

Resumo

21 de abril de 2023 00:01

Variável Estatística

→ Qualitativa:

→ Nominal:

→ Nomes sem uma ordem de categorias.

→ Ordinal:

→ Nomes com uma ordem de categorias.

→ Quantitativa:

→ Discreta (Dominio = conj. (in)finito de inteiros):

→ Associa a contagens.

→ Contínua (Dominio = conj. de n^o reais):

→ Associa a medidas.

Tabelas de Freqüências

i	x_i	m_i	N_i	f_i	f_i
Nº Linha	Valor ou Classe	Freq. Absoluta (contagem)	Freq. Absolutas Acumuladas	Freq. Relativas $(\frac{m_i}{m})$	Freq. Relativas Acumuladas

→ Sem classes:

→ V. Qualitativas;

→ V. Quantitativas Discretas (^{Pouca}_{Muita} Distinção).

→ Com classes:

→ V. Quantitativas contínuas;

→ V. Quantitativas Discretas (^{Muita}_{Muita} Distinção).

→ N.º de classes: $k = 1 + \log_2 m = 1 + \frac{\ln m}{\ln 2}$

→ Amplitude das classes: $h = \frac{\max(x_i) - \min(x_i)}{k}$

Graficos

→ Barra:

→ Dados Qualitativos;

→ Dados Quantitativos Discretos.

→ Circulares:

→ Dados Qualitativos;

→ Dados Quantitativos Discretos (^{Menos}_{Muito} Usado).

→ Histogramas:

→ Dados Quantitativos Contínuos
(em classes).

Métricas de Localização

→ Central:

→ Mediana (maior n; 1 mais repetido)

→ Média ($\text{mean}()$) | $\bar{x} = (\sum_{i=1}^n x_i) / n$)

→ Mediana ($\text{median}()$)

No caso de V. nominais afenos se contabiliza a moda!

→ Não central:

→ Quantis:

$$mf = m \times 0.5:$$

- Se mf é inteiro: $\tilde{x} = \frac{x_{(mf)} + x_{(mf+1)}}{2}$
- Se não: $\tilde{x} = x_{([mf]+1)}$ Mediana

$$Q_1 \rightarrow mf = m \times 0.25$$

$$Q_2 \rightarrow \tilde{x}$$

$$Q_3 \rightarrow mf = m \times 0.75$$

Métricas de Dispersão

→ Dispersão Absoluta:

→ Amplitude Total: ($\text{max}() - \text{min}()$)

→ Amplitude Interquartil: $AIQ = Q_3 - Q_1 = \underline{IQR()}$

→ Variância: $\sigma^2 = \underline{\text{Var}()}$

→ Desvio Padrão: $\sigma = \underline{s() = \sqrt{\sigma^2}}$

→ Dispersão Relativa:

→ Coeficiente de Variação: $cV = \frac{\sigma}{\bar{x}} \times 100\%$

Métricas de Simetria

→ $b_s = \text{Skewness}() = \frac{1}{n} \cdot \sum_{i=1}^n \left(\frac{x_i - \bar{x}}{\sigma} \right)^3$

• $b_s = 0 \rightarrow$ Simétrica

• $b_s > 0 \rightarrow$ Assimétrica positiva / para a direita

• $b_s < 0 \rightarrow$ Assimétrica negativa / para a esquerda

V. Al. Discretas

$X - N: \mathbb{Z} \dots \dots$ (contagens)

$f(n) = P(X = n) \rightarrow$ função Probabilística

↳ Nota:

$\rightarrow f(n) \geq 0, \forall n$

$$\rightarrow \sum f(n) = 1 \quad \rightarrow f(n) = \begin{cases} 0, n \leq \text{Min} \\ \dots, \text{Min} < n < \text{Max} \\ 1, n \geq \text{Max} \end{cases}$$

n	Min	...	Max
$f(n)$

$F(n) = P(X \leq n) \rightarrow$ função Distribuição

$$P(a < X \leq b) = F(b) - F(a)$$

$$P(\bar{a}) = 1 - P(a)$$

$$P(a|b) = \frac{P(a \cap b)}{P(b)}$$

$$\mathbb{E}[X] = \mu = \sum n \cdot f(n)$$

$$\mathbb{E}[X^2] = \sum n^2 \cdot f(n)$$

$$\mathbb{E}[ax + b] = a \cdot \mathbb{E}[x] + b$$

$$\mathbb{E}[x+y] = \mathbb{E}[x] + \mathbb{E}[y]$$

Valor Esperado

$$\mathbb{E}[a] = a$$

$$\mathbb{V}[X] = \mathbb{E}[X^2] - \mathbb{E}^2[X] \quad \mathbb{V}[a] = \emptyset$$

