Estatística Paramétrica : Estimação Pontual
Intervalos de confiança

Contexto: Se assume q' $x_1, \dots, x_n$ é uma amostra de $X_1, \dots, X_n$ desde $f(\cdot;\theta)$ onde a forma de f é conhecida e θ é desconhecido mas θ é escalar.

Tipos de Estimação $P|\theta$ com base em $x_1, \dots, x_n$.

(i) Estimação pontual: é quando uma função estatística $l(X_1, \dots, X_n)$ e $l(X_1, \dots, X_n)$ estimará o valor desconhecido de $\tau(\theta)$ (função do parâmetro θ desconhecido)

(ii) Estimação por intervalo: são definidas duas estatísticas $l_1(X_1, \dots, X_n)$ $l_2(X_1, \dots, X_n)$ tais q' $l_1 < l_2$; assim $(l_1(X_1, \dots, X_n); l_2(X_1, \dots, X_n))$ constitue um intervalo a onde se estima. pertence $\tau(\theta)$.

Exemplo simples: se em n ensaios independentes, com probabilidade de sucesso constante p, (p desconhecida e objeto da estimação) observamos k sucessos e $n-k$ fracassos, a nossa estimativa pontual de p.

pode ser: $\frac{k}{n}$ = frequência relativa de sucessos.

Métodos para calcular estimadores :

Contexto : $X_1, \dots, X_n$ amostra de $f(\cdot\,; \theta)$, forma do f conhecida
θ desconhecido.

θ : pode ser vetor $\theta = (\theta_1, \dots, \theta_k)$

(H) : "espaço paramétrico" = conjunto de possíveis valores assumidos por θ.

Metodologia geral p/ estimar θ :

$$\boxed{\text{baseados na amostra } x_1, \dots, x_n \text{ , podemos estimar } (\theta_1, \dots, \theta_k) = \theta \text{ ou } \tau_1(\theta), \dots, \tau_r(\theta) \text{ construindo estatísticas } l_1(x_1, \dots, x_n), \dots, l_r(x_1, \dots, x_n)}$$

Def: se $l(x_1, \dots, x_n)$ pode ser usado p/ estimar $\tau(\theta)$, dizemos q' l é um estimador de $\tau(\theta)$.
Uma estimativa de $\tau(\theta)$ é obtida na hora q' a amostra $x_1, \dots, x_n$ é fixada
$l(x_1, \dots, x_n)$ = estimativa de $\tau(\theta)$.

1) Método dos Momentos : Seja $f(\cdot\,; \theta_1, \dots, \theta_k)$ a densidade da v.a. X
Seja: $\mu'_r = E(X^r)$ r-ésimo momento central de X. Em geral $\mu'_r = \mu'_r(\theta_1, \dots, \theta_k)$
Se $X_1, \dots, X_n$ é uma amostra de $f(\cdot, \theta_1, \dots, \theta_k)$ e M'_j é o j-ésimo momento amostral $M'_j = \frac{1}{n} \sum_{i=1}^{n} X_i^j$

estimador

os M'_j $j = 1, \dots, k$ constituem k equações $M'_j = \mu'_j(\hat{\theta}_1, \dots, \hat{\theta}_k)$
com k incógnitas; as soluções de $\{M'_j\}_{j=1}^{k}$: $\hat{\theta}_1, \dots, \hat{\theta}_k$ são denominados estimadores obtidos pelo método dos momentos (estimadores de $\theta_1, \dots, \theta_k$).

Exemplo $\{x_i\}_{i=1}^{n}$ amostra desde $N(\mu, \sigma^2)$.

