Módulo básico - Tópicos de Estatística e Probabilidade
ONS – 2006/2007 - Profa. Mônica Barros
LISTA DE EXERCÍCIOS # 3

PROBLEMA 1

Toma-se duas amostras de engenheiros formados há 5 anos por duas Universidades e faz-se uma pesquisa salarial, cujos resultados estão a seguir.

	Universidade 1	Universidade 2
Tamanho da amostra	20	12
Salário médio na amostra (por ano)	R$ 50000	R$ 60000
Desvio padrão dos salários na amostra	R$ 8000	R$ 12000

a) Encontre um intervalo de confiança 95% para mu1 – mu2 onde mu1 é o salário médio real (e desconhecido) dos engenheiros formados na universidade 1 e mu2 é a mesma coisa para os engenheiros formados na universidade 2.
b) Com 95% de probabilidade existe a chance de mu1 e mu2 serem iguais? Por que?
c) Encontre um intervalo de confiança 95% para a razão das variâncias $\frac{\sigma_2^2}{\sigma_1^2}$. As variâncias das duas amostras podem ser iguais com este grau de confiança?

Solução

a) Seja: $R = \sqrt{\left(\frac{1}{n} + \frac{1}{m}\right)\left(\frac{(m-1)S_1^2 + (n-1)S_2^2}{n+m-2}\right)}$

O IC 100(1-α)% para mu1 – mu2 (supondo que as variâncias das duas amostras são iguais) é:
$$\left((\overline{X} - \overline{Y}) - bR;\ (\overline{X} - \overline{Y}) + bR\right)$$

Onde b é obtido da distribuição com m + n -2 graus de liberdade tal que Pr(T < b) = 1 – α/2.

Neste caso, a distribuição apropriada é uma t com 30 graus de liberdade e b é tal que Pr(T < b) = 97.5%, e então, da tabela: b = 2.042.

$$R = \sqrt{\left(\frac{1}{12} + \frac{1}{20}\right)\left(\frac{(19)(8000)^2 + (11)(12000)^2}{30}\right)} = 3527.6684$$

b.R = 7203.50
$$\left((\overline{X} - \overline{Y}) - bR;\ (\overline{X} - \overline{Y}) + bR\right) = (-10000 - 7203.50, -10000 + 7205.50) = (-17203.50, -2796.50)$$

b) NÃO, POIS O INTERVALO DE CONFIANÇA NÃO INCLUI ZERO.

c) Sabemos que:

$$\frac{(m-1)S_1^2}{\sigma_1^2} \sim \chi^2_{m-1} \quad \text{e} \quad \frac{(n-1)S_2^2}{\sigma_2^2} \sim \chi^2_{n-1}$$

são independentes, e assim a razão destas variáveis (divididas por seus graus de liberdade) tem distribuição F(m-1, n-1).

$$F = \frac{\dfrac{(m-1)S_1^2}{\sigma_1^2}/(m-1)}{\dfrac{(n-1)S_2^2}{\sigma_2^2}/(n-1)} = \frac{\dfrac{S_1^2}{\sigma_1^2}}{\dfrac{S_2^2}{\sigma_2^2}} = \frac{S_1^2}{S_2^2} \cdot \frac{\sigma_2^2}{\sigma_1^2}$$

NOTA: A FUNÇÃO A SER USADA DO EXCEL, É, NA VERSÃO EM INGLÊS, FINV – EM PORTUGUÊS DEVE SER INVF. OS PARÂMETROS SÃO: FINV(PROB; DF1; DF2) ONDE PROB É A PROBABILIDADE DE ESTAR ACIMA DO PONTO, DF1 E DF2 SÃO RESPECTIVAMENTE OS GRAUS DE LIBERDADE DO NUMERADOR E DO DENOMINADOR.

Logo, deve-se achar a e b da densidade F(m-1, n-1) tais que Pr(a < F < b) = 100(1-α) e Pr(F < a) = α/2 = Pr(F > b).