$$\downarrow \sigma^2 = \mathbb{E}[(X-\mu)^2]$$

$$= \sum (n-\mu)^2 \cdot f(n)$$

Variância

$$\mathbb{V}[ax+b] = a^2 \cdot \mathbb{V}[x]$$

$$\sigma = \sqrt{\mathbb{V}[X]} \rightarrow$$
 Desvio Padrão

Intervalo de
Resultados:
 $0 \leq P(a) \leq 1$

U → em
^ → e
I → Sabemos que
S → se

V. Al. Contínuas

$x - \dots \dots$

$f(x) = P(x=n) \rightarrow f.$ Densidade e Probabilidade

$$\hookrightarrow f(x) = f'(x)$$

\hookrightarrow Afirmas é F.D.P. Se:

$$\rightarrow f(x) \geq 0, \forall x$$

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

$F(x) = P(X \leq x) = \int_{-\infty}^x f(t) dt \rightarrow f.$ Distribuição

$$P(a < x < b) = F(b) - F(a)$$

$$E[x] = \int_{-\infty}^{+\infty} x \cdot f(x) dx \quad \text{Valor Esperado}$$

$$E[x^2] = \int_{-\infty}^{+\infty} x^2 \cdot f(x) dx$$

$$E[a] = a$$

$$E[ax + b] = a \cdot E[x] + b$$

Primitivas Importantes

$$P(x^k) = \frac{x^{k+1}}{k+1}, k \neq -1$$

$$P(k) = kx$$

Exemplo de uma Integral:

$$\begin{aligned} \int_2^4 (1 - \frac{x}{4}) dx &= \int_2^4 1 dx - \int_2^4 \frac{x}{4} dx \\ &= \left[x \right]_2^4 - \frac{1}{4} \left[\frac{x^2}{2} \right]_2^4 \frac{1}{4} P(x) \\ &= (4 - 2) - \frac{1}{4} (\frac{4^2}{2} - \frac{2^2}{2}) \\ &= 2 - \frac{1}{4} (\frac{16}{2} - \frac{4}{2}) = 2 - \frac{12}{8} \\ &= 1/8 \end{aligned}$$

$$V[x] = E[x^2] - E^2[x] = \sigma^2 \rightarrow \text{Variância}$$

$$V[a] = \phi$$

$$\sigma = \sqrt{\sigma^2} = \sqrt{V[x]}$$

$$V[ax + b] = a^2 \cdot V[x]$$

\rightarrow Desvio Padrão

Distribuições Tóricas Zombeiras

- V. Al. Discretas:

→ Uniforme Discreta:

- Para situações em que todos os valores possíveis têm a mesma probabilidade de ocorrer.
- Aplica-se apenas a conjuntos de val. discretos.

$$X \sim U(n)$$

$n = b - a + 1 \Rightarrow N$ elementos do domínio.

$$f(x) = P(X=x) = \frac{1}{n} = \begin{cases} 1/n, & \forall x \in D_x \\ 0, & \text{c.c.} \end{cases}$$

- $D_x = \text{Inteiros consecutivos:}$

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

$$V[x] = \frac{(b-a+1)^2 - 1}{12}$$

- $D_x \neq \text{Inteiros consecutivos:}$

$$E[x] = \frac{1}{n} \sum_{i=1}^n x_i$$

$$V[x] = \frac{1}{n} \sum_{i=1}^n x_i^2 - E^2[x]$$

$$D_x = \{a, a+1, a+2, \dots, b\}$$

↓ 1º do domínio ↓ Último do domínio

→ Binomial:

→ Para situações em que há 2 resultados possíveis e interessante calcular a probabilidade de m-2e sucessos. $E[x] = m \cdot f$

$$X \sim B(n, f) \rightarrow P(\text{Sucesso}) \quad V[x] = m \cdot f \cdot (1-f)$$

$X \sim N$: 2e Sucessos em n foras.

$$f(x) = P(x = x) \Rightarrow \underline{c_{\text{binom}}}(x, n, f)$$

$$F(x) = P(x \leq x) \Rightarrow \underline{f_{\text{binom}}}(x, n, f)$$

$$F(k) = \text{prob} (\Rightarrow k = F^{-1}(\text{prob})) \Rightarrow \underline{g_{\text{binom}}}(\text{prob}, n, f)$$

$$D_x = \{0, 1, 2, \dots\}$$

$$m = m_{x_1} + m_{x_2}$$

$$\text{Atribuição da Binomial: } Y = X_1 + X_2 \sim B(m, f)$$

$$= \text{em todos} \leftarrow$$

→ Poisson:

→ Ocorrência de eventos raros num intervalo de tempo.

