CE-003: Estatística II - Turma: K/O, Prova Final (20/03/2013)

GRR: _______________ **Nome:** ______________________________________ **Turma:** ___

1. Um professor preparou 40 versões diferentes de uma lista de exercícios. As listas são atribuídas ao acaso para estudantes sorteando-se para cada um um número de 1 a 40 que identifica a lista. Se um grupo de três colegas decide fazer as listas juntos, que a probabilidade de que ao menos dois deles recebam a mesma versão?

   Solução:

$$\begin{aligned}
\text{Espaço Amostral: } & S \text{ todas possíveis atribuições de 40 listas para 3 estudantes} \\
n(S) &= 40 \cdot 40 \cdot 40 \\
\text{Evento: } & E \text{ coincidência de lista em ao menos 2 estudantes} \\
& \overline{E} \text{ sem coincidência de listas} \\
n(\overline{E}) &= 40 \cdot 39 \cdot 38 \\
P[E] &= 1 - P[\overline{E}] = 1 - \frac{n(\overline{E})}{n(S)} = 1 - \frac{40 \cdot 39 \cdot 38}{40^3} = 0.0737.
\end{aligned}$$

   Nota: este exercício é semelhante ao problema da *concidência de aniversários* em um grupo de pessoas.

---

2. Um determinado componente de computador é despachados de uma fábrica para entrega em lotes de 70 unidades. Antes do despacho de um lote, 25 das unidades escolhidas ao acaso passam por um teste detalhado. Se alguma destas unidades falha no teste, todo o lote é rejeitado.

   (a) Qual a probabilidade de um lote contendo exatamente quatro unidades falhas ser despachado?

   (b) Indique os cálculos necessários para obter o número de unidades que deveriam ser testadas para que esta probabilidade não ultrapasse 0,02?

   Solução:

$$X : \text{número de unidades que apresentam falhas entre as unidades testadas}$$

$$X \sim HG(N = 70, K = 4, n = 25) \quad P[X = x] = \frac{\binom{K}{x}\binom{N-K}{n-x}}{\binom{N}{n}}$$

   (a) $P[X = 0] = \frac{\binom{4}{0}\binom{66}{25}}{\binom{70}{25}} = 0.1625$

   (b) $n = ?|P[X = 0] \leq 0,02 \longrightarrow \frac{\binom{4}{0}\binom{66}{n}}{\binom{70}{n}} \leq 0,02 \longrightarrow \frac{(70-n)!}{(66-n)! \prod_{i=67}^{70} i} \overset{\text{sol. numérica}}{\longrightarrow} n \geq 43$

---

3. Sabe-se que em um sistema de transmissão de dados, uma tempestade causa, em média, a falha de transmissão de um *pacote* em cada 200. Transmitindo 500 *pacotes* nestas condições, qual a probabilidade que:

   (a) no máximo três *pacotes* não sejam transmitidos
   (b) todos sejam transmitidos

Solução:

$$X : \text{número falhas de transmissão em 500 pacotes}$$
$$X \sim B(n = 500, p = 1/200) \quad P[X = x] = \binom{n}{x} p^x (1-p)^{(} n - x)$$
$$X \approx P(\lambda = 500 \cdot (1/200)) \quad P[X = x] = \frac{e^{-\lambda}\lambda^x}{x!}$$

   (a) $P[X \leq 3] = 0.7578$
   (b) $P[X = 0] = 0.0816$

---

4. Sabe-se que o tempo de uso de um determinado serviço pode ser descrito pela distribuição exponencial e tem média de 30 segundos. Qual a probabilidade de um acesso:

   (a) ter o tempo inferior a 12 segundos;
   (b) durar entre 12 e 30 segundos;
   (c) ultrapassar 42 segundos, sabendo-se que já chegou a 30 segundos?

   Os acessos são divididos em três categorias: rápidos (com tempo inferior a 12 segundos), médios (com tempo entre 12 e 30 segundos), longos (com tempos entre 30 e 40 segundos) e demorados (com tempo superior a 42 segundos). Os custos e cada categoria são definidos como sendo, respectivamente 1,5; 2,8; 5,0 e 10,0 unidades de custo.

   (d) Qual o custo médio dos acessos?

   Se for tomada uma amostra de 100 tempos, qual a probabilidade de que o tempo médio desta amostra:

   (e) ultrapasse 35 segundos;
   (f) esteja entre 28 e 35 segundos?
   (g) Qual deve ser o tamanho da amostra para que a probabilidade do tempo médio da amostra ultrapassar 35 segundos seja no máximo de 0,01?

   Desconfia-se que houve alguma mudança no padrão dos tempos de acesso e para verificar isto tomou-se uma amostra de 100 tempos de acesso. A seguir estão alguns desses valores.

```
> cat(paste(tempos[1:10], "  ...    ..."),tempos[96:100])

16.8   ...    ... 2.2   ...    ... 16.1   ...    ... 9.5   ...    ... 19.5   ...    ... 1.5   .
```

   A média dos tempos amostrados é de 32.71.

   (h) Obtenha um intervalo de confiança (95%) para o tempo médio.
   (i) Voce diria que há evidências fortes de que houve uma mudança no comportamento dos tempos de acesso? Justifique.

