

## MAE5911/IME: Fundamentos de Estatística e Machine Learning. Prof.: Alexandre Galvão Patriota

**Questão 01:** Seja  $(Z_1, \dots, Z_n)$  uma amostra aleatória de  $Z \sim \text{Ber}(\theta)$ ,  $\theta \in (0, 1)$ .

- (a) Encontre o EMV para  $g(\theta) = P_\theta(Z = 0)$ .
- (b) Encontre o EMV para  $g(\theta) = \text{Var}_\theta(Z)$ .
- (c) Considere que os dados foram observados  $(0, 0, 1, 0, 0, 1)$ . Encontre as estimativas de MV nos itens acima.
- (d) Construa o valor-p para a hipótese  $H: \theta = 0.1$  usando os dados do item anterior.

Podemos calcular os itens (a) e (b) utilizando a propriedade de invariância do estimador de máxima verossimilhança (EMV).

**Teorema (Invariância do EMV).** Seja  $\hat{\theta}_{MV}$  o estimador de máxima verossimilhança de um parâmetro  $\theta \in \Theta$  e seja  $g: \Theta \rightarrow \mathcal{G}$  uma função. Então o estimador de máxima verossimilhança de

$$\tau = g(\theta)$$

é dado por

$$\hat{\tau}_{MV} = g(\hat{\theta}_{MV}).$$

Isto é, para obter o EMV de qualquer função de  $\theta$ , basta aplicar essa função ao EMV de  $\theta$ .

### Cálculo do EMV de $\theta$ ( $\hat{\theta}_{MV}$ )

#### Função de verossimilhança

Seja  $z_1, \dots, z_n$  uma amostra aleatória de  $Z \sim \text{Bernoulli}(\theta)$ ,  $\theta \in (0, 1)$ . Com:

$$P_\theta(z_i = 1) = \theta, \quad P_\theta(z_i = 0) = 1 - \theta.$$

A função de probabilidade para  $Z = z_i$  é então:

$$f(z_i; \theta) = \theta^{z_i} (1 - \theta)^{1-z_i}.$$

Como  $z_1, \dots, z_n$  são independentes, a probabilidade conjunta, definida como a função de verossimilhança é o produto das densidades individuais:

$$\ell(\theta, z_1, \dots, z_n) = \prod_{i=1}^n \theta^{z_i} (1 - \theta)^{1-z_i}.$$

Utilizando a propriedade de produto de exponenciais, temos:

$$\ell(\theta, z_1, \dots, z_n) = \theta^{\sum_{i=1}^n z_i} (1 - \theta)^{\sum_{i=1}^n (1-z_i)} = \theta^{\sum_{i=1}^n z_i} (1 - \theta)^{n - \sum_{i=1}^n z_i}.$$

Seja

$$S = \sum_{i=1}^n z_i,$$

Podemos escrever:

$$\ell(\theta, S) = \theta^S (1 - \theta)^{n-S}.$$

## Função Log-verossimilhança

Podemos simplificar os cálculos trabalhando com o logaritmo desta função, mantendo o estimador de máxima verossimilhança inalterado, uma vez que este trata-se do ponto crítico desta função. Maximizar  $\ell(\theta)$  ou  $\mathcal{L}(\theta) = \log \ell(\theta)$  é equivalente, devido ao fato do logaritmo ser estritamente crescente. Portanto, aplicando o logaritmo, simplificamos para:

$$\mathcal{L}(\theta, S) = \log \ell(\theta, S) = \log (\theta^S (1 - \theta)^{n-S}) = S \log \theta + (n - S) \log(1 - \theta).$$

## Ponto crítico da log-verossimilhança

Calculando a derivada de  $\mathcal{L}(\theta, S)$  em relação a  $\theta$ :

$$\frac{\partial \mathcal{L}(\theta, S)}{\partial \theta} = \frac{\partial}{\partial \theta} [S \log \theta + (n - S) \log(1 - \theta)] = S \cdot \frac{1}{\theta} + (n - S) \cdot \left( -\frac{1}{1 - \theta} \right) = \frac{S}{\theta} - \frac{n - S}{1 - \theta}.$$

Seja a condição de máximo onde está definido  $\hat{\theta}_{MV}$ :

$$\frac{\partial \mathcal{L}(\theta, S)}{\partial \theta} = 0 \iff \frac{S}{\hat{\theta}_{MV}} - \frac{n - S}{1 - \hat{\theta}_{MV}} = 0.$$

$$\frac{S}{\hat{\theta}_{MV}} = \frac{n - S}{1 - \hat{\theta}_{MV}}.$$

$$S(1 - \hat{\theta}_{MV}) = \hat{\theta}_{MV}(n - S).$$

$$S - S\hat{\theta}_{MV} = n\hat{\theta}_{MV} - S\hat{\theta}_{MV}.$$

Os termos  $-S\hat{\theta}_{MV}$  cancelam em ambos os lados, restando:

$$S = n\hat{\theta}_{MV}.$$

Isolando  $\hat{\theta}_{MV}$ :

$$\boxed{\hat{\theta}_{MV} = \frac{S}{n}}.$$

## Estimador de máxima verossimilhança

Substituindo de volta em função da variável aleatória  $z_i$ , obtemos:

$$\hat{\theta}_{MV} = \frac{1}{n} \sum_{i=1}^n z_i = \bar{Z}.$$

## Teste de hipótese

O teste de hipótese é formulado como um framework de decisão. No processo de decisão existem dois possíveis cenários de enganos, sendo que ambos não podem ser reduzidos simultaneamente, a menos do aumento da amostra, e se um cenário de engano diminui o seu complementar aumenta. Escolhe-se por arbitrariedade priorizar a mitigação do erro **falso positivo**, em detrimento do erro **falso negativo**. Desta forma a hipótese nula é definida como a hipótese que se deseja refutar. Entretanto, o teste é construído para que haja apenas uma baixa probabilidade, comumente 5%, de que  $H_0$  seja refutada, ou seja, a sua rejeição é definida como um evento raro. Consequentemente, quando  $H_0$  rejeitada, considera-se que o processo de decisão está com uma margem de erro suficientemente reduzida, de acordo com o nível de significância escolhido, que reflete a magnitude do erro tolerado.

Deste framework, o valor-p é a métrica que caracteriza a amostra observada no cenário em que a hipótese nula é verdadeira. O valor-p então calculado é comparado com o nível de significância escolhido para que a decisão de rejeitar ou não a hipótese nula seja tomada.

(a)  $g(\theta) = P_\theta(Z = 0)$

$$g(\theta) = P_\theta(Z = 0) = 1 - \theta.$$

Pelo teorema da **invariância do EMV**:

$$\hat{g}_{MV} = g(\hat{\theta}_{MV}) = 1 - \hat{\theta}_{MV}.$$

Já encontramos que  $\hat{\theta}_{MV} = \bar{Z}$ , então:

$$\boxed{\hat{g}_{MV} = 1 - \bar{Z}}.$$

(b)  $g(\theta) = \text{Var}_\theta(Z)$

A variância é definida por:

$$\text{Var}(Z) = E(Z^2) - E(Z)^2.$$

Pela definição de esperança para variáveis discretas,

$$E(Z) = \sum_{z \in \{0,1\}} z P(Z = z).$$

Mas  $P(Z = 1) = \theta$  e  $P(Z = 0) = 1 - \theta$ , então:

$$E(Z) = 0 \cdot (1 - \theta) + 1 \cdot \theta = \theta.$$

$E(Z^2) = E(Z) = \theta$  (pois  $Z \in \{0,1\}$ ), então:

$$\text{Var}_\theta(Z) = \theta - \theta^2 = \theta(1 - \theta).$$

Portanto:

$$g(\theta) = \text{Var}_\theta(Z) = \theta(1 - \theta).$$

Pela **invariância do EMV**:

$$\hat{g}_{MV} = g(\hat{\theta}_{MV}) = \hat{\theta}_{MV}(1 - \hat{\theta}_{MV}).$$

Para  $\hat{\theta}_{MV} = \bar{Z}$ , encontramos:

$$\boxed{\hat{g}_{MV} = \bar{Z}(1 - \bar{Z})}.$$

(c) **Sejam os dados observados:** (0, 0, 1, 0, 0, 1)

Para o cálculo da média:  $n = 6$  e  $\sum z_i = 2$ , portanto

$$\hat{\theta}_{MV} = \bar{Z} = \frac{2}{6} = \frac{1}{3}.$$

Assim:

$$(a) \quad \hat{g}_{MV} = \widehat{P_\theta(Z = 0)} = 1 - \bar{Z} = \frac{2}{3}$$

$$(b) \quad \hat{g}_{MV} = \widehat{\text{Var}_\theta(Z)} = \bar{Z}(1 - \bar{Z}) = \frac{1}{3} \cdot \frac{2}{3} = \frac{2}{9}$$

**(d) Teste de hipótese  $H_0 : \theta = 0.1$**

Na distribuição de Bernoulli, o parâmetro  $\theta$  representa a probabilidade de sucesso, equivalente à média. Se supormos que a hipótese nula é verdadeira, o valor-p dos dados observados em relação a  $H_0$  é:

