Universidade de São Paulo
Instituto de Matemática e Estatística
Aluna: Nara Avila Moraes
Nro. USP: 5716734

2o. semestre, 2025
Exercício 02 - 15/10

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

Questão 01: Seja $(Z_1, \ldots, Z_n)$ 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.

O estimador via função de máxima verossimilhança foi proposto por Fisher, que demonstrou a superioridade do deste método referente a outros métodos estimadores, como o método de momentos. Este estimador é utilizado sobretudo quando a informação da distribuição de probabilidade do modelo é conhecida. Ela tem propriedades como eficiência assintótica, consistência, distribuição normal e invariancia [Notas de aula pg. 60]. O estimador de máxima verossimilhança representa uma estimativa para o valor do parâmetro θ que torna os dados observados o mais plausível possível.

Devido a propriedade de invariancia do EMV, basta encontrarmos EMV para a letra (a) e depois calcular a letra (b) utilizando este estimador como parâmetro.

Seja $(Z_1, \ldots, Z_n)$ a amostra aleatória de

$$Z_i \sim \text{Bernoulli}(\theta), \quad \theta \in (0, 1).$$

A função de verossimilhança, que é a função de densidade de probabilidade conjunta do modelo, é dada por

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

Para encontrar o estimador de máxima verossimilhança, precisamos maximizar $L(\theta)$ em relação a θ. Mas como esta função se trata de um produtório, é mais fácil maximizar a função log-verossimilhança, que por ser monotonicamente crescente, preserva o valor de theta no ponto máximo $L(\theta)$.

Tomando o logaritmo natural em ambos os lados:

$$\ell(\theta) = \log L(\theta; z_1, \ldots, z_n) = \log\left(\prod_{i=1}^{n} \theta^{z_i}(1-\theta)^{1-z_i}\right).$$

Sabendo que $\log(ab) = \log a + \log b$ e que $\log(\prod_i a_i) = \sum_i \log a_i$, temos:

$$\ell(\theta) = \sum_{i=1}^{n} \log\big(\theta^{z_i}(1-\theta)^{1-z_i}\big).$$

Aplicando novamente a propriedade do logaritmo de um produto:

$$\ell(\theta) = \sum_{i=1}^{n} \big[\log(\theta^{z_i}) + \log((1-\theta)^{1-z_i})\big].$$

Sabendo que $\log(a^b) = b \log a$, obtemos:

$$\ell(\theta) = \sum_{i=1}^{n} \big[z_i \log\theta + (1-z_i)\log(1-\theta)\big].$$

Por fim, separando as somas:

$$\ell(\theta) = \left(\sum_{i=1}^{n} z_i\right)\log\theta + \left(\sum_{i=1}^{n}(1-z_i)\right)\log(1-\theta).$$

A log-verossimilhança correspondente é

$$\boxed{\ell(\theta) = \sum_{i=1}^{n} \big[z_i \log\theta + (1-z_i)\log(1-\theta)\big].}$$

Para derivar em relação a θ, sabemos que $\dfrac{d}{d\theta}\log\theta = \dfrac{1}{\theta}$ e $\dfrac{d}{d\theta}\log(1-\theta) = -\dfrac{1}{1-\theta}$, obtemos, termo a termo:

$$\frac{d}{d\theta}\Big[z_i\log\theta\Big] = z_i\,\frac{1}{\theta} \qquad \text{e} \qquad \frac{d}{d\theta}\Big[(1-z_i)\log(1-\theta)\Big] = (1-z_i)\Big(-\frac{1}{1-\theta}\Big) = -\frac{1-z_i}{1-\theta}.$$

Logo,

$$\ell'(\theta) = \sum_{i=1}^{n}\left(\frac{z_i}{\theta} - \frac{1-z_i}{1-\theta}\right) = \frac{\sum_{i=1}^{n} z_i}{\theta} - \frac{\sum_{i=1}^{n}(1-z_i)}{1-\theta} = \frac{\sum_i z_i}{\theta} - \frac{n-\sum_i z_i}{1-\theta}.$$

Igualando a zero para encontrar o **ponto crítico.**

$$\ell'(\theta) = 0 \iff \frac{\sum_i z_i}{\theta} = \frac{n-\sum_i z_i}{1-\theta} \iff (1-\theta)\sum_i z_i = \theta\,(n - \sum_i z_i).$$

Expandindo e rearranjando:

$$\sum_i z_i - \theta\sum_i z_i = \theta n - \theta\sum_i z_i \iff \sum_i z_i = \theta n \iff \widehat{\theta}_{MV} = \frac{1}{n}\sum_{i=1}^{n} z_i = \overline{Z}.$$

Obtivemos o estimador de máxima verossimilhança para a probabilidade de sucesso θ:

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

O teste de hipótese, mede o quão improvável seria observar $\hat{\theta}_{MV}$ se a hipótese nula fosse verdadeira. O valor-p é esta medida de improbabilidade.

Em condições de regularidade, o valor-p segue uma distribuição uniforme em $(0,1)$. Isso significa que, se o seu valor é muito pequeno, por exemplo $p \leq 0{,}01$, então o resultado observado é um evento raro sob a hipótese nula H_0, ou, alternativamente, que a hipótese nula está incorreta. Nessa situação, existem evidências para rejeitar H_0.

Resposta:

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

1. Seja $Z \sim \text{Bernoulli}(\theta)$, temos:

$$P_\theta(Z=z) = \theta^z(1-\theta)^{1-z}, \quad z \in \{0,1\}.$$

2. Substituindo $z=0$:

$$P_\theta(Z=0) = \theta^0(1-\theta)^{1-0} = 1-\theta.$$

3. Portanto,

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

4. Pela **invariância do EMV**, temos:

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

5. E como $\widehat{\theta}_{MV} = \overline{Z}$, obtemos:

$$\boxed{\widehat{g}_{MV} = 1 - \overline{Z}.}$$

—

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

1. Para uma Bernoulli, a variância é dada por:

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

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

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

3. Portanto,

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

4. Pela **invariância do EMV**:

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

5. Como $\widehat{\theta}_{MV} = \overline{Z}$, segue:

$$\boxed{\widehat{g}_{MV} = \overline{Z}(1 - \overline{Z}).}$$

—

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

Temos $n = 6$ e $\sum z_i = 2$, portanto

$$\hat{\theta}_{MV} = \bar{Z} = \frac{2}{6} = \frac{1}{3} \approx 0.3333.$$

Assim:

$$P(\widehat{Z = 0}) = 1 - \hat{\theta}_{MV} = \frac{2}{3} \approx 0.6667,$$

$$\widehat{\mathrm{Var}}(Z) = \hat{\theta}_{MV}(1 - \hat{\theta}_{MV}) = \frac{1}{3} \cdot \frac{2}{3} = \frac{2}{9} \approx 0.2222.$$

—

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

Queremos verificar se temos evidência a partir dos dados observados para rejeitar a hipótese nula de que a probabilidade de sucesso é $\theta = 0.1$.

