UNIVERSIDADE DE SÃO PAULO
Programa de Pos-Graduação em Estatística

Alex Monito Nhancololo

1 Exercício 1: Seja $(Z_1, ..., Z_n)$ uma amostra aleatória de $Z \sim$ Ber(θ), $\theta \in (0,1)$

a) EMV para $g(\theta) = P_\theta(Z=0)$ Como $Z \sim \text{Ber}(\theta)$, temos que:

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

$$L_{z_n}(\theta) = \prod_{i=1}^{n} P_\theta(Z=z_i) = \prod_{i=1}^{n} \theta^{z_i}(1-\theta)^{1-z_i} = \theta^{\sum_{i=1}^n z_i}(1-\theta)^{n-\sum_{i=1}^n z_i}$$
$$\ell(\theta) = \log\left(\theta^{\sum_{i=1}^n z_i}(1-\theta)^{n-\sum_{i=1}^n z_i}\right) = \left(\sum_{i=1}^n z_i\right)\log\theta + \left(n - \sum_{i=1}^n z_i\right)\log(1-\theta)$$
$$\frac{d\ell}{d\theta} = \frac{\sum_{i=1}^n z_i}{\theta} - \frac{n-\sum_{i=1}^n z_i}{1-\theta} = U_n(\theta)$$
$$\frac{d\ell}{d\theta} = 0 \Longrightarrow \frac{\sum_{i=1}^n z_i}{\hat\theta} - \frac{n-\sum_{i=1}^n z_i}{1-\hat\theta} = 0 \Longleftrightarrow \frac{\sum_{i=1}^n z_i}{\hat\theta} = \frac{n-\sum_{i=1}^n z_i}{1-\hat\theta}$$
$$\left(\sum_{i=1}^n z_i\right)\left(1-\hat\theta\right) = \hat\theta\left(n - \sum_{i=1}^n z_i\right)$$
$$\sum_{i=1}^n z_i - \cancel{\hat\theta\sum_{i=1}^n z_i} = \hat\theta n - \cancel{\hat\theta\sum_{i=1}^n z_i}$$
$$\sum_{i=1}^n z_i = \hat\theta n$$
$$\hat\theta = \frac{\sum_{i=1}^n z_i}{n}$$
$$\hat\theta_{MV}(z_n) = \bar{Z}, \text{ pois } \frac{dU_n(\theta)}{d(\theta)} = -\frac{\sum_{i=1}^n z_i}{\theta^2} - \frac{n-\sum_{i=1}^n z_i}{(1-\theta)^2} < 0$$

Pelo princípio da invariância do EMV, o EMV para $g(\theta) = 1-\theta$ é: $\widehat{g(\theta)}_{MV} = 1 - \hat\theta_{MV}(z_n) = 1 - \bar{Z}$.

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

Como $Z \sim \text{Ber}(\theta)$, e $\text{Var}_\theta(Z) = \theta(1-\theta)$, e já calculou-se o EMV para θ, temos que pelo princípio da invariância, o EMV para $g(\theta) = \text{Var}_\theta(Z)$ é: $\widehat{g(\theta)}_{MV} = \hat\theta_{MV}(z_n)(1-\hat\theta_{MV}(z_n)) = \bar{Z}(1-\bar{Z})$.

c) Estimativas para os dados $(0,0,1,0,0,1)$
$\Longrightarrow$ Para $g(\theta) = \theta$

$$\hat\theta_{MV}(z_n) = \frac{\sum_{i=1}^n z_i}{n} = \frac{0+0+1+0+0+1}{6} = \frac{2}{6} = \frac{1}{3}, n = 6.$$

$\Longrightarrow$ Para $g(\theta) = P_\theta(Z=0) = 1-\theta$:

$$\widehat{g(\theta)}_{MV} = 1 - \hat\theta_{MV}(z_n) = 1 - \bar{z} = 1 - \frac{1}{3} = \frac{2}{3}.$$

$\Longrightarrow$ Para $g(\theta) = \text{Var}_\theta(Z) = \theta(1-\theta)$:

$$\widehat{g(\theta)} = \hat\theta_{MV}(z_n)(1-\hat\theta_{MV}(z_n)) = \bar{z}(1-\bar{z}) = \frac{1}{3}\cdot\left(1-\frac{1}{3}\right) = \frac{1}{3}\cdot\frac{2}{3} = \frac{2}{9}.$$

d) Valor-p para $H_0 : \theta = 0.1$

Do exercício 1a), temos: $L_{z_n}(\theta) = \theta^{\sum_{i=1}^n z_i}(1-\theta)^{n-\sum_{i=1}^n z_i}$. Pelo teorema da fatoração de Neyman, vemos que a verossimilhança depende dos dados apenas através da quantidade $\sum_{i=1}^n z_i$, portanto, $\sum_{i=1}^n z_i$ contém toda a informação da amostra relevante para inferência sobre o parâmetro θ. Isto é, a estatística suficiente é $T = \sum_{i=1}^n z_i$

Assim, sob H_0, $T = \sum_{i=1}^n z_i \sim \text{Bin}(n = 6, \theta = 0.1)$. O valor observado é $T = \sum_{i=1}^6 z_i = 0 + 0 + 1 + 0 + 0 + 1 = 2$. O valor-p bilateral é definido como:

$$\text{Valor-p}(H_0, \boldsymbol{z}_n) = 2 \cdot \min \left\{ \sup_{\theta \in \Theta_0} P_\theta^{(n)}(T_{H_0}(\boldsymbol{Z}_n) \geq T_{H_0}(\boldsymbol{z}_n)),\ \sup_{\theta \in \Theta_0} P_\theta^{(n)}(T_{H_0}(\boldsymbol{Z}_n) \leq T_{H_0}(\boldsymbol{z}_n)) \right\}.$$

