Inferência Estatística

Introdução

E.F.T[1]

[1]EACH-USP
Universidade de São Paulo

ACH2053

Outline

Outline

Definição de a.a.

As varíaveis aleatórias $X_1, ..., X_n$ são chamada de *amostra aleatória* de tamanho n da população com distribuição $f(x|\theta)$ se:

- as variáveis $X_1, ..., X_n$ são mutuamente independentes e
- a f.d.p. ou f.p. de cada marginal X_i for a mesma $(f(x_i|\theta))$ para cada $i = 1, ..., n$.

A partir da definição acima, a f.d.p. conjunta ou a f.p.conjunta de $X_1, ..., X_n$ é dada por

$$f(X_1, ..., X_n|\theta) = f(\mathbf{X}|\theta) = f(X_1|\theta)...f(X_n|\theta) = \prod_{i=1}^{n} f(X_i|\theta) \qquad (1)$$

Definição de a.a.

As varíaveis aleatórias $X_1, ..., X_n$ são chamada de *amostra aleatória* de tamanho n da população com distribuição $f(x|\theta)$ se:

- as variáveis $X_1, ..., X_n$ são mutuamente independentes e
- a f.d.p. ou f.p. de cada marginal X_i for a mesma ($f(x_i|\theta)$) para cada $i = 1, ..., n$.

A partir da definição acima, a f.d.p. conjunta ou a f.p.conjunta de $X_1, ..., X_n$ é dada por

$$f(X_1, ..., X_n|\theta) = f(\mathbf{X}|\theta) = f(X_1|\theta)...f(X_n|\theta) = \prod_{i=1}^{n} f(X_i|\theta) \qquad (1)$$

Outline

Amostragem de uma Bernoulli

Suponha que as v.a. $X_1, ..., X_n$ formam uma a.a. de uma distribuição de Bernoulli com parâmetro desconhecido θ $(0 \leq \theta \leq 1)$. Para quaiquer valores observados $x_1, ..., x_n$ onde a f.p. de cada x_i é:

$$f(x_i|\theta) = \theta^{x_i}(1-\theta)^{1-x_i}$$

onde x_i só pode ser 0 ou 1.
Desta forma, a f.p.conjunta de $x_1, ..., x_n$ será:

$$f(x_1, ..., x_n|\theta) = f_n(\mathbf{x}|\theta) = \prod_{i=1}^{n} f(x_i|\theta) = \prod_{i=1}^{n} \theta^{x_i}(1-\theta)^{1-x_i}$$

Se denotarmos por $y = \sum_{i=1}^{n} x_i$ teremos:

$$f_n(\mathbf{x}|\theta) = \prod_{i=1}^{n} f(x_i|\theta) = \theta^{\sum_{i=1}^{n} x_i}(1-\theta)^{n-\sum_{i=1}^{n} x_i} = \theta^y(1-\theta)^{n-y}$$

Amostragem de uma Bernoulli

Suponha que as v.a. $X_1, ..., X_n$ formam uma a.a. de uma distribuição de Bernoulli com parâmetro desconhecido θ $(0 \leq \theta \leq 1)$. Para quaiquer valores observados $x_1, ..., x_n$ onde a f.p. de cada x_i é:

$$f(x_i|\theta) = \theta^{x_i}(1-\theta)^{1-x_i}$$

onde x_i só pode ser 0 ou 1.
Desta forma, a f.p.conjunta de $x_1, ..., x_n$ será:

$$f(x_1, ..., x_n|\theta) = f_n(\mathbf{x}|\theta) = \prod_{i=1}^{n} f(x_i|\theta) = \prod_{i=1}^{n} \theta^{x_i}(1-\theta)^{1-x_i}$$

Se denotarmos por $y = \sum_{i=1}^{n} x_i$ teremos:

$$f_n(\mathbf{x}|\theta) = \prod_{i=1}^{n} f(x_i|\theta) = \theta^{\sum_{i=1}^{n} x_i}(1-\theta)^{n-\sum_{i=1}^{n} x_i} = \theta^{y}(1-\theta)^{n-y}$$

Amostragem de uma Bernoulli

Suponha que as v.a. $X_1, ..., X_n$ formam uma a.a. de uma distribuição de Bernoulli com parâmetro desconhecido θ $(0 \leq \theta \leq 1)$. Para quaiquer valores observados $x_1, ..., x_n$ onde a f.p. de cada x_i é:

$$f(x_i|\theta) = \theta^{x_i}(1-\theta)^{1-x_i}$$

onde x_i só pode ser 0 ou 1.
Desta forma, a f.p.conjunta de $x_1, ..., x_n$ será:

$$f(x_1, ..., x_n|\theta) = f_n(\mathbf{x}|\theta) = \prod_{i=1}^{n} f(x_i|\theta) = \prod_{i=1}^{n} \theta^{x_i}(1-\theta)^{1-x_i}$$

Se denotarmos por $y = \sum_{i=1}^{n} x_i$ teremos:

$$f_n(\mathbf{x}|\theta) = \prod_{i=1}^{n} f(x_i|\theta) = \theta^{\sum_{i=1}^{n} x_i}(1-\theta)^{n-\sum_{i=1}^{n} x_i} = \theta^{y}(1-\theta)^{n-y}$$

Amostragem de uma Uniforme

Suponha que as v.a. $X_1, ..., X_n$ formam uma a.a. de uma distribuição Uniforme no intervalo $(0, \theta)$ com parâmetro desconhecido θ $(\theta > 0)$. Para quaiquer valores observados $x_1, ..., x_n$ onde a f.d.p. de cada x_i é:

$$f(x_i|\theta) = \begin{cases} \frac{1}{\theta} & \text{para} \quad 0 \leq x_i \leq \theta \\ 0 & \text{c.c} \end{cases}$$

Desta forma, a f.d.p.conjunta de $x_1, ..., x_n$ será:

$$f(x_1, ..., x_n|\theta) = f_n(\mathbf{x}|\theta) = \prod_{i=1}^{n} f(x_i|\theta) = \begin{cases} \frac{1}{\theta^n} & \text{para} \quad 0 \leq x_i \leq \theta \\ 0 & \text{c.c} \end{cases}$$