Dados amostrados:

$$(0, 0, 1, 0, 0, 1)$$

Sob a hipótese nula  $\theta = 0.1$ , construímos  $X$ :

$$X \sim \text{Binomial}(n = 6, \theta = 0.1).$$

A estatística de interesse é a média amostral ou probabilidade de sucessos. Em relação à amostra observada, é:

$$\bar{Z} = \frac{1}{n} \sum_{i=1}^n z_i = \frac{1}{6} \sum_{i=1}^6 z_i = \frac{1}{6} * 2 = \frac{1}{3}.$$

Vamos analisar a média da amostra contra a média testada (a média da hipótese nula). O valor-p então é definido a partir de: no cenário em que  $H_0$  é válida, ou seja, temos uma distribuição de Bernoulli com média = 0.1, qual a probabilidade de observar média  $\bar{Z} = 1/3$  ou maior? O posicionamento da média observada em relação à média testada indica a direção do teste.

$$p = P_{H_0} \left( \bar{Z} \geq \frac{1}{3} \right).$$

Como  $X = \sum_{i=1}^6 z_i$ ,  $\bar{Z} = \frac{1}{n} \sum_{i=1}^n z_i$  e  $n = 6$ :

$$\bar{Z} = \frac{X}{6}.$$

Então,

$$\bar{Z} \geq \frac{1}{3} \iff \frac{X}{6} \geq \frac{1}{3} \iff X \geq 2.$$

Logo, o valor- $p$  é:

$$p = P_{H_0}(X \geq 2),$$

$$p = P(X = 2) + P(X = 3) + P(X = 4) + P(X = 5) + P(X = 6).$$

Cada termo é dado por:

$$P(X = k) = \binom{6}{k} (0.1)^k (0.9)^{6-k}.$$

Portanto:

$$P(X = 2) = \binom{6}{2} (0.1)^2 (0.9)^4 = 15 \cdot 0.01 \cdot 0.6561 = 0.098415.$$

$$P(X = 3) = \binom{6}{3} (0.1)^3 (0.9)^3 = 20 \cdot 0.001 \cdot 0.729 = 0.01458.$$

$$P(X = 4) = \binom{6}{4}(0.1)^4(0.9)^2 = 15 \cdot 0.0001 \cdot 0.81 = 0.001215.$$

$$P(X = 5) = \binom{6}{5}(0.1)^5(0.9) = 6 \cdot 0.00001 \cdot 0.9 = 0.000054.$$

$$P(X = 6) = \binom{6}{6}(0.1)^6 = 1 \cdot 0.000001 = 0.000001.$$

Assim, o valor-p é:

$$p = 0.098415 + 0.01458 + 0.001215 + 0.000054 + 0.000001 = 0.114265.$$

Pela conservação da probabilidade, seria mais simples calcular o valor-p pelo seu valor complementar:

$$\bar{p} = P_{H0}(X < 2) = (1 - p)$$

$$P(X = 0) = (0.9)^6 = 0.531441, \quad P(X = 1) = \binom{6}{1}(0.1)(0.9)^5 = 0.354294.$$

Logo:

$$\bar{p} = 1 - p = P(X = 0) + P(X = 1) = 0.531441 + 0.354294 = 0.885735.$$

$\text{valor-p} = 1 - 0.885735 = 0.114265$

Considerando um nível de significância  $\alpha = 0.05$  não há evidências suficientes para rejeitar  $H_0$ , isto é, para supor que a média seja diferente de 0.1.

**Questão 02:** Seja  $(Z_1, \dots, Z_n)$  uma amostra aleatória de  $Z \sim \text{Exp}(\theta)$ ,  $\theta \in (0, \infty)$ .

- (a) Encontre o EMV para  $g(\theta) = P_\theta(Z > 1)$ .
- (b) Encontre o EMV para  $g(\theta) = P_\theta(0.1 < Z < 1)$ .
- (c) Encontre o EMV para  $g(\theta) = \text{Var}_\theta(Z)$ .
- (d) Considere que os dados foram observados (0.2, 0.6, 0.3, 0.2, 0.8, 0.12). Encontre um IC aproximado de 95% de confiança para  $g(\theta)$  nos itens acima.
- (e) Faça uma simulação de Monte Carlo para verificar se os IC's aproximados obtidos no passo anterior têm cobertura próxima do nível de confiança estabelecido. Caso não tenham, proponha um tamanho amostral que produza IC's mais confiáveis para cada caso.

Podemos calcular os itens (a), (b) e (c) utilizando a propriedade de invariância do estimador de máxima verossimilhança (EMV), calculando uma única vez  $\hat{\theta}_{MV}$ .

### Cálculo do EMV de $\theta$ ( $\hat{\theta}_{MV}$ ) para a distribuição Exponencial

#### Função de verossimilhança

Seja  $(z_1, \dots, z_n)$  uma amostra aleatória de  $Z \sim \text{Exp}(\theta)$ , com  $\theta \in (0, \infty)$ , a função densidade de probabilidade para  $z_i$  é então:

$$f(z_i; \theta) = \theta e^{-\theta z_i}, \quad z_i > 0.$$

Como  $z_1, \dots, z_n$  são independentes, a probabilidade conjunta, definida como a função de verossimilhança é o

produto das densidades individuais:

$$\ell(\theta; z_1, \dots, z_n) = \prod_{i=1}^n f(z_i | \theta) = \prod_{i=1}^n \theta e^{-\theta z_i} \mathbf{1}_{\{z_i > 0\}}.$$

O termo indicador  $\prod_{i=1}^n \mathbf{1}_{\{z_i > 0\}}$  não depende de  $\theta$ , logo não influencia a maximização e pode ser ignorado. Assim,

$$\ell(\theta; z_1, \dots, z_n) = \prod_{i=1}^n \theta e^{-\theta z_i}.$$

Aplicando a propriedade distributiva do produto e de produto de exponenciais, obtemos:

$$\ell(\theta; z_1, \dots, z_n) = \left( \prod_{i=1}^n \theta \right) \left( \prod_{i=1}^n e^{-\theta z_i} \right) = \theta^n \exp \left( -\theta \sum_{i=1}^n z_i \right).$$

Seja

$$S = \sum_{i=1}^n z_i,$$

Podemos escrever:

$$\ell(\theta, S) = \theta^n \exp(-\theta S).$$

### Função log-verossimilhança

Para simplificar os cálculos, trabalhamos com a função logaritmo da verossimilhança  $\mathcal{L}(\theta, S)$ :

$$\mathcal{L}(\theta, S) = \log \ell(\theta, S) = \log(\theta^n) + \log(\exp(-\theta S)).$$

Usando as propriedades  $\log(a^b) = b \log a$  e  $\log(e^x) = x$ , obtemos

$$\mathcal{L}(\theta, S) = n \log \theta - \theta S.$$

### Ponto crítico da log-verossimilhança

Calculando a derivada de  $\mathcal{L}(\theta, S)$  em relação a  $\theta$ :

$$\frac{\partial \mathcal{L}(\theta, S)}{\partial \theta} = \frac{\partial}{\partial \theta} [n \log \theta - \theta S] = n \cdot \frac{1}{\theta} - S \cdot 1 = \frac{n}{\theta} - S.$$

Seja a condição de máximo onde está definido  $\hat{\theta}_{MV}$ :

$$\frac{\partial \mathcal{L}(\theta, S)}{\partial \theta} = 0 \iff \frac{n}{\hat{\theta}_{MV}} - S = 0 \iff \frac{n}{\hat{\theta}_{MV}} = S.$$

Isolando  $\hat{\theta}_{MV}$ :

$$\boxed{\hat{\theta}_{MV} = \frac{n}{S}},$$

### Estimador de máxima verossimilhança

Substituindo de volta em função da variável aleatória  $z_i$ , obtemos:

$$\hat{\theta}_{MV} = \frac{n}{\sum_{i=1}^n z_i} = \frac{1}{\bar{Z}},$$

## Intervalo de confiança

Conhecendo a **distribuição de probabilidade de  $\hat{\theta}$**  é possível construir um intervalo

$$\hat{\theta}_1 \leq \theta \leq \hat{\theta}_2$$

que contém  $\theta$ . Esta técnica diferencia-se da estimação "por ponto", onde se calcula um único valor para o parâmetro populacional. No caso do intervalo de confiança busca-se um seguimento, ou intervalo  $\hat{\theta}_1 : \hat{\theta}_2$  que contém o parâmetro desconhecido (FONSECA, 2008, p. 186).

