Disciplina: MAE0212: Introdução a Probabilidade e Estatística II
Professora: Silvia Ferrari
Nome: **NUSP**: **Nota**:

Use os espaços em branco, inclusive o verso das páginas, para responder às questões. Mostre os desenvolvimentos, não apenas as respostas.

Prova 1

1. De estudos anteriores sabe-se que as alturas das plantas de um determinado híbrido de milho são normalmente distribuídos com desvio padrão de $\sigma =$0,80 m.

   a) (0.5 ponto) Se a altura média (μ) fosse conhecida e $\mu =$2,00 m, qual seria a probabilidade da média de uma amostra de 81 plantas desse híbrido ser superior a $2,20\ m$?

   b) (1 ponto) Selecionada uma amostra de 81 plantas desse híbrido, observou-se média amostral de 2,10 m. Construa um intervalo de confiança de 92% de confiança para a média populacional (μ) desconhecida.

   c) (1 ponto) Que tamanho deveria ter a amostra para que o intervalo [2,10-0,10; 2,10+0,10] tivesse 99% de confiança?

   Solução:

   Considere a variável aleatória X que representa a altura das plantas de determinado híbrido de milho e observer que $X \sim N(\mu, 0.80^2)$. Além disso, sabemos que, neste caso, $\bar{X} \sim N(\mu, 0.80^2/n)$.

   a) Suponha que $\mu = 2$ e, assim, $\bar{X} \sim N(2, 0.80^2/n)$. Estamos interessados na probabilidade da média de uma amostra de 81 plantas desse híbrido ser superior a 2.20, isto é,

   $$\mathbb{P}(\bar{X} > 2.20) = 1 - \mathbb{P}(\bar{X} \le 2.20) = 1 - \mathbb{P}\left(\frac{\bar{X} - 2}{0.8/\sqrt{81}} \le \frac{2.20 - 2}{0.8/\sqrt{81}}\right)$$
   $$= 1 - \mathbb{P}\left(Z \le 2.25\right) = 1 - 0.987 = 0.013.$$

   b) Considere uma amostra aleatória simples de tamanho 81 cujo valor observado para a média amostral ($\bar{x}$) foi 2.10. Assim, O intervalo de confiança observado para μ com coeficiente de confiança γ é dado por

   $$\text{IC}(\mu; \gamma) = \left[\bar{x} - z_\gamma \frac{\sigma}{\sqrt{n}}, \bar{x} + z_\gamma \frac{\sigma}{\sqrt{n}}\right],$$

   em que z_γ é tal que $\mathbb{P}(-z_\gamma \le Z \le z_\gamma) = \gamma$, sendo $Z \sim N(0,1)$. Para $\gamma = 0.92\%$, temos que $z_\gamma = 1.75$, assim temos que o intervalo de confiança para μ com $\gamma = 92\%$ é dado por

   $$\left[2.10 - 1.75\frac{0.8}{\sqrt{81}}, 2.10 + 1.75\frac{0.8}{\sqrt{81}}\right] = [1.94, 2.25].$$

c) Queremos obter o valor de n para que o intervalo $[2.10 - 0.10, 2.10 + 0.10]$ tenha 99% de confiança. Note que o intervalo de confiança para μ é dado por

$$\text{IC}(\mu; \gamma) = \left[\bar{X} - z_\gamma \frac{\sigma}{\sqrt{n}}, \bar{X} + z_\gamma \frac{\sigma}{\sqrt{n}}\right],$$

em que z_γ é tal que $\mathbb{P}(-z_\gamma \leq Z \leq z_\gamma) = \gamma$, sendo $Z \sim N(0,1)$. Queremos $\gamma = 99\%$, assim, $z_\gamma = 2.57$. Dessa forma, queremos obter o valor de n tal que

$$z_\gamma \frac{\sigma}{\sqrt{n}} = 0.10 \quad \Rightarrow 2.57 \frac{0.8}{\sqrt{n}} = 0.10 \quad \Rightarrow n = 422.7.$$

Assim, o tamanho amostral necessário para que o intervalo tenha 99% de confiança é $n = 423$.

