Anotaçõess de Inferência Frequentista

Gustavo S. Garone

2025-02-01

Índice

Preface

This is a Quarto book.

To learn more about Quarto books visit https://quarto.org/docs/books.

1 Introdução

Este livrete é um compilado de minhas anotações para as Disciplinas MAE0225 e MAE301, ambas ministradas pelo Professor Alexandre Galvão Patriota.

Para quaisquer erros, sugestões e críticas, contactar em gustavo.garone arroba usp.br

2 Modelos Estatísticos na abordagem clássica

Em teoria de probabilidades, conhecemos a medida de probabilidade, logo, fazemos descrições probabilísticas.

$$(\Omega, \mathcal{A}, P) \overset{\text{X}}{\to} (\mathbb{R}, \mathcal{B}, P_X)$$

Na prática, contudo, não conhecemos a medida P.

Definimos então uma família de medidas de probabilidades que possivelmente descrevem o comportamento aleatório dos dados.

O *Modelo Estatístico* é definido pela trinca

$$(\Omega, \mathcal{A}, \mathcal{P})$$

em que Ω é o espaço amostral (evento certo), $\mathcal{A}$ é a Sigma álgebra, uma família de subconjuntos ou eventos em Ω, e $\mathcal{P}$ é uma família de medidas de probabilidade que possivelmente descrevem o comportamento aleatório dos dados ou eventos sobre investigação.

2.1 Modelo Estatístico Paramétrico

Se $\mathcal{P} = \{P_\theta : \theta \in \Theta\}$, em que $\Theta \subseteq \mathbb{R}^P$ e $p \in \mathbb{N}$, então dizemos que $(\Omega, \mathcal{A}, \mathcal{P})$ é um modelo estatístico **Paramétrico**.

Caso não exista $\Theta \subseteq \mathbb{R}^P$ fixo, então dizemos que o modelo é não-paramétrico. *Obs*: Θ é o espaço paramétrico e θ é o vetor de parâmetros. θ **Não*** é variável aleatória, apenas indexa as medidas de probabilidade.

2.1.1 Exemplos

2.1.1.1 Exemplo de Bernoulli

Considere um Ensaio de Bernoulli

$$\Omega = \{S, F\}, \mathcal{A} = 2^\Omega$$

Temos algum conhecimento prévio que sugere que as probabilidades de sucesso podem ser $0.1, 0.5, 0.9$

Nesse caso,

$$\mathcal{P} = \{P_1, P_2, P_3\}$$

em que

$$\begin{cases} P_1(\{S\}) = 0.1;\ P_1(\{F\}) = 0.9;\ P_1(\Omega) = 1 \\ P_2(\{S\}) = 0.5;\ P_2(\{F\}) = 0.5;\ P_2(\Omega) = 1 \\ P_3(\{S\}) = 0.9;\ P_3(\{F\}) = 0.1;\ P_3(\Omega) = 1 \end{cases}$$

Observe que

$$\mathcal{P} = \{P_\theta : \theta \in \Theta\}$$

em que $\Theta = \{1, 2, 3\} \subseteq \mathbb{R}$ Portanto, $(\Omega, \mathcal{A}, \mathcal{P})$ é um modelo paramétrico.

2.1.1.2 Exemplo de Exponencial

Seja $X : \Omega \to \mathbb{R}$ tal que

$$X(\omega) = \begin{cases} 1, \omega = \mathrm{S} \\ 0, c.c. \end{cases}$$

$$E_\theta(X) = \sum_{x=0}^{1} x P_\theta(X = x)$$

$$\begin{cases} \theta = 1 \Rightarrow E_1(X) = 0.1 \\ \theta = 2 \Rightarrow E_2(X) = 0.5 \\ \theta = 3 \Rightarrow E_3(X) = 0.9 \end{cases}$$

