit{e, notando que } s^2-2\sigma ts=(s-\sigma t)^2-\sigma^2t^2, \\
&= \frac{e^{t\mu}}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} e^{-\frac{1}{2}\left[(s-\sigma t)^2-\sigma^2t^2\right]}ds \\
&= \frac{e^{t\mu}}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} e^{-\frac{1}{2}(s-\sigma t)^2}e^{\frac{\sigma^2t^2}{2}}ds \\
&= \frac{e^{t\mu}e^{\frac{t^2\sigma^2}{2}}}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} e^{-\frac{1}{2}(s-\sigma t)^2}ds \\
&= e^{t\mu}e^{\frac{\sigma^2t^2}{2}}\underbrace{\int_{-\infty}^{+\infty}\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}(s-\sigma t)^2}ds}_{f.d.p \quad de \quad uma \quad N(\sigma t,1)}
\end{aligned}$$

$$M_X(t) = e^{t\mu+\frac{\sigma^2t^2}{2}}$$

Exemplo 11.6. $X \sim Gama(\alpha,\beta)$

$$\begin{aligned}
M_X(t) &= \int_0^{\infty} e^{tx}\frac{\beta^\alpha}{\Gamma(\alpha)}x^{\alpha-1}e^{-\beta x}dx \\
&= \frac{\beta^\alpha}{\Gamma(\alpha)}\int_0^{\infty} x^{\alpha-1}e^{-\beta x+tx}dx = \frac{\beta^\alpha}{\Gamma(\alpha)}\int_0^{\infty} x^{\alpha-1}e^{-x(\beta-t)}dx
\end{aligned}$$

A integral só converge se $\beta - t > 0$, ou seja, $t < \beta$. Neste caso, fazendo $u = \beta - t$, temos

$$\begin{aligned}
M_X(t) &= \frac{\beta^\alpha}{\Gamma(\alpha)}\int_0^{\infty} x^{\alpha-1}e^{-ux}dx = \frac{\beta^\alpha}{u^\alpha}\overbrace{\underbrace{\frac{u^\alpha}{\Gamma(\alpha)}\int_0^{\infty} x^{\alpha-1}e^{-ux}dx}_{Gama(\alpha,u)}}^{=1} \\
&= \left(\frac{\beta}{u}\right)^\alpha = \left(\frac{\beta}{\beta-t}\right)^\alpha
\end{aligned}$$

Portanto,

$$M_X(t) = \left(\frac{\beta}{\beta-t}\right)^\alpha \quad ; \quad t<\beta$$

Observação 11.3.

i) Se $\alpha = 1 \Longrightarrow M_X(t) = \frac{\beta}{\beta-t} \Longrightarrow X \sim Exp(\beta)$

ii) Seja $X \sim Gama(\alpha,\beta)$, *se* $\alpha = \frac{n}{2}$ *e* $\beta = \frac{1}{2}$, *então* $X \sim \chi^2_{(n)}$.

Exemplo 11.7. $X \sim \chi^2_{(n)}$

$$M_X(t) = \left(\frac{\frac{1}{2}}{\frac{1}{2}-t}\right)^{\frac{n}{2}} = \left(\frac{1}{1-2t}\right)^{\frac{n}{2}} = (1-2t)^{-\frac{n}{2}}$$

11.2 Momentos

Definição 11.2. *Defini-se o Momento de Ordem n da variável aleatória X, em relação a zero, a quantidade $E(X^n)$.*

Observação 11.4.

i) Podemos notar que o primeiro momento de X é $E(X)$.

ii) A variância de X é uma função dos dois primeiros momentos: $Var(X) = E(X^2) - \mu^2$

iii) Todos os momentos, se existirem, podem ser obtidos via fgm.

iv) Quando a função é definida por $E(t^X)$ ela recebe o nome de Função Geradora de Probabilidades

v) Quando a função é definida por $E(e^{itX})$, onde $i = \sqrt{-1}$, ela recebe o nome de Função Característica

Primeiro momento:

$$M'_X(t) = \frac{dM_X(t)}{dt} = \frac{d}{dt}E(e^{tX}) = E\left(\frac{d}{dt}e^{tX}\right) = E(Xe^{tX})$$

Agora, avaliando no ponto $t = 0$,

$$M'_X(0) = \frac{dM_X(t)}{dt} \mid_{t=0} = E(Xe^{0\times x}) = E(X)$$

Segundo momento:

$$M''_X(t) = \frac{d^2M_X(t)}{dt^2} = \frac{d^2}{dt}E(e^{tX}) = \frac{d}{dt}E(Xe^{tX}) = E\left(\frac{d}{dt}Xe^{tX}\right) = E(X^2e^{tX})$$

Agora, avaliando no ponto $t = 0$,

$$M''(0) = \frac{d^2M_X(t)}{dt^2} \mid_{t=0} = E(X^2e^{0\times X}) = E(X^2)$$

3) *Admitindo que a derivada n-ésima de $M_X(t)$ exista, teremos:*

$$M^{(n)}(0) = E(X^n)$$

Exemplo 11.8. *$X \sim B(n,p)$. Obter $E(X)$ e $Var(X)$.*

$$M_X(t) = [pe^t + 1 - p]^n$$

$$\begin{aligned} E(X) &= M'(0) \\ Var(X) &= E(X^2) - E((X))^2 = M''_X(0) - (M'_X(0))^2 \end{aligned}$$

Temos que,

$$M^{'}(t) = n[pe^t + 1 - p]^{n-1}pe^t = npe^t[pe^t + 1 - p]^{n-1}$$

E, avaliando no ponto $t = 0$,

$$E(X) = M_X^{'}(0) = npe^0[pe^0 + 1 - p]^{n-1} = np$$

Vamos agora obter o segundo momento de X:

$$\begin{aligned} M^{''}(t) &= \frac{d}{dt}[npe^t(pe^t + 1 - p)^{n-1}] \\ &= npe^t(pe^t + 1 - p)^{n-1} + npe^t(n-1)(pe^t + 1 - p)^{n-2}pe^t \\ &= npe^t(pe^t + 1 - p)^{n-1} + n(n-1)(pe^t)^2(pe^t + 1 - p)^{n-2} \end{aligned}$$

$$\begin{aligned} E(X^2) &= M_X^{''}(0) = npe^0(pe^0 + 1 - p)^{n-1} + n(n-1)(pe^0)^2(pe^0 + 1 - p)^{n-2} \\ &= np + n(n-1)p^2 = np + n^2p^2 - np^2 \end{aligned}$$

Com isso,

$$Var(X) = np + n^2p^2 - np^2 - n^2p^2 = np - np^2 = np(1-p)$$

Exemplo 11.9. $X \sim N(\alpha, \beta^2)$

$$M_X(t) = \exp\{t\alpha + \frac{\beta^2t^2}{2}\}$$

$$M^{'}(t) = e^{\left(t\alpha + \frac{\beta^2t^2}{2}\right)}(\alpha + \beta^2 t)$$

$$M^{''}(t) = (\alpha + \beta^2 t)^2 e^{\left(t\alpha + \frac{\beta^2t^2}{2}\right)} = e^{\left(t\alpha + \frac{\beta^2t^2}{2}\right)}\beta^2$$

$$E(X) = M_X^{'}(0) = e^{(0+0)}(\alpha + \beta^2 0) = \alpha$$

$$E(X^2) = M_X^{''}(0)(\alpha + 0)^2 e^0 + e^0\beta^2 = \alpha^2 + \beta^2$$

$$Var(X) = E(X^2) - (E(X))^2 = \alpha^2 + \beta^2 - \alpha^2 = \beta^2$$

11.3 Transformações via fgm

Teorema 33. *Suponha que a v.a X tenha f.g.m M_X. A f.g.m da v.a $Y = \alpha X + \beta$ será dada por:*

$$M_Y(t) = e^{\beta t}M_X(\alpha t)$$

Demostração:

$$\begin{aligned} M_Y(t) &= E(e^{tY}) = E(e^{t(\alpha X + \beta)}) = E(e^{\alpha t X + \beta t}) = e^{\beta t}E(e^{\alpha t X}) \\ &= e^{\beta t}M_X(\alpha t) \end{aligned}$$

Teorema 34. *Sejam X e Y duas v.a's com f.g.m $M_X(t)$ e $M_Y(t)$, respectivamente. Se $M_X(t) = M_Y(t)$ para todos os valores de t, então X e Y terão a mesma distribuição de probabilidade.*

Exemplo 11.10. *$X \sim N(\mu, \sigma^2)$. Usando a f.g.m prove que $Y = \alpha X + \beta$ terá distribuição normal com média $\alpha\mu + \beta$ e variância $\alpha^2\sigma^2$.*

$$X \sim N(\mu, \sigma^2) \Longrightarrow M_X(t) = e^{\mu t + \frac{\sigma^2 t^2}{2}}$$

$$\begin{aligned} M_Y(t) &= e^{\beta t} M_x(\alpha t) = e^{\beta t}[e^{\mu\alpha t + \frac{\sigma^2}{2}(\alpha t)^2}] \\ &= e^{\beta t + \alpha\mu t + \frac{(\alpha\sigma)^2 t^2}{2}} = e^{(\alpha\mu+\beta)t + \frac{(\alpha\sigma)^2 t^2}{2}} \end{aligned}$$

$$\Longrightarrow Y \sim N(\alpha\mu + \beta; \alpha^2\sigma^2)$$

Teorema 35. *Sejam X e Y v.a's independentes e $Z = X + Y$. Sejam $M_X(t)$, $M_Y(t)$ e $M_Z(t)$ as f.g.m das v.a's X, Y e Z, respectivamente. Então:*

$$M_Z(t) = M_X(t)M_Y(t)$$

Demostração:

$$\begin{aligned} M_Z(t) = E(e^{tZ}) = E(e^{t(X+Y)}) = E(e^{tX+tY}) = E(e^{tX}e^{tY}) = \\ = E(e^{tX})E(e^{tY}) = M_X(t)M_Y(t) \end{aligned}$$

Observação 11.5. *Este teorema pode ser generalizado para uma soma de n v.a's independentes.*

$$\begin{cases} X_1 : & M_1(t) \\ X_2 : & M_2(t) \\ \vdots & \vdots \\ X_n : & M_n(t) \end{cases}$$

$$Z = X_1 + X_2 + \ldots + X_n \qquad M_Z(t) = M_1(t) \times \ldots \times M_n(t) = \prod_{i=1}^{n} M_i(t)$$

Teorema 36. *Sejam $X_1, X_2, \ldots, X_n$ v.a's independentes com distribuição $N(\mu_i, \sigma_i^2)$, $i = 1, 2, \ldots, n$. Então, a v.a $Z = X_1 + X_2 + \ldots + X_n$ terá distribuição $N\left(\sum_{i=1}^{n}\mu_i, \sum_{i=1}^{n}\sigma_i^2\right)$.*

Demostração:

$$M_{X_i}(t) = e^{\mu_i t + \sigma_i^2 \frac{t^2}{2}} \qquad ; \qquad i = 1, 2, \ldots, n$$

$$\begin{aligned} M_Z(t) &= M_{X_1}(t)M_{X_2}(t) \times \ldots \times M_{X_n}(t) \\ &= e^{\mu_1 t + \sigma_1^2 \frac{t^2}{2}} \times e^{\mu_2 t + \sigma_2^2 \frac{t^2}{2}} \times \ldots \times e^{\mu_n t + \sigma_n^2 \frac{t^2}{2}} \\ &= e^{(\mu_1 + \mu_2 + \ldots + \mu_n)t + (\sigma_1^2 + \sigma_2^2 + \cdots + \sigma_n^2)\frac{t^2}{2}} \\ &= e^{(\sum_{i=1}^{n}\mu_i)t + (\sum_{i=1}^{n}\sigma_i^2)\frac{t^2}{2}} \end{aligned}$$

$$\Longrightarrow Z \sim N\left(\sum_{i=1}^{n}\mu_i; \sum_{i=1}^{n}\sigma_i^2\right)$$

Teorema 37. *Sejam $X_1, X_2, \ldots, X_n$ v.a's independentes. Suponha que $X_i \sim P(\lambda_i)$, $i = 1, 2, \ldots, n$. Então, a v.a $Z = X_1 + X_2 + \cdots + X_n$ terá distribuição de Poisson com parâmetro $\lambda = \lambda_1 + \lambda_2 + \ldots + \lambda_n$.*

Demonstração:

$$M_{X_i}(t) \;=\; e^{\lambda_i(e^{t-1})} \qquad ; \qquad i = 1, 2, \ldots, n$$

Logo,

$$\begin{aligned} M_Z(t) &= e^{\lambda_1(e^t-1)} \times e^{\lambda_2(e^t-1)} \times \cdots \times e^{\lambda_n(e^t-1)} = e^{(\lambda_1+\lambda_2+\ldots+\lambda_n)(e^t-1)} \\ &= e^{(\sum_{i=1}^n \lambda_i)(e^t-1)} \Longrightarrow Z \sim P(\sum_{i=1}^n \lambda_i) \end{aligned}$$

Teorema 38. *Suponha que $X_i \sim \chi^2_{n_i}$, $i = 1, 2, \ldots, k$, onde os X_i são v.a's independentes. Então, a v.a $Z = X_1 + X_2 + \ldots + X_n$ terá distribuição $\chi^2_{(n)}$, onde $n = \sum_{i=1}^n n_i$.*

