UNIVERSIDADE FEDERAL DE SANTA MARIA
DEPARTAMENTO DE ESTATÍSTICA
Disciplina: Estatística Inferencial
Professora: Laís Helen Loose

Lista 2 - Estimação e Distribuição amostral da média

1. Responda as seguintes perguntas e dê a devida justificativa:

   (a) Qual é a diferença entre um estimador e uma estatística?
   Estatística é qualquer função da amostra aleatória e que não envolve parâmetro. Estimador é uma estatística em que os seus possíveis valores são associados ao espaço paramétrico de um parâmetro que se quer estimar.

   (b) É certo dizer que toda estatística é um estimador?
   Não. Uma estatística só é um estimador se seus possíveis valores estão associados ao espaço paramétrico de um parâmetro que se quer estimar.

   (c) É certo que todo estimador é uma estatística?
   Sim.

   (d) É certo que toda função de uma amostra aleatória é uma estatística?
   Não necessariamente, podemos ter uma função da amostra aleatória envolvendo o parâmetro, o que não é uma estatística.

   (e) É certo que toda função de uma estatística é uma estatística?
   Sim.

2. Seja $X_1, X_2, \ldots, X_n$ uma amostra aleatória de $X \sim f(x;\theta)$ com θ desconhecido. Determine se as expressões das alternativas dadas abaixo são estatísticas:

   (a) $T = X_1$ Estatística.

   (b) $T = (X_1 + X_2)/2$ Estatística.

   (c) $T = X_1 + 2X_2 + \ldots + nX_n$ Estatística.

   (d) $T = (X_1 + 2X_2 + \ldots + nX_n) - \theta$ Não é uma estatística.

   (e) $T = \theta(X_{(n)} - X_{(1)})$ Não é uma estatística.

   (f) $T = X_1^2 + \ldots + X_n^2$ Estatística.

   (g) $T = (X_1 + \ldots + X_n)^2$ Estatística.

   (h) $T = \frac{1}{n}\sum_{i=1}^{n} \frac{X_i - E(X_i)}{\sqrt{Var(X_i)}}$ Não é uma estatística.

   (i) $T = \frac{1}{n}\sum_{i=1}^{n} \frac{X_i - \overline{X})}{\sqrt{S^2}}$ Estatística.

3. Foram sorteadas 15 famílias com filhos num certo bairro e observado o número de crianças de cada família, matriculadas na escola. Os dados foram 1, 1, 2, 0, 2, 0, 2, 3, 4, 1, 1, 2, 0, 0, e 2. Obtenha as estimativas correspondentes aos seguintes estimadores da média de crianças na escola nesse bairro,

$$\begin{aligned} \hat{\mu}_2 &= \frac{X_1 + X_2}{2}, \\ \hat{\mu}_3 &= \overline{X}_n. \end{aligned}$$

   Qual deles é o melhor estimador da média e por quê?

   Consideramos X o número de crianças de cada família matriculadas na escola, podemos assumir que $X \sim \text{Poisson}(\theta)$, $E(X) = \theta$ e $Var(X) = \theta$. Supondo que $X_1, \ldots, X_n$ é uma amostra aleatória de $X \sim \text{Poisson}(\theta)$.

   Para verificar qual é melhor é preciso avaliar as propriedades dos dois estimadores.

- Média

$$E(\hat{\mu}_2) = E\left(\frac{X_1 + X_2}{2}\right) = \frac{E(X_1) + E(X_2)}{2} = \frac{2E(X)}{2} = E(X) = \theta$$

$$E(\hat{\mu}_3) = E\left(\frac{X_1 + X_2 + \ldots X_n}{n}\right) = \frac{E(X_1) + E(X_2) + \ldots E(X_n)}{n} = \frac{nE(X)}{n} = E(X) = \theta$$

- Viés

$$B(\hat{\mu}_2) = E(\hat{\mu}_2) - \theta = \theta - \theta = 0$$
$$B(\hat{\mu}_3) = E(\hat{\mu}_3) - \theta = \theta - \theta = 0$$

  Os dois estimadores são não viesados. Temos que avaliar a variância deles para poder indicar qual é o melhor.