Como $\Theta_0 = \{\theta = 0.1\}$ e $T = T_{H_0}(\boldsymbol{z}_n) = \sum_{i=1}^n z_i = 2$, tem-se:

$$\sup_{\theta \in \Theta_0} P_\theta^{(n)}(T_{H_0}(\boldsymbol{Z}_n) \geq 2) = P_{0.1}(T \geq 2),$$
$$\sup_{\theta \in \Theta_0} P_\theta^{(n)}(T_{H_0}(\boldsymbol{Z}_n) \leq 2) = P_{0.1}(T \leq 2).$$

Sob H_0 a função de probabilidade é binomial (soma de Z experimentos Bernoulli), dada por:

$$\begin{aligned}
P_\theta(T = k) = \binom{n}{k}\theta^k(1-\theta)^{n-k} = P_{0.1}(T = k) &= \binom{6}{k}(0.1)^k(0.9)^{6-k}, \quad k = 0, 1, \ldots, 6. \\
P_{0.1}(T = 0) &= \binom{6}{0}(0.1)^0(0.9)^6 = 1 \cdot 1 \cdot 0.9^6 = 0.531441, \\
P_{0.1}(T = 1) &= \binom{6}{1}(0.1)^1(0.9)^5 = 6 \cdot 0.1 \cdot 0.9^5 = 0.354294, \\
P_{0.1}(T = 2) &= \binom{6}{2}(0.1)^2(0.9)^4 = 15 \cdot 0.01 \cdot 0.9^4 = 0.098415, \\
P_{0.1}(T = 3) &= \binom{6}{3}(0.1)^3(0.9)^3 = 20 \cdot 0.001 \cdot 0.9^3 = 0.01458, \\
P_{0.1}(T = 4) &= \binom{6}{4}(0.1)^4(0.9)^2 = 15 \cdot 0.0001 \cdot 0.9^2 = 0.001215, \\
P_{0.1}(T = 5) &= \binom{6}{5}(0.1)^5(0.9)^1 = 6 \cdot 0.00001 \cdot 0.9 = 0.000054, \\
P_{0.1}(T = 6) &= \binom{6}{6}(0.1)^6(0.9)^0 = 1 \cdot 0.000001 \cdot 1 = 0.000001.
\end{aligned}$$

Assim, temos:

$$\begin{aligned}
P_{0.1}(T \geq 2) &= P(T = 2) + P(T = 3) + P(T = 4) + P(T = 5) + P(T = 6) \\
&= 0.098415 + 0.01458 + 0.001215 + 0.000054 + 0.000001 = 0.114265, \\
P_{0.1}(T \leq 2) &= P(T = 0) + P(T = 1) + P(T = 2) = 0.531441 + 0.354294 + 0.098415 = 0.98415.
\end{aligned}$$

$$\text{Valor-p}(H_0, \boldsymbol{z}_n) = 2 \cdot \min\left(P_{0.1}(T \geq 2), P_{0.1}(T \leq 2)\right) = 2 \cdot \min(0.114265, 0.98415) = 2 \cdot 0.114265 = 0.22853.$$

Nota: Não foi especificada a H_a, no caso de ser unilateral, bastaria apenas um Valor-p$(H_0, \boldsymbol{z}_n) = 0.114265$, análogo à fórmula apresentada nas notas de aula.

2 Exercício: Seja $(Z_1, ..., Z_n)$ uma amostra aleatória de $Z \sim \exp(\theta), \theta \in (0, \infty)$

a) EMV para $g(\theta) = P_\theta(Z > 1)$
Como $Z \sim \exp(\theta)$, a função de densidade de Z é:

$$\begin{aligned}
&P_\theta(Z = z) = f_\theta(z) = \theta e^{-\theta z},\ z > 0 \\
&P_\theta(Z > 1) = \int_1^\infty \theta e^{-\theta z} dz = \lim_{a \to \infty} \int_1^a \theta e^{-\theta z} dz = \lim_{a \to \infty} \int_1^a -d(e^{-\theta z}) = \lim_{a \to \infty} (-) \exp(-\theta z)\Big|_1^a \\
&\qquad = -\lim_{a \to \infty} [\exp(-\theta \times 1) - \exp(-\theta \times a)] = \exp(-\theta) - \underbrace{\cancel{\exp(-\theta \times \infty)}}_{0} \\
&\qquad = \exp(-\theta) = g(\theta)
\end{aligned}$$

A função de verossimilhança é:

$$
\begin{aligned}
L_{z_n}(\theta) &= \prod_{i=1}^{n} P_\theta(Z = z_i) = \prod_{i=1}^{n} \theta e^{-\theta z_i} = \theta^n e^{-\theta \sum_{i=1}^n z_i},\ z_i > 0 \\
\ell(\theta) &= \log(L_{z_n}(\theta)) = \log\left(\theta^n e^{-\theta \sum_{i=1}^n z_i}\right) = n \log \theta - \theta \sum_{i=1}^{n} z_i = U_n(\theta) \\
\frac{d\ell}{d\theta} &= \frac{n}{\theta} - \sum_{i=1}^{n} z_i \\
\frac{d\ell}{d\theta} &= 0 \Longrightarrow \frac{n}{\hat{\theta}} - \sum_{i=1}^{n} z_i = 0 \Longleftrightarrow \sum_{i=1}^{n} z_i = \frac{n}{\hat{\theta}} \Longleftrightarrow \hat{\theta} \sum_{i=1}^{n} z_i = n \\
\hat{\theta} &= \frac{n}{\sum_{i=1}^n z_i} \times \frac{\frac{1}{n}}{\frac{1}{n}} = \frac{1}{\bar{Z}} \\
\hat{\theta}_{MV}(z_n) &= \frac{1}{\bar{z}},\ \text{ pois, } \frac{d\,U_n(\theta)}{d\theta} = \frac{d^2\ell}{d\theta^2} = -\frac{n}{\theta^2} < 0
\end{aligned}
$$