Demonstração:

$$M_{X_i}(t) = (1-2t)^{-\frac{n_i}{2}} \qquad ; \qquad i = 1, 2, \ldots, k.$$

Portanto,

$$\begin{aligned} M_Z(t) &= M_{X_1}(t) M_{X_2}(t) \ldots M_{X_n}(t) = (1-2t)^{\frac{-n_1}{2}} (1-2t)^{\frac{-n_2}{2}} \ldots, (1-2t)^{-\frac{n_k}{2}} \\ &= (1-2t)^{-\frac{n_1+n_2+\ldots+n_k}{2}} = (1-2t)^{-\frac{n}{2}}, \end{aligned}$$

onde

$$n = \sum_{i=1}^k n_i \Longrightarrow Z \sim \chi^2_{(n)}$$

Teorema 39. *Seja $Z = X_1 + X_2 + \ldots + X_r$, onde os X_i são r v.a's independentes e indenticamente distribuídas, cada uma tendo distribuição exponencial com o mesmo parâmetro α. Então, Z terá distribuição Gama com parâmetros r e α.*

Demonstração:

$$M_{X_i}(t) = \frac{\alpha}{\alpha - t} \qquad ; \quad i = 1, 2, \ldots, r \quad ; \quad t < \alpha$$

$$M_z(t) = \left(\frac{\alpha}{\alpha - t}\right)^r \Longrightarrow Z \sim Gama(r, \alpha)$$

a) Só vale para exponenciais de mesmo parâmetro;

b) $W = 2\alpha Z \sim \chi^2_{(2r)}$

$$M_W(t) = M_Z(2\alpha t) = \left(\frac{\alpha}{\alpha - 2\alpha t}\right)^r = \left(\frac{1}{1-2t}\right)^r = (1-2t)^{\frac{-2r}{2}}$$

Portanto, na necessidade de alguma probabilidade associada a uma distribuição Gama, fazemos a transformação e usamos tabela da distribuição χ^2. Por exemplo,

$$P(Z \leq 3) = P(2\alpha Z \leq 6\alpha) \qquad (usar \quad a \quad \chi^2_{(2r)})!$$

Exemplo 11.11. *Suponha que X tenha a seguinte fdp:*

$$f_X(x) = \lambda e^{-\lambda(x-\alpha)} \qquad , \qquad x \geq \alpha.$$

a) Determine a fgm de X.

b) Empregando a fgm, ache a $E(X)$ e $Var(X)$.

Resposta:

a)

$$M_X(t) = \frac{\lambda e^{\alpha t}}{\lambda - t}, \qquad t \leq \lambda$$

b)

$$E(X) = \frac{\alpha\lambda + 1}{\lambda}, \qquad Var(X) = \frac{1}{\lambda^2}$$

Observação 11.6. *Verificar que $Y = X + \alpha$, onde $Y \sim Exp(\lambda)$.*

Exemplo 11.12. *Alguns resistores R_i, $i = 1, 2, \ldots, n$ são montados em série em um circuito. Suponha que a resistência de cada um seja normalmente distribuída com $E(R_i) = 10$ ohms e $Var(R_i) = 0,16$ $ohms^2$*

a) Se $n = 5$, qual será a probabilidade de que a resistência do circuito exceda 49 ohms ?

b) Quantos resistores devemos escolher (ou seja, n) de forma que a probabilidade de que a resistência total exceda 100 ohms seja de aproximadamente 0,05?

Solução:

a)

$$n = 5, \quad R = R_1 + R_2 + \ldots + R_5 \quad \Longrightarrow R \sim N(50; 0,8)$$

$$P(R > 49) = P(Z > \frac{49-50}{\sqrt{0,8}}) = P(Z > -1,12) = 0,8686$$

b)

$$P(R > 100) = 0,05 \quad , \quad n = ?$$

$$R = R_1 + R_2 + \ldots + R_n \Longrightarrow R \sim N(10n, 0,16n)$$

$$P(R > 100) = 0,05 \Longrightarrow P\left(Z > \frac{100-10n}{0,4\sqrt{n}}\right) = 0,05 \Longrightarrow \frac{100-10n}{0,4\sqrt{n}} = 0,65$$

$$\Longrightarrow \frac{(100-10n)^2}{0,16n} = 2,7225 \Longrightarrow 10000 - 2000n + 100n^2 = 0,4356n$$

$$\Longrightarrow 100n^2 - 2000,4356n + 10000 = 0 \qquad \begin{cases} n_1 = 9,8 \\ n_2 = 10,2 \end{cases} \quad \Rightarrow n = 10$$

Exemplo 11.13. *Em um circuito, n resistores são montados em séries. Suponha que a resistência de cada um seja uniformemente distribuída sobre [0,1] e suponha, também, que todas as resistências sejam independentes. Seja R a resistência total.*

a) Estabeleça a fgm de R.

b) Empregando a fgm, obtenha a $E(R)$ e $Var(R)$

Solução:

R_i = *Resistência do i-ésimo resistor.* $(i = 1, 2, \ldots, n)$
$R_i \sim U(0, 1)$ *independentes.*
R = *Resistência total.* $R = R_1 + R_2 + \ldots + R_n$

a) $M_R(t) = M_{R_1}(t)M_{R_2}(t)\ldots M_{R_n}(t)$

$$M_{R_1}(t) = E(e^{tR_1}) = \int_0^1 e^{tr}dr = \frac{e^{tr}}{t}\ |_0^1 = \frac{e^t - 1}{t}$$

$$M_R(t) = \left(\frac{e^t - 1}{t}\right)^n$$

b) $E(R) =?\qquad Var(R) =?$

$$M^{'}(t) = n\left(\frac{e^t - 1}{t}\right)^{n-1}\left(\frac{te^t - (e^{t-1})}{t^2}\right) = n\left(\frac{e^t - 1}{t}\right)^{n-1}\left(\frac{te^t - e^t + 1}{t^2}\right)$$

$$M^{'}_X(0) = \frac{0}{0}\qquad (Indeterminacao)$$

Aplicando a regra de L'hopital, temos:

$$E(X) = \lim_{t\to 0} M^{'}(t) = \left[n\left(\frac{e^t - 1}{t}\right)^{n-1}\left(\frac{te^2 - e^t + 1}{t^2}\right)\right]$$

$$= \lim_{t\to 0}\left[n\left(\frac{e^t - 1}{t}\right)^{n-1}\right]\lim_{t\to 0}\left[\frac{te^t - e^t + 1}{t^2}\right]$$

Notando, agora, que

$$\lim_{t\to 0} n\left(\frac{e^t - 1}{t}\right)^{n-1} = n\lim_{t\to 0}\left(\frac{e^t - 1}{t}\right)^{n-1} = n\left(\lim_{t\to}\frac{e^t - 1}{t}\right)^{n-1} = n\left(\lim_{t\to} e^t\right) = n$$

e

$$\lim_{t\to 0}\left(\frac{te^t - e^t + 1}{t^2}\right) = \lim_{t\to 0}\left(\frac{e^t + te^t - e^t}{2t}\right) = \lim_{t\to}\frac{te^t}{2t} = \frac{1}{2}$$

temos

$$E(X) = n\frac{1}{2} = \frac{n}{2}$$

Exemplo 11.14. *Suponha que a duração da vida de uma peça seja exponencialmente distribuída, com média 2. Suponha que 10 dessas peças sejam instaladas sucessivamente, de modo que a i-ésima peça seja instalada imediatamente depois que a ordem* $(i-1)$ *tenha falhado. Seja* T_i *a duração até falhar da i-ésima peça,* $i = 1, 2, \ldots, 10$*, sempre medida a partir do instante de instalação. Portanto,* $S_{10} = T_1 + \ldots + T_{10}$ *representa o tempo total de funcionamento das 10 peças. Admitindo que os* T_i *sejam independentes, calcule* $P(S_{10} \geq 15, 5)$.

Solução:

D_1 = *Duração de vida*
$D \sim Exp(0, 5)$
T_i = *Duração até falhar da i-ésima peça* $i = 1, 2, \ldots, 10$
$S = T_1 + T_2 + \cdots + T_{10}$ = *Tempo total de funcionamento das 10 peças.*

$$P(S > 15, 5) = ?$$

$$S \sim Gama(10; 0, 5) \qquad \begin{cases} r = 10, \\ \alpha = 0, 5, \end{cases}$$

$$2\alpha S = S \sim \chi^2_{20}$$

$$P(S > 15, 5) = P(\chi^2_{20} > 15, 5) = 0, 75$$

Capítulo **12**

Teoremas limite e convergência

12.1 Lei dos grandes números

Definição 12.1 (A lei dos grandes números - LGN - Formulação de Bernoulli)**.** *Seja E um experimento e A um evento associado a E. Considere n repetições independentes de E, sendo n_A o nº de vezes em que A ocorre nas n repetições, e faça $f_n(A) = \dfrac{n_A}{n}$. Seja $P(A) = p$ (a qual se admite seja a mesma para todas as repetições). Então, para todo $\varepsilon > 0$, teremos*

$$P\,[\ |f_n(A) - p\ | \geq \varepsilon\ ] \leq \frac{p\,(1-p)}{n\varepsilon^2}$$

ou, equivalentemente,

$$P\,[\ |f_n(A) - p\ | < \varepsilon\ ] \geq 1 - \frac{p\,(1-p)}{n\varepsilon^2}$$

Demonstração:

$$n_A \sim B(n,p) \Rightarrow \begin{cases} E(n_A) = np \\ Var\ (n_A) = np\ (1-p) \end{cases}$$

$$Mas,\ f_n(A) = \tfrac{n_A}{n} \Rightarrow \begin{cases} E(f_n(A)) = p \\ Var\ (f_n(A)) = \frac{1}{n^2} \cdot np\ (1-p) = \frac{p\ (1-p)}{n} \end{cases}$$

Usando a desigualdade de Tchebycheff, temos:

$$P\left[\ |f_n(A) - p\ | < k\sqrt{\tfrac{p\ (1-p)}{n}}\ \right] \geq 1 - \tfrac{1}{k^2}$$

$$Tomando \quad \varepsilon = k\sqrt{\tfrac{p\ (1-p)}{n}} \quad \Rightarrow \quad k = \tfrac{\varepsilon\sqrt{n}}{\sqrt{p\ (1-p)}} \quad \Rightarrow \quad k^2 = \tfrac{n\varepsilon^2}{p\ (1-p)}$$

$$\Rightarrow P\,[|f_n(A) - p\ | < \varepsilon\ ] \geq 1 - \tfrac{p\ (1-p)}{n\varepsilon^2}$$

Obs:

1) $\lim\limits_{n\to\infty} P\,[|f_A - p\ | < \varepsilon] = 1$ *para todo* $\varepsilon > 0$ *;*

($f_n(A)$ converge *para todo* $p = P(A)$ *)*

2) Esta é uma "convergência em probabilidade", no sentido de que a probabilidade do evento

$$\left[\,\left|\frac{n_A}{n} - P(A)\right| < \varepsilon\right]$$

pode se tornar arbitrariamente próxima da unidade, ao se tornar n suficientemente grande;

Exemplo 12.1. *Quantas repetições de E deveremos realizar, a fim de termos uma probabilidade de ao menos $0,95$ para que a frequência relativa difira de $p = P(A)$ por menos do que $0,01$?*

$$P\left[\,|f_A - p\,| < 0,01\right] \geq 0,95$$

Para $\varepsilon = 0,01$ qual o valor de n que satisfaz a

$$1 - \frac{p\,(1-p)}{n\,\varepsilon^2} = 0,95 \quad \Longleftrightarrow \quad \frac{p\,(1-p)}{n\,\varepsilon^2} = 0,05 \quad \Longleftrightarrow \quad n = \frac{p\,(1-p)}{(0,01)^2\,0,05}$$

Portanto,

$$P\left[\,|f_A - p\,| < \varepsilon\,\right] \geq \underbrace{1-\delta}_{\gamma} \quad \text{sempre que} \quad n \geq \frac{p\;(1-p)}{\varepsilon^2\,\delta}$$

Exemplo 12.2. *Quantas vezes deveremos jogar um dado equilibrado de maneira a ficarmos ao menos 95% certos de que a frequência relativa de tirarmos um seis, fique ao menos de $0,01$ da probabilidade teórica $1/6$?*

$p = \frac{1}{6}$ $\qquad \varepsilon = 0,01$ $\qquad \delta = 0,05$

$n \geq \frac{\frac{1}{6}\times\frac{5}{6}}{(0,01)^2\cdot\,0,05} = 27.777,8 \simeq 27.778$

Observação 12.1. *Quando não conhecermos o valor de p, podemos empregar o fato de que $p\;(1-p)$ é máximo quando $p = \frac{1}{2}$, e esse valor máximo é $1/4$. Logo,*

$$P\left[\,|f_A - p\,| < \varepsilon\,\right] \geq 1-\delta \quad \text{sempre que } n \geq \frac{1}{4\,\varepsilon^2\,\delta},$$

pois,

$f(p) = p\,(1-p) = p - p^2 \qquad ; \qquad 0 < p < 1$

$f'(p) = 1 - 2p \qquad \Rightarrow \qquad f''(p) = -2 < 0$

$f'(p) = 0 \quad \Rightarrow \quad 1 - 2p = 0 \quad \Rightarrow \quad p = 1/2$ *(ponto de máximo)*