(a)  $g(\theta) = P_{\theta}(Z > 1)$

Integrando a função densidade de probabilidade  $f(z_i, \theta)$  para encontrar  $P_{\theta}(Z > 1)$ :

$$g(\theta) = P_{\theta}(Z > 1) = \int_1^{\infty} f(z, \theta) dz = \int_1^{\infty} \theta e^{-\theta z} dz.$$

Calculando a integral:

$$\int_1^{\infty} \theta e^{-\theta z} dz = \left[ -e^{-\theta z} \right]_{z=1}^{\infty} = 0 - (-e^{-\theta}) = e^{-\theta}.$$

Portanto,

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

Pela **invariância do EMV**, o estimador de máxima verossimilhança de  $g(\theta)$  é

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

Já encontramos que  $\hat{\theta}_{MV} = \frac{1}{Z}$  para distribuição Exponencial, então:

$$\hat{g}_{MV} = e^{-\hat{\theta}_{MV}} = \exp\left(-\frac{1}{Z}\right).$$

(b)  $g(\theta) = P_{\theta}(0.1 < Z < 1)$

Integrando a função densidade de probabilidade  $f(z_i, \theta)$  para encontrar  $P_{\theta}(0.1 < Z < 1)$ :

$$g(\theta) = P_{\theta}(0.1 < Z < 1) = \int_{0.1}^1 f(z, \theta) dz = \int_{0.1}^1 \theta e^{-\theta z} dz.$$

Calculando a integral:

$$\int_{0.1}^1 \theta e^{-\theta z} dz = \left[ -e^{-\theta z} \right]_{z=0.1}^1 = -e^{-\theta} - (-e^{-0.1\theta}) = e^{-0.1\theta} - e^{-\theta}.$$

Logo,

$$g(\theta) = e^{-0.1\theta} - e^{-\theta}.$$

Pela **invariância do EMV**,

$$\hat{g}_{MV} = g(\hat{\theta}_{MV}) = e^{-0.1\hat{\theta}_{MV}} - e^{-\hat{\theta}_{MV}}.$$

Substituindo  $\hat{\theta}_{MV} = \frac{1}{\bar{Z}}$ , obtemos:

$$\hat{g}_{MV} = \exp\left(-\frac{0.1}{\bar{Z}}\right) - \exp\left(-\frac{1}{\bar{Z}}\right).$$

(c)  $g(\theta) = \text{Var}_{\theta}(Z)$

A variância é definida por:

$$\text{Var}(Z) = E_{\theta}(Z^2) - (E_{\theta}(Z))^2.$$

Pela definição de esperança para variáveis contínuas, sendo  $h(z)$  a função de interesse:

$$E_{\theta}(h(Z)) = \int_{-\infty}^{\infty} h(z) f(z; \theta) dz.$$

Em particular, a esperança do primeiro e segundo momentos de  $Z$ , são:

$$E_{\theta}(Z) = \int_0^{\infty} z \theta e^{-\theta z} dz = [-ze^{-\theta z}]_0^{\infty} + \int_0^{\infty} e^{-\theta z} dz = 0 + \left[-\frac{1}{\theta} e^{-\theta z}\right]_0^{\infty} = 0 + \left(0 - \left(-\frac{1}{\theta}\right)\right) = \frac{1}{\theta}.$$

e

$$E_{\theta}(Z^2) = \int_0^{\infty} z^2 \theta e^{-\theta z} dz = \int_0^{\infty} \left(\frac{y}{\theta}\right)^2 \theta e^{-y} \frac{dy}{\theta} = \frac{1}{\theta^2} \int_0^{\infty} y^2 e^{-y} dy = \frac{1}{\theta^2} \cdot 2! = \frac{2}{\theta^2}.$$

Então:

$$\text{Var}_{\theta}(Z) = E(Z^2) - [E(Z)]^2 = \frac{2}{\theta^2} - \left(\frac{1}{\theta}\right)^2 = \frac{1}{\theta^2}.$$

Portanto:

$$g(\theta) = \text{Var}_{\theta}(Z) = \frac{1}{\theta^2}.$$

Pela **invariância do EMV**,

$$\hat{g}_{MV} = g(\hat{\theta}_{MV}) = \frac{1}{\hat{\theta}_{MV}^2}.$$

Substituindo  $\hat{\theta}_{MV} = \frac{1}{\bar{Z}}$ , obtemos:

$$\hat{g}_{MV} = \bar{Z}^2.$$

(d) **Sejam os dados observados:** (0,2, 0,6, 0,3, 0,2, 0,8, 0,12), **encontrar IC para itens (a), (b) e (c)**

Para a distribuição exponencial é possível encontrar o intervalo de confiança exato para  $\theta$  utilizando  $\chi^2$ , já que

$$S \sim \text{Gamma}(n, \theta),$$

E portanto podemos expressar o intervalo de confiança em termos de  $\chi_{2n}^2$ :

$$2\theta S \sim \chi_{2n}^2.$$

Chegando a:

$$P\left(\chi_{2n, \alpha/2}^2 \leq 2\theta S \leq \chi_{2n, 1-\alpha/2}^2\right) = 1 - \alpha.$$



Dividindo tudo por  $2S > 0$ :

$$P\left(\frac{\chi_{2n,\alpha/2}^2}{2S} \leq \theta \leq \frac{\chi_{2n,1-\alpha/2}^2}{2S}\right) = 1 - \alpha.$$

Mas vamos utilizar o método aproximado que é útil na maior parte das aplicações onde a distribuição não terá possibilidade de cálculo do intervalo de confiança exato. Segue o desenvolvimento assintótico do EMV para a distribuição Exponencial.

## Distribuição aproximada de $\hat{\theta}_{MV}$

Para um modelo paramétrico com densidade  $f(z, \theta)$ , a informação de Fisher é:

$$I(\theta) = \mathbb{E}_\theta \left[ \left( \frac{\partial}{\partial \theta} \log f(z, \theta) \right)^2 \right] = -\mathbb{E}_\theta \left[ \frac{\partial^2}{\partial \theta^2} \log f(z, \theta) \right].$$

A segunda derivada da função log-verossimilhança  $\mathcal{L}(\theta, S)$  é

$$-\mathcal{L}''(\theta, S) = \frac{n}{\theta^2},$$

logo a Informação de Fisher é

$$I(\theta) = \frac{n}{\theta^2}.$$

Assintoticamente:

$$\hat{\theta}_{MV} \approx N\left(\theta, \frac{1}{nI(\theta)}\right) \iff \hat{\theta}_{MV} \approx N\left(\theta, \frac{\theta^2}{n}\right) \iff \frac{\hat{\theta}_{MV} - \theta}{\theta/\sqrt{n}} \approx N(0, 1).$$

Como  $\theta$  é desconhecido, podemos utilizar o valor de Máxima Verossimilhança  $\hat{\theta}_{MV}$  que é um parâmetro consistente e se aproxima de  $\theta$  quando  $n$  é grande. Então substituímos  $\theta$  por  $\hat{\theta}_{MV}$  no denominador:

$$\frac{\hat{\theta}_{MV} - \theta}{\hat{\theta}_{MV}/\sqrt{n}} \approx N(0, 1).$$

Para um nível de confiança  $1 - \alpha$ , o intervalo de confiança da Normal Padrão  $N(0, 1)$  é:

$$P\left(-z_{1-\alpha/2} \leq \frac{\hat{\theta}_{MV} - \theta}{\hat{\theta}_{MV}/\sqrt{n}} \leq z_{1-\alpha/2}\right) \approx 1 - \alpha.$$

$$P\left(-z_{1-\alpha/2} \frac{\hat{\theta}_{MV}}{\sqrt{n}} \leq \hat{\theta}_{MV} - \theta \leq z_{1-\alpha/2} \frac{\hat{\theta}_{MV}}{\sqrt{n}}\right) \approx 1 - \alpha.$$

$$P\left(-z_{1-\alpha/2} \frac{\hat{\theta}_{MV}}{\sqrt{n}} - \hat{\theta}_{MV} \leq -\theta \leq z_{1-\alpha/2} \frac{\hat{\theta}_{MV}}{\sqrt{n}} - \hat{\theta}_{MV}\right) \approx 1 - \alpha.$$

$$P\left(\hat{\theta}_{MV} - z_{1-\alpha/2} \frac{\hat{\theta}_{MV}}{\sqrt{n}} \leq \theta \leq \hat{\theta}_{MV} + z_{1-\alpha/2} \frac{\hat{\theta}_{MV}}{\sqrt{n}}\right) \approx 1 - \alpha.$$

$$P\left(\hat{\theta}_{MV} \left(1 - \frac{z_{1-\alpha/2}}{\sqrt{n}}\right) \leq \theta \leq \hat{\theta}_{MV} \left(1 + \frac{z_{1-\alpha/2}}{\sqrt{n}}\right)\right) \approx 1 - \alpha.$$

$$\boxed{\theta \in \left[ \hat{\theta}_{MV} \left(1 - \frac{z_{1-\alpha/2}}{\sqrt{n}}\right), \hat{\theta}_{MV} \left(1 + \frac{z_{1-\alpha/2}}{\sqrt{n}}\right) \right]}$$

É útil nomear os limites inferior e superior para referência:

$$L_\theta = \hat{\theta}_{MV} \left( 1 - \frac{z_{1-\alpha/2}}{\sqrt{n}} \right), \quad U_\theta = \hat{\theta}_{MV} \left( 1 + \frac{z_{1-\alpha/2}}{\sqrt{n}} \right).$$