2. (2 pontos) Seja X uma variável aleatória com função densidade de probabilidade

$$f(x;\mu) = \frac{4}{\mu^2} x e^{-2x/\mu}, \quad x > 0.$$

A média de X é $\mathbb{E}(X) = \mu$ e sua variância é $Var(X) = \mu^2/2$. Seja $(X_1, \ldots, X_n)$ uma A.A.S. de X.

a) (1.5 ponto) Mostre que $\hat{\mu} = \overline{X}$ é o estimador de máxima verossimilhança de μ. Este estimador é não viciado? Por que? Usando o Teorema Limite Central, qual é a distribuição aproximada de $\hat{\mu}$ quando n é grande?

b) (1 ponto) Contrua um intervalo de confiança aproximada para μ supondo que n seja grande. Obtenha o intervalo de confiança se $n = 100$, $\overline{x} = 15$ e $\gamma = 90\%$.

Solução:

Seja X uma variável aleatória com função densidade de probabilidade $f(x;\mu)$.

a) A função de verossimilhança de μ, dada a amostra observada $(x_1, \ldots, x_n)$, é dada por

$$\begin{aligned}
L(\mu) &= \mathbb{P}_\lambda(X_1 = x_1, \ldots, X_n = x_n) \\
&= \mathbb{P}_\mu(X_1 = x_1) \cdots \mathbb{P}_\mu(X_n = x_n) \quad \text{(pois } X_1, \ldots, X_n \text{ são independentes)} \\
&= \prod_{i=1}^{n} \mathbb{P}_\mu(X = x_i) \quad \text{(pois } X_1, \ldots, X_n \text{ têm a mesma distribuição de X)} \\
&= \prod_{i=1}^{n} \frac{4}{\mu^2} x_i e^{-2x_i/\mu} \\
&= \frac{4^n}{\mu^{2n}} e^{-2\sum_{i=1}^{n} x_i/\mu}.
\end{aligned}$$

Assim, o logaritmo da função de verossimilhança é dado por

$$\ell(\mu) = n\log(4) - 2n\log(\mu) - \frac{2}{\mu}\sum_{i=1}^{n} x_i.$$

Note que o objetivo é encontrar o valor $\widehat{\mu}$ que maximiza $\ell(\mu)$. Dessa forma, temos que

$$\frac{\mathrm{d}\ell(\mu)}{\mathrm{d}\mu} = -\frac{2n}{\mu} + \frac{2}{\mu^2}\sum_{i=1}^{n} x_i.$$

A estimativa de máxima verossimilhança de μ, $\widehat{\mu}$, é tal que $\mathrm{d}\ell(\mu)/\mathrm{d}\mu$ avaliada em $\mu = \widehat{\mu}$ é igual a zero. Assim, temos que

$$-\frac{2n}{\mu} + \frac{2}{\mu^2}\sum_{i=1}^{n} x_i = 0 \quad \Rightarrow \quad \frac{1}{\mu}\sum_{i=1}^{n} x_i = n \quad \Rightarrow \quad \widehat{\mu} = \bar{x}.$$

Além disso,

$$\frac{\mathrm{d}^2\ell(\mu)}{\mathrm{d}\mu^2} = \frac{2n}{\mu^2} - \frac{4}{\mu^3}n\bar{x}.$$

Avaliando $\frac{\mathrm{d}^2\ell(\mu)}{\mathrm{d}\mu^2}$ em $\mu = \widehat{\mu} = \bar{x}$, e observando que $x_i > 0, \forall i = 1, \ldots, n$, temos que

$$\frac{\mathrm{d}^2\ell(\mu)}{\mathrm{d}\mu^2}\Big|_{\mu=\bar{x}} = \frac{2n}{\bar{x}^2} - \frac{4}{\bar{x}^3}n\bar{x} = \frac{2n}{\bar{x}^2} - \frac{4n}{\bar{x}^2} = -\frac{2n}{\bar{x}} < 0.$$