Assim, pelo princípio da invariância do EMV, o EMV para $g(\theta) = \exp(-\theta)$ é:
$\widehat{g(\theta)}_{MV} = g(\hat{\theta}_{MV}(z_n)) = \exp(-\hat{\theta}_{MV}(z_n)) = \exp(-1/\bar{Z})$

b) EMV para $g(\theta) = P_\theta(0.1 < Z < 1)$

$$
g(\theta) = P_\theta(0.1 < Z < 1) = \int_{0.1}^{1} \theta \exp(-\theta \cdot z) dz = [-\exp(-\theta \cdot z)]_{0.1}^{1} = \exp(-0.1 \cdot \theta) - \exp(-\theta)
$$

Pelo princípio da invariância:

$$
\widehat{g(\theta)}_{MV} = g(\hat{\theta}_{MV}(z_n)) = \exp(-0.1/\bar{Z}) - \exp(-1/\bar{Z}).
$$

c) EMV para $g(\theta) = \mathrm{Var}_\theta(Z)$
Como $Z \sim \mathrm{Exp}(\theta)$, temos que $\mathrm{Var}_\theta(Z) = \frac{1}{\theta^2}$. Pelo princípio da invariância:

$$
\widehat{g(\theta)}_{MV} = \frac{1}{\hat{\theta}^2_{MV}(z_n)} = \frac{1}{\left(\frac{1}{\bar{z}}\right)^2} = \bar{Z}^2.
$$

d) IC aproximado de 95% para $g(\theta)$ com dados $(0.2, 0.6, 0.3, 0.2, 0.8, 0.12)$
Pela teoria assintótica do EMV, temos: $\sqrt{n}(\hat{\theta}_n - \theta) \xrightarrow{D} N(0, V_\theta)$, onde V_θ (a variância assintótica). Se V_θ é contínua e com derivada contínua em relação a θ, temos: $\sqrt{n}V_\theta^{-\frac{1}{2}}(\hat{\theta} - \theta) \xrightarrow{D} N_p(0, I_p(\theta))$ e $Q_n(\mathbf{Z}_n) = n(\hat{\theta}_n - \theta)^\top V_\theta^{-1}(\hat{\theta}_n - \theta) \xrightarrow{D} \chi^2_p$.
Assim, o IC aproximado de γ para θ é: $C_\gamma(\mathbf{Z}_n) = \{\theta \in \Theta : Q_n(\mathbf{Z}_n) \leq q_{\gamma,p}\}$, em que $q_{\gamma,p}$ é o quantil γ de uma χ^2_p.
Pelo exercício 2a) , sabe-se que a função score é: $U_n(\theta) = \frac{d\ell}{d\theta} = \frac{n}{\theta} - \sum_{i=1}^n z_i$. Assim, a informação de Fisher é dada por:

$$
I_n(\theta) = -E\left[\frac{d^2\,\ell}{d\,\theta^2}\right] = -E\left[\frac{d\,U_n(\theta)}{d\,\theta}\right] = -E\left[-\frac{n}{\theta^2}\right] = \frac{n}{\theta^2}.
$$

Portanto, a variância assintótica do EMV $\hat{\theta}_n$ é: $\mathrm{Var}_{\text{ass}}(\hat{\theta}_n) = I_n(\theta)^{-1} = \frac{\theta^2}{n}$.
Pelo que, $V_\theta = \mathrm{Var}\left(\sqrt{n}(\hat{\theta}_n - \theta)\right) = n \cdot \mathrm{Var}_{\text{ass}}(\hat{\theta}_n) = n \cdot \frac{\theta^2}{n} = \theta^2$. Assim, $V_\theta = \theta^2$, e $\sqrt{n}(\hat{\theta}_n - \theta) \xrightarrow{D} N\left(0, \theta^2\right)$.
Para uma função $g : \Theta \to \mathbb{R}$ diferenciável, pelo método delta, temos que: $\sqrt{n}(g(\hat{\theta}_n) - g(\theta)) \xrightarrow{D} N(0, [g'(\theta)]^2 V_\theta)$ e $\frac{\sqrt{n}(g(\hat{\theta}_n) - g(\theta))}{\sqrt{[g'(\theta)]^2 V_\theta}} = \frac{g(\hat{\theta}_n) - g(\theta)}{\sqrt{\frac{[g'(\theta)]^2 V_\theta}{n}}} = \frac{\sqrt{n}\left(g(\hat{\theta}_n) - g(\theta)\right)}{\sqrt{[g'(\theta)]^2\theta^2}} = \frac{\sqrt{n}\left(g(\hat{\theta}_n) - g(\theta)\right)}{|g'(\theta)\theta|} \xrightarrow{D} N(0, 1)$
Um IC aproximado de γ para $g(\theta)$ é: $C_\gamma^{(g)}(\mathbf{Z}_n) = \left\{g(\theta) : \left|\frac{g(\hat{\theta}_n) - g(\theta)}{\frac{|g'(\hat{\theta}_n)|\hat{\theta}_n}{\sqrt{n}}}\right| \leq z_{(1+\gamma)/2}\right\}$, onde $z_{(1+\gamma)/2}$ é o quantil da normal padrão tal que $P(|Z| \leq z_{(1+\gamma)/2}) = \gamma$.

$$
\begin{aligned}
C_\gamma^{(g)}(\mathbf{Z}_n) &= \left\{g(\theta) : \left|\frac{g(\hat{\theta}_n) - g(\theta)}{\frac{|g'(\hat{\theta}_n)|\hat{\theta}_n}{\sqrt{n}}}\right| \leq z_{(1+\gamma)/2}\right\} = \left\{g(\theta) : \left|g(\hat{\theta}_n) - g(\theta)\right| \leq \frac{z_{\left(\frac{1+\gamma}{2}\right)} \times \left|g'(\hat{\theta}_n)\right|\hat{\theta}_n}{\sqrt{n}}\right\} \\
&= \left\{g(\theta) : g(\hat{\theta}_n) \pm \frac{z_{\left(\frac{1+\gamma}{2}\right)} \times \left|g'(\hat{\theta}_n)\right|\hat{\theta}_n}{\sqrt{n}}\right\} \\
&= \left\{g(\theta) : g(\hat{\theta}_n) \pm 1.96 \cdot \frac{|g'(\hat{\theta}_n)|\hat{\theta}_n}{\sqrt{n}}\right\}
\end{aligned}
$$

onde $z_{(1+\gamma)/2}$ é o quantil da normal padrão tal que $P(|Z| \leq z_{(1+\gamma)/2}) = \gamma$, com $Z \sim N(0,1)$. Para $\gamma = 0.95$, $z_{(1+\gamma)/2} = z_{0.975} \approx 1.96$.