**Cálculo do intervalo aproximado para o Item (a):**  $g(\theta) = P_\theta(Z > 1)$

Para a Exponencial, temos:

$$g(\theta) = P_\theta(Z > 1) = \int_1^\infty \theta e^{-\theta z} dz = e^{-\theta}.$$

A função  $g(\theta) = e^{-\theta}$  é estritamente decrescente em  $\theta$ . Para  $L_\theta \leq \theta \leq U_\theta$ , como  $g$  é decrescente,

$$g(L_\theta) \geq g(\theta) \geq g(U_\theta).$$

Substituímos  $g(x) = e^{-x}$ :

$$e^{-L_\theta} \geq e^{-\theta} \geq e^{-U_\theta} \iff e^{-U_\theta} \leq g(\theta) \leq e^{-L_\theta}.$$

Logo, o IC aproximado para  $g(\theta)$  é:

$$\boxed{g(\theta) \in [e^{-U_\theta}, e^{-L_\theta}].}$$

Para os dados amostrados:

$$(0.2, 0.6, 0.3, 0.2, 0.8, 0.12).$$

$$n = 6, \quad S = \sum z_i = 2.22, \quad \bar{Z} = \frac{S}{6} = 0.37.$$

$$\hat{\theta}_{MV} = \frac{1}{\bar{Z}} = \frac{1}{0.37} \approx 2.703.$$

Com  $z_{0.975} = 1.96$  e  $n = 6$ :

$$\frac{z_{0.975}}{\sqrt{n}} = \frac{1.96}{\sqrt{6}} \approx 0.800.$$

$$L_\theta = 2.703(1 - 0.800) \approx 0.541, \quad U_\theta = 2.703(1 + 0.800) \approx 4.864.$$

Portanto, o IC aproximado de 95% para  $\theta$  é:

$$\theta \in [0.541, 4.864]$$

$$g_{\min} = e^{-U_\theta} = e^{-4.864} \approx 0.0077, \quad g_{\max} = e^{-L_\theta} = e^{-0.541} \approx 0.583.$$

$$\boxed{g(\theta) = P_\theta(Z > 1) \in [0.0077, 0.583].}$$

**Cálculo do intervalo aproximado para o Item (b):**  $g(\theta) = P_\theta(0,1 < Z < 1)$

Para a Exponencial, temos

$$f(z; \theta) = \theta e^{-\theta z}, \quad z > 0, \theta > 0.$$

Então

$$g(\theta) = P_\theta(0,1 < Z < 1) = \int_{0.1}^1 \theta e^{-\theta z} dz = \left[ -e^{-\theta z} \right]_{z=0.1}^{z=1} = e^{-0.1\theta} - e^{-\theta}.$$

Do item (a), usando a teoria assintótica do EMV e a Informação de Fisher (obtendo  $\hat{\theta}_{MV} \approx N(\theta, \theta^2/n)$ ), para um nível de confiança  $1 - \alpha$  temos

$$P\left(-z_{1-\alpha/2} \leq \frac{\hat{\theta}_{MV} - \theta}{\hat{\theta}_{MV}/\sqrt{n}} \leq z_{1-\alpha/2}\right) \approx 1 - \alpha.$$

Manipulando a inequação, obtemos o IC aproximado

$$\theta \in \left[ \hat{\theta}_{MV} \left(1 - \frac{z_{1-\alpha/2}}{\sqrt{n}}\right), \hat{\theta}_{MV} \left(1 + \frac{z_{1-\alpha/2}}{\sqrt{n}}\right) \right] =: [L_\theta, U_\theta].$$

**Análise de  $g(\theta)$**  . Derivando,

$$g'(\theta) = -0,1 e^{-0.1\theta} + e^{-\theta}.$$

O ponto crítico é dado por  $g'(\theta^*) = 0$ :

$$-0,1 e^{-0.1\theta^*} + e^{-\theta^*} = 0 \iff ; e^{-\theta^*} = 0,1 e^{-0.1\theta^*} \iff ; e^{-0.9\theta^*} = 0,1 \iff \theta^* = \frac{\log 10}{0.9}.$$

Logo,  $g(\theta)$  cresce em  $0 \leq \theta \leq \theta^*$  e decresce em  $\theta \geq \theta^*$  (único máximo em  $\theta^*$ ).

**Intervalo para  $g(\theta)$  a partir do IC de  $\theta$**

Dado o IC  $[L_\theta, U_\theta]$  para  $\theta$ , o intervalo correspondente para  $g(\theta)$  é o intervalo que contém todos os valores

$$\{g(\theta) : \theta \in [L_\theta, U_\theta]\}.$$

Como  $g$  é crescente até  $\theta^*$  e decrescente depois, se  $\theta^* \in [L_\theta, U_\theta]$  temos

$$g_{\min} = \min\{g(L_\theta), g(U_\theta)\}, \quad g_{\max} = g(\theta^*),$$

e, portanto,

$$g(\theta) \in [g_{\min}, g_{\max}].$$

Os dados observados são (0,2, 0,6, 0,3, 0,2, 0,8, 0,12), para os quais já havíamos obtido, no item (a):

$$n = 6, \quad \hat{\theta}_{MV} = \frac{1}{\bar{Z}} \approx 2,703,$$

e, com  $z_{0.975} = 1,96$ ,

$$L_\theta = \hat{\theta}_{MV} \left(1 - \frac{z_{0.975}}{\sqrt{n}}\right) \approx 0,541, \quad U_\theta = \hat{\theta}_{MV} \left(1 + \frac{z_{0.975}}{\sqrt{n}}\right) \approx 4,864.$$

O ponto de máximo é

$$\theta^* = \frac{\log 10}{0.9} \approx 2,558,$$

que de fato pertence ao intervalo  $[L_\theta, U_\theta]$ .

Calculando:

$$g(L_\theta) = e^{-0.1L_\theta} - e^{-L_\theta} \approx 0,365,$$

$$g(U_\theta) = e^{-0.1U_\theta} - e^{-U_\theta} \approx 0,607,$$

$$g(\theta^*) = e^{-0.1\theta^*} - e^{-\theta^*} \approx 0,697.$$

Logo,

$$g_{\min} = \min\{g(L_\theta), g(U_\theta)\} \approx 0,365, \quad g_{\max} = g(\theta^*) \approx 0,697.$$

Portanto, o **intervalo de confiança aproximado de 95%** para

$$g(\theta) = P_\theta(0,1 < Z < 1)$$

é

$$g(\theta) = P_\theta(0,1 < Z < 1) \in [0,365, 0,697].$$

**Cálculo do intervalo aproximado para o Item (c):**  $g(\theta) = \text{Var}\theta(Z)$

Para  $Z \sim \text{Exp}(\theta)$  com  $\theta > 0$ , sabe-se que

$$g(\theta) = \text{Var}\theta(Z) = \frac{1}{\theta^2}.$$

Na letra (d) queremos um IC aproximado de 95% para  $g(\theta)$  usando os dados

$$(0,2, 0,6, 0,3, 0,2, 0,8, 0,12).$$

Primeiro, relembremos o IC aproximado para  $\theta$  obtido pela assintótica Normal do EMV.

A média amostral é

$$\bar{Z} = \frac{1}{6} \sum_{i=1}^6 z_i = \frac{2,22}{6} \approx 0,37,$$

e o EMV de  $\theta$  é

$$\hat{\theta}_{MV} = \frac{1}{\bar{Z}} \approx \frac{1}{0,37} \approx 2,703.$$

Pela teoria assintótica do EMV,

$$\hat{\theta}_{MV} \approx N\left(\theta, \frac{\theta^2}{n}\right),$$

e, substituindo  $\theta$  por  $\hat{\theta}_{MV}$  no desvio-padrão, obtemos o IC aproximado de 95% para  $\theta$ :

$$P\left(-z_{0,975} \leq \frac{\hat{\theta}_{MV} - \theta}{\hat{\theta}_{MV}/\sqrt{n}} \leq z_{0,975}\right) \approx 0,95,$$

o que é equivalente a

$$\theta \in \left[\hat{\theta}_{MV}\left(1 - \frac{z_{0,975}}{\sqrt{n}}\right), \hat{\theta}_{MV}\left(1 + \frac{z_{0,975}}{\sqrt{n}}\right)\right].$$

Com  $z_{0,975} = 1,96$  e  $n = 6$ ,

$$\frac{z_{0,975}}{\sqrt{n}} = \frac{1,96}{\sqrt{6}} \approx 0,800,$$

logo

$$L_\theta = \hat{\theta}_{MV}(1 - 0,800) \approx 2,703 \times 0,200 \approx 0,541,$$

$$U_\theta = \hat{\theta}_{MV}(1 + 0,800) \approx 2,703 \times 1,800 \approx 4,864.$$