Amostragem de uma Uniforme

Suponha que as v.a. $X_1, ..., X_n$ formam uma a.a. de uma distribuição Uniforme no intervalo $(0, \theta)$ com parâmetro desconhecido θ ($\theta > 0$). Para quaiquer valores observados $x_1, ..., x_n$ onde a f.d.p. de cada x_i é:

$$f(x_i|\theta) = \begin{cases} \frac{1}{\theta} & \text{para} \quad 0 \leq x_i \leq \theta \\ 0 & \text{c.c} \end{cases}$$

Desta forma, a f.d.p.conjunta de $x_1, ..., x_n$ será:

$$f(x_1, ..., x_n|\theta) = f_n(\mathbf{x}|\theta) = \prod_{i=1}^{n} f(x_i|\theta) = \begin{cases} \frac{1}{\theta^n} & \text{para} \quad 0 \leq x_i \leq \theta \\ 0 & \text{c.c} \end{cases}$$

Amostragem de uma Uniforme

Suponha que as v.a. $X_1, ..., X_n$ formam uma a.a. de uma distribuição Uniforme no intervalo $(0, \theta)$ com parâmetro desconhecido θ $(\theta > 0)$. Para quaiquer valores observados $x_1, ..., x_n$ onde a f.d.p. de cada x_i é:

$$f(x_i|\theta) = \begin{cases} \frac{1}{\theta} & \text{para} \quad 0 \leq x_i \leq \theta \\ 0 & \text{c.c} \end{cases}$$

Desta forma, a f.d.p.conjunta de $x_1, ..., x_n$ será:

$$f(x_1, ..., x_n|\theta) = f_n(\mathbf{x}|\theta) = \prod_{i=1}^{n} f(x_i|\theta) = \begin{cases} \frac{1}{\theta^n} & \text{para} \quad 0 \leq x_i \leq \theta \\ 0 & \text{c.c} \end{cases}$$

Amostragem de uma Uniforme

Suponha que as v.a. $X_1, ..., X_n$ formam uma a.a. de uma distribuição Uniforme no intervalo $(0, \theta)$ com parâmetro desconhecido θ $(\theta > 0)$. Para quaiquer valores observados $x_1, ..., x_n$ onde a f.d.p. de cada x_i é:

$$f(x_i|\theta) = \begin{cases} \frac{1}{\theta} & \text{para} \quad 0 \leq x_i \leq \theta \\ 0 & \text{c.c} \end{cases}$$

Desta forma, a f.d.p.conjunta de $x_1, ..., x_n$ será:

$$f(x_1, ..., x_n|\theta) = f_n(\mathbf{x}|\theta) = \prod_{i=1}^{n} f(x_i|\theta) = \begin{cases} \frac{1}{\theta^n} & \text{para} \quad 0 \leq x_i \leq \theta \\ 0 & \text{c.c} \end{cases}$$

Amostragem de uma Poisson

Suponha que as v.a. $X_1, ..., X_n$ formam uma a.a. de uma distribuição de Poisson com parâmetro desconhecido θ ($\theta > 0$). Para quaiquer valores observados $x_1, ..., x_n$ onde a f.p. de cada x_i é:

$$f(x_i|\theta) = \frac{e^{-\theta}\theta^{x_i}}{x_i!}$$

onde x_i pode assumir valores inteiros

Desta forma, a f.p.conjunta de $x_1, ..., x_n$ será:

$$f(x_1, ..., x_n|\theta) = f_n(\mathbf{x}|\theta) = \prod_{i=1}^{n} f(x_i|\theta) = \prod_{i=1}^{n} \frac{e^{-\theta}\theta^{x_i}}{x_i!}$$

Se denotarmos por $y = \sum_{i=1}^{n} x_i$ teremos:

$$f_n(\mathbf{x}|\theta) = \prod_{i=1}^{n} f(x_i|\theta) = (\prod_{i=1}^{n} \frac{1}{x_i!})e^{-n\theta}\theta^{\sum_{i=1}^{n} x_i} = (\prod_{i=1}^{n} \frac{1}{x_i!})e^{-n\theta}\theta^{y}$$

Amostragem de uma Poisson

Suponha que as v.a. $X_1, ..., X_n$ formam uma a.a. de uma distribuição de Poisson com parâmetro desconhecido θ ($\theta > 0$). Para quaiquer valores observados $x_1, ..., x_n$ onde a f.p. de cada x_i é:

$$f(x_i|\theta) = \frac{e^{-\theta}\theta^{x_i}}{x_i!}$$

onde x_i pode assumir valores inteiros
Desta forma, a f.p.conjunta de $x_1, ..., x_n$ será:

$$f(x_1, ..., x_n|\theta) = f_n(\mathbf{x}|\theta) = \prod_{i=1}^{n} f(x_i|\theta) = \prod_{i=1}^{n} \frac{e^{-\theta}\theta^{x_i}}{x_i!}$$

Se denotarmos por $y = \sum_{i=1}^{n} x_i$ teremos:

$$f_n(\mathbf{x}|\theta) = \prod_{i=1}^{n} f(x_i|\theta) = (\prod_{i=1}^{n} \frac{1}{x_i!})e^{-n\theta}\theta^{\sum_{i=1}^{n} x_i} = (\prod_{i=1}^{n} \frac{1}{x_i!})e^{-n\theta}\theta^{y}$$

Amostragem de uma Poisson

Suponha que as v.a. $X_1, ..., X_n$ formam uma a.a. de uma distribuição de Poisson com parâmetro desconhecido θ ($\theta > 0$). Para quaiquer valores observados $x_1, ..., x_n$ onde a f.p. de cada x_i é:

$$f(x_i|\theta) = \frac{e^{-\theta}\theta^{x_i}}{x_i!}$$

onde x_i pode assumir valores inteiros
Desta forma, a f.p.conjunta de $x_1, ..., x_n$ será:

$$f(x_1, ..., x_n|\theta) = f_n(\mathbf{x}|\theta) = \prod_{i=1}^{n} f(x_i|\theta) = \prod_{i=1}^{n} \frac{e^{-\theta}\theta^{x_i}}{x_i!}$$