Então, b é encontrado diretamente no Excel: b = FINV(0.025; 19; 11) = 3.2428
E a é tal que: Pr(F < a) = 2.5% e então Pr(F > a) = 97.5% e assim a = FINV(0.975;19; 11) = 0.3617.

$$\Pr\left(0.3617 < \frac{(8000)^2 \sigma_2^2}{(12000)^2 \sigma_1^2} < 3.2428\right) = 0.95$$

E então o IC é:

$$\left(0.3617\left(\frac{12}{8}\right)^2, 3.2428\left(\frac{12}{8}\right)^2\right) = (0.8139, 7.2964)$$ Note que o intervalo inclui 1, e portanto as variâncias das duas amostras podem ser iguais.

Problema 2

Você ganhou uma caixa com 30 bombons. O peso médio dos bombons da caixa foi 22g, e o desvio padrão dos pesos observados foi 0.5g. Suponha que o peso dos bombons segue uma distribuição Normal.

Encontre um intervalo de confiança 90% para a variância do peso dos bombons.

Solução

O IC tem a forma:

$$\left(\frac{(n-1)S^2}{b}, \frac{(n-1)S^2}{a} \right)$$ onde a e b são obtidos da distribuição Qui-quadrado com n-1 graus de liberdade tais que: Pr(a < X < b) = 100(1-α) e Pr(X < a) = α/2 = Pr(X > b) onde X é a variável Qui-quadrado.

NOTA: A FUNÇÃO CHIINV DO EXCEL (EM INGLÊS) NOS DÁ OS PONTOS PERCENTUAIS DA QUI-QUADRADO. A SINTAXE É CHIINV(PROB; DF) ONDE PROB É A PROBABILIDADE DE ESTAR ACIMA DO PONTO, E DF SÃO OS GRAUS DE LIBERDADE.

POR EXEMPLO, A SEGUIR, a = CHIINV(0.95;29) = 17.708 E b = CHIINV(0.05;29)
Neste caso, o IC é 90%, e a densidade e Qui-quadrado com 29 graus de liberdade. Da tabela temos:
b = 42.557 e a = 17.708.

O IC é:

$$\left(\frac{29(0.5)^2}{42.557}, \frac{29(0.5)^2}{17.708} \right) = (0.1704, 0.4094)$$

PROBLEMA 3

As notas num certo exame padronizado têm média 450 e desvio padrão 50. Uma nota acima de 480 é considerada muito boa. Uma pessoa consegue entrar um MBA de prestígio se ela obtém acima de 480 neste exame.

Numa certa sala onde o exame foi aplicado, 25 pessoas fizeram o teste. A nota média destas pessoas foi 490. Isso é estranho? Você acha que deve haver algum tipo de investigação para tentar detector alguma fraude? Dica – use o Teorema Central do Limite.

SOLUÇÃO

Seja X a nota no teste. Pelo enunciado do problema, X tem média 450 e desvio padrão 50. Logo, a média amostral das notas das 25 pessoas daquela sala (supondo que as notas são iid) é uma variável com média 450 e variância $(50)^2/25$. Então, pelo teorema central do limite:

$$\frac{\overline{X} - 450}{\sqrt{\frac{(50)^2}{25}}} = \frac{\overline{X} - 450}{\frac{50}{5}} = \frac{\overline{X} - 450}{10}$$ é aproximadamente N(0,1).

$$\Pr(\overline{X} > 490) = \Pr\left(\frac{\overline{X} - 450}{10} > \frac{490 - 450}{10}\right) = 1 - \Phi(4) = 0$$ Logo, é absolutamente improvável que a nota média daquelas 25 pessoas tenha superado 490, um indício claro de fraude no teste, que deverá ser investigado.

PROBLEMA 4

Toma-se duas amostras de marcas de pneus para testar a sua durabilidade média (em milhares de km).Os resultados estão a seguir.

	Marca 1	Marca 2
Tamanho da amostra	15	13
Durabilidade média do pneu (em mil km)	50	45
Desvio padrão da durabilidade média (em mil km)	9	13

a) Encontre um intervalo de confiança 95% para mu1 – mu2 onde mu1 é a durabilidade média dos pneus da marca 1 e mu2 é a mesma coisa para a marca 2.
b) Com 95% de probabilidade existe a chance de mu1 e mu2 serem iguais? Por que?

