CC0288 - Inferência Estatística I

Método da Máxima Verossimilhança - 10/04/2023.

Prof. Maurício

1. Estimadores de Máxima Verossimilhança

O Novo Dicionário Aurélio da Língua Portuguesa (2ł edição, 1986) define verossímil (ou verossimilhante) aquilo que é semelhante à verdade, provável, e verossimilhança(ou verossimilidade, ou ainda verossimilitude), à qualidade ou caráter de verossímil. O que seria uma amostra verossímil?

Seria uma amostra que fornecesse a melhor informação possível sobre um parâmetro de interesse da população, desconhecido, e que desejamos estimar.

O princípio da verossimilhança afirma que devemos escolher aquele valor do parâmetro desconhecido que maximiza a probabilidade de obter a amostra particular observada, ou seja, o valor que torna aquela amostra a mais provável.

O uso desse princípio conduz a um método de estimação pelo qual se obtêm os chamados estimadores de máxima verossimilhança que, em geral, têm propriedades muito boas.

Esse princípio foi enunciado por Fisher pela primeira vez em 1912 e, em 1922, deu-lhe forma mais completa, introduzindo a expressão "likelihood"(verossimilhança). Veja Fisher (1935) para mais detalhes.

Vamos começar com um exemplo.

Exemplo 11.10 Suponha que temos n provas de Bernoulli com $P(sucesso) = p,\ 0 < p < 1$ e X =número de sucessos. Devemos tomar como estimador aquele valor de p que torna a amostra observada a mais provável de ocorrer.

Suponha, por exemplo, que $n = 3$ e obtemos dois sucessos e um fracasso. A função de verossimilhança é

$$L(p) = P(2 \text{ sucessos e } 1 \text{ fracasso}) = p^2(1-p).$$

Maximizando essa função em relação a p, obtemos

$$L'(p) = 2p(1-p) - p^2 = 2p - 3p^2.$$

A derivada segunda de $L(p)$ é dada por:

$$L''(p) = 2 - 6p.$$

Igualando a derivada primeira a zero temos:

$$L'(p) = p(2-3p) = 0,$$

do que seguem $p = 0$ ou $p = \frac{2}{3}$.

Note que

$$L(0) = 0 \quad e \quad L\left(\frac{2}{3}\right) = \frac{4}{9} \times \frac{1}{3} = \frac{4}{27}.$$

É fácil ver que o ponto máximo é $\frac{2}{3}$,

pois

$$L''(\frac{2}{3}) = 2 - 6 \times \frac{2}{3} = 2 - 4 = -2 < 0.$$

que é a estimativa de máxima verossimilhança (EMV) de p.

. De modo geral, o EMV do parâmetro p de uma distribuição Bernoulli é e o parâmetro p de uma distribuição binomial é

$$\hat{p} = \frac{X}{n} \qquad (11.29)$$

que é o estimador usado anteriormente no Exemplo 11.1.

Para chegar a (11.29), observe que a função de verossimilhança nesse caso é

$$L(p) = p^x(1-p)^{n-x},$$

que é a probabilidade de se obter x sucessos e $n - x$fracassos.

O máximo dessa função ocorre no mesmo ponto

$$l(p) = log(L(p)),$$

em $log(a)$ é o logaritmo natural ou neperiano de a.

Assim,

$$l(p) = log\left(p^x(1-p)^{n-x}\right) = log\left(p^x\right) + log\left((1-p)^{n-x}\right),$$

$$l(p) = x\ log(p) + (n-x)\ log(1-p)$$

Derivando temos:

$$l^{'}(p) = \ \frac{x}{p} - \frac{n-x}{1-p}$$

e igualando a zero

$$\frac{x}{p} - \frac{n-x}{1-p} = 0$$

$$\frac{x}{p} = \frac{n-x}{1-p}$$

que pode ser posta na forma:

$$\frac{p}{1-p} = \frac{x}{n-x}$$

Aplicando propriedades de proporção temos:

$$\frac{p}{p+1-p} = \frac{x}{x+n-x},$$

Assim

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

O estimador de MV de p é dado por:
obtemos

$$\widehat{p}_{MV} = \frac{X}{n}.$$

Na realidade o livro passa por cima de vários detalhes que serão agora feitos:
A derivada segunda de $L(p)$ é dada por