Portanto, o IC aproximado de 95% para  $\theta$  é

$$\theta \in [0,541, 4,864].$$

Agora transformamos esse intervalo para  $g(\theta) = \text{Var}\theta(Z) = 1/\theta^2$ . Analisando  $g(\theta)$ :

$$g(\theta) = \frac{1}{\theta^2}, \quad \theta > 0,$$

é estritamente decrescente em  $\theta$ . Assim, se

$$L_\theta \leq \theta \leq U_\theta,$$

então, aplicando  $g$  e invertendo a ordem das desigualdades,

$$g(U_\theta) \leq g(\theta) \leq g(L_\theta),$$

isto é,

$$\frac{1}{U_\theta^2} \leq \text{Var}\theta(Z) \leq \frac{1}{L_\theta^2}.$$

Logo, um IC aproximado de 95% para  $g(\theta) = \text{Var}\theta(Z)$  é

$$g(\theta) = \text{Var}\theta(Z) \in \left[ \frac{1}{U_\theta^2}, \frac{1}{L_\theta^2} \right].$$

Substituindo os valores numéricos  $L_\theta \approx 0,541$  e  $U_\theta \approx 4,864$ ,

$$\frac{1}{U_\theta^2} \approx \frac{1}{4,864^2} \approx 0,042, \quad \frac{1}{L_\theta^2} \approx \frac{1}{0,541^2} \approx 3,42.$$

Portanto, o intervalo de confiança aproximado de 95% para a variância é

$$\text{Var}\theta(Z) \in [0,042, 3,42].$$

### (e) Teste de confiança do IC's via simulação Monte Carlo

- (a) No experimento de Monte Carlo associado ao Listing 1, com  $M = 1000$  simulações, a taxa de cobertura obtida foi 0.962. Assim, o método utilizado para encontrar o intervalo aproximado para  $g(\theta) = e^{-\theta}$  cobriu o valor verdadeiro em 96.2% das repetições, desempenho compatível com o nível nominal de 95% mesmo para  $n = 6$ .
- (b) Para o intervalo construído no Listing 2, a cobertura obtida foi 0.981. O intervalo aproximado para  $g(\theta) = P_\theta(0.1 < Z < 1)$  apresentou desempenho compatível com a cobertura de 95%, mesmo para amostra pequena.
- (c) No experimento correspondente ao Listing 3, referente a  $g(\theta) = \text{Var}_\theta(Z) = 1/\theta^2$ , a taxa de cobertura obtida foi 0.962. O método aproximado forneceu um intervalo com desempenho acima do nível nominal de 95%, mesmo sob amostra pequena.

```
1 set.seed(123)
2
3 # Parâmetros do experimento
4 theta0 <- 2.703 # "verdadeiro" theta
5 n <- 6 # tamanho da amostra
6 M <- 1000 # número de simulações
7 z_0975 <- 1.96 # quantil 0.975 da N(0,1)
8
9 # valor verdadeiro de g(theta) = P(Z > 1)
10 g_true <- exp(-theta0)
11
12 # vetor para guardar se o IC cobriu ou não
13 covered <- logical(M)
14
15 for (m in 1:M) {
16   # 1) cria uma amostra com distribuição Exponencial(theta0)
```

```

17 z <- rexp(n, theta0)
18
19 # 2) EMV de theta: 1 / media
20 theta_hat <- 1 / mean(z)
21
22 # 3) IC aproximado para theta
23 L_theta <- theta_hat * (1 - z_0975 / sqrt(n))
24 U_theta <- theta_hat * (1 + z_0975 / sqrt(n))
25
26 # 4) transforma para g(theta) = e^{-theta}
27 L_g <- exp(-U_theta)
28 U_g <- exp(-L_theta)
29
30 # 5) verifica se g_true está no intervalo
31 covered[m] <- (L_g <= g_true) && (g_true <= U_g)
32 }
33
34 # Estimativa da cobertura
35 mean(covered)

```

Listing 1: Simulação de Monte Carlo para testar o intervalo de confiança aproximado obtido para o item (a),  $g(\theta) = P_{\theta}(Z > 1)$ .

```

1 set.seed(123)
2
3 # Parâmetros do experimento
4 theta0 <- 2.703 # "verdadeiro" theta (mesmo do item a)
5 n <- 6 # tamanho da amostra
6 M <- 10000 # número de simulações
7 z_0975 <- 1.96 # quantil 0.975 da N(0,1)
8
9 # Função g(theta) do item (b):
10 # g(theta) = P(0.1 < Z < 1) = exp(-0.1*theta) - exp(-theta)
11 g_fun <- function(theta) exp(-0.1 * theta) - exp(-theta)
12 g_true <- g_fun(theta0)
13
14 # Ponto onde g(theta) atinge seu máximo
15 theta_star <- log(10) / 0.9
16
17 covered <- logical(M)
18
19 for (m in 1:M) {
20
21 # 1) Amostra Z_1, ..., Z_n ~ Exponencial(theta0)
22 z <- rexp(n, rate = theta0)
23
24 # 2) EMV de theta: 1 / media
25 theta_hat <- 1 / mean(z)
26
27 # 3) IC aproximado para theta
28 L_theta <- theta_hat * (1 - z_0975 / sqrt(n))
29 U_theta <- theta_hat * (1 + z_0975 / sqrt(n))
30
31 # 4) Intervalo para g(theta)
32 g_L <- g_fun(L_theta)
33 g_U <- g_fun(U_theta)
34
35 # Como g(theta) NÃO é monotona, verificamos se o máximo está no intervalo
36 if (L_theta <= theta_star && theta_star <= U_theta) {
37 g_max <- g_fun(theta_star)
38 g_min <- min(g_L, g_U)
39 } else {
40 g_min <- min(g_L, g_U)
41 g_max <- max(g_L, g_U)
42 }
43
44 # 5) Verifica cobertura
45 covered[m] <- (g_min <= g_true) && (g_true <= g_max)

```

```

46 }
47
48 # Estimativa da cobertura
49 mean(covered)

```

Listing 2: Simulação de Monte Carlo para testar o intervalo de confiança aproximado obtido para o item (b),  $g(\theta) = P_\theta(0.1 < Z < 1)$ .

```

1  set.seed(123)
2
3  # Parametros do experimento
4  theta0 <- 2.703      # "verdadeiro" valor de theta
5  n      <- 6          # tamanho da amostra
6  M      <- 1000       # numero de simulacoes
7  z_0975 <- 1.96       # quantil 0.975 da N(0,1)
8
9  # valor verdadeiro de g(theta) = Var(Z) = 1/theta^2
10 g_true <- 1 / theta0^2
11
12 # vetor logico: se o IC cobriu (TRUE) ou nao (FALSE)
13 covered <- logical(M)
14
15 for (m in 1:M) {
16   # 1) gera uma amostra Z_1,...,Z_n ~ Exp(theta0)
17   z <- rexp(n, rate = theta0)
18
19   # 2) EMV de theta: 1 / media
20   theta_hat <- 1 / mean(z)
21
22   # 3) IC aproximado para theta (mesma formula usada na letra (a))
23   L_theta <- theta_hat * (1 - z_0975 / sqrt(n))
24   U_theta <- theta_hat * (1 + z_0975 / sqrt(n))
25
26   # 4) transforma o IC de theta para g(theta)=1/theta^2
27   L_g <- 1 / (U_theta^2) # menor valor de Var(Z)
28   U_g <- 1 / (L_theta^2) # maior valor de Var(Z)
29
30   # 5) verifica se o verdadeiro g_true esta no intervalo
31   covered[m] <- (L_g <= g_true) && (g_true <= U_g)
32 }
33
34 # Estimativa da cobertura do IC aproximado
35 mean(covered)

```

Listing 3: Simulação de Monte Carlo para testar o intervalo de confiança aproximado obtido para o item (c),  $g(\theta) = \text{Var}_\theta(Z)$

**Questão 03:** Seja  $(Z_1, \dots, Z_n)$  uma amostra aleatória de  $Z \sim N(\mu, \sigma^2)$ ,  $\theta = (\mu, \sigma^2) \in (-\infty, \infty) \times (0, \infty)$ .

- (a) Encontre o EMV para  $g(\theta) = E_\theta(Z)$ .
- (b) Encontre o EMV para  $g(\theta) = P_\theta(Z < 2)$ .
- (c) Encontre o EMV para  $g(\theta) = P_\theta(2.6 < Z < 4)$ .
- (d) Encontre o EMV para  $g(\theta) = \text{Var}_\theta(Z)$ .
- (e) Considere que os dados foram observados (2.4, 2.7, 2.3, 2, 2.5, 2.6). Encontre as estimativas de MV nos itens acima.

Podemos calcular os itens (a), (b), (c) e (d) utilizando a propriedade de invariância do estimador de máxima verossimilhança (EMV), calculando uma única vez  $\hat{\theta}_{MV}$ .

### Cálculo do EMV de $\theta$ ( $\hat{\theta}_{MV}$ ) para a distribuição Normal

#### Função de verossimilhança

Seja  $(z_1, \dots, z_n)$  uma amostra aleatória de  $Z \sim N(\mu, \sigma^2)$ , com parâmetro  $\theta = (\mu, \sigma^2) \in (-\infty, \infty) \times (0, \infty)$ .