Se denotarmos por $y = \sum_{i=1}^{n} x_i$ teremos:

$$f_n(\mathbf{x}|\theta) = \prod_{i=1}^{n} f(x_i|\theta) = (\prod_{i=1}^{n} \frac{1}{x_i!})e^{-n\theta}\theta^{\sum_{i=1}^{n} x_i} = (\prod_{i=1}^{n} \frac{1}{x_i!})e^{-n\theta}\theta^{y}$$

Amostragem de uma Exponencial

Suponha que as v.a. $X_1, ..., X_n$ formam uma a.a. de uma distribuição Exponencial com parâmetro desconhecido θ ($\theta > 0$). Para quaiquer valores observados $x_1, ..., x_n$ onde a f.d.p. de cada x_i é:

$$f(x_i|\theta) = \theta e^{-\theta x_i}$$

Desta forma, a f.d.p.conjunta de $x_1, ..., x_n$ será:

$$f(x_1, ..., x_n|\theta) = f_n(\mathbf{x}|\theta) = \prod_{i=1}^{n} f(x_i|\theta) = \prod_{i=1}^{n} \theta e^{-\theta x_i}$$

Se denotarmos por $y = \sum_{i=1}^{n} x_i$ teremos:

$$f_n(\mathbf{x}|\theta) = \prod_{i=1}^{n} f(x_i|\theta) = \theta^n e^{-\theta \sum_{i=1}^{n} x_i} = \theta^n e^{-\theta y}$$

Amostragem de uma Exponencial

Suponha que as v.a. $X_1, ..., X_n$ formam uma a.a. de uma distribuição Exponencial com parâmetro desconhecido θ ($\theta > 0$). Para quaiquer valores observados $x_1, ..., x_n$ onde a f.d.p. de cada x_i é:

$$f(x_i|\theta) = \theta e^{-\theta x_i}$$

Desta forma, a f.d.p.conjunta de $x_1, ..., x_n$ será:

$$f(x_1, ..., x_n|\theta) = f_n(\mathbf{x}|\theta) = \prod_{i=1}^{n} f(x_i|\theta) = \prod_{i=1}^{n} \theta e^{-\theta x_i}$$

Se denotarmos por $y = \sum_{i=1}^{n} x_i$ teremos:

$$f_n(\mathbf{x}|\theta) = \prod_{i=1}^{n} f(x_i|\theta) = \theta^n e^{-\theta \sum_{i=1}^{n} x_i} = \theta^n e^{-\theta y}$$

Amostragem de uma Exponencial

Suponha que as v.a. $X_1, ..., X_n$ formam uma a.a. de uma distribuição Exponencial com parâmetro desconhecido θ ($\theta > 0$). Para quaiquer valores observados $x_1, ..., x_n$ onde a f.d.p. de cada x_i é:

$$f(x_i|\theta) = \theta e^{-\theta x_i}$$

Desta forma, a f.d.p.conjunta de $x_1, ..., x_n$ será:

$$f(x_1, ..., x_n|\theta) = f_n(\mathbf{x}|\theta) = \prod_{i=1}^{n} f(x_i|\theta) = \prod_{i=1}^{n} \theta e^{-\theta x_i}$$

Se denotarmos por $y = \sum_{i=1}^{n} x_i$ teremos:

$$f_n(\mathbf{x}|\theta) = \prod_{i=1}^{n} f(x_i|\theta) = \theta^n e^{-\theta \sum_{i=1}^{n} x_i} = \theta^n e^{-\theta y}$$

Amostragem de uma Normal

Suponha que as v.a. $X_1, ..., X_n$ formam uma a.a. de uma distribuição Normal com média desconhecido θ $(-\infty < \theta < \infty)$ e variância σ^2 conhecida. Para quaiquer valores observados $x_1, ..., x_n$ onde a f.d.p. de cada x_i é:

$$f(x_i|\theta) = \frac{1}{(2\pi)^{1/2}\sigma} e^{[-\frac{(x_i-\theta)^2}{2\sigma^2}]}$$

para $-\infty < x_i < \infty$.
Desta forma, a f.d.p.conjunta de $x_1, ..., x_n$ será:

$$f(x_1, ..., x_n|\theta) = f_n(\mathbf{x}|\theta) = \prod_{i=1}^{n} \frac{1}{(2\pi)^{1/2}\sigma} e^{[-\frac{(x_i-\theta)^2}{2\sigma^2}]} \quad (2)$$

Amostragem de uma Normal

Suponha que as v.a. $X_1, ..., X_n$ formam uma a.a. de uma distribuição Normal com média desconhecido θ $(-\infty < \theta < \infty)$ e variância σ^2 conhecida. Para quaiquer valores observados $x_1, ..., x_n$ onde a f.d.p. de cada x_i é:

$$f(x_i|\theta) = \frac{1}{(2\pi)^{1/2}\sigma} e^{[-\frac{(x_i-\theta)^2}{2\sigma^2}]}$$

para $-\infty < x_i < \infty$.
Desta forma, a f.d.p.conjunta de $x_1, ..., x_n$ será:

$$f(x_1, ..., x_n|\theta) = f_n(\mathbf{x}|\theta) = \prod_{i=1}^{n} \frac{1}{(2\pi)^{1/2}\sigma} e^{[-\frac{(x_i-\theta)^2}{2\sigma^2}]} \qquad (2)$$

Amostragem de uma Normal

Suponha que as v.a. $X_1, ..., X_n$ formam uma a.a. de uma distribuição Normal com média desconhecido θ $(-\infty < \theta < \infty)$ e variância σ^2 conhecida. Para quaiquer valores observados $x_1, ..., x_n$ onde a f.d.p. de cada x_i é:

$$f(x_i|\theta) = \frac{1}{(2\pi)^{1/2}\sigma} e^{[-\frac{(x_i-\theta)^2}{2\sigma^2}]}$$

para $-\infty < x_i < \infty$.
Desta forma, a f.d.p.conjunta de $x_1, ..., x_n$ será:

$$f(x_1, ..., x_n|\theta) = f_n(\mathbf{x}|\theta) = \prod_{i=1}^{n} \frac{1}{(2\pi)^{1/2}\sigma} e^{[-\frac{(x_i-\theta)^2}{2\sigma^2}]} \tag{2}$$

