Resumo Prob.Est.

Capítulo 1: Calculo de Probabilidade

Espaço Amostral (Ω) : Enumeração (finita ou infinita) de todos os resultados possíveis.

$\Omega = A1, A2, A3, ...$

Evento (A): Resultados ou conjunto de resultados possíveis. Chamamos 'evento' qualquer subconjunto do espaço amostral.

Evento Impossível (ø): Conjunto Vazio, pois ele nunca acontecerá.

Probabilidade (P(A)): Probabilidade de um evento **A** ocorrer.

$$P(A) = \frac{A}{\Omega}$$

União - (A ∪ B)

> Pelo menos um ocorre

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

$P(A \cup B) = P(A) + P(B)$, Para eventos mutuamente exclusivos.

Interseção - (A ∩ B)

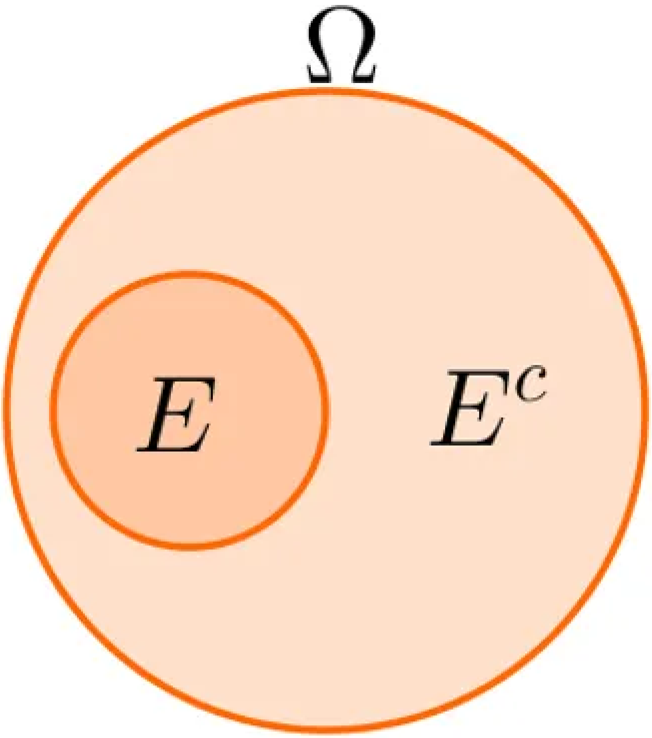
> A **e** B ocorrem

$P(A \cap B) = P(A) \cdot P(B)$ - Eventos Independentes

$P(A \cap B) = P(A) \cdot P(B|A)$ - Eventos Dependentes

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

Evento Complementar (A^c)



$P(A^c) = 1 - P(A)$

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

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

Permutação e Combinação

Uma *permutação de k elementos* é quando a ordem de sorteio importa, e a quantidade de possíveis permutações é dado por

$$P_{n,k} = n(n-1)\ldots(n-k+1) = \frac{n!}{(n-k)!}$$

Uma *combinação de k elementos* é quando a ordem não importa, e a quantidade de possíveis combinações é dada por

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

Axiomas de Probabilidade

Seja Ω um espaço amostral. Uma *probabilidade* é uma função $\mathbb{P}$ que atribui a eventos $A \subseteq \Omega$ um número real $\mathbb{P}(A)$ e satisfaz os seguintes axiomas:

- $0 \leq \mathbb{P}(A) \leq 1$ ou $\mathbb{P}(A) \in [0, 1]$
- $\mathbb{P}(\Omega) = 1$
- Para $A_1, A_2, \ldots, A_n$ disjuntos e tomados 2 a 2:

$$\mathbb{P}(\cup_{i=1}^{\infty} A_i) = \sum_{n=1}^{\infty} \mathbb{P}(A_n)$$

Propriedades de Probabilidade

1. $\mathbb{P}(A^c) = 1 - \mathbb{P}(A)$
2. $A \subset B \to \mathbb{P}(A) \leq \mathbb{P}(B)$
3. $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$
4. Princípio da inclusão-exclusão

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

5. Leis de Morgan

$$\mathbb{P}(\cup_{i=1}^{n} A_i) = 1 - \mathbb{P}(\cap_{i=1}^{n} A_i^c)$$

Capítulo 2: Dependência e condicionamento

Probabilidade Condicional

Para eventos A e B, a probabilidade condicional de *A dado B* é definida como

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

- $P(A|B) = 1 - P(A^c|B)$
- $P(A_1 \cup A_2|B) = P(A_1|B) + P(A_2|B) - P(A_1 \cap A_2|B)$

Independência

Dois eventos A e B são independentes se

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

Independência é o oposto de mutuamente exclusivos (disjuntos)!
Obs: $P(A|B) = P(A)$

Teorema de Bayes

Se A e B são eventos com probabilidade positiva, então

$$P(A|B) = \frac{(P(A) \cdot P(B|A))}{P(B)}$$

obs: $P(A \cap B) = P(A) \cdot (B|A)$
$P(A|B) + P(A|B^c) = 1$

Capítulo 3: Variáveis aleatórias discretas

Variável Aleatória

Seja Ω um espaço amostral. Uma variável aleatória (v.a) é uma função

$$X : w \in \Omega \rightarrow X(w) \in \mathbb{R}$$

Varáveis aleatórias são *características numéricas* de um experimento aleatório representado por w.