Seja $\Omega = (0, \infty)$ e $\mathcal{A}$ uma sigma-álgebra de Ω (Sigma-Álgebra de Borel) Ω representa o tempo até a ocorrência de um evento (uma reclamação, por exemplo) Temos conhecimento prévio de que as funções densidade de probabilidade que possivelmente descrevem esse evento são:

$$f_1(\omega) = \begin{cases} \mathrm{e}^{-1\omega}, \omega > 0 \\ 0, c.c. \end{cases}$$

$$f_2(\omega) = \begin{cases} 2\mathrm{e}^{-2\omega}, \omega > 0 \\ 0, c.c. \end{cases}$$

$$f_3(\omega) = \begin{cases} \frac{1}{2}\mathrm{e}^{-\frac{1}{2}\omega}, \omega > 0 \\ 0, c.c. \end{cases}$$

$$f_4(\omega) = \begin{cases} \frac{1}{10}\mathrm{e}^{-\frac{1}{10}\omega}, \omega > 0 \\ 0, c.c. \end{cases}$$

e Ps,

$$P_1(A) = \int_A f_1(\omega) d\omega$$
$$P_2(A) = \int_A f_2(\omega) d\omega$$
$$P_3(A) = \int_A f_3(\omega) d\omega$$
$$P_4(A) = \int_A f_4(\omega) d\omega$$

Dessa forma,

$$P_\theta = \int_A f_\theta(\omega) d\omega, \omega \in \Omega$$

e $\Omega = \{1, 2, 3, 4\}$ Seja $X : \Omega \to \mathbb{R}$ tal que

$$X(\omega) = \omega$$

Note que

$$E_\theta(X) = \int_{-\infty}^{\infty} x f_\theta(x) dx, \theta \in \{1, 2, 3, 4\}$$

$$E_\theta(X) = \begin{cases} 1, \ \theta = 1 \\ \frac{1}{2}, \ \theta = 2 \\ 2, \ \theta = 3 \\ 10, \ \theta = 4 \end{cases}$$

2.2 Principais Modelos Estatísticos

2.2.1 Modelo Estatístico de Bernoulli

Dizemos que X é uma variável aleatória com modelo de Bernoulli se, e somente se,

$$P_\theta(X = x) = \begin{cases} \theta^x \cdot (1-\theta)^{1-x}, x \in \{0, 1\} \\ 0, x \notin \{0, 1\} \end{cases}$$

em que $\theta \in \Theta$ e $\Theta = (0, 1) \subseteq \mathbb{R}$ O parâmetro é a probabilidade de sucesso

$$\begin{cases} E_\theta(X) = \theta \\ \mathrm{Var}_\theta(X) = \theta(1-\theta) \\ P_\theta(X = 1) = \theta, P_\theta(X = 0) = 1 - \theta \end{cases}$$

2.2.2 Modelo Estatístico Binomial

Dizemos que X é uma variável aleatória com modelo binomial se, e somente se,

$$P_\theta(X = x) = \begin{cases} \binom{n}{x} \cdot \theta^x \cdot (1-\theta)^{n-x}, x \in \{0, 1, \dots, n\} \\ 0, x \notin \{0, 1, \dots, n\} \end{cases}$$

em que n é conhecido e fixado previamente, θ é a probabilidade de sucesso (parâmetro do modelo) e $\Theta = (0, 1)$ é o espaço paramétrico

$$\begin{cases} E_\theta(X) = n\theta \\ \mathrm{Var}_\theta(X) = n\theta(1-\theta) \\ P_\theta(X = 0) = (1-\theta)^n, \dots, P_\theta(X = n) = \theta^n \end{cases}$$

2.2.3 Modelo Estatístico Geométrico

Dizemos que X é uma variável aleatória com modelo estatístico geométrico se, e somente se,

$$P_\theta(X = x) = \begin{cases} \theta(1-\theta)^{x-1}, x \in \{1, \dots\} \\ 0, x \notin \{1, \dots\} \end{cases}$$

em que θ é o parâmetro do modelo (probabilidade de sucesso) e $\Theta = (0, 1)$ é o espaço paramétrico

$$\begin{cases} E_\theta(X) = \frac{1}{\theta} \\ \mathrm{Var}_\theta(X) = \frac{1-\theta}{\theta^2} \end{cases}$$

2.2.4 Modelo de Poisson

Dizemos que X é uma variável aleatória com modelo estatístico Poisson, se, e somente se,

$$P_\theta(X = x) = \begin{cases} \mathrm{e}^{-\theta} \cdot \frac{\theta^x}{x!}, x \in \{0, 1, \dots\} \\ 0, x \notin \{0, 1, \dots\} \end{cases}$$

em que θ é a taxa média de ocorrência do evento (parâmetro do modelo) e $\Theta = (0, \infty)$, o espaço paramétrico.