Assim, $\bar{x}$ maximiza a função $\ell(\mu)$ e, portanto, o estimador de máxima verossimilhança de μ é $\bar{X}$. Adicionalmente, note que

$$\mathbb{E}_\mu(\bar{X}) = \mathbb{E}_\mu(X) = \mu.$$

Logo, o estimador de máxima verossimilhança $\bar{X}$ é não viciado para μ.

Adicionalmente, Pelo Teorema Limite Central, quando n é suficientemente grande, temos que $\bar{X} \sim N\left(\mu, \frac{\mu^2/2}{n}\right)$, de modo que

$$\frac{\bar{X} - \mu}{\sqrt{\mu^2/2n}} \sim N(0, 1).$$

b) Baseando-se na distribuição amostral de $\bar{X}$ para grandes amostras, prosseguimos com a construção do intervalo de confiança para μ com coeficiente de confiança γ.

$$\mathbb{P}\left(-z_\gamma \leq \frac{\bar{X} - \mu}{\sqrt{\mu^2/2n}} \leq z_\gamma\right) = \gamma,$$

em que z_γ é tal que $\mathbb{P}(-z_\gamma \leq Z \leq z_\gamma) = \gamma$, sendo $Z \sim N(0, 1)$. Assim,

$$\begin{aligned} &\mathbb{P}\left(-z_\gamma \leq \frac{\bar{X} - \mu}{\sqrt{\mu^2/2n}} \leq z_\gamma\right) = \gamma \\ \Rightarrow\ &\mathbb{P}\left(-z_\gamma\sqrt{\frac{\mu^2}{2n}} \leq \bar{X} - \mu \leq z_\gamma\sqrt{\frac{\mu^2}{2n}}\right) = \gamma \\ \Rightarrow\ &\mathbb{P}\left(\bar{X} - z_\gamma\sqrt{\frac{\mu^2}{2n}} \leq \mu \leq \bar{X} + z_\gamma\sqrt{\frac{\mu^2}{2n}}\right) = \gamma. \end{aligned}$$

Substituindo μ por $\widehat{\mu} = \bar{X}$ nos extremos do intervalo, temos

$$\mathrm{IC}(\lambda; \gamma) = \left[\bar{X} - z_\gamma\sqrt{\frac{\bar{X}^2}{2n}}, \bar{X} + z_\gamma\sqrt{\frac{\bar{X}^2}{2n}}\right].$$

Considerando que de uma amostra de tamanho 100 foi observado $\bar{x} = 15$, então a estimativa de máxima verossimilhança de μ é 15. Ainda, o intervalo de confiança para μ considerando essa amostra observada e $\gamma = 90\%$ é dado por

$$\left[15 - 1.64 \cdot \sqrt{1.125}, 15 + 1.64 \cdot \sqrt{1.125}\right] = [13.15, 16.84].$$

3. Um biólogo pretende estimar a proporção p de camarões marinhos que apresentam comprimento superior a 6.0 cm. O biólogo propõe observar uma amostra aleatória de camarões e calcular a proporção de camarões com comprimento superior a 6,0 cm.

   a) (1.0 ponto) Qual deve ser o tamanho da amostra, para que o erro da estimativa seja no máximo 0,06, com probabilidade 0,92?

   b) (1.5 ponto) De 120 camarões observados, 36 apresentam comprimento superior a 6.0 cm. Dê uma estimativa pontual para p e, com base nela, construa um intervalo de 95% de confiança para p.