Amostragem de uma Gamma

Suponha que as v.a. $X_1, ..., X_n$ formam uma a.a. de uma distribuição Gamma com parâmetros α e β com ($\alpha > 0$ e $\beta > 0$). Para quaiquer valores observados $x_1, ..., x_n$ onde a f.d.p. de cada x_i é:

$$f(x_i|\alpha,\beta) = \begin{cases} \frac{\beta^\alpha}{\Gamma(\alpha)} x_i^{\alpha-1} e^{-\beta x_i} & \text{para} \quad x > 0 \\ 0 & \text{para} \quad x \leq 0 \end{cases} \tag{3}$$

Sabe-se que para esta distribuição, o valor esperado e a variância para qualquer variável X_i é

$$E[X] = \frac{\alpha}{\beta} \tag{4}$$

$$Var(X) = \frac{\alpha}{\beta^2} \tag{5}$$

Amostragem de uma Gamma

Suponha que as v.a. $X_1, ..., X_n$ formam uma a.a. de uma distribuição Gamma com parâmetros α e β com ($\alpha > 0$ e $\beta > 0$). Para quaiquer valores observados $x_1, ..., x_n$ onde a f.d.p. de cada x_i é:

$$f(x_i|\alpha, \beta) = \begin{cases} \frac{\beta^\alpha}{\Gamma(\alpha)} x_i^{\alpha-1} e^{-\beta x_i} & \text{para} \quad x > 0 \\ 0 & \text{para} \quad x \leq 0 \end{cases} \tag{3}$$

Sabe-se que para esta distribuição, o valor esperado e a variância para qualquer variável X_i é

$$E[X] = \frac{\alpha}{\beta} \tag{4}$$

$$Var(X) = \frac{\alpha}{\beta^2} \tag{5}$$

Amostragem de uma Gamma

Desta forma, a f.d.p.conjunta de $x_1, ..., x_n$ será:

$$\begin{aligned} f(x_1, ..., x_n|\theta) &= f_n(\mathbf{x}|\theta) \\ &= \prod_{i=1}^{n} \frac{\beta^{\alpha}}{\Gamma(\alpha)} x_i^{\alpha-1} e^{-\beta x_i} \\ &= \frac{\beta^{n\alpha}}{\Gamma^{n}(\alpha)} \left(\prod_{i=1}^{n} x_i\right)^{\alpha-1} e^{-\beta \sum_{i=1}^{n} x_i} \end{aligned} \tag{6}$$

Amostragem de uma Gamma

Desta forma, a f.d.p.conjunta de $x_1, ..., x_n$ será:

$$\begin{aligned} f(x_1, ..., x_n|\theta) &= f_n(\mathbf{x}|\theta) \\ &= \prod_{i=1}^{n} \frac{\beta^\alpha}{\Gamma(\alpha)} x_i^{\alpha-1} e^{-\beta x_i} \\ &= \frac{\beta^{n\alpha}}{\Gamma^n(\alpha)} \left(\prod_{i=1}^{n} x_i\right)^{\alpha-1} e^{-\beta \sum_{i=1}^{n} x_i} \end{aligned} \tag{6}$$

Amostragem de uma Gamma

Desta forma, a f.d.p.conjunta de $x_1, ..., x_n$ será:

$$\begin{aligned} f(x_1, ..., x_n|\theta) &= f_n(\mathbf{x}|\theta) \\ &= \prod_{i=1}^{n} \frac{\beta^\alpha}{\Gamma(\alpha)} x_i^{\alpha-1} e^{-\beta x_i} \\ &= \frac{\beta^{n\alpha}}{\Gamma^n(\alpha)} (\prod_{i=1}^{n} x_i)^{\alpha-1} e^{-\beta \sum_{i=1}^{n} x_i} \qquad (6) \end{aligned}$$

Amostragem de uma Beta

Suponha que as v.a. $X_1, ..., X_n$ formam uma a.a. de uma distribuição Beta com parâmetros α e β com ($\alpha > 0$ e $\beta > 0$). Para quaiquer valores observados $x_1, ..., x_n$ a f.d.p. de cada x_i é:

$$f(x_i|\alpha,\beta) = \begin{cases} \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x_i^{\alpha-1}(1-x_i)^{\beta-1} & \text{para} \quad 0 < x < 1 \\ 0 & \text{c.c} \end{cases} \tag{7}$$

Sabe-se que para esta distribuição, o valor esperado e a variância para qualquer variável X_i é

$$E[X] = \frac{\alpha}{\alpha+\beta} \tag{8}$$

$$Var(X) = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)} \tag{9}$$

Amostragem de uma Beta

Suponha que as v.a. $X_1, ..., X_n$ formam uma a.a. de uma distribuição Beta com parâmetros α e β com ($\alpha > 0$ e $\beta > 0$). Para quaiquer valores observados $x_1, ..., x_n$ a f.d.p. de cada x_i é:

$$f(x_i|\alpha, \beta) = \begin{cases} \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x_i^{\alpha-1}(1-x_i)^{\beta-1} & \text{para} \quad 0 < x < 1 \\ 0 & \text{c.c} \end{cases} \tag{7}$$

Sabe-se que para esta distribuição, o valor esperado e a variância para qualquer variável X_i é

$$E[X] = \frac{\alpha}{\alpha+\beta} \tag{8}$$

$$Var(X) = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)} \tag{9}$$

Outline

Definição de Estatística

As v.a. $X_1, ..., X_n$ formam uma a.a. Sabe-se que a sua distribuição conjunta $f_n(\mathbf{x}|\theta)$ tem a seguinte forma para algum valor particular de $\theta \in \Omega$:

$$f_n(\mathbf{x}|\theta) = f(x_1|\theta)...f(x_n|\theta) \tag{10}$$

Em outras palavras, sabe-se que a conjunta de $X_1, ..., X_n$ é membro da família que contém todas as f.d.p do tipo $f_n(\mathbf{x}|\theta)$ para todos os possíveis valores de $\theta \in \Omega$. O problema de estimar o valor de θ pode ser visto como a seleção por inferência de uma particular distribuição nesta família que gerou as observações $X_1, ..., X_n$.