→ É calcular a probabilidade de ocorrência de eventos raros num intervalo de tempo.

$$X \sim P(\lambda) \quad \lambda = E[x] = V[x]$$

$X \sim N$: 2e ocorrências num intervalo de medida.

$$f(x) = P(x = x) \Rightarrow \underline{c_{\text{pois}}}(x, \lambda)$$

$$F(x) = P(x \leq x) \Rightarrow \underline{f_{\text{pois}}}(x, \lambda)$$

$$F(x) = \text{prob} (\Rightarrow x = F^{-1}(\text{prob})) \Rightarrow \underline{g_{\text{pois}}}(\text{prob}, \lambda)$$

$$D_x = \{0, 1, 2, \dots\}$$

$$\lambda = \lambda_{x_1} + \lambda_{x_2} \leftarrow$$

$$\text{Atribuição da Poisson: } Y = X_1 + X_2 \sim NP(\lambda)$$

• V. Al. Contínuas:

→ Uniforme Contínua:

→ É igual à Uniforme Discreta, mas para intervalos contínuos.

$$X \sim U(a, b) \quad D_x = [a, b]$$

$$f(x) \rightarrow \underline{d_{unif}}(x, a, b) \quad a \rightarrow 1^{\text{º}} \text{ valor mínimo}$$

$b \rightarrow$ último valor mínimo

$$F(x) \rightarrow \underline{f_{unif}}(x, a, b) \rightarrow P(X \leq x) = \frac{x-a}{b-a}$$

$$F^{-1}(x) \rightarrow \underline{q_{unif}}(f^{-1}(x), a, b)$$

$$\mathbb{E}[X] = \frac{a+b}{2} \quad V[X] = \frac{(b-a)^2}{12}$$

→ Exponencial:

→ Tempo/dist. entre ocorrências sucessivas.

→ É usada para calcular a probabilidade de um evento ocorrer num intervalo de tempo.

$$X \sim Exp(\sigma)$$

$$D_x = [0, +\infty[$$

$$\mathbb{E}[X] = \sigma$$

$$V[X] = \sigma^2$$

Até a 1^a ocorrência,
entre 2 ocorrências

Falta 2 Memória: $P(X \geq a+b | X \geq a) =$

$$= P(X \geq b)$$

$$f(x) \rightarrow \underline{d_{exp}}(x, 1/\sigma)$$

$$F(x) \rightarrow \underline{f_{exp}}(x, 1/\sigma) \rightarrow P(X \leq x) =$$

$$F^{-1}(x) \rightarrow \underline{q_{exp}}(f^{-1}(x), 1/\sigma) = 1 - e^{-\frac{x}{\sigma}}, x \geq 0$$

→ Normal:

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

$$f(x) \rightarrow \underline{c_{norm}}(x, \mu, \sigma)$$

$\mu \rightarrow$ Média
 $\sigma \rightarrow$ Desvio

$$F(x) \rightarrow \underline{f_{norm}}(x, \mu, \sigma) \rightarrow P(x \leq | < | n)$$

$$F^{-1}(n) \rightarrow \underline{f_{norm}}(f_{rob}, \mu, \sigma)$$

$$\mathbb{E}[x] = \mu$$

$$\mathbb{V}[x] = \sigma^2$$

$$D_x = \mathbb{R}$$

→ Normal Standard / Regulada:

$$Z = \frac{x - \mu}{\sigma} \sim N(0, 1)$$

$$\rightarrow P(Z \leq | < | n) = \underline{\Phi}(n)$$

$$\Phi(-n) = 1 - \underline{\Phi}(n)$$

$$f(x) \rightarrow \underline{c_{norm}}(x)$$

$$\rightarrow \underline{\Phi}(n) \rightarrow \underline{f_{norm}}(n)$$

$$\Phi^{-1}(n) = \underline{f}_{(n)} \rightarrow \underline{f_{norm}}(f_{rob})$$

$$P(x \leq | < | k) = \alpha$$

$$\underline{\Phi}(k) = \alpha$$

$$(=) k = \Phi^{-1}(\alpha) = \underline{f}_{(\alpha)}$$

Dados no

Exercícios

• Exemplos entre Normal e Normal Reluzida:

Exemplo 25 dos slides

X - Altura a que crescem festeiros, em m.

$$\mu = ?$$

$$\sigma = 1.1$$

$$E[X] = ? = \mu = 17.41\text{m}$$

$$P(X \geq 16) = 0.9 (=) 1 - P(X < 16) = 0.9 (=)$$

$$F(16) = 0.1 (=) \Rightarrow Z = \frac{x - \mu}{\sigma} \sim N(0,1)$$

$$P(Z < \frac{16 - \mu}{\sigma}) = 0.1 (=)$$

$$\Phi\left(\frac{16 - \mu}{1.1}\right) = 0.1 (=) \quad \Phi(x) \text{ é } \sigma F(x) \text{ de uma normal } \phi 1.$$

$$\frac{16 - \mu}{1.1} = \Phi^{-1}(0.1) = \underline{\text{qnorm}}(0.1) = -1.282 (=)$$

$$16 - \mu = -1.282 (=) \mu = 17.41\text{m},$$

• Exemplos entre Normal e Binomial:

X - Comprimentos de um salto feito por um atleta.

$$X \sim N(7.23, 0.33)$$

Qual a probabilidade de, em 5 saltos, haver 2 com mais de 7.5m?

$$Y \sim B(n, p) (=) Y \sim B(5, 0.2066)$$

n = 5 saltos

$$p = P(\text{Sucesso}) = P(X > 7.5) = 1 - P(X \leq 7.5) \\ = 1 - \text{pnorm}(7.5, 7.23, 0.33) = 0.2066,$$

$$P(Y=2) = \text{binom}(2, 5, 0.2066) = 0.2132,$$

- Propriedades da Normal (V. Indefensantes):

- Aritimetica da Normal:

$$y = \sum_{i=1}^k x_i \sim N\left(\sum_{i=1}^k \mu_i, \sqrt{\sum_{i=1}^k \sigma_i^2}\right)$$

- Combinações Lineares da Normal:

$$y = \sum_{i=1}^k a_i x_i \sim N\left(\sum_{i=1}^k a_i \mu_i, \sqrt{\sum_{i=1}^k a_i^2 \sigma_i^2}\right)$$

- Teorema do Limite Central:

$$T = \sum_{i=1}^n x_i \sim N(n\mu, \sqrt{n\sigma^2})$$

(=)

É Parecido à Aritimetica, mas é para distribuições desconhecidas!

$$T = \sum_{i=1}^n x_i \sim N(n\mu, \sigma \sqrt{n})$$

Para $n \geq 30!$

→ Aprox. da Binomial pela Normal:

$$X \sim B(n, p) \xrightarrow{n \rightarrow +\infty} X \sim (np, \sqrt{npq})$$

→ Aprox. da Poisson pela Normal:

$$X \sim P(\lambda) \xrightarrow{\lambda \rightarrow +\infty} X \sim N(\lambda, \sqrt{\lambda})$$

- Exemplos das Propriedades:

- Exemplos 26:

x_1 - Pontuação do fogador 1

x_2 - Pontuação do fogador 2

$$x_1 \sim N(84, 5) \quad x_2 \sim N(85, 8)$$

1) Atividade

$$P(x_1 + x_2 > 176)$$

$$Y = x_1 + x_2 \sim N(84 + 85, \sqrt{5^2 + 8^2})$$

$$Y \sim N(169, \sqrt{89})$$

$$P(Y > 176) = 1 - f(176)$$

$$= 1 - \text{fnorm}(176, 169, \sqrt{89})$$

$$= 0.229,,$$

2) Combinação Linear

$$P(x_2 > x_1) (=) P(x_1 - x_2 < 0)$$

$$W = x_1 - x_2 \sim N(84 - 85, \sqrt{5^2 - (-1)^2 8^2})$$

$$W \sim N(-1, \sqrt{89})$$

$$P(W < 0) = \text{fnorm}(0, -1, \sqrt{89})$$

$$= 0.5422,,$$

• Ejemplo 28: T. Límite Central

x_i - Contenido, en l., la garrafa i.

$$i = [1, 500] \quad \mu_i = 1 \quad \sqrt{v[x]} = 0.0201$$

\bar{T} - Zafacilade é um recipiente com 500 garrafas, em l..