Nota: Aqui fiz considerando que $Z \sim N(0,1)$, porém, sabe-se que $Z^2 \sim \chi^2_1$, então poderia usar teste Qui-quadrado.

$\Longrightarrow$ Dados: $(0.2, 0.6, 0.3, 0.2, 0.8, 0.12)$, $n = 6$

$$\bar{z} = \frac{\sum_{i=1}^{n}}{n} = \frac{0.2 + 0.6 + 0.3 + 0.2 + 0.8 + 0.12}{6} = 0.37, \quad \hat{\theta}_n = \theta_{MV}(z_n) = \frac{1}{0.37} \approx 2.7027$$

a) Para $g(\theta) = \exp(-\theta) = P_\theta(Z > 1)$

$$\begin{aligned}
\widehat{g(\theta)} &= \exp(-\hat{\theta}_{MV}(z_n)) = \exp(-1/\bar{z}) = e^{-2.7027} \approx 0.0671 \\
g'(\theta) &= -\exp(-\theta) \Rightarrow g'(\hat{\theta}_n) = -\widehat{g(\theta)} \approx -0.0671 \\
\frac{|g'(\hat{\theta}_n)|\hat{\theta}_n}{\sqrt{n}} &= \frac{0.0671 \times 2.7027}{\sqrt{6}} \approx 0.0740 \\
\text{IC}_{95\%} &= 0.0671 \pm 1.96 \times 0.0740 = (-0.0779, 0.2121)
\end{aligned}$$

Nota: O IC inclui valores negativos, o que não é válido para uma probabilidade. Truncar o intervalo para $[0, 0.2121]$ pode ser ideal. No entanto, isso introduz um viés, pois a cobertura real do IC pode ser menor que o nível nominal. Uma abordagem mais rigorosa seria usar métodos de bootstrap ou transformações que garantam limites válidos.

Como $g(\theta) = P_\theta(Z > 1) = e^{-\theta} > 0$, usamos a transformação logarítmica.

Item (a): IC com transformação logarítmica para $g(\theta) = e^{-\theta}$

$$\begin{aligned}
h(\theta) &= \log(g(\theta)) = -\theta \\
\widehat{h(\theta)} &= -\hat{\theta} = -2.7027 \\
h'(\theta) &= -1 \\
\text{Var}(\widehat{h(\theta)}) &\approx [h'(\theta)]^2 \cdot \frac{\hat{\theta}^2}{n} = (-1)^2 \cdot \frac{(2.7027)^2}{6} \approx \frac{7.3046}{6} \approx 1.2174 \\
\text{EP} &= \sqrt{\text{Var}(\widehat{h(\theta)})} \approx \sqrt{1.2174} \approx 1.1034 \\
\text{IC para } h(\theta) &: \widehat{h(\theta)} \pm z_{0.975} \cdot \text{EP} = -2.7027 \pm 1.96 \times 1.1034 = [-4.8654, -0.5400] \\
\text{IC para } g(\theta) &: \left[e^{-4.8654}, e^{-0.5400}\right] \approx [0.0077, 0.5828]
\end{aligned}$$

Justificativa: A transformação logarítmica garante que o intervalo esteja contido em $(0, \infty)$, evitando valores negativos inadequados para uma probabilidade. O IC resultante $[0.0077, 0.5828]$ é assintoticamente válido e preserva a interpretação probabilística.

b) Para $g(\theta) = P_\theta(0.1 < Z < 1) = \exp(-0.1 \cdot \theta) - \exp(-\theta)$

$$\begin{aligned}
\widehat{g(\theta)} &= \exp(-0.1/\bar{z}) - \exp(-1/\bar{z}) = e^{-0.27027} - e^{-2.7027} \approx 0.7632 - 0.0671 = 0.6961 \\
g'(\theta) &= -0.1\exp(-0.1 \cdot \theta) + \exp(-\theta) \\
g'(\hat{\theta}_n) &= -0.1 \times \exp(-0.1/\bar{z}) + \exp(-1/\bar{z}) \approx -0.1 \times 0.7632 + 0.0671 = -0.00922 \\
\frac{|g'(\hat{\theta}_n)|\hat{\theta}_n}{\sqrt{n}} &= \frac{0.00922 \times 2.7027}{\sqrt{6}} \approx 0.01017 \\
\text{IC}_{95\%} &= 0.6961 \pm 1.96 \times 0.01017 = (0.6762, 0.7160)
\end{aligned}$$

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

$$\begin{aligned}
\widehat{g(\theta)} &= \frac{1}{\hat{\theta}^2_{MV}(z_n)} = \frac{1}{2.7027^2} \approx 0.1369 \\
g'(\theta) &= -\frac{2}{\theta^3} \Rightarrow g'(\hat{\theta}_n) \approx -\frac{2}{2.7027^3} \approx -0.1012 \\
\frac{|g'(\hat{\theta}_n)|\hat{\theta}_n}{\sqrt{n}} &= \frac{0.1012 \times 2.7027}{\sqrt{6}} \approx 0.1116 \\
\text{IC}_{95\%} &= 0.1369 \pm 1.96 \times 0.1116 = (-0.0818, 0.3556)
\end{aligned}$$

Nota: O IC inclui valores negativos, o que não é válido para uma probabilidade. Truncar o intervalo para $[0, 0.3556]$ pode ser ideal. No entanto, isso introduz um viés, pois a cobertura real do IC pode ser menor que o nível nominal.

Uma abordagem mais rigorosa seria usar métodos de bootstrap ou transformações que garantam limites válidos.