- Variância

$$Var(\hat{\mu}_2) = Var\left(\frac{X_1 + X_2}{2}\right) = \frac{Var(X_1) + Var(X_2)}{2^2} = \frac{2Var(X)}{4} = \frac{Var(X)}{2} = \frac{\theta}{2}$$

$$\begin{aligned} Var(\hat{\mu}_3) =& Var\left(\frac{X_1 + X_2 + \ldots + X_n}{n}\right) = \frac{Var(X_1) + Var(X_2) + \ldots + Var(X_n)}{n^2} \\ =& \frac{nVar(X)}{n^2} = \frac{Var(X)}{n} = \frac{\theta}{n} \end{aligned}$$

  Para $n > 2$ o estimador $\widehat{\mu}_3$ tem menor variância, logo, ele é melhor que o estimador $\widehat{\mu}_2$.

4. Para estimar a média θ desconhecida de uma população, foram propostos dois estimadores não viesados independentes, $\hat{\theta}_1$ e $\hat{\theta}_2$, de tal foram que $Var(\hat{\theta}_1) = Var(\hat{\theta}_2)/3$. Considere os seguintes estimadores ponderados de θ:
   (i) $T_1 = (\hat{\theta}_1 + \hat{\theta}_2)/2$;
   (ii) $T_2 = (4\hat{\theta}_1 + \hat{\theta}_2)/5$;
   (iii) $T_3 = \hat{\theta}_1$.

   (i) Quais estimadores são não viesados?
   Se os dois estimadores propostos para estimar θ são não viesados, quer dizer que $E(\hat{\theta}_1) = \theta$ e $E(\hat{\theta}_2) = \theta$, agora, basta verficiar se T_1, T_2 e T_3 são não viesados.

$$E(T_1) = E\left(\frac{\hat{\theta}_1 + \hat{\theta}_2}{2}\right) \frac{E(\hat{\theta}_1) + E(\hat{\theta}_2)}{2} = \frac{2\theta}{2} = \theta$$

$$E(T_2) = E\left(\frac{4\hat{\theta}_1 + \hat{\theta}_2}{5}\right) \frac{4E(\hat{\theta}_1) + E(\hat{\theta}_2)}{5} = \frac{5\theta}{5} = \theta$$

$$E(T_3) = E(\hat{\theta}_1) = \theta$$

   Como $E(T_i) = \theta$ para $i \in \{1, 2, 3\}$, os três estimadores são não viesados.

   (ii) Disponha esses estimadores em ordem crescente de eficiência (quanto menor a variância, mais eficiente é o estimador).
   Da questão temos: $Var(\hat{\theta}_1) = Var(\hat{\theta}_2)/3 \Longrightarrow 3Var(\hat{\theta}_1) = Var(\hat{\theta}_2)$

$$Var(T_1) = Var\left(\frac{\hat{\theta}_1 + \hat{\theta}_2}{2}\right) = \frac{Var(\hat{\theta}_1) + 3Var(\hat{\theta}_1)}{2^2} = \frac{4Var(\hat{\theta}_1)}{4} = Var(\hat{\theta}_1)$$

$$Var(T_2) = Var\left(\frac{4\hat{\theta}_1 + \hat{\theta}_2}{5}\right) = \frac{4^2Var(\hat{\theta}_1) + 3Var(\hat{\theta}_1)}{5^2} = \frac{19Var(\hat{\theta}_1)}{25}$$

$$Var(T_3) = Var(\hat{\theta}_1)$$

Assim, temos que

$$Var(T_2) < Var(T_1) = Var(T_3).$$

Portanto, o estimador mais eficiente é o T_2.