$$l^{''}(p) = \ -\frac{x}{p^2} - \frac{n-x}{(1-p)^2}$$

Note que

$$n\,\hat{p} = x \quad e\ \ n - x = n - n\,\hat{p} = n(1-\hat{p}).$$

$$l^{''}(\hat{p}) = \ -\frac{x}{\hat{p}^2} - \frac{n-x}{(1-\hat{p})^2}$$

$$l^{''}(\hat{p}) = \ -\frac{n\,\hat{p}}{\hat{p}^2} - \frac{n(1-\hat{p}}{(1-\hat{p})^2}$$

$$l^{''}(\hat{p}) = \ -n\left[\frac{1}{\hat{p}} + \frac{1}{1-\hat{p}}\right]$$

$$l^{''}(\hat{p}) = -\frac{n}{\hat{p}(1-\hat{p})} < 0,$$

para $0 < \hat{p} < 1$. A condição do máximo está satisfeita.
Note para $x = 0$ temos

$$L(p) = (1-p)^n$$

que é uma função sempre decrescente em p . O máximo ocorre em $p = 0 = \frac{0}{n}$.
E para $x = n$ temos

$$L(p) = p^n$$

que é uma função sempre crescente em p . O máximo ocorre em $p = 1 = \frac{n}{n}$.
Falta mais um lembrete. Para acharmos o máximo procuramos os pontos em que:

$$L'(p) = 0.$$

No caso em que

$$l(p) = log(L(p)).$$

$$l'(p) = \frac{L'(p)}{L(p)} = 0,$$

portanto $l'(p) = 0$ quando $L'(p) = 0$.

Na realidade este estimador obtido é da Bernoulli de parâmetro p. Veja:

$$f(x, p) = p^x (1-p)^{1-x}\ I_A(x), A = \{0, 1\}.$$

Seja $(x_1, x_2, \ldots, x_n)$ uma amostra aleatória de X. A função de verossimilhança de p é dada por:

$$L(p) = \prod_{i=1}^{n} f(x_i, p) = \prod_{i=1}^{n} p^{x_i}(1-p)^{1-x_i}$$

$$L(p) = p^{\displaystyle\sum_{i=1}^{n} x_i} (1-p)^{n-\displaystyle\sum_{i=1}^{n} x_i}$$

Fazendo

$$s = \sum_{i=1}^{n} x_i$$

temos

$$l(p) = p^s(1-p)^{n-s},$$

que é a verossimilhança anterior.

Vamos continuar no livro:

O procedimento, pois, é obter a função de verossimilhança, que depende dos parâmetros desconhecidos e dos valores amostrais, e depois maximizar essa função ou o logaritmo dela, o que pode ser mais conveniente em determinadas situações.

Chamando de

$$L(\theta; X_1, \ldots X_n)$$

a função de verossimilhança,

O logaritmo da função de verossimilhança

$$l\left(\theta; X_1, \ldots X_n\right) = log\left(L(\theta; X_1, \ldots X_n)\right).$$

No caso de variáveis contínuas, a função de verossimilhança é definida da seguinte maneira. Suponha que a v.a. X tenha densidade $f(x;\theta)$, onde destacamos a dependência do parâmetro θ desconhecido.

Retiramos uma amostra de X, de tamanho n , $(X_1,\ldots,X_n)$, e sejam $(x_1,\ldots,x_n)$ os valores efetivamente observados.

Definição.

A função de verossimilhança é definida por

$$L\left(\theta;x_1,\ldots x_n\right)=f(x_1;\theta)\ldots f(x_n;\theta)=\prod_{i=1}^{n}f(x_i;\theta) \qquad (11.30)$$

que deve ser encarada como uma função de θ. O estimador de máxima verossimilhança de θ é o valor θ_{MV} que maximiza $L(\theta;x_1,\ldots,x_n)$.

Se indicarmos por $\mathbf{x}=(\mathbf{x_1},\ldots,\mathbf{x_n})^{\mathbf{t}}$ o vetor contendo a amostra, é costume

denotar a verossimilhança por $L(\theta|\mathbf{x})$ e a log-verossimilhança por $l(\theta|\mathbf{x})$.

O parâmetro θ pode ser um vetor,

como no caso de querermos estimar a média μ e a variância σ^2 de uma normal.