Também podemos usar a função $\mathbb{P}_x : A \rightarrow [0, 1]$ definida por

$$\mathbb{P}_x(A) = \mathbb{P}(X \in A)$$

para calcula probabilidades
$\mathbb{P}_x$ é chama de distribuição de probabilidade da v.a. X.

Variável Aleatória Discreta

Uma v.a. X é discreta se o conjunto $\mathit{\Omega}_X \subset R$ de *todos os valores possíveis de X* (não confundir com Ω!) for enumerável.
A função massa de probabilidade (f.m.p.) de uma v.a. X discreta é a função $p_X : \mathit{\Omega} \rightarrow [0, 1]$ dada por

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

> Associa a cada valor possível da variável aleatória discreta suas respectiva probabilidade

Tal que,

$$p(x) = \begin{cases} P(X = x_i), & i = 1, 2, 3, \ldots \\ 0, & \text{caso contrário} \end{cases}$$

Satisfazendo, $p(x) \geq 0$ e $\sum_{x \in Rx} p(x) = 1$

Modelos de Variáveis Aleatórias Discretas

Modelo Bernoulli

> Sucesso ou Fracasso

$$X \sim Ber(p) \qquad (0 < p < 1)$$

$$\mathbb{p}_X(x) = p^x(1-p)^{1-x} \quad \text{para } x \in \{0, 1\}$$

- $E[X] = P$
- $Var(X) = P(1 - P)$

Modelo Binomial

$X \sim Bin(n, p)$

> Chama-se de experimento binomial ao experimento que
>
> - consiste em n ensaios de Bernoulli
> - cujo ensaios são independentes, e
> - para qual a probabilidade de sucessos em casa ensaio é sempre igual a p $(0 < p < 1)$

$$p_X = \mathbb{P}(X = x) = \binom{n}{x} \cdot p^x \cdot (1-p)^{n-x},$$

$$\binom{n}{x} = \frac{n!}{x!(n-x)!} \qquad\qquad x \in \{0, 1, 2, \ldots, n\}$$

- $E[X] = n \cdot p$
- $Var(X) = n \cdot p \cdot (1-p)$

Modelo Hipergeométrico

$X \sim Hip(m, n, k)$

> m $\longrightarrow$ Sucessos
>
> n $\longrightarrow$ Fracassos (N-m)
>
> k $\longrightarrow$ Tamanho da amostra
>
> m+n $\longrightarrow$ N (Ω)

$$p_X = \mathbb{P}(X = x) = \frac{\binom{m}{x} \cdot \binom{n}{k-x}}{\binom{m+n}{k}}$$

- $E[X] = \frac{k \cdot m}{m+n} = k \cdot \frac{m}{N} = kp$
- $Var(X) = \frac{k \cdot m}{m+n} \left[\frac{(k-1)(m-1)}{m+n-1} + 1 \right] = np(1-p)\frac{N-k}{N-1}$

Modelo Geométrico

$X \sim Geom(p)$

> Número de repetições de um ensaio de Bernoulli com probabilidade de sucesso $(0 < p < 1)$ até ocorrer o primeiro sucesso

$$\mathbb{P}(X = x) = p \cdot (1-p)^x \qquad , x \in \mathbb{N}$$

- $E[X] = \frac{1}{p}$
- $Var(X) = \frac{1-P}{p^2}$

Modelo Binomial Negativo

> Tentativas independentes com mesma probabilidade de sucesso $p \in [0, 1]$, sejam realizadas até que se acumule um total de r sucessos.

$$\mathbb{P}(X = x) = \binom{x-1}{r-1} p^r \cdot (1-p)^{x-r}, \qquad x \in \{r, r+1, \dots\}$$

- $E[X] = \frac{r}{p}$
- $Var(X) = \frac{r(1-p)}{p^2}$

Modelo Poisson

$X \sim Poi(\lambda)$

> Eventos Raros

$$p(x) = e^{-\lambda} \cdot \frac{\lambda^x}{x!}, \quad x = \{0, 1, 2, ..\}$$

$E[X] = Var(x) = \lambda$

Capítulo 4: Esperança e variância

Valor Esperado (Esperança, média)

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

Onde, x é o valor de X, e $p(x)$ é a probabilidade de X.
O valor esperado é uma **constante**
O valor esperado é uma medida de centralidade. Esse valor depende somente da distribuição da v.a. X, isto é, da f.m.p. p_X.