5. Seja X_1, X_2, X_3 uma amostra aleatória de X com distribuição exponencial, $X \sim \text{Exp}(\theta)$, em que $E(X) = \dfrac{1}{\theta}$. Considere os estimadores

$$\begin{aligned} \hat{\theta}_1 &= \overline{X}_n, \\ \hat{\theta}_2 &= X_2, \\ \hat{\theta}_3 &= \frac{X_1 + X_2}{2}. \end{aligned}$$

   (i) Mostre que o viés dos três estimadores é zero;

$$E(\hat{\theta}_1) = E(\overline{X}) = E(X) = \frac{1}{\theta}$$

$$B(\hat{\theta}_1) = E(\hat{\theta}_1) - \frac{1}{\theta} = \frac{1}{\theta} - \frac{1}{\theta} = 0$$

$$E(\hat{\theta}_2) = E(X_2) = E(X) = \frac{1}{\theta}$$

$$B(\hat{\theta}_2) = E(\hat{\theta}_2) - \frac{1}{\theta} = \frac{1}{\theta} - \frac{1}{\theta} = 0$$

$$E(\hat{\theta}_3) = E\left(\frac{X_1 + X_2}{2}\right) = E(X) = \frac{1}{\theta}$$

$$B(\hat{\theta}_3) = E(\hat{\theta}_3) - \frac{1}{\theta} = \frac{1}{\theta} - \frac{1}{\theta} = 0$$

   (ii) Qual dos estimadores tem menor variância? Lembrar que no caso exponencial $V(X) = \dfrac{1}{\theta^2}$.

$$Var(\hat{\theta}_1) = Var(\overline{X}) = \frac{Var(X)}{n} = \frac{1}{n\theta^2}$$

$$Var(\hat{\theta}_2) = Var(X_2) = Var(X) = \frac{1}{\theta^2}$$

$$Var(\hat{\theta}_3) = Var\left(\frac{X_1 + X_2}{2}\right) = \frac{Var(X_1) + Var(X_2)}{2^2} = \frac{2Var(X)}{4} = \frac{1}{2\theta^2}.$$

   Para $n > 2$ o estimador $\hat{\theta}_1$ é o que tem a menor variância.

6. Considere um experimento consistindo de n provas de Bernoulli, com probabilidade de sucesso θ. Seja X o número de sucessos e considere os estimadores

   (i) $\hat{\theta}_1 = X/n$;

   (ii) $\hat{\theta}_2 = \begin{cases} 1, & \text{se sucesso na primeira prova} \\ 0, & \text{se fracasso na primeira prova} \end{cases}.$

(i) Determine a esperança e variância de cada estimador. Por que $\hat{\theta}_2$ não é um "bom" estimador?

Vamos definir inicialmente $Y \sim \text{Bernoulli}(\theta)$ e $X = \sum_{i=1}^{n} Y_i \sim \text{Binomial}(n, \theta)$. Temos que $E(Y) = \theta$ e $Var(Y) = \theta(1-\theta)$, $E(X) = n\theta$ e $Var(X) = n\theta(1-\theta)$.

$$E(\hat{\theta}_1) = E(X/n) = E(X)/n = n\theta/n = \theta.$$

$$E(\hat{\theta}_2) = E(Y) = \theta.$$

$$Var(\hat{\theta}_1) = Var(X/n) = Var(X)/n^2 = n\theta(1-\theta)/n^2 = \theta(1-\theta)/n.$$

$$Var(\hat{\theta}_2) = Var(Y) = \theta(1-\theta).$$

Para $n > 1$ o estimador $\hat{\theta}_1$ será melhor, pois tem menor variância.

(ii) Verifique se $\hat{\theta}_1$ e $\hat{\theta}_2$ são consistentes;

Para que o estimador seja consistente

$$\lim_{n\to\infty} EQM(\widehat{\theta}) = 0$$

Assim, calcularemos primeiro o EQM, que é dado por $EQM(\widehat{\theta}) = Var(\widehat{\theta}) + [B(\widehat{\theta})]^2$

$$B(\hat{\theta}_1) = E(\hat{\theta}_1) - \theta = \theta - \theta = 0$$

$$B(\hat{\theta}_2) = E(\hat{\theta}_2) - \theta = \theta - \theta = 0$$

$$EQM(\hat{\theta}_1) = Var(\hat{\theta}_1) + [B(\hat{\theta}_1)]^2 = \theta(1-\theta)/n + 0^2$$

$$EQM(\hat{\theta}_2) = Var(\hat{\theta}_2) + [B(\hat{\theta}_2)]^2 = \theta(1-\theta) + 0^2$$

$$\lim_{n\to\infty} EQM(\widehat{\theta}_1) = \lim_{n\to\infty} \theta(1-\theta)/n = 0$$

$$\lim_{n\to\infty} EQM(\widehat{\theta}_2) = \lim_{n\to\infty} \theta(1-\theta) = \theta(1-\theta)$$

Assim, $\widehat{\theta}_1$ é consistente, já $\widehat{\theta}_1$ não é.