Como $g(\theta) = g(\theta) = \text{Var}_\theta(Z) = 1/\theta^2 > 0$, usamos a transformação logarítmica.

Item (c): IC com Transformação Logarítmica para $g(\theta) = \frac{1}{\theta^2}$

$$h(\theta) = \ln(g(\theta)) = -2\ln(\theta)$$
$$\widehat{h(\theta)} = -2\ln(2.7027) \approx -2 \times 0.9943 = -1.9886$$
$$h'(\theta) = -\frac{2}{\theta}$$
$$\text{Var}(\widehat{h(\theta)}) \approx [h'(\theta)]^2 \cdot \frac{\hat{\theta}^2}{n} = \left(-\frac{2}{\theta}\right)^2 \cdot \frac{\hat{\theta}^2}{n} = \frac{4}{n} = \frac{4}{6} \approx 0.6667$$
$$\text{EP} = \sqrt{\text{Var}(\widehat{h(\theta)})} \approx \sqrt{0.6667} \approx 0.8165$$
$$\text{IC para } h(\theta) : \widehat{h(\theta)} \pm z_{0.975} \cdot \text{EP} = -1.9886 \pm 1.96 \times 0.8165 = [-3.5889, -0.3883]$$
$$\text{IC para } g(\theta) : \left[e^{-3.5889}, e^{-0.3883}\right] \approx [0.0276, 0.6783]$$

Justificativa: A transformação logarítmica assegura que a variância estimada seja positiva e que o IC para $g(\theta)$ esteja contido em $(0, \infty)$, corrigindo a limitação do método delta direto. O IC $[0.0276, 0.6783]$ é adequado para uma variância.

e) Com base na simulação de Monte Carlo, os intervalos de confiança (ICs) aproximados para o tamanho amostral original de n=6 não são confiáveis. A Figura 1 demonstra que para esta amostra pequena a cobertura empírica se desvia consideravelmente do nível nominal de 95%, os intervalos são, em média, muito longos e os estimadores apresentam um viés elevado. Para obter resultados mais confiáveis, a análise da Figura 2 é fundamental, pois mostra que tanto a cobertura quanto o comprimento dos ICs melhoram com o aumento do tamanho amostral. Observa-se que a cobertura converge para o valor esperado de 95% e os intervalos se tornam mais precisos à medida que n aumenta. Portanto, propõe-se um tamanho amostral de n=50, pois a partir desse ponto a simulação indica que a cobertura se estabiliza no nível desejado, produzindo ICs aproximados mais confiáveis para todos os casos.

Figura 1: Análise de cobertura dos intervalos de confiança

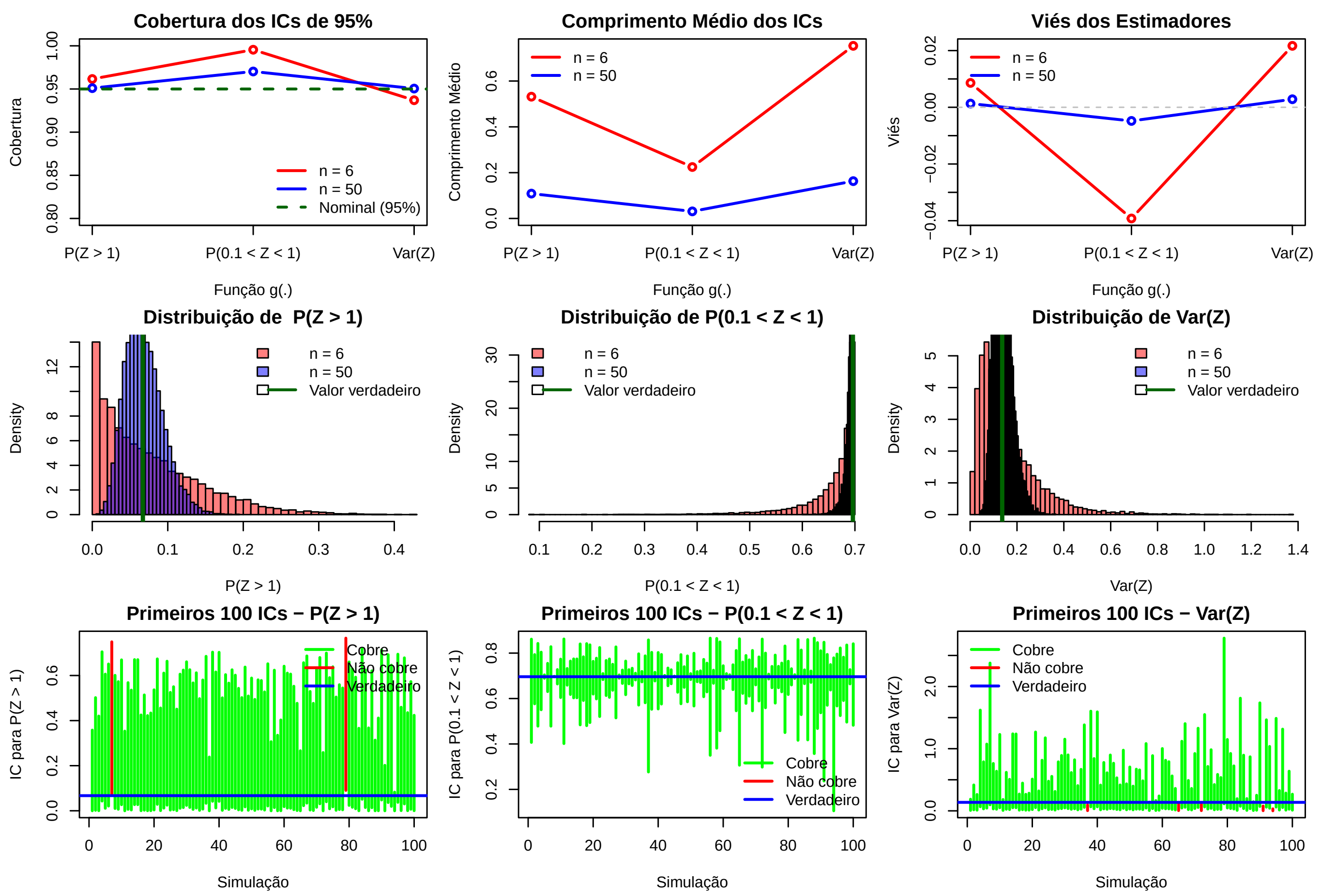
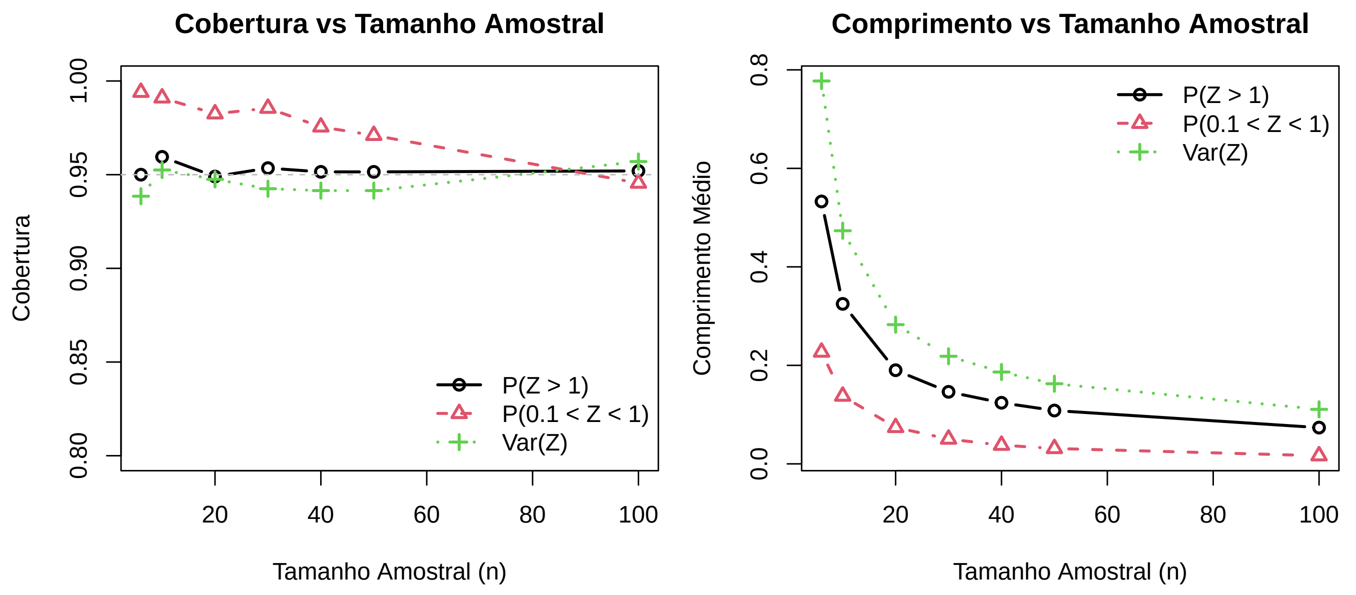