Nesse caso,

$$\theta=(\mu,\sigma^2)^t.$$

Exemplo 11.11 Suponha que a v.a. X tenha distribuição exponencial, com parâmetro $\alpha>0$, desconhecido, e queremos obter o EMV desse parâmetro. A densidade de X é dada por 7.26

$$f(x,\ \alpha)=\frac{1}{\alpha}\ e^{-\frac{x}{\alpha}}\ I_A(x),\ \ a=(0.\infty)$$

Então, a verossimilhança é dada por

$$L(\alpha|\mathbf{x})=\prod_{i=1}^{n}\frac{1}{\alpha}\ e^{-\frac{x_i}{\alpha}}=\frac{1}{\alpha^n}\ e^{-\frac{s}{\alpha}},$$

em que

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

e o logaritmo da função de verossimilhança fica

$$l(\alpha|\mathbf{x})=-n\ log(\alpha)-\frac{s}{\alpha}.$$

Derivando

$$l^{'}(\alpha|\mathbf{x})=-\frac{\mathbf{n}}{\alpha}+\frac{\mathbf{s}}{\alpha^{\mathbf{2}}}=-\mathbf{n}\ (\alpha)^{-\mathbf{1}}+\mathbf{s}\ (\alpha)^{-\mathbf{2}}.$$

A derivada segunda é dada por:

$$l^{''}(\alpha|\mathbf{x}) = n\ (\alpha)^{-2} - 2\ s\ (\alpha)^{-3}.$$

De $l^{'}(\alpha|\mathbf{x}) = \mathbf{0}$ temos:

$$-\ \frac{n}{\alpha} + \frac{s}{\alpha^2} = 0$$

$$-n\ \alpha + s = 0$$

$$\alpha = \frac{s}{n}.$$

Note que:

$$l^{''}(\frac{s}{n}) = n\ (\frac{s}{n})^{-2} - 2\ s\ (\frac{s}{n})^{-3} = n\ \frac{n^2}{s^2} - 2s\ \frac{n^3}{s^3}.$$

$$l^{''}(\frac{s}{n}) = \frac{n^3}{s^2} - 2\ \frac{n^3}{s^2} = -\frac{n^3}{s^2} < 0.$$

O EMV de α é

$$\hat{\alpha}_{MV} = \frac{\sum_{i=1}^{n} X_i}{n} = \bar{X} \qquad (11.31).$$

que nada mais é do que a média amostral. Lembremos que na distribuição exponencial $E(X) = \alpha$, e portanto o estimador obtido é o esperado pelo senso comum.

Na nossa disciplina usamos a seguinte parametrização da distribuição exponencial:

$$f(x,\theta) = \theta\ e^{-\theta x}\ \ I_A(x), \quad A = (0,\infty).$$

Vamos achar o EMV de θ:

$$L(\theta) = \prod_{i=1}^{n} \theta\ e^{-\theta x_i} = \theta^n\ e^{-\theta \sum_{i=1}^{n} x_i} = \theta^n\ e^{-\theta\ s}$$

$$l(\theta) = n\ log(\theta) - \theta s$$

A derivada primeira é dada por:

$$l^{'}(\theta) = \frac{n}{\theta} - s.$$

A derivada segunda é dada por:

$$l^{''}(\theta) = -\ \frac{n}{\theta^2} < 0.$$

Fazendo $l^{'}(\theta) = 0$ temos:

$$\frac{n}{\theta} = s.$$

$$\theta = \frac{s}{n}.$$

O estimador de MV de θ é dado por:

$$\hat{\theta}_{MV} = \frac{n}{\sum_{i=1}^{n} X_i} = \frac{n}{n\,\bar{X}} = \frac{1}{\bar{X}}.$$

Note uma coisa interessante o valor esperado de X vale

$$E(X) = g(\theta) = \frac{1}{\theta} = \alpha.$$

O estimador de máxima verossimilhança de $E(X)$ é dado por:

$$\hat{E}(X) = \hat{\alpha}_{MV} = \bar{X}.$$

Que também pode ser obtido através da propriedade da Invariância dos estimadores de Máxima Verossimilhança que diz:

$$\hat{g(\theta)} = g(\hat{\theta}).$$

Para

$$\mu = E(X) = g(\theta) = \frac{1}{\theta}.$$

$$\hat{\mu} = g\left(\frac{1}{\bar{X}}\right) = \bar{X}.$$

$$E(X) = g(\theta) = \frac{1}{\theta} = \alpha.$$

O estimador de máxima verossimilhança de $\sigma^2 = V(X) = \frac{1}{\theta^2}$ é dado por:
Para

$$\sigma^2 = V(X) = g(\theta) = \frac{1}{\theta^2}.$$

$$\hat{\sigma}^2 = g\left(\frac{1}{\bar{X}}\right) = \bar{X}^2.$$

A única condição exigida sobre $g(\theta)$ é que ela seja uma função mensurável.

No caso discreto, a função de verossimilhança pode ser escrita na forma

$$L\left(\theta; x_1, \dots x_n\right) = P(X_1 = x_1|\theta) \dots P(X_n = x_1|\theta) = f(x_1||\theta) \dots f(x_n|\theta) = \prod_{i=1}^{n} f(x_i|\theta)$$

Na realidade não tem diferença usamos a função densidade de probabilidade no caso contínuo ou a função de probabilidade no caso discreto.

