Inferência Estatística

$X \sim N(\mu, \sigma^2)$ $\quad E(X) = \mu$ $\quad Var(X) = \sigma^2$

$Y \sim Bernoulli(p)$ $\quad E(Y) = p$ $\quad Var(Y) = p(1-p)$

$W \sim Exp(\alpha)$ $\quad E(W) = \dfrac{1}{\alpha}$ $\quad Var(W) = \dfrac{1}{\alpha^2}$

Geralmente $\mu, \sigma^2, p, \alpha \rightarrow$ parâmetros populacionais desconhecidos

Inferência Estatística - Conjunto de técnicas destinadas a tirar conclusões sobre uma população (ou sobre seus parâmetros) com base em evidências fornecidas pela amostra.

Quando utilizamos uma distribuição de probabilidades para modelar a variável de interesse, a característica populacional desconhecida (média, variância populacionais, proporção populacional) se torna um parâmetro de uma distribuição de probabilidades (p. ex: μ, σ^2, p, α)

Ex: X - altura de indivíduos de uma população

$E(X) = \mu$ - altura média dos indivíduos dessa população

$$Y = \begin{cases} 1 & \text{se o aluno vai fazer vestibular} \\ 0 & \text{caso contrário} \end{cases}$$

$Y \sim \text{Bernoulli}(p)$

p - proporção de alunos na população que vai fazer vestibular

$\mu, p \rightarrow$ parâmetros desconhecidos

Procedimentos Inferenciais: Estimação

Testes de Hipóteses

Tomada uma amostra aleatória (a.c.s. com rep.) da população, os estimadores de μ, σ^2 e p são respectivamente

$$\bar{X} = \frac{X_1 + \cdots + X_n}{n} \quad \text{média amostral}$$

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

$\hat{p}$ - proporção amostral

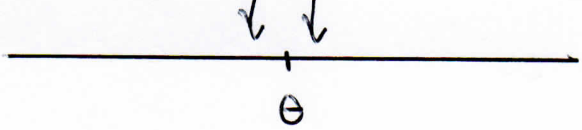
Definição 6 - Consistência

Um estimador $\hat{\theta}$ é consistente se

i) $\lim_{n \to \infty} E(\hat{\theta}) = \theta$ (o estimador é assintoticamente não viciado)

e

ii) $\lim_{n \to \infty} Var(\hat{\theta}) = 0$ $\rightarrow$ Para n grande, os valores de $\hat{\theta}$ são altamente concentrados em torno de sua média que, pelo item a), é aproximadamente θ.



Obs:

Se $\hat{\theta}$ é não viciado, $E(\hat{\theta}) = \theta \quad \forall n$ e

$\therefore \lim\limits_{n \to \infty} E(\hat{\theta}) = \theta$

Se $\hat{\theta}$ é viciado mas é consistente então será não viciado apenas para n grande.

No exemplo 3) de estimação

$E(X) = \mu \qquad Var(X) = \sigma^2$

$X_1, X_2 \dots X_n$ amostra aleatória de X

$\Rightarrow \bar{X} = \dfrac{\sum_{i=1}^{n} X_i}{n}$ é não viciado para μ e

$$Var(\bar{X}) = \frac{\sigma^2}{n}$$

$$\lim_{n \to \infty} Var(\bar{X}) = \lim_{n \to \infty} \frac{\sigma^2}{n} = 0 \qquad (1)$$

$$e \quad \lim_{n \to \infty} E(\bar{X}) = \lim_{n \to \infty} \mu = \mu \quad (2)$$

$\overset{(1)\,e\,(2)}{\Longrightarrow}$ $\bar{X}$ é um estimador consistente de μ.

Para n grande, $\bar{X}$ é altamente concentrado em torno de sua média, que é μ, o parâmetro que desejamos estimar.