A LNG pode ser usada de maneira útil na seguinte situação. Suponha que queiramos saber quantas repetições de um experimento de Bernoulli devemos realizar a fim de que K/n difira do valor teórico p (desconhecido) de menos de ε, com probabilidade maior ou igual a γ. Ou seja, queremos determinar n, tal que

$$P\left\{\left|\frac{K}{n} - p\right| < \varepsilon\right\} \geq \gamma$$

$$P\left\{\left|\frac{K}{n} - p\right| < \varepsilon\right\} \geq 1 - \frac{p\;(1-p)}{n\;\varepsilon^2};$$

Logo, comparando, temos que n *deve satisfazer*

$$1-\frac{p\,(1-p)}{n\;\varepsilon^2}=\gamma \quad\Rightarrow\quad n=\frac{p\,(1-p)}{\delta\;\varepsilon^2},\ \ onde\ \delta=1-\gamma.$$

Como não conhecemos p *, usamos o fato de que* $p\,(1-p)\leq 1/4$*. Logo, basta tomar* n *tal que*

$$n=1/(4\delta\varepsilon^2).$$

Exemplo 12.3. *Suponha que a proporção de fumantes de uma população seja* p*, desconhecida. Queremos determinar* p *com um erro de, no máximo,* $0,05$*. Qual deve ser o tamanho da amostra* n *a ser escolhida com reposição, se* $\gamma=0,95$*?*

Definição 12.2 (Outra formulação da Lei dos Grandes Números)**.** *Sejam* $X_1,\cdots,X_n$ *v.a.'s i.i.d com* $E(X_i)=\mu$ *e* $Var(X_i)=\sigma^2$*. Seja* $\overline{X}=\frac{X_1+X_2+\ldots+X_n}{n}$ *(Média aritimética de* $X_1,\cdots,X_n$ *). Então,*

$$E(\overline{X})=\mu \quad e \quad Var(\overline{X})=\frac{n\,\sigma^2}{n^2}=\frac{\sigma^2}{n}$$

Pela desigualdade de Tchebycheff, temos:

$P\left[\ \left|\overline{X}-\mu\ \right|<\frac{k\,\sigma}{\sqrt{n}}\ \right]\geq 1-\frac{1}{k^2}$

Seja $\quad \varepsilon=\frac{k\,\sigma}{\sqrt{n}} \quad\Rightarrow\quad k=\frac{\varepsilon\sqrt{n}}{\sigma}$*, logo*

$P\left[\ \left|\overline{X}-\mu\ \right|<\varepsilon\ \right]\geq 1-\frac{\sigma^2}{n\,\varepsilon^2}$

$\lim_{n\to\infty} P\left[\ \left|\overline{X}-\mu\ \right|<\varepsilon\right]=1$

$\therefore$ *A média aritmética* $\overline{X}$ *é uma v.a. que* converge *para* $E(X)$*.*

12.1.1 Aproximação Normal da Distribuição Binomial

Consideremos uma v.a. X *com distribuição Binomial, ou seja,* $X\sim B(n,p)$*, com fp dada por*

$$P(X=k)=\binom{n}{k}p^k(1-p)^{n-k}.$$

No entanto, se n *for grande será bastante complicado calcular os fatoriais, de forma que precisamos de alguma simplificação ou aproximação.*

Fórmulas de Stirling

Para n *grande,*

$$n!\simeq\sqrt{2\pi}\ \ e^{-n}n^{n+\frac{1}{2}}$$

Empregando a fórmula de Stirling aos vários fatoriais, obtemos

$$P(X = k) \simeq \frac{1}{\sqrt{2\pi np(1-p)}} \ \exp \left\{ -\frac{1}{2} \left[\frac{k - np}{\sqrt{np(1-p)}} \right]^2 \right\}$$

Aproximação de Demoire - Laplace para a Distribuição Binomial

Se $X \sim B(n,p)$ e se $Z = \frac{X-np}{\sqrt{np(1-p)}}$, então, para n grande, Z terá aproximadamente distribuição normal padrão ($\mu = 0$ e $\sigma^2 = 1$).

Observação 12.2. *Esta aproximação será válida para $n > 10$, desde que p seja próximo de $\frac{1}{2}$. Se p for próximo de 0 ou 1, n deverá ser um tanto maior para garantir uma boa aproximação.*

Correção de Continuidade:

Ao aproximar uma v.a. discreta por uma contínua, estamos aproximando uma integral pela soma das áreas de retângulos, de forma que precisamos de cuidados adicionais. Algumas correções de continuidade devem ser utilizadas para melhorar a aproximação. Estas correções consistem em adicionar ou subtrair 0,5 de forma conveniente.

Exemplo 12.4. *Seja $X \sim B(10; 0,5)$. Obter $P(X \geq 7)$ e $P(X < 3)$ utilizando a correção de continuidade.*

$$Y \sim N(np\ ;\ np\,(1-p)) \quad \Longrightarrow \quad Y \sim N(5\ ;\ 2,5)$$

a) $P(X \geq 7) \simeq P(Y \geq 6,5)$

$$\begin{cases} P(X \geq 7) = 0,172 \\ P(Y \geq 6,5) = P\left(Z \geq \frac{6,5-5}{\sqrt{2,5}}\right) = P(Z \geq 0,94) = 0,1736 \end{cases}$$

b) $P(X < 3) \simeq P(-0,5 < Y < 2,5)$

$P(X < 3) = 0,055$

$$\begin{aligned} P(-0,5 < Y < 2,5) &= P\left(\frac{-0,5-5}{\sqrt{2,5}} < Z < \frac{2,5-5}{\sqrt{2,5}}\right) \\ &= P(-3,48 < Z < -1,58) = 0,49975 - 0,44295 = 0,0568 \end{aligned}$$

No geral, devemos aplicar as seguintes correções, como exemplo,

a) $P(X = k) \simeq P\left(k - \frac{1}{2} \leq Y \leq k + \frac{1}{2}\right)$

b) $P(a \leq X \leq b) \simeq P\left(a - \frac{1}{2} \leq Y \leq b + \frac{1}{2}\right)$

c) $P(X \leq C) \simeq P\left(-\frac{1}{2} \leq Y \leq c + \frac{1}{2}\right)$

Exemplo 12.5. *Um sistema complexo é constituído de 100 componentes que funcionam independentemente. A probabilidade de que qualquer um dos componentes venha a falhar durante o período de operação é igual a 0,10. A fim de que o sistema completo funcione, pelo menos 85 dos componentes devem funcionar perfeitamente. Calcule a probabilidade de que isso aconteça.*

Como $X \sim B(100\ ;\ 0,9)$, *podemos adotar* $Y \sim N(90\ ;\ 9)$. *Com isso,*

$$P(X \geq 85) \simeq P(Y \geq 84,5) = P\left(Z \geq \tfrac{84,5-90}{3}\right) = P(Z \geq -1,83) = 0,96638$$

Exemplo 12.6. *Suponha que o sistema do Exercício 12.5 seja constituído de* n *componentes cada um deles tendo uma confiabilidade de 0,90. O sistema funcionará se ao menos 80 por cento dos componentes funcionarem adequadamente. Determine* n *de maneira que o sistema tenha uma confiabilidade de 0,95.*

$$P(T \geq 0,8n) = 0,95 \qquad T \sim Bin(n; 0,90)$$

$$P(T \geq 0,8n) \simeq P\left(Z \geq \tfrac{0,8n-0,9n}{\sqrt{0,09n}}\right) = P\left(Z \geq \tfrac{-0,33n}{\sqrt{n}}\right) = 0,95 \quad \Rightarrow \quad n = 24,21$$

Portanto, devemos adotar $n = 25$.

12.2 O Teorema Central do Limite

Teorema 40. *Sejam* $X_1, \cdots, X_n$, *variáveis aleatórias independentes, todas com a mesma distribuição. Sejam* $\mu = E(X_i)$ *e* $\sigma^2 = Var(X_i)$ *a esperança e a variância comuns. Façamos* $S_n = \sum_{i=1}^{n} X_i$. *Então,* $E(S_n) = n\,\mu$ *e* $Var(S_n) = n\,\sigma^2$ *e, para* n *grande temos que* $T_n = (S_n - n\,\mu)/\sqrt{n}\,\sigma$ *terá aproximadamente a distribuição* $N(0,1)$, *no sentido de que* $\lim\limits_{n\to\infty} P(T_n \leq t) = \Phi(t)$.

Demonstração:

$M_X(t):\quad$ *f.m.g das* $X_i^{\prime}s$
$M_{S_n}(t):\quad$ *f.m.g de* S_n
$M_{S_n}(t) = [M_X(t)]^n$

Como T_n *é uma função linear de* S_n, *temos que*

$$M_{T_n}(t) = e^{\frac{-\sqrt{n}\mu}{\sigma}t}\left[M_X\left(\frac{t}{\sqrt{n}\ \sigma}\right)\right]^n$$

Portanto,

$$\ln M_{T_n}(t) = \frac{-\sqrt{n}\mu\, t}{\sigma} + n \ln M_X\left(\frac{t}{\sqrt{n}\ \sigma}\right)$$

Desenvolvendo $M_X(t)$ em série de Mclaurin, obtemos

$M(t) = 1 + M'(0)t + \frac{M''(0)t^2}{2} + R \quad \rightarrow \quad$ *onde R é o resto da série*

$\quad\quad = 1 + \mu t + \frac{(\mu^2+\sigma^2)\,t^2}{2} + R$

Logo,

$$\ln M_{T_n}(t) = \frac{-\sqrt{n}\;\mu\,t}{\sigma} + n\,\ln\left[1 + \frac{\mu\,t}{\sqrt{n}\;\sigma} + \frac{(\mu^2+\sigma^2)\,t^2}{2n\sigma^2} + R\right]$$

Agora, empregando o desenvolvimento de Mclaurin para $\ln(1+x)$;

$$\ln(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} + \ldots \qquad (\,|x| < 1)$$

Em nosso caso, $x = \frac{\mu\,t}{\sqrt{n}\;\sigma} + \frac{(\mu^2+\sigma^2)\,t^2}{2n\sigma^2} + R$ *, e para n suficientemente grande, o valor absoluto será menor que a unidade. Com isso,*

$$\ln M_{T_n}(t) = -\frac{\sqrt{n}\;\mu\,t}{\sigma} + n\left[\left(\frac{\mu\,t}{\sqrt{n}\;\sigma} + \frac{(\mu^2+\sigma^2)\,t^2}{2n\sigma^2} + R\right) - \frac{1}{2}\left(\frac{\mu t}{\sqrt{n}\;\sigma} + \frac{(\mu^2+\sigma^2)\,t^2}{2n\sigma^2} + R\right)^2 + \cdots\right]$$

Após algumas transformações, teremos

$$\lim_{n\to\infty} \ln M_{T_n}(t) = \frac{t^2}{2}$$

$\Rightarrow \lim_{n\to\infty} M_{T_n}(t) = e^{\frac{t^2}{2}}$ *que é a fgm de uma N(0,1)*

□

Exemplo 12.7. *Suponha que X_i, $i = 1, 2, \cdots, 50$ sejam v.a's independentes, cada uma delas tendo distribuição de Poisson com parâmetro $\lambda = 0,03$. Faça $S = X_1 + \cdots + X_{50}$.*

a) Empregando o Teorema Central do Limite, calcule $P(S \geq 3)$.

b) Compare a resposta de (a) com o valor exato dessa probabilidade.

