Resolução Provinha II - MAE 1512

Ex.1
i) Vamos calcular a média amostral $\bar{X}$, com $n = 9$.

$$\bar{X} = \frac{4.9 + 7.0 + 8.1 + 4.5 + 5.6 + 6.8 + 7.2 + 5.7 + 6.2}{9} =$$

$$\bar{X} = \frac{56}{9} = 6.222$$

Assumo o valor de $\hat{\mu} = \bar{X}$, logo $\hat{\mu} = 6.222$.

Vamos calcular o intervalo de confiança IC, sabendo que o desvio padrão é $\sigma = 2$ e o tamanho da amostra é $n = 9$. Usaremos $\gamma = 0.95$, ou seja teremos uma chance de 95% de encontrar a média "real" entre os pontos. Logo, o intervalo de confiança fica:

$$\bar{X} \pm Z_\gamma \frac{\sigma}{\sqrt{n}}$$

Sabendo que $Z_\gamma = 1.96$ pela tabela da normal padrão, $N\sim(0,1)$, temos que:

$$= 6.222 \pm 1.96 \frac{2}{\sqrt{9}}$$

Ponto Inferior do IC:

$$= 6.222 + 1.96 \frac{2}{\sqrt{9}} = 4.915$$

Ponto Superior do IC:

$$= 6.222 - 1.96 \frac{2}{\sqrt{9}} = 7.529$$

Assim o IC (intervalo de confiança), é:

$$IC = \{4.915, 7.529\}$$

ii) Queremos um n para metade do intervalo em i). Sabemos que o comprimento do intervalo em i) é de duas vezes o desvio padrão dividido pela raiz do tamanho da amostra, vezes o Z_γ. Assim, temos:
Comprimento total do intervalo:

$$= 2\left(Z_\gamma \frac{\sigma}{\sqrt{n}}\right) = 2\frac{2}{\sqrt{9}}1.96 = 2.613$$

Outra maneira de calucular o tamanho do intervalo é subtraindo o ponto superior($psup$) pelo inferior($pinf$):

$$psup - pinf = 7.529 - 4.915 = 2.613$$

Queremos a metade da metade deste intervalo:

$$\frac{2.613}{4} = 0.653$$

Aplicando esse resultado para encontrarmos n, temos:

$$\begin{aligned}
Z_\gamma \frac{\sigma}{\sqrt{n}} &= 0.653 && (1)\\
Z_\gamma \frac{\sigma}{0.653} &= \sqrt{n} && (2)\\
\sqrt{n} &= Z_\gamma \frac{\sigma}{0.653} && (3)\\
n &= Z_\gamma^2 \frac{\sigma^2}{0.653^2} && (4)\\
n &= 1.96^2 \frac{2^2}{0.653^2} && (5)\\
n &= 36 && (6)
\end{aligned}$$

Assim, nossa nova amostra deve conter $n = 36$ elementos.

iii) Queremos que o erro cometido seja de:

$$P(|\bar{X} - \mu| \leq 0.1) = 0.90$$

Para isso temos que achar um n tal que:

$$P\left(\frac{|\bar{X}-\mu|}{\sigma/\sqrt{n}} \leq Z_\gamma\right) = 0.90$$

Temos que Z_γ na normal $N{\sim}(0,1)$ com 0.90 de probabilidade é de:

$$Z_\gamma = 1.64$$

Assim temos que:

$$P\left(|\bar{X}-\mu| \leq \frac{1.64}{\sqrt{n}}\sigma\right) = 0.90$$

E pela condição inicial e sabendo que $\sigma = 2$:

$$\frac{1.64}{\sqrt{n}}2 = 0.1$$

$$n = \frac{1.64^2}{0.1^2}2^2$$

$$n = 1075.8$$

Uma amostra de $n = 1075$ elementos seria necessária para um erro de 0.1 entre a média amostral $\bar{X}$ e média polucional μ, com 90% de probabilidade.

