CE-009: Introdução à Estatística - Turma: A, Prova Final (05/07/2017)

GRR: ______________ **Nome:** ____ **Turma:** ________

1. Um lote é formado por 10 animais sadios, quatro com problams menores e dois com problemas graves. Todos os animais são numerados e é feita a escolha de um animal ao acaso.

   (a) Tal situação pode ser considerada um *experimento aleatório*? Justifique.

   (b) Defina o espaço amostral.

   Encontre a probabilidade de que o animal selecionado:

   (c) não tenha problemas,

   (d) não tenha problemas graves,

   (e) seja sadio ou tenha problemas graves.

   Se do lote de animais descrito anteriormente forem selecionados **dois** animais

   (f) qual será o espaço amostral?

   E calcule a probabilidade que:

   (g) ambos sejam sadios,

   (h) ao menos um seja sadio,

   (i) no máximo um seja sadio,

   (j) exatamente um seja sadio,

   (k) nenhum deles seja sadio

   Solução:

   (a) Sim, pois a escolha do animal é feita ao acaso, e portanto não há como saber qual será o estado do animal selecionado.

   (b)

   $$\text{Notação: } S: \text{sadio}, \; M: \text{com problemas menores}, \; G: \text{problemas graves}$$
   $$\Omega = \{S, M, G\}$$

   (c) $P[S] = \frac{10}{16} = 0,625$

   (d) $P[\overline{G}] = 1 - P[G] = 1 - \frac{2}{16} = \frac{14}{16} = 0,875$

   (e) $P[\overline{S \cup G}] = P[S] + P[G] - P[S \cap G] \overset{M.E.}{=} \frac{10}{16} + \frac{2}{16} - 0 = \frac{12}{16} = 0,75$

   (f) $\Omega = \{(SS), (SM), (SG), (MM), (MG), (GG)\}$

   (g) $P[SS] = \frac{10}{16} \cdot \frac{9}{15} = \frac{90}{240} = 3/8 = 0,375$

   (h) $P[SS] + P[SM] + P[SG] = \frac{10}{16} \cdot \frac{9}{15} + 2 \cdot \frac{10}{16} \cdot \frac{4}{15} + 2 \cdot \frac{10}{16} \cdot \frac{2}{15} = \frac{210}{240} = 7/8 = 0,875$

(i) $P[\overline{SS}] = 1 - P[SS] = 1 - \frac{10}{16} \cdot \frac{9}{15} = \frac{150}{240} = 5/8 = 0,625$

(j) $P[SM] + P[SG] = 2 \cdot \frac{10}{16} \cdot \frac{4}{15} + 2 \cdot \frac{10}{16} \cdot \frac{2}{15} = \frac{120}{240} = 1/2 = 0,5$

(k) $P[\overline{SS}] = \frac{6}{16} \cdot \frac{5}{15} = \frac{30}{240} = 1/8 = 0,125$

---

2. Um contador eletrônico de bactérias registra, em média, 2,4 bactérias por cm^3 de um líquido. Admita que a váriável tem distribuição adequada dentre as vistas no curso e obtenha as probabilidades:

   (a) de que se encontre ao menos três bactérias em uma amostra de 1 cm^3,

   (b) de que se encontre ao menos três bactérias em uma amostra de 2 cm^3,

   Solução:

$$\begin{aligned} X &: \text{número de bactérias por volume de líquido} \\ x &\in \{0, 1, 2, \ldots\} \\ X &\sim \mathrm{P}(\lambda) \\ P[X = x] &= \frac{e^{-\lambda}\lambda^x}{x!} \end{aligned}$$

   (a) $\lambda = 2,4 \longrightarrow P[X \geq 3] = 1 - \{P[X = 0] + P[X = 1] + P[X = 2]\} = 1 - \{\frac{e^{-2,4}2,4^0}{0!} + \frac{e^{-2,4}2,4^1}{1!} + \frac{e^{-2,4}2,4^2}{2!}\} = 1 - e^{-2,4}\{1 + 2,4 + \frac{2,4^2}{2}\} = 1 - 0,57 = 0,43$

   (b) $\lambda = 4,8 \longrightarrow P[X \geq 3] = 1 - \{P[X = 0] + P[X = 1] + P[X = 2]\} = 1 - \{\frac{e^{-4,8}4,8^0}{0!} + \frac{e^{-4,8}4,8^1}{1!} + \frac{e^{-4,8}4,8^2}{2!}\} = 1 - e^{-4,8}\{1 + 4,8 + \frac{4,8^2}{2}\} = 1 - 0,14 = 0,86$

---

3. A observação dos pesos, X, de um grande número de espigas de milho mostrou que essa variável é normalmente distribuída com média $\mu = 90g$ e desvio padrão $\sigma = 7g$. Em um programa de melhoramento, entre outras características, uma linhagem, para continuar sendo avaliada no programa, deve satisfazer a condição $78g < X < 104g$. Nestas condições, em um programa de melhoramento com 200 linhagens pergunta-se.

   (a) Qual a probabilidade de uma linhagem ser aceita?

   (b) Qual o número de linhagens consideradas que deverá então serem descartadas do programa?

   (c) Qual o peso acima do qual espera-se encontrar 15% das linhagens?