Veja o Problema 37 para o caso de termos mais de um parâmetro.

Exemplo 1 Seja $X \sim Bin(5, \theta)$. Qual o estimador de MV para θ baseado em uma amostra aleatória $X_1, X_2, \ldots, X_n$?

Solução: A f.p. de X é dada por:

$$f(x|\theta) = \binom{5}{x} \theta^x (1-\theta)^{5-x} \ I_A(x), \quad A = \{0, 1, 2, 3, 4, 5\},$$

com

$$\mu = E(X) = 5\theta \quad e \quad V(X) = 5\theta(1-\theta).$$

A função de verossimilhança é dada por:

$$L(\theta) = \prod_{i=1}^{n} \binom{5}{x_i} \theta^{x_i} (1-\theta)^{5-x_i}.$$

$$L(\theta) = \prod_{i=1}^{n} \binom{5}{x_i} \times \theta^s \ \times (1-\theta)^{5n-s},$$

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

O logaritmo da função de verossimilhança é dado por:

$$l(\theta) = \sum_{i=1}^{n} log\left(\binom{5}{x_i}\right) + s \ log(\theta) + (5n-s) \ log(1-\theta)$$

A derivada primeira é dada por:

$$l^{'}(\theta) = \frac{s}{\theta} - \frac{5n-s}{1-\theta}.$$

Fazendo $l^{'}(\theta) = 0$ temos:

$$\frac{s}{\theta} = \frac{5n-s}{1-\theta}.$$

$$\frac{\theta}{1-\theta} = \frac{s}{5n-s}.$$

$$\frac{\theta}{\theta+1-\theta} = \frac{s}{s+5n-s}.$$

$$\theta = \frac{s}{5n} = \frac{n\,\bar{x}}{5n} = \frac{\bar{x}}{5}.$$

A derivada segunda é dada por:

$$l''(\theta) = -\frac{s}{\theta^2} - \frac{5n - s}{(1-\theta)^2}.$$

Assim,

$$l''\left(\frac{s}{5n}\right) = -\frac{s}{\frac{s^2}{25n^2}} - \frac{5n - s}{(1 - \frac{s}{5n})^2}.$$

$$l''\left(\frac{s}{5n}\right) = -\frac{25n^2}{s} - \frac{25n^2}{5n - s}$$

$$l''\left(\frac{s}{5n}\right) = -25n^2\left[\frac{1}{s} + \frac{1}{5n - s}\right] = -\frac{125n^3}{s(5n - s)} < 0, \quad s \neq 0,\ 5n.$$

O estimador de MV T de θ é dado por:

$$T = \frac{\bar{X}}{5}.$$

Note que:

$$E(T) = \frac{E(\bar{X})}{5} = \frac{\mu}{5} = \frac{5\theta}{5} = \theta$$

assim T é não viciado para θ.

$$V(T) = \frac{V(\bar{X})}{25} = \frac{\frac{\sigma^2}{n}}{25} = \frac{\sigma^2}{25n} = \frac{5\theta(1-\theta)}{25n} = \frac{\theta(1-\theta)}{5n}.$$

$$\lim_{n\to\infty} V(T) = \frac{\theta(1-\theta)}{5} \lim_{n\to\infty} \frac{1}{n} = 0.$$

T é um estimador consistente de θ.

Em algumas situações práticas função de verossimilhança é uma função crescente ou decrescente do parâmetro θ. Assim não podemos derivá-la para obter o estimador de MV.

Exemplo: 2 Seja X uma variável aleatória contínua com f.d.p. dada por:

$$f(x;\theta) = e^{-(x-\theta)}\ \ I_A(x), \quad A = [\theta, \infty), \theta > 0.$$

Note que o suporte depende de θ.

Assim na verossimilhança não podemos descartar o suporte.

