MQI 2003 – Estatística para Metrologia – semestre 2008.01

Lista 4

Profa. Mônica Barros

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).

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

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 <u>a</u> e <u>b</u> 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 <u>a</u> 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

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:

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

$$= [-3.085, 13.085]$$

PROBLEMA 3

A proporção de cura para uma certa doença através do tratamento padrão é 45%. Dr. Exxottericc propõe um novo tratamento baseado em ervas milagrosas e afirma que tratou 50 pacientes, tendo curado 25 deles. O que você acha, o tratamento do Dr. Exxottericc é melhor? Construa um Intervalo de Confiança 95% para responder esta pergunta.

Solução

O IC para a proporção Binomial é aproximado e baseado na distribuição Normal. Sua forma é: $\left(\hat{p} - z.\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}, \hat{p} + z.\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\right)$

No caso de um IC 95%, z = 1.96. Aqui, n = 50 e a proporção de curas observadas na amostra é 0.50. Logo, o IC é:

$$\left(0.50 - 1.96\sqrt{\left(\frac{(0.5)(0.5)}{50}\right)}, 0.5 + 1.96\sqrt{\left(\frac{(0.5)(0.5)}{50}\right)}\right) = (0.361, 0.639)$$

O tratamento pode ou não ser melhor que o outro– veja como o IC é amplo, e inclui valores de p abaixo de 45%.

PROBLEMA 4

Você trabalha no departamento de marketing de uma empresa farmacêutica que acabou de lançar um remédio que auxilia na redução de peso. Numa amostra de 36 mulheres que tomaram o remédio, a perda de peso média em três meses foi de 8.4 kg, com desvio padrão 2.6 kg.

Para comparação, tomou-se uma outra amostra, de 26 mulheres, que só fizeram dieta. Nesta segunda amostra, a perda de peso média em três meses foi de 6.2 kg, com desvio padrão 2.4 kg.

Encontre um IC 95% para a diferença de perda de peso entre os dois grupos. Você pode concluir, com probabilidade 95% que o remédio qua a sua empresa está lançando é eficaz para a dieta?

Nota: suponha que as perdas de peso são variáveis Normais.

Solução

Para começar a resolver o problema, precisamos decidir qual IC usar. A forma do IC desejado é:

$$\left(\overline{X} - \overline{Y} - bR, \overline{X} - \overline{Y} + bR\right) \text{ onde } R = \sqrt{\left(\frac{1}{n} + \frac{1}{m}\right)\left(\frac{(m-1)S_1^2 + (n-1)S_2^2}{n+m-2}\right)}$$

O ponto percentual da densidade t é obtido da densidade com n+m-2 = 36+26-2 = 60 graus de liberdade. Na tabela da densidade t, este ponto para um IC 95% é: 2.00 (apenas para comparação, o mesmo percentil para a densidade N(0,1) é 1.96). O IC desejado é:

$$\left(8.4 - 6.2 \mp 2.R\right) = 2.2 \mp 2.R \quad \text{onde } R = \sqrt{\left(\frac{1}{36} + \frac{1}{26}\right)\left(\frac{35(2.6)^2 + (25)(2.4)^2}{60}\right)} = \sqrt{0.4202} = 0.6482$$

$$\left(2.2 - 2 * 0.6482, 2.2 + 2 * 0.6482\right) = \left(0.9036, \quad 3.4964\right)$$

Note que o IC só inclui valores positivos, portanto a média de perda de peso com o remédio é superior à média sem o remédio, e o tratamento é eficaz.

PROBLEMA 5

Você é o Secretário de Fazenda de uma cidade e acabou de tomar posse. No passado, você notou que, a partir de uma amostra das maiores 256 empresas na cidade, a arrecadação de tributos por empresa era, em média, R$ 1 milhão, com desvio padrão (amostral) R$ 200 mil.

Você decide implementar uma grande mudança na secretaria de Fazenda, intensificando a fiscalização. Logo no primeiro mês da sua gestão, o valor médio arrecadado nas 256 maiores empresas passou para R$ 1.2 milhões.

Construa IC 95% e 99% para verificar se as medidas propostas por você resultaram num aumento significativo da arrecadação.

Solução

Para começar a resolver o problema, precisamos decidir qual IC usar. Este é um caso típico de grandes amostras – note que, em nenhum lugar do enunciado diz-se que a distribuição é Normal, mas fala-se do tamanho da amostra (256), bastante grande.

A forma do IC desejado é:

$$\left(\overline{X} - z.\frac{S}{\sqrt{n}}, \overline{X} + z.\frac{S}{\sqrt{n}}\right) = \left(1200000 - z\frac{200000}{\sqrt{256}}, 1200000 + z\frac{200000}{\sqrt{256}}\right) =$$
$$= \left(1{,}200{,}000 - z(12500), 1{,}200{,}000 + z(12500)\right)$$

Para um IC 95%: z = 1.96 e então o IC torna-se (1,175,500 , 1,224,499)

Para um IC 99%: z = 2.576 e então o IC torna-se (1,167,802 , 1,232,197)

Em ambos os casos, a arrecadação anterior (R$ 1 milhão) está fora do IC, indicando que as medidas tomadas para aumentar a arrecadação surtiram efeito.

PROBLEMA 6

Um país tem que decidir uma questão importante num referendo. Você trabalha numa empresa de pesquisas, e faz uma pesquisa de "boca de urna" para medir a proporção de eleitores que dizem ter votado a favor da questão no referendo, tentando inferir sobre a real proporção de eleitores na população favoráveis à tal proposta. A sua amostra tem 625 eleitores, dos quais 300 foram favoráveis à questão. Encontre um IC 90% aproximado para o valor verdadeiro de p, a proporção de eleitores na população favoráveis à proposta.

Solução

A forma do IC aproximado para p é:

$$\left(\hat{p} - z_{1-\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}, \hat{p} + z_{1-\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\right)$$

Onde $z_{1-\alpha/2}$ obtido da densidade N(0,1) tal que Pr(Z < $z_{1-\alpha/2}$) = 1-α/2

Neste caso:

$\hat{p} = \dfrac{300}{625} = 0.48$ e para um IC 90%, z = 1.645. Logo, o IC 90% para p é:

$$\left(0.48 - 1.645\sqrt{\frac{0.48(0.52)}{625}}, 0.48 + 1.645\sqrt{\frac{0.48(0.52)}{625}}\right) = (0.48 - 0{,}0329, 0.48 + 0{,}0329) = (0.4471,\ \ 0.5129)$$

Então, com base na nossa "boca de urna" não é possível saber se a proposta será ou não aprovada, pois o IC inclui valores abaixo e acima de 50%.

PROBLEMA 7

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:

$\left(\dfrac{(n-1)S^2}{b}, \dfrac{(n-1)S^2}{a}\right)$ onde a e b são obtidos da distribuição Qui-quadrado com n-1 graus de liberdade tais que: Pr(a < X < b) = 100(1-) e Pr(X < a) = /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 temos:
b = 39.364 e a = 12.401.

O IC é:

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

Ou seja, os desvios padrão estarão no intervalo aproximado: (46.9, 83.5)