   Solução:

$$\begin{aligned} X &: \text{peso de espigas} \\ X &\sim \mathrm{N}(\mu = 90, \sigma^2 = 7^2) \\ \mathrm{C} : \text{Condição} &: 78g < X < 104g \end{aligned}$$

   (a) $P[78 < X < 104] = P[\frac{78-90}{7} < Z < \frac{104-90}{7}] = P[-1,71 < Z < 2] = 0,934$

   (b) $200 \cdot \{1 - P[78 < X < 104]\} = 200 \cdot \{1 - 0,934\} = 13$

   (c) $P[X > k] = 0,15 \longrightarrow z = \frac{k-90}{7} = 1,04 \longrightarrow k = 90 + 7 \cdot 1,04 = 97,3$

---

4. Suponha agora, no contexto do problema anterior, que uma única linhagem foi selecionada para ser testada em maior detalhe e foram obtidos as seguintes medidas de pesos de suas espigas:
`102 89 85 91 88 95 99 87 91 95 93 89 92 97.`

(a) Obtenha duas medidas de posição dos dados.

(b) Obtenha duas medidas de dispersão dos dados.

(c) Construa um *boxplot* das observações.

(d) Obtenha um intervalo de confiança (95%) para a produtividade média.

(e) Baseado nas análises acima, pode-se qfirmar que a linhagem selecionada apresenta um comportamento *estatísticamente* diferente da população de pesos mencionada na questão anterior? Justifique sua resposta.

Solução:

(a) A média é $\overline{x} = 92,4$ e a mediana é $md(x) = 91,5$

(b) O desvio padrão é $S = 4,81$ e a amplitude interquartílica é $AI = 6$

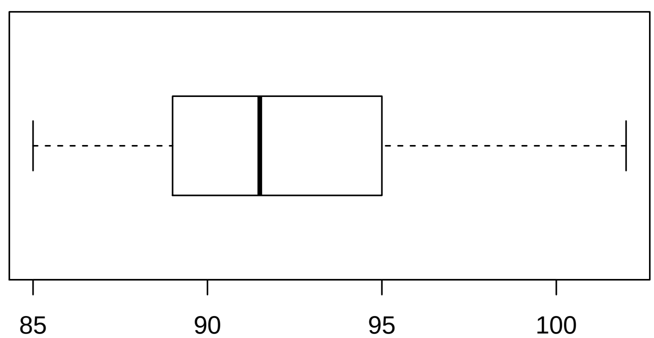

Figura 1: Box-plot dos dados de peso de espigas
não inclui

(c)

(d)

$$\overline{x} \pm t_{0.95,14} \frac{S}{\sqrt{n}}$$
$$92,36 \pm 0,826 \cdot 1,29$$
$$92,4 \pm 1,06$$
$$(91,29 \quad ; \quad 93,42)$$

(e) A linhagem selecionada apresenta uma média *estatisticamente* diferente da anterior (90) já que seu IC o valor 90.

---

5. Foram escolhidos ao acaso 500 animais (bovinos) de uma região para estimar a proporção de com propenção à uma certa doença. Destes, 120 testaram positivo.

(a) Obtenha a estimativa pontual do percentual de susceptíveis na população.

(b) Obtenha a estimativa intervalar (com confiança de 95%) do percentual de susceptíveis na população.

(c) Idem anterior porém com confiança de 80%.

(d) Deseja-se extender o levantamento para obter uma margem de erro de 1,5% para 95% de confiança. Quantos animais adicionais devem ser selecionados e testados?

Solução:

População:

$$\begin{aligned} &\text{X : propenção à doença, 0 - não propenso, 1 - propenso}\\ &\text{X} \sim B(p)\\ &p \text{ : parâmetro desconhecido}\end{aligned}$$

Amostra:

Dados:

$$n = 500 \;\; ; \;\; \sum_{i=1}^{500} x_i = 120$$

Estimador:

$$\hat{p} = \frac{\sum_{i=1}^{500} x_i}{n}$$

Estimativa:

$$\hat{p} = \frac{120}{500} = 0,24$$

Distribuição amostral:

$$\hat{p} \sim \text{N}(\mu_{\hat{p}} = p, \sigma^2_{\hat{p}} = \frac{p(1-p)}{n})$$

(a) $\hat{p} = 0,24$

(b) Intervalo assintótico (tomando $p = \hat{p}$)

$$\begin{aligned} &\hat{p} \pm z_{0.95}\sqrt{\frac{p(1-p)}{n}}\\ &0,24 \pm 1,96 \cdot 0,019\\ &0,24 \pm 0,0374\\ &(0,203 \;\; ; \;\; 0,277)\end{aligned}$$

(c)

$$\begin{aligned} &\hat{p} \pm z_{0,80}\sqrt{\frac{p(1-p)}{n}}\\ &0,24 \pm 1,28 \cdot 0,019\\ &0,24 \pm 0,0245\\ &(0,216 \;\; ; \;\; 0,264)\end{aligned}$$

(d)

$$
\begin{aligned}
\mathrm{ME} &= z_{0.95}\sqrt{\frac{p(1-p)}{n}} \\
0,015 &= 1,96^2\sqrt{\frac{0.24(1-0.24)}{n}} \\
n &= \lceil 1,96^2 \frac{0.24(1-0.24)}{0,015^2} \\
n &= 3115
\end{aligned}
$$

Portanto 2615 novas amostras.

Soluções computacionais com o programa R:

```
> prop.test(120, 500, conf=0.95)$conf

[1] 0,2037 0,2804
attr(,"conf.level")
[1] 0,95

> prop.test(120, 500, conf=0.80)$conf

[1] 0,2154 0,2663
attr(,"conf.level")
[1] 0,8
```