   Solução:

   Considere a variável aleatória X que assume o valor 1 se um camarão marinho selecionada ao acaso apresenta compprimento superior a 6.0 cm e 0 caso contrário. Note que $X \sim Bernoulli(p)$ e que, pelo Teorema Limite Central, temos que $\bar{X} \sim N(p, p(1-p)/n)$.

   a) Queremos obter o tamanho da amostra para que o erro da estimativa seja no máximo 0.06 com probabilidade 0.92. Isto é,

   $$\mathbb{P}(\mid \widehat{p} - p \mid \leq 0.06) = 0.92.$$

   Se n for grande o suficiente podemos aproximar a distribuição de $\widehat{p}$ utilizando o Teorema Limite Central, de modo que

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

   Então,

   $$\begin{aligned}
   \mathbb{P}(\mid \widehat{p} - p \mid \leq 0.06) &= 0.92 \\
   \Rightarrow \mathbb{P}\left(\frac{-0.06}{\sqrt{\dfrac{p(1-p)}{n}}} \leq \frac{\widehat{p} - p}{\sqrt{\dfrac{p(1-p)}{n}}} \leq \frac{0.06}{\sqrt{\dfrac{p(1-p)}{n}}}\right) &= 0.92 \\
   \Rightarrow \mathbb{P}\left(\frac{-0.06}{\sqrt{\dfrac{p(1-p)}{n}}} \leq Z \leq \frac{0.06}{\sqrt{\dfrac{p(1-p)}{n}}}\right) &= 0.92.
   \end{aligned}$$

   Logo,

   $$\frac{0.06}{\sqrt{\dfrac{p(1-p)}{n}}} = 1.75,$$

e

$$n = \frac{p(1-p)1.75^2}{0.06^2} = p(1-p) \cdot 981.78.$$

Como p é desconhecido, substituimos $p(1-p)$ por seu valor máximo, 0.25. Note que o gráfico de $g(p) = p(1-p)$ é uma parábola com concavidade para baixo e o máximo ocorre em $p = 0.5$. Assim, tomamos

$$n = \frac{0.25 \cdot 1.75^2}{0.06^2} = 212.67.$$

Assim, o tamanho amostral para que o erro da estimativa seja no máximo 0.06 com probabilidade 0.92 é 213.

b) Uma estimativa pontual para p é $\widehat{p} = 36/120 = 0.3$. Um intervalo de confiança para p com 93% de confiança é dado por

$$\left[0.3 - 1.81\sqrt{\frac{0.3 \cdot 0.7}{120}}, 0.2 + 1.81\sqrt{\frac{0.3 \cdot 0.7}{120}}\right] = [0.22, 0.37].$$

4. Uma balança para encher pacotes de sementes está programada para produzir pacotes com pesos normalmente distribuídos com média 20 kg e desvio padrão 0.20 kg. Periodicamente é feita uma inspeção para verificar se o peso médio está sob controle. Deseja-se testar se a balança se desregulou e está produzindo um peso médio inferior a 20 kg. Para este fim, foi selecionada uma amostra de oito pacotes de sementes, cujos resultados foram: 20,3, 19,8, 20,3, 19,7, 19,8, 19,8, 19,8 e 19,7.

   a) (1.0 ponto) Formule o problema com um problema de teste de hipóteses enunciando as hipóteses nula e alternativa.

   b) (1.5 ponto) Contrua a região crítica adotando $\alpha = 4\%$. Qual é a conclusão do teste com base na amostra observada?

   Solução:

   Considere a variável aleatória X que representa o peso de um pacote de sementes selecionado ao acaso. Pelo enunciado, temos que $X \sim N(\mu, 0.2^2)$. Consideremos a construção do teste seguindo os 5 passos.

   a) *Passo 1*: Note que o interese é verificar se o peso médio é inferior a 20 kg. Dessa forma, consideremos as seguintes hipóteses:

   $$\mathcal{H}_0 : \mu = 20 \quad vs \quad \mathcal{H}_1 : \mu < 20.$$

   b) *Passo 2*: Inicialmente, vamos assumir que, quando a balança está desregulada, apenas a média sai de controle, isto é, $\sigma^2 = 0.2^2$. Dessa forma, para todo μ, a média ($\bar{X}$) de 8 pacotes terá distribuição $N(\mu, 0.2^2/8)$. Isto é, o desvio padrão de $\bar{X}$ é $\sigma = 0.2/\sqrt{8}$. Vamos utilizar a distribuição de $\bar{X}$ para construir o teste.

