CE-003: Estatística II - Turma: K/O, 3^a Prova (28/11/2018)

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

1. Assume-se que em uma rede com um grande número de nós 20% deles podem estar inacessíveis a qualquer tempo. São feitas inspeções periódicas em 400 nós da rede escolhidos ao acaso e se 90 (22,5%) ou mais desses estão inacessíveis, é feito um rastreamento detalhado para verificação e detecção de problemas.

   (a) Identifique no contexto: a população, a amostra, o estimador, a distribuição amostral e a região crítica.

   (b) Qual a probabilidade de encontrar mais que 85 nós inacessíveis em uma inspeção?

   (c) Se a rede está normal, qual a probabilidade de encontrar 100 ou mais nós inacessíveis em uma inspeção?

   (d) Qual os valores de proporções ao redor do valor médio (20%) de nós inacessíveis, dentre os quais devem-se estar 80% das inspeções.

   (e) Qual a probabilidade de fazer um rastreamento desnecessário?

   (f) Quantos nós deveriam ser inspecionados para que a chance de fazer um rastreamento desnecessário não ultrapasse 1%?

Solução:

$$
\begin{aligned}
X &: \text{estado do nó } (0 \text{ - } \text{ acessível}, 1: \text{ inacessível}) \\
X &\sim B(\mu = E[X] = p = 0,20, \sigma^2 = \text{Var}[X] = p(1-p) = 0,2 \cdot 0,8 = 0,16) \\
\hat{p} = \overline{X} &: \text{proporção de acessíveis entre n (400) nós} \\
\hat{p} = \overline{X} &\sim N(\mu_{\hat{p}} = p = 0,20, \sigma^2_{\hat{p}} = \frac{p(1-p)}{n} = \frac{0,2(1-0,2)}{400} = 0,0004 = 0,02^2)
\end{aligned}
$$

(a) ...

(b) $P[\hat{p} > 85/400] = P[Z > \frac{85/400 - 0,20}{0,02}] = P[Z > 0,625] = 0,5 - P[0 < Z < 0,625] = 0,5 - 0,234 = 0,266$

(c) $P[\hat{p} > 100/400] = P[Z > \frac{100/400 - 0,20}{0,02}] = P[Z > 2,5] = 0,5 - P[0 < Z < 2,5] = 0,5 - 0,4938 = 0,00621$

(d)

$$
\begin{aligned}
P[p - \Delta_{\hat{p}} < X < p + \Delta_{\hat{p}}] &= 0,80 \\
z = \frac{p + \Delta_{\hat{p}} - p}{0,02} &= 1,28 \\
\Delta_{\hat{p}} &= 0,0256 \\
P[0,174 < X < 0,226] &= 0,80
\end{aligned}
$$

(e) $P[\hat{p} > 0,225] = P[Z > \frac{0,225 - 0,20}{0,02}] = P[Z > 1,25] = 0,5 - P[0 < Z < 1,25] = 0,5 - 0,3944 = 0,1056$

(f)

$$
\begin{aligned}
z_{0,99} = 2,33 &= \frac{0,225 - 0,20}{0,16/n} \\
n &= \frac{2,33^2}{(0,225 - 0,20)^2} 0,16 = 1386
\end{aligned}
$$

Solução Computacional:

```
> sP <- sqrt(.2*.8/400)
> (p2.a <- pnorm(85/400, m=0.20, sd=sP, low=F))

[1] 0,266

> (p2.b <- pnorm(100/400, m=0.20, sd=sP, low=F))

[1] 0,00621

> (q2.c <- qnorm(c(0.10, 0.90), m=0.20, sd=sP))

[1] 0,1744 0,2256

> (p2.d <- pnorm(0.225, m=0.20, sd=sP, low=F))

[1] 0,1056

> (n.e <- ceiling(qnorm(0.99)^2 * (0.2*.8)/0.025^2))

[1] 1386
```

---

2. O tempo de processamento (em segundos) de certo tipo de requisição a um servidor tem distribuição aproximadamente normal. Foram anotados os seguintes tempos em uma amostra de oito requisições: 15,2; 16,0; 14,5; 13,1; 15,4; 17,0; 16,2 e 14,5. Obtenha intervalos de confiança (90%) para a média e para a variância dos tempos de processamento.

   Solução:
   IC para média: $\bar{x} \pm t_{\alpha/2} s/\sqrt{n}$
   IC para variância: $\left(\frac{(n-1)s^2}{\chi^2_{0,95}}; \frac{(n-1)s^2}{\chi^2_{0,05}} \right)$

```
> tp <- c(15.2, 16.0, 14.5, 13.1, 15.4, 17.0, 16.2, 14.5)
> xbar <- mean(tp)
> xbar

[1] 15,24

> s2 <- var(tp)
> s2

[1] 1,471

> n <- length(tp)
> t.test(tp, conf.level=0.90)$conf.int

[1] 14,43 16,05
attr(,"conf.level")
[1] 0,9

> c((n-1)*s2/qchisq(0.95, df=n-1), (n-1)*s2/qchisq(0.05, df=n-1))

[1] 0,7321 4,7518
```

---

3. Uma companhia de cigarros anuncia que o índice médio de nicotina dos seus cigarros apresenta-se abaixo de 26 mg. Um laboratório realiza 6 análises, obtendo as seguintes quantidades de nicotina: 27, 24, 21, 25, 26, 22. Sabe-se que o índice de nicotina se distribui normalmente. Pode-se aceitar, ao nível de 10% de significância, a afirmação do fabricante? E ao nível de 5%?

Solução:

$$H_0 : \mu \geq 26 \text{ vs } H_a : \mu < 26$$
$$\alpha = 0,10 \text{ ou } \alpha = 0,05$$
$$t_c = \frac{\overline{x} - \mu_0}{S/\sqrt{n}} = \frac{24,17 - 26}{2,317/\sqrt{6}} = -1,938$$

(a) Para $\alpha = 0,10$

$$RC : \{t < -1,48\}$$
$$Rejeita - seH0...$$

(b) Para $\alpha = 0,05$

$$RC : \{t < -2,02\}$$
$$NaoRejeita - seH0...$$

Soluções computacionais (linguagem R):

```
> y <- c(27,  24,  21,  25,  26,  22)
> (m.y <- mean(y))

[1] 24,17

> (v.y <- var(y))

[1] 5,367

> (RC0.10 <- qt(0.10, df=length(y)-1))

[1] -1,476

> (RC0.05 <- qt(0.05, df=length(y)-1))

[1] -2,015

> (tc <- (mean(y)-26)/sqrt(var(y)/length(y)))

[1] -1,938

> ## solução "direta" usando função apropriada do R
> (T1 <- t.test(y, alternative="less", mu=26, conf=0.90))

One Sample t-test

data:  y
t = -1,9, df = 5, p-value = 0,06
alternative hypothesis: true mean is less than 26
90 percent confidence interval:
  -Inf 25,56
sample estimates:
mean of x
    24,17

> (T2 <- t.test(y, alternative="less", mu=26, conf=0.95))
```

```
One Sample t-test

data:  y
t = -1,9, df = 5, p-value = 0,06
alternative hypothesis: true mean is less than 26
95 percent confidence interval:
  -Inf 26,07
sample estimates:
mean of x
    24,17
```

---

4. Discuta os conceitos a seguir no contexto de inferência estatística.

   (a) Estimador *viciado* e *não viciado*.

   (b) Método de estimação.

   (c) Incerteza na estimação.

   (d) Distribuição amostral.