CE-003: Estatística II, 3^a Prova - 2^o semestre 2011 (21/12/2011)

1. O tempo de afastamentos por motivo de saúde solicitadas em um orgão público em um mês tem distribuição normal de média 250 horas e desvio padrão de 25 horas. Qual a probabilidade de que o tempo de afastamento no próximo mês:

   (a) fique entre 200 e 300 horas;

   (b) não ultrapasse 280 horas;

   (c) se afaste de média por mais de 40 horas.

   Qual o tempo de afastamento que

   (d) é superado com probabilidade de apenas 0,05;

   (e) deve se superado em 90% dos meses.

   Solução:

$$X : \text{tempo de afastamento}$$
$$X \sim N(\mu = 250, \sigma^2 = 25^2)$$

   (a) $P[200 < X < 300] = P[\frac{200-250}{25} < Z < \frac{300-250}{25}] = P[-2 < Z < 2] = 2 \cdot P[0 < Z < 2] = 2 \cdot 0,4772 = 0,9545$

   (b) $P[X \leq 280] = P[Z < \frac{280-250}{25}] = P[Z < 1,2] = 0,5 + P[0 < Z < 1,2] = 0,5 + 0,3849 = 0,8849$

   (c) $P[|X-250| > 40] = P[X < 210] + P[X > 290] = P[Z < \frac{210-250}{25}] + P[Z > \frac{290-250}{25}] = P[Z < -1,6] + P[Z > 1,6] = 2 \cdot (0,5 - 0,4452) = 0,1096$

   (d) $P[X > K_1] = 0,05 \longrightarrow z = 1,645 \longrightarrow \frac{K_1-250}{25} = 1,645 \longrightarrow K_1 = 291,1.$

   (e) $P[X > K_2] = 0,90 \longrightarrow z = -1,282 \longrightarrow \frac{K_2-250}{25} = -1,282 \longrightarrow K_2 = 218.$

```
> ## a)
> diff(pnorm(c(200,300), m=250, sd=25))

[1] 0,9544997

> ## b)
> pnorm(280, m=250, sd=25)

[1] 0,8849303

> ## c)
> pnorm(210, m=250, sd=25) + pnorm(290, m=250, sd=25, low=FALSE)

[1] 0,1095986

> ## d)
> qnorm(0.95, m=250, sd=25)

[1] 291,1213

> ## e)
> qnorm(0.10, m=250, sd=25)

[1] 217,9612
```

2. O tempo de atendimento a usuários em uma central de atendimento tem distribuição normal de média 12 minutos de desvio padrão de 2 minutos. Qual a probabilide de que uma chamada dure:

(a) dure menos que 10 minutos;

(b) mais que 15 minutos.

Se um atendente recebe 36 chamadas em um dia, qual a probabilidade de que:

(c) o tempo total de atendimento ultrapasse 7 horas;

(d) o tempo médio de atendimento ultrapasse 12,5 minutos.

Pretende-se reduzir o tempo total médio de atendimento. Encontre quantas chamadas por dia o atendente deve receber para que

(e) o tempo total médio seja de 6 horas;

(f) o tempo total não ultrapasse 6 horas com 0,90 de probabilidade.

Se um atendente trabalhar um período de 4 horas qual a probabilidade de

(g) conseguir atender ao menos 18 chamadas.

Solução:

$$X : \text{tempo de atendimento (min)}$$
$$X \sim N(\mu = 12, \sigma^2 = 2^2)$$

(a) $P[X < 10] = P[Z < \frac{10-12}{2}] = P[Z < -1] = 0,5 - 0,3413 = 0,1587$

(b) $P[X > 15] = P[Z > \frac{15-12}{2}] = P[Z > 1,5] = 0,5 - 0,4332 = 0,0668$

$$\bar{X}_{36} : \text{tempo médio de atendimento de 36 chamadas (min)}$$
$$\bar{X}_{36} \sim N(\mu = 12, \sigma^2 = 2^2/36)$$

(c) $P[\sum_{i=1}^{36} X_i > 420] = P[\bar{X}_{36} > 420/36] = P[Z > \frac{(420/36)-12}{2/\sqrt{36}}] = P[Z > -1] = 0,8413.$

(d) $P[\bar{X}_{36} > 12,5] = P[Z > \frac{12,5-12}{2/\sqrt{36}}] = P[Z > 1,5] = 0,0668.$

$$S_k = \sum_{i=1}^{k} x_i : \text{tempo total de atendimento de k chamadas (min)}$$
$$S_k \sim N(\mu = 12 \cdot k, \sigma^2 = k \cdot 2^2)$$

(e) $E[S_k] = k \cdot E[X] = k \cdot 12 = 360 \longrightarrow k = 30$ chamadas

(f) $P[S_k < 360] = 0,90 \longrightarrow z = 1,282 \longrightarrow \frac{360-12k}{\sqrt{k}\cdot 2} = 1,282 \longrightarrow k = 29$

$$Y : \text{número de atendimento em 4 horas}$$
$$Y \approx N(\mu = 20, \sigma^2 = 20)$$

(g) $P[Y > 18] = P[Z > \frac{18-20}{\sqrt{20}}] = 0,681$

```
> # a)
> pnorm(10, m=12, sd=2)
```

```
[1] 0,1586553

> # b)
> pnorm(15, m=12, sd=2, lower=FALSE)

[1] 0,0668072

> # c)
> pnorm(420/36, m=12, sd=2/6, low=F)

[1] 0,8413447

> # d)
> pnorm(12.5, m=12, sd=2/6, low=F)

[1] 0,0668072

> # e)
> f <- function(k) (360 - 12 *k) - 2*sqrt(k) * qnorm(0.90)
> ceiling(uniroot(f, c(1, 40))$root)

[1] 29

> # f)
> pnorm(18, m=20, sd=sqrt(18), lower=FALSE)

[1] 0,6813241
```

---

3. Uma amostra aleatória de 60 impressoras foi submetida a testes de impressão e em 12 delas algum defeito relevante foi observado.