Logo

$$L(\theta) = \prod_{i=1}^{n} f(x_i, \theta) = \prod_{i=1}^{n} e^{-(x_i - \theta)}\ \ I_A(x_i)$$

$$L(\theta) = e^{-\sum_{i=1}^{n} x_i} \; e^{n\theta} \prod_{i=1}^{n} I_A(x_i)$$

$$L(\theta) = e^{-s} \; e^{n\theta} \prod_{i=1}^{n} I_A(x_i)$$

Mas

$$\prod_{i=1}^{n} I_A(x_i) = 1.$$

Quando

$$I_A(x_1) = 1, \;\; I_A(x_2) = 1, \ldots, I_A(x_n) = 1.$$

Ou

$$x_1 \geq \theta, x_2 \geq \theta, \ldots, x_n \geq \theta.$$

Ou

$$y_1 = min(x_1, x_2, \ldots, x_n) \geq \theta.$$

Ou

$$0 < \theta \leq y_1$$

ou

$$I_{(0,y_1]} \; (\theta) = 1.$$

Logo a verossimilhança fica:

$$L(\theta) = e^{-s} \; e^{n\theta} \; I_{([0,y_1]} \; (\theta).$$

Maximizar $L(\theta)$ é maximizar $g(\theta) = e^{n\theta}$.
Mas

$$g'(\theta) = n \; e^{n\theta} > 0.$$

Logo $g(\theta)$ é uma função sempre crescente . Seu máximo ocorre no limite superior do seu domínio.
Assim

$$\hat{\theta} = y_1 = min(x_1, x_2, \ldots, x_n).$$

Assim o EMV de θ é:

$$Y_1 = min(X_1, X_2, \ldots, X_n).$$

A f.d.p. de Y_1 é dada por:

$$g_{Y_1}(y) = n\,[1 - F_X(y)]^{n-1}\, f(y) = n[S(y)]^{n-1}\, f(y).$$

Mas para $y > \theta$

$$S(y) = \int_y^\infty e^{-(t-\theta)}\, dt = e^{-(y-\theta)}.$$

$$g_{Y_1}(y) = n\,[1 - F_X(y)]^{n-1}\, f(y) = n[e^{-(y-\theta)}]^{n-1}\, e^{-(y-\theta)}\, I_{[\theta,\infty]}(y).$$

$$g_{Y_1}(y) = n\, e^{-n(y-\theta)}\, I_{[\theta,\infty]}(y).$$

O valor esperado de Y_1 é dado por:

$$E(Y_1) = \int_\theta^\infty ny\, e^{-n(y-\theta)}\, dy = \int_\theta^\infty y\, e^{-n(y-\theta)}\, n\, dy$$

Fazendo a mudança de variável:

$$u = n(y-\theta) \quad , \quad du = ndy.$$

logo

$$y = \frac{u}{n} + \theta.$$

Assim

$$E(Y_1) = \int_0^\infty \left(\frac{u}{n} + \theta\right) e^{-u}\, du$$

temos

$$E(Y_1) = \frac{1}{n}\int_0^\infty u\, e^{-u}\, du + \theta \int_0^\infty e^{-u}\, du$$

$$E(Y_1) = \frac{1}{n}\,\Gamma(2) + \theta\,\Gamma(1) = \frac{1}{n} + \theta,$$

pois $\Gamma(2) = 1! = 1$ *e* $\Gamma(1) = 0! = 1$.

Logo

$$T = Y_1 - \frac{1}{n}$$

é um estimador não viciado de θ.

Note que

$$Var(T) = Var\left(Y_1 - \frac{1}{n}\right) = Var(Y_1).$$

O valor esperado de Y_1^2 é dado por:

$$E(Y_1^2) = \int_\theta^\infty ny^2\ e^{-n(y-\theta)}\ dy = \int_\theta^\infty y^2\ e^{-n(y-\theta)}\ n\ dy$$

Fazendo a mesma mudança de variável temos:

$$E(Y_1^2) = \int_0^\infty \left(\frac{u}{n} + \theta\right)^2\ e^{-u}\ du$$

temos

$$E(Y_1^2) = \frac{1}{n^2}\int_0^\infty u^2\ e^{-u}\ du + \frac{2\theta}{n}\int_0^\infty u\ e^{-u}\ du + \theta^2 \int_0^\infty e^{-u}\ du$$

$$E(Y_1^2) = \frac{1}{n^2}\Gamma(3) + \frac{2\theta}{n}\Gamma(2) + \theta^2\ \Gamma(1).$$

$$E(Y_1^2) = \frac{2}{n^2} + \frac{2\theta}{n} + \theta^2,$$

pois $\gamma(3) = 2! = 2$.

$$E(Y_1^2) = \theta^2 + \frac{2\theta}{n} + \frac{1}{n^2} + \frac{1}{n^2},$$

$$E(Y_1^2) = \left[\theta + \frac{1}{n}\right]^2 + \frac{1}{n^2}.$$

Finalmente temos:

$$V(Y_1) = E(Y_1^2) - E^2(Y_1) = \left[\theta + \frac{1}{n}\right]^2 + \frac{1}{n^2} - \left[\theta + \frac{1}{n}\right]^2 = \frac{1}{n^2}$$

Note que Y_1 não é um estimador não viciado de θ mas assintoticamente ele é:

$$\lim_{n\to\infty} E(Y_1) = \lim_{n\to\infty}\left(\theta + \frac{1}{n}\right) = \theta + 0 = \theta.$$

Além disso

$$\lim_{n\to\infty} V(Y_1) = \lim_{n\to\infty}\ \frac{1}{n^2} = 0,$$

Assim Y_1 é um estimador consistente de θ.