Solução:

a) Temos que $E(S) = Var(S) = 50 \times \lambda = 1,5$. Logo, $S \sim N(1,5\;;\;1,5)$

$P(S \geq 3) = P(Z \geq 1,22) = 0,11123$

b) $S \sim P(1,5)$

$P(S \geq 3) = 1 - P(S < 3) = 1 - P(S = 0) - P(S = 1) - P(S = 2)$

$\quad\quad = 1 - e^{-1,5} - e^{-1,5} \times (1,5) - \frac{e^{-1,5}(1,5)^2}{2} = 0,19115$

2) (Lista # 4- 3)

$$\begin{cases} n \ \ bolas & \Rightarrow \quad r^n \quad possibilidades \\ r \ \ caixas & \end{cases}$$

$$X_i = \begin{cases} 1, & se \ a \ caixa \ i \ estiver \ vazia \\ 0, & c.c \end{cases}$$

a) $E(X_i) = ?$

$P(X_i = 1) = \frac{(r-1)^n}{r^n} = \left(\frac{r-1}{r}\right)^n = \left(1 - \frac{1}{r}\right)^n$

$P(X_i = 0) = 1 - \left(1 - \frac{1}{r}\right)^n$

$E(X_i) = 1 \times P(X_i = 1) + 0 \times P(X_i = 0) = \left(1 - \frac{1}{r}\right)^n$

b) $E(X_i Y_j) = ? \quad i \neq j$

$$X_i X_j = \begin{cases} 0; & se \ X_i = 0 \ \ e \ \ X_j = 0 \quad ou \quad X_i = 0 \ \ e \ \ X_j = 1 \quad ou \quad X_i = 1 \ \ e \ \ X_j = 0 \\ 1; & se \ X_i = 1 \ e \ X_j = 1 \end{cases}$$

$E(X_i X_j) = 0 \times P(X_i X_j = 0) + 1 \times P(X_i X_j = 1) = P(X_i X_j = 1)$

$E(X_i X_j) = P(X_i X_j = 1) = \frac{(r-2)^n}{r^n} = \left(1 - \frac{2}{r}\right)^n$

c) $S_r = n^o$ *de caixas vazias.*

$S_r = X_1 + \ldots\ldots + X_r$

$E(S_r) = E(X_1) + \ldots\ldots + E(X_r) = r \cdot \left(1 - \frac{1}{r}\right)^n$

d) $Var(S_r) = ?$

$Var(X \pm Y) = Var(X) + Var(Y) \pm 2 \ cov(X, Y) \quad ; \quad cov(X, Y) = E(XY) - E(X)E(Y)$

$Var(X_1 + \ldots\ldots + X_n) = \sum_{i=1}^{n} Var(X_i) + 2 \sum_{i \neq j \, = \, 1}^{n} cov(X_i, X_j)$

$Var(S_r) = Var(X_1 + X_2 + \ldots\ldots + X_r)$

$\qquad = \sum_{i=1}^{r} Var(X_i) + 2 \sum_{i \neq j \, = \, 1}^{r} Cov(X_i, X_j)$

$Var(X_i) = E(X_i^2) - (E(X_i))^2 \qquad ; \qquad i = 1, \ldots\ldots, r$

$E(X_i^2) = 1^2 \times P(X_i = 1) + 0^2 P(X_i = 0) = \left(1 - \frac{1}{r}\right)^n$

$Var(X_i) = \left(1 - \frac{1}{r}\right)^n - \left(1 - \frac{1}{r}\right)^{2n}$

$\qquad = \left(1 - \frac{1}{r}\right)^n \left[1 - \left(1 - \frac{1}{r}\right)^n\right]$

$Cov(X_i, X_j) = E(X_i X_j) - E(X_i)E(X_j) \qquad ; \qquad i \neq j$

$\qquad = \left(1 - \frac{2}{r}\right)^n - \left[1 - \left(1 - \frac{1}{r}\right)^n\right]^2$

$$=\left(1-\tfrac{2}{r}\right)^n - \left(1-\tfrac{1}{r}\right)^{2n}$$

$$Var(S_r) = r\left(1-\tfrac{1}{r}\right)^n\left[1-\left(1-\tfrac{1}{r}\right)^n\right] + 2\times\left(\begin{array}{c} r \\ 2 \end{array}\right)\left[\left(1-\tfrac{2}{r}\right)^n - \left(1-\tfrac{1}{r}\right)^{2n}\right]$$
$$= r\left(1-\tfrac{1}{r}\right)^n\left[1-\left(1-\tfrac{1}{r}\right)^n\right] + r(r-1)\left[\left(1-\tfrac{2}{r}\right)^n - \left(1-\tfrac{1}{r}\right)^{2n}\right]$$

3) (Lista #4 - 9)

$X \sim P(\lambda)$

a) $P\left(X \leq \frac{\lambda}{2}\right) \leq \frac{4}{\lambda}$

$$\begin{aligned} P(X \leq \tfrac{\lambda}{2}) &= 1 - P(X > \tfrac{\lambda}{2}) = 1 - P(X-\lambda > \tfrac{\lambda}{2} - \lambda) \\ &= 1 - P(X-\lambda > -\tfrac{\lambda}{2}) \\ &= 1 - \left[P\left(|X-\lambda| < \tfrac{\lambda}{2}\right) + P(X-\lambda > \tfrac{\lambda}{2})\right] \\ &= 1 - P\left(|X-\lambda| < \tfrac{\lambda}{2}\right) - P\left(X > \tfrac{3}{2}\lambda\right) \\ &= P\left(|X-\lambda| \geq \tfrac{\lambda}{2}\right) - P\left(X > \tfrac{3}{2}\lambda\right) \\ &\leq P(|X-\lambda| \geq \tfrac{\lambda}{2}) \leq \tfrac{Var(X)}{(\lambda/2)^2} = \tfrac{\lambda}{\lambda^2/4} = \tfrac{4}{\lambda} \end{aligned}$$

b) $P(X \geq 2\lambda) \leq \frac{1}{\lambda}$

$$\begin{aligned} P(X \geq 2\lambda) &= P(X-\lambda \geq \lambda) = 1 - P(X-\lambda < \lambda) \\ &= 1 - \left[P(|X-\lambda| < \lambda) + P(X-\lambda < -\lambda)\right] \\ &= P(|X-\lambda| < \lambda) - P(X-\lambda < -\lambda) \\ &= P(|X-\lambda| \geq \lambda) - P(X-\lambda < -\lambda) \\ &\leq P(|X-\lambda| \geq \lambda) \leq \tfrac{Var(X)}{\lambda^2} = \tfrac{\lambda}{\lambda^2} = \tfrac{1}{\lambda} \end{aligned}$$

Capítulo **9**

Cadeias de Markov

9.1 Conceitos iniciais

Neste capítulo trataremos de aplicações da Probabilidade que serve para modelar muitos fenômenos, tais como a evolução de uma colônia de bactérias, um incêndio em uma floresta, movimentação de moléculas de um gás, dentre outros. Consideraremos uma sequência de variáveis aleatórias medidas ao longo do tempo, espaço ou qualquer outra dimensão. Esta sequência recebe o nome de Processos Estocásticos*. A seguir formalizaremos a definição deste processo.*

Definição 9.1. *Um Processo Estocástico (PE) é uma família de v.a. $\{X_t,\ t \in T\}$. Para cada $t \in T$ (Espaço do Parâmetro) X_t representa a "posição" de um sistema que evolui aleatoriamente no tempo.*

O Espaço do Parâmetro pode ser discreto ou contínuo. T é discreto se for enumerável e contínuo se for um subconjunto dos reais. Quando T é discreto o Processo será representado por $\{X_n, n = 0, 1, 2, \cdots\}$. Neste caso particular em que $n \in I\!N$ diremos que estamos em processo estocástico a Tempo Discreto. *Um Processo $\{X_t,\ t \in I\!R\}$ será chamado de Processo Estocástico a* Tempo Contínuo.

O conjunto de valores que as variáveis X_n assumem será chamado de Espaço de Estados *do processo, representado por $I\!E$. Em princípio, o conjunto $I\!E$ será um subconjunto dos inteiros. Se $\{X_n = i\}$, diremos que o processo está no estado i no tempo n.*

Veremos, inicialmente, que estamos tratando de uma sequencia de variáveis aleatórias dependentes, mas com uma dependência bem restrita.

Definição 9.2. *Uma sequência de variáveis aleatórias $X_0, X_1, X_2, \cdots$ definidas em um espago amostral com valores em um conjunto enumerável $I\!E$ é uma Cadeia de Markov (CM) se, para todo $n, n \geq 1$, e estados $i_0, i_1, \cdots, i_{n-1}, i, j$ de $I\!E$, tem-se:*

$$P(X_{n+1} = j|X_n = i, X_{n-1} = i_{i-1}, X_{n-2} = i_{i-2}, \cdots, X_0 = i_0) = P(X_{n+1} = j|X_n = i). \tag{9.1}$$

Esta propriedade nos diz que para conhecer a vida futura (X_{n+1}) do processo, conhecidos o presente (X_n) e o passado ($X_{n-1}, X_{n-2}, \cdots, X_0$), é suficiente conhecer o presente, não importando o passado. Ou seja, não importa por quais estados ele passou antes de chegar ao estado i. Em outras palavras, o Presente já contém toda a informação do passado.

Em algumas situações a probabilidade de passar do estado i para o estado j depende da etapa n, ou seja,

$$P(X_{n+1} = j|X_n = i) = P_{ij}(n).$$

No entanto, o caso mais comum é quando não há essa dependência. Suponha, então, que quando o processo está no estado i no tempo n ele poderá passar ao estado j no tempo $n+1$ com probabilidade P_{ij}, independente de n,

$$P(X_{n+1} = j|X_n = i) = P_{ij}.$$

então este processo será chamado de Cadeia de Markov com probabilidade de transição estacionária*. O valor P_{ij} será chamado de* Probabilidade de Transição*, satisfazendo*

$$P_{ij} \geq 0,\ \forall i, j \geq 0; \quad \sum_{j=0}^{\infty} P_{ij} = 1.$$

A Matriz de Probabilidade de Transição em um passo *será denotada por $\boldsymbol{P}$, com*

$$\boldsymbol{P} = \begin{pmatrix} P_{00} & P_{01} & P_{02} & \cdots & P_{0n} \\ P_{10} & P_{11} & P_{12} & \cdots & P_{1n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ P_{n0} & P_{n1} & P_{n2} & \cdots & P_{nn} \end{pmatrix},$$

Exemplo 9.1. *Consideremos uma cadeia com 5 estados: $\mathbb{E} = \{0, 1, 2, 3, 4\}$ e matriz de transição dada por*

$$\mathbf{P} = \begin{array}{c} \\ 0 \\ 1 \\ 2 \\ 3 \\ 4 \end{array} \begin{array}{c} \begin{array}{ccccc} \;0\; & \;\;1\;\; & \;\;2\;\; & \;\;3\;\; & \;\;4\; \end{array} \\ \begin{pmatrix} 1 & 1 & 0 & 0 & 0 \\ 1/2 & 0 & 1/2 & 0 & 0 \\ 0 & 1/2 & 0 & 1/2 & 0 \\ 0 & 0 & 1/2 & 0 & 1/2 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \end{array}.$$

Quando necessário, usaremos um gráfico para representar o movimento da cadeia, do seguinte tipo:

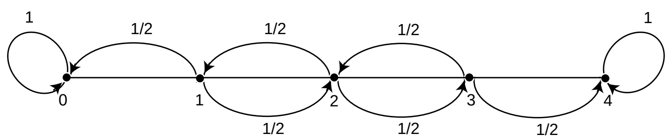


Figura 9.1 *Representação Gráfica de uma CM.*

Exemplo 9.2 (Passeio Aleatório em $\mathbb{Z}$)**.** *. Consideremos uma partícula inicialmente na origem. A cada instante a partícula dá um salto para a direita com probabilidade p ou para a esquerda com probabilidade $1-p$. Seja X_n a posição da partícula no instante n (ou seja, após n saltos). Temos que*

$$X_1 = \begin{cases} -1, & q \\ 1, & p \end{cases} \qquad X_2 = \begin{cases} -2, & q^2 \\ 0, & 2pq \\ 2, & p^2 \end{cases}$$

Esta é um Processo Estocástico com $T = \mathbb{N}$ e $\mathbb{E} = \mathbb{Z}$.
Pergunta: quanto vale $P(X_n = k | X_0 = 0)$ e $P(X_n = k | X_0 = i)$?

Exemplo 9.3 (Soma de v.a.i.i.d.)**.** *Seja $X_0, X_1, X_2, \cdots$ uma sequência de v.a. independentes e identicamente distribuídas, definidas em um espaço S e assumindo valores em $\mathbb{E}$. Seja, ainda*

$$Y_n = X_0 + X_1 + \cdots + X_n.$$

Então, $\{Y_n, n = 0, 1, 2, \cdots\}$ é uma cadeia de Markov.

Para verificar esta afirmação, temos que mostrar que a Equação (9.1) vale neste caso. De fato, temos que:

$$Y_{n+1} = Y_n + X_{n+1}.$$

Observemos que dar a sequência $Y_0 = i_0, Y_1 = i_1, \cdots, Y_n = i_n$ *é equivalente a dar a sequência* $X_0 = i_0, X_1 = i_1 - i_0, .., X_n = i_n - i_{n-1}$. *Asim,*

$$\begin{aligned}
&P[Y_{n+1} = i_{n+1} | Y_n = i_n,\ Y_{n-1} = i_{n-1},\ ..,\ Y_0 = i_0] = \\
&P[Y_n + X_{n+1} = i_{n+1} | Y_n = i_n,\ Y_{n-1} = i_{n-1},\ ..,\ Y_0 = i_0] = \\
&P[X_{n+1} = i_{n+1} - i_n | X_n = i_n - i_{n-1},\ ..,\ X_1 = i_1 - i_0,\ X_0 = i_0] = \\
&P[X_{n+1} = i_{n+1} - i_n]
\end{aligned}$$

que decorre do fato das variáveis da sequência $X_0, X_1, ...$ *são independentes. Decorre que:*

$$P[Y_{n+1} = i_{n+1} | Y_n = i_n] = P[X_{n+1} = i_{n+1} - i_n],$$

o que prova nossa afirmação. □

Exemplo 9.4 (Ruina do jogador). *O Jogador A participa de uma sequência de jogos com o jogador B. As jogadas são independentes e em cada uma delas o jogador A ganha uma unidade monetária (digamos, um dolar) de B com probabilidade* p *e perde uma unidade monetária com probabilidade* $1 - p$. *Se designarmos por* X_n *o ganho do jogador A na n-ésima jogada, temos uma situação particular do exemplo anterior, em que as variáveis aleatórias* X_i *assumem os valores* $+1$ *e* -1 *com probabilidades* p *e* $1-p$, *respectivamente. Assim, a sucessão de fortunas do jogador A é uma CM, com* $\mathbb{E} = \{0, 1, 2, \cdots, N\}$, T *discreto e cujos elementos da matriz* $\boldsymbol{P}$ *são dados por*

$$P_{ij} = \begin{cases} 1 & \textit{se i=j=0 ou i=j=N} \\ p & \textit{se j=i+1} \\ q & \textit{se j=i-1} \end{cases}$$

Ou seja,

$$\mathbf{P} = \begin{array}{c} \\ 0 \\ 1 \\ \vdots \\ N-1 \\ N \end{array} \begin{array}{c} \begin{array}{ccccccc} 0 & 1 & 2 & \cdots & N-2 & N-1 & N \end{array} \\ \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 & 0 & 0 \\ q & 0 & p & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & q & 0 & p \\ 0 & 0 & 0 & \cdots & 0 & 0 & 1 \end{pmatrix} \end{array}$$

Exemplo 9.5 (Previsão do tempo). *Consideremos que o fato de chover, ou não, hoje depende apenas do fato de ter chovido, ou não, ontem. Suponhamos que se choveu ontem, choverá hoje com probabilidade* α; *se não choveu ontem, não choverá hoje com probabilidade* β. *Este problema pode ser formulado em termos de uma Cadeia de Markov? Por que? Em caso positivo, qual a matriz de probabilidades de transição* $\boldsymbol{P}$.