(iii) Considere um estimador para a variância da população $\sigma^2 = \theta(1-\theta)$, dado por $\hat{\theta}_1(1-\hat{\theta}_1)$. Em média este estimador acertaria a verdadeira variância da população?

$$\begin{aligned}
E(\hat{\sigma}^2) &= E[\hat{\theta}_1(1-\hat{\theta}_1)] \\
&= E[\hat{\theta}_1 - \hat{\theta}_1^2] \\
&= E(\hat{\theta}_1) - E(\hat{\theta}_1^2) \\
&= E(\hat{\theta}_1) - Var(\hat{\theta}_1) - [E(\hat{\theta}_1)]^2 \\
&= \theta - \frac{\theta(1-\theta)}{n} - \theta^2 \\
&= \theta(1-\theta) - \frac{\theta(1-\theta)}{n} \\
&= \theta(1-\theta)\left(1 - \frac{1}{n}\right) \\
&= \theta(1-\theta)\left(\frac{n-1}{n}\right)
\end{aligned}$$

O estimador $\hat{\theta}_1(1-\hat{\theta}_1)$ não acertaria a variância da população.

(iv) Sugira um estimador para a variância da população diferente do sugerido em (iii) e justifique.
Um estimador dado por

$$\hat{\theta}_1(1-\hat{\theta}_1)\frac{n}{n-1}$$

em média acertaria a variância da população e portanto seria não viesado.

7. Para cada um dos problemas listados abaixo estime as quantidades de interesse utilizando o método dos momentos.

$$S_x^2 = \frac{\sum_{i=1}^{n} X_i^2 - \frac{\left(\sum_{i=1}^{n} X_i\right)^2}{n}}{n}$$

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

(a) Considere que o tempo necessário para um medicamento fazer efeito, em minutos, seja modelado segundo uma distribuição uniforme contínua no intervalo $[a, b]$. Estime os limitantes a e b dessa distribuição a partir da amostra observada $\{9, 10, 14, 14, 12, 11, 13, 15, 13, 12, 16, 15, 13\}$. Temos $n = 13$, $\sum_{i=1}^{13} x_i = 167$ e $\sum_{i=1}^{13} x_i^2 = 2195$.

$$s_x^2 = \frac{2195 - \frac{(167)^2}{13}}{13} = 3.822$$

$$\overline{x}_n = \frac{167}{13} = 12.846.$$

Os estimadores são: $\hat{a} = \overline{X}_n - \sqrt{3S_x^2}$ e $\hat{b} = \overline{X}_n + \sqrt{3S_x^2}$.
As estimativas são:

$$\hat{a} = 12.846 - \sqrt{3 \times 3.822} = 9.460$$

$$\hat{b} = 12.846 + \sqrt{3 \times 3.822} = 16.232$$

(b) Assuma que o tempo de vida de lâmpadas de determinada marca tenha distribuição exponencial ($f_\theta(x) = \frac{1}{\theta}\exp(-x/\theta)$, se $x > 0$ e zero caso contrário). Estime o parâmetro θ considerando a amostra observada $\{4, 6, 13, 12, 8, 5, 7, 10, 11, 9\}$. Temos $n = 10$ e $\sum_{i=1}^{10} x_i = 85$.