$$\mu_1 = E(X) = \mu \xleftarrow{1^a\ m.} \qquad \mu_1 = \mu_1(\mu, \sigma^2)$$

$$\mu_2 = E(X^2) = \sigma^2 + \mu^2 \xleftarrow{2^{da}\ m.} \qquad \mu_2 = \mu_2(\mu, \sigma^2)$$

$$M_1 = \frac{1}{n}\sum_{i=1}^{n} x_i \longleftarrow \text{1ª m. amostral}$$

$$M_2 = \frac{1}{n}\sum_{i=1}^{n} x_i^2 \longleftarrow \text{2da m. amostral}$$

$$\begin{cases} M_1 = \mu_1(\hat{\mu}, \hat{\sigma}^2) \\ M_2 = \mu_2(\hat{\mu}, \hat{\sigma}^2) \end{cases} ; \quad \frac{1}{n}\sum_{i=1}^{n} x_i = \hat{\mu} \quad ; \quad \frac{1}{n}\sum_{i=1}^{n} x_i^2 = \hat{\sigma}^2 + \hat{\mu}^2 \quad ; \quad \hat{\mu}_M = \bar{x}_n \quad ; \quad \hat{\sigma}^2_M = \frac{1}{n}\sum^{n} x_i^2 - \bar{x}_n^2$$

$$\boxed{\hat{\mu}_M = \bar{x}_n \quad ; \quad \hat{\sigma}^2_M \overset{***}{=} \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x}_n)^2}$$

$*$ $\frac{1}{n}\sum x_i^2 = \hat{\sigma}^2 + \hat{\mu}^2 \Longleftrightarrow \frac{1}{n}\sum x_i^2 - \hat{\mu}^2 = \hat{\sigma}^2$ ie $\frac{1}{n}\sum x_i^2 - (\bar{x})^2 = \hat{\sigma}^2$

$**$ note que $\frac{1}{n}\sum (x_i - \bar{x})^2 = \frac{1}{n}\sum\{x_i^2 - 2x_i\bar{x} + \bar{x}^2\} =$

$= \frac{1}{n}\sum x_i^2 - 2\frac{\bar{x}}{n}\sum x_i + \frac{n}{n}\bar{x}^2 = \frac{1}{n}\sum x_i^2 - 2\bar{x}^2 + \bar{x}^2 = \frac{1}{n}\sum x_i^2 - \bar{x}^2$

ou seja $***$ é valido

2) Método de Máximo Verossimilhança:

Def: Uma função de densidade de n variáveis $x_1 \ldots, x_n$ e um parâmetro θ (pode ser k-dimensional), digamos $f(x_1, \ldots, x_n; \theta)$ é chamado verossimilhança. (a função é expressa como os dados (amostra) se relaciona com o parâmetro θ)

Nota 1: Se $x_1, \ldots, x_n$ é uma amostra $f(x_1, \ldots, x_n; \theta) = \prod_{i=1}^{n} f(x_i; \theta)$

Nota 2: $f(x_1, \ldots, x_n; \theta)$ é comumente denotado por $L(\cdot\,; x_1, \ldots, x_n)$ e vista como função de θ para uma sequência $x_1, \ldots, x_n$ fixa.

Def: <u>Estimador de máximo verossimilhança</u>.

Se $L(\theta) = L(\theta; x_1, \dots, x_n)$ onde $(x_1, \dots, x_n)$ é o valor assumido por $(X_1, \dots, X_n)$; então $\hat{\theta} = \hat{\theta}(x_1, \dots, x_n)$ é o estimador de máx. verossimilhança se $\hat{\theta}$ maximiza a função $L(\cdot)$. (ou maximiza $\log(L(\cdot))$).

ie se $\theta = (\theta_1, \dots, \theta_k)$ e ~~$(x_1, \dots, x_n)$ é uma amostra aleat.~~

L satisfaz condições de regularidade, $\hat{\theta}$ é tq.
mais adiante

$$\frac{\partial L}{\partial \theta_j}(\hat{\theta}) = 0 \quad \forall j = 1, \dots, k \; ; \; L(\hat{\theta}) \geq L(\theta) \quad \forall \theta \in \Theta$$ (espaço param. met de θ)

i.e. $\hat{\theta} \doteq \left(\hat{\theta}_1(x_1, \dots, x_n), \hat{\theta}_2(x_1, \dots, x_n), \dots, \hat{\theta}_k(x_1, \dots, x_n)\right)$

<u>Exemplo</u> $x_1, \dots, x_n$ amostra de $N(\mu, \sigma^2)$ $\underbrace{(\mu, \sigma^2)}_{\theta}$ dos conhecidos.