Seja $\{X_n\}_{n \geq 0}$ *o estado climático no n-ésimo dia, com* $\mathbb{E} = \{0, 1\}$, *e* T *discreto, onde o estado 0 representa* chuva *(C) e 1 representa* não chuva *(*$\overline{C}$*). Pelo próprio enunciado já vemos que é uma CM,* *pois a situação em um dia só depende do dia anterior. A matriz de probabilidades de transição é dada por*

$$\boldsymbol{P} = \begin{array}{c} \\ C \\ \overline{C} \end{array} \begin{array}{c} \begin{array}{cc} C & \overline{C} \end{array} \\ \begin{pmatrix} \alpha & 1-\alpha \\ 1-\beta & \beta \end{pmatrix} \end{array}$$

Exemplo 9.6 (Previsão do tempo II)**.** *Suponha que o fato de chover hoje dependa das condições climáticas de dois dias anteriores. Especificamente, suponha que se choveu ontem e hoje, amanhã irá chover com probabilidade 0,7; se choveu hoje mas ontem não, amanhã irá chover com probabilidade 0,5; se choveu ontem mas hoje não então amanhã irá chover com probabilidade 0,4; se não choveu nem ontem nem hoje, amanhã irá chover com probabilidade 0,2.*

Considerando os estados 0 (CC). 1 ($C\overline{C}$), 2 ($\overline{C}C$) e 3 ($\overline{CC}$), vemos que está satisfeita a condição markoviana. Além disso, do estado 0 só podemos passar para os estados 0 e 1; do estado 1 só podemos passar para os estados 2 e 3, do estado 2 só podemos passar aos estados 0 e 1, e do estado 3 para os estados 2 e 3. A matriz de transição é dada por

$$\boldsymbol{P} = \begin{array}{c} \\ 0 \\ 1 \\ 2 \\ 3 \end{array} \begin{array}{c} \begin{array}{cccc} 0 & 1 & 2 & 3 \end{array} \\ \begin{pmatrix} 0,7 & 0,3 & 0 & 0 \\ 0 & 0 & 0,4 & 0,6 \\ 0,5 & 0,5 & 0 & 0 \\ 0 & 0 & 0,8 & 0,2 \end{pmatrix} \end{array}$$

Exemplo 9.7 (Modelo de Humor)**.** *Um dia na vida de Gary pode se animado (C), mais-ou-menos(S) ou triste (G). Se o dia de hoje é animado então o de amanha sera C, S ou G com respectivas probabilidades 0,5, 0,4, 0,1. Se hoje ele está mais-ou-menos, amanhã será C, S ou G com respectivas probabilidades 0,3, 0,4, 0,3. E se hoje ele está triste amanhã será C, S ou G com respectivas probabilidades 0,2, 0,3, 0,5. Chamando X_n o humor de Gary no n-ésimo dia, entao $\{X_n, n \geq 0\}$ é uma cadeia de Markov de 3 estados (estado 0 = C, estado 1 = S, estado 2 = G) com matriz de probabilidade de transicao de estados dada por*

$$\boldsymbol{P} = \begin{array}{c} \\ C \\ S \\ G \end{array} \begin{array}{c} \begin{array}{ccc} C & S & G \end{array} \\ \begin{pmatrix} 0,5 & 0,4 & 0,1 \\ 0,3 & 0,4 & 0,3 \\ 0,2 & 0,3 & 0,5 \end{pmatrix} \end{array}$$

Exemplo 9.8 (Modelo de Estoques)**.** *Consideremos que as demandas D de um certo produto A em uma determinada loja são v.a.i.i.d. com distribuição Poisson(λ). Designemos por $X_n, n \geq 1$, os níveis de estoque ao final do n-ésimo dia e por S a capacidade do estoque. A reposição do estoque é feita da seguinte forma: se no final do dia i o nível de estoque X_n é menor que s (um número pré-determinado), então são ordenadas $S - X_n$ peças para reposição, de modo que no dia seguinte o estoque inicial será S; se tivermos $s \leq X_n \leq S$, então não é feita nenhuma reposição de estoque. Temos, portanto,*

$$X_{n+1} = \begin{cases} X_n - D_{n+1} & ; \quad s \leq X_n \leq S \\ S - D_{n+1} & ; \quad X_n < s \end{cases} \tag{9.2}$$

Seja X_0 o nível de estoque no primeiro dia de atividade da loja. $\{X_n\}_{n\geq 0}$ é uma CM? Se for, determine sua matriz de probabilidades de transição.
Obs: *Estoques negativos no final de uma dia serão interpretados como peças que seriam vendidas caso houvessem peças suficientes no estoque.*

Para verificar a veracidade da afirmação, basta notar que $X_{n+1} = X_n - D_{n+1}$.
Seja X_0 o nivel de estoque no primeiro dia de atividade da loja. A sequência de estoques $X_0, X_1, X_2, \cdots$ é uma Cadeia de Markov. De fato, a distribuição de probabilidade de X_{n+1} só depende de X_n e da distribuição de probabilidade de D_{n+1}, que é independente de $X_1, X_2, \cdots, X_n$. Temos, usando (9.2):

Para $s \leq j \leq S$

$$\begin{aligned} P[X_{n+1} = k|X_n = j] &= P[X_n - D_{n+1} = k|X_n = j] = P[D_{n+1} = X_n - k|X_n = j] \\ &= P[D_{n+1} = j - k] = \frac{e^{-\lambda}\lambda^{j-k}}{(j-k)!} \end{aligned}$$

Para $j < s$

$$\begin{aligned} P[X_{n+1} = k|X_n = j] &= P[S - D_{n+1} = k|X_n = j] = P[S - D_{n+1} = k] \\ &= P[D_{n+1} = S - k] = \frac{e^{-\lambda}\lambda^{S-k}}{(S-k)!}, \end{aligned}$$

que são os elementos da matriz de probabilidades de transição da Cadeia.
O modelo que descreverenos nesse exemplo ocorre em várias areas, tais como: física, biologia, genética, pesquisa operacional para citar algumas. Mas descreveremos em termos de partícula. Para aplicá-lo à genética basta substituir partícula por gen.

Exemplo 9.9 (**Processo de Ramificação**). *Seja X_0 o número de partículas que formam a Geração inicial, ou Geração Zero. Cada partícula pode produzir novas partículas descendentes. Designemos por Z_i, a variável aleatória igual ao número de descendentes da particula i. Suporemos que essas variáveis aleatórias são independentes e identicamente distribuidas. Com isso, seja X_n o número de partículas que formam a Geração n. Pode-se dizer que $\{X_n\}_{n\geq 0}$ é uma CM?*

Considando a sequência $X_0, X_1, \cdots, X_n, \cdots$, teremos que X_n é uma Cadeia de Markov. De fato, decorre da hipótese de independência dos dependentes de diferentes partículas que se soubermos que $X_n = i$, não importando como o processo chegou a i, a probabilidade de termos j partículas na geração $n+1$ será:

$$P_{ij} = P[X_{n+1} = j|X_n = i] = P[Z_1 + Z_2 + \ldots + Z_i = j]$$

Como a distribuição dos Z_i é conhecida e eles são independentes, nos podemos determinar a distribuição da soma e obter a matriz de probabilidades de transição.

Exemplo 9.10 (**Modelo de Ehrenfest**). *P. e T. Ehrenfest propuseram um modelo, que descreveremos de modo simplificado, para representar as trocas de calor entre duas regiões em um volume de um gás. Aqui as regiões são representadas por duas urnas e as moléculas do gás serão representadas por bolas. As urnas A e B contém um total de $2N$ bolas. Assim, se a urna A contém k bolas, então a urna B conterá $2N-k$. Uma bola é sorteada ao acaso entre as $2N$. Se a bola sorteada estiver na urna A ela é transferida para a urna B e se estiver na urna B é transferida para A. Seja X_0 o número inicial de bolas na urna A e designemos por X_n o numero de bolas nessa urna após o n-ésimo sorteio.*

$\{X_n,\ n \geq 0\}$ é uma Cadeia de Markov cujo espaço de estados é o conjunto $\mathbb{E} = \{0, 1, 2, \cdots, 2N\}$. De fato, basta observar que se numa dada etapa existem i bolas na urna A então na próxima etapa número de bolas nessa urna será $i-1$ com probabilidade $i/2N$ e $i+1$ com probabilidade $1 - i/2N$, independentemente de como chegou-se a essa situação. A matriz de probabilidadae de transição da Cadeia de Ehrenfest é dada por:

$$\boldsymbol{P} = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 & 0 & 0 \\ \frac{1}{2N} & 0 & 1-\frac{1}{2N} & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1-\frac{1}{2N} & 0 & \frac{1}{2N} \\ 0 & 0 & 0 & \cdots & 0 & 1 & 0 \end{pmatrix},$$

9.2 Probabilidades de transição em n etapas

Seja $X_0, X_1, \cdots$ una cadeia de Markov com matriz de probabilidades de transição $\boldsymbol{P}$. Vamos determinar as probabilidades de transição dessa cadeia em n etapas, isto é, $P[X_n = j|X_0 = i]$, que denotaremos por $P_{ij}^{(n)}$. Mostraremos que para todo $n \geq 2$

$$P_{ij}^{(n)} = \sum_{k \in E} P_{ik}^{(n-1)} P_{kj}$$

Para $n = 2$,

$$\begin{aligned} P_{ij}^{(2)} &= P[X_2 = j|X_0 = i] \\ &= \sum_{k \in E} P[X_2 = j,\ X_1 = k|X_0 = i] \\ &= \sum_k P[X_1 = k|X_0 = i] P[X_2 = j|X_1 = k,\ X_0 = i] \\ &= \sum_k P[X_1 = k|X_0 = i] P[X_2 = j|X_1 = k] \\ &= \sum_k P_{ik} P_{kj} \end{aligned}$$

Notemos que $P_{ij}^{(2)} = P_{ij}^2$, que é o elemento ij do quadrado da matriz $\boldsymbol{P}$. Para $n > 2$, temos:

$$\begin{aligned} P_{ij}^{(n)} &= P[X_n = j|X_0 = i] \\ &= \sum_k P[X_n = j,\ X_{n-1} = k|X_0 = i] \\ &= \sum_k P[X_{n-1} = k|X_0 = i] P[X_n = j|X_{n-1} = k,\ X_0 = i] \\ &= \sum_k P[X_{n-1} = k|X_0 = i] P[X_n = j|X_{n-1} = k] \\ &= \sum_k P_{ik}^{(n-1)} P_{kj} \qquad (9.3) \end{aligned}$$

Ponhamos $n = 3$ na expressão (9.3). Como $P_{ij}^{(2)} = P_{ij}^2$ segue-se dessa expressão que $P_{ij}^{(3)} = P_{ij}^3$. Prosseguindo iteradamente, vemos que para todo $n \geq 2$, $P_{ij}^{(n)} = P_{ij}^n$, ou seja, que a matriz de probabilidades de transição em n etapas é $\boldsymbol{P}^n$, a n-ésima potência da matriz de probabilidades de transição em uma etapa.

Lema 1. *Para quaisquer inteiros n e m, tem-se:*

$$P_{ij}^{(n+m)} = \sum_k P_{ik}^{(n)} P_{kj}^{(m)} \tag{9.4}$$

Para provar vamos repeteir o raciocínio utilizado para verificar a Expresão (9.3). De fato,

$$\begin{aligned} P[X_{n+m} = j | X_0 = i] &= \sum_k P[X_{n+m} = j,\ X_n = k | X_0 = i] \\ &= \sum_k P[X_{n+m} = j | X_n = k,\ X_0 = i] P[X_n = k | X_0 = i] \\ &= \sum_k P[X_{n+m} = j | X_n = k] P[X_n = k | X_0 = i] \\ &= \sum_k P_{ik}^{(n)} P_{kj}^{(m)} \end{aligned}$$

A Expressão (9.4) é denominada Equações de Chapman-Kolmogorov.
Observe que em termos matriciais a equação de Chapman-Kolmogorov é a expressão de $\boldsymbol{P}^{n+m}$ como o produto de $\boldsymbol{P}^n$ por $\boldsymbol{P}^m$. Probabilisticamente a equação de Chapman-Kolmogorov diz que: para a cadeia fazer uma transição do estado i ela precisa fazer uma transiçao em n etapas para algum estado k seguida de uma outra transição em m etapas do estado k para o estado j.