O estimador é $\hat{\theta} = \overline{X}_n$ e a estimativa é:

$$\hat{\theta} = \overline{x}_n = \frac{85}{10} = 8.5$$

(c) Assuma que a altura das pessoas seja modelada segundo uma distribuição normal. Estime a média e a variância populacional da altura das pessoas quando se observa a seguinte amostra $\{1.86, 1.67, 1.37, 1.53, 1.97, 1.77, 1.48, 1.96, 1.65, 1.58, 1.78, 1.74, 1.53, 1.73, 1.36\}$. Temos $n = 15$, $\sum_{i=1}^{15} x_i = 24.98$ e $\sum_{i=1}^{13} x_i^2 = 42.11$.

Os estimadores são $\hat{\mu} = \overline{X}_n$ e $\hat{\sigma}^2 = S_x^2$, as estimativas são dadas a seguir:

$$\hat{\mu} = \overline{x}_n = \frac{24.98}{15} = 1.665$$

$$\hat{\sigma}^2 = s_x^2 = \frac{42.11 - \frac{(24.98)^2}{15}}{15} = 0.034$$

(d) Para os itens (b) e (c) quais seriam as estimativas de máxima verossimilhança?

8. Estamos interessados em estimar a proporção θ de mulheres que frequentam um estádio de futebol. Para isso foi realizado um experimento durante n jogos consecutivos observando-se a variável aleatória $X_1, \ldots, X_n$ de X, sendo X = "O número de homens que chegam ao estádio até o aparecimento da primeira mulher". Temos que $E(X) = \frac{1-\theta}{\theta}$.

   (a) Obtenha o estimador de momentos de θ.

   Estimador de Momentos é tal que $E(X) = \overline{X}_n$, como $E(X) = \dfrac{1-\theta}{\theta}$, temos

$$\frac{1-\hat{\theta}}{\hat{\theta}} = \overline{X}_n$$
$$\Longrightarrow \hat{\theta} = \frac{1}{1+\overline{X}_n}$$

   (b) Seja $\widehat{\theta} = \frac{1}{1+\overline{X}_n}$ o estimador de máxima verossimilhança para o parâmetro θ. Calcule as estimativas de θ com base nos estimadores de momentos e de máxima verossimilhança, considerando a amostra observada: $\{8, 20, 30, 15, 12, 10, 18, 25\}$. O que esse valor quer dizer?

A estimativa é dada abaixo:

$$\hat{\theta} = \frac{1}{1+\overline{x}_n} = \frac{1}{1+17.25} = 0.055$$

O valor estimado de $\widehat{\theta} = 0.055$ representa a proporção de mulheres que frequentam estágios de futebol.

9. Seja 2.53, 0.49, 1.12, 0.17, 2.71, 0.52, 1.42, 2.72, 2.45, 1.56, 0.40, 0.05, 2.44, 2.41, 2.84, uma amostra aleatória da variável aleatória X com distribuição uniforme em $[0, b]$, ou seja $X \sim U(0, b)$.

   Encontre a estimativa de b pelo método dos momentos.

   (a) Obtenha o estimador de momentos de b.
   Estimador de momentos de b

$$\overline{X}_n = \frac{\hat{a}+\hat{b}}{2} = \frac{0+\hat{b}}{2}$$
$$\Longrightarrow \hat{b} = 2\overline{X}_n$$

   (b) Obtenha a estimativa de momentos de b.

Temos $n = 15$ e $\sum_{i=1}^{15} x_i = 23.83$.

Estimativa de momentos de b

$$\overline{x}_n = \frac{23.83}{15} = 1.589$$
$$\hat{b} = 2\overline{x}_n = 2 \times 1.589 = 3.178$$

10. Seja $X_1, \ldots, X_n$ uma amostra aleatória da distribuição da variável aleatória X cuja função densidade é dada por:

$$f_\theta(x) = \begin{cases} \dfrac{\beta^\alpha x^{\alpha-1} e^{-\beta x}}{\Gamma(\alpha)}, & x \geq 0, \\ 0, & \text{caso contrário,} \end{cases}$$

em que $\alpha > 0$, $\beta > 0$ são os parâmetros da distribuição. A esperança e variância são dadas por $E(X) = \dfrac{\alpha}{\beta}$ e $V(X) = \dfrac{\alpha}{\beta^2}$.