Ex.2

Com desvio padrão (σ) desconhecido, devemos levar em conta o desvio padrão amostral, S. Para calcularmos S, antes devemos calcular S^2, a variância amostral, definida abaixo:

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(xi-\bar{X})^2$$

Em i) $\bar{X} = 6.222$, calculando a variância acima obtemos $S^2 = 1.35$ e o intervalo de confianą é dado por:

$$\bar{X} \pm t_\gamma\sqrt{\frac{S^2}{n}}$$

Como a variância é amostral não sabemos exatamente como se comporta sua distribuição, assim vamos aproximar usando a tabela $t - Student$, com $(n-1)$ graus de liberdade e $\gamma = 0.95$, obtendo:

$$t_\gamma = 2.306$$

Assim temos que:

$$6.222 \pm 2.306\sqrt{\frac{1.35}{9}}$$

Ponto Inferior do IC:

$$= 6.222 + 2.306\sqrt{\frac{1.35}{9}} = 5.329$$

Ponto Superior do IC:

$$= 6.222 - 2.306\sqrt{\frac{1.35}{9}} = 7.115$$

Assim o IC com Variância desconhecida, é:

$$IC = \{5.329, 7.115\}$$

Ex.3

Temos que a proporção na cidade A é de:

$$\hat{P}_A = \frac{X_A}{n_A} = \frac{180}{400} = 0.450$$

E a proporção na cidade B é de:

$$\hat{P}_B = \frac{X_B}{n_B} = \frac{350}{600} = 0.583$$

Sabemos que o interlavo de confiança definida para proporção é:

$$\hat{P}_A - \hat{P}_B \pm Z_\gamma\sqrt{\frac{\hat{P}_A(1-\hat{P}_A)}{n_A} + \frac{\hat{P}_B(1-\hat{P}_B)}{n_B}}$$

Então, para $Z_\gamma = 1.96$:

$$0.450 - 0.583 \pm 1.96\sqrt{\frac{0.45(0.55)}{400} + \frac{0.583(0.417)}{600}}$$

Montamos o IC para proporção:
Ponto Inferior do IC:

$$-0.133 - 0.0627 = -0.196$$

Ponto Superior do IC:

$$-0.133 + 0.0627 = -0.070$$

Assim o IC com Variância desconhecida, é:

$$IC = \{-0.196, -0.070\}$$

Ex.4
Queremos saber se existe independcia entre a utilização do seguro e o sexo do segurado, logo temos duas hipóteses:

$$H_0 = \text{existe independência}$$
$$H_1 = \text{não existe independência}$$

Vamos testar H_0, temos uma tabela de contingência de 2×2 com GL graus de liberdade com $l = 2$ linhas e $c = 2$ colunas:

$$GL = (l-1)(c-1) = 1$$

Nossa tabela observa está abaixo:
Sendo U utliza e $\bar{U}$ não utiliza.

	U	$\bar{U}$	
H	100	900	1000
M	150	850	1000
	250	1750	2000

Vamos fazer a tabela com o valor esperado,sendo i quantidade de linhas e j colunas:

$$e_{ij} = \frac{totallinha_i \times totalcoluna_j}{total_{geral}}$$

A tabela do valor esperado fica:

	U	$\bar{U}$	
H	125	875	1000
M	125	875	1000
	250	1750	2000

Usamos o teste Qui-quadrado χ^2, para verficar a hipótese:

$$\chi^2 = \sum_{i=1}^{n} \sum_{j=1}^{k} \frac{(o_{ij} - e_{ij})^2}{e_{ij}}$$

$$\chi^2 = \frac{(100 - 125)^2}{100} + \frac{(900 - 875)^2}{875} + \frac{(150 - 125)^2}{125} + \frac{(850 - 875)^2}{875} = 11.43$$

Grau de Liberdade $(L - 1)(C - 1) = (2 - 1)(2 - 1) = 1$.

Com $\alpha = 0.05$, temos que o valor na reta é de 3.841.

Como $3.841 \leq 11.43$. O valor encontrado está na região de rejeição, logo rejeitamos a hipótese nula, e concluímos que existe dependeência.