Figura 2: Análise de cobertura para diferentes tamanhos da amostra



3 Exercício: Seja $(Z_1, \ldots, Z_n)$ uma amostra aleatória de $Z \sim N(\mu, \sigma^2)$, com $\theta = (\mu, \sigma^2) \in (-\infty, \infty) \times (0, \infty)$

a) EMV para $g(\theta) = E_\theta(Z)$

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

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

$$L_{z_n}(\mu, \sigma^2) = \prod_{i=1}^{n} f(z_i|\mu, \sigma^2) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(z_i-\mu)^2}{2\sigma^2}\right) = \left(2\pi\sigma^2\right)^{-\frac{n}{2}} \exp\left(-\frac{\sum_{i=1}^{n}(z_i-\mu)^2}{2\sigma^2}\right)$$

$$\ell(\mu, \sigma^2) = \log L_{z_n}(\mu, \sigma^2) = \log\left[\prod_{i=1}^{n} f(z_i|\mu, \sigma^2)\right] = \log\left[\left(2\pi\sigma^2\right)^{-\frac{n}{2}} \exp\left(-\frac{\sum_{i=1}^{n}(z_i-\mu)^2}{2\sigma^2}\right)\right]$$

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

$$\frac{d\,\ell(\mu, \sigma^2)}{d\,\mu} = \frac{1}{\sigma^2}\sum_{i=1}^{n}(z_i-\mu) = U_n(\mu)$$

$$\frac{d\,\ell(\mu, \sigma^2)}{d\,\mu} = 0 \implies \sum_{i=1}^{n}(z_i-\hat{\mu}) = 0 \implies \hat{\mu}_{MV}(z_n) = \frac{1}{n}\sum_{i=1}^{n} z_i = \bar{Z}, \quad \text{pois } \frac{d\,U_n(\mu)}{d\,\mu} = -\frac{n}{\sigma^2} < 0$$

$$\widehat{g(\theta)} = \hat{\mu}_{MV}(z_n) = \bar{Z}$$

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

$$g(\theta) = P_\theta(Z<2) = P_\theta(\underbrace{\frac{Z-\mu}{\sigma}}_{\tilde{Z}\sim N(0,1)} < \frac{2-\mu}{\sigma}) = P_\theta(\tilde{Z} < \frac{2-\mu}{\sigma}) = \Phi\left(\frac{2-\mu}{\sigma}\right),$$

onde Φ é a função de distribuição acumulada da $\tilde{Z} \sim N(0,1)$

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

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

$$\frac{d\,\ell(\mu,\sigma^2)}{d\,\sigma^2} = 0 \implies -\frac{n}{2\widehat{\sigma^2}} + \frac{1}{2\left(\widehat{\sigma^2}\right)^2}\sum_{i=1}^n (z_i-\hat{\mu})^2 = 0 \iff \frac{n}{2\widehat{\sigma^2}} = \frac{1}{2\left(\widehat{\sigma^2}\right)^2}\sum_{i=1}^n (z_i-\hat{\mu})^2$$

