

Guia Teórico Completo de Probabilidade e Estatística

Vicente Duarte

Ano Letivo 2024/2025

Resumo

Este guia teórico apresenta uma visão abrangente dos conceitos de Probabilidade e Estatística abordados nas séries de problemas do curso. O documento está estruturado de forma progressiva, abordando desde conceitos básicos de probabilidade até distribuições de variáveis aleatórias, probabilidade condicional e técnicas de inferência estatística. É especialmente útil como material de apoio ao estudo e à resolução dos exercícios propostos.

Conteúdo

1	Conceitos Fundamentais de Probabilidade	4
1.1	Espaço Amostral e Eventos	4
1.2	Definição de Probabilidade	4
1.3	Operações com Eventos	4
1.4	Probabilidade Condicional	5
1.5	Independência de Eventos	5
1.6	Teorema da Probabilidade Total	6
1.7	Teorema de Bayes	6

2	Variáveis Aleatórias	7
2.1	Definição e Tipos	7
2.2	Função de Probabilidade (V.A. Discretas)	7
2.3	Função Densidade de Probabilidade (V.A. Contínuas)	8
2.4	Função de Distribuição Acumulada	8
2.5	Valor Esperado (Média)	9
2.6	Variância e Desvio Padrão	9
3	Distribuições Discretas Importantes	10
3.1	Distribuição de Bernoulli	10
3.2	Distribuição Binomial	10
3.3	Distribuição Geométrica	11
3.4	Distribuição de Poisson	12
4	Distribuições Contínuas Importantes	13
4.1	Distribuição Uniforme Contínua	13
4.2	Distribuição Exponencial	13
4.3	Distribuição Normal	15
5	Variáveis Aleatórias Multidimensionais	16
5.1	Distribuição Conjunta	16
5.2	Distribuições Marginais	17
5.3	Distribuições Condicionais	17
5.4	Independência de Variáveis Aleatórias	17
5.5	Covariância e Correlação	18
5.6	Esperança Condicional	18
6	Teoremas Limite e Aproximações	19
6.1	Lei dos Grandes Números	19
6.2	Teorema Central do Limite	20
6.3	Aproximação da Distribuição Binomial pela Normal	21

7	Estatística Inferencial Básica	22
7.1	Estimação Pontual	22
7.2	Intervalos de Confiança	23
7.3	Testes de Hipóteses	24
8	Exemplo de Aplicação Completa	25
8.1	Contexto do Problema	25
8.2	Formulação do Teste de Hipóteses	25
8.3	Cálculo da Estatística de Teste	25
8.4	Decisão	26
8.5	Cálculo do Valor-p	26
8.6	Interpretação	26
8.7	Intervalo de Confiança	26
8.8	Interpretação do Intervalo	27
8.9	Conclusão	27

1 Conceitos Fundamentais de Probabilidade

1.1 Espaço Amostral e Eventos

O espaço amostral, geralmente representado por Ω , é o conjunto de todos os resultados possíveis de uma experiência aleatória. Um evento A é um subconjunto do espaço amostral, ou seja, $A \subseteq \Omega$.

Exemplo

Ao lançar um dado de seis faces, o espaço amostral é $\Omega = \{1, 2, 3, 4, 5, 6\}$. O evento "sair um número par" corresponde ao conjunto $A = \{2, 4, 6\}$.

1.2 Definição de Probabilidade

A probabilidade de um evento A , denotada por $P(A)$, é uma medida da possibilidade de A ocorrer. Pela definição axiomática de Kolmogorov, a probabilidade deve satisfazer:

- (i) $P(A) \geq 0$ para qualquer evento A
- (ii) $P(\Omega) = 1$
- (iii) Se A_1, A_2, \dots são eventos mutuamente exclusivos, então $P(A_1 \cup A_2 \cup \dots) = P(A_1) + P(A_2) + \dots$

Exemplo

No lançamento de uma moeda equilibrada, $P(\text{cara}) = P(\text{coroa}) = \frac{1}{2}$.
Para o lançamento de um dado equilibrado, $P(\text{sair } 3) = \frac{1}{6}$.

1.3 Operações com Eventos

As operações básicas com eventos incluem:

- **União** ($A \cup B$): evento que ocorre quando A ou B (ou ambos) ocorrem

- **Interseção** ($A \cap B$): evento que ocorre quando A e B ocorrem simultaneamente
- **Complemento** (\bar{A} ou A^c): evento que ocorre quando A não ocorre

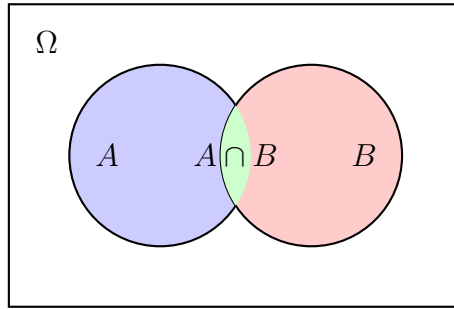


Figura 1: Diagrama de Venn representando eventos A e B , sua interseção e o espaço amostral Ω

1.4 Probabilidade Condicional

A probabilidade condicional de um evento A dado que o evento B ocorreu, denotada por $P(A|B)$, é definida como:

$$P(A|B) = \frac{P(A \cap B)}{P(B)}, \quad \text{onde } P(B) > 0 \quad (1)$$

Fórmula Fundamental

A fórmula da probabilidade condicional pode ser reorganizada para obter:

$$P(A \cap B) = P(B) \times P(A|B) = P(A) \times P(B|A) \quad (2)$$

Esta é uma das fórmulas mais importantes e frequentemente utilizadas em problemas de probabilidade.