A função densidade de probabilidade para  $z_i$  é

$$f(z_i; \mu, \sigma^2) = \frac{1}{\sqrt{2\pi} \sigma} \exp\left(-\frac{(z_i - \mu)^2}{2\sigma^2}\right).$$

Como  $z_1, \dots, z_n$  são independentes, a função de verossimilhança é o produto das densidades individuais:

$$\ell(\mu, \sigma^2; z_1, \dots, z_n) = \prod_{i=1}^n f(z_i; \mu, \sigma^2).$$

Aplicando a propriedade distributiva do produto e de produto de exponenciais, obtemos:

$$\ell(\mu, \sigma^2; z_1, \dots, z_n) = \left(\frac{1}{\sqrt{2\pi} \sigma}\right)^n \exp\left(-\frac{1}{2\sigma^2} \sum_{i=1}^n (z_i - \mu)^2\right).$$

Seja

$$S = \sum_{i=1}^n z_i \quad e \quad Q = \sum_{i=1}^n z_i^2, \quad \Longleftrightarrow \quad \sum_{i=1}^n (z_i - \mu)^2 = \sum_{i=1}^n (z_i^2 - 2\mu z_i + \mu^2) = Q - 2\mu S + n\mu^2.$$

Podemos escrever:

$$\ell(\mu, \sigma^2, S, Q) = \left(\frac{1}{\sqrt{2\pi} \sigma}\right)^n \exp(Q - 2\mu S + n\mu^2).$$

Para simplificar os cálculos, trabalhamos com a função logaritmo da verossimilhança  $\mathcal{L}(\mu, \sigma^2, S, Q)$ :

$$\mathcal{L}(\mu, \sigma^2, S, Q) = \log \ell(\mu, \sigma^2, S, Q) = \log \left[ \left(\frac{1}{\sqrt{2\pi} \sigma}\right)^n \right] + \log \left[ \exp\left(-\frac{1}{2\sigma^2}(Q - 2\mu S + n\mu^2)\right) \right],$$

$$\mathcal{L}(\mu, \sigma^2, S, Q) = -\frac{n}{2} \log(2\pi) - \frac{n}{2} \log(\sigma^2) - \frac{1}{2\sigma^2} (Q - 2\mu S + n\mu^2).$$

Calculando a derivada de  $\mathcal{L}(\mu, \sigma^2, S, Q)$  em relação a  $\mu$ :

$$\frac{\partial \mathcal{L}(\mu, \sigma^2, S, Q)}{\partial \mu} = -\frac{1}{2\sigma^2} \frac{\partial}{\partial \mu} [Q - 2\mu S + n\mu^2] = -\frac{1}{2\sigma^2} (-2S + 2n\mu) = \frac{S - n\mu}{\sigma^2}.$$

Seja a condição de máximo onde está definido  $\hat{\mu}_{MV}$ :

$$\frac{\partial \mathcal{L}(\mu, \sigma^2, S, Q)}{\partial \mu} = 0 \Longleftrightarrow \frac{S - n\hat{\mu}_{MV}}{\sigma^2} = 0 \Longleftrightarrow S - n\hat{\mu}_{MV} = 0.$$

Isolando  $\hat{\mu}_{MV}$ :

$$\boxed{\hat{\mu}_{MV} = \frac{S}{n}.$$

Calculando a derivada de  $\mathcal{L}(\mu, \sigma^2, S, Q)$  em relação a  $\sigma^2$ :

$$\frac{\partial \mathcal{L}(\mu, \sigma^2, S, Q)}{\partial \sigma^2} = -\frac{n}{2} \frac{1}{\sigma^2} - \frac{\partial}{\partial \sigma^2} \left[ \frac{1}{2\sigma^2} (Q - 2\mu S + n\mu^2) \right] = -\frac{n}{2\sigma^2} + \frac{1}{2} \frac{Q - 2\mu S + n\mu^2}{(\sigma^2)^2}.$$



$$\frac{\partial \mathcal{L}(\mu, \sigma^2, S, Q)}{\partial \sigma^2} = \frac{-n(\sigma^2) + (Q - 2\mu S + n\mu^2)}{2(\sigma^2)^2}.$$

Seja a condição de máximo onde está definido  $\hat{\sigma}_{MV}^2$ :

$$\frac{\partial \mathcal{L}(\hat{\mu}_{MV}, \sigma^2, S, Q)}{\partial \sigma^2} = 0 \iff -n(\hat{\sigma}_{MV}^2) + (Q - 2\hat{\mu}_{MV}S + n\hat{\mu}_{MV}^2) = 0.$$

Isolando  $\hat{\sigma}_{MV}^2$ :

$$n\hat{\sigma}_{MV}^2 = Q - 2\hat{\mu}_{MV}S + n\hat{\mu}_{MV}^2 = Q - \frac{S^2}{n}, \quad \text{para} \quad \hat{\mu}_{MV} = \frac{S}{n}.$$

Assim,

$$n\hat{\sigma}_{MV}^2 = Q - \frac{S^2}{n}.$$

Finalmente:

$$\boxed{\hat{\sigma}_{MV}^2 = \frac{1}{n} \left( Q - \frac{S^2}{n} \right)}$$

#### Estimador de máxima verossimilhança

Substituindo de volta em função da variável aleatória  $z_i$ , obtemos:

$$\hat{\mu}_{MV} = \frac{S}{n} = \frac{1}{n} \sum_{i=1}^n z_i = \bar{Z}$$

$$\boxed{\hat{\mu}_{MV} = \bar{Z}}$$

$$\hat{\sigma}_{MV}^2 = \frac{1}{n} \left( Q - \frac{S^2}{n} \right) = \frac{1}{n} \left( \sum_{i=1}^n z_i^2 - \frac{(n\bar{Z})^2}{n} \right) = \frac{1}{n} \left( \sum_{i=1}^n z_i^2 - n\bar{Z}^2 \right) = \frac{1}{n} \left( \sum_{i=1}^n z_i^2 - 2\bar{Z}n\bar{Z} + n\bar{Z}^2 \right)$$

$$\hat{\sigma}_{MV}^2 = \frac{1}{n} \left( \sum_{i=1}^n z_i^2 - 2\bar{Z} \sum_{i=1}^n z_i + n\bar{Z}^2 \right) = \frac{1}{n} \sum_{i=1}^n (z_i^2 - 2z_i\bar{Z} + \bar{Z}^2) = \frac{1}{n} \sum_{i=1}^n (z_i - \bar{Z})^2$$

$$\boxed{\hat{\sigma}_{MV}^2 = \frac{1}{n} \sum_{i=1}^n (z_i - \bar{Z})^2}$$

(a)  $g(\theta) = E_{\theta}(Z)$

Pela definição de esperança para variáveis contínuas, sendo  $h(z)$  a função de interesse:

$$E_{\theta}(h(Z)) = \int_{-\infty}^{\infty} h(z) f(z; \theta) dz.$$

Em particular, a esperança do primeiro momento de  $Z$ , é:

$$E_{\theta}(Z) = \int_{-\infty}^{\infty} z \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(z-\mu)^2}{2\sigma^2}} dz = \int_{-\infty}^{\infty} (\mu + y) \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{y^2}{2\sigma^2}} dy = \mu,$$

$$E_{\theta}(Z) = \mu \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{y^2}{2\sigma^2}} dy + \int_{-\infty}^{\infty} y \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{y^2}{2\sigma^2}} dy$$

$$E_{\theta}(Z) = \mu \int_{-\infty}^{\infty} \phi_{\sigma}(y) dy + \int_{-\infty}^{\infty} y \phi_{\sigma}(y) dy = \mu \cdot 1 + 0 = \mu.$$

Então:

$$g(\theta) = g(\mu, \sigma^2) = E_{\theta}(Z) = \mu.$$

Pela **invariância do EMV**, o estimador de máxima verossimilhança de  $g(\theta)$  é:

$$\hat{g}_{MV} = g(\hat{\mu}_{MV}, \hat{\sigma}_{MV}^2) = \hat{\mu}_{MV} = \bar{Z}.$$

$$\boxed{\hat{g}_{MV} = \bar{Z}.}$$

(b)  $g(\theta) = P_{\theta}(Z < 2)$

Seja a função distribuição acumulada da Normal padrão  $N(0, 1)$  definida por:

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-t^2/2} dt.$$

Integrando a função densidade de probabilidade  $f(z_i; \mu, \sigma^2)$  para encontrar  $P_{\theta}(Z < 2)$ , obtemos:

$$P_{\theta}(Z < 2) = \int_{-\infty}^2 \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(z-\mu)^2/(2\sigma^2)} dz = \int_{-\infty}^{\frac{2-\mu}{\sqrt{\sigma^2}}} \frac{1}{\sqrt{2\pi}} e^{-y^2/2} dy = \Phi\left(\frac{2-\mu}{\sqrt{\sigma^2}}\right),$$

Assim, expressamos o resultado em termos de  $\Phi$ :

$$g(\theta) = g(\mu, \sigma^2) = \Phi\left(\frac{2-\mu}{\sqrt{\sigma^2}}\right).$$