Linearidade da Esperança

Se X é v.a., então para todos os números reais a e b

$$\mathbb{E}(\alpha X + b) = \alpha \mathbb{E}(X) + b$$
$$\mathbb{E}(\alpha X + \beta Y) = \alpha \mathbb{E}(X) + \beta \mathbb{Y}$$

Esperança de Função de v.a.

$$\mathbb{E}(g(x)) = \sum_{x \in \Omega_X} g(X) \cdot \mathbb{P}(X = x)$$

Variância e Desvio Padrão

- Variância

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

- Desvio Padrão

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

Capítulo 5: Distribuições de probabilidades conjuntas

Distribuição conjunta

Sejam X e Y v.a.s definidas no mesmo espaço amostral. O par (X,Y) é chamado vetor aleatório bidimensional

O vetor aleatório bidimensional (X,Y) é chamado discreto se X e Y são v.a.s discretas
A função de massa de probabilidade conjunta do v.a. (X,Y) discreto é a função $p_{X,Y}(x,y)$ definida por

$$p_{X,Y}(x,y) = \mathbb{P}(X = x, Y = y) \text{ para todo } (x,y) \in \Omega_X \times \Omega_Y$$

$$\begin{cases} p_{X,Y}(x,y) = \mathbb{P}(X = x, Y = y) \text{ para todo } (x,y) \in \Omega_X \times \Omega_Y \\ \sum_{x \in \Omega_X} \sum_{y \in \Omega_Y} p_{X,Y} = 1 \end{cases}$$

Distribuição Marginal

$$p_X(x) = \sum_{y \in \Omega_Y} \mathbb{P}(X = x, Y = y) \text{ para todo } x \in \Omega_X$$

$$p_Y(y) = \sum_{x \in \Omega_X} \mathbb{P}(X = x, Y = y) \text{ para todo } y \in \Omega_Y$$

Independência

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

Distribuição Condicional

$$p_{X|Y}(x|y) = \frac{p_{X,Y}(x,y)}{p_Y(y)} \quad \text{para todo} \quad x \in \Omega_X$$

Note que a f.m.p. condicional $p_{X|Y}$ satisfaz

$$\begin{cases} p_{X|Y}(x|y) \geq 0 \quad \text{para todo} \quad (x,y) \in \Omega_X \times \Omega_Y \\ \sum\limits_{x \in \Omega_X} p_{X|Y}(x|y) = 1 \quad \text{para cada} \quad y \in \Omega_Y \end{cases}$$

Covariância

$$\begin{aligned} \text{Cov}(X,Y) &= \mathbb{E}[(X - \mathbb{E}(X))(Y - \mathbb{E}(Y))] \\ &= \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y] \quad \text{(Linearidade da esperança)} \end{aligned}$$

As seguintes propriedades são imediatas da definição:

1. $\text{Cov}(X,X) = \text{Var}(X))$
2. $\text{Cov}(X,Y) = \text{Cov}(Y,X)$
3. $\text{Cov}(X,Z) = 0$, se Z é uma variável aleatória constante com probabilidade 1
4. $\text{Cov}(aX,Y) = a \cdot \text{Cov}(X,Y), \quad \text{para } a \in \mathbb{R}$
5. Para quaisquer números reais a, b, c e d,

$$\text{Cov}(aX + bY, cX + dY) = ac \cdot \text{Var}(X) + bd \cdot \text{Var}(Y) + (ad + bc) \cdot \text{Cov}(X, Y)$$

Capítulo 6: Variáveis Aleatórias Contínuas

> Uma $f(x)$ definida sobre o espaço amostral (Ω) e assumindo valores num intervalo de número reais, é dita uma variável aleatória contínua. Uma v.a. é contínua se existir $F_x : \mathbb{R} \to \mathbb{R}$, denominada **função de densidade de probabilidade (f.d.p.)**, satisfazendo:

$$f(x) \geq 0, \forall x \in \mathbb{R}$$
$$P(a \leq x \leq b) = \int_a^b f(x)dx$$

OBS: Satisfazendo $\int_{-\infty}^{+\infty} f(x)dx = 1$

> Variável Aleatória Discreta $\rightarrow$ **Contagem**
> Variável Aleatória Contínua $\rightarrow$ **Medição**

- Função de Distribuição Acumulada (f.d.a.)

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(t)dt, \quad x \subset \mathbb{R} \ \ \forall f(x) = F'(x)$$

$E[X] = \int_{-\infty}^{+\infty} x \cdot f(x)dx$

$Var(X) = \int_{\mathbb{R}} (x - E[X])^2 \cdot f(x)dx = E[X^2] \cdot E[X]^2$

Modelos Probabilísticos

Modelo Uniforme Contínuo

$X \sim Uniforme(a, b)$

> Dizemos que X é uma variável uniforme no intervalo $[a.b], (a, b) \in \mathbb{R}, a < b$, se a função de densidade de probabilidade da variável x é constante nesse intervalo e nula fora dele

- Função de densidade de probabilidade

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

- Função de distribuição

$$F(x) = \begin{cases} 0, & x \leq a \\ \frac{x-a}{b-a}, & a \leq x \leq b \\ 1, & x \geq b \end{cases}$$

- Esperança

$$E[X] = \frac{a+b}{2}$$

- Variância

$$Var(x) = \frac{(a+b)^2}{12}$$

Modelo Exponencial

$X \sim Exp(\lambda)$

> Dizemos que X é uma variável exponencial com parâmetro λ, $\lambda > 0$, se a função de densidade de X é dada por:

$$f(x) = \begin{cases} \lambda \cdot e^{-\lambda x}, & x \geq 0 \\ 0, & x < 0 \end{cases}$$

- Função de distribuição

$$F(x) = \begin{cases} 0, & x < 0 \\ 1 - e^{-\lambda x}, & x \geq 0 \end{cases}$$