1.5 Independência de Eventos

Dois eventos A e B são independentes se e somente se:

$$P(A \cap B) = P(A) \times P(B) \quad (3)$$

Equivalentemente, A e B são independentes se $P(A|B) = P(A)$ ou $P(B|A) = P(B)$.

Exemplo

Ao lançar um dado duas vezes, o evento "tirar um número par no primeiro lançamento" é independente do evento "tirar um número menor que 3 no segundo lançamento".

1.6 Teorema da Probabilidade Total

Se B_1, B_2, \dots, B_n constituem uma partição do espaço amostral Ω (ou seja, são mutuamente exclusivos e $B_1 \cup B_2 \cup \dots \cup B_n = \Omega$), então para qualquer evento A :

$$P(A) = \sum_{i=1}^n P(A|B_i) \times P(B_i) \quad (4)$$

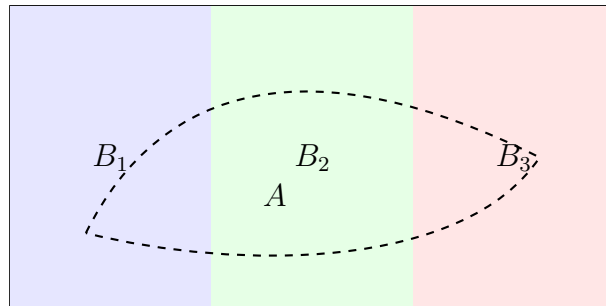


Figura 2: Ilustração do Teorema da Probabilidade Total: evento A distribuído entre a partição $\{B_1, B_2, B_3\}$

1.7 Teorema de Bayes

O Teorema de Bayes permite inverter condições em probabilidades condicionais:

$$P(B_i|A) = \frac{P(A|B_i) \times P(B_i)}{\sum_{j=1}^n P(A|B_j) \times P(B_j)} \quad (5)$$

Numa forma mais simplificada, para dois eventos A e B :

$$P(B|A) = \frac{P(A|B) \times P(B)}{P(A)} \quad (6)$$

Dica para Aplicar o Teorema de Bayes

O Teorema de Bayes é particularmente útil em problemas onde:

- São fornecidas probabilidades do tipo $P(A|B)$
- Pede-se para calcular probabilidades do tipo $P(B|A)$
- Existe informação adicional que surge após uma observação inicial

2 Variáveis Aleatórias

2.1 Definição e Tipos

Uma variável aleatória é uma função que associa um valor numérico a cada resultado do espaço amostral. As variáveis aleatórias podem ser:

- **Discretas:** assumem valores em conjunto finito ou enumerável (e.g., número de caras em 10 lançamentos de moeda)
- **Contínuas:** podem assumir qualquer valor em um intervalo contínuo (e.g., tempo de espera, altura de uma pessoa)

2.2 Função de Probabilidade (V.A. Discretas)

Para uma variável aleatória discreta X , a função de probabilidade $p_X(x)$ (também chamada de função massa de probabilidade) é definida como:

$$p_X(x) = P(X = x) \quad (7)$$

Propriedades:

- (i) $p_X(x) \geq 0$ para todo x
- (ii) $\sum_x p_X(x) = 1$

2.3 Função Densidade de Probabilidade (V.A. Contínuas)

Para uma variável aleatória contínua X , a função densidade de probabilidade $f_X(x)$ satisfaz:

$$P(a \leq X \leq b) = \int_a^b f_X(x) dx \quad (8)$$

Propriedades:

- (i) $f_X(x) \geq 0$ para todo x
- (ii) $\int_{-\infty}^{\infty} f_X(x) dx = 1$

2.4 Função de Distribuição Acumulada

A função de distribuição acumulada (FDA) $F_X(x)$ de uma variável aleatória X é definida como:

$$F_X(x) = P(X \leq x) \quad (9)$$

Para variáveis aleatórias discretas:

$$F_X(x) = \sum_{t \leq x} p_X(t) \quad (10)$$

Para variáveis aleatórias contínuas:

$$F_X(x) = \int_{-\infty}^x f_X(t) dt \quad (11)$$

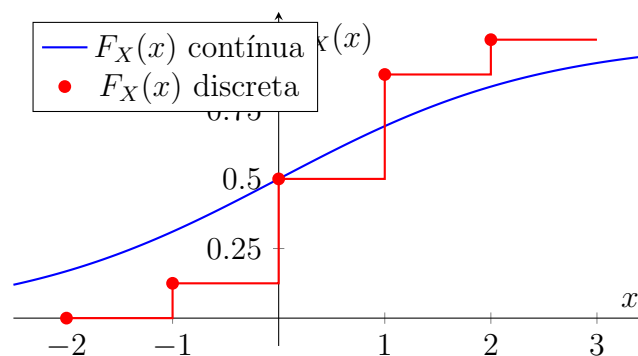


Figura 3: Comparação entre funções de distribuição acumulada para variáveis aleatórias contínuas e discretas

2.5 Valor Esperado (Média)

O valor esperado de uma variável aleatória X , denotado por $E[X]$ ou μ_X , representa o valor médio que esperamos obter se repetirmos a experiência muitas vezes.