Definição de Estatística

Qualquer função a valores reais $T = r(X_1, ..., X_n)$ das observações na a.a. é chamada de *estatística*.

As estatísticas mais comuns são:

$$\begin{aligned}
\bar{X}_n &= \frac{1}{n}\sum_{i=1}^{n} X_i : \quad \text{Média da amostra} \\
S^2 &= \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X}_n)^2 : \quad \text{Variância da amostra} \\
X_{(1)} &= min(X_1, ..., X_n) : \quad \text{o menor valor da amostra} \\
X_{(n)} &= max(X_1, ..., X_n) : \quad \text{o maior valor da amostra} \\
X_{(i)} &= \text{i-ésima maior observação da amostra}
\end{aligned}$$

Definição de Estatística

Qualquer função a valores reais $T = r(X_1, ..., X_n)$ das observações na a.a. é chamada de *estatística*.
As estatísticas mais comuns são:

$$
\begin{aligned}
\bar{X}_n &= \frac{1}{n}\sum_{i=1}^{n} X_i : \quad \text{Média da amostra} \\
S^2 &= \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X}_n)^2 : \quad \text{Variância da amostra} \\
X_{(1)} &= min(X_1, ..., X_n) : \quad \text{o menor valor da amostra} \\
X_{(n)} &= max(X_1, ..., X_n) : \quad \text{o maior valor da amostra} \\
X_{(i)} &= \text{i-ésima maior observação da amostra}
\end{aligned}
$$

Definição de Estatística

Para qualquer valor fixado de $\theta \in \Omega$, a distribuição de qualquer estatística T pode ser derivada da conjunta de $X_1, ..., X_n$. Em geral, esta distribuição dependerá de θ, desta maneira, haverá uma família de distribuições para T correspondentes aos diferentes possíveis valores de $\theta \in \Omega$.

Definição de Estatística

Para qualquer valor fixado de $\theta \in \Omega$, a distribuição de qualquer estatística T pode ser derivada da conjunta de $X_1, ..., X_n$. Em geral, esta distribuição dependerá de θ, desta maneira, haverá uma família de distribuições para T correspondentes aos diferentes possíveis valores de $\theta \in \Omega$.

Definição de Estatística Suficiente

Se dois indivíduos A e B devem estimar o valor de um parâmetro θ, A pode observar os valores de $X_1, ..., X_n$ e B não pode observar os mesmos, mas apenas o valor de certa estatística $T = r(X_1, ..., X_n)$.

Em geral A será capaz de achar melhores estimadores do que B pois pode escolher qualquer função das observações $X_1, ..., X_n$ como estimador de θ.

Definição de Estatística Suficiente

Se dois indivíduos A e B devem estimar o valor de um parâmetro θ, A pode observar os valores de $X_1, ..., X_n$ e B não pode observar os mesmos, mas apenas o valor de certa estatística $T = r(X_1, ..., X_n)$.
Em geral A será capaz de achar melhores estimadores do que B pois pode escolher qualquer função das observações $X_1, ..., X_n$ como estimador de θ.

Definição de Estatística Suficiente

Se dois indivíduos A e B devem estimar o valor de um parâmetro θ, A pode observar os valores de $X_1, ..., X_n$ e B não pode observar os mesmos, mas apenas o valor de certa estatística $T = r(X_1, ..., X_n)$.
Em geral A será capaz de achar melhores estimadores do que B pois pode escolher qualquer função das observações $X_1, ..., X_n$ como estimador de θ.

Definição de Estatística Suficiente

No entanto, algumas vezes B será capaz de estimar como A. Em esse problema, a função $T = r(X_1, ..., X_n)$ em algum sentido, sumariza toda a informação contida na a.a. tal que o conhecimento dos valores individuais de $X_1, ..., X_n$ será irrelevante na busca de um bom estimador de θ. A estatística T tendo esta propriedade é chamada de *estatística suficiente*.

Definição de Estatística Suficiente

Formalmente, se T é uma estatística e t é um valor particular de T, então a distribuição condicional de $X_1, ..., X_n$ dado que $T = t$, pode ser calculada da equação 10 que em geral depende de θ. Desta forma, para cada valor de t existe uma família de possíveis distribuições condicionais correspondentes aos diferentes valores possíveis de $\theta \in \Omega$. Poderia acontecer no entanto, que para cada valor de t, a distribuição condicional de $X_1, ..., X_n$ dado que $T = t$ é a mesma para todos os valores de θ. Neste caso, é dito que T é uma *estatística suficiente para o parâmetro* θ.

Definição de Estatística Suficiente

Uma estatística suficiente T é considerada uma sumarização de toda a informação relevante sobre θ contida na amostra $X_1, ..., X_n$. Se o indivíduo B pode observar T (e T é uma estatística suficiente) e não os valores individuais de $X_1, ..., X_n$, então, a distribuição condicional de $X_1, ..., X_n$ dado que $T = t$ é completamente conhecida para qualquer valor observado t e não depende do valor desconhecido θ. Portanto, para qualquer valor de t que poderia ser observado, o indivíduo B pode gerar n variáveis aleatórias $X_1^{'}, ..., X_n^{'}$ de acordo com esta distribuição condicional conjunta. Este processo de gerar as variáveis $X_1^{'}, ..., X_n^{'}$ é chamado de *aleatorização auxiliar*.

Definição de Estatística Suficiente

Quando é observado T e logo gerado $X_1', ..., X_n'$ de acordo com uma especificada distribuição condicional conjunta, segue que para qualquer valor de $\theta \in \Omega$, a marginal conjunta de $X_1', ..., X_n'$ será a mesma da conjunta $X_1, ..., X_n$.