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

$$\text{pois } \frac{d\,U_n(\sigma^2)}{d\,\sigma^2} = \frac{n}{2(\sigma^2)^2} - \frac{\sum_{i=1}^n (z_i-\bar{Z})^2}{(\sigma^2)^3} < 0$$

Assim, como EMV para $\mu : \hat{\mu}_{MV}(z_n) = \bar{Z}$ e EMV para $\sigma^2 : \widehat{\sigma^2}_{MV}(z_n) = \frac{1}{n}\sum_{i=1}^n (z_i-\bar{Z})^2$, pelo princípio da invariância

$$\widehat{g(\theta)} = \Phi\left(\frac{2-\hat{\mu}_{MV}(z_n)}{\widehat{\sigma}_{MV}(z_n)}\right) = \Phi\left(\frac{2-\bar{Z}}{\sqrt{\widehat{\sigma^2}_{MV}(z_n)}}\right)$$

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

$$g(\theta) = P_\theta(2.6<Z<4) = P_\theta\left(\frac{2.6-\mu}{\sigma} < \underbrace{\frac{Z-\mu}{\sigma}}_{\tilde{Z}\sim N(0,1)} < \frac{4-\mu}{\sigma}\right) = \Phi\left(\frac{4-\mu}{\sigma}\right) - \Phi\left(\frac{2.6-\mu}{\sigma}\right)$$

Pelo princípio da invariância:

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

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

$g(\theta) = \mathrm{Var}_\theta(Z) = \sigma^2$ e pelos cálculos acima, EMV para σ^2 é $\widehat{\sigma}^2_{MV}(z_n) = \frac{1}{n}\sum_{i=1}^n (z_i-\bar{Z})^2$. Assim, pelo princípio da invariância do EMV, temos

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

e) Estimativas de MV com dados $(2.4, 2.7, 2.3, 2, 2.5, 2.6)$

$$
\begin{aligned}
n &= 6 \\
\bar{z} &= \frac{2.4 + 2.7 + 2.3 + 2 + 2.5 + 2.6}{6} = \frac{14.5}{6} \approx 2.4167
\end{aligned}
$$

$$
\begin{aligned}
\widehat{\sigma}^2_{MV}(z_n) &= \frac{1}{n}\sum_{i=1}^{n}(z_i - \bar{Z})^2 \\
&= \frac{1}{6}\sum_{i=1}^{n}(z_i - \bar{z})^2 \\
&= \frac{1}{6}\left[(2.4 - 2.4167)^2 + (2.7 - 2.4167)^2 + (2.3 - 2.4167)^2 + (2 - 2.4167)^2 + (2.5 - 2.4167)^2 + (2.6 - 2.4167)^2\right] \\
&= \frac{1}{6}\left[(-0.0167)^2 + (0.2833)^2 + (-0.1167)^2 + (-0.4167)^2 + (0.0833)^2 + (0.1833)^2\right] \\
&= \frac{1}{6}\left[0.00027889 + 0.08026889 + 0.01361889 + 0.17363889 + 0.00693889 + 0.03359889\right] \\
&= \frac{1}{6} \times 0.30834334 \approx 0.05139 \\
\widehat{\sigma}_{MV}(z_n) &= \sqrt{0.05139} \approx 0.2267
\end{aligned}
$$

$\Longrightarrow$ Estimativas para cada item:
a) EMV para $g(\theta) = E_\theta(Z)$:

$$\widehat{g(\theta)} = \bar{z} = 2.4167$$

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

$$
\begin{aligned}
\widehat{g(\theta)} &= \Phi\left(\frac{2 - \hat{\mu}_{MV}(z_n)}{\hat{\sigma}_{MV}(z_n)}\right) = \Phi\left(\frac{2 - 2.4167}{0.2267}\right) = \Phi\left(\frac{-0.4167}{0.2267}\right) \approx \Phi(-1.838) \\
&= 1 - \Phi(1.838)
\end{aligned}
$$

$\Phi(1.838) \approx 0.9669$ (da tabela da normal padrão) **ver aqui**

$$\Rightarrow \widehat{g(\theta)} \approx 1 - 0.9669 = 0.0331$$

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

$$
\begin{aligned}
\widehat{g(\theta)} &= \Phi\left(\frac{4 - \bar{Z}}{\sqrt{\widehat{\sigma}^2_{MV}(z_n)}}\right) - \Phi\left(\frac{2.6 - \bar{Z}}{\sqrt{\widehat{\sigma}^2_{MV}(z_n)}}\right) = \Phi\left(\frac{4 - 2.4167}{0.2267}\right) - \Phi\left(\frac{2.6 - 2.4167}{0.2267}\right) \\
&= \Phi\left(\frac{1.5833}{0.2267}\right) - \Phi\left(\frac{0.1833}{0.2267}\right) \\
&= \Phi(6.984) - \Phi(0.808)
\end{aligned}
$$

$\Phi(6.984) \approx 1.0000$

$\Phi(0.808) \approx 0.7905$ (da tabela da normal padrão) **ver aqui**

$$\Rightarrow \widehat{g(\theta)} \approx 1.0000 - 0.7905 = 0.2095$$

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

$$\widehat{g(\theta)} = \widehat{\sigma}^2_{MV}(z_n) = \frac{1}{n}\sum_{i=1}^{n}(z_i - \bar{Z})^2 = 0.05139, \text{ Cálculos no início da 3e)}$$

4 Exercício

Seja $(Z_1, \ldots, Z_n)$ uma amostra aleatória de $Z \sim f_\theta$, com $\theta \in (0, \infty)$, onde a função densidade de probabilidade é dada por: $f_\theta(x) = \theta x^{\theta-1}$, para $x \in (0, 1)$, e $f_\theta(x) = 0$ caso contrário.