   Passo 3: Fixe $\alpha = 4\%$ e, sob a hipótese nula, $\bar{X} \sim N(20, 0.02^2)$. Queremos encontrar uma região crítica da forma $RC = \{\bar{x} \leq c\}$, em que c é tal que $\mathbb{P}(erroI) = 0.04$. Dessa forma, temos que

   $$\mathbb{P}(\bar{X} \leq c \mid \mu = 20) = 0.04 \Rightarrow \mathbb{P}\left(Z \leq \frac{c-20}{0.2/\sqrt{8}}\right) = 0.04 \Rightarrow \frac{c-20}{0.2/\sqrt{8}} = -1.75$$

   $$\Rightarrow c - 20 = -1.750.2/\sqrt{8} \Rightarrow c = 19.88.$$

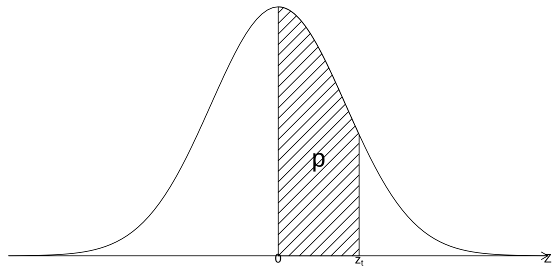
   Assim, a região crítica é

   $$RC = \{\bar{x} \leq 19.88\}.$$

   Passo 4: A média amostral observada na amostra foi $\bar{x} = 19.9$.

Passo 5: Como o valor observado de $\bar{x}$ não pertence à região crítica, não temos evidência amostral suficiente para rejeitar a hipótese nula, considerando $\alpha = 4\%$. Isto é, não temos evidência para afimar que o peso médio esteja fora controle.