Para variáveis aleatórias discretas:

$$E[X] = \sum_x x \cdot p_X(x) \quad (12)$$

Para variáveis aleatórias contínuas:

$$E[X] = \int_{-\infty}^{\infty} x \cdot f_X(x) dx \quad (13)$$

2.6 Variância e Desvio Padrão

A variância de uma variável aleatória X , denotada por $\text{Var}(X)$ ou σ_X^2 , mede a dispersão dos valores em torno da média:

$$\text{Var}(X) = E[(X - \mu_X)^2] = E[X^2] - (E[X])^2 \quad (14)$$

O desvio padrão σ_X é definido como a raiz quadrada da variância:

$$\sigma_X = \sqrt{\text{Var}(X)} \quad (15)$$

Propriedades do Valor Esperado e Variância

Para constantes a e b e variáveis aleatórias X e Y :

$$E[aX + b] = aE[X] + b \quad (16)$$

$$\text{Var}(aX + b) = a^2 \text{Var}(X) \quad (17)$$

$$(18)$$

Se X e Y são independentes:

$$E[X \cdot Y] = E[X] \cdot E[Y] \quad (19)$$

$$\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y) \quad (20)$$

3 Distribuições Discretas Importantes

3.1 Distribuição de Bernoulli

A distribuição de Bernoulli modela experiências com apenas dois resultados possíveis (sucesso ou fracasso). Se X é uma variável aleatória de Bernoulli com parâmetro p (probabilidade de sucesso), então:

$$P(X = x) = \begin{cases} p, & \text{se } x = 1 \\ 1 - p, & \text{se } x = 0 \end{cases} \quad (21)$$

Média: $E[X] = p$

Variância: $\text{Var}(X) = p(1 - p)$

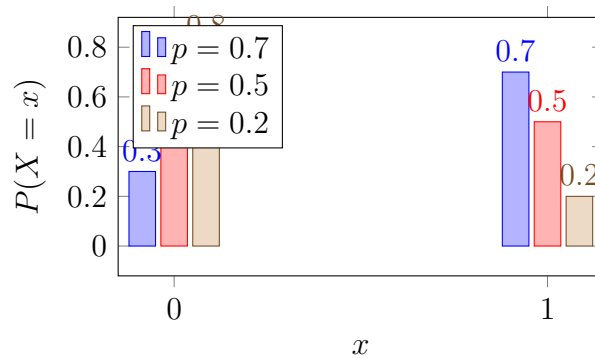


Figura 4: Função de massa de probabilidade da distribuição de Bernoulli para diferentes valores de p

3.2 Distribuição Binomial

A distribuição binomial modela o número de sucessos em n tentativas independentes de Bernoulli, cada uma com probabilidade de sucesso p . Se X é uma variável aleatória binomial com parâmetros n e p , denotamos $X \sim \text{Bin}(n, p)$ e:

$$P(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}, \quad k = 0, 1, 2, \dots, n \quad (22)$$

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

Média: $E[X] = np$

Variância: $\text{Var}(X) = np(1 - p)$

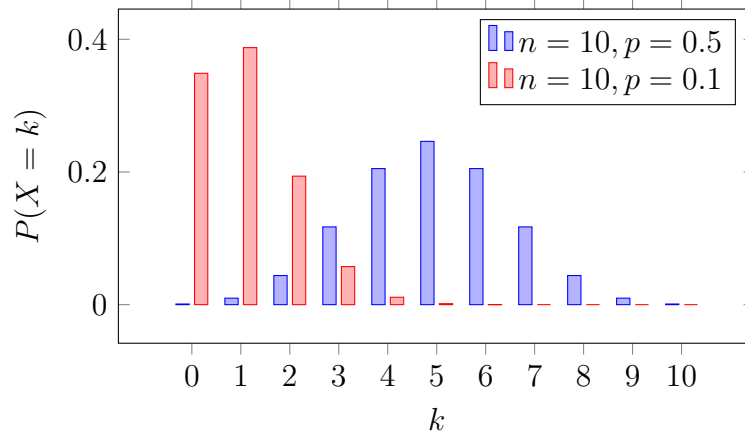


Figura 5: Função de massa de probabilidade da distribuição binomial para diferentes valores de p com $n = 10$

3.3 Distribuição Geométrica

A distribuição geométrica modela o número de tentativas necessárias até o primeiro sucesso em tentativas independentes de Bernoulli. Se X é uma variável aleatória geométrica com parâmetro p (probabilidade de sucesso em cada tentativa), então:

$$P(X = k) = (1 - p)^{k-1}p, \quad k = 1, 2, 3, \dots \quad (23)$$

Média: $E[X] = \frac{1}{p}$

Variância: $\text{Var}(X) = \frac{1-p}{p^2}$

Propriedade de "Falta de Memória"

A distribuição geométrica é a única distribuição discreta que possui a propriedade de "falta de memória":

$$P(X > n + m | X > n) = P(X > m) \quad (24)$$

Esta propriedade significa que, se já esperamos n tentativas sem sucesso, a probabilidade de precisarmos de mais m tentativas é igual à probabilidade original de precisarmos de m tentativas.