A diferença entre uma estatística suficiente de uma não suficiente pode ser explicado da seguinte forma: A *aleatorização auxiliar* usada para gerar as variáveis $X_1', ..., X_n'$ após a observação da estatística suficiente T não requer conhecimento sobre o valor de θ. Se T não for suficiente, a *aleatorização auxiliar* poderia não acontecer , pois a distribuição condicional conjunta de $X_1, ..., X_n$ para um dado valor de T envolve o valor de θ e este valor é desconhecido.

Definição de Estatística Suficiente

Quando é observado T e logo gerado $X_1^{'}, ..., X_n^{'}$ de acordo com uma especificada distribuição condicional conjunta, segue que para qualquer valor de $\theta \in \Omega$, a marginal conjunta de $X_1^{'}, ..., X_n^{'}$ será a mesma da conjunta $X_1, ..., X_n$.
A diferença entre uma estatística suficiente de uma não suficiente pode ser explicado da seguinte forma: A *aleatorização auxiliar* usada para gerar as variáveis $X_1^{'}, ..., X_n^{'}$ após a observação da estatística suficiente T não requer conhecimento sobre o valor de θ. Se T não for suficiente, a *aleatorização auxiliar* poderia não acontecer , pois a distribuição condicional conjunta de $X_1, ..., X_n$ para um dado valor de T envolve o valor de θ e este valor é desconhecido.

Definição de Estatística Suficiente

Suponha que o indivíduo A quem observa os valores de $X_1, ..., X_n$ planeja usar um estimador particular $\delta(X_1, ..., X_n)$ para estimar θ, e B observa o valor de T e gera $X_1^{'}, ..., X_n^{'}$ que tem a mesma distribuição conjunta de $X_1, ..., X_n$, . Portanto, se B usa o estimador $\delta(X_1^{'}, ..., X_n^{'})$, segue que a distribuição de probabilidade do estimador de B será a mesma da distribuição de probabilidade do estimador de A.

Critério de Fatorização

Seja $X_1, ..., X_n$ uma a.a de uma distribuição com f.d.p ou f.p $f(x|\theta)$, onde $\theta \in \Omega$ é desconhecido. Uma estatística $T = r(X_1, ..., X_n$ é suficiente para θ se e somente se a distribuição conjunta $f_n(\mathbf{x}|\theta)$ pode ser fatorada como segue para todos os valores de $\mathbf{x} = (x_1, ..., x_n)$ e todos os valores de $\theta \in \Omega$:

$$f_n(\mathbf{x}|\theta) = \mathbf{u}(\mathbf{x})v[r(\mathbf{x}), \theta] \tag{11}$$

$\mathbf{u}$ depende de $\mathbf{x}$ mas não de θ e a função v depende de θ mas a dependência é nos valores observados $\mathbf{x}$ a través da estatística $r(\mathbf{x})$.

Exemplo: Amostragem de uma Distribuição de Poisson

Suponha que as v.a. $X_1, ..., X_n$ formam uma a.a. de uma distribuição de Poisson com média θ desconhecida ($\theta > 0$). Mostraremos que $T = \sum_{i=1}^{n} X_i$ é uma estatística suficiente para θ.
Para qualquer conjunto de inteiros não negativos $x_1, ..., x_n$ a f.p conjunta $f_n(\mathbf{x}|\theta)$ de $X_1, ..., X_n$ é:

$$f_n(\mathbf{x}|\theta) = \prod_{i=1}^{n} f(x_i|\theta) = (\prod_{i=1}^{n} \frac{1}{x_i!})e^{-n\theta}\theta^{\sum_{i=1}^{n} x_i} = (\prod_{i=1}^{n} \frac{1}{x_i!})e^{-n\theta}\theta^{y}$$

onde $y = \sum_{i=1}^{n} x_i$.

Exemplo: Amostragem de uma Distribuição de Poisson

Suponha que as v.a. $X_1, ..., X_n$ formam uma a.a. de uma distribuição de Poisson com média θ desconhecida ($\theta > 0$). Mostraremos que $T = \sum_{i=1}^{n} X_i$ é uma estatística suficiente para θ.
Para qualquer conjunto de inteiros não negativos $x_1, ..., x_n$ a f.p conjunta $f_n(\mathbf{x}|\theta)$ de $X_1, ..., X_n$ é:

$$f_n(\mathbf{x}|\theta) = \prod_{i=1}^{n} f(x_i|\theta) = (\prod_{i=1}^{n} \frac{1}{x_i!})e^{-n\theta}\theta^{\sum_{i=1}^{n} x_i} = (\prod_{i=1}^{n} \frac{1}{x_i!})e^{-n\theta}\theta^{y}$$

onde $y = \sum_{i=1}^{n} x_i$.

Exemplo: Amostragem de uma Distribuição de Poisson

Suponha que as v.a. $X_1, ..., X_n$ formam uma a.a. de uma distribuição de Poisson com média θ desconhecida ($\theta > 0$).
Mostraremos que $T = \sum_{i=1}^{n} X_i$ é uma estatística suficiente para θ.
Para qualquer conjunto de inteiros não negativos $x_1, ..., x_n$ a f.p conjunta $f_n(\mathbf{x}|\theta)$ de $X_1, ..., X_n$ é:

$$f_n(\mathbf{x}|\theta) = \prod_{i=1}^{n} f(x_i|\theta) = (\prod_{i=1}^{n} \frac{1}{x_i!})e^{-n\theta}\theta^{\sum_{i=1}^{n} x_i} = (\prod_{i=1}^{n} \frac{1}{x_i!})e^{-n\theta}\theta^{y}$$

onde $y = \sum_{i=1}^{n} x_i$.

Exemplo: Amostragem de uma Distribuição de Poisson

Pode-se ver que $f_n(\mathbf{x}|\theta)$ pode ser expressado como a equação 11 como o produto de uma função que não depende de θ e uma função que depende de θ mas esta dependência é no vetor observado $\mathbf{x}$ através do valor de y. Segue que $T = \sum_{i=1}^{n} X_i$ é uma estatística suficiente para θ.