Distribuição Normal



	0	1	2	3	4	5	6	7	8	9
0,0	0.00000	0.00399	0.00798	0.01197	0.01595	0.01994	0.02392	0.02790	0.03188	0.03586
0,1	0.03983	0.04380	0.04776	0.05172	0.05567	0.05962	0.06356	0.06749	0.07142	0.07535
0,2	0.07926	0.08317	0.08706	0.09095	0.09483	0.09871	0.10257	0.10642	0.11026	0.11409
0,3	0.11791	0.12172	0.12552	0.12930	0.13307	0.13683	0.14058	0.14431	0.14803	0.15173
0,4	0.15542	0.15910	0.16276	0.16640	0.17003	0.17364	0.17724	0.18082	0.18439	0.18793
0,5	0.19146	0.19497	0.19847	0.20194	0.20540	0.20884	0.21226	0.21566	0.21904	0.22240
0,6	0.22575	0.22907	0.23237	0.23565	0.23891	0.24215	0.24537	0.24857	0.25175	0.25490
0,7	0.25804	0.26115	0.26424	0.26730	0.27035	0.27337	0.27637	0.27935	0.28230	0.28524
0,8	0.28814	0.29103	0.29389	0.29673	0.29955	0.30234	0.30511	0.30785	0.31057	0.31327
0,9	0.31594	0.31859	0.32121	0.32381	0.32639	0.32894	0.33147	0.33398	0.33646	0.33891
1,0	0.34134	0.34375	0.34614	0.34849	0.35083	0.35314	0.35543	0.35769	0.35993	0.36214
1,1	0.36433	0.36650	0.36864	0.37076	0.37286	0.37493	0.37698	0.37900	0.38100	0.38298
1,2	0.38493	0.38686	0.38877	0.39065	0.39251	0.39435	0.39617	0.39796	0.39973	0.40147
1,3	0.40320	0.40490	0.40658	0.40824	0.40988	0.41149	0.41309	0.41466	0.41621	0.41774
1,4	0.41924	0.42073	0.42220	0.42364	0.42507	0.42647	0.42785	0.42922	0.43056	0.43189
1,5	0.43319	0.43448	0.43574	0.43699	0.43822	0.43943	0.44062	0.44179	0.44295	0.44408
1,6	0.44520	0.44630	0.44738	0.44845	0.44950	0.45053	0.45154	0.45254	0.45352	0.45449
1,7	0.45543	0.45637	0.45728	0.45818	0.45907	0.45994	0.46080	0.46164	0.46246	0.46327
1,8	0.46407	0.46485	0.46562	0.46638	0.46712	0.46784	0.46856	0.46926	0.46995	0.47062
1,9	0.47128	0.47193	0.47257	0.47320	0.47381	0.47441	0.47500	0.47558	0.47615	0.47670
2,0	0.47725	0.47778	0.47831	0.47882	0.47932	0.47982	0.48030	0.48077	0.48124	0.48169
2,1	0.48214	0.48257	0.48300	0.48341	0.48382	0.48422	0.48461	0.48500	0.48537	0.48574
2,2	0.48610	0.48645	0.48679	0.48713	0.48745	0.48778	0.48809	0.48840	0.48870	0.48899
2,3	0.48928	0.48956	0.48983	0.49010	0.49036	0.49061	0.49086	0.49111	0.49134	0.49158
2,4	0.49180	0.49202	0.49224	0.49245	0.49266	0.49286	0.49305	0.49324	0.49343	0.49361
2,5	0.49379	0.49396	0.49413	0.49430	0.49446	0.49461	0.49477	0.49492	0.49506	0.49520
2,6	0.49534	0.49547	0.49560	0.49573	0.49585	0.49598	0.49609	0.49621	0.49632	0.49643
2,7	0.49653	0.49664	0.49674	0.49683	0.49693	0.49702	0.49711	0.49720	0.49728	0.49736
2,8	0.49744	0.49752	0.49760	0.49767	0.49774	0.49781	0.49788	0.49795	0.49801	0.49807
2,9	0.49813	0.49819	0.49825	0.49831	0.49836	0.49841	0.49846	0.49851	0.49856	0.49861
3,0	0.49865	0.49869	0.49874	0.49878	0.49882	0.49886	0.49889	0.49893	0.49896	0.49900
3,1	0.49903	0.49906	0.49910	0.49913	0.49916	0.49918	0.49921	0.49924	0.49926	0.49929
3,2	0.49931	0.49934	0.49936	0.49938	0.49940	0.49942	0.49944	0.49946	0.49948	0.49950
3,3	0.49952	0.49953	0.49955	0.49957	0.49958	0.49960	0.49961	0.49962	0.49964	0.49965
3,4	0.49966	0.49968	0.49969	0.49970	0.49971	0.49972	0.49973	0.49974	0.49975	0.49976
3,5	0.49977	0.49978	0.49978	0.49979	0.49980	0.49981	0.49981	0.49982	0.49983	0.49983
3,6	0.49984	0.49985	0.49985	0.49986	0.49986	0.49987	0.49987	0.49988	0.49988	0.49989
3,7	0.49989	0.49990	0.49990	0.49990	0.49991	0.49991	0.49992	0.49992	0.49992	0.49992
3,8	0.49993	0.49993	0.49993	0.49994	0.49994	0.49994	0.49994	0.49995	0.49995	0.49995
3,9	0.49995	0.49995	0.49996	0.49996	0.49996	0.49996	0.49996	0.49996	0.49997	0.49997

Tabela 1: Probabilidades $p = P[0 \leq Z \leq Z_t]$ da Distribuição Normal padrão com valores de Z_t dados nas margens da tabela