$$\bar{T} = \sum_{i=1}^{500} x_i \sim N(500, 0.0201\sqrt{500})$$

$$P(\bar{T} > 500.1) = 1 - F_{\bar{T}}(500.1) =$$

$$1 - F_{\text{norm}}(500.1, 500, 0.0201 \times \sqrt{500}) = 0.412,$$

\rightarrow Qui-Quadrado: \rightarrow T-Student:

$$X \sim \chi^2(n)$$

$$X \sim t(n)$$

$$f(x) \rightarrow \underline{\text{chisq}}(n, n)$$

$$f(x) \rightarrow \underline{t}(n, n)$$

$$f^{-1}(n) \rightarrow \underline{\text{fchisq}}(fob, n)$$

$$f^{-1}(n) \rightarrow \underline{ft}(fob, n)$$

$$f^{-1}(n) \rightarrow \underline{qchisq}(fob, n)$$

$$f^{-1}(n) \rightarrow \underline{qt}(fob, n)$$

\rightarrow F de Snedecor:

$$X \sim F(m, n)$$

$$f(x) \rightarrow \underline{f}(n, m, n)$$

$$f(x) \rightarrow \underline{ff}(n, m, n)$$

$$f^{-1}(n) \rightarrow \underline{ff}(fob, m, n)$$

Relações Entre Distribuições

• Exponencial e Poisson

Se

$$X \sim P(\lambda)$$

$X \rightarrow N$: λ ocorrências num intervalo de tempo t .

$\lambda \rightarrow N$: Média de ocorrências num intervalo de tempo t .

e

$$Y \sim E\text{nf}(\theta)$$

$Y \rightarrow$ Tempo de espera entre ocorrências sucessivas.

$\sigma \rightarrow$ Tempo de espera médio entre oe. sucessivas.

então

$$\theta = \frac{t}{\lambda} \quad (\Rightarrow) \quad \lambda = \frac{t}{\theta}.$$

Exemplos com a Regra de 3 Simples:

• P \rightarrow Cnf: (Ex. 21)

$X \sim N$: 2 avarias em 8h.

2 Avarias - 8h.
1 Avaria - σ

$$X \sim P(2)$$

Y - Tempo entre avarias consecutivas.

$$Y \sim E\text{nf}(\sigma) = E\text{nf}\left(\frac{10}{2}\right) = E\text{nf}(4)$$

• Enf \rightarrow P: (Ex. 22)

X - Tempo entre chegadas consecutivas, em s.

$$X \sim E\text{nf}(90)$$

$Y \sim N$: 2 utentes que chegam em 6 minutos.

$$Y \sim P(\lambda) = P\left(\frac{10}{1.5}\right) = P(4)$$

Testes

• Paramétricos:

→ Servem para testar:

→ μ → Média

→ σ^2 → Variância

→ f → Proportion

→ $\mu_1 - \mu_2$ → Diferença de Médias

→ σ_1^2 / σ_2^2 → Quociente de Variâncias

→ $f_1 - f_2$ → Diferença de Proportiones

• Não Paramétricos:

→ Testes Ajustamentos:

→ Qui-Quadrado

→ Kolmogorov-Smirnov

→ Lillie-Fors

→ SH

→ Diferença de Medianas:

→ Wilcoxon

→ Mann-Whitney

• Escolha do Tipo de Teste

H_0	H_1	Tipo de Teste
$\sigma = \sigma_0$	$\sigma \neq \sigma_0$	Bilateral
	$\sigma > \sigma_0$	unilateral direito
	$\sigma < \sigma_0$	unilateral esquerdo
$\sigma \geq \sigma_0$	$\sigma < \sigma_0$	unilateral esquerdo
$\sigma \leq \sigma_0$	$\sigma > \sigma_0$	unilateral direito

Passos Para os Testes e Hipóteses:

- 1º- Formular hipóteses;
- 2º- Definir a distribuição amostral e a estatística de teste;
- 3º- Tomar a decisão: Rejeita-se H_0 se:
 - Pela Região Crítica: ET pertencer à RC.
 - Pelo P-Value: $P\text{-Value} \leq \alpha$
- 4º- Fazer a conclusão

Testes para Diferenças em Populações

- Testes para Diferenças de Médias:
 - Há normalidade / independência?

Vê-se quais testes e ajustamentos:

- Kolmogorov-Smirnov;
- Lilliefors;
- Shapiro-Wilk.

→ Sim (Cap. 5):

Teste Paramétrico ($\mu_1 - \mu_2$)

→ Não: → Não (Cap. 6.2):

$$n \geq 30 \quad n < 30$$

Teste Não Paramétricos:

- Wilcoxon (Vars. Emparelhadas)
- Mann-Whitney (Vars. Independentes)