Solução:

$$X: \text{tempo de acesso ao serviço}$$
$$X \sim \mathrm{E}(\lambda = 1/30) \quad f(x) = \frac{1}{30}\exp\{-x/30\} \quad F(x) = 1 - \exp\{-x/30\}$$

(a) $P[X < 12] = \int_{x=0}^{12} f(x)\mathrm{d}x = F(12) = 0.3297$

(b) $P[12 < X < 30] = \int_{x=12}^{30} f(x)\mathrm{d}x = F(30) - F(12) = 0.3024$

(c) Pela propriedade de falta de memória de exponencial,
$P[X > 42|X > 30] = P[X > 12] = 1 - P[X < 12] = 1 - \int_{x=0}^{12} f(x)\mathrm{d}x = 1 - F(12) = 0.6703$

(d) Y : custo de acesso

y	$1,5$	$2,8$	$5,0$	10
$P[Y = y]$	$P[X < 12] = 0.33$	$P[12 < X < 30] = 0.302$	$P[30 < X < 42] = 0.121$	$P[X > 42] = 0.247$

$\mathrm{E}(X) = \sum_i y_i P[Y = y_i] = 1,5 \cdot 0.33 + 2,8 \cdot 0.302 + 5,0 \cdot 0.121 + 10,0 \cdot 0.247 = 4.4$ unidades de custo

$$X \sim \mathrm{Exp}(\lambda = 1/30) \quad \mathrm{E}(X) = 1/\lambda \quad \mathrm{Var}(X) = 1/\lambda^2$$
$$\overline{X}_n \approx N(\mu = 1/\lambda = 30, \sigma^2 = (1/\lambda^2)/n = 30^2/n) \quad n = 100$$

(e) $P[\overline{X} > 35] = P[Z > \frac{35-30}{30/10}] = P[Z > 1.667] = 0.048$

(f) $P[28 < \overline{X} < 35] = P[\frac{28-30}{30/10} < Z < \frac{35-30}{30/10}] = P[-0.667 < Z < 1.667] = 0.7$

(g)

$$P[\overline{X} > 35] \leq 0,01$$
$$P[Z > \frac{35-30}{30/\sqrt{n}}] \leq 0,01$$
$$\frac{35-30}{30/\sqrt{n}} \geq 2.326$$
$$n \geq \frac{2.326^2 30^2}{(35-30)^2} = 195$$

$$\text{Assumindo que } \overline{X}_n \approx N(\mu = 1/\hat{\lambda} = 30, \sigma^2 = (1/\hat{\lambda}^2)/n = 30^2/n)$$

(h) $\hat{\lambda} \pm z_{0,95}\hat{\lambda}/\sqrt{n} \longrightarrow 32.71 \pm 1.96\frac{32.71}{10} \longrightarrow (26.3 \;\; ; \;\; 39.12)$

(i) Justifique.

---

5. Considere os dados a seguir.

`13 4 5 6 5 7 7 4 17 6 17 3 8 18 5 8 10 6 5 4 4 8 3 8 18`

(a) Calcule a média e mediana dos dados.

(b) Calcule o desvio padrão, coeficiente de variação.

(c) Faça um histograma dos dados.

(d) Faça um gráfico *box-plot*.

(e) Faça um diagrama *ramo-e-folhas*.

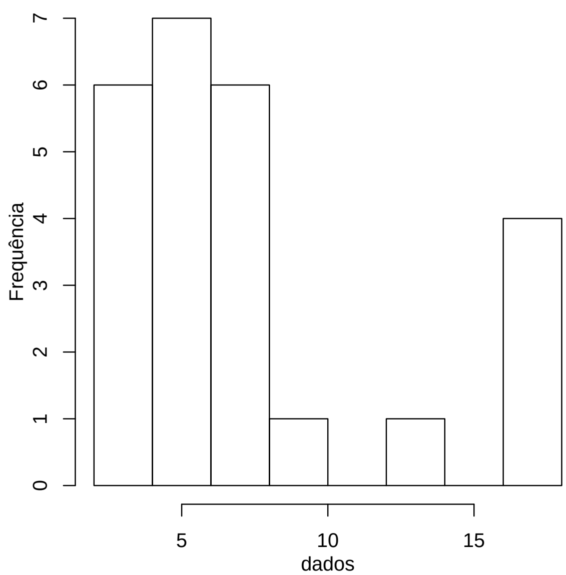
(f) Caracterize/descreva a distribuição dos dados.

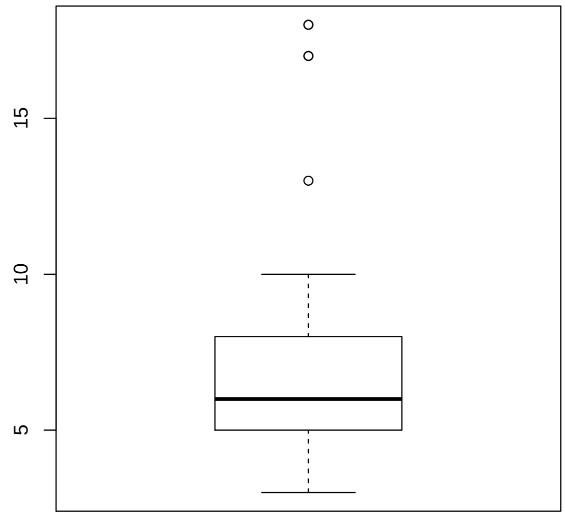
Solução:

(a) $\overline{x} = 8 \quad \text{md} = 6$

(b) $S = 4.8 \quad \text{CV} = 80.1$

c) d)





e)
```
The decimal point is 1 digit(s) to the right of the |

0 | 334444
0 | 5555666778888
1 | 03
1 | 7788
```