Solução

a) Seja: $$R = \sqrt{\left(\frac{1}{n} + \frac{1}{m}\right)\left(\frac{(m-1)S_1^2 + (n-1)S_2^2}{n+m-2}\right)}$$

O IC 100(1-α)% para mu1 – mu2 (supondo que as variâncias das duas amostras são iguais) é:

$$\left((\overline{X} - \overline{Y}) - bR;\ (\overline{X} - \overline{Y}) + bR\right)$$

Onde b é obtido da distribuição com m + n -2 graus de liberdade tal que Pr(T < b) = 1 – α/2.

Neste caso, a distribuição apropriada é uma t com 26 graus de liberdade e b é tal que Pr(T < b) = 97.5%, e então, da tabela: b = 2.056.

$$R = \sqrt{\left(\frac{1}{15} + \frac{1}{13}\right)\left(\frac{14(9)^2 + 12(13)^2}{26}\right)} = \sqrt{17.4627} = 4.1788$$

b.R = 8.5917

$$\left((\overline{X} - \overline{Y}) - bR, (\overline{X} - \overline{Y}) - bR\right) = (50 - 45 - 8.5917, 50 - 45 + 8.5917) = (-3.5917, +13.5917)$$

b) **SIM,** POIS O INTERVALO DE CONFIANÇA **INCLUI ZERO**.

PROBLEMA 5

Uma linha de produção produz pacotes de café cujo peso nominal é 1 kg. Toma-se uma amostra de 25 pacotes e observa-se que o peso médio na amostra é 985g e o desvio padrão dos pesos é 60g. Encontre um intervalo de confiança 95% para a variância dos pesos dos pacotes supondo que os pesos têm distribuição Normal.

Solução

O IC tem a forma:

$\left(\frac{(n-1)S^2}{b}, \frac{(n-1)S^2}{a}\right)$ onde a e b são obtidos da distribuição Qui-quadrado com n-1 graus de liberdade tais que: Pr(a < X < b) = 100(1-α) e Pr(X < a) = α/2 = Pr(X > b) onde X é a variável Qui-quadrado.

Neste caso, o IC é 95%, e a densidade e Qui-quadrado com 24 graus de liberdade. Da tabela Qu-quadrado temos:
b = 39.364 e a = 12.401.

O IC é:

$$\left(\frac{24(60)^2}{39.364}, \frac{24(60)^2}{12.401}\right) = (2194.90, 6967.18)$$

Problema 6

A mesma prova foi aplicada em duas turmas, com os resultados descritos a seguir.

	Turma A	Turma B
Média	64	69
Desvio Padrão	15	20
Número de Alunos	40	32

Encontre um intervalo de confiança 95% para a diferença das médias μ_B - μ_A. Use uma aproximação Normal, já que o número de graus de liberdade da distribuição t é grande.

Solução

O resultado fundamental aqui é:

$$T = \frac{Z}{\sqrt{\frac{V}{n+m-2}}} = \frac{\overline{X} - \overline{Y} - (\mu_1 - \mu_2)}{\sqrt{\left(\frac{1}{n} + \frac{1}{m}\right)\left(\frac{(m-1)S_1^2 + (n-1)S_2^2}{n+m-2}\right)}} \sim t_{n+m-2}$$

Aqui podemos supor que a distribuição é Normal, pois se trata de uma t com 40 + 32 – 2 = 70 graus de liberdade.

A variância amostral combinada é:

$$S_p^2 = \left(\frac{(m-1)S_1^2 + (n-1)S_2^2}{n+m-2} \right) = \frac{39(15)^2 + 31(20)^2}{70} = 302.50 \Rightarrow S_p = \sqrt{302.5} = 17.3925$$

O IC 95% para μ_B - μ_A é:

$$\overline{Y} - \overline{X} \pm 1.96 \sqrt{\left(\frac{1}{n} + \frac{1}{m} \right) \left(\frac{(m-1)S_1^2 + (n-1)S_2^2}{n+m-2} \right)} = (69-64) \pm 1.96 \sqrt{\left(\frac{1}{40} + \frac{1}{32} \right) 302.5} = 5 \pm 8.0850 =$$
$$= [-3.085, 13.085]$$