a) EMV para $g(\theta) = E_\theta(Z)$

$$
\begin{aligned}
g(\theta) &= E_\theta(Z) = \int_0^1 x \cdot \theta x^{\theta-1}\,dx = \theta \int_0^1 x^\theta\,dx = \theta \left[\frac{x^{\theta+1}}{\theta+1}\right]_0^1 = \frac{\theta}{\theta+1}.\\
L(\theta) &= \prod_{i=1}^n f_\theta(z_i) = \prod_{i=1}^n \theta z_i^{\theta-1} = \theta^n \left(\prod_{i=1}^n z_i\right)^{\theta-1}.\\
\ell(\theta) &= \log L(\theta) = n\log\theta + (\theta-1)\sum_{i=1}^n \log z_i.\\
\frac{d\ell}{d\theta} &= \frac{n}{\theta} + \sum_{i=1}^n \log z_i = U_n(\theta)
\end{aligned}
$$

$$
\frac{n}{\hat\theta} + \sum_{i=1}^n \log z_i = 0 \implies \hat\theta_{MV}(z_n) = -\frac{n}{\sum_{i=1}^n \log z_i}, \text{ pois }, \frac{d\,U_n(\theta)}{d\,\theta} = -\frac{n}{\theta^2} < 0
$$

Pelo princípio da invariância do EMV $\widehat{g(\theta)} = \frac{\hat\theta_{MV}(z_n)}{\hat\theta_{MV}(z_n)+1}$.
b) EMV para $g(\theta) = P_\theta(Z > 0.3)$

$$
g(\theta) = P_\theta(Z > 0.3) = \int_{0.3}^1 \theta x^{\theta-1}\,dx = \left[x^\theta\right]_{0.3}^1 = 1 - (0.3)^\theta.
$$

Pelo princípio da invariância do EMV $\widehat{g(\theta)} = 1 - (0.3)^{\hat\theta_{MV}(z_n)}$.

c) EMV para $g(\theta) = P_\theta(0 < Z < 0.1)$

$$
g(\theta) = P_\theta(0 < Z < 0.1) = \int_0^{0.1} \theta x^{\theta-1}\,dx = \left[x^\theta\right]_0^{0.1} = (0.1)^\theta.
$$

Pelo princípio da invariância do EMV $\widehat{g(\theta)} = (0.1)^{\hat\theta_{MV}(z_n)}$.

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

$$
\begin{aligned}
E_\theta(Z) &= \frac{\theta}{\theta+1},\\
E_\theta(Z^2) &= \int_0^1 x^2 \cdot \theta x^{\theta-1}\,dx = \theta\int_0^1 x^{\theta+1}\,dx = \theta\left[\frac{x^{\theta+2}}{\theta+2}\right]_0^1 = \frac{\theta}{\theta+2},\\
\mathrm{Var}_\theta(Z) &= E_\theta(Z^2) - [E_\theta(Z)]^2 = \frac{\theta}{\theta+2} - \left(\frac{\theta}{\theta+1}\right)^2.
\end{aligned}
$$

Pelo princípio da invariância do EMV $\widehat{g(\theta)} = \frac{\hat\theta_{MV}(z_n)}{\hat\theta_{MV}(z_n)+2} - \left(\frac{\hat\theta_{MV}(z_n)}{\hat\theta_{MV}(z_n)+1}\right)^2$.

e) Estimativas de MV com dados $(0.12, 0.50, 0.20, 0.23, 0.30, 0.11)$

$$
\begin{aligned}
n &= 6,\\
\sum_{i=1}^n \log z_i &= \log(0.12) + \log(0.50) + \log(0.20) + \log(0.23) + \log(0.30) + \log(0.11)\\
&\approx -2.1203 - 0.6931 - 1.6094 - 1.4697 - 1.2040 - 2.2073 = -9.3038,\\
\hat\theta_{MV}(z_n) &= -\frac{n}{\sum_{i=1}^n \log z_i} = -\frac{6}{-9.3038} \approx 0.6449.
\end{aligned}
$$

a) Estimativa para $g(\theta) = E_\theta(Z)$:

$$
\widehat{g(\theta)} = \frac{\hat\theta_{MV}(z_n)}{\hat\theta_{MV}(z_n)+1} = \frac{0.6449}{0.6449+1} \approx \frac{0.6449}{1.6449} \approx 0.3921.
$$

b) Estimativa para $g(\theta) = P_\theta(Z > 0.3)$:

$$
\begin{aligned}
(0.3)^{\hat\theta_{MV}(z_n)} &= \exp(0.6449 \cdot \log(0.3)) \approx \exp(0.6449 \cdot (-1.20397)) \approx \exp(-0.776) \approx 0.4603,\\
\widehat{g(\theta)} &= 1 - (0.3)^{\hat\theta_{MV}(z_n)} = 1 - 0.4603 \approx 0.5397.
\end{aligned}
$$

c) Estimativa para $g(\theta) = P_\theta(0 < Z < 0.1)$:

$$(0.1)^{\hat{\theta}_{MV}(z_n)} = \exp(0.6449 \cdot \log(0.1)) \approx \exp(0.6449 \cdot (-2.302585)) \approx \exp(-1.484) \approx 0.2269,$$
$$\widehat{g(\theta)} \approx 0.2269.$$

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

$$\frac{\hat{\theta}_{MV}(z_n)}{\hat{\theta}_{MV}(z_n) + 2} \approx \frac{0.6449}{2.6449} \approx 0.2438,$$
$$\left(\frac{\hat{\theta}_{MV}(z_n)}{\hat{\theta}_{MV}(z_n) + 1}\right)^2 \approx \left(\frac{0.6449}{1.6449}\right)^2 \approx (0.3921)^2 \approx 0.1537,$$
$$\widehat{g(\theta)} = \frac{\hat{\theta}_{MV}(z_n)}{\hat{\theta}_{MV}(z_n) + 2} - \left(\frac{\hat{\theta}_{MV}(z_n)}{\hat{\theta}_{MV}(z_n) + 1}\right)^2 = \widehat{g(\theta)} \approx 0.2438 - 0.1537 = 0.0901.$$