CE-003: Estatística II - Turma: K/O, Prova Final (08/12/2014)

GRR: ______________ **Nome:** ______________________________________ **Turma:** ___

1. O escore de um estudante em um determinado exame é padronizado e expresso por um número entre 0 e 1. Suponha que um estudante é aprovado se obtiver um escore de pelo menos 0,65. Suponha ainda que o resultado pode ser modelado por uma variável aleatória com função de densidade de probabilidades dada por:

$$f(x) = \begin{cases} \frac{10}{3}x & \text{se } 0 < x < 0,6 \\ 5 - 5x & \text{se } 0,6 < x < 1 \\ 0 & \text{caso contrário} \end{cases}$$

   (a) Qual a probabilidade que um estudante não seja aprovado no exame?

   (b) Qual o escore mediano?

   (c) Qual o escore médio?

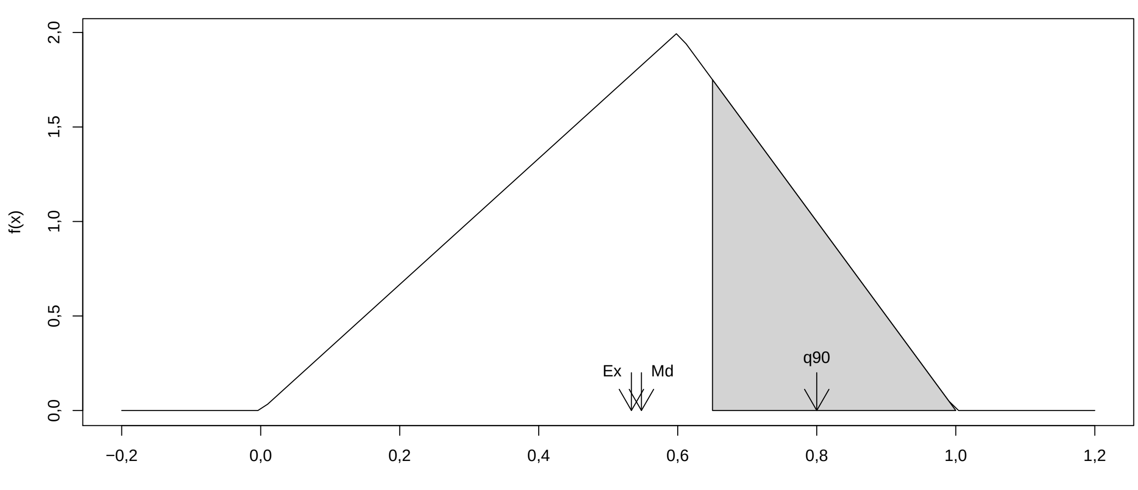
   (d) Qual o escore será obtido com apenas 10% de chance de haver escores acima dele?

   (e) Se um estudante foi aprovado, qual a probabilidade de que seu escore tenha sido superior a 0,80?

   (f) Suponha agora que são sorteados 5 estudantes ao acaso deste grupo para fazer uma entrevista de avaliação. Faça suposições adequadas e obtenha a probabilidade de que: (i) nenhum deles tenha sido aprovado; (ii) ao menos um tenha disto reprovado.

   (g) Suponha ainda que agora vai se sortear estudantes ao acaso para entrega de um prêmio de participação para dois aprovados. Qual a probabilidade de que seja necessário fazer mais de três sorteios (i.e. quatro ou mais sorteios) até se sortear o segundo aprovado?