$$\begin{cases} E_\theta(X) = \theta \\ \mathrm{Var}_\theta(X) = \theta \end{cases}$$

2.2.5 Modelo Multinomial

Dizemos que X é um Vetor Aleatório com modelo estatístico Multinomial se, e somente se a função de probabilidade é

$$P_\theta(X_1 = x_1, \dots, X_k = x_k) = \begin{cases} \frac{n!}{x_1!x_2!\dots x_k!} \cdot \theta_1^{x_1} \cdot, \cdots, \cdot \theta_k^{x_n} & x_1 + \cdots + x_k = n \\ 0, c.c \end{cases}$$

em que $\theta_1 + \cdots + \theta_k = 1$ e $0 \leq \theta_i \leq 1$, $\forall i = 1, 2, \dots, k$, $\Theta = \{(\theta_1, , \dots, \theta_k) \in \mathbb{R}^k : 0 \leq \theta_i \leq 1, i = 1, \dots, k, \theta_1 + \cdots + \theta_k = 1\}$

$$\begin{cases} E_\theta(X_i) = n\theta_i \\ \mathrm{Var}_\theta(X_i) = n\theta i(1 - \theta_i) \\ \mathrm{Cov}(X_i X_j) = -n\theta_i\theta_j \end{cases}$$

Esse modelo tem aplicação em modelos de linguagem como o ChatGPT. (k como tamanho do vocabulário, $n = 1$, $\theta_1 =$ probabilidade de escolher o primeiro elemento do vocabulário e assim por diante.)

2.2.6 Modelo Uniforme contínuo

Dizemos que X é uma variável aleatória contínua com modelo estatístico Uniforme em $(\theta_1, \theta_2), \theta_2 > \theta_1$, se, e somente se, a sua Função Densidade de Probabilidade é

$$f_\theta(x) = \begin{cases} \frac{1}{\theta_2 - \theta_1}, x \in (\theta_1, \theta_2) \\ 0, c.c. \end{cases}$$

em que $\theta = (\theta_1, \theta_2)$ vetor, $\Theta = \{\theta \in \mathbb{R}^2 : \theta_2 > \theta_1\}$

$$\begin{cases} E_\theta(X) = \frac{b+a}{2} \\ \mathrm{Var}_\theta(X) = \frac{(b-a)^2}{12} \end{cases}$$

$$X \sim U(\theta_1, \theta_2), \theta = (\theta_1, \theta_2) \text{ Vetor de Parâmetros}$$

2.2.7 Modelo Exponencial

Dizemos que X é uma variável aleatória contínua com modelo estatístico Exponencial se, e somente se, a sua Função Densidade de Probabilidade é dada por

$$f_\theta(x) = \begin{cases} \theta \mathrm{e}^{-\theta x}, x > 0 \\ 0, c.c. \end{cases}$$

em que $\theta > 0, \Theta = \{\theta \in \mathbb{R} : \theta > 0\}$

$$\begin{cases} E_\theta(X) = \frac{1}{\theta} \\ \mathrm{Var}_\theta(X) = \frac{1}{\theta^2} \end{cases}$$

$$X \sim \mathrm{Exp}(\theta), \theta > 0$$

2.2.8 Modelo Normal

Dizemos que X é uma variável aleatória contínua com modelo estatístico Normal com média μ e variância σ^2 se, e somente se, a sua Função Densidade de Probabilidade é dada por

$$f_\theta(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \cdot \mathrm{e}^{-\frac{1}{2\sigma^2}(x-\mu)^2}, x \in \mathbb{R}$$

em que $\theta = (\mu, \sigma^2) \in \Theta = \{(\mu, \sigma^2) \in \mathbb{R} \times \mathbb{R}^+\}$

$$\begin{cases} E_\theta(X) = \mu \\ \mathrm{Var}_\theta(X) = \sigma^2 \end{cases}$$

$$X \sim N(\mu, \sigma^2), \theta = (\mu, \sigma^2) \text{ Vetor de Parâmetros}$$

3 População e Amostra

Veja: Modelo Estatístico para definições dos modelos estatísticos paramétricos.