Distribuição em grandes amostras

No caso em que o tamanho da amostra é grande e as condições de regularidade estão satisfeitas teremos:

$$\sqrt{n}\left(\widehat{\theta}-\theta\right) \overset{P}{\sim} N\left(0\ ,\ \frac{1}{I_F(\theta)}\right),$$

e

$$\sqrt{n}\left(g(\widehat{\theta})-g(\theta)\right) \overset{a}{\sim} N\left(0\ ,\ \frac{(g'(\theta))^2}{I_F(\theta)}\right).$$

Exemplo 1:

Seja $X \sim Ber(\theta)$. Seja $X_1, X_2, \ldots, X_n$ uma amostra aleatória de X.

Qual o estimador de máxima verossimilhança de $g(\theta) = \theta(1-\theta) = \theta - \theta^2$?

Qual a distribuição assintótica deste estimador?

O estimador de máxima verossimilhança de θ é dado por:

$$\widehat{\theta} = \bar{X}.$$

Pela propriedade da invariância temos que o estimador de máxima verossimilhança de $g(\theta)$,T, é dado por:

$$T = g(\bar{X}) = \bar{X}(1-\bar{X}).$$

Temos que

$$g'(\theta) = 1 - 2\theta.$$

Além disso

$$I_F(\theta) = \frac{1}{\theta(1-\theta)}.$$

Logo

$$\frac{(g'(\theta))^2}{I_F(\theta)} = (1-2\theta)^2\,\theta(1-\theta)$$

$$\sqrt{n}\left(g(\widehat{\theta})-\theta(1-\theta)\right) \overset{a}{\sim} N\left(0\ ,\ (1-2\theta)^2\,\theta(1-\theta)\right).$$

Exemplo 2:

Seja $X \sim Poisson(\theta)$. Seja $X_1, X_2, \ldots, X_n$ uma amostra aleatória de X.

Qual o estimador de máxima verossimilhança de $g(\theta) = e^{-\theta}$?

Qual a distribuição assintótica deste estimador?

O estimador de máxima verossimilhança de θ é dado por:

$$\widehat{\theta} = \bar{X}.$$

Pela propriedade da invariância temos que o estimador de máxima verossimilhança de $g(\theta)$,T, é dado por:

$$T = g(\bar{X}) = e^{-\bar{X}}.$$

Temos que

$$g'(\theta) = -e^{-\theta}.$$

$$(g'(\theta))^2 = e^{-2\theta}.$$

Além disso

$$I_F(\theta) = \theta\ e^{-2\theta}$$

Logo

$$\frac{(g'(\theta))^2}{I_F(\theta)} = \ \theta\ e^{-2\theta}$$

$$\sqrt{n}\left(g(\widehat{\theta}) - \theta(1-\theta)\right) \overset{a}{\sim} N\left(0\ ,\ \theta\ e^{-2\theta}\right).$$

Qual a média de T?

$$E(T) = E\left(e^{-\bar{X}}\right) = M_{\bar{X}}\ (-1)$$

Sabemos que

$$S = \sum_{i=1}^{n} X_i \sim Poisson(n\theta).$$

A função geradora de momentos de S é dada por:

$$M_S(t) = e^{n\theta(e^t-1)}$$

$$M_{\bar{X}}(t) = M_S(t/n) = e^{n\theta(e^{t/n}-1)}$$

Logo,

$$E(T) = e^{n\theta(e^{-1/n}-1)} = e^{-\theta\ \ n\left(e^{-1/n}-1\right)}$$

Logo T é viciado .

Ele é assintoticamente não viciado?