Solução:



   (a) $P[X < 0,65] = 1 - P[X > 0,65] = 1 - \frac{(1-0,65)(5-5*0,65)}{2} = 0,694$

(b) $\mathrm{Md}(x) = q_{0,50} \longrightarrow P[X < q_{0,50}] = 0,50 \rightarrow \frac{(q_{0,50})((10/3)q_{0,50})}{2} = 0,50 \rightarrow q_{0,50} = 0,548$

(c) $\mathrm{E}[X] = \int_0^1 f(x)\mathrm{d}x = 0,533$

(d) $P[X < q_{0,90}] = 0,90 \rightarrow \frac{(1-q_{0,90})(5-5q_{0,90})}{2} = 0,90 \rightarrow q_{0,90} = 0,8$

(e) $P[X > 0,80|X > 0,65] = \frac{P[X>0,80]}{P[X>0,65]} = \frac{(1-0,80)(5-5*0,65)/2}{(1-0,65)(5-5*0,65/2)} = 0,327$

(f) Supondo sorteio com reposição ou que a população de estudantes é "grande":

$$Y : \text{número de estudantes aprovados entre os 5 escolhidos}$$
$$Y \sim \mathrm{B}(n = 5, p = P[X > 0,65] = 0,306)$$
$$(i) P[Y = 0] = \binom{5}{0}(0,306)^0(1-0,306)^5 = 0,161$$
$$(ii) P[Y \leq 4] = 1 - P[Y = 5] = 1 - \binom{5}{5}(0,306)^5(1-0,306)^0 = 0,997$$

(g) Supondo sorteio com reposição ou que a população de estudantes é "grande":

$$Y : \text{número de estudantes sorteados "reprovados" até obter o segundo aprovado}$$
$$Y \sim \mathrm{BN}(r = 2, p = P[X > 0,65] = 0,306)$$
$$(i) P[Y \geq 2] = 1 - P[Y \leq 0] - P[Y \leq 1] =$$
$$1 - \binom{0+2-1}{2-1}(1-0,306)^0(0,306)^2 - \binom{1+2-1}{2-1}(1-0,306)^1(0,306)^2 = 0,776$$

---

2. O diagrama ramo-e-folhas abaixo mostra medidas do fluxo anual do rio Nilo próximo à cidade de Ashwan no período de 1871-1970.

```
A casa decimal esté 2 digitos à direita de |

  4 | 6
  5 |
  6 | 5899
  7 | 000123444455667778
  8 | 0000112222333445555556667779
  9 | 0011222244466678899
 10 | 0122234455
 11 | 00012244566678
 12 | 112356
 13 | 7
```

(a) Obtenha a mediana e quartis dos dados

(b) Obtenha o máximo, mínimo, primeiro e nono decis

(c) Faça um diagrama *box-plot* dos dados

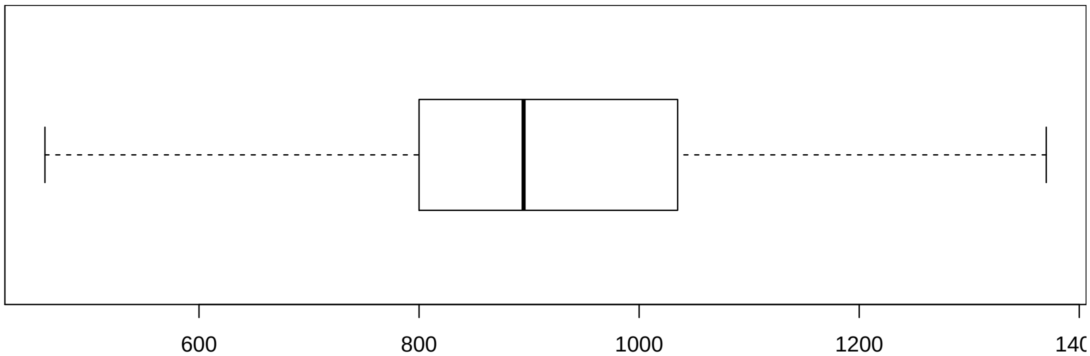
(d) O que pode ser dito da distribuição dos dados baseando-se nos gráficos e medidas?

Solução:

(a)
```
Mediana 1o Quartil 3o Quartil
    895        800       1035
```

(b)
```
  Min      Max 1o decil 9o decil
  460     1370      725     1160
```

(c) boxplot:



3. 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.

```
16,8   2,2   16,1   9,5   19,5   1,5   26   6,4   18,9   25,6   ...   ... 26,1   9,5   10,
```

   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.