(a) Obtenha os estimadores de momentos para α e β.

$$\overline{X}_n = \frac{\hat{\alpha}}{\hat{\beta}} \Longrightarrow \hat{\alpha} = \overline{X}_n \hat{\beta} \tag{1}$$

$$S_x^2 = \frac{\hat{\alpha}}{\hat{\beta}^2} \tag{2}$$

Substituindo $\hat{\alpha}$ da equaçao 1 na equação 2, temos:

$$S_x^2 = \frac{\overline{X}_n \hat{\beta}}{\hat{\beta}^2} = \frac{\overline{X}_n}{\hat{\beta}}$$

$$\Longrightarrow \hat{\beta} = \frac{\overline{X}_n}{S_x^2}$$

Agora substituindo $\hat{\beta}$ na equação 1, temos:

$$\hat{\alpha} = \overline{X}_n \hat{\beta} = \overline{X}_n \frac{\overline{X}_n}{S_x^2} = \frac{\overline{X}_n^2}{S_x^2}$$

(b) Para a amostra observada $\{1.4, 2.5, 1.6, 0.8, 3.4, 2.7\}$ calcule as estimativas de momentos para α e β.

As estimativas são calculadas a seguir:

$$\overline{x}_n = \frac{\sum_{i=1}^{n} x_i}{n} = \frac{12.4}{6} = 2.067$$

$$s_x^2 = \frac{\sum_{i=1}^{n} x_i^2 - \dfrac{\left(\sum_{i=1}^{n} x_i\right)^2}{n}}{n} = \frac{30.26 - \dfrac{(12.4)^2}{6}}{6} = 0.772$$

$$\hat{\alpha} = \frac{\overline{x}^2}{s_x^2} = \frac{2.067^2}{0.772} = 5.534$$

$$\hat{\beta} = \frac{\overline{x}}{s_x^2} = \frac{2.067}{0.772} = 2.677$$

11. No monitoramento da qualidade da produção de uma indústria, amostras de nove unidades de cilindros metálicos são selecionadas periodicamente, e uma das variáveis avaliadas é o diâmetro dos cilindros. Considere que, quando operando "sob controle" (sem ocorrência de problemas identificáveis), os diâmetros dos cilindros produzidos têm distribuição normal, com média $\mu = 10cm$ e desvio padrão $\sigma = 0.3cm$.

$$E(\overline{X}_n) = E(X) \quad \text{e} \quad V(\overline{X}_n) = \frac{V(X)}{n}$$

(a) Qual a distribuição de $\overline{X}_n$, da média amostral de 9 diâmetros de cilindros metálicos?

$$\overline{X}_n \sim N\left(10; \frac{0.3^2}{9}\right)$$
$$\overline{X}_n \sim N(10; 0.01)$$

(b) Qual a probabilidade da média de uma amostra exceder 10.12cm?

$$\begin{aligned} P(\overline{X}_n > 10.12) &= P\left(\sqrt{n}\frac{\overline{X}_n - \mu}{\sqrt{\sigma^2}} > \sqrt{9}\frac{(10.12 - 10)}{\sqrt{0.3^2}}\right) \\ &= P(Z > 1.2) \\ &= 0.115 \end{aligned}$$

(c) Qual a probabilidade da média de uma amostra ser inferior a 9.7cm ou superior a 10.3cm? Você diria que um valor fora desse intervalo seria típico de um processo operando sob controle??

$A = \{$a média de uma amostra ser inferior a 9.7cm ou superior a 10.3cm$\}$

$$\begin{aligned} P(A) &= P(\overline{X}_n < 9.7) + P(\overline{X}_n > 10.3) \\ &= P\left(Z < \sqrt{9}\frac{(9.7 - 10)}{\sqrt{0.3^2}}\right) + P\left(Z > \sqrt{9}\frac{(10.3 - 10)}{\sqrt{0.3^2}}\right) \\ &= P(Z < -3) + P(Z > 3) \\ &= 0.003 \end{aligned}$$

(d) Suponha que a média do processo, devido a um problema de manutenção das máquinas, seja alterada para 9.8cm. Recalcule a probabilidade solicitada no ítem anterior, agora considerando essa nova configuração do processo. O que muda?