Devemos mostrar que:

$$\lim_{n\to\infty} n\,\left(e^{-1/n}-1\right)=1.$$

Note que:

$$e^a=\sum_{i=0}^{\infty}\frac{a^i}{i!}=1+a+\sum_{i=2}^{\infty}\frac{a^i}{i!}.$$

$$e^{-1/n}=1-\frac{1}{n}+\sum_{n=2}^{\infty}\frac{(-1)^i}{n^i\;i!}.$$

$$e^{-1/n}-1=\sum_{i=2}^{\infty}\frac{(-1)^i}{n^i\;i!},\quad .$$

$$e^{-1/n}-1=-\frac{1}{n}+\sum_{n=2}^{\infty}\frac{(-1)^n}{n^n\;n!},\quad .$$

Multiplicando por (-1)

$$1-e^{-1/n}=\frac{1}{n}-\sum_{n=2}^{\infty}\frac{(-1)^n}{n^n\;n!},\quad .$$

Multiplicando por (n)

$$n\,\left(1-e^{-1/n}\right)=1-n\,\sum_{n=2}^{\infty}\frac{(-1)^n}{n^n\;n!},\quad .$$

$$n\,\left(1-e^{-1/n}\right)=1-\,\sum_{n=2}^{\infty}\frac{(-1)^n}{n^{n-1}\;n!},\quad .$$

$$\lim_{n\to\infty} n\,\left(e^{-1/n}-1\right)=1,$$

pois

$$\lim_{n\to\infty}\frac{(-1)^n}{n^{n-1}\;n!},=0,\;\;n\geq 2.$$

Vamos calcular $E(T^2)$:

$$E(T^2)=E\left(e^{-2\bar{X}}\right)=M_{\bar{X}}\,(-2)=M_S(-2/n)$$

$$E(T^2)=e^{n\theta(e^{-2/n}-1)}.$$

$$Var(T)=e^{n\theta(e^{-2/n}-1)}-e^{-2\,n\,\theta\;\left(e^{-1/n}-1\right)}.$$

Problemas

10. Na função de verossimilhança $L(p)$ da binomial, suponha que $n = 5$ e $x = 3$. Construa o gráfico da função para os possíveis valores de
$$p = \frac{1}{5}, \frac{2}{5}, \frac{3}{5}, \frac{4}{5},$$
e verifique que o máximo ocorre realmente para $p = \frac{3}{5}$.

11. Observa-se uma sequência de ensaios de Bernoulli, independentes, com parâmetro p, até a ocorrência do primeiro sucesso. Se X indicar o número de ensaios necessários:

(a) Mostre que
$$P(X = x) = p\ (1-p)^{x-1} I_A(x), \;\; a = \{1, 2, \ldots\},$$
(distribuição geométrica).

(b) Repetiu-se esse experimento n vezes e, em cada um deles, o número de ensaios necessários foram $x_1, x_2, \ldots, x_n$. Encontre o EMV para p.

(c) Usando uma moeda, repetiu-se esse experimento 5 vezes, e o número de ensaios necessários até a ocorrência da primeira coroa foi 2, 3, 1, 4, 1, respectivamente. Qual a estimativa de MV para $p =$ probabilidade de ocorrência de coroa nessa moeda?

Existiria outra maneira de estimar p?

12. Suponha que X seja uma v.a. com distribuição normal, com média μ e variância 1. Obtenha o EMV de μ , para uma amostra de tamanho n, $x_1, \ldots, x_n$).

13. Considere Y uma v.a. com distribuição de Poisson, com parâmetro $\lambda > 0$. Obtenha a EMV de λ, baseado numa amostra de tamanho n.

2. Exercícios do Heleno e Mônica:

   1. (Exercício 3.1) Seja $X_1, X_2, \ldots, X_n$ uma amostra aleatória da variável aleatória X com função densidade de probabilidade
   $$f(x|\theta) = \frac{\theta}{x^2}\ I_A(x), \;\; A = [\theta, \infty), \theta > 0.$$
   Encontre o estimador de **MV** de θ e de $g(\theta) = E[1/X]$.

   2. (Exercício 3.2) Seja $X_1, X_2, \ldots, X_n$ uma amostra aleatória da variável aleatória X com função densidade de probabilidade
   $$f(x|\theta) = \theta\ x^{\theta-1}\ I_A(x), \;\; A = (0, 1), \theta > 0.$$
      i. Encontre o estimador de **MV** de θ .
      ii. Encontre o estimador de **MV** de
      $$g(\theta) = E(X) = \frac{\theta}{1+\theta}.$$

   iii. Encontre a distribuição aproximada dos estimadores de **MV** obtidos em **(i)** e **(ii)** quando n é grande.

3. (Exercício 3.3) Seja $X_1, X_2, \ldots, X_n$ uma amostra aleatória da variável aleatória $X \sim N(\mu, 1)$. Encontre o estimador de **MV** de $g(\mu) = P(X > 0)$ e sua distribuição aproximada quando n é grande.