Sabemos que, sob H_0, a estatística

$$X = \sum_{i=1}^{n} Z_i$$

segue uma distribuição Binomial(n, θ).

Com $n = 6$ e $x_{\text{obs}} = 2$ sucessos observados, temos:

$$X \sim \mathrm{Binomial}(6, 0.1).$$

—

O valor-p é a probabilidade de observar um valor da estatística $T(Z_n)$ *tão ou mais extremo quanto o observado*, assumindo que H_0 é verdadeira:

$$\boxed{\text{valor-}p(H_0, \mathbf{z}_n) = \sup_{\theta\in\Theta_0} P_\theta^{(n)}\big(T_{H_0}(\mathbf{Z}_n) \geq T_{H_0}(\mathbf{z}_n)\big).}$$

No contexto binomial, a estatística de teste é o número de sucessos X. Vamos avaliar a probabilidade de eventos com parâmetro de sucesso a partir da amostra observada até eventos mais extremos uma vez que a amostra observada já está acima do esperado, em relação à hipotese nula. Portanto vamos avaliar:

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

—

Calculando o valor-p pelo seu valor complementar:

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

Logo, a probabilidade testada é o complemento da soma dessas probabilidades:

$$valor - p = 1 - [P(X = 0) + P(X = 1)] = 1 - (0.531441 + 0.354294) = 0.114265.$$

$$\boxed{\text{valor-p} = 0.114265}$$

—

Este resultado **não é improvável** sob H_0 (<0.05). Portanto, não há evidência para rejeitar H_0, isto é, não evidências de que $\theta = 0.1$ seja incompatível com os dados.

$$\boxed{\text{Conclusão: Não há evidências suficientes para rejeitar } H_0 : \theta = 0.1.}$$

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.

Seja $Z_1, \dots, Z_n \overset{\text{iid}}{\sim} \text{Exp}(\theta)$, com $\theta > 0$. A função densidade de probabilidade é

$$f(z \mid \theta) = \theta\, e^{-\theta z}\, \mathbf{1}_{\{z\geq 0\}},$$

onde $\mathbf{1}_{\{z\geq 0\}}$ é a função indicadora (vale 1 se $z \geq 0$ e 0 caso contrário).

Como as observações são independentes e identicamente distribuídas, a **função de verossimilhança**, que é a função de densidade de probabilidade conjunta, é o produto das densidades individuais:

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

Aplicando a propriedade distributiva do produto, separamos os fatores:

$$L(\theta; z) = \Big(\prod_{i=1}^{n} \mathbf{1}_{\{z_i \geq 0\}}\Big)\Big(\prod_{i=1}^{n} \theta e^{-\theta z_i}\Big).$$

Como o termo indicador $\prod \mathbf{1}_{\{z_i \geq 0\}}$ não depende de θ, ele não influencia a maximização e pode ser ignorado. Assim:

$$L(\theta; z) = \prod_{i=1}^{n} (\theta e^{-\theta z_i}).$$

Usando a propriedade $\prod(ab) = \prod a \cdot \prod b$, obtemos:

$$L(\theta; z) = \Big(\prod_{i=1}^{n} \theta\Big)\Big(\prod_{i=1}^{n} e^{-\theta z_i}\Big).$$

Aplicando: 1. $\prod_{i=1}^{n} \theta = \theta^n$ (produto de n fatores iguais), 2. $\prod e^{a_i} = e^{\sum a_i}$ (exponencial do somatório), segue que a função de verossimilhança é da forma:

$$\boxed{L(\theta; z) = \theta^n \exp\left(-\theta \sum_{i=1}^{n} z_i\right).}$$

—

Para encontrar a função **Log-verossimilhança**, aplicamos log em ambos os lados, e usamos as propriedades do logaritmo:

1. $\log(ab) = \log a + \log b$ 2. $\log(a^b) = b \log a$ 3. $\log(e^x) = x$ 4. $\log(\prod a_i) = \sum \log a_i$

$$\ell(\theta) = \log L(\theta; z) = \log(\theta^n) + \log\left(\exp\left[-\theta \sum_{i=1}^{n} z_i\right]\right).$$

Aplicando $\log(a^b) = b \log a$ no primeiro termo e $\log(e^x) = x$ no segundo termo, obtemos:

$$\ell(\theta) = n \log \theta - \theta \sum_{i=1}^{n} z_i.$$

Para maximizar, derivamos em relação a θ, aplicando as regras de derivação:

$$\frac{d}{d\theta} \log \theta = \frac{1}{\theta}, \qquad \frac{d}{d\theta}(a\theta) = a.$$

$$\ell'(\theta) = \frac{d}{d\theta}[n \log \theta - \theta \sum_{i=1}^{n} z_i] = n\frac{1}{\theta} - \sum_{i=1}^{n} z_i = \frac{n}{\theta} - \sum_{i=1}^{n} z_i.$$

Logo:

$$\boxed{\ell'(\theta) = \frac{n}{\theta} - \sum_{i=1}^{n} z_i.}$$

Igualamos a zero para encontrar o **ponto crítico** que maximiza a função.

$$\ell'(\theta) = 0 \iff \frac{n}{\theta} - \sum_{i=1}^{n} z_i = 0 \iff \frac{n}{\theta} = \sum_{i=1}^{n} z_i \iff \widehat{\theta}_{MV} = \frac{n}{\sum_{i=1}^{n} z_i} = \frac{1}{\overline{Z}}.$$

—

Para assegurar que se trata de um ponto de máximo, derivamos novamente:

$$\ell''(\theta) = \frac{d}{d\theta}\left(\frac{n}{\theta} - \sum_{i=1}^{n} z_i\right) = -\frac{n}{\theta^2}.$$

Como $\ell''(\theta) < 0$ para $\theta > 0$, o ponto crítico corresponde a um máximo.

