CE-003: Estatística II - Turma: AMB, 2^a Prova (25/07/2013)

GRR: ________________ **Nome:** ____ **Turma:** ___

1. Em cada uma das situações a seguir: identifique a variável aleatória, a distribuição adequada e calcule a probabilidade solicitada. Em um estudo de ecologia definem-se como unidades de investigação parcelas de $25m^2$. O interesse é verificar a ocorrência de uma determinada espécie vegetal em parcelas alocadas aleatoriamente em uma área de vegetação natural. Espera-se que 70% das parcelas contenham a espécie com um número médio de 1,21 plantas por parcela da espécie.

   (a) Qual a probabilidade de encontrar a espécie em ao menos três dentre cinco parcelas observadas?

   (b) Se parcelas forem inspecionadas até encontrar três delas com presença da espécie, qual a probabilidade de ser necessário examinar mais que cinco parcelas?

   (c) Se um estudo preliminar registrou a presença em 35 entre 100 parcelas examinadas e uma outra equipe revisita cinco das parcelas, qual a probabilidade de encontrar a espécie em no máximo uma delas?

   (d) Qual a probabilidade de se encontrar duas ou mais plantas em uma parcela?

Solução:

(a)

$$\begin{aligned}\text{Evento } X_A &: \text{ número de parcelas com a espécie entre cinco parcelas}\\ X_A &\sim \text{Bin}(n=5, p=0,70)\\ P[X_A \geq 3] &= P[X_A = 3] + P[X_A = 4] + P[X_A = 5] = 0.837\end{aligned}$$

(b)

$$\begin{aligned}\text{Evento } X_B &: \text{ número de parcelas inspecionadas até encontrar a } 3^a \text{ com a espécie}\\ X_B &\sim \text{NB}(k=2, p=0,15)\\ P[X_B \geq 5] &= 1 - P[X_B = 3] - P[X_B = 4] - P[X_B = 5] = 0.163\end{aligned}$$

(c)

$$\begin{aligned}\text{Evento } X_C &: \text{ número de parcelas com a espécie dentre 5 revisitadas}\\ X_C &\sim \text{HG}(N=100, K=35, n=5)\\ P[X_C \leq 1] &= P[X_C = 0] + P[X_C = 1] = 0.424\end{aligned}$$

(d)

$$\begin{aligned}\text{Evento } X_D &: \text{ número de plantas por parcela}\\ X_D &\sim \text{P}(\lambda = 1,21)\\ P[X_D \geq 2] &= 1 - P[X_D = 0] - P[X_D = 1] = 0.341\end{aligned}$$

---

2. Seja a função $f(x) = \exp\{-(x-K)\}$ para $4 < x < +\infty$.

   (a) qual deve ser o valor de k para que $f(x)$ seja função de densidade de probabilidade?

   (b) calcule $P[X > 6]$

   (c) calcule $P[X < 8|X > 5]$

   (d) encontre os quartis desta distribuição.

   (e) encontre o valor esperado desta distribuiçao

Solução:

(a) $\int_4^\infty f(x)\mathrm{d}x = 1 \longrightarrow K = 4$

(b) $P[X > 6] = \int_6^\infty f(x)\mathrm{d}x = 0.135$

(c) calcule $P[X < 8|X > 5] = \frac{P[5<X<8}{P[X>5]} = \frac{\int_5^8 f(x)\mathrm{d}x}{\int_5^\infty f(x)\mathrm{d}x} = 0.95$

(d)

$$\int_4^{Q_1} f(x)\mathrm{d}x = 0,25 \longrightarrow Q_1 = 4.3$$
$$\int_4^{Q_2} f(x)\mathrm{d}x = 0,50 \longrightarrow Q_2 = 4.7$$
$$\int_4^{Q_3} f(x)\mathrm{d}x = 0,75 \longrightarrow Q_3 = 5.4$$

(e) $E[X] = \int_4^\infty x \cdot f(x)\mathrm{d}x = 5$

Complementos e soluções alternativas:

$$
\begin{aligned}
F(x) &= \int_0^x f(x)\mathrm{d}x = \frac{1}{6}(5x - 2x^2) \\
(\text{Quantis})x_q &= \frac{5 - \sqrt{25 - 24p}}{2}(\text{obtido com } F^{-1}(x)) \\
P[X > 6] &= 1 - Fx(6) = 0.135 \\
P[X < 8|X > 5] &= \frac{F(8) - F(5)}{1 - F(5)} = 0.95 \\
Q_1 &= x_{0,25} = F^{-1}(0,25) = 4.29 \\
Q_2 &= x_{0,50} = F^{-1}(0,50) = 4.69 \\
Q_3 &= x_{0,75} = F^{-1}(0,75) = 5.39
\end{aligned}
$$

---

3. Assume-se que a geração diária de energia (em escala logarítmica) por uma turbina eólica possui distribuição normal de média 2,5 e desvio padrão 0,60 unidades.