$$L(\mu, \sigma^2) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x_i-\mu)^2}{2\sigma^2}} = \left(\frac{1}{2\pi\sigma^2}\right)^{n/2} \cdot \exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2\right\}$$

$$\log L(\mu, \sigma^2) = -\frac{n}{2} \cdot \log(2\pi) - \frac{n}{2}\log(\sigma^2) - \frac{1}{2\sigma^2} \cdot \sum_{i=1}^{n}(x_i-\mu)^2 \; ; \; \Theta = \left\{(\mu, \sigma^2), \mu \in \mathbb{R}, \sigma^2 > 0\right\}$$

$$\frac{\partial \log L}{\partial \mu} = \frac{1}{\sigma^2}\sum_{i=1}^{n}(x_i-\mu) \; ; \; \frac{\partial \log L}{\partial \sigma^2} = -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4}\sum_{i=1}^{n}(x_i-\mu)^2$$

$$\boxed{\hat{\mu} = \bar{x} \; ; \; \hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2}$$ · estimadores de MV.

ver EP 3

$$\hat{\theta} = (\hat{\mu}, \hat{\sigma}^2)$$

$$\frac{d}{dx}\left(-\frac{1}{2x}\sum(x_i-\mu)^2\right) = -\frac{\sum(x_i-\mu)^2}{2}(-1)x^{-2}$$

Veja 4 EP

$$0 = \frac{1}{\sigma^2} \sum (x_i - \mu) \quad \text{sea} \quad \sum x_i = n\mu \quad \text{sea} \quad \hat{\mu} = \frac{\sum x_i}{n}$$

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

ie $\hat{\sigma}^2 = \dfrac{\sum (x_i - \hat{\mu})^2}{n}$

Estimação Pontual, casos específicos.

Dist Binomial.
" Poisson.
" Normal
" exponencial.

Binomial (1) Suponhamos q' uma urna contem um nº de bolas pretas e um nº de bolas brancas, sabe q' conhecemos o radio 3/1, mas não conhecemos se são 3 pretas 1 branca ou 3 brancas e 1 preta. Isto é a prob de extrair uma bola preta da urna pode ser $\frac{3}{4}$ ou $\frac{1}{4}$. (se 3 é o nº de bolas pretas na urna ou 1 respectivamente).

Se n bolas são extraídas da urna com reposição e X = # de bolas extraídas ao final das n extracções. $X \sim Bin(p, n)$

ie $$f(x;p) = \binom{n}{x} p^x q^{n-x} \qquad x = 0, 1, 2, \ldots, n.$$

Se tiramos uma amostra de tamanho $n=3$. e pretendemos estimar o valor de p. (ele pode ser $\frac{3}{4}$ ou $\frac{1}{4}$). tendemos

x	0	1	2	3
$f(x;\frac{3}{4})$	1/64	9/64	27/64	27/64
$f(x;\frac{1}{4})$	27/64	27/64	9/64	1/64

Então se $x=0$ $\hat{p}(x=0) = \frac{1}{4}$ pois $f(0;\frac{1}{4}) > f(0;\frac{3}{4})$.

Si $x=1$ $\hat{p}(x=1) = \frac{1}{4}$ " $f(1;\frac{1}{4}) > f(1,\frac{3}{4})$

⋮

x	$\hat{p}$
0	0,25
1	0,25
2	0,75
3	0,75

Claramente temos escolhido $\hat{p}$ em função da amostra, olhando a probabilidade maior.

Esse procedimento corresponde a Estimação por máxima verossimilhança.

$$f(x;p) = \binom{n}{x} p^x q^{n-x} \qquad x = 0, 1, \ldots, n$$

uma vez conhecido o valor da amostra x

$$L(p) = \binom{n}{x} p^x (1-p)^{n-x} \qquad x \text{ fixo.}$$

$$\frac{\partial L}{\partial p} = x \cdot p^{x-1} \cdot (1-p)^{n-x} \cdot \binom{n}{x} + \binom{n}{x} p^x \cdot (n-x) \cdot (1-p)^{n-x-1} (-1)$$

$$= \binom{n}{x} p^{x-1} \cdot (1-p)^{n-x-1} \left\{ x(1-p) - p(n-x) \right\} = 0.$$

se $p=0$, $p=1$ ou $x(1-p) - p \cdot (n-x) = 0$

$$x - \cancel{xp} + \cancel{px} - np$$

$$x - pn = 0.$$

$$\hat{p} = \frac{x}{n}$$

$$\therefore \hat{p} = \frac{x}{n} = \frac{\#\text{ de bolas pretas encontradas no experimento}}{\#\text{ total de bolas}}.$$

: estimador de máxima verossimilhança.