- Esperança

$$E[X] = \frac{1}{\lambda}$$

- Variância

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

Modelo Normal

$X \sim N(\mu, \sigma^2)$

$$f(x) = \frac{1}{\sigma . \sqrt{2\pi}} \cdot e^{-\frac{1}{2}(\frac{x-\mu}{\sigma})}, \quad \forall x \in \mathbb{R}$$

Propriedades

- Esperança: $E[X] = \mu$
- Variância: $Var(x) = \sigma^2$
- A f.d.p. de X é simétrica com respeito a $X = \mu$, logo, $f_x(\mu + x) = f_x(\mu - x)$
- Em particular, $\phi(-z) = 1 - \phi(z)$, para todo $z \in R$
- $aX + b \sim N(a\mu + b, a^2\sigma^2)$

Padronização

> Seja X uma v.a. tal que $E[X] = \mu$ e $Var(x) = \sigma^2$, para $N(\mu, \sigma^2)$
> Queremos $N(0,1)$, Então: $E[Z] = 0$ e $Var(Z) = 1$

$$X\ N(\mu, \sigma^2) \rightarrow Z = \frac{X - \mu}{\sigma}\ N(0,1)$$

- $P(a \leq X \leq b) = P\left(\frac{a-\mu}{\sigma} \leq Z \leq \frac{b-\mu}{\sigma}\right) = \phi\left(\frac{b-\mu}{\sigma}\right) - \phi\left(\frac{a-\mu}{\sigma}\right)$
- $P(X \leq a) = P\left(Z \leq \frac{a-\mu}{\sigma}\right) = \phi\left(\frac{a-\mu}{\sigma}\right)$
- $P(X \geq a) = P\left(Z \geq \frac{a-\mu}{\sigma}\right) = 1 - \phi\left(\frac{a-\mu}{\sigma}\right)$

Capítulo 7: Teorema Central do Limite e Lei dos Grandes Números

Lei dos Grandes Números

> Sejam $X_1, X_2, \ldots$ v.a. independentes e identicamente distribuídas (i.i.d.) com média μ e variância σ^2 e $S_n := \sum_{i=1}^{n} X_i$. Sabemos que

$$E(S_n) = n\mu \qquad \mathrm{Var}(S_n) = n\sigma^2$$

Se considerarmos a média $M_n := \frac{S_n}{n}$, temos

$$E(M_n) = \mu \qquad \mathrm{Var}(M_n) = \frac{\sigma^2}{n}$$

Lei dos Grandes Números:

> Para todo $\epsilon > 0$, no limite $n \rightarrow \infty$

$$P(|M_n - \mu| \geq \epsilon) \rightarrow 0$$

Soma de variáveis aleatórias

> Dadas n variáveis aleatórias $X_1, \ldots X_n$, podemos definir uma nova variável aleatória como

$$S_n := \sum_{i=1}^{n} X_i$$

Distribuições da soma de variáveis aleatórias independentes

$$
\begin{aligned}
X_i \sim Ber(p) \quad &\rightarrow \quad \sum_{i=1}^{n} X_i \sim Bin(n,p) \\
X_i \sim Bin(m_i,p) \quad &\rightarrow \quad \sum_{i=1}^{n} X_i \sim Bin\Big(\sum_{i=1}^{n} m_i, p\Big) \\
X_i \sim Poi(\lambda_i) \quad &\rightarrow \quad \sum_{i=1}^{n} X_i \sim Poi\Big(\sum_{i=1}^{n} \lambda_i\Big) \\
X_i \sim N(\mu_i, \sigma_i^2) \quad &\rightarrow \quad \sum_{i=1}^{n} X_i \sim N\Big(\sum_{i=1}^{n} \mu_i, \sum_{i=1}^{n} \sigma_i^2\Big)
\end{aligned}
$$

Teorema Central do Limite

Sejam $X_1, X_2, \ldots$ v.a. independentes e identicamente distribuídas (i.i.d) com média μ e variância σ^2 e $S_n := \sum_{i=1}^{n} X_i$.
Consideramos uma nova variável aleatória

$$Z_n = \frac{S_n - E(S_n)}{\sqrt{\mathrm{Var}(S_n)}} = \frac{S_n - n\mu}{\sigma\sqrt{n}}$$

Temos que $E(Z_n) = 0$ e Var$(Z_n) = 1$. Isso significa que Z_n tem uma distribuição que não se concentra ao redor do valor médio (a variância não vai para zero com n)
O TCL permite conhecer a distribuição limite de Z_n (quando n $\rightarrow \infty$)
Teorema Central do Limite: A distribuição Z_n se aproxima de uma normal padrão $Z \sim N(0,1)$, ou seja, para todos $x \in R$,

$$P\left(\frac{S_n - n\mu}{\sigma\sqrt{n}} \geq x\right) \rightarrow P(Z \geq x) = 1 - \phi(x)$$

$X_1, X_2, \ldots$ i.i.d com média μ e variância σ^2, $S_n = \sum_{i=1}^{n} X_i$

$$P(S_n \leq c) = P\left(\frac{S_n - n\mu}{\sqrt{n}\sigma} \leq \frac{c - n\mu}{\sqrt{n}\sigma}\right) \approx_{TCL} \phi\left(\frac{c-\mu}{\sqrt{n}\sigma}\right)$$

Onde $Z \sim N(0,1)$ e a aproximação melhora a medida que n cresce.