(a) Qual a probabilidade da geração ser inferior a 2 unidades em um particular dia?

(b) Qual a probabilidade da geração diária ficar entre 2 e 3 unidades?

(c) Define-se como o *o valor mínimo de referência* o valor gerado em pelo menos 80% dos dias? Qual é este valor?

(d) Qual a probabilidade de gerar mais que 18 unidades em uma semana (7 dias)?

(e) Qual a probabilidade da geração diária média em um mês (30 dias) ultrapassar 2,6 unidades?

Solução:

$$
\begin{array}{ll}
X : \text{geração diária (log)} & X \sim \mathrm{N}(2,5 \ ; \ 0,6^2) \\
\overline{X}_n : \text{média da geração diária em } n \text{ dias} & \overline{X}_n \sim \mathrm{N}(2,5 \ ; \ 0,6^2/n)
\end{array}
$$

(a) $P[X < 2] = P[Z < \frac{2-2,5}{0,6}] = P[Z < -0.833] = 0.202$

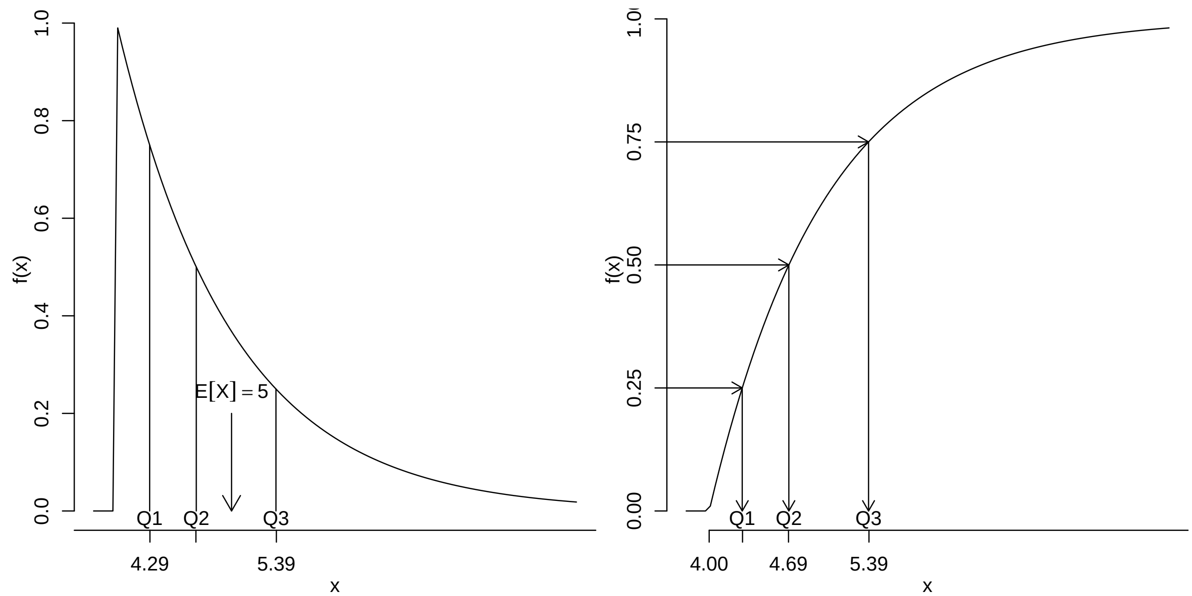

Figura 1: Função de densidade de probabilidades (esquerda) e de densidade acumulada (direita). Valores marcados são os quartis e esperança da v.a. X.