(2) Apresentamos a seguir a estimativa de p proposta pelo método dos momentos

Como é simples de notar, $X \sim Bin(p, n)$, $X = \sum_{i=1}^{n} X_i$; $X_i \sim Ber(p)$ indep

Então se considerarmos $X_1, \ldots, X_n$, n realizações indep de Y

$Y \sim Ber(p)$ temos $\mu_1 = E(Y) = p$

Tendo n realiz de Y: $x_1 \dots x_n$

$$\underbrace{\text{o momento amostral}}_{=M_1} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

logo $\boxed{\hat{p}_{M_1} = \frac{\sum_{i=1}^{n} x_i}{n} = \frac{x}{n}}$ x valor assumido pelo $X \sim Bin(n, p)$, $x_i \sim Ber(p)$

Note que ambos métodos oferecem o mesmo estimador de

$$p: \hat{p} = \bar{x} \quad , \bar{x} = \frac{\sum^{n} x_i}{n} \; ; x_i \sim Ber(p)$$

---

(3)

Propriedades do estimador de p: $\bar{x}$

<u>Viés</u>: $E(\hat{p}) = E(\bar{x}) = E\left(\frac{\sum^{n} X_i}{n}\right) = \frac{1}{n}\sum^{n} \underbrace{E(X_i)}_{p} = p$ $\quad X_i \sim Ber(p)$.

$\therefore \boxed{\hat{p} = \bar{x} \text{ é não enviesado}}$

(4) Se o objetivo for estimar $p(1-p)$: função não 1:1 de p $\therefore$ não vale o 1º teo do anulo anterior, se vale o seguinte.

(seja $\hat{\theta}$ est MV. de θ na densidade $f(\cdot, \theta)$ se $\tau(\theta)$ = for uma transf do espaço paramétrico, então o est de MV. de $\tau(\theta)$ é $\tau(\hat{\theta})$)

$\boxed{\text{Então o estimador de MV de } p(1-p) \text{ é } \bar{x}(1-\bar{x})}$

(5) Estatística suficiente e critério da fatoração

$$f(x_i) = p^{x_i}(1-p)^{1-x_i} \quad x_i = 0,1 \quad 0 \le p \le 1$$

$$\Rightarrow f(\theta, x_1, \dots, x_n) = p^{\sum x_i}(1-p)^{n-\sum x_i}$$

$\Rightarrow \boxed{\sum_{i=1}^{n} x_i \text{ é est suficiente p/ } p \text{ pelo teorema da fatoração}}$

Continuando com a estimativa não enviesada de p:

$$\hat{p}_n = \bar{x}_n \qquad \bar{x}_n = \frac{\sum^{n} x_i}{n}$$

$x_i \sim Ber(p)$

$$Var(\hat{p}_n) = \frac{1}{n^2} \cdot n\, Var(x_i) = \frac{Var(x_i)}{n} = \frac{p \cdot (1-p)}{n} \xrightarrow[n \to \infty]{} 0$$

$$\boxed{\hat{p}_n = \bar{x}_n \Rightarrow Var(\hat{p}_n) \xrightarrow[n \to \infty]{} 0}$$

$\therefore$ Se consideramos uma sequencia de estimadores de p. baseados em um nº (variável) de realizações de Bernoullis indep $x_1, \dots, x_n \dots$

Onde $\left\{ \hat{p}_n = \bar{x}_n \right\}_{n=1}^{\infty}$ a sequencia é erro quadrático consistente

pois $\hat{p}_n$ é estimador não enviesado de p e $Var(\hat{p}_n) \xrightarrow[n \to \infty]{} 0$.

Então $\hat{p}_n$ é consistente simples. estimador de p