—

O **estimador de máxima verossimilhança** (EMV) é, portanto:

$$\boxed{\widehat{\theta}_{MV} = \frac{n}{\sum_{i=1}^{n} z_i} = \frac{1}{\overline{Z}}},$$

É válido a propriedade da **Invariância do EMV**: para $g(\theta)$, o EMV é $\widehat{g} = g(\widehat{\theta}_{MV})$.

Composição da distribuição Exponencial contínua acumulada. Separando a distribuição exponencial em função de distribuição acumulada até a observação de interesse e o restante da distribuição acumulada denominada a cauda da distribuição, temos:

Expressão para densidade total:

$$f(z \mid \theta) = \theta e^{-\theta z}\mathbf{1}_{\{z \geq 0\}}.$$

Função de distribuição acumulada:

$$F(z) = P_\theta(Z \leq z) = \int_0^z \theta e^{-\theta t}\,dt = \big[-e^{-\theta t}\big]_0^z = 1 - e^{-\theta z}.$$

Cauda:

$$P_\theta(Z > z) = 1 - F(z) = e^{-\theta z}.$$

Resposta:

(a) $g(\theta) = P_\theta(Z > 1)$**.**

1. Pela fórmula da cauda,
$$g(\theta) = P_\theta(Z > 1) = e^{-\theta\cdot 1} = e^{-\theta}.$$
2. Pela **invariância do EMV**,
$$\widehat{g}_{MV} = g(\widehat{\theta}_{MV}) = e^{-\widehat{\theta}_{MV}}.$$
3. Usando $\widehat{\theta}_{MV} = \dfrac{1}{\overline{Z}}$,
$$\boxed{\widehat{g}_{(a)} = e^{-\widehat{\theta}_{MV}} = \exp\Big(-\frac{1}{\overline{Z}}\Big).}$$

(b) $g(\theta) = P_\theta(0.1 < Z < 1)$**.**

1. Para $0 < a < b$, utilizando a expressão para função de distribuição acumulada já calculada,
$$P_\theta(a < Z < b) = F(b) - F(a) = (1 - e^{-\theta b}) - (1 - e^{-\theta a}) = e^{-\theta a} - e^{-\theta b}.$$

2. Com $a = 0.1$ e $b = 1$,

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

3. Pela invariância,

$$\widehat{g}_{(b)} = e^{-0.1\,\widehat{\theta}_{MV}} - e^{-\widehat{\theta}_{MV}} = \exp\Big(-\frac{0.1}{\overline{Z}}\Big) - \exp\Big(-\frac{1}{\overline{Z}}\Big).$$

$$\boxed{\widehat{g}_{(b)} = \exp\Big(-\frac{0.1}{\overline{Z}}\Big) - \exp\Big(-\frac{1}{\overline{Z}}\Big).}$$

(c) $g(\theta) = \mathrm{Var}_\theta(Z)$**.**

1. Para $Z \sim \mathrm{Exp}(\theta)$, $E(Z) = \dfrac{1}{\theta}$ e $E(Z^2) = \displaystyle\int_0^\infty z^2\theta e^{-\theta z}dz = \dfrac{2}{\theta^2}$. Assim,

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

2. Logo $g(\theta) = \theta^{-2}$ e, por invariância,

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

3. Como $\widehat{\theta}_{MV} = 1/\overline{Z}$,

$$\boxed{\widehat{g}_{(c)} = \frac{1}{(1/\overline{Z})^2} = \overline{Z}^2.}$$

(d) ICs aproximados de 95% para $g(\theta)$ (usando o método delta). Para a Exponencial, a informação de Fisher é $I(\theta) = \frac{n}{\theta^2}$, de modo que

$$\widehat{\theta}_{MV} \,\dot{\sim}\, N\Big(\theta,\ \frac{\theta^2}{n}\Big).$$

Pelo método delta, para g diferenciável:

$$\widehat{g} \,\dot{\sim}\, N\Big(g(\theta),\ \frac{\theta^2}{n}\,[g'(\theta)]^2\Big), \quad \text{e usamos } \theta \leftarrow \widehat{\theta}_{MV}.$$

Assim:

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

$$\widehat{\mathrm{Var}}(\widehat{g}_{(a)}) = \frac{\widehat{\theta}^2_{MV}}{n}\,e^{-2\widehat{\theta}_{MV}}, \qquad \mathrm{IC}_{95\%}:\ \widehat{g}_{(a)}\ \pm\ 1.96\sqrt{\frac{\widehat{\theta}^2_{MV}}{n}\,e^{-2\widehat{\theta}_{MV}}}.$$

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

$$\widehat{\mathrm{Var}}(\widehat{g}_{(b)}) = \frac{\widehat{\theta}^2_{MV}}{n}\left[-0.1\,e^{-0.1\widehat{\theta}_{MV}} + e^{-\widehat{\theta}_{MV}}\right]^2,$$

$$\mathrm{IC}_{95\%}:\ \widehat{g}_{(b)}\ \pm\ 1.96\sqrt{\frac{\widehat{\theta}^2_{MV}}{n}\left[-0.1\,e^{-0.1\widehat{\theta}_{MV}} + e^{-\widehat{\theta}_{MV}}\right]^2}.$$

(c) $g(\theta) = 1/\theta^2$, $g'(\theta) = -2/\theta^3$. Então

$$\widehat{\mathrm{Var}}(\widehat{g}_{(c)}) = \frac{4}{n\,\widehat{\theta}^4_{MV}}, \qquad \mathrm{IC}_{95\%}:\ \widehat{g}_{(c)}\ \pm\ 1.96\sqrt{\frac{4}{n\,\widehat{\theta}^4_{MV}}}.$$

Aplicando aos dados $(0.2, 0.6, 0.3, 0.2, 0.8, 0.12)$:

$$n = 6, \quad \sum z_i = 2.22, \quad \overline{Z} = 0.37, \quad \widehat{\theta}_{MV} = \frac{n}{\sum z_i} = \frac{6}{2.22} \approx 2.7027.$$

Estimativas plug-in:

$$\widehat{g}_{(a)} = e^{-\widehat{\theta}_{MV}} \approx 0.0668, \qquad \widehat{g}_{(b)} = e^{-0.1\widehat{\theta}_{MV}} - e^{-\widehat{\theta}_{MV}} \approx 0.6963, \qquad \widehat{g}_{(c)} = \overline{Z}^2 = 0.1369.$$

ICs (delta, 95%):

$$\text{(a)} \;\; [0,\ 0.211] \;\; (\text{truncado a } [0,1]), \qquad \text{(b)} \;\; [0.676,\ 0.717], \qquad \text{(c)} \;\; [0,\ 0.356] \;\; (\text{truncado a } [0,\infty)).$$

Obs.: Com $n = 6$ o delta pode produzir limites fora do espaço paramétrico; é padrão truncar aos limites naturais.