Capítulo 8: Gráficos e Estatística Descritiva (Estatística)

> Estatística é um conjunto de conceitos e métodos para coletar, organizar, analisar e interpretar dados

Estatística descritiva: se preocupa com a organização e apresentação dos dados observados (tabelas, gráficos, medidas descritivas como média e variância...)
Inferência estatística: se preocupa de como dar informação sobre um universo (população) a partir de um conjunto de dados observados (amostra)

Estatística Descritiva

População (N): conjuntos de todos os elementos sob investigação
Amostra (n): subconjuntos finitos da população
Parâmetro: característica numérica de uma população

Exemplo: Pesquisa eleitoral no estado do Rio de Janeiro

$$\begin{array}{ccc} \text{população} & \rightarrow & \text{todos os eleitores} \\ \text{amostra} & \rightarrow & \text{1000 eleitores entrevistados} \\ \text{parâmetro} & \rightarrow & \text{idade média da população} \end{array}$$

Tipologia de Variáveis

Dada uma população ou amostra, podemos estar interessados em várias características dos elementos constituintes; essas características são chamadas variáveis

- *Quantitativa*: assume um valor numérico
  - discreta: número finito ou enumerável
  - contínua: número não enumerável
- *Qualitativa*: classificada em categorias
  - nominal: categorias não-ordenadas
  - ordinal: categorias ordenadas

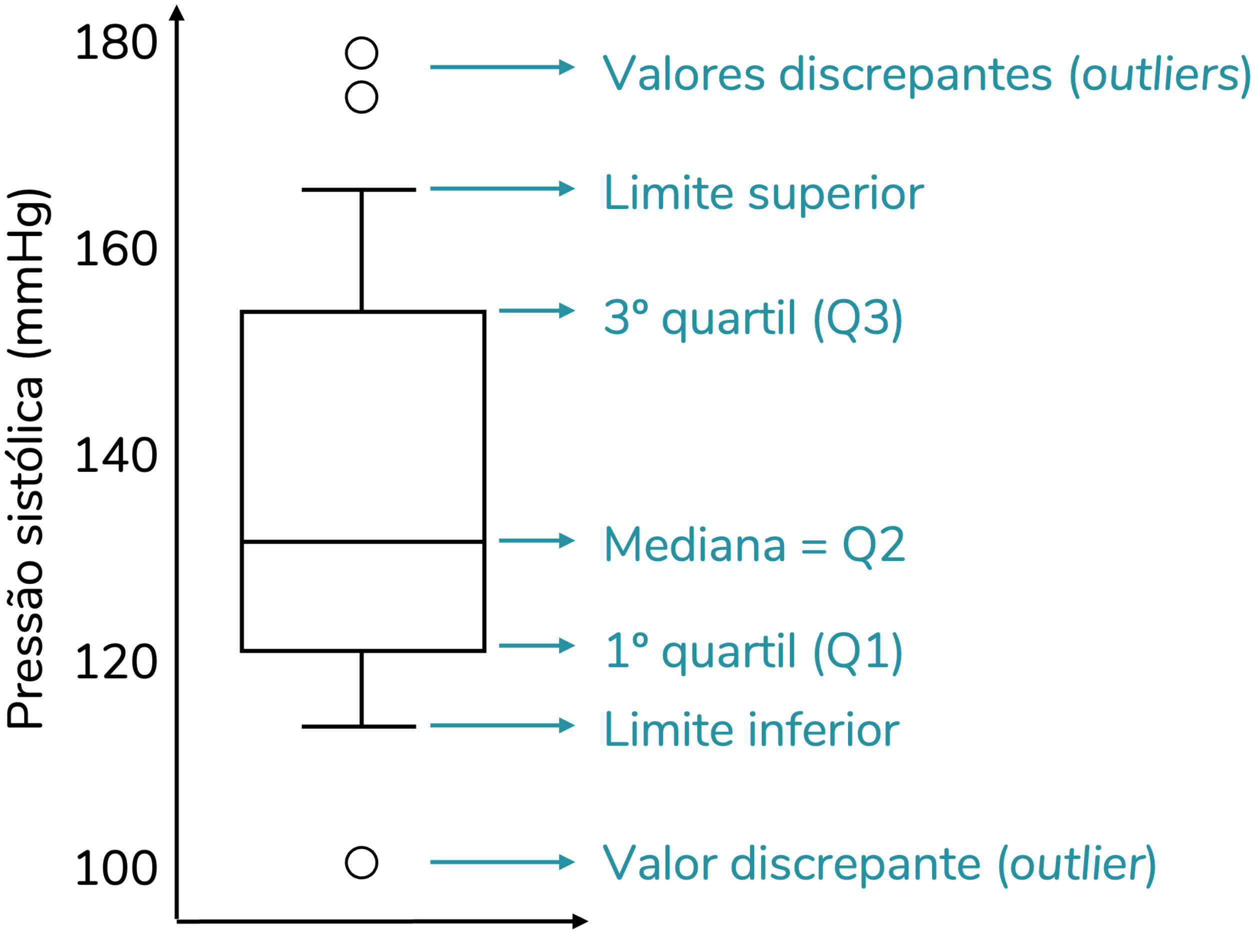
Distribuições de Frequências

Distribuições de Frequências: tabelas e gráficos
Frequência absoluta: número de vezes que cada valor é observado
Frequência relativa: número de vezes que cada valor é observado dividido pelo tamanho da amostra
Frequência acumulada: soma das frequências absolutas dos valores inferiores ou iguais ao valor dado