3.4 Distribuição de Poisson

A distribuição de Poisson modela o número de ocorrências de um evento em um intervalo fixo, quando esses eventos ocorrem a uma taxa média constante e independentemente uns dos outros. Se X é uma variável aleatória de Poisson com parâmetro λ (taxa média de ocorrências), denotamos $X \sim \text{Poisson}(\lambda)$ e:

$$P(X = k) = \frac{e^{-\lambda} \lambda^k}{k!}, \quad k = 0, 1, 2, \dots \quad (25)$$

Média: $E[X] = \lambda$

Variância: $\text{Var}(X) = \lambda$

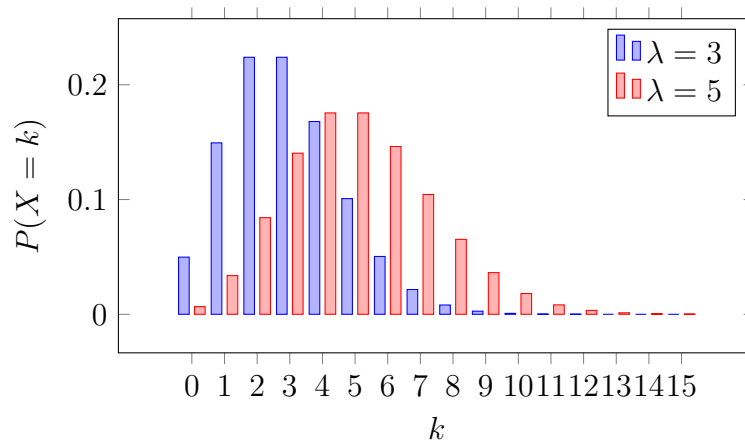


Figura 6: Função de massa de probabilidade da distribuição de Poisson para diferentes valores de λ

4 Distribuições Contínuas Importantes

4.1 Distribuição Uniforme Contínua

A distribuição uniforme contínua modela situações onde todos os valores em um intervalo $[a, b]$ têm a mesma probabilidade de ocorrer. Se X é uma variável aleatória uniforme no intervalo $[a, b]$, denotamos $X \sim \text{Uniforme}(a, b)$ e sua função densidade de probabilidade é:

$$f_X(x) = \begin{cases} \frac{1}{b-a}, & \text{se } a \leq x \leq b \\ 0, & \text{caso contrário} \end{cases} \quad (26)$$

Função de distribuição acumulada:

$$F_X(x) = \begin{cases} 0, & \text{se } x < a \\ \frac{x-a}{b-a}, & \text{se } a \leq x \leq b \\ 1, & \text{se } x > b \end{cases} \quad (27)$$

Média: $E[X] = \frac{a+b}{2}$

Variância: $\text{Var}(X) = \frac{(b-a)^2}{12}$



Figura 7: Funções de densidade e de distribuição da distribuição uniforme contínua no intervalo $[0, 5]$

4.2 Distribuição Exponencial

A distribuição exponencial é frequentemente usada para modelar o tempo até a ocorrência de um evento, quando a taxa de ocorrência é constante. Se X é uma variável aleatória exponencial com parâmetro $\lambda > 0$ (taxa), denotamos $X \sim \text{Exp}(\lambda)$ e sua função densidade de probabilidade é:

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x}, & \text{se } x \geq 0 \\ 0, & \text{se } x < 0 \end{cases} \quad (28)$$

Função de distribuição acumulada:

$$F_X(x) = \begin{cases} 1 - e^{-\lambda x}, & \text{se } x \geq 0 \\ 0, & \text{se } x < 0 \end{cases} \quad (29)$$

Média: $E[X] = \frac{1}{\lambda}$

Variância: $\text{Var}(X) = \frac{1}{\lambda^2}$

Propriedade de "Falta de Memória"

Assim como a distribuição geométrica, a distribuição exponencial possui a propriedade de "falta de memória":

$$P(X > s + t | X > t) = P(X > s) \quad (30)$$

Isso significa que, se um componente já funcionou por t unidades de tempo sem falhar, a probabilidade de ele funcionar por mais s unidades de tempo é igual à probabilidade original de funcionar por s unidades de tempo.

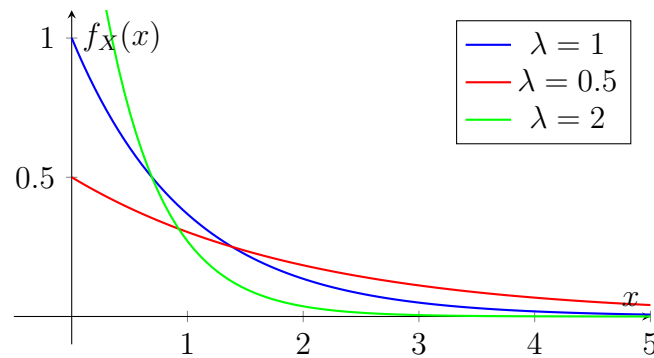


Figura 8: Função densidade de probabilidade da distribuição exponencial para diferentes valores de λ

4.3 Distribuição Normal

A distribuição normal (ou gaussiana) é uma das distribuições mais importantes em estatística. Se X é uma variável aleatória normal com média μ e variância σ^2 , denotamos $X \sim N(\mu, \sigma^2)$ e sua função densidade de probabilidade é:

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad -\infty < x < \infty \quad (31)$$

A distribuição normal padrão, denotada por $Z \sim N(0, 1)$, tem média zero e variância unitária. Qualquer variável aleatória normal pode ser padronizada através da transformação $Z = \frac{X-\mu}{\sigma}$.