Pela **invariância do EMV**, o estimador de máxima verossimilhança de  $g(\theta)$  é:

$$\hat{g}_{MV} = g(\hat{\mu}_{MV}, \hat{\sigma}_{MV}^2) = \Phi\left(\frac{2-\hat{\mu}_{MV}}{\sqrt{\hat{\sigma}_{MV}^2}}\right) = \Phi\left(\frac{2-\bar{Z}}{\sqrt{\hat{\sigma}_{MV}^2}}\right) = \Phi\left(\frac{(2-\bar{Z})\sqrt{n}}{\sqrt{\sum_{i=1}^n (z_i - \bar{Z})^2}}\right)$$

$$\boxed{\hat{g}_{MV} = \Phi\left(\frac{(2-\bar{Z})\sqrt{n}}{\sqrt{\sum_{i=1}^n (z_i - \bar{Z})^2}}\right)}$$

(c)  $g(\theta) = P_\theta(2.6 < Z < 4)$

De forma similar, podemos decompor a probabilidade em termos da função de distribuição acumulada da Normal padrão  $N(0, 1)$  :

$$P_\theta(2.6 < Z < 4) = P_\theta(Z < 4) - P_\theta(Z \leq 2.6) = \Phi\left(\frac{4 - \mu}{\sqrt{\sigma^2}}\right) - \Phi\left(\frac{2.6 - \mu}{\sqrt{\sigma^2}}\right).$$

Logo,

$$g(\theta) = g(\mu, \sigma^2) = \Phi\left(\frac{4 - \mu}{\sqrt{\sigma^2}}\right) - \Phi\left(\frac{2.6 - \mu}{\sqrt{\sigma^2}}\right).$$

Pela **invariância do EMV**, o estimador de máxima verossimilhança de  $g(\theta)$  é

$$\hat{g}_{MV} = g(\hat{\mu}_{MV}, \hat{\sigma}_{MV}^2) = \Phi\left(\frac{4 - \bar{Z}}{\sqrt{\hat{\sigma}_{MV}^2}}\right) - \Phi\left(\frac{2.6 - \bar{Z}}{\sqrt{\hat{\sigma}_{MV}^2}}\right),$$

$$\boxed{\hat{g}_{MV} = \Phi\left(\frac{(4 - \bar{Z})\sqrt{n}}{\sqrt{\sum_{i=1}^n (z_i - \bar{Z})^2}}\right) - \Phi\left(\frac{(2.6 - \bar{Z})\sqrt{n}}{\sqrt{\sum_{i=1}^n (z_i - \bar{Z})^2}}\right)}.$$

(d)  $g(\theta) = \text{Var}_\theta(Z)$

A variância é definida por:

$$\text{Var}(Z) = E_\theta(Z^2) - (E_\theta(Z))^2.$$

Pela definição de esperança para variáveis contínuas, sendo  $h(z)$  a função de interesse:

$$E_\theta(h(Z)) = \int_{-\infty}^{\infty} h(z) f(z; \theta) dz.$$

Em particular, a esperança do primeiro e segundo momentos de  $Z$ , para a distribuição Normal, são:

$$E_\theta(Z) = \mu.$$

e

$$\begin{aligned} E_\theta(Z^2) &= \int_{-\infty}^{\infty} z^2 \frac{1}{\sqrt{2\pi}\sigma^2} \exp\left(-\frac{(z - \mu)^2}{2\sigma^2}\right) dz = \int_{-\infty}^{\infty} (\mu + \sigma y)^2 \frac{1}{\sqrt{2\pi}} e^{-y^2/2} dy. \\ E_\theta(Z^2) &= \mu^2 \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-y^2/2} dy + 2\mu\sigma \int_{-\infty}^{\infty} y \frac{1}{\sqrt{2\pi}} e^{-y^2/2} dy + \sigma^2 \int_{-\infty}^{\infty} y^2 \frac{1}{\sqrt{2\pi}} e^{-y^2/2} dy. \\ E_\theta(Z^2) &= \mu^2 \cdot 1 + 2\mu\sigma \cdot 0 + \sigma^2 \cdot 1 = \mu^2 + \sigma^2. \end{aligned}$$

Então:

$$\text{Var}_\theta(Z) = E(Z^2) - [E(Z)]^2 = \mu^2 + \sigma^2 - \mu^2 = \sigma^2.$$

Portanto,

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

Pela **invariância do EMV**, o estimador de máxima verossimilhança de  $g(\theta)$  é:

$$\hat{g}_{MV} = g(\hat{\mu}_{MV}, \hat{\sigma}_{MV}^2) = \hat{\sigma}_{MV}^2 = \frac{1}{n} \sum_{i=1}^n (z_i - \bar{Z})^2.$$

$$\hat{g}_{MV} = \frac{1}{n} \sum_{i=1}^n (z_i - \bar{Z})^2.$$

(e) **Sejam os dados observados:** (0,12, 0,50, 0,20, 0,23, 0,30, 0,11)

Para o cálculo da média:  $n = 6$  e  $\sum_{i=1}^n z_i = 0,12 + 0,50 + 0,20 + 0,23 + 0,30 + 0,11 = 1,46$ . Portanto:

$$\hat{\mu}_{MV} = \bar{Z} = \frac{1,46}{6} \approx 0,2433.$$

$$\hat{\mu}_{MV} \approx 0,2433$$

Para a variância de máxima verossimilhança:

$$\hat{\sigma}_{MV}^2 = \frac{1}{n} \sum_{i=1}^n (z_i - \bar{Z})^2$$

$$\hat{\sigma}_{MV}^2 = (0,12 - 0,2433)^2 + (0,50 - 0,2433)^2 + (0,20 - 0,2433)^2 + (0,23 - 0,2433)^2 + (0,30 - 0,2433)^2 + (0,11 - 0,2433)^2$$

$$\hat{\sigma}_{MV}^2 \approx 0,01736$$

Por fim, o desvio-padrão de máxima verossimilhança é

$$\hat{\sigma}_{MV} = \sqrt{\hat{\sigma}_{MV}^2} \approx \sqrt{0,01736} \approx 0,1317.$$

Assim:

$$(a) \hat{g}_{MV} = \widehat{E_{\theta}(Z)} = \hat{\mu}_{MV} = \bar{Z} \approx 0,2433.$$

$$(b) \hat{g}_{MV} = P_{\theta}(\widehat{Z} < 2) = \Phi\left(\frac{2 - \hat{\mu}_{MV}}{\hat{\sigma}_{MV}}\right) = \Phi\left(\frac{2 - \bar{Z}}{\hat{\sigma}_{MV}}\right) \approx \Phi(13,33) \approx 1.$$

$$(c) \hat{g}_{MV} = P_{\theta}(2,6 < \widehat{Z} < 4) = \Phi\left(\frac{4 - \bar{Z}}{\hat{\sigma}_{MV}}\right) - \Phi\left(\frac{2,6 - \bar{Z}}{\hat{\sigma}_{MV}}\right) \approx \Phi(28,52) - \Phi(17,89) \approx 0.$$

$$(d) \hat{g}_{MV} = \widehat{\text{Var}_{\theta}(Z)} = \hat{\sigma}_{MV}^2 \approx 0,01736.$$

**Questão 04:** Seja  $(Z_1, \dots, Z_n)$  uma amostra aleatória de  $Z \sim f_{\theta}$ ,  $\theta \in (0, \infty)$ , tal que a função densidade de probabilidade é dada por

$$f_{\theta}(x) = \theta x^{\theta-1}, \quad x \in (0, 1),$$

e  $f_{\theta}(x) = 0$ , caso contrário.

- (a) Encontre o EMV para  $g(\theta) = E_{\theta}(Z)$ .
- (b) Encontre o EMV para  $g(\theta) = P_{\theta}(Z > 0.3)$ .
- (c) Encontre o EMV para  $g(\theta) = P_{\theta}(0 < Z < 0.1)$ .
- (d) Encontre o EMV para  $g(\theta) = \text{Var}_{\theta}(Z)$ .

- (e) Considere que os dados foram observados (0.12, 0.50, 0.20, 0.23, 0.30, 0.11). Encontre as estimativas de MV para os itens acima.

Podemos calcular os itens (a), (b), (c) e (d) utilizando a propriedade de invariância do estimador de máxima verossimilhança (EMV), calculando uma única vez  $\hat{\theta}_{MV}$ .

## Cálculo do EMV de $\theta$ ( $\hat{\theta}_{MV}$ ) para a distribuição Beta

### Função de verossimilhança

Seja  $(z_1, \dots, z_n)$  uma amostra aleatória de  $Z \sim f_\theta$ ,  $\theta \in (0, \infty)$ , com função densidade de probabilidade:

$$f(z_i, \theta) = \theta z_i^{\theta-1}, \quad 0 < z_i < 1,$$

e  $f(z_i, \theta) = 0$  caso contrário.