Gráficos para variáveis qualitativas/quantitativas

- Gráfico de Barra (Barras, Histograma)
- Gráfico de Setores (de Pizza)
- Diagrama de Disperssão
- Boxplot

Construção de Boxplot



1. Encontre os 3 quartis
    1. Quartil 1 (probabilidade 25%): Q1 - valor que deixa 1/4 das observações à esquerda
    2. Quartil 2, ou mediana (probabilidade 50%): Q2 - valor que deixa 2/4 das observações à esquerda
    3. Quartil 3 (probabilidade 75%): Q3 - valor que deixa 3/4 das observações à esquerda
    4. Distância interquartílica: Q3 − Q1 - medida de dispersão
2. Classifica-se como outliers á direita, pontos que se satisfazem o seguinte
    1. $X_i > LS$ (Limite Superior), para $LS = Q_3 + 1.5 \times (Q_3 - Q_1)$
3. Classifica-se como outliers à esquerda, pontos que se satisfazem o seguinte
    1. $X_i < LI$ (Limite Inferior), para $LI = Q_1 - 1.5 \times (Q_3 - Q_1)$
4. A linha se estende até os valores mais extremos do conjunto de dados que não sejam outliers

Medidas de Centralidade

Dada uma coleção de valores de uma variável quantitativa é útil definir formas de resumir esses dados. Uma maneira de fazer isso é através de *medidas de centralidade*.

- Média aritmética: dados valores $x_1, \ldots, x_n$, definimos

$$\overline{x} := \frac{\sum_{i=1}^{n} x_i}{n}$$

- Mediana: dados valores $x_1, \ldots x_n$, sejam $x_1, x_2, \ldots, x_n$ os mesmos valores ordenados. Definimos $Q2$ como

$$Q2 := \begin{cases} x_{\left(\frac{n+1}{n}\right)} & \text{se n é ímpar (valor na posição central)} \\ \frac{x_{\left(\frac{n}{2}\right)}+x_{\left(\frac{n}{2}+1\right)}}{2} & \text{se n é par (média dos valores nas posições centrais)} \end{cases}$$

- Moda: é aquele valor que ocorre com mais frequência

Medidas de Dispersão

Para ter uma caracterização melhor outras medidas, chamada *medidas de dispersão* são introduzidas. Essas são indicadores do grau de espalhamento dos valores em torno da média Dados os valores $x_1, \ldots, x_n$, definimos

- Variância Amostral

$$S^2 := \frac{\sum_{i=1}^{n}(x_i - \overline{x})^2}{n-1} = \frac{\sum_{i=1}^{n} x_i^2 - n\overline{x}^2}{n-1}$$

  - $E(S^2) = \sigma^2$
- Desvio Padrão Amostral

$$s := \sqrt{\frac{\sum_{i=1}^{n}(x_i - \overline{x})^2}{n-1}}$$

- Coeficiente de variação amostral:

$$cv := \frac{s}{\overline{x}}$$

Capítulo 9: Intervalos de Confiança

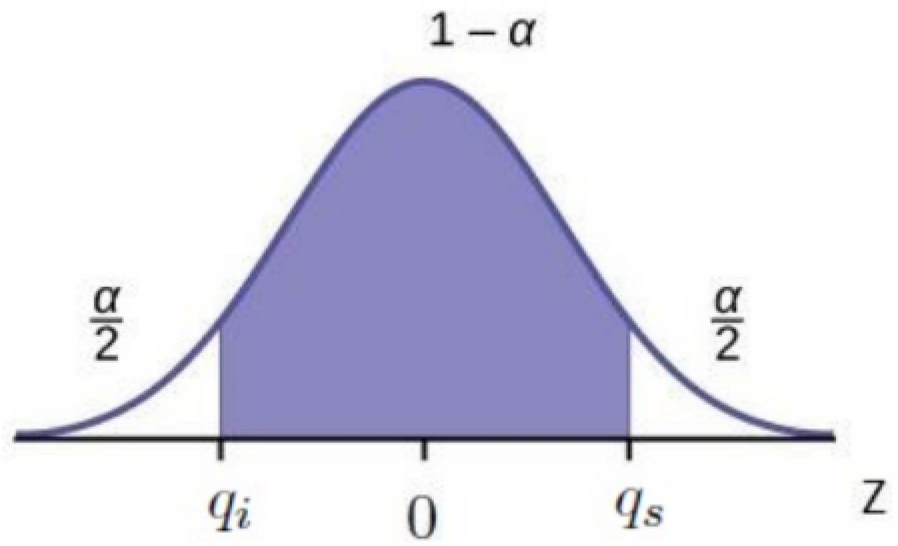
Estimativa Pontual

1. Uso $\overline{x}$ para estimar $E(x)$
2. Uso $S^2 = \frac{1}{n-1}\sum(x_i - \overline{x})^2$ para estimar $\mathrm{Var}(x) = \sigma^2$
3. Uso $S = \sqrt{S^2}$ para estimar $DP = \sqrt{\mathrm{Var}(x)} = \sigma$

Caso σ conhecido

Se

$$X \sim N(\mu, \frac{\sigma^2}{n}) \to Z = \frac{\overline{X} - \mu}{\frac{\sigma}{\sqrt{n}}} \sim N(0,1)$$



Dado $\alpha \in (0, 1)$, o valor de $(1 - \alpha = \gamma)$ é o que chamamos de nível de confiança do intervalo $IC_\mu(\gamma)$. Em geral, γ será um valor próximo de 1, como por exemplo $0.95, 0.98$ ou até 0.99.