Exemplo

$X \sim N(\mu, \sigma^2)$. Amostra aleatória de tamanho n.

$$\left.\begin{array}{l} E(\bar{X}) = \mu \\ Var(\bar{X}) = \dfrac{\sigma^2}{n} \end{array}\right\} \bar{X} \text{ é não viciado e consistente para } \mu.$$

$\hat{\mu}_1 = \bar{X}$

Seja $\hat{\mu}_2 = Med(X_1, X_2 \ldots X_n) \to$ Mediana

Verifica-se que $\begin{cases} E(\hat{\mu}_2) = \mu \text{ (não viciado para } \mu) \\ Var(\hat{\mu}_2) = \dfrac{\pi}{2}\dfrac{\sigma^2}{n} \end{cases}$

Como $\lim_{n \to \infty} Var(\hat{\mu}_2) = \lim_{n \to \infty} \frac{\pi}{2} \frac{\sigma^2}{n} = 0$,

$\hat{\mu}_2 = Med(X_1, X_2, \ldots, X_n)$ é um estimador consistente para μ.

Mas $$\frac{Var(\hat{\mu}_1)}{Var(\hat{\mu}_2)} = \frac{\frac{\sigma^2}{n}}{\frac{\pi}{2}\frac{\sigma^2}{n}} = \frac{2}{\pi} = 0{,}63 \Rightarrow$$

$Var(\hat{\mu}_1) < Var(\hat{\mu}_2) \Rightarrow \hat{\mu}_1$ é mais eficiente que $\hat{\mu}_2$.

Estimadores para média, proporção e Variância

Parâmetro	Estimador	Propriedades
$\mu = E(X)$	$\bar{X} = \frac{X_1 + X_2 + \cdots + X_n}{n}$	Não viciado e Consistente
p	$\hat{p}$ = proporção amostral	não viciado e Consistente
σ^2	$S^2 = \frac{1}{n-1}\left(\sum X_i^2 - n\bar{X}^2\right)$	não viciado e Consistente

→ Variância amostral

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2$$ é viciado e consistente

$$E(\hat{\sigma}^2) = \frac{n-1}{n}\sigma^2$$

Demonstração: Pag 232 Magalhães e Lima

S^2 e $\hat{\sigma}^2$ são consistentes

Como S^2 é não viciado para σ^2, é usualmente adotado para estimar σ^2.

Obs:

- A variância ou o desvio padrão de um estimador não viciado fornecem uma ideia de sua precisão, por isso, o desvio padrão do estimador é denominado erro padrão.

$$\text{Erro Padrão}(\bar{X}) = \frac{\sigma}{\sqrt{n}}.$$

- Fórmula para cálculo

$$S^2 = \frac{1}{n-1}\left[\sum_{i=1}^{n} X_i^2 - n\bar{X}^2\right].$$

Métodos de Obtenção de Estimadores

- Momentos
- Máxima Verossimilhança
- Mínimos Quadrados

Método da Máxima Verossimilhança

X variável aleatória com distribuição de probabilidades que depende de um parâmetro desconhecido θ.

Tomada uma amostra aleatória da v.a. X, $X_1, X_2 \ldots, X_n$, a função de verossimilhança é definida como

$$L(\theta, x_1, x_2 \ldots x_n) = P(X_1 = x_1)\, P(X_2 = x_2) \ldots \times P(X_n = x_n)$$

se X for discreta e

$$L(\theta, x_1, x_2 \ldots x_n) = f(x_1, \theta)\, f(x_2, \theta) \ldots f(x_n, \theta)$$

$f(x, \theta)$ função densidade se X for contínua

Vista como função de θ.

Definição

O estimador de máxima verossimilhança de θ é o valor $\hat{\theta}_{mv}$ que maximiza $L(\theta, x_1, x_2 \dots x_n)$.

Exemplo 1

$X \sim \text{Bernoulli}(p)$

fracasso, sucesso

x	0	1
$P(X=x)$	$1-p$	p

$$P(X=x) = p^x (1-p)^{1-x}$$

$x=0 \quad p^0(1-p)^{1-0} = 1-p$

$x=1 \quad p^1(1-p)^{1-1} = p$

forma condensada de escrever a distribuição de Bernoulli

Se p é desconhecido, tomada uma amostra aleatória $X_1, X_2 \dots X_n$ dessa distribuição

$$L(\theta, x_1, x_2 \dots x_n) = P(X_1 = x_1)\, P(X_2 = x_2) \times \dots \times P(X_n = x_n)$$