Lema 2. *Uma distribuição inicial para os estados de uma cadeia de Markov e sua matriz de probabilidades de transição determinam para todo inteiro $n \geq 1$ as distribuições de dimensão finita da Cadeia.*

Seja $\pi_i^{(0)} = P[X_0 = i]$, para todo $i \in E$ seja $\boldsymbol{P} = (P_{ij})$ para $(i,j) \in E \times E$ a matriz de probabilidades de transição da Cadeia.

$$\begin{aligned} &P[X_0 = i_0,\ X_1 = i_1,\ \ldots,\ X_n = i_n] = \\ &= P[X_0 = i_0] P[X_1 = i_1 | X_0 = i_0] \cdots P[X_n = i_n | X_0 = i_0,\ X_1 = i_1,\ ..X_{n-1} = i_{n-1}] \\ &= P[X_0 = i_0] P[X_1 = i_1 | X_0 = i_0] \cdots P[X_n = i_n | X_{n-1} = i_{n-1}] \\ &= \pi_{i_0}^{(0)} P_{i_0,i_1} \cdots P_{i_{n-1},i_n} \end{aligned}$$

Vamos, a titulo de exemplo, calcular algumas potências de uma matriz de probabilidades de transição de uma cadeia de Markov com dois estados. O que observamos nesse exemplo será discutido em profundidade mais adiante.

Exemplo 9.11. *Uma cadeia de Markov com dois estados 0 e 1 é utilizada como modelo para representar mudanças atmosféricas. O estado 0 designa um dia em que não chove, e o estado 1 um dia que chove. A matriz de probabilidades de transição é dada por:*

$$\boldsymbol{P} = \begin{pmatrix} \frac{3}{4} & \frac{1}{4} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}$$

Calculando-se o produto da matriz $\boldsymbol{P}$ *por* $\boldsymbol{P}$*, linha por coluna obtemos* $\boldsymbol{P}^2$*, multiplicando- se* $\boldsymbol{P}^2$ *por* $\boldsymbol{P}$ *obtemos* $\boldsymbol{P}^3$*, multiplicando-se* $\boldsymbol{P}^3$ *por* $\boldsymbol{P}^2$ *obtemos* $\boldsymbol{P}^5$e *multiplicando-se* $\boldsymbol{P}^5$ *por si mesma obtemos* $\boldsymbol{P}^{10}$

$$\boldsymbol{P}^2 = \begin{pmatrix} \frac{11}{16} & \frac{5}{16} \\ \frac{5}{8} & \frac{3}{8} \end{pmatrix}$$

$$\boldsymbol{P}^3 = \begin{pmatrix} \frac{43}{64} & \frac{21}{64} \\ \frac{21}{32} & \frac{11}{32} \end{pmatrix}$$

$$\boldsymbol{P}^5 = \begin{pmatrix} 0.6689 & 0.3330 \\ 0.6660 & 0.3339 \end{pmatrix}$$

$$\boldsymbol{P}^5 = \begin{pmatrix} 0.6667 & 0.3333 \\ 0.6667 & 0.3333 \end{pmatrix}$$

A partir da distribuição inicial $p_i, i \in E$*, e da matriz de probabilidades de transição podemos determinar a distribuição de probabilidade de* X_n*.*

$$\begin{aligned} P[X_n = j] &= P[\bigcup_i \{X_n = j,\ X_0 = i\}] = \sum_i P[X_n = j,\ X_0 = i] \\ &= \sum_i P[X_n = j | X_0 = i] P[X_0 = i] = \sum_i \pi_i^{(0)} P_{ij}^{(n)}. \end{aligned} \tag{9.5}$$

9.3 Estrutura de uma cadeia de Markov

Definição 9.3. *O estado* i *conduz ao estado* j*, que será denotado por* $i \to j$*, se e somente se, existe um* $n \geq 1$ *tal que* $P_{ij}^n > 0$

Definição 9.4. *O estado* i *se comunica com o estado* j *se* $i \to j$ *e* $j \to i$*. Usaremos a notação* $i \leftrightarrow j$ *para indicar que* i *se comunica com* j*.*

A relação $i \to j$ *é transitiva, ou seja, se* i *conduz a* j *e* j *conduz a* k*, então* i *conduz a* k*. De fato, da equação de Chapman-Kolmogorov segue que:*

$$P_{ik}^{(n+m)} \geq P_{ij}^{(n)} P_{jk}^{(m)}$$

Decorre desse fato que a relação $i \leftrightarrow j$ *é também transitiva. Segue da definição que esta relação é ainda simétrica e reflexiva, sendo assim uma relação de equivalência no conjunto dos estados da Cadeia de Markov. Essa relação de equivalência divide o conjunto dos estados em classes de equivalência e nosso objetivo agora é descrever o comportamento dos estados que pertencem a essas classes de equivalência.*

9.3.1 Probabilidade de primeira passagem ou retorno

Para um estado genérico j *vamos introduzir a variável aleatória* T_j *que é o instante da primeira visita a* j*, quando a Cadeia parte de um estado* $i, i \neq j,$ *e é igual ao instante do primeiro retorno a* j *quando a Cadeia parte de* j*.*

$$T_j = \min\{n \geq 1 | X_n = j\}$$

Se não existir um inteiro n *satisfazendo essa condição põe-se* $T_j = \infty$*. Vamos denotar por* P_i *a* *distribuição de probabilidade associada à Cadeia que inicia o seu movimento no estado i. Com essa notação temos a igualdade :*

$$P_i[X_n = j] = P[X_n = j|X_0 = i]$$

Vamos denotar por $f_{ij}^{(n)}$ a probabilidade que a Cadeia que parte do estado i visite o estado j pela primeira vez no instante n, isto é :

$$f_{ij}^{(n)} = P_i[T_j = n]$$

Denotemos por f_{ij} a probabilidade que a cadeia que parte de i visite j em algum instante. Temos:

$$f_{ij} = \sum_{n=1}^{\infty} f_{ij}^{(n)} = \sum_{n=1}^{\infty} P_i[T_j = n] = P[T_j < \infty]$$

Observemos que:

(1) Se a Cadeia parte de i ela nunca visita j com probabilidade $P_i[T_j = \infty] = 1 - f_{ij}$

(2) $f_{ij}^{(1)} = P_i[X_1 = j] = P_{ij}$

(3) $f_{ij}^{(n)} = P_i[X_r \neq j,\ 1 \leq r \leq (n-1),\ X_n = j]$

(4) Como a cadeia tem probabilidades de transição estacionárias segue-se que:

$$f_{ij}^{(n)} = P[X_{m+r} \neq j,\ 1 \leq r \leq n-1,\ X_{n+m} = j|X_m = i]$$

para todo $m \geq 1$ e todo $n \geq 1$.

Teorema 41. *Para todo i e j de $\mathbb{E}$ e todo n, $1 \leq n \leq \infty$, temos:*

$$P_{ij}^{(n)} = \sum_{r=1}^{n} f_{ij}^{(r)} P_{jj}^{(n-r)}$$

Demonstração:

$$\begin{aligned}
P_{ij}^{(n)} &= P_i[X_n = j] = P_i[(\bigcup_{r=1}^{n}[T_j = r]) \cap X_n = j] \\
&= \sum_{r=1}^{n} P_i[T_j = r,\ X_n = j] = \sum_{r=1}^{n} P_i[T_j = r] P_i[X_n = j|T_j = r] \\
&= \sum_{r=1}^{n} P_i[T_j = r] P_i[X_n = j|X_1 \neq j,\ \ldots,\ X_{r-1} \neq j,\ X_r = j] \\
&= \sum_{r=1}^{n} P_i[T_i = r] P_j[X_{n-r} = j]
\end{aligned}$$

Assim,

$$P_{ij}^{(n)} = \sum_{r=1}^{n} f_{ij}^{(r)} P_{jj}^{(n-r)}$$

9.4 classificação de estados de uma Cadeia de Markov

Seja $\{X_n\}$ uma CM com espaço de estados $\mathbb{E}$ (finito ou não) e matriz de transição de probabilidades $\boldsymbol{P}$.

Definição 9.5. *Um estado i de uma Cadeia de Markov é dito* ***recorrente*** *se a Cadeia partindo de i voltar a ele com probabilielade 1.*

Com a notação que introduzimos na seção anterior, o estado i ser recorrente significa que $f_{ii} = 1$.

Definição 9.6. *O estado i de uma Cadeia de Markov é dito* ***absorvente*** *se $P_{ii} = 1$. Ou seja, depois que a cadeia entra nele não conseguie mais sair.*

Definição 9.7. *O estado i de uma Cadeia de Markov é dito* ***transitório*** *se a Cadeia partindo de i voltar a i com probabilidade estritamente menor que 1.*

Definição 9.8. *O estado i de uma Cadeia de Markov é dito ter* ***período*** *d se a Cadeia partindo de i voltar a i em tempos múltiplos de d. Quando não existir período definido a cadeia é dita ser* ***aperiódica****.*

Exemplo 9.12.

$$\mathbf{P} = \begin{array}{c} \\ 0 \\ 1 \\ 2 \end{array} \begin{array}{c} \begin{array}{ccc} 0 & 1 & 2 \end{array} \\ \begin{pmatrix} 1/2 & 1/2 & 0 \\ 1/2 & 1/4 & 1/4 \\ 0 & 1/3 & 2/3 \end{pmatrix} \end{array}$$

Exemplo 9.13.

$$\mathbf{P} = \begin{array}{c} \\ 0 \\ 1 \\ 2 \\ 3 \end{array} \begin{array}{c} \begin{array}{cccc} 0 & 1 & 2 & 3 \end{array} \\ \begin{pmatrix} 1/2 & 1/2 & 0 & 0 \\ 1/2 & 1/2 & 0 & 0 \\ 1/4 & 1/4 & 1/4 & 1/4 \\ 0 & 0 & 0 & 1 \end{pmatrix} \end{array}$$

Neste exemplo podemos formar três classes: $C_1 = \{0, 1\}$, $C_2 = \{2\}$ e $C_3 = \{3\}$.

Exemplo 9.14. *Exemplo de uma cadeia periódica com $d = 2$.*

$$\mathbf{P} = \begin{array}{c} \\ 1 \\ 2 \end{array} \begin{array}{c} \begin{array}{cc} 1 & 2 \end{array} \\ \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \end{array}$$

Com a notação introdnzida tem-se para um estado transitório $i, f_{ii} < 1$.

Proposição 1. *Se o estado i é transitório, a Cadeia partindo de i volta a i um número finito de vezes.*

De fato, cada vez que a cadeia volta ao estado i, a evolução futura é independente do passado e existe probabilidade $1 - f_{ii}$ da cadeia nunca mais voltar a i, logo a cadeia volta a i exatamente n vêzes com probabilidade $(f_{ii})^n(1 - f_{ii})$, para $n = 1, 2, \ldots$. Observemos que se o estado i é recorrente, a cadeia partindo de i volta a i infinitas vezes com probabilidade 1. De fato, como a cadeia partindo de i volta a i com probabilidade um, e pela propriedade Markoviana a cada retorno é como se estivesse iniciando o movimento, segue-se que ela volta infinitas vezes com probabilidade 1.

Definição 9.9. *Um conjunto A de estados de uma Cadeia de Markov é dito* ***fechado*** *se nenhum estado de A conduz a algum estado de A^c.*

Se o conjunto fechado A é formado por um único estado i entao este estado é dito absorvente.

Definição 9.10. *Uma Cadeia de Markov é dita* ***irredutível*** *se não existe nenhum suconjunto fechado diferente do conjunto de todos os estados. Ou seja, se todos os estados se comunicam.*

Proposição 2. *Se i é um estado recorrente $\sum_{n=1}^{\infty} P_{ii}^{(n)} = \infty$. Se i é transitório $\sum_{n=1}^{\infty} P_{ii}^{(n)} < \infty$.*

Demonstração. Denotemos por N_{ii} número de retornos a i partindo de i. Pela observação vimos que N_{ii} é infinito com probabilidade 1 e portanto $E_i(N_{ii}) = \infty$.

Mas $N_{ii} = \sum_{n=1}^{\infty} I_{\{i\}}(X_n)$, onde $I_{\{i\}}(X_n) = 1$ se $X_n = i$ e 0 se $X_n \neq i$. Com isso,

$$E_i(N_{ii}) = \sum_{n=1}^{\infty} E_i(I_{\{i\}}(X_n)) = \sum_{n=1}^{\infty} P_i[X_n = i] = \sum_{n=1}^{\infty} P_{ii}^n$$

o que mostra que $\sum_{n=1}^{\infty} P_{ii}^n = \infty$.