3.1 Variável Populacional

Pela teoria estatística, população é o conjunto sob investigação de todos os potenciais elementos.

A *Variável Populacional* representa os valores numéricos de cada elemento da população:

$$X \sim f_{\theta,} \theta \in \Theta$$

em que f_θ é a Função Densidade de Probabilidade da Variável Aleatória populacional. θ é o vetor de parâmetros (desconhecido) e Θ é o espaço paramétrico

3.2 Amostra (Teórica)

É uma parte ou subconjunto da população.

3.2.1 Amostra Aleatória

Dizemos que $(X_1, \dots, X_n)$ é uma amostra aleatória de X (v.a. populacional) se $X_1, \dots, X_n$ forem independentes e identicamente distribuídas de acordo com a distribuição de X Ou seja,

$$\text{Independentes} = \begin{cases} X_1 \sim f_{\theta,} \theta \in \Theta \\ . \\ . \\ . \\ X_n \sim f_\theta, \theta \in \Theta \end{cases}$$

3.3 Amostra (Observada)

É formada por valores numéricos após utilizar um procedimento de amostragem.

$$x_1, \ldots, x_n$$

em que n é o tamanho amostral.

4 Quantidade de Interesse

É uma quantidade relacionada com a distribuição da variável aleatória populacional.

$$g(\theta)$$

Como $g(\theta) = E_\theta(X), g(\theta) = \mathrm{Var}_\theta(X), g(\theta) = P_\theta(X \geq 1)$ e até $g(\theta) = \theta$

5 Distribuição Amostral

Seja $(X_1, \dots, X_n)$ amostra aleatória (a.a.) de $X \sim f_{\theta,} \theta \in \Theta$

a Função Densidade de Probabilidade conjunta de $(X_1, X_2, \dots, X_n)$ é $\forall \theta \in \Theta$, no caso discreto:

$$P_\theta(X_1 = k_1, X_2 = k_2, \dots, X_n = k_n) \stackrel{\text{ind}}{=} \prod_{i=1}^{n} P_\theta(X_i = k_i) \stackrel{\text{id}}{=} \prod_{i=1}^{n} P_\theta(X = k_i) \stackrel{\text{def}}{=} \prod_{i=1}^{n} f_\theta(k_i)$$

no caso contínuo:

$$f_\theta^{(x)}(k_1, k_2, \dots, k_n) \stackrel{\text{i.i.d}}{=} \prod_{i=1}^{n} f_\theta(k_i)$$

5.1 Exemplos

5.1.1 Exemplo um

Seja (X_1, X_2, X_3) uma a.a de $X \sim \text{Ber}(\theta), \theta \in (0,1)$ 1. Especifique o espaço paramétrico 2. Calcule a função de probabilidade da amostra 3. Encontre as seguintes quantidade de interesses em função de θ 1. $g(\theta) = E_\theta(X)$ 2. $g(\theta) = P_\theta(X = 0)$ 3. $g(\theta) = \text{CV}_\theta(X)$ Resolução 1. $\Theta = (0,1)$ 2. Duas resoluções possíveis 1. Dando valores à amostra

(X_1, X_2, X_3)	$P(X_1 = k_1, X_2 = k_2, X_3 = k_3) = \prod_{i=1}^{3} P_\theta(X = k_1)$
$(0,0,0)$	$(1-\theta)^3$
$(0,0,1)$	$(1-\theta)^2\theta$
$(0,1,0)$	$(1-\theta)^2\theta$
$(1,0,0)$	$(1-\theta)^2\theta$
$(0,1,1)$	$(1-\theta)\theta^2$
$(1,0,1)$	$(1-\theta)\theta^2$
$(1,1,0)$	$(1-\theta)\theta^2$
$(1,1,1)$	θ^3

2. Enunciando a função Observe que, se $k \in \{0,1\}$,

$$P_\theta(X = k) = \theta^k (1-\theta)^{1-k} \cdot \mathbb{1}_{\{0,1\}}(k) \Rightarrow$$

$$P_\theta(X_1 = k, X_2 = k_2, X_3 = k_3) \overset{\text{i.i.d}}{=} \prod_{i=1}^{3} \{\theta^{k_i}(1-\theta)^{1-k_1}\mathbb{1}_{\{0,1\}}(k_1)\} =$$

$$= \theta^{\sum_{i=1}^{3} k_1}(1-\theta)^{3-\sum_{i=1}^{3} k_1} \prod_{i=1}^{3} \mathbb{1}_{\{0,1\}}(k_i)$$