$$= p^{x_1}(1-p)^{1-x_1}\, p^{x_2}(1-p)^{1-x_2} \dots p^{x_n}(1-p)^{1-x_n}$$

$$= p^{\sum x_i}(1-p)^{n-\sum x_i} = L(p)$$

$$\ln L(p) = \sum x_i \ln p + (n - \sum x_i) \ln(1-p)$$

$$\frac{\partial \ln L(p)}{\partial p} = \frac{\sum x_i}{p} + (-1)\frac{n - \sum x_i}{1-p}$$

$$\frac{\partial \ln L(p)}{\partial p} = 0 \implies$$

$$\frac{\sum x_i}{p} - \frac{n - \sum x_i}{1-p} = 0$$

$$\frac{(1-p)\sum x_i - p(n - \sum x_i)}{p(1-p)} = 0$$

$$-np = -\sum x_i$$

$$\hat{p} = \frac{\sum X_i}{n} = \frac{\text{nº de sucessos nas } n \text{ repetições}}{n}$$

Prova-se que $\dfrac{\partial^2 \ln L(p)}{\partial p^2} < 0$

Exemplo 2

$X_1, X_2 \ldots X_n$ amostra aleatória da dist. exponencial

$$f(x) = \alpha e^{-\alpha x} \qquad x \geq 0$$

$$L(\alpha, x_1, x_2 \ldots x_n) = \alpha e^{-\alpha x_1} \, \alpha e^{-\alpha x_2} \ldots \times \alpha e^{-\alpha x_n}$$

$$= \alpha^n e^{-\alpha \sum x_i}$$

$$\ln L(\alpha) = n \ln \alpha + \ln e^{-\alpha \sum x_i} = n \ln \alpha - \alpha \sum x_i$$

$$\frac{\partial \ln L}{\partial \alpha} = \frac{n}{\alpha} - \sum x_i = 0$$

$$n - \alpha \sum x_i = 0$$

$$\hat{\alpha} = \frac{n}{\sum X_i} = \frac{1}{\overline{X}}$$

$$\frac{\partial^2 \ln L}{\partial \alpha^2} = \frac{-n}{\alpha^2} < 0 \quad \Rightarrow \hat{\alpha} \text{ é ponto de máximo}$$

$X \sim Exp(\alpha) \qquad E(X) = \mu = \dfrac{1}{\alpha}$

$\therefore \hat{\mu} = \dfrac{1}{\hat{\alpha}} = \dfrac{1}{n/\sum X_i} = \dfrac{\sum X_i}{n} = \overline{X}$ é o est. m.v. de μ

Exemplo 3

$X_1, X_2 \dots X_n$ amostra aleatória da distribuição

$N(\mu, \sigma^2)$, μ e σ^2 desconhecidos.

Obter os estimadores de máxima verossimilhança de μ e σ^2.

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(x-\mu)^2}$$

$$L(\mu, \sigma^2, x_1, x_2 \dots x_n) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{1}{2\sigma^2}(x_i-\mu)^2}$$

$$= \left(\frac{1}{2\pi\sigma^2}\right)^{n/2} \exp\left[-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2\right]$$

$$l(\mu, \sigma^2) = \ln L(\mu, \sigma^2, x_1, x_2 \dots x_n) =$$

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

$$\frac{\partial l(\mu, \sigma^2)}{\partial \mu} = -\frac{1}{2\sigma^2}(-1)\sum 2(x_i-\mu) = 0$$

$\Longrightarrow$

$$\sum(x_i - \mu) = 0 \quad \left[\sum x_i - n\mu = 0 \right]$$

$$\hat{\mu} = \frac{\sum X_i}{n} = \bar{X}$$ não viciado para μ

$$\frac{\partial l(\mu, \sigma^2)}{\partial \sigma^2} = -\frac{n}{2} \cdot \frac{1}{\sigma^2} + \frac{1}{2\sigma^4} \sum (x_i - \mu)^2$$

$$\frac{-n}{2\sigma^2} + \frac{1}{2\sigma^4} \sum (x_i - \mu)^2 = 0$$

$$\frac{-n\sigma^2 + \sum (x_i - \mu)^2}{2\sigma^4} = 0$$

$$\hat{\sigma}^2 = \frac{\sum (X_i - \bar{X})^2}{n}$$ viciado para σ^2

Aproximação da Binomial pela Normal

A aproximação é boa se $np(1-p) > 3$.

Bussab e Morettin $np > 5$

$n(1-p) > 5$