Se i é transitório observemos que:

$$\begin{aligned} E_i(N_{ii}) &= \sum_{n=1}^{\infty} n(f_{ii})^n(1-f_{ii}) = (1-f_{ii})f_{ii}\sum_{n=1}^{\infty} n(f_{ii})^{n-1} \\ &= \frac{(1-f_{ii})f_{ii}}{(1-f_{ii})^2} = \frac{f_{ii}}{(1-f_{ii})} < \infty. \end{aligned}$$

□

Teorema 42. *Se $\boldsymbol{P}$ é uma matriz de transição aperiódica, então:*

(i) As potências $\boldsymbol{P}^n$ convergem para uma matriz A;

(ii) As linhas de A são iguais.

9.5 Distribuição estacionária

Nesta seção determinaremos as medidas de probabilidades invariantes, ou de equilíbio, ou estacionárias de uma cadeia de Markov.

Definição 9.11. *Seja $\{X_n\}$ uma cadeia de Markov com $\mathbb{E} = \{0, 1, \cdots \mathbb{N}\}$ (finito) e matriz de transição $\boldsymbol{P}$. Um vetor de probabilidades $\boldsymbol{\pi} = (\pi_0, \pi_1, \cdots, \pi_N)$ é dito ser a* ***distribuição estacionária*** *para a cadeia se*

$$\boldsymbol{\pi}\boldsymbol{P} = \boldsymbol{\pi}, \tag{9.6}$$

Ou seja, se $\pi_j = \sum_{i \in \mathbb{E}} P_{ij}\pi_i$, sujeto às restrições $\pi_j \geq 0$ e $\sum_{j \in \mathbb{E}} \pi_i = 1$.

Sendo $\boldsymbol{\pi}$ *estacionária, então*

$$
\begin{aligned}
\boldsymbol{\pi}\boldsymbol{P}^2 &= \boldsymbol{\pi}(\boldsymbol{P}\boldsymbol{P}) = (\boldsymbol{\pi}\boldsymbol{P})\boldsymbol{P} = \boldsymbol{\pi}\boldsymbol{P} = \boldsymbol{\pi} \\
\boldsymbol{\pi}\boldsymbol{P}^3 &= \boldsymbol{\pi}(\boldsymbol{P}^2\boldsymbol{P}) = (\boldsymbol{\pi}\boldsymbol{P}^2)\boldsymbol{P} = \boldsymbol{\pi}\boldsymbol{P} = \boldsymbol{\pi} \\
&\vdots \\
\boldsymbol{\pi}\boldsymbol{P}^n &= \boldsymbol{\pi}(\boldsymbol{P}^{n-1}\boldsymbol{P}) = (\boldsymbol{\pi}\boldsymbol{P}^{n-1})\boldsymbol{P} = \boldsymbol{\pi}\boldsymbol{P} = \boldsymbol{\pi}
\end{aligned}
$$

Portanto, $\boldsymbol{\pi} = \boldsymbol{\pi}\boldsymbol{P}^n$, $\forall n \geq 1$. *E já vimos que* $\boldsymbol{P}^n$ *converge para uma matriz cujas linhas são todas iguais. Estas linhas são justamente iguais a* $\boldsymbol{\pi}$ *e cada elemento representa o tempo médio que a CM passa em cada estado.*

Observação 9.1.

i) Cada probabilidade π_j *é usualmente identificado como a proporção do tempo que o processo permanece no estado* j*.*

ii) Usando a notação vista anteriormente, temos que

$$\pi_j = \lim_{n\to\infty} \pi_j^{(n)}, \qquad \textit{onde } \pi_j^{(n)} = P(X_n = j)$$

$$\boldsymbol{\pi} = \lim_{n\to\infty} \boldsymbol{\pi}_n \qquad \textit{onde } \boldsymbol{\pi}_n = (\pi_0^{(n)}, \pi_1^{(n)}, \cdots, \pi_N^{(n)})$$

iii) sempre podemos ignorar uma das equações, substituindo-a por $\displaystyle\sum_{j\in\mathbb{E}} \pi_i = 1$.

Exemplo 9.15. *Consideremos a CM com matriz de transição abaixo. Obter a distribuição estacionaria.*

$$\mathbf{P} = \begin{array}{c c} & \begin{array}{cc} 1 & 2 \end{array} \\ \begin{array}{c} 1 \\ 2 \end{array} & \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \end{array}$$

Já vimos que esta cadeia é periódica com $d = 2$. *Com isso, a cadeia passa metade do tempo em cada um dos estados, o que resultará em* $\pi_0 = \pi_1 = \frac{1}{2}$. *Temos que resolveer o sistema*

$$(\pi_0, \pi_1) = (\pi_0, \pi_1)\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

que resultará no sistema $\begin{cases} \pi_0 = \pi_1 \\ \pi_0 + \pi_1 = 1 \end{cases}$ *cuja solução é* $\pi_0 = \pi_1 = \frac{1}{2}$.

Exemplo 9.16. *Consideremos a CM para o modelo de chuvas com matriz de transição abaixo. Obter a distribuição estacionaria.*

$$\mathbf{P} = \begin{array}{c c} & \begin{array}{cc} 1 & 2 \end{array} \\ \begin{array}{c} 1 \\ 2 \end{array} & \begin{pmatrix} \alpha & 1-\alpha \\ 1-\beta & \beta \end{pmatrix} \end{array}$$

Temos que resolveer o sistema

$$(\pi_0, \pi_1) = (\pi_0, \pi_1) \begin{pmatrix} \alpha & 1-\alpha \\ 1-\beta & \beta \end{pmatrix}$$

que resultará no sistema $\begin{cases} \pi_0 = \alpha\pi_0 + (1-\beta)\pi_1 \\ \pi_0 + \pi_1 = 1 \end{cases}$

Para o exemplo citado, com a matriz a seguir, a solução é

$$\mathbf{P} = \begin{array}{c} \\ 1 \\ 2 \end{array} \begin{array}{c} \begin{array}{cc} 1 & 2 \end{array} \\ \begin{pmatrix} 0,7 & 0,3 \\ 0,4 & 0,6 \end{pmatrix} \end{array}$$

Exercício 9.5.1. *Determinar a distribuição estacionária para o modelo de humor com 3 estados apresentado no Exemplo 9.7.*

Definição 9.12. *Seja j um estado da CM e m_{jj} o número de transições esperado da cadeia, partindo de j, retornar a j. Então,*

$$\pi_j = \frac{1}{m_{jj}}$$

Exemplo 9.17 (Modelo de chuva)**.** *Determinar a distribuição estacionária via tempo médio de retorno.*

Seja $T_j = \min\{n \geq 1 | X_n = j | X_0 = 0\}$ *o primeiro instante em que a cadeia retorna ao estado 0. Temos que*

$$\begin{aligned} P(T_0 = 1) &= \alpha \\ P(T_0 = 2) &= (1-\alpha)(1-\beta) \\ P(T_0 = 3) &= (1-\alpha)\beta(1-\beta) \\ \vdots &= \vdots \\ P(T_0 = k) &= (1-\alpha)\beta^{k-2}(1-\beta),\ k \geq 2 \end{aligned}$$

Logo,

$$\begin{aligned}
m_{00} &= E(T_0) = \sum_{k=1}^{\infty} kP(T_0 = k) \\
&= P(T_0 = 1) + \sum_{k=2}^{\infty} kP(T_0 = k) \\
&= \alpha + \sum_{k=2}^{\infty} k(1-\alpha)\beta^{k-2}(1-\beta) \quad \textit{e fazendo } k-1 = j, \\
&= \alpha + \sum_{j=1}^{\infty} (j+1)(1-\alpha)\beta^{j-1}(1-\beta) \\
&= \alpha + (1-\alpha)\sum_{j=1}^{\infty} (j+1)\beta^{j-1}(1-\beta) \\
&= \alpha + (1-\alpha)\left[\sum_{j=1}^{\infty} j\beta^{j-1}(1-\beta) + \sum_{j=1}^{\infty} \beta^{j-1}(1-\beta)\right] \\
&= \alpha + (1-\alpha)\left[\frac{1}{1-\beta} + 1\right] \\
&= \frac{2-\beta-\alpha}{1-\beta}
\end{aligned}$$

E, consequentemente, $\pi_0 = \frac{1-\beta}{2-\beta-\alpha}$ *e* $\pi_1 = \frac{1-\alpha}{2-\beta-\alpha}$.

Exercícios

Capítulo **10**

A distribuição exponencial e o processo de Poisson

Neste capítulo trataremos de um processo estocástico a parâmetro contínuo, ao contrário da cadeia de markov que foi tratado a parâmetro discreto. Muitos fenômenos são modelados dessa forma, tais como: (a) a chegada de uma mensagem em centro de computação operando on-line, (b) uma chamada de telefone em um centro de reservas de passagens, (c) uma interrupção do final de uma fila, (d) a ocorrência de uma falha de um hardware ou software em um computador.

10.1 Processo de Contagem

Um processo estocástico $\{N(t), t \geq 0\}$ é dito ser um ***Processo de Contagem*** *se $N(t)$ representa o número total de eventos que ocorreram até o tempo t. Portanto, um processo de contagem deve satisfazer:*

i) $N(t) \geq 0$

ii) $N(t)$ *é assume valores inteiros*

iii) Se $s < t$, então $N(s) \leq N(t)$

iv) Para $s < t$, $N(t) - N(s)$ é o número de eventos ocorridos no intervalo $(s, t]$.

Um processo de contagem é dito possuir ***incrementos independentes*** *se o número de eventos que ocorrem um intervalos de tempo disjuntos são independentes. Isso significa, por exemplo, que o número de eventos que ocorre no instante t ($N(t)$) deve ser independente do número de eventos que ocorre entre os instantes t e $t + s$ ($N(t + s) - N(t)$).*

Um processo de contagem édito ter ***incrementos estacionários*** *se a distribuição do número de eventos que ocorrem em qualquer intervalo de tempo dependa somente do comprimento do intervalo de tempo. Ou seja, o processo tem incrementos estacionários se número de eventos no intervalo $(s, s + t)$ tem a mesma ditribuição para todo s.*

Exemplo 10.1. *Considerando $N(t)$ o numero de pessoas que entram em uma loja até o instante t, então $\{N(t), t > 0\}$ é um processo de contagem no qual um evento corresponde a uma pessoa entrando na loja. Note que se $N(t)$ é o numero de pessoas dentro da loja no instante t, entao $\{N(t), t > 0\}$ nao será um processo de contagem, pois não há garantias da condição (iii).*

10.2 Processo de Poisson

O exemplo mais importante de processo de contagem é o processo de Poisson, e a distribuição mais importante associada a esse fenômeno é a distribuição de Poisson.

Definição 10.1. *Um processo de contagem $\{N(t), t \geq 0\}$ é dito ser um processo de Poisson com taxa $\lambda > 0$ se for verdade que:*

i) $N(0) = 0$

ii) O processo tem incrementos independentes

iii) O numero de eventos em qualquer intervalo de comprimento t é distribuido como Poisson com media λt. Isto e, para todo s e $t > 0$,

$$P[N(t+s) - N(t) = n] = e^{-\lambda t}\frac{(\lambda t)^n}{n!}, \quad n = 0, 1, 2, \cdots$$

Notemos de (iii) que o processo de poisson $N(t)$ tem incrementos estacionários e que

$$E[N(t)] = \lambda t,$$

de forma que a condição (ii) é bastante razoável, assim como a condição (i). No entanto, a condição (iii) precisa ser mais explorada, por isso uma definição equivalente costuma ser bastante útil. Antes desta, definiremos uma função especial chamada de $o(h)$ que indica a taxa de convergência da função.

Definição 10.2. *A função f é dita ser $o(h)$ se*

$$\lim_{h \to 0} \frac{f(h)}{h} = 0.$$

Definição 10.3. *Um processo de contagem $\{N(t), t \geq 0\}$ é dito ser um processo de Poisson com taxa $\lambda > 0$ se for verdade que:*

i) $N(0) = 0$.

ii) O processo tem incrementos independentes e estacionários.

iii) A probabilidade de que exatamente um evento ocorra em qualquer intervalo de tempo de comprimento h é $\lambda h + o(h)$, isto é,

$$P(N(h) = 1) = \lambda h + o(h)$$

iv) A probabilidade de que mais que um evento ocorra em qualquer intervalo de tempo de comprimento h é $o(h)$, isto é,

$$P(N(h) \geq 2) = o(h)$$

Em suma, em um intervalo de tempo relativamente pequeno, a probabilidade de uma ocorrência do evento é proporcional ao tamanho do intervalo, enquento a probabilidade de duas ou maus ocorrências decresce rapidamente. A probabilidade de nenhuma ocorrência será dada por $P(N(h) = 0) = 1 - \lambda h + o(h)$.

Teorema 43. *As definições 10.1 e 10.4 são equivalentes.*

Portanto, o número médio de eventos ocorrendo em qualquer intervalo de tempo de comprimento t é λt. Esse teorema é surpreendente porque mostra como três condições físicas simples e naturais em $N(t)$ podem caracterizar as funções de probabilidade das variáveis aleatórias. Além disso, é o único parâmetro não conhecido dessas funções de probabilidade. É importante notar que de acordo com o Teorema 43, o número de eventos ocorrendo em cada intervalo de tempo de comprimento t tem uma distribuição de Poisson com valor médio λ. Portanto, é o número médio de eventos ocorrendo por unidade de tempo. A próxima seção caracteriza um outro atributo importante de um processo de Poisson.