Construindo o intervalo

Digamos que $\gamma = 0.95$ e $\frac{\alpha}{2} = q$:

$$\begin{aligned}\gamma = 0.95 &= P(-q < Z < q) \\ &= P(-q < \frac{\overline{X} - \mu}{\frac{\sigma}{\sqrt{n}}} < q) \\ &= P(-q\frac{\sigma}{\sqrt{n}} < \overline{X} - \mu < q\frac{\sigma}{\sqrt{n}}) \\ &= p(-\overline{X} - q\frac{\sigma}{\sqrt{n}} < \mu < -\overline{X} + q\frac{\sigma}{\sqrt{n}}) \\ &= p(\overline{X} - q\frac{\sigma}{\sqrt{n}} < \mu < \overline{X} + q\frac{\sigma}{\sqrt{n}})\end{aligned}$$

Logo,

$$IC_\mu(\gamma) = \left[\overline{X} - q\frac{\sigma}{\sqrt{n}}; \overline{X} + q\frac{\sigma}{\sqrt{n}}\right]$$

Com $\gamma = 0.95$ temos, q = 1.96.
(Procurar na tabela normal $P(Z < q) = \frac{\gamma+1}{2} = 0.975 \rightarrow q$)

Caso σ desconhecido

Estimamos σ^2 usando S^2 e padronizamos $\overline{X}$ por:

$$T = \frac{\overline{X} - \mu}{\frac{S}{\sqrt{n}}}$$

A variável aleatória T é diferente de Z, e sua distribuição deixa de ser normal e passa a ser t-Student:

$$T \sim (n - 1)$$

Onde, $(n-1)$ são graus de liberdade, e n é o tamanho da amostra.
Portanto,

$$IC_{\mu}(\gamma) = \left[\overline{X} - t\frac{S}{\sqrt{n}}; \overline{X} + t\frac{S}{\sqrt{n}}\right]$$

Intervalos de Confiança para Proporção

Dada uma amostra $X_1, \ldots, X_n \sim Bern(p)$:

$$\overline{X} = \frac{X_1 + \cdots + X_i}{n} = \hat{p} \approx N(E(\overline{X}), \mathrm{Var}(\overline{X})) \therefore N\left(p, \frac{p(1-p)}{n}\right)$$

- $E(\overline{X}) = E(X_i) = p$
- $\mathrm{Var}(\overline{X}) = \frac{\mathrm{Var}(X_i)}{n} = \frac{p(1-p)}{n}$
- Z = $\frac{\hat{p}-p}{\sqrt{\frac{p(1-p)}{n}}}$ $\approx$ N(0,1)

  Temos que,

$$IC_p^*(\gamma) = P\left(\hat{p} - q\sqrt{\frac{p(1-p)}{n}} < p < \hat{p} + q\sqrt{\frac{p(1-p)}{n}}\right)$$

  Onde $q = \frac{1+\gamma}{2}$.
  Abordagem não conservadora: Substituímos p na equação por $\hat{p} \sim N(0,1) = \gamma$
  Abordagem conservadora: Consideramos um γ acima do nível de confiança fixado inicialmente

$$IC_{\mu}(\gamma) = \left[\hat{p} - q\frac{1}{2\sqrt{n}}; \hat{p} + q\frac{1}{2\sqrt{n}}\right]$$

Capítulo 10: Testes de Hipóteses

Construção de Testes de Hipóteses

Considere $X_1, \ldots, X_n$ uma amostra aleatória com $E(X_i) = \mu$ e $\mathrm{Var}(X_i) = \sigma^2$
Formulamos duas hipóteses:

$$\begin{cases} H_0 : \mu = x, & \text{Hipótese Nula} \\ H_1 : \mu \neq x, & \text{Hipótese Alternativa} \end{cases}$$

Calcule a média amostral: $\overline{X}$

Tipos de Erro

- Tipo 1: Rejeitar H_0 dado que H_0 é verdade
- Tipo 2: Não rejeitar H_0 dado que H_0 é falsa (i.e. dado H_1)

Realização de Testes de Hipóteses

1. Definir α: nível de significância, como 1%, 5%, ou 10%.
2. Formular Hipóteses H_0 e H_1
3. Estabelecer a estatística de teste: determina plausibilidade de H_0.
    1. t-Student: $t(n-1,\gamma)$
    2. Normal: $N(\mu,\sigma^2) = N(\mu,\frac{s}{\sqrt{n}})$
4. Identificar a região crítica (RC): valores da estatística incompatíveis com H_0.