Como  $z_1, \dots, z_n$  são independentes, a função de verossimilhança é o produto das densidades individuais:

$$\ell(\theta; z_1, \dots, z_n) = \prod_{i=1}^n f_\theta(z_i) = \prod_{i=1}^n \theta z_i^{\theta-1}.$$

Aplicando a propriedade distributiva do produto, obtemos:

$$\ell(\theta; z_1, \dots, z_n) = \theta^n \prod_{i=1}^n z_i^{\theta-1}.$$

Seja, e utilizando a identidade  $a^b = e^{b \log a}$ :

$$T = \sum_{i=1}^n \log z_i \quad e \quad \prod_{i=1}^n z_i^{\theta-1} = \exp\left((\theta-1) \sum_{i=1}^n \log z_i\right) = \exp((\theta-1)T),$$

Podemos escrever:

$$\ell(\theta, T) = \theta^n \exp((\theta-1)T).$$

Para simplificar os cálculos, trabalhamos com a função logaritmo da verossimilhança  $\mathcal{L}(\theta; T)$ :

$$\mathcal{L}(\theta; T) = \log \ell(\theta; T) = \log(\theta^n) + \log[\exp((\theta-1)T)] = n \log \theta + (\theta-1)T.$$

Calculando a derivada de  $\mathcal{L}(\theta, T)$  em relação a  $\theta$ :

$$\frac{\partial \mathcal{L}(\theta, T)}{\partial \theta} = \frac{\partial}{\partial \theta} [n \log \theta + (\theta-1)T] = \frac{n}{\theta} + T.$$

Seja a condição de máximo onde está definido  $\hat{\theta}_{MV}$ :

$$\frac{\partial \mathcal{L}(\theta, T)}{\partial \theta} = 0 \iff \frac{n}{\hat{\theta}_{MV}} + T = 0 \iff \frac{n}{\hat{\theta}_{MV}} = -T$$

Isolando  $\hat{\theta}_{MV}$ :

$$\boxed{\hat{\theta}_{MV} = -\frac{n}{T}}.$$

## Estimador de máxima verossimilhança

Substituindo de volta em função da variável aleatória  $z_i$ , obtemos:

$$\hat{\theta}_{MV} = -\frac{n}{\sum_{i=1}^n \log z_i}.$$

(a)  $g(\theta) = \mathbb{E}_\theta(Z)$ .

Pela definição de esperança para variáveis contínuas, sendo  $h(z)$  a função de interesse:

$$E_\theta(h(Z)) = \int_{-\infty}^{\infty} h(z) f(z; \theta) dz.$$

Portanto, a esperança do primeiro momento de  $Z$ , para a distribuição  $\text{Beta}(\theta, 1)$ , é:

$$E_\theta(Z) = \int_0^1 z \theta z^{\theta-1} dz = \theta \int_0^1 z^\theta dz = \theta \left[ \frac{z^{\theta+1}}{\theta+1} \right]_0^1 = \frac{\theta}{\theta+1},$$

Então:

$$g(\theta) = \mathbb{E}_\theta(Z) = \frac{\theta}{\theta+1}.$$

Pela **invariância do EMV**, o estimador de máxima verossimilhança de  $g(\theta)$  é:

$$\hat{g}_{MV} = g(\hat{\theta}_{MV}) = \frac{\hat{\theta}_{MV}}{\hat{\theta}_{MV} + 1} = \frac{-\sum_{i=1}^n \log z_i}{1 - \sum_{i=1}^n \log z_i} = \frac{n}{n - \sum_{i=1}^n \log z_i}.$$

$$\hat{g}_{MV} = \frac{n}{n - \sum_{i=1}^n \log z_i}.$$

(b)  $g(\theta) = P_\theta(Z > 0,3)$

Integrando a função densidade de probabilidade  $f(z, \theta)$  para encontrar  $P_\theta(Z > 0,3)$ , temos:

$$g(\theta) = P_\theta(Z > 0,3) = \int_{0,3}^1 f(z, \theta) dz = \int_{0,3}^1 \theta z^{\theta-1} dz.$$

Calculando a integral:

$$\int_{0,3}^1 \theta z^{\theta-1} dz = \theta \left[ \frac{z^\theta}{\theta} \right]_{0,3}^1 = [z^\theta]_{0,3}^1 = 1 - (0,3)^\theta.$$

Logo,

$$g(\theta) = 1 - (0,3)^\theta.$$

Pela **invariância do EMV**, o estimador de máxima verossimilhança de  $g(\theta)$  é:

$$\hat{g}_{MV} = g(\hat{\theta}_{MV}) = 1 - (0,3)^{\hat{\theta}_{MV}} = 1 - (0,3)^{-\sum_{i=1}^n \log z_i}$$

$$\hat{g}_{MV} = 1 - (0,3)^{-\sum_{i=1}^n \frac{1}{\log z_i}}.$$

(c)  $g(\theta) = P_\theta(0 < Z < 0,1)$

Integrando a função densidade de probabilidade  $f(z, \theta)$  para encontrar  $P_\theta(0 < Z < 0,1)$ , temos:

$$g(\theta) = P_\theta(0 < Z < 0,1) = \int_0^{0,1} f(z, \theta) dz = \int_0^{0,1} \theta z^{\theta-1} dz.$$

Calculando a integral:

$$\int_0^{0,1} \theta z^{\theta-1} dz = \theta \left[ \frac{z^\theta}{\theta} \right]_0^{0,1} = [z^\theta]_0^{0,1} = (0,1)^\theta - 0 = (0,1)^\theta.$$

Logo,

$$g(\theta) = (0,1)^\theta.$$

Pela **invariância do EMV**, o estimador de máxima verossimilhança de  $g(\theta)$  é:

$$\hat{g}_{MV} = g(\hat{\theta}_{MV}) = (0,1)^{\hat{\theta}_{MV}} = (0,1)^{-\sum_{i=1}^n \frac{1}{\log z_i}}.$$

$$\hat{g}_{MV} = (0,1)^{-\sum_{i=1}^n \frac{1}{\log z_i}}.$$

(d)  $g(\theta) = \text{Var}_\theta(Z)$ .

A variância é definida por:

$$\text{Var}(Z) = E_\theta(Z^2) - (E_\theta(Z))^2.$$

Pela definição de esperança para variáveis contínuas, sendo  $h(z)$  a função de interesse:

$$E_\theta(h(Z)) = \int_{-\infty}^{\infty} h(z) f(z; \theta) dz.$$

Em particular, a esperança do primeiro e segundo momentos de  $Z$ , para a  $\text{Beta}(\theta, 1)$ , são:

$$E_\theta(Z) = \frac{\theta}{\theta + 1},$$

e

$$E_\theta(Z^2) = \int_0^1 z^2 \theta z^{\theta-1} dz = \theta \int_0^1 z^{\theta+1} dz = \theta \left[ \frac{z^{\theta+2}}{\theta+2} \right]_0^1 = \frac{\theta}{\theta+2}.$$

Então:

$$\text{Var}_\theta(Z) = E_\theta(Z^2) - (E_\theta(Z))^2 = \frac{\theta}{\theta+2} - \left( \frac{\theta}{\theta+1} \right)^2 = \frac{\theta}{(\theta+1)^2(\theta+2)}.$$

Portanto,

$$g(\theta) = \frac{\theta}{(\theta+1)^2(\theta+2)}.$$

Pela **invariância do EMV**, o estimador de máxima verossimilhança de  $g(\theta)$  é:

$$\hat{g}_{MV}(\theta) = g(\hat{\theta}_{MV}) = \frac{\hat{\theta}_{MV}}{(\hat{\theta}_{MV} + 1)^2(\hat{\theta}_{MV} + 2)}.$$

$$\hat{g}_{MV} = \frac{\hat{\theta}_{MV}}{(\hat{\theta}_{MV} + 1)^2(\hat{\theta}_{MV} + 2)}.$$

(e) **Sejam os dados observados:** (0.12, 0.50, 0.20, 0.23, 0.30, 0.11)

Com  $n = 6$ ,

$$\sum_{i=1}^n \log z_i = \log(0.12) + \log(0.50) + \log(0.20) + \log(0.23) + \log(0.30) + \log(0.11) \approx -9.3038.$$

Logo,

$$\hat{\theta}_{MV} = -\frac{6}{-9.3038} \approx 0.6449.$$

Assim:

$$(a) \hat{g}_{MV} = \frac{\hat{\theta}_{MV}}{1 + \hat{\theta}_{MV}} \approx \frac{0.6449}{1.6449} \approx 0.3921$$

$$(b) \hat{g}_{MV} = 1 - (0.3)^{\hat{\theta}_{MV}} \approx 1 - (0.3)^{0.6449} \approx 0.5400$$

$$(c) \hat{g}_{MV} = (0.1)^{\hat{\theta}_{MV}} \approx (0.1)^{0.6449} \approx 0.2265$$

$$(d) \hat{g}_{MV} = \frac{\hat{\theta}_{MV}}{(\hat{\theta}_{MV} + 1)^2(\hat{\theta}_{MV} + 2)} \approx \frac{0.6449}{(1.6449)^2(2.6449)} \approx 0.0901$$

## Referências

Jairo Simon Fonseca. *Curso de Estatística*. Atlas, São Paulo, 6 edition, 2008.

Alexandre Patriota. Notas de aula de fundamentos de estatística e machine learning. Notas de aula da disciplina MAE5911, 2025.