Média: $E[X] = \mu$

Variância: $\text{Var}(X) = \sigma^2$

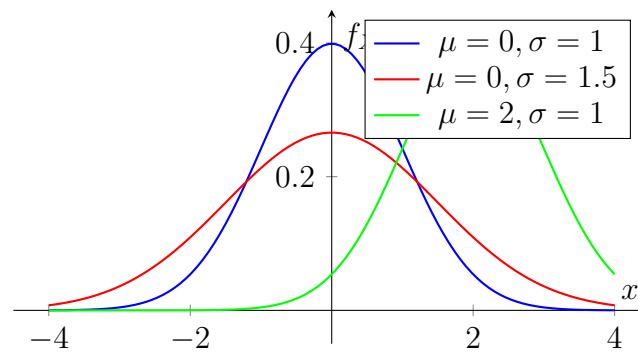


Figura 9: Função densidade de probabilidade da distribuição normal para diferentes valores de μ e σ

Propriedades da Distribuição Normal

- É simétrica em torno da média μ
- A moda e a mediana coincidem com a média
- Tem forma de sino, com ponto de inflexão em $\mu \pm \sigma$
- Aproximadamente 68,27% dos valores estão a uma distância de até σ da média
- Aproximadamente 95,45% dos valores estão a uma distância de até 2σ da média
- Aproximadamente 99,73% dos valores estão a uma distância de até 3σ da média (regra 3-sigma)

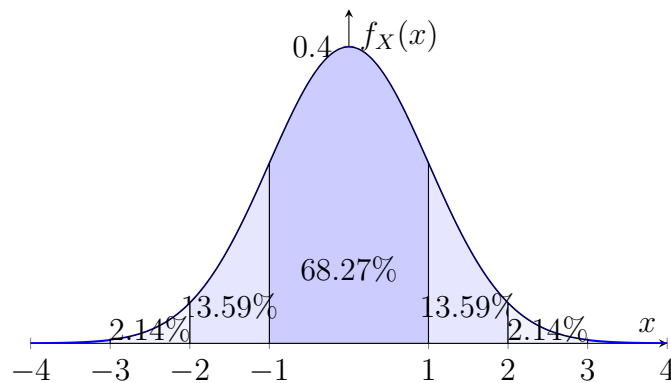


Figura 10: Distribuição normal padrão com áreas correspondentes a diferentes intervalos de desvios padrão

5 Variáveis Aleatórias Multidimensionais

5.1 Distribuição Conjunta

Para variáveis aleatórias discretas X e Y , a função de probabilidade conjunta é:

$$p_{X,Y}(x,y) = P(X = x, Y = y) \quad (32)$$

Para variáveis aleatórias contínuas X e Y , a função densidade de probabilidade conjunta satisfaz:

$$P((X, Y) \in A) = \iint_A f_{X,Y}(x, y) dx dy \quad (33)$$

5.2 Distribuições Marginais

Para variáveis aleatórias discretas:

$$p_X(x) = \sum_y p_{X,Y}(x, y) \quad (34)$$

$$p_Y(y) = \sum_x p_{X,Y}(x, y) \quad (35)$$

Para variáveis aleatórias contínuas:

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x, y) dy \quad (36)$$

$$f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x, y) dx \quad (37)$$

5.3 Distribuições Condicionais

Para variáveis aleatórias discretas:

$$p_{X|Y}(x|y) = \frac{p_{X,Y}(x, y)}{p_Y(y)}, \quad \text{se } p_Y(y) > 0 \quad (38)$$

$$p_{Y|X}(y|x) = \frac{p_{X,Y}(x, y)}{p_X(x)}, \quad \text{se } p_X(x) > 0 \quad (39)$$

Para variáveis aleatórias contínuas:

$$f_{X|Y}(x|y) = \frac{f_{X,Y}(x, y)}{f_Y(y)}, \quad \text{se } f_Y(y) > 0 \quad (40)$$

$$f_{Y|X}(y|x) = \frac{f_{X,Y}(x, y)}{f_X(x)}, \quad \text{se } f_X(x) > 0 \quad (41)$$

5.4 Independência de Variáveis Aleatórias

Duas variáveis aleatórias X e Y são independentes se e somente se:

Para variáveis aleatórias discretas:

$$p_{X,Y}(x, y) = p_X(x) \cdot p_Y(y), \quad \text{para todo } x, y \quad (42)$$

Para variáveis aleatórias contínuas:

$$f_{X,Y}(x, y) = f_X(x) \cdot f_Y(y), \quad \text{para todo } x, y \quad (43)$$

5.5 Covariância e Correlação

A covariância entre duas variáveis aleatórias X e Y é definida como:

$$\text{Cov}(X, Y) = E[(X - \mu_X)(Y - \mu_Y)] = E[XY] - E[X]E[Y] \quad (44)$$

onde $\mu_X = E[X]$ e $\mu_Y = E[Y]$.