(e) Cobertura por Monte Carlo (plano de simulação). Para verificar a cobertura empírica dos ICs acima:

1. Fixe um valor de θ (por ex., $\theta = \widehat{\theta}_{MV} = 2.7027$) e tamanhos $n \in \{6, 10, 20, 30, 50\}$.
2. Para cada par (θ, n), repita B vezes (ex.: $B = 10{,}000$):
   (a) Gere $Z_1, \ldots, Z_n \overset{iid}{\sim} \text{Exp}(\theta)$.
   (b) Calcule $\widehat{\theta}_{MV}$, as estimativas $\widehat{g}$ e os ICs (delta) de (a)–(c).
   (c) Registre se o verdadeiro $g(\theta)$ caiu dentro do IC.
3. Estime a cobertura como a frequência relativa de acertos. Compare com 95%.
4. Se a cobertura ficar abaixo de 95%, aumente n até estabilizar próximo de 95%.

Expectativa: Para $n = 6$, (a) e (c) tendem a cobertura abaixo de 95% (assimetria/limites fora do espaço). Cobertura melhora sensivelmente para $n \gtrsim 30$.

Questão 03: Seja $(Z_1, \ldots, Z_n)$ 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.

A distribuição Normal surge como limite assintótico da soma de variáveis aleatórias independentes com variância finita, conforme estabelecido pelo **Teorema Central do Limite (TCL)**.

TCL: Seja $\{X_1, X_2, \ldots, X_n\}$ uma sequência de variáveis aleatórias independentes e identicamente distribuídas (i.i.d.) com média $\mu = \mathbb{E}[X_i]$ e variância $\sigma^2 = \text{Var}(X_i) < \infty$. Definimos a média amostral

$$\bar{X}_n = \frac{1}{n} \sum_{i=1}^{n} X_i.$$

Então, a variável associada à média amostral centralizada e reescalada para ter média 0 e variância 1

$$Z_n = \frac{\bar{X}_n - \mu}{\sigma / \sqrt{n}}$$

converge em distribuição para uma Normal padrão, isto é,

$$Z_n \xrightarrow{d} \mathcal{N}(0, 1) \qquad \text{quando } n \to \infty.$$

Em outras palavras, independentemente da distribuição original das X_i, a média amostral tende, para amostras grandes, a seguir aproximadamente uma **distribuição Normal** com média μ e variância σ^2/n.

—

Para modelar matematicamente a função densidade da distribuição Normal, buscamos uma função contínua, que satisfaça as condições de simetria e normalização, da forma:

$$f(z) = A\, e^{-k(z-\mu)^2},$$

onde $A > 0$ e $k > 0$ são constantes a determinar. Essa escolha se justifica porque o termo $e^{-k(z-\mu)^2}$ é simétrico e decai rapidamente conforme $|z-\mu|$ aumenta, representando a concentração de probabilidade em torno de μ.

Impondo a condição de normalização $\int_{-\infty}^{+\infty} f(z)\, dz = 1$, temos:

$$A \int_{-\infty}^{+\infty} e^{-k(z-\mu)^2}\, dz = 1.$$

Usando a mudança de variável $x = \sqrt{k}\,(z-\mu)$, obtemos $dz = dx/\sqrt{k}$, e assim:

$$A \int_{-\infty}^{+\infty} e^{-x^2} \frac{dx}{\sqrt{k}} = 1.$$

Sabendo que $\displaystyle\int_{-\infty}^{+\infty} e^{-x^2}\, dx = \sqrt{\pi}$, segue que

$$A = \sqrt{\frac{k}{\pi}}.$$

Chamamos μ de **média** (ou valor esperado) e associamos k à **variância** σ^2 por meio da relação

$$k = \frac{1}{2\sigma^2},$$

que garante que $\mathrm{Var}(Z) = \sigma^2$. Substituindo A e k, obtemos:

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

Essa é a **função densidade de probabilidade** da distribuição Normal, denotada por $Z \sim \mathcal{N}(\mu, \sigma^2)$.

—

Agora, considerando uma amostra aleatória $(Z_1, \ldots, Z_n)$ de $Z \sim \mathcal{N}(\mu, \sigma^2)$, temos a função de densidade conjunta, a **função de verossimilhança**, como produtório das densidades individuais:

$$L(\mu, \sigma^2; z_1, \ldots, z_n) = \prod_{i=1}^{n} f(z_i \mid \mu, \sigma^2) = (2\pi\sigma^2)^{-n/2} \exp\left[-\frac{1}{2\sigma^2} \sum_{i=1}^{n} (z_i - \mu)^2\right].$$

Calculando a função **log verossimilhança**, para simplificar a maximização:

$$\ell(\mu, \sigma^2) = \log L(\mu, \sigma^2) = -\frac{n}{2}\log(2\pi) - \frac{n}{2}\log(\sigma^2) - \frac{1}{2\sigma^2} \sum_{i=1}^{n} (z_i - \mu)^2.$$

Sabendo que $\dfrac{d}{d\mu}(z_i-\mu)^2 = -2(z_i-\mu)$ e que a derivada é linear em relação à soma, derivamos em relação a μ:

$$\frac{\partial \ell}{\partial \mu} = -\frac{1}{2\sigma^2}\sum_{i=1}^{n}[-2(z_i-\mu)] = \frac{1}{\sigma^2}\sum_{i=1}^{n}(z_i-\mu).$$