$B = \{$a média de uma amostra ser inferior a 9.7cm ou superior a 10.3cm$\}$

$$\begin{aligned} P(B) &= P(\overline{X}_n < 9.7) + P(\overline{X}_n > 10.3) \\ &= P\left(Z < \sqrt{9}\frac{(9.7 - 9.8)}{\sqrt{0.3^2}}\right) + P\left(Z > \sqrt{9}\frac{(10.3 - 9.8)}{\sqrt{0.3^2}}\right) \\ &= P(Z < -1) + P(Z > 5) \\ &= 0.159 \end{aligned}$$

12. A duração do “tonner” de uma máquina de fotocópias pode ser modelada por uma distribuição normal com média 15 e desvio padrão 2 (em milhares de cópias). Para uma amostra de 12 fotocopiadoras a duração do “tonner” será observada, qual a probabilidade da média amostral do “tonner” durar:

(a) Menos de 16 mil cópias?

$A = \{$média amostral do "tonner"durar menos de 16 mil cópias$\}$

$$\begin{aligned} P(A) &= P(\overline{X}_n < 16) \\ &= P\left(Z < \sqrt{12}\frac{(16-15)}{\sqrt{2^2}}\right) \\ &= P(Z < 1.732) \\ &= 0.958 \end{aligned}$$

(b) Mais de 13 mil cópias?

$B = \{$média amostral do "tonner"durar mais de 13 mil cópias$\}$

$$\begin{aligned} P(A) &= P(\overline{X}_n > 13) \\ &= P\left(Z < \sqrt{12}\frac{(13-15)}{\sqrt{2^2}}\right) \\ &= P(Z > -3.464) \\ &= 1 \end{aligned}$$

(c) Entre 12 e 14 mil cópias?

$C = \{$média amostral do "tonner"durar entre 12 e 14 mil cópias$\}$

$$\begin{aligned} P(C) &= P(12 < \overline{X}_n < 14) \\ &= P\left(\sqrt{12}\frac{(12-15)}{\sqrt{2^2}} < Z < \sqrt{12}\frac{(14-15)}{\sqrt{2^2}}\right) \\ &= P(-5.196 < Z < -1.732) \\ &= P(Z < -1.732) - P(Z < -5.196) \\ &= 0.042 \end{aligned}$$

13. Um fabricante afirma que sua vacina contra gripe imuniza em 80% dos casos. Uma amostra de 25 indivíduos que tomaram a vacina foi sorteada e testes foram feitos para verificar a imunização ou não desses indivíduos. Se o fabricante estiver correto, qual é a probabilidade de a proporção de imunizados na amostra ser inferior a 0.75? E superior a 0.85?

Pelo teorema central do limite, $\hat{p}$ terá distribuição aproximadamente normal, com média p e variância $\frac{p(1-p)}{n}$, ou seja,

$$\hat{p} \sim N\left(p, \frac{p(1-p)}{n}\right).$$

Assim, temos que

$$Z = \frac{\hat{p} - p}{\sqrt{\frac{p(1-p)}{n}}} \sim N(0,1).$$

$A = \{$proporção de imunizados na amostra ser inferior a 0.75$\}$

$$\begin{aligned} P(A) &= P(\hat{p} < 0.75) \\ &= P\left(\frac{\hat{p} - p}{\sqrt{\frac{p(1-p)}{n}}} < \frac{0.75 - 0.8}{\sqrt{\frac{0.8(1-0.8)}{25}}}\right) \\ &= P(Z < -0.625) \\ &= 0.266 \end{aligned}$$

$A = \{\text{proporção de imunizados na amostra ser superior a } 0.85\}$

$$\begin{aligned}
P(A) &= P(\hat{p} > 0.85) \\
&= P\left(Z > \frac{0.85 - 0.8}{\sqrt{\frac{0.8(1-0.8)}{25}}}\right) \\
&= P(Z > 0.625) \\
&= 0.266
\end{aligned}$$