A correlação entre X e Y é:

$$\rho_{X,Y} = \frac{\text{Cov}(X, Y)}{\sigma_X \sigma_Y} \quad (45)$$

onde $\sigma_X = \sqrt{\text{Var}(X)}$ e $\sigma_Y = \sqrt{\text{Var}(Y)}$.

Propriedades da Correlação

- $-1 \leq \rho_{X,Y} \leq 1$
- $\rho_{X,Y} = 1$ ou $\rho_{X,Y} = -1$ se e somente se existe uma relação linear perfeita entre X e Y
- $\rho_{X,Y} = 0$ se X e Y são independentes (a recíproca não é necessariamente verdadeira)

5.6 Esperança Condicional

A esperança condicional de uma variável aleatória X dado $Y = y$ é definida como:

Para variáveis aleatórias discretas:

$$E[X|Y = y] = \sum_x x \cdot p_{X|Y}(x|y) \quad (46)$$

Para variáveis aleatórias contínuas:

$$E[X|Y = y] = \int_{-\infty}^{\infty} x \cdot f_{X|Y}(x|y) dx \quad (47)$$

A esperança condicional $E[X|Y]$ pode ser vista como uma função de Y .

Lei da Esperança Total

$$E[X] = E[E[X|Y]] \quad (48)$$

Ou seja, a esperança incondicional de X pode ser obtida tomando a média ponderada das esperanças condicionais.

6 Teoremas Limite e Aproximações

6.1 Lei dos Grandes Números

A Lei dos Grandes Números assegura que a média aritmética de um grande número de variáveis aleatórias independentes e identicamente distribuídas converge para a esperança comum dessas variáveis.

Seja X_1, X_2, \dots, X_n uma sequência de variáveis aleatórias independentes e identicamente distribuídas com $E[X_i] = \mu$. Defina $\bar{X}_n = \frac{1}{n}(X_1 + X_2 + \dots + X_n)$. Então:

$$\bar{X}_n \xrightarrow{P} \mu \quad \text{quando } n \rightarrow \infty \quad (49)$$

onde \xrightarrow{P} denota convergência em probabilidade.

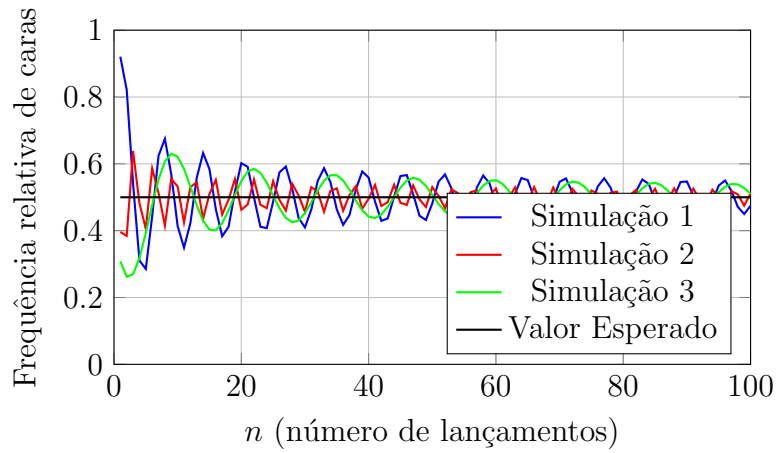


Figura 11: Ilustração da Lei dos Grandes Números: frequência relativa de caras em n lançamentos de uma moeda equilibrada

6.2 Teorema Central do Limite

O Teorema Central do Limite estabelece que a soma de um grande número de variáveis aleatórias independentes e identicamente distribuídas segue aproximadamente uma distribuição normal, independentemente da distribuição original das variáveis.

Seja X_1, X_2, \dots, X_n uma sequência de variáveis aleatórias independentes e identicamente distribuídas com $E[X_i] = \mu$ e $\text{Var}(X_i) = \sigma^2 < \infty$. Defina $S_n = X_1 + X_2 + \dots + X_n$. Então, para n suficientemente grande:

$$\frac{S_n - n\mu}{\sigma\sqrt{n}} \xrightarrow{d} N(0, 1) \quad \text{quando } n \rightarrow \infty \quad (50)$$

onde \xrightarrow{d} denota convergência em distribuição.

Aplicação do Teorema Central do Limite

O TCL permite aproximar a distribuição da soma de várias variáveis aleatórias por uma distribuição normal, facilitando enormemente os cálculos de probabilidades. Por exemplo, para n grande:

$$P\left(a \leq \frac{S_n - n\mu}{\sigma\sqrt{n}} \leq b\right) \approx \Phi(b) - \Phi(a) \quad (51)$$

onde Φ é a função de distribuição acumulada da normal padrão.