Igualando a zero para encontrar o ponto crítico:

$$\frac{\partial \ell}{\partial \mu} = 0 \iff \sum_{i=1}^{n}(z_i-\mu) = 0 \iff n\mu = \sum_{i=1}^{n} z_i.$$

Assim, o **estimador de máxima verossimilhança** para μ é

$$\boxed{\widehat{\mu}_{MV} = \bar{Z} = \frac{1}{n}\sum_{i=1}^{n} Z_i.}$$

—

Agora derivamos a **log-verossimilhança** em relação a σ^2:

$$\frac{\partial \ell}{\partial(\sigma^2)} = -\frac{n}{2}\frac{1}{\sigma^2} + \frac{1}{2(\sigma^2)^2}\sum_{i=1}^{n}(z_i-\mu)^2.$$

Igualando a zero e substituindo $\mu = \widehat{\mu}_{MV}$:

$$0 = -\frac{n}{2\sigma^2} + \frac{1}{2(\sigma^2)^2}\sum_{i=1}^{n}(z_i-\bar{Z})^2.$$

Multiplicando ambos os lados por $2(\sigma^2)^2$ e rearranjando os termos:

$$n\sigma^2 = \sum_{i=1}^{n}(z_i-\bar{Z})^2 \quad\Longrightarrow\quad \boxed{\widehat{\sigma^2}_{MV} = \frac{1}{n}\sum_{i=1}^{n}(Z_i-\bar{Z})^2.}$$

—

Observação importante: O estimador de máxima verossimilhança da variância é *viciado* para amostras finitas, pois subestima a verdadeira dispersão. Calculando a esperança para o estimador obtido, tem-se que:

$$E[\widehat{\sigma^2}_{MV}] = \frac{n-1}{n}\sigma^2.$$

Esse viés ocorre porque a média amostral $\bar{Z}$ é calculada a partir dos próprios dados, introduzindo uma restrição linear entre os desvios $(Z_i-\bar{Z})$. Assim, apenas $n-1$ deles são linearmente independentes, o que caracteriza a perda de um grau de liberdade.

$$\sum_{i=1}^{n}(Z_i-\bar{Z}) = 0 \quad\Longrightarrow\quad (Z_n-\bar{Z}) = -\sum_{i=1}^{n-1}(Z_i-\bar{Z}).$$

Para eliminar o viés, multiplicamos o estimador por $\frac{n}{n-1}$, restaurando a estimativa não viciada da variância populacional:

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(Z_i-\bar{Z})^2, \qquad \text{para o qual} \quad E[S^2] = \sigma^2.$$

Essa correção garante que o estimador seja não viciado. Quando n é grande, a diferença entre os estimadores é desprezível, e $\widehat{\sigma^2}_{MV}$ é assintoticamente não viciado e consistente.

—

Derivando novamente para confirmar que o ponto crítico é de máximo:

$$\frac{\partial^2 \ell}{\partial \mu^2} = -\frac{n}{\sigma^2} < 0, \qquad \frac{\partial^2 \ell}{\partial (\sigma^2)^2} = \frac{n}{2(\sigma^2)^2} - \frac{1}{(\sigma^2)^3} \sum_{i=1}^{n} (z_i - \mu)^2.$$

Substituindo $\sum (z_i - \widehat{\mu}_{MV})^2 = n\widehat{\sigma}^2_{MV}$, obtemos valores negativos, confirmando o máximo.

—

Os estimadores de máxima verossimilhança para os parâmetros da distribuição Normal são:

$$\boxed{\widehat{\mu}_{MV} = \bar{Z}, \qquad \widehat{\sigma^2}_{MV} = \frac{1}{n} \sum_{i=1}^{n} (Z_i - \bar{Z})^2.}$$

E pela **propriedade de invariância do EMV**, para qualquer função $g(\mu, \sigma^2)$, temos:

$$\boxed{\widehat{g}_{MV} = g(\widehat{\mu}_{MV}, \widehat{\sigma}^2_{MV}).}$$

Resposta:

(a) $\quad g(\theta) = E_\theta(Z)$

Esperança matemática. A esperança (ou valor esperado) de uma variável aleatória é o análogo contínuo da média ponderada. Ela representa o valor médio que esperaríamos observar após infinitas repetições do experimento.

- **Caso discreto:** se Z assume valores z_i com probabilidades p_i, então
$$E(Z) = \sum_i z_i p_i.$$
- **Caso contínuo:** se Z tem densidade $f(z)$, então
$$E(Z) = \int_{-\infty}^{+\infty} z\, f(z)\, dz.$$

Exemplo: Normal $N(\mu, \sigma^2)$. A densidade é

$$f(z \mid \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(z-\mu)^2}{2\sigma^2} \right].$$

Então:

$$E(Z) = \int_{-\infty}^{+\infty} z\, f(z \mid \mu, \sigma^2)\, dz.$$

Fazendo a mudança de variável $x = \frac{z-\mu}{\sigma}$, obtemos

$$E(Z) = \int_{-\infty}^{+\infty} (\mu + \sigma x) \frac{1}{\sqrt{2\pi}} e^{-x^2/2}\, dx = \mu \underbrace{\int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-x^2/2}\, dx}_{=1} + \sigma \underbrace{\int_{-\infty}^{+\infty} \frac{x}{\sqrt{2\pi}} e^{-x^2/2}\, dx}_{=0}.$$

Logo,

$$\boxed{E(Z) = \mu.}$$

Fato conhecido da Normal: se $Z \sim N(\mu, \sigma^2)$ então $E(Z) = \mu$.

- ▷ **Como justificar rapidamente:** pela *linearidade da esperança* e pela *padronização* $Z = \mu + \sigma T$ com $T \sim N(0,1)$,
$$E(Z) = E(\mu + \sigma T) = \mu + \sigma E(T) = \mu + \sigma \cdot 0 = \mu.$$

Logo, $g(\theta) = \mu$ e, pela **invariância do EMV**,

$$\boxed{\widehat{g}_{(a)} = \widehat{\mu}_{MV} = \overline{Z}}$$

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

Passo 1 — Padronização (transformação linear): Se $Z \sim N(\mu, \sigma^2)$ então

$$T = \frac{Z - \mu}{\sigma} \sim N(0,1).$$

Passo 2 — Troca de variável na probabilidade (monotonicidade):

$$P_\theta(Z < 2) = P\left(\frac{Z-\mu}{\sigma} < \frac{2-\mu}{\sigma}\right) = \Phi\left(\frac{2-\mu}{\sigma}\right).$$

(Usamos que a função $z \mapsto (z-\mu)/\sigma$ é estritamente crescente quando $\sigma > 0$.)