(b) $P[2 < X < 3] = P[\frac{2-2,5}{0,6} < Z < \frac{3-2,5}{0,6}] = P[-0.833 < Z < 0.833] = 0.595$

(c)

$$
\begin{aligned}
P[X > x] &= P[Z > z] = 0,80 \\
z &= \frac{x-2,5}{0,6} = -0.8416 \\
x &= 2,5 + 0,6 \cdot (-0.8416) = 2.00
\end{aligned}
$$

(d) $P[\sum i = 1_7 X_i > 18] = P[\overline{X}_7 > 18/7] = P[Z < \frac{(18/7)-2,5}{0,6}] = P[Z < 0.119] = 0.376$

(e) $P[\overline{X}_{30} > 2,6] = P[Z < \frac{2,5-2,5}{0,6/\sqrt{30}}] = P[Z < 0.913] = 0.181$

---

4. Foi feita uma pesquisa com 1.500 pessoas que acessam um site para se estimar a proporção de mulheres que acessam o site. Encontrou-se que 1150 acessos eram de mulheres.

   (a) Obtenha a margem de erro para pesquisa para confiança de 95%

   (b) Qual seria esta margem e erro para uma amostra de 500 pessoas?

   (c) Qual deveria ser o tamanho da amostra para que a margem de erro não ultrapasse 1%

   (d) Teste, com base na pesquisa feita, a hipótese de que as mulheres representam mais que 75% dos acessos.

Solução:

(a)

$$
\begin{aligned}
\text{M.E.} &= z\sqrt{\frac{p(1-p)}{n}} \\
\text{M.E. (assintótica)} &= z\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = 1.96\sqrt{\frac{0.77(1-0.77)}{1500}} = 0.021 \\
\text{M.E. (conservadora)} &= z\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = 1.96\sqrt{\frac{1}{4 \cdot 1500}} = 0.025
\end{aligned}
$$

(b)

$$\text{M.E.} = z\sqrt{\frac{p(1-p)}{n}}$$
$$\text{M.E. (assintótica)} = z\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = 1.96\sqrt{\frac{0.77(1-0.77)}{500}} = 0.037$$
$$\text{M.E. (conservadora)} = z\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = 1.96\sqrt{\frac{1}{4\cdot 500}} = 0.044$$

(c) Qual deveria ser o tamanho da amostra para que a margem de erro não ultrapasse 1%

$$\text{M.E.} = z\sqrt{\frac{p(1-p)}{n}}$$
$$n = \frac{z^2}{\text{M.E.}^2}p(1-p)$$
$$\text{n (assintótico)} = \frac{1.96^2}{0,01^2}0.77(1-0.77) = 6872$$
$$\text{n (assintótico)} = \frac{1.96^2}{0,01^2}0,5(1-0,5) = 9604$$

(d)

$$H_0 : p = 0,75 \quad vs \quad H_a : p > 0,75$$
$$\alpha = 0,95$$
$$z_c = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} = \frac{0.77 - 0,75}{\sqrt{\frac{0,75(1-0,75)}{1500}}} = 1.526$$
$$z_{0,95} = 0.6745$$
$$\text{p} - \text{valor} = 0.0635$$
$$\text{Conclusão : Não Rejeita H0}$$

---

5. Foi tomada uma amostra para estimar a biomassa em uma área de regeneração a partir de uma amostra de um conjunto de parcelas, obtendo os valores a seguir. Faça as suposições necessárias e obtenha intervalos e confiança (90%) para a média e para a variância da biomassa. Teste a hipótese de que a biomassa está abaixo do valor esperado de 157 unidades.

```
156,4   155,7   157,1   155,9   156,9  160,1   154,9   156,1   157,4   158,5   155,4   157,1
```

Solução:

(a)

$$\text{I.C.}_{90\%}(\mu):$$

$$\overline{x} \pm t_{(n-1),90\%}\frac{S}{\sqrt{n}}$$

$$156.8 \pm 0.819\frac{1.43}{\sqrt{12}}$$

$$156.8 \pm 0.743$$

$$(156.0 \;\; ; \;\; 157.5)$$

$$\text{I.C.}_{90\%}(\sigma^2):$$

$$\left(\frac{(n-1)S^2}{\chi^2_{sup}} \;\; ; \;\; \frac{(n-1)S^2}{\chi^2_{inf}}\right)$$

$$\left(\frac{(12-1)2.05}{19.7} \;\; ; \;\; \frac{(12-1)2.05}{4.57}\right)$$

$$(1.147 \;\; ; \;\; 4.933)$$

(b)

$$H_0 : \mu = 157 \quad vs \quad H_a : \mu < 157$$

$$\alpha = 0,10$$

$$t_c = \frac{\overline{x} - \mu_0}{S/\sqrt{n}} = \frac{156.8 - 157}{1.43/12} = -0.5038$$

$$t_{0,90} = -1.363$$

$$\text{p} - \text{valor} = 0.312$$

$$\text{Conclusão : Não Rejeita H0}$$