4. (Exercício 3.4) Seja $X_1, X_2, \ldots, X_n$ uma amostra aleatória da variável aleatória X com função densidade de probabilidade

$$f(x|\theta) = \frac{x}{\theta^2}\, e^{-x/\theta}\, I_A(x), \quad A = (0, \infty), \theta > 0.$$

   i. Encontre o estimador de **MV** de θ e verifique que ele é eficiente.
   ii. Encontre o estimador de **MV** de

$$g(\theta) = Var(X)$$

   e sua distribuição aproximada em grandes amostras.

5. (Exercício 3.5) Encontre a distribuição aproximada para grandes amostras do estimador de máxima verossimilhança de

$$g(u) = P(X < 0) = \Phi(-\mu)$$

considerado no exercício 3.3.

6. (Exercício 3.6) Encontre o estimador de **MV** de

$$g(\mu) = \mu^2$$

considerado no exercício 3.3 e compare seu erro médio quadrático com o do estimador eficiente(?)

$$T = \bar{X}^2 - \frac{1}{n}.$$

7. (Exercício 3.7) Considere uma amostra aleatória de tamanho n da distribuição da variável aleatória X onde cada observação apresenta um de três resultados possíveis (por exemplo, favorável, contra e indiferente) que denotamos por "0", "1"e "2". Suponhamos que:

$$P(X = 0) = p_1 = \frac{1-\theta}{2}\ ; \ \ P(X = 1) = p_2 = \frac{1}{2}\ \ ; \ P(X = 2) = p_3 = \frac{\theta}{2}.$$

Seja n_1 o número de vezes que "0";ocorre, seja n_2 o número de vezes que "1";ocorre e seja n_3 o número de vezes que "2"; ocorre.

   i. Encontre, como função de n_1, n_2 e n_3, uma estatística suficiente para θ.
   ii. Encontre o estimador de **MV** de θ.

8. (Exercício 3.8) Seja $X_1, X_2, \ldots, X_n$ uma amostra aleatória da variável aleatória X com função densidade de probabilidade

$$f(x|\theta) = \theta\ (\theta + 1)\ x^{\theta - 1}(1 - \theta)\ I_A(x), \quad A = (0, 1), \theta > 0.$$

i. Encontre, usando o método dos momentos, um estimador para θ .

ii. Encontre o estimador de **MV** de θ
e sua distribuição aproximada para grandes amostras

9. (Exercício 3.10) Seja $X_1, X_2, \ldots, X_n$ uma amostra aleatória da variável aleatória X com função densidade de probabilidade

$$f(x|\theta) = \frac{x+1}{\theta\,(\theta+1)}\, e^{-x/\theta}\, I_A(x), \quad A = (0, \infty), \theta > 0.$$

i. Encontre o estimador de **MV** de θ
e sua distribuição aproximada para grandes amostras.

ii. Encontre, usando o método dos momentos, um estimador para θ.

10. (Exercício 3.11) Refaça o exercício 3.7, supondo agora que

$$p_1 = \theta^2 \quad p_2 = 2\theta(1-\theta) \quad e \quad p_3 = (1-\theta)^2, \quad 0 < \theta < 1.$$

11. (Exercício 3.12) Seja $X_1, X_2, \ldots, X_n$ uma amostra aleatória da variável aleatória $X \sim N(0, \sigma^2)$. Encontre o estimador de **MV** de $g(\sigma) = \sigma$ e sua distribuição aproximada em grandes amostras
.

12. (Exercício 3.13) Seja $X_1, X_2, \ldots, X_n$ uma amostra aleatória da variável aleatória $X \sim Exp(\theta)$.

Encontre o estimador de **MV** de $g(theta) = P(X > 1)$ e sua distribuição aproximada quando n for grande.

13. (Exercício 3.18) Sejam $Y_1, Y_2, \ldots, Y_n$ variáveis aleatórias independentes com

$$Y_i \sim N(\beta_0 + \beta_1 X_i, \sigma^2), \quad i = 1, 2, \ldots, n,$$

com $X_i,\ i = 1, 2, \ldots, n$ conhecidos.

Encontre os estimadores de máxima verossimilhança de

$$\theta = (\beta_0, \beta_1, \sigma^2)'.$$

14. (Exercício 3.19) Sejam $Y_1, Y_2, \ldots, Y_n$ variáveis aleatórias independentes com

$$Y_i \sim N(\beta\, X_i, \sigma^2\, X_i), \quad i = 1, 2, \ldots, n,$$

com $X_i,\ i = 1, 2, \ldots, n$ conhecidos.

Encontre os estimadores de máxima verossimilhança de β e σ^2.

Encontre os estimadores de mínimos quadrados.