Passo 3 — Invariância do EMV (plug-in):

$$\boxed{\widehat{g}_{(b)} = \Phi\left(\frac{2 - \widehat{\mu}_{MV}}{\widehat{\sigma}_{MV}}\right)}$$

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

Passo 1 — Padronização:

$$T = \frac{Z - \mu}{\sigma} \sim N(0,1).$$

Passo 2 — Regra da janela + CDF da Normal padrão:

$$P_\theta(2.6 < Z < 4) = P\left(\frac{2.6-\mu}{\sigma} < T < \frac{4-\mu}{\sigma}\right) = \Phi\left(\frac{4-\mu}{\sigma}\right) - \Phi\left(\frac{2.6-\mu}{\sigma}\right).$$

Passo 3 — Invariância do EMV (plug-in):

$$\boxed{\widehat{g}_{(c)} = \Phi\left(\frac{4 - \widehat{\mu}_{MV}}{\widehat{\sigma}_{MV}}\right) - \Phi\left(\frac{2.6 - \widehat{\mu}_{MV}}{\widehat{\sigma}_{MV}}\right)}$$

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

Variância. A **variância** de uma variável aleatória Z mede a dispersão dos valores em torno da média $E(Z)$. Formalmente, ela é definida por

$$\mathrm{Var}(Z) = E\big[(Z - E(Z))^2\big].$$

Definição expandida. Usando a propriedade de linearidade da esperança e a expansão do quadrado,

$$(Z - E(Z))^2 = Z^2 - 2Z\,E(Z) + E(Z)^2,$$

temos

$$\mathrm{Var}(Z) = E(Z^2) - 2E(Z)E(Z) + [E(Z)]^2 = E(Z^2) - [E(Z)]^2.$$

Portanto,

$$\boxed{\mathrm{Var}(Z) = E(Z^2) - [E(Z)]^2.}$$

Cálculo para a Normal. Se $Z \sim \mathcal{N}(\mu, \sigma^2)$, já sabemos que $E(Z) = \mu$. Logo precisamos calcular $E(Z^2)$.

Por definição:

$$E(Z^2) = \int_{-\infty}^{+\infty} z^2\, f(z \mid \mu, \sigma^2)\, dz, \quad \text{onde} \quad f(z \mid \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(z-\mu)^2}{2\sigma^2}\right].$$

Mudança de variável (padronização). Definimos $x = \frac{z-\mu}{\sigma}$, de modo que $z = \mu + \sigma x$ e $dz = \sigma dx$. Substituímos na integral:

$$E(Z^2) = \int_{-\infty}^{+\infty} (\mu + \sigma x)^2 \frac{1}{\sqrt{2\pi}} e^{-x^2/2}\, dx.$$

Expandindo o quadrado:

$$(\mu + \sigma x)^2 = \mu^2 + 2\mu\sigma x + \sigma^2 x^2.$$

Substituindo e separando termos:

$$E(Z^2) = \mu^2 \underbrace{\int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-x^2/2}\, dx}_{=1} + 2\mu\sigma \underbrace{\int_{-\infty}^{+\infty} \frac{x}{\sqrt{2\pi}} e^{-x^2/2}\, dx}_{=0} + \sigma^2 \underbrace{\int_{-\infty}^{+\infty} \frac{x^2}{\sqrt{2\pi}} e^{-x^2/2}\, dx}_{=1}.$$

Usamos propriedades da Normal padrão $X \sim N(0,1)$:

$$E(X) = 0, \quad E(X^2) = 1.$$

Logo:

$$E(Z^2) = \mu^2 + \sigma^2.$$

Substituindo na definição da variância:

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

Conclusão.

$$\boxed{\mathrm{Var}(Z) = \sigma^2.}$$

Assim, o parâmetro σ^2 da distribuição normal é, por definição, a variância populacional — ele controla a dispersão dos valores de Z em torno da média μ. **Fato conhecido da Normal:** se $Z \sim N(\mu, \sigma^2)$, então $\mathrm{Var}(Z) = \sigma^2$.

- ▷ **Justificativa curta:** pela padronização $Z = \mu + \sigma T$ com $T \sim N(0,1)$, usando *homogeneidade da variância* para fatores constantes,

$$\mathrm{Var}(Z) = \mathrm{Var}(\mu + \sigma T) = \sigma^2\, \mathrm{Var}(T) = \sigma^2 \cdot 1 = \sigma^2.$$

Logo, $g(\theta) = \sigma^2$ e, pela **invariância do EMV**, com o EMV já obtido para σ^2,

$$\boxed{\widehat{g}_{(d)} = \widehat{\sigma}^2_{MV} = \frac{1}{n}\sum_{i=1}^{n}(Z_i - \overline{Z})^2}$$

(Recordando: para a Normal com μ desconhecido, o EMV de σ^2 usa divisor n, não $n-1$.)

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

(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.

Objetivo. Procuramos uma densidade $f(x)$ em $(0,1)$, flexível o bastante para modelar proporções, com dois parâmetros de forma que controlem o comportamento perto de 0 e 1.

1) Escolha do núcleo. Uma família natural em $(0,1)$ é

$$f(x) \;\propto\; x^{\alpha-1}(1-x)^{\beta-1}, \qquad 0 < x < 1,$$

com $\alpha > 0,\ \beta > 0$. *Justificativas:* (i) $x^{\alpha-1}$ controla a massa perto de 0; (ii) $(1-x)^{\beta-1}$ controla a massa perto de 1; (iii) quando $\alpha = \beta$ a forma é simétrica; (iv) a família é conjugada à Binomial.

Logo, tomamos

$$f(x) = C\, x^{\alpha-1}(1-x)^{\beta-1}\, \mathbf{1}_{(0,1)}(x),$$

e determinamos a constante de normalização C impondo $\int_0^1 f(x)\,dx = 1$.