Exemplo: Amostragem de uma Distribuição contínua

Suponha que as v.a. $X_1, ..., X_n$ formam uma a.a. de uma distribuição contínua com a seguinte f.d.p.

$$f(x|\theta) = \begin{cases} \theta x^{\theta-1} & \text{para} \quad 0 < x < 1 \\ 0 & \text{c.c.} \end{cases} \tag{12}$$

assuma que o valor do parâmetro θ é desconhecido ($\theta > 0$). Mostraremos que $T = \prod_{i=1}^{n} X_i$ é uma estatística suficiente para θ.

Exemplo: Amostragem de uma Distribuição contínua

Para $0 < x_i < 1$ $(i = 1, ..., n)$ a f.d.p. conjunta $f_n(\mathbf{x}|\theta)$ de $X_1, ..., X_n$ é:

$$f_n(\mathbf{x}|\theta) = \theta^n (\prod_{i=1}^{n} x_i)^{\theta-1} \tag{13}$$

Se, pelo menos um valor de x_i estiver fora do intervalo $0 < x_i < 1$, então $f_n(\mathbf{x}|\theta) = 0$ para cada $\theta \in \Omega$. Observamos que o lado direito da equação 13 depende de $\mathbf{x}$ através do valor do produto $\prod_{i=1}^{n} x_i$. Portanto, se considerarmos $u(\mathbf{x}) = 1$ e $r(\mathbf{x}) = \prod_{i=1}^{n} x_i$, então $f_n(\mathbf{x}|\theta)$ pode ser considerada fatorada. Segue-se então que a estatística $T = \prod_{i=1}^{n} X_i$ é suficiente para θ.

Exemplo: Amostragem de uma Normal

Suponha que as v.a. $X_1, ..., X_n$ formam uma a.a. de uma distribuição Normal com média desconhecido μ $(-\infty < \theta < \infty)$ e variância σ^2 conhecida. Mostramos que $T = sum_{i=1}^{n} X_i$ é uma estatística suficiente para μ.

Para $-\infty < x_i < \infty$ $(i = 1, ..., n)$ a f.d.p.conjunta de $X_1, ..., X_n$ é

$$f_n(\mathbf{x}|\mu) = \prod_{i=1}^{n} \frac{1}{(2\pi)^{1/2}\sigma} exp[-\frac{(x_i - \mu)^2}{2\sigma^2}] \tag{14}$$

que pode ser rescrita como

$$f_n(\mathbf{x}|\mu) = \frac{1}{(2\pi)^{n/2}\sigma^n} exp(-\frac{1}{2\sigma^2}\sum_{i=1}^{n} x_i^2) exp(\frac{\mu}{\sigma^2}\sum_{i=1}^{n} x_i - \frac{n\mu^2}{2\sigma^2}) \tag{15}$$

Exemplo: Amostragem de uma Normal

Suponha que as v.a. $X_1, ..., X_n$ formam uma a.a. de uma distribuição Normal com média desconhecido μ $(-\infty < \theta < \infty)$ e variância σ^2 conhecida. Mostramos que $T = sum_{i=1}^{n} X_i$ é uma estatística suficiente para μ.

Para $-\infty < x_i < \infty$ $(i = 1, ..., n)$ a f.d.p.conjunta de $X_1, ..., X_n$ é

$$f_n(\mathbf{x}|\mu) = \prod_{i=1}^{n} \frac{1}{(2\pi)^{1/2}\sigma} exp[-\frac{(x_i - \mu)^2}{2\sigma^2}] \quad (14)$$

que pode ser rescrita como

$$f_n(\mathbf{x}|\mu) = \frac{1}{(2\pi)^{n/2}\sigma^n} exp(-\frac{1}{2\sigma^2}\sum_{i=1}^{n} x_i^2) exp(\frac{\mu}{\sigma^2}\sum_{i=1}^{n} x_i - \frac{n\mu^2}{2\sigma^2}) \quad (15)$$

Exemplo: Amostragem de uma Normal

Suponha que as v.a. $X_1, ..., X_n$ formam uma a.a. de uma distribuição Normal com média desconhecido μ $(-\infty < \theta < \infty)$ e variância σ^2 conhecida. Mostramos que $T = sum_{i=1}^{n} X_i$ é uma estatística suficiente para μ.

Para $-\infty < x_i < \infty$ $(i = 1, ..., n)$ a f.d.p.conjunta de $X_1, ..., X_n$ é

$$f_n(\mathbf{x}|\mu) = \prod_{i=1}^{n} \frac{1}{(2\pi)^{1/2}\sigma} exp[-\frac{(x_i - \mu)^2}{2\sigma^2}] \qquad (14)$$

que pode ser rescrita como

$$f_n(\mathbf{x}|\mu) = \frac{1}{(2\pi)^{n/2}\sigma^n} exp(-\frac{1}{2\sigma^2}\sum_{i=1}^{n} x_i^2) exp(\frac{\mu}{\sigma^2}\sum_{i=1}^{n} x_i - \frac{n\mu^2}{2\sigma^2}) \qquad (15)$$

Exemplo: Amostragem de uma Normal

pode se notar que $f_n(\mathbf{x}|\mu)$ pode ser expressado como o produto de uma função que não depende de μ e uma função que depende de **x** através do valor de $\sum_{i=1}^{n} x_i$. Portanto, pelo critério de fatorização, $T = \sum_{i=1} nX_i$ é uma estatística suficiente para μ.

Desde que $T = \sum_{i=1} nX_i = n\bar{x}_n$ podemos estabelecer de forma equivalente, que fator final da equação 15 depende de $x_1, ..., x_n$ através do valor de $\bar{x}_n$, o que implica que $T = \bar{X}_n$ é uma estatística suficiente para μ.

Exemplo: Amostragem de uma Normal

pode se notar que $f_n(\mathbf{x}|\mu)$ pode ser expressado como o produto de uma função que não depende de μ e uma função que depende de **x** através do valor de $\sum_{i=1}^{n} x_i$. Portanto, pelo critério de fatorização, $T = \sum_{i=1} nX_i$ é uma estatística suficiente para μ.
Desde que $T = \sum_{i=1} nX_i = n\bar{x}_n$ podemos estabelecer de forma equivalente, que fator final da equação 15 depende de $x_1, ..., x_n$ através do valor de $\bar{x}_n$, o que implica que $T = \bar{X}_n$ é uma estatística suficiente para μ.