    Unilateral:

    α = P(Rejeitar $H_0 \,|\, H_0$ verdade) = $P\left(\left|\frac{\bar{X}-\mu}{\frac{S}{\sqrt{n}}}\right| < q|\mu = x\right) \rightarrow$ RC $= Z_n < q$

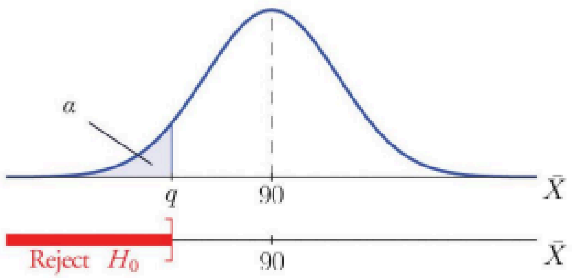
    Bilateral :

    α = P(Rejeitar $H_0 \,|\, H_0$ verdade) = $P\left(\left|\frac{\bar{X}-\mu}{\frac{S}{\sqrt{n}}}\right| \neq q|\mu = x\right) \rightarrow$ RC $= -q < Z_n < q$

5. Conclusão:

    - Rejeitar H_0 se a estatística de teste estiver na RC.
    - Não rejeitar H_0 caso contrário.

1. Defina o nível de significância α (pequeno: 1%, 5%, 10%).
2. Se vale $\mathcal{H}_1 : \mu < 90$, então valores muito menores do que 90 para $\bar{x}$ estarão de acordo com $\mathcal{H}_1$.
3. Defina quais valores de $\bar{x}$ são plausíveis dado H_0 e quais não são.



4. Se $\bar{x} < q$, considero $\bar{x}$ incompatível com $\mu = 90$ (é muito improvável que $\bar{x} < q$ se $\mu = 90$).
5. Região crítica: $RC := \{\bar{x} \in \mathbb{R} : \ \bar{x} < q\}$
6. Rejeita-se $\mathcal{H}_0$ se observarmos $\bar{x} \ \in \ RC$

Teste para Proporções

$X_1,\ldots,X_n$: Amostra aleatória $X_i \sim Bern(p)$

- Seja $X_1, \ldots, X_n$ uma amostra de tamanho n e $\hat{p} = \frac{1}{n}\sum_{i=1}^{n} X_i$ a proporção amostral (estimador de p)
- TCL $\Longrightarrow \dfrac{\hat{p}-p}{\sqrt{\frac{p(1-p)}{n}}} \approx N(0,1)$, se n é suficientemente grande $(np(1-p) \geq 3)$
- Temos aqui também três testes de interesse:

$$\begin{cases} \mathcal{H}_0: & p \geq p_0 \\ \mathcal{H}_1: & p < p_0 \end{cases}, \quad \begin{cases} \mathcal{H}_0: & p \leq p_0 \\ \mathcal{H}_1: & p > p_0 \end{cases} \quad \text{e} \quad \begin{cases} \mathcal{H}_0: & p = p_0 \\ \mathcal{H}_1: & p \neq p_0 \end{cases}$$

Estatística de Teste:

$$Z = \left(\left| \frac{\hat{p}-p}{\sqrt{\frac{p(1-p)}{n}}} \right| \neq q | \mu = \overline{x} \right)$$

P-Valor

P-valor é a probabilidade sob H_0 de observarmos um valor mais extremo (de acordo com H_1) para estatística de teste do que aquele observado na amostra.

Regra de Teste: Rejeitamos H_0 se P-valor $< \alpha$

Quanto menor o p-valor mais evidência temos contrária a H_0

Portanto, mais geralmente, calculamos o p-valor das seguintes formas:

- Teste unilateral $\begin{cases} \mathcal{H}_0: & \mu \geq \mu_0 \\ \mathcal{H}_1: & \mu < \mu_0 \end{cases} \Longrightarrow \tilde{\alpha} = P(\overline{X} < \overline{x} \mid \mu = \mu_0)$
- Teste unilateral $\begin{cases} \mathcal{H}_0: & \mu \leq \mu_0 \\ \mathcal{H}_1: & \mu > \mu_0 \end{cases} \Longrightarrow \tilde{\alpha} = P(\overline{X} > \overline{x} \mid \mu = \mu_0)$
- Teste bilateral $\begin{cases} \mathcal{H}_0: & \mu = \mu_0 \\ \mathcal{H}_1: & \mu \neq \mu_0 \end{cases} \Longrightarrow \tilde{\alpha} = P(|\overline{X} - \mu_0| > |\overline{x} - \mu_0| \mid \mu = \mu_0)$

Define o nível de significância mínimo (α) necessário para rejeitar H_0.

Cálculo:

1. Construa o teste de hipótese
2. Defina o nível de significância (α)
3. Tome uma amostra e calcule sua estatística de teste T

$Z_{teste} = \frac{\hat{p}-p}{\sqrt{\frac{p(1-p)}{n}}}$

4. $p - valor$: p($H_1 | H_0$)

1. Teste unilateral

$Z \sim N(0,1) = 1 - \phi(Z)$

3. Teste bilateral

$Z \sim N(0,1) = 2 \cdot (1 - \phi(Z))$

1. Se

1. $p - valor < \alpha \rightarrow$ Rejeito H_0
2. $p - valor > \alpha \rightarrow$ Não Rejeito H_0