2) Normalização (Integral Beta). Defina a *função Beta* de Euler

$$B(\alpha,\beta) = \int_0^1 x^{\alpha-1}(1-x)^{\beta-1}\,dx, \qquad \alpha,\beta > 0.$$

Então

$$1 = \int_0^1 f(x)\,dx = C\int_0^1 x^{\alpha-1}(1-x)^{\beta-1}dx = C\,B(\alpha,\beta) \implies C = \frac{1}{B(\alpha,\beta)}.$$

Assim, a densidade fica

$$\boxed{f(x;\alpha,\beta) = \frac{1}{B(\alpha,\beta)}\, x^{\alpha-1}(1-x)^{\beta-1}, \quad 0 < x < 1.}$$

3) Ligação com Γ. Recorde a função Gama: $\Gamma(t) = \displaystyle\int_0^\infty u^{t-1}e^{-u}\,du$. Vale a identidade clássica

$$B(\alpha,\beta) = \frac{\Gamma(\alpha)\,\Gamma(\beta)}{\Gamma(\alpha+\beta)}.$$

Logo,

$$\boxed{f(x;\alpha,\beta) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\, x^{\alpha-1}(1-x)^{\beta-1}, \quad 0 < x < 1.}$$

4) Construção alternativa (razão de Gammas). Se $U \sim \text{Gamma}(\alpha, 1)$ e $V \sim \text{Gamma}(\beta, 1)$ são independentes e definimos

$$X = \frac{U}{U+V} \in (0,1), \qquad T = U + V \in (0, \infty),$$

então $X \sim \text{Beta}(\alpha, \beta)$.

Esboço da prova: O jacobiano da transformação $(u, v) \mapsto (x, t)$ é $\left|\frac{\partial(u,v)}{\partial(x,t)}\right| = t$. O conjunto transformado é $\{0 < x < 1,\ t > 0\}$. A densidade conjunta de (X, T) é

$$f_{X,T}(x,t) = \frac{t^{\alpha+\beta-1}}{\Gamma(\alpha)\Gamma(\beta)}\, x^{\alpha-1}(1-x)^{\beta-1} e^{-t}.$$

Integrando t em $(0, \infty)$,

$$f_X(x) = \int_0^\infty f_{X,T}(x,t)\, dt = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\, x^{\alpha-1}(1-x)^{\beta-1} = \frac{1}{B(\alpha,\beta)}\, x^{\alpha-1}(1-x)^{\beta-1}.$$

5) Momentos (fórmulas conhecidas). Para $X \sim \text{Beta}(\alpha, \beta)$,

$$E[X] = \frac{\alpha}{\alpha+\beta}, \qquad \text{Var}(X) = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}.$$

No caso particular $\text{Beta}(\theta, 1)$, obtém-se

$$E[X] = \frac{\theta}{\theta+1}, \qquad \text{Var}(X) = \frac{\theta}{(\theta+1)^2(\theta+2)}.$$

Distribuição Beta. A distribuição Beta é contínua no intervalo $(0, 1)$, com densidade

$$f(x; \alpha, \beta) = \frac{1}{B(\alpha,\beta)}\, x^{\alpha-1}(1-x)^{\beta-1}, \quad 0 < x < 1,$$

onde $B(\alpha, \beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$ é a função Beta. No caso particular $\text{Beta}(\theta, 1)$,

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

e a normalização é verificada por

$$\int_0^1 \theta x^{\theta-1}\, dx = 1.$$

Modelo. Se $X \sim \text{Beta}(\theta, 1)$, então a densidade é

$$f_\theta(x) = \theta\, x^{\theta-1}\, \mathbf{1}_{(0,1)}(x), \qquad \theta > 0.$$

Esperança $E[X]$. Usamos a definição (caso contínuo) e regras elementares de integração:

$$E[X] = \int_0^1 x\, f_\theta(x)\, dx = \int_0^1 x\, \theta\, x^{\theta-1}\, dx = \theta \int_0^1 x^\theta\, dx.$$

Propriedade usada: $\displaystyle\int_0^1 x^a\, dx = \frac{1}{a+1}$ para $a > -1$.

Como $\theta > 0$, segue que

$$E[X] = \theta \cdot \frac{1}{\theta+1} = \boxed{\frac{\theta}{\theta+1}}.$$

Segundo momento $E[X^2]$. De modo análogo,

$$E[X^2] = \int_0^1 x^2 f_\theta(x)\,dx = \int_0^1 x^2\,\theta\,x^{\theta-1}\,dx = \theta\int_0^1 x^{\theta+1}\,dx = \theta\cdot\frac{1}{\theta+2} = \boxed{\frac{\theta}{\theta+2}}.$$

Variância $\mathrm{Var}(X)$. Pela definição, $\mathrm{Var}(X) = E[X^2] - (E[X])^2$:

$$\mathrm{Var}(X) = \frac{\theta}{\theta+2} - \left(\frac{\theta}{\theta+1}\right)^2 = \frac{\theta(\theta+1)^2 - \theta^2(\theta+2)}{(\theta+2)(\theta+1)^2}.$$

Álgebra no numerador:

$$\theta(\theta+1)^2 - \theta^2(\theta+2) = \theta\big(\theta^2+2\theta+1\big) - \theta^2(\theta+2) = \theta^3 + 2\theta^2 + \theta - \theta^3 - 2\theta^2 = \theta.$$

Logo,

$$\mathrm{Var}(X) = \boxed{\frac{\theta}{(\theta+1)^2(\theta+2)}}.$$

O parâmetro θ controla a concentração: valores $\theta < 1$ favorecem x próximos de 0, $\theta > 1$ favorecem x próximos de 1, e $\theta = 1$ dá a distribuição uniforme.