3. Em função de θ: 1. $g(\theta) = E_\theta(X) = \theta$ 2. $g(\theta) = P_\theta(X = 0) = 1 - \theta$ 3. $g(\theta) = \mathrm{CV}_\theta(X) = \frac{\sqrt{\theta(1-\theta)}}{\theta}$

5.1.2 Exemplo dois

Seja $(X_1, \dots, X_n)$ uma a.a de $X \sim \mathrm{Ber}(\theta), \theta \in (0,1)$, encontre a f.p conjunta da amostra.

$$P_\theta(X_1 = k_1, \dots, X_n = k_n) \overset{\text{iid}}{=} \prod_{i=1}^{n} P_\theta(X = k_i) = \prod_{i=1}^{n} \{\theta^{k_i}(1-\theta)^{1-k_i}\mathbb{1}_{\{0,1\}}(k_i)\}$$

$$\Rightarrow P_\theta(X_1 = k_1, \dots, X_n = k_n) = \theta^{\sum_{i=1}^{n} k_1} \cdot \theta^{n-\sum_{i=i}^{n} k_i} \cdot \mathbb{1}_{\{0,1\}}(k_i)$$

5.1.3 Exemplo três

Seja $(X_1, \dots, X_n)$ uma a.a de $X \sim \mathrm{Pois}(\theta), \theta \in (0, \infty)$, encontre a f.p. conjunta da amostra. Como esse vetor é uma a.a. (ou seja, variáveis aleatórias independentes e identicamente distribuídas), temos que

$$P_\theta(X_1 = k_1, \dots, X_n = k_n) \overset{\text{iid}}{=} \prod_{i=1}^{n} P_\theta(X = k_i) = \prod_{i=1}^{n} \{\mathrm{e}^{-\theta} \cdot \frac{\theta^{k_i}}{k_i!}\}$$

Sempre que $k_i \in \{0, 1, \dots\}, \forall i = 1, \dots, n$

$$\Rightarrow P_\theta(\dots) = \mathrm{e}^{-n\theta} \cdot \frac{\theta^{\sum_{i=1}^{n} k_i}}{\prod_{i=1}^{n}(k_i)!}$$

5.1.4 Exemplo contínuo um

Seja $(X_1, \dots, X_n)$ uma a.a de $X \sim \text{Exp}(\theta), \theta \in (0, \infty)$, encontre a função densidade de probabilidade (f.d.p.) conjunta da amostra.

$$F_\theta^{(*)}(k_1, \dots, k_n) \overset{\text{iid}}{=} \prod_{i=1}^{n} f_\theta(k_i) = \prod_{i=1}^{n} \{\theta \mathrm{e}^{-\theta k_i} \cdot \mathbb{1}_{(0,\infty)}(k_i)\}$$

$$\Rightarrow f_\theta^{(*)}(\dots) = \theta^n \cdot \mathrm{e}^{-\theta \sum\limits_{i=1}^{n} k_i} \cdot \prod_{i=1}^{n} \mathbb{1}_{(0,\infty)}(k)$$

5.1.5 Exemplo contínuo dois

Seja $(X_1, \dots, X_n)$ uma a.a. (i.i.d) de $X \sim N(\mu, \sigma^2)$ em que $\theta = (\mu, \sigma^2) \in \Theta = \mathbb{R} \times \mathbb{R}_+$. Considere $\underset{\sim}{x} = (x_1, \dots, x_n)$ a amostra observada.

$$L_{\underset{\sim}{X}} \overset{\text{iid}}{=} \prod_{i=1}^{n} f_\theta(x_i) = \prod_{i=1}^{n} \left\{ \frac{1}{\sqrt{2\pi\sigma^2}} \exp\{-\frac{1}{2\sigma^2}(x_1 - \mu)^2\} \right\}$$

$$= \frac{1}{(2\pi\sigma^2)^{\frac{x}{2}}} \cdot \exp\{-\frac{1}{2\sigma^2} \sum_{i=1}^{n} (x_i - \mu)^2\}$$