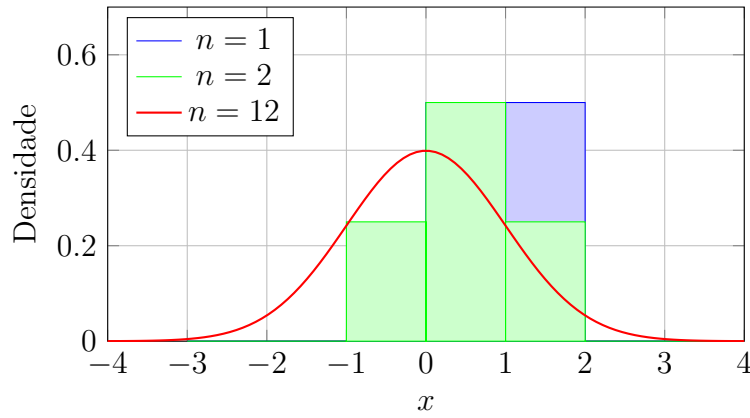


Figura 12: Ilustração do Teorema Central do Limite: a soma de n variáveis aleatórias independentes e identicamente distribuídas aproxima-se de uma distribuição normal conforme n aumenta

6.3 Aproximação da Distribuição Binomial pela Normal

Para valores grandes de n e valores de p não muito próximos de 0 ou 1, a distribuição binomial $\text{Bin}(n, p)$ pode ser aproximada por uma distribuição normal com média np e variância $np(1 - p)$:

$$\text{Bin}(n, p) \approx N(np, np(1 - p)) \quad (52)$$

Regra prática: a aproximação é considerada boa quando $np > 5$ e $n(1 - p) > 5$.

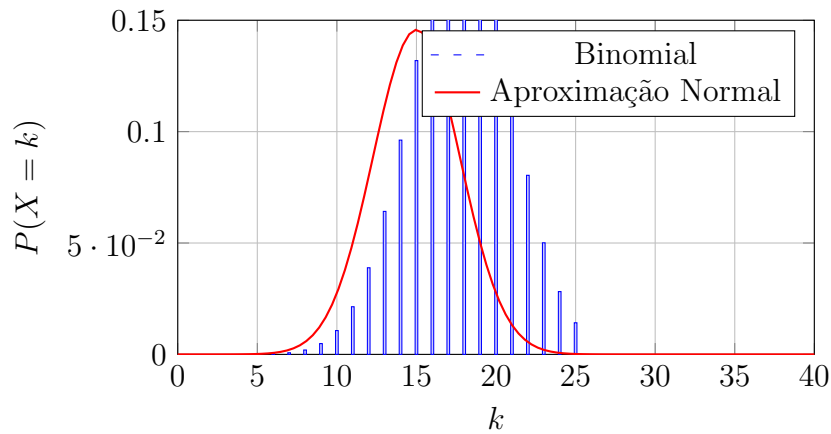


Figura 13: Aproximação da distribuição binomial pela distribuição normal

7 Estatística Inferencial Básica

7.1 Estimação Pontual

A estimação pontual consiste em utilizar uma única estatística (estimador) para estimar um parâmetro desconhecido da população.

- **Estimador:** função dos dados da amostra
- **Estimativa:** valor específico do estimador calculado a partir de dados amostrais

Propriedades desejáveis de um estimador:

- **Não-viesado:** $E[\hat{\theta}] = \theta$ (o valor esperado do estimador é igual ao parâmetro)
- **Consistente:** $\hat{\theta} \xrightarrow{P} \theta$ quando $n \rightarrow \infty$ (o estimador converge em probabilidade para o parâmetro)
- **Eficiente:** tem a menor variância possível entre os estimadores não-viesados

Estimadores comuns:

$$\text{Média amostral: } \bar{X} = \frac{1}{n} \sum_{i=1}^n X_i \quad (53)$$

$$\text{Variância amostral: } S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2 \quad (54)$$

$$\text{Proporção amostral: } \hat{p} = \frac{X}{n} \quad (55)$$

7.2 Intervalos de Confiança

Um intervalo de confiança é um intervalo de valores que, com uma certa probabilidade (nível de confiança), contém o verdadeiro valor do parâmetro populacional.

Para a média populacional μ de uma distribuição normal com variância σ^2 conhecida, um intervalo de confiança de nível $(1 - \alpha)$ é:

$$\bar{X} \pm z_{\alpha/2} \frac{\sigma}{\sqrt{n}} \quad (56)$$

onde $z_{\alpha/2}$ é o quantil de ordem $1 - \alpha/2$ da distribuição normal padrão.

Se a variância σ^2 é desconhecida e a amostra é de uma distribuição normal, utiliza-se a distribuição t de Student:

$$\bar{X} \pm t_{\alpha/2, n-1} \frac{S}{\sqrt{n}} \quad (57)$$

onde $t_{\alpha/2, n-1}$ é o quantil de ordem $1 - \alpha/2$ da distribuição t com $n - 1$ graus de liberdade.

Interpretação Correta de um Intervalo de Confiança

Um intervalo de confiança de 95% não significa que a probabilidade do parâmetro estar no intervalo é 95%. Significa que, se muitas amostras fossem coletadas e intervalos de confiança calculados, 95% desses intervalos conteriam o verdadeiro valor do parâmetro.

7.3 Testes de Hipóteses

Um teste de hipóteses é um procedimento para decidir se uma hipótese sobre a população é plausível, com base em dados amostrais.