10.3 Tempo entre chegadas e tempo de espera

Teorema 44. *Seja $\{N(t), t > 0\}$ um processo de Poisson com taxa λ. Seja $0 < t_1 < t_2 < t_3 < ...$ os tempos sucessivos de ocorrência dos eventos, e seja os tempos entre chegadas (entre ocorrências) definidos por $\tau_1 = t_1, \tau_2 = t_2 - t_1, \cdots, \tau_k = t_k - t_{k-1}, \cdots$ Então o conjunto de tempos entre chegadas são variáveis aleatórias mutualmente independentes distribuídas identicamente e exponencialmente.*

O inverso desse teorema também é verdade (admitimos ambos, o teorema e o seu reverso sem provas). Quando ocorre algum evento, tal como a chegada de uma mensagem em um sistema computacional, a chegada de clientes em um banco, etc, eles podem ser descritos por um processo de Poisson. Entretanto o processo de Poisson é um caso especial de um tipo de processo estocástico mais geral, conhecido como processo de nascimento e morte. *Com efeito, o processo de Poisson é também conhecido como processo de nascimento puro e a partir desse raciocínio é que estaremos aptos a generalizar esse processo e chegar a importantes conclusões sobre uma gama de aplicações desses processo estocásticos de tempo contínuo.*

Para darmos uma idéia da prova, condideremos X_1 o tempo até a primeira ocorrência e X_n o tempo entre a n-ésima e $(n-1)$-ésima ocorrências do processo. A distribuição de X_1 é facilmente obtida observando-se que

$$P(X_1 > t) = P(N(t) = 0) = e^{-\lambda t}\frac{(\lambda t)^0}{0!} = e^{-\lambda t}.$$

Daí, $F_1(t) = P(X_1 \leq t) = 1 - e^{-\lambda t}$, de forma que X_1 tem distribuição exponencial com parâmetro λ. Para os demais intervalos o cálculo é similar, usando que o processo tem incrementos independentes e estacionários:

$$\begin{aligned} P(X_2 > t|X_1 = s) &= P(\textit{nenhum evento em } (s, s+t]|X_1 = s) \\ &= P(\textit{nenhum evento em } (s, s+t]) \\ &= P(\textit{nenhum evento em } (0, t]) \\ &= P(N(t) = 0) = e^{-\lambda t}, \end{aligned}$$

de onde concluímos que X_2 também é exponencialmente distribuída com média $1/\lambda$ e, sobretudo, que X_2 é independente de X_1. Repetindo o argumeto temos que todos os intervalos entre chegadas têm distribuição exponencial de mesmo parâmetro e são independentes.

Uma outra característica de interesse é o tempo até a ocorrência da n-ésimo evento:

$$S_n = \sum_{i=1}^{n} X_i, \quad n \geq 1.$$

Portanto, $S_n \sim Gama(n, \lambda)$ com densidade

$$f(t) = \lambda e^{-\lambda t}\frac{(\lambda t)^{n-1}}{n-1!}.$$

Este tempo de espera *também tem a seguinte propriedade: o tempo até a ocorrência do n-ésimo evento é inferior a t se e somente se no tempo t temos pelo menos n ocorrências, ou seja,*

$$\{S_n \leq t\} = \{N(t) \geq n\}$$

Daí,

$$F_{S_n}(t) = P\{S_n \leq t\} = P\{N(t) \geq n\} = \sum_{j=n}^{\infty} e^{-\lambda t}\frac{(\lambda t)^j}{j!}$$

que tomando a derivada de ambos os lados nos dá o resultado.

10.4 Distribuição condicional dos tempos de chegadas

Conhecendo que até o tempo t houve uma ocorrência do processo, podemos nos perguntar se é mais provável que ocorrência tenha de dado no início ou no final do intervalo. Para verificarmos isso, notemos que

$$\begin{aligned}
P(X_1 < t|N(t) = 1) &= \frac{P\{X_1 < t, N(t) = 1\}}{P(N(t) = 1)} \\
&= \frac{P(\text{1 evento em } (0,s], \text{0 eventos em (s,t]})}{P(N(t) = 1)} \\
&= \frac{P(\text{1 evento em } (0,s])P(\text{0 eventos em (s,t]})}{P(N(t) = 1)} \\
&= \frac{\lambda s e^{-\lambda s} e^{-\lambda(t-s)}}{\lambda t e^{-\lambda t}} \\
&= \frac{s}{t}.
\end{aligned}$$

Assim, concluímos que a distribuição de $X_1|N(t) = 1$ é uniformemente distribuída no intervalo $(0,t]$. E este resultado pode ser ser generalizado para qualquer intervalo.

Teorema 45. *Dado que $N(t) = n$, os tempos de chegada $S_1, \cdots, S_n$ tem a mesma distribuição que as estatísticas de ordem correspondentes a n v.a.i.i.d.Uniformes em $(0,t)$, com densidade,*

$$f(t_1, \cdots, t_n) = \frac{n!}{t^n}, \quad 0 < t_1 < \cdots < t_n.$$

Exemplo 10.2. *Suponha que clientes cheguem a uma estação de trem de acordo com um processo de poisson de taxa λ. Se o trem parte ao tempo t, qual o valo esperado da soma dos tempos de espera dos clientes que chegaram no intervalo $(0,t)$.*

Devemos calcular $E[\sum_{i=1}^{N(t)}(t - S_i)]$. Condicionando em $N(t)$ obtemos

$$\begin{aligned}
E\left[\sum_{i=1}^{N(t)}(t - S_i)|N(t) = n\right] &= E\left[\sum_{i=1}^{n}(t - S_i)|N(t) = n\right] \\
&= nt - E\left[\sum_{i=1}^{n} S_i|N(t) = n\right].
\end{aligned}$$

Agora, se $U_i, i = 1, \cdots, n$ *denotam um conjunto de* n *v.a.* $U(0,t)$*, então*

$$\begin{aligned} E\left[\sum_{i=1}^{N(t)} S_i | N(t) = n\right] &= E\left[\sum_{i=1}^{N(t)} U_{(i)}\right] \quad \text{(pelo Teorema 45)} \\ &= E\left[\sum_{i=1}^{N(t)} U_i\right] \quad \left(\text{pois } \sum_{i=1}^{N(t)} U_{(i)} = \sum_{i=1}^{N(t)} U_i\right) \\ &= \frac{nt}{2}. \end{aligned}$$

Daí,

$$E\left[\sum_{i=1}^{N(t)} (t - S_i) | N(t) = n\right] = nt - \frac{nt}{2} = \frac{nt}{2},$$

de onde obtemos que

$$E\left[\sum_{i=1}^{N(t)} (t - S_i) | N(t)\right] = \frac{tN(t)}{2},$$

e, consequentemente,

$$E\left[\sum_{i=1}^{N(t)} (t - S_i)\right] = E\left(\left[\sum_{i=1}^{N(t)} (t - S_i) | N(t)\right]\right) = \frac{t}{2} E[N(t)] = \frac{t}{2}\lambda t = \frac{\lambda t^2}{2}.$$

10.5 Processe de Poisson não homogêneo

Em algumas situações não é possível que as taxas de ocorrência do processo se mantém constante ao longo do dia, mes ou ano, por exemplo, de forma que não temos a propriedade de Estacionariedade. Exemplos desse caso são chamadas telefônicas ou entradas de clientes em um supermeercado. No geral, para pequenos intervalos de tempo é possível assumir a homogeneidade. Nesta seção trabalharemos com a taxa $\lambda(t)$ *sendo função do tempo, o que generaliza o caso homogêneo, mas muito da construção do caso homog eneo se mantém. Enunciaremos basicamente a definição do processo.*

Definição 10.4. *Um processo de contagem* $\{N(t), t \geq 0\}$ *é dito ser um processo de Poisson não estacionário ou não homogêneo com função de intensidade* $\lambda(t) > 0$ *se:*

i) $N(0) = 0$.

ii) $\{N(t), t \geq 0\}$ *tem incrementos independentes.*

iii) $P\{N(t + h) - N(t) = 1\} = \lambda(t)h + o(h)$

iv) $P(N(t + h) - N(t) \geq 2) = o(h)$.

Fazendo $m(t) = \int_0^t \lambda(s)ds$*, pode ser mostrado que*

$$P\{N(t + s) - N(t) = n\} = e^{-[m(t+s)-m(t)]} \frac{[m(t + s) - m(t)]^n}{n!}, \quad n \geq 0.$$

Ou seja, $N(t+s)-N(t)$ *tem distribuição Poisson com média* $m(t+s)-m(t)$*. Por isso a função* $m(t)$ *é chamada* função valor médio.

Exemplo 10.3. *Norbert tem um carro de cachorro quente que abre às 8h. Das 8h às 11h clientes chegam, em média, a uma taxa crescente que começa com 8 e termina com 20 pessoas por hora. Das 11h às 13h a taxa fica constante em 20 clientes por hora. Depois a taxa cai linearmente até as 17h, quando encerra, com 12 cliente por hora. Assumindo que os clientes chegam de forma independente em intervalos disjuntos, o modelo acima parece razoável? Qual o número esperado de chegadas neste período? Qual a probabilidade de que nenhum cliente chegue entre 8h30 e 9h30?*

O modelo de Poisson não-homogêneo parece bom para a situação apresentada, com taxa:

$$\lambda(t) = \begin{cases} 5+5t, & 0 \le t \le 3 \\ 20, & 3 \le t \le 5 \\ 20 - 2(t-5), & 5 \le t \le 9. \end{cases}$$

Com isso, o número de chegadas entre 8h30 e 9h30 será Poisson com média $m(\frac{3}{2}) - m(\frac{1}{2})$*. O número esperado de chegadas nesse intervalo será*

$$\int_{1/2}^{3/2} (5+5t)dt = 10,$$

e este número será zero com probabilidade

$$\exp\left\{ -\int_{1/2}^{3/2} (5+5t)dt \right\} = e^{-10}.$$

10.6 Processo de Poisson composto

O processo de Poisson composto é usado quando cada ocorrência do Processo de Poisson está relacionada a outra variável. Por exemplo, em um supermeercado, cada cliente que entra no supermercado gasta um valor Y_i*. Portanto, o processo de poisson composto pode ser assim enunciado:*

Definição 10.5. *Um processo de contagem* $\{X(t), t \ge 0\}$ *é dito ser um processo de Poisson composto se ele puder se representado, para cada* $t \ge 0$*, por*

$$X(t) = \sum_{i=1}^{N(t)} Y_i,$$

onde $\{N(t), t \ge 0\}$ *é um processo de Poisson e* $\{Y_i, i = 1, 2, \cdots\}$ *é uma família de v.a. independentes e identicamente distribuídas. O processo* $\{N(t), t \ge 0\}$ *e a sequência* $\{Y_i, i = 1, 2, \cdots\}$ *são assumidos serem independentes.*

Vamos calcular a esperança e variância deste processo. Para calcular $E[X(t)]$*, temos que condicionar em* $N(t)$*:*

$$E[X(t)] = E(E[X(t)|N(t)]).$$

Agora,

$$\begin{aligned} E[X(t)|N(t)=n] &= E\left[\sum_{i=1}^{N(t)} Y_i|N(t)=n\right] = E\left[\sum_{i=1}^{n} Y_i|N(t)=n\right] \\ &= E\left[\sum_{i=1}^{n} Y_i\right] = \sum_{i=1}^{n} E[Y_i] = nE(Y_1). \end{aligned}$$

Portanto,

$$E[X(t)|N(t)] = N(t)E(Y_1).$$

E, consequentemente,

$$E[X(t)] = E(E[X(t)|N(t)]) = E[N(t)E(Y_1)] = E[N(t)][E(Y_1)] = \lambda t E(Y_1)]$$

Para o cálculo da variância procedemos de forma similar, obtendo:

$$Var[X(t)|N(t)] = N(t)Var(Y_1)$$

e, portnto,

$$\begin{aligned} Var[X(t)] &= E[N(t)Var(Y_1)] + Var[N(t)E(Y_1)] \\ &= \lambda t Var(Y_1) + (E(Y_1))^2 Var[N(t)] \\ &= \lambda t Var(Y_1) + (E(Y_1))^2 \lambda t \\ &= \lambda t[Var(Y_1) + (E(Y_1))^2] \\ &= \lambda t E(Y_1^2). \end{aligned}$$

Exemplo 10.4. *Suponha que famílias migrem para uma área de acordo com um processo de Poisson de taxa $\lambda = 2$ por semana. Se o número de pessoas em cada família é independente e toma valores valores 1,2,3 ou 4 com probabilidades $\frac{1}{6}, \frac{1}{3}, \frac{1}{3}, \frac{1}{6}$, respectivamente, então qual o número de indivíduos migrando para aquela área durante 5 semanas?*

Seja Y_i o número de pessoas na i-ésima família. Temos que

$$E(Y_1) = \frac{5}{2} \quad e \quad E(Y_1^2) = \frac{43}{6}.$$

Seja $X(5)$ o núnero de imigrantes durante as 5 semanas. Então,

$$\begin{aligned} E[X(5)] &= 2 \times 5 \times \frac{5}{2} = 25 \\ Var[X(t)] &= 2 \times 5 \times \frac{43}{6} = \frac{215}{3}. \end{aligned}$$