   (a) Encontre um intervalo de confiança (95%) para a proporção p de defeitos verdadeira (populacional).

   (b) Uma impressora é considerada aceitável se a proporção p é menor que 0,25 (25%). Use um teste de hipótese adequado para decidir ($\alpha = 0,01$) se a impressora é aceitável.

   (c) Usando a estimativa da proporção obtida na amostra como sendo o verdadeiro valor da proporção p, calcule um novo tamanho de amostra de impressoras a serem testadas para que a verdadeira proporção seja estimada com uma margem de erro de 0,025.

   (d) Quanto deveria ser o novo tamanho da amostra no item anterior para que a margem de erro fosse de no máximo 0,025 independentemente do valor da verdadeira proporção?

Solução:

$$\begin{aligned} &X : \text{defeito} \\ &X \sim Ber(p) \\ &\hat{p} \sim N(p, \frac{p(1-p)}{n}) \\ &n = 60 \quad \hat{p} = 12/60 = 0,2 \end{aligned}$$

(a) IC assintótico:
$\hat{p} \pm z\sqrt{\frac{p(1-p)}{n}} \longrightarrow 0,20 \pm 1,96\sqrt{\frac{0,2\cdot(1-0,2)}{60}} \longrightarrow 0,20 \pm 0,101 \longrightarrow (0,099; 0,301)$
IC conservador:
$\hat{p} \pm z\sqrt{\frac{1}{4n}} \longrightarrow 0,20 \pm 1,96\sqrt{\frac{1}{4\cdot 60}} \longrightarrow 0,20 \pm 0,127 \longrightarrow (0,073; 0,327)$

(b)

$$H_0 : p \geq 0,25 \text{ (não aceitável)} \quad vs \quad H_0 : p < 0,25 \text{ (aceitável)}$$
$$\alpha = 0,01 \qquad RC : \{z > -2,326\}$$
$$z_c = \frac{0,20 - 0,25}{\sqrt{\frac{0,25(1-0,25)}{60}}} = -0,894$$
$$\text{valor-p} = 0,18555$$

Conclusão: $z_c \notin RC$ ou valor-p $> \alpha(0,01)$. Não rejeita-se H_0 para o nível de significância de 1%, ou seja, não há evidências suficientes de que a proporção seja inferior a 25% e a impressora é considerada não aceitável.

(c) $z\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = 0,025 \longrightarrow n = \lceil \frac{z^2\hat{p}(1-\hat{p})}{0,025^2} = \lceil \frac{1,96^2(0,2)(0,8)}{0,025^2} = 984$

(d) $z\sqrt{\frac{1}{4n}} = 0,025 \longrightarrow n = \lceil \frac{z^2}{4 \cdot 0,025^2} = \lceil \frac{1,96^2}{4 \cdot 0,025^2} = 1537$

---

4. Mostre que, assumindo-se que uma variável possa ser descrita por um valor médio μ mais um desvio aleatório tal que $y_i = \mu + \epsilon_i$, o estimador $\hat{\mu} = \bar{x} = \sum_{i=1}^n x_i$ é um estimador de mínimos quadrados de μ.

$$Q(\mu) = \sum_{i=1}^n \epsilon_i^2 = \sum_{i=1}^n (y_i - \mu)^2$$
$$\frac{\mathrm{d}Q(\mu)}{\mathrm{d}\mu} = 0 \longrightarrow \sum_{i=1}^n (-2)(y_i - \hat{\mu}) = 0 \longrightarrow \sum_{i=1}^n y_i - n\hat{\mu} = 0$$
$$\hat{\mu} = \frac{\sum_{i=1}^n y_i}{n}$$

---

5. Encontre a expressão do estimador de máxima verosimilhança da parâmetro da distribuição exponencial. Usando este estimador obtenha a estimativa do parâmetro para uma amostra que forneceu os seguinte valores: `23,7; 12,8; 11,1; 5,0; 7,0; 10,1; 2,3; 5,4; 15,2; 8,5; 4,7; 1,5.`

$$X \sim \text{Exp}(\lambda)$$
$$f(x) = \lambda e^{-\lambda x} \mathrm{I}_{(0,\infty)}(x)$$
$$l(\lambda) = \log \prod_{i=1}^n f(x_i) = \sum_{i=1}^n \log\{f(x_i)\} =$$
$$= \sum_{i=1}^n (\log(\lambda) - \lambda x_i) = n\log(\lambda) - \lambda \sum_{i=1}^n x_i$$
$$\frac{\mathrm{d}l(\lambda)}{\mathrm{d}\lambda} = 0 \longrightarrow \frac{n}{\hat{\lambda}} - \sum_{i=1}^n x_i = 0 \longrightarrow \hat{\lambda} = \frac{n}{\sum_{i=1}^n x_i} = \frac{1}{\bar{x}}$$

Para os dados da amostra a estimativa é:

$$\hat{\lambda} = 1/\bar{x} = 0,11$$