Outline

Limitações no uso de Estatísticas Suficientes

A existência e a forma da estatística suficiente em um determinado problema depende de *maneira crítica* da forma da função assumida pela f.d.p. ou f.p. Uma estatística que é suficiente quando a f.d.p. é $f(x|\theta)$ pode não ser suficiente quando é assumido que a f.d.p. é $g(x|\theta)$, mesmo que $g(x|\theta)$ possa ser similar a $f(x|\theta)$ para cada valor de $\theta \in \Omega$.

Limitações no uso de estatísticas suficientes

Suponha que um pesquisador está em dúvida quanto à forma exata da f.d.p. em um problema específico, mas assume por conveniência que a f.d.p é $f(x|\theta)$, suponha também que a estatística T é suficiente sob essa suposição. Por causa da incerteza do pesquisador quanto à forma exata da f.d.p. ele desejaria usar um estimador de θ que "funciona" razoávelmente bem para uma ampla variedade de f.d.p´s., mesmo que o estimador selecionado não apresente o requerimento que deveria depender nas observações somente através da estatística T.

Um estimador que "funciona" bem para uma ampla variedade de possíveis f.d.p´s., mesmo que não seja necessáriamente o melhor estimador disponível para qualquer particular família de f.d.p´s. é chamado de *estimador robusto*.

Limitações no uso de estatísticas suficientes

Suponha que um pesquisador está em dúvida quanto à forma exata da f.d.p. em um problema específico, mas assume por conveniência que a f.d.p é $f(x|\theta)$, suponha também que a estatística T é suficiente sob essa suposição. Por causa da incerteza do pesquisador quanto à forma exata da f.d.p. ele desejaria usar um estimador de θ que "funciona" razoávelmente bem para uma ampla variedade de f.d.p´s., mesmo que o estimador selecionado não apresente o requerimento que deveria depender nas observações somente através da estatística T.
Um estimador que "funciona" bem para uma ampla variedade de possíveis f.d.p´s., mesmo que não seja necessáriamente o melhor estimador disponível para qualquer particular família de f.d.p´s. é chamado de *estimador robusto*.

Outline

1. Relembrando Amostras Aleatórias
   - Definição de a.a.
   - Exemplos de a.a.
2. Estatísticas Suficientes
   - Estatísticas Suficientes
   - Limitações no uso de estatísticas suficientes
   - Limitações no uso de estatísticas suficientes
3. Estimadores Não Viesados
   - Estimadores Não Viesados

Definição

Seja $X_1, ..., X_n$ uma a.a de uma distribuição com f.d.p ou f.p $f(x|\theta)$, onde $\theta \in \Omega$ é desconhecido. Neste tipo de problemas, é desejável usar um estimador $\delta(X_1, ..., X_n)$ que, com alta probabilidade, seja proximo ao valor de θ, i.e, um estimador δ cujo valor mude com o valor de θ, de tal forma que sem interesar qual seja o valor verdadeiro de θ, a distribuição de probabilidade de δ está concentrado ao redor deste valor.

Definição

Um estimador $\delta(X_1, ..., X_n)$ é um estimador não viesado do parâmetro θ se $E_\theta[\delta(X_1, ..., X_n)] = \theta$, para cada possível valor de θ.
Em outras palavras, um estimador de um parâmetro θ é não viesado se sua esperança é igual ao valor desconhecido do verdadeiro valor de θ.

Exemplo

$\bar{X}_n$ é um estimador não viesado da média desconhecida θ de uma normal, pois $E_\theta[\bar{X}_n] = \theta$ para $-\infty < \theta < \infty$.

Se $X_1, ..., X_n$ forma uma a.a. de uma distribuição arbitrária, para a qual a média μ é desconhecida, a média amostral sempre será um estimador não viesado de μ pois é sempre verdade que $E(\bar{X}_n) = \mu$.

Exemplo

$\bar{X}_n$ é um estimador não viesado da média desconhecida θ de uma normal, pois $E_\theta[\bar{X}_n] = \theta$ para $-\infty < \theta < \infty$.
Se $X_1, ..., X_n$ forma uma a.a. de uma distribuição arbitrária, para a qual a média μ é desconhecida, a média amostral sempre será um estimador não viesado de μ pois é sempre verdade que $E(\bar{X}_n) = \mu$.

Exemplo

Se definirmos

$$S_0^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X}_n)^2 \tag{16}$$

para tentar demonstrar se é ou não um estimador não viesado de σ^2, então usaremos a identidades:

$$\sum_{i=1}^{n}(X_i - \mu)^2 = \sum_{i=1}^{n}(X_i - \bar{X}_n)^2 + n(\bar{X}_n - \mu)^2 \tag{17}$$

então segue que

$$E[S_0^2] = E[\frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X}_n)^2] \tag{18}$$

$$= E[\frac{1}{n}\sum_{i=1}^{n}(X_i - \mu)^2] - E[(\bar{X}_n - \mu)^2] \tag{19}$$

Exemplo

Como cada observação X_i tem média μ e variância σ^2, então: $E[(X_i-\mu)^2]=\sigma^2$ para $i=1,...,n$. Desta forma:

$$E[\frac{1}{n}\sum_{i=1}^{n}(X_i-\mu)^2]=\frac{1}{n}\sum_{i=1}^{n}E[(X_i-\mu)^2]=\frac{1}{n}n\sigma^2=\sigma^2 \tag{20}$$

Como $\bar{X}_n$ tem média μ e variância σ^2/n

$$E[(\bar{X}_n-\mu)^2]=var(\bar{X}_n)=\sigma^2/n \tag{21}$$

Segue-se então que

$$E(S_0^2)=\sigma^2-\frac{1}{n}\sigma^2=\frac{n-1}{n}\sigma^2 \tag{22}$$

Exemplo

O estimador não viesado de σ^2 será então:

$$S_1^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X}_n)^2$$