- **Hipótese nula** (H_0): hipótese que está sendo testada
- **Hipótese alternativa** (H_1 ou H_a): hipótese que será aceita se H_0 for rejeitada
- **Estatística de teste**: função dos dados que é usada para decidir se H_0 deve ser rejeitada
- **Região crítica**: conjunto de valores da estatística de teste que levam à rejeição de H_0
- **Nível de significância** (α): probabilidade de rejeitar H_0 quando ela é verdadeira (erro Tipo I)
- **Valor-p**: probabilidade de observar um valor da estatística de teste pelo menos tão extremo quanto o observado, supondo que H_0 seja verdadeira

Para testar $H_0 : \mu = \mu_0$ contra $H_1 : \mu \neq \mu_0$ (teste bilateral) para uma distribuição normal com σ conhecido:

$$Z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \quad (58)$$

Rejeita-se H_0 ao nível de significância α se $|Z| > z_{\alpha/2}$.

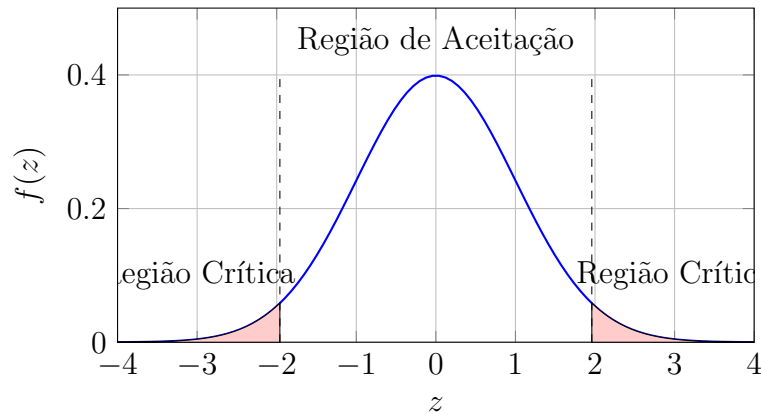


Figura 14: Regiões de aceitação e crítica para um teste bilateral ao nível de significância $\alpha = 5\%$

8 Exemplo de Aplicação Completa

8.1 Contexto do Problema

Suponha que uma empresa de semicondutores quer testar se um novo método de fabricação melhora a vida útil dos seus chips. A vida útil dos chips com o método atual segue aproximadamente uma distribuição normal com média $\mu = 1200$ horas e desvio padrão $\sigma = 150$ horas.

A empresa testa 36 chips produzidos com o novo método e obtém uma vida útil média de 1245 horas. Eles querem saber se há evidência estatística de que o novo método melhora a vida útil.

8.2 Formulação do Teste de Hipóteses

$$H_0 : \mu = 1200 \text{ (o novo método não melhora a vida útil)} \quad (59)$$

$$H_1 : \mu > 1200 \text{ (o novo método melhora a vida útil)} \quad (60)$$

8.3 Cálculo da Estatística de Teste

Considerando o desvio padrão conhecido:

$$Z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} = \frac{1245 - 1200}{150/\sqrt{36}} = \frac{45}{25} = 1.8 \quad (61)$$

8.4 Decisão

Para um teste unilateral com nível de significância $\alpha = 0.05$, o valor crítico é $z_{0.05} = 1.645$.

Como $Z = 1.8 > 1.645$, rejeitamos H_0 ao nível de significância de 5%.

8.5 Cálculo do Valor-p

$$\text{Valor-p} = P(Z > 1.8) = 1 - \Phi(1.8) \approx 0.0359 \quad (62)$$

onde Φ é a função de distribuição acumulada da normal padrão.

8.6 Interpretação

O valor-p de 0.0359 é menor que o nível de significância de 0.05, o que confirma nossa decisão de rejeitar H_0 . Há evidência estatística para concluir que o novo método de fabricação melhora a vida útil dos chips.

8.7 Intervalo de Confiança

Um intervalo de confiança de 95% para a média da vida útil com o novo método é:

$$\bar{X} \pm z_{0.025} \frac{\sigma}{\sqrt{n}} = 1245 \pm 1.96 \cdot \frac{150}{\sqrt{36}} \quad (63)$$

$$= 1245 \pm 1.96 \cdot 25 \quad (64)$$

$$= 1245 \pm 49 \quad (65)$$

$$= [1196, 1294] \quad (66)$$

8.8 Interpretação do Intervalo

O intervalo de confiança de 95% para a média da vida útil dos chips produzidos com o novo método vai de 1196 a 1294 horas. Como este intervalo contém o valor 1200 (embora por uma margem pequena), isso sugere que a melhoria pode não ser tão substancial quanto se esperava inicialmente.

8.9 Conclusão

Embora haja evidência estatística para concluir que o novo método melhora a vida útil dos chips (teste de hipóteses), o intervalo de confiança sugere que a melhoria pode ser modesta. A empresa deve considerar outros fatores, como o custo da implementação do novo método, antes de tomar uma decisão final.

Dica de Interpretação

Sempre considere tanto o resultado do teste de hipóteses quanto o intervalo de confiança ao tomar decisões baseadas em dados estatísticos. O teste de hipóteses fornece uma decisão binária (rejeitar ou não rejeitar), enquanto o intervalo de confiança fornece uma gama de valores plausíveis para o parâmetro, o que pode ser mais informativo para a tomada de decisões práticas.