Função de verossimilhança. Como as observações são independentes e identicamente distribuídas, a densidade conjunta é o produto das individuais:

$$L(\theta; x_1, \dots, x_n) = \prod_{i=1}^n f_\theta(x_i) = \prod_{i=1}^n \theta x_i^{\theta-1}.$$

Aplicando a propriedade distributiva do produto:

$$L(\theta; x) = \theta^n \prod_{i=1}^n x_i^{\theta-1} = \theta^n \exp\left\{(\theta-1)\sum_{i=1}^n \log x_i\right\}.$$

—

Função log-verossimilhança. Usando as propriedades dos logaritmos:

$$\begin{aligned} &\text{(i)}\ \log(ab) = \log a + \log b,\\ &\text{(ii)}\ \log(a^b) = b\log a,\\ &\text{(iii)}\ \log\Big(\prod_i a_i\Big) = \sum_i \log a_i, \end{aligned}$$

temos:

$$\ell(\theta) = \log L(\theta; x) = n\log\theta + (\theta-1)\sum_{i=1}^n \log x_i.$$

—

Derivada primeira. Aplicando as propriedades diferenciais:

$$\frac{d}{d\theta}\log\theta = \frac{1}{\theta}, \qquad \frac{d}{d\theta}\big[(\theta-1)c\big] = c,$$

obtemos:

$$\ell'(\theta) = \frac{n}{\theta} + \sum_{i=1}^n \log x_i.$$

—

Ponto crítico (máxima verossimilhança). Igualando a derivada a zero:

$$\ell'(\theta) = 0 \iff \frac{n}{\theta} + \sum_{i=1}^{n} \log x_i = 0 \iff \boxed{\widehat{\theta}_{MV} = -\frac{n}{\displaystyle\sum_{i=1}^{n} \log x_i}}.$$

Como $0 < x_i < 1$, temos $\log x_i < 0$, logo o estimador é positivo.

—

Segunda derivada.

$$\ell''(\theta) = \frac{d}{d\theta}\left(\frac{n}{\theta}\right) = -\frac{n}{\theta^2} < 0,$$

garantindo que o ponto crítico é um *máximo*.

—

Resultado final (EMV).

$$\boxed{\widehat{\theta}_{MV} = -\frac{n}{\displaystyle\sum_{i=1}^{n} \log X_i}}, \qquad X_i \in (0,1).$$

—

Observações.

- O estimador de máxima verossimilhança é sempre positivo, pois $\log X_i < 0$.
- A estatística suficiente é $T = -\sum_{i=1}^{n} \log X_i$, que segue uma distribuição Gamma$(n, \text{taxa} = \theta)$.
- Pela *invariância do EMV*, para qualquer função $g(\theta)$,

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

  Por exemplo, para $g(\theta) = P_\theta(X > c) = 1 - c^\theta$, tem-se

  $$\widehat{g}_{MV} = 1 - c^{\widehat{\theta}_{MV}}.$$

Resposta:

Questão 04. Seja $(Z_1, \ldots, Z_n)$ uma amostra i.i.d. de densidade

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

e $f_\theta(x) = 0$ caso contrário. (Trata-se de uma Beta$(\theta, 1)$.)

EMV de θ. A verossimilhança é

$$L(\theta; \mathbf{z}) = \prod_{i=1}^{n} \theta\, z_i^{\theta-1} = \theta^n \prod_{i=1}^{n} z_i^{\theta-1} = \theta^n \Big(\prod_{i=1}^{n} z_i\Big)^{\theta-1}.$$

A log-verossimilhança é

$$\ell(\theta) = \log L(\theta; \mathbf{z}) = n \log\theta + (\theta - 1) \sum_{i=1}^{n} \log z_i.$$

Derivando e igualando a zero,

$$\ell'(\theta) = \frac{n}{\theta} + \sum_{i=1}^{n} \log z_i = 0 \quad \Longrightarrow \quad \widehat{\theta}_{MV} = -\frac{n}{\sum_{i=1}^{n} \log Z_i}.$$

A segunda derivada é

$$\ell''(\theta) = -\frac{n}{\theta^2} < 0 \quad (\theta > 0),$$

logo o ponto crítico é máximo global. Assim,

$$\boxed{\widehat{\theta}_{MV} = -\frac{n}{\sum_{i=1}^n \log Z_i}}.$$

Momentos e probabilidades. Para $X \sim f_\theta$, para $a > -\theta$,

$$\mathbb{E}_\theta(X^a) = \int_0^1 x^a\,\theta x^{\theta-1}dx = \theta\int_0^1 x^{a+\theta-1}dx = \frac{\theta}{a+\theta}.$$

Em particular,

$$\mathbb{E}_\theta(X) = \frac{\theta}{\theta+1}, \qquad \mathbb{E}_\theta(X^2) = \frac{\theta}{\theta+2}, \qquad \mathrm{Var}_\theta(X) = \frac{\theta}{(\theta+1)^2(\theta+2)}.$$

Para $0 < a < b \leq 1$,

$$P_\theta(a < X < b) = \int_a^b \theta x^{\theta-1}dx = \Big[x^\theta\Big]_a^b = b^\theta - a^\theta.$$

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

$$\mathbb{E}_\theta(Z) = \frac{\theta}{\theta+1} \quad \Longrightarrow \quad \boxed{\widehat{g}_{(a)} = \frac{\widehat{\theta}_{MV}}{\widehat{\theta}_{MV}+1}}.$$

(b) $g(\theta) = P_\theta(Z > 0.3)$**.**

$$P_\theta(Z > 0.3) = 1 - P_\theta(0 < Z \leq 0.3) = 1 - (0.3)^\theta \quad \Longrightarrow \quad \boxed{\widehat{g}_{(b)} = 1 - (0.3)^{\widehat{\theta}_{MV}}}.$$

(c) $g(\theta) = P_\theta(0 < Z < 0.1)$**.**

$$P_\theta(0 < Z < 0.1) = (0.1)^\theta \quad \Longrightarrow \quad \boxed{\widehat{g}_{(c)} = (0.1)^{\widehat{\theta}_{MV}}}.$$

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

$$\mathrm{Var}_\theta(Z) = \frac{\theta}{(\theta+1)^2(\theta+2)} \quad \Longrightarrow \quad \boxed{\widehat{g}_{(d)} = \frac{\widehat{\theta}_{MV}}{(\widehat{\theta}_{MV}+1)^2(\widehat{\theta}_{MV}+2)}}.$$

(e) Estimativas numéricas. Dados: (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,

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

Então,

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

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

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

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

Resumo: $\widehat{\theta}_{MV} = -n/\sum \log Z_i \approx 0.6449$ e, por invariância, as estimativas de (a)–(d) são os valores de $g(\theta)$ avaliados em $\widehat{\theta}$ como mostrado acima.