ICMC – USP – SME 320

Estatística I – 2ª PROVA – 2007/1 – 29/6/2007

S O L U Ç Â O

1. Denotamos por X_1 e X_2 a viscosidade (em u.v.) antes e após a mudança, respectivamente. De acordo com o enunciado, temos duas amostras independentes, a distribuição destas variáveis é normal e a variância (desconhecida) é a mesma (σ^2). Representando a média das variáveis por μ_1 e μ_2, a pergunta formulada pode ser respondida testando $\text{H}_0 : \mu_1 - \mu_2 = 0$ *versus* $\text{H}_1 : \mu_1 - \mu_2 < 0$. A hipótese alternativa é unilateral à esquerda (região crítica à esquerda). A partir dos dados coletados temos:

$$\text{Antes da mudança:} \quad n_1 = 15, \quad \overline{X}_1 = \frac{1}{n_1}\sum_{i=1}^{n_1} X_{1i} = 750,2$$

$$\text{e} \quad S_1^2 = \frac{1}{n_1 - 1}\sum_{i=1}^{n_1}(X_{1i} - \overline{X}_1)^2 = 365,9;$$

$$\text{Após a mudança:} \quad n_2 = 8, \quad \overline{X}_2 = \frac{1}{n_2}\sum_{i=1}^{n_2} X_{2i} = 756,9$$

$$\text{e} \quad S_2^2 = \frac{1}{n_2 - 1}\sum_{i=1}^{n_2}(X_{2i} - \overline{X}_2)^2 = 453,0,$$

de forma que a variância (σ^2) é estimada como

$$S_p^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2} = 394,9,$$

e a estatística utilizada no teste é

$$t_0 = \frac{\overline{X}_1 - \overline{X}_2}{S_p\sqrt{1/n_1 + 1/n_2}} = -0,767.$$

Com 5% de significância e $n_1 + n_2 - 2 = 21$ graus de liberdade, consultamos a tabela da distribuição t de Student e obtemos $t_c = -1,720$. Se $t_0 < t_c$ devemos rejeitar H_0. Como $t_0 > t_c$, não rejeitamos H_0. Logo, com um nível de significância de 5%, concluímos que o objetivo de aumentar a viscosidade média não foi atingido.

2. Denotamos por X_1 e X_2 a taxa de ataque (em 10^{-3} pol/min) conforme as soluções 1 e 2, respectivamente. Representamos a média das variáveis por μ_1 e μ_2. De acordo com o enunciado, temos duas amostras independentes e as variâncias (σ_1^2 e σ_2^2) são diferentes e desconhecidas.

   (a) A pergunta formulada pode ser respondida testando $\text{H}_0 : \mu_1 - \mu_2 = 0$ *versus* $\text{H}_1 : \mu_1 - \mu_2 \neq 0$. A hipótese alternativa é bilateral (região crítica bilateral). A partir dos dados coletados temos:

$$\text{Solução 1:} \quad n_1 = 10, \quad \overline{X}_1 = \frac{1}{n_1}\sum_{i=1}^{n_1} X_{1i} = 9,97$$

$$\text{e} \quad S_1^2 = \frac{1}{n_1 - 1}\sum_{i=1}^{n_1}(X_{1i} - \overline{X}_1)^2 = 0,178;$$

$$\text{Solução 2:} \quad n_2 = 10, \quad \overline{X}_2 = \frac{1}{n_2}\sum_{i=1}^{n_2} X_{2i} = 10,4$$

$$\text{e} \quad S_2^2 = \frac{1}{n_2 - 1}\sum_{i=1}^{n_2}(X_{2i} - \overline{X}_2)^2 = 0,0533.$$

De posse das estimativas das variâncias, os graus de liberdade são dados por

$$g = \frac{(S_1^2/n_1 + S_2^2/n_2)^2}{\frac{(S_1^2/n_1)^2}{n_1-1} + \frac{(S_2^2/n_2)^2}{n_2-1}} \cong 14.$$

A estatística utilizada no teste é

$$t_0 = \frac{\overline{X}_1 - \overline{X}_2}{\sqrt{S_1^2/n_1 + S_2^2/n_2}} = -2,828.$$

Com 1% de significância e $g = 14$ graus de liberdade, consultamos a tabela da distribuição t de Student e obtemos $t_c = 2,977$. Se $|t_0| > t_c$ devemos rejeitar H_0. Como $|t_0| < t_c$, não rejeitamos H_0. Logo, com um nível de significância de 1%, concluímos que a taxa média de ataque é a mesma para as duas soluções.

(b) Um intervalo de confiança de 99% para $\mu_1 - \mu_2$ é dado por

$$\overline{X}_1 - \overline{X}_2 \mp t_c\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}} = -0,43 \mp 0,453 = [-0,883; 0,023].$$

Nesta solução supomos que sejam válidas as condições que permitem aplicar o teorema central do limite, justificando os resultados utilizados.

3. Segundo o enunciado, trata-se de um problema sobre uma proporção (precisamente, de itens defeituosos produzidos), denotada por θ. A pergunta formulada pode ser respondida testando $\mathrm{H}_0 : \theta = \theta_0$ *versus* $\mathrm{H}_1 : \theta > \theta_0$, sendo que $\theta_0 = 1/100$. A hipótese alternativa é unilateral à direita (região crítica à direita). De acordo com os dados coletados ($n = 800$ e 10 componentes defeituosos), a proporção amostral é $\widehat{\theta} = 10/800 = 0,0125$. A estatística utilizada no teste é

$$z_0 = \frac{\widehat{\theta} - \theta_0}{\sqrt{\theta_0(1 - \theta_0)/n}} = 0,711.$$

Com 5% de significância, consultamos a tabela da distribuição normal padrão e obtemos $z_c = 1,65$. Se $z_0 > z_c$ devemos rejeitar H_0. Como $z_0 < z_c$, não rejeitamos H_0. Logo, com um nível de significância de 5%, os dados não confirmam a suspeita de que a fração de itens defeituosos produzidos excede 1/100.

Nesta solução recorremos ao teorema central do limite.

4. A variável aleatória quantidade de xarope liberada pela máquina (em cm^3) é representada por X, com média igual a μ. De uma amostra com $n = 25$ garrafas envasadas foram obtidas média amostral $\overline{X} = 32,4$ e desvio padrão amostral $s = 0,74$.

   (a) A pergunta formulada pode ser respondida testando $\text{H}_0 : \mu = \mu_0$ *versus* $\text{H}_1 : \mu \neq \mu_0$, com $\mu_0 = 29,5$. A hipótese alternativa é bilateral (região crítica bilateral). A estatística utilizada no teste é

   $$t_0 = \frac{\sqrt{n}\,(\overline{X} - \mu_0)}{s} = 19,6.$$

   Com qualquer nível de significância usual (1% ou 5%, por exemplo), rejeitamos H_0, pois $|t_0|$ ultrapassa os respectivos valores t_c da tabela da distribuição t de Student com $n - 1 = 24$ graus de liberdade. Logo, os dados sugerem que a quantidade média envasada não é 29,5 cm^3.

   (b) Pretende-se detectar uma diferença $\delta = 0,3$ cm^3 (em relação a 29,5 cm^3) com poder $1-\beta = 0,9$. Na solução abaixo o desvio padrão de X (σ, desconhecido) é substituído pelo desvio padrão amostral (s). A expressão do tamanho amostral (n) para atender ao que se pretende é dada por

   $$n = \frac{(z_{\alpha/2} + z_\beta)^2\,\sigma^2}{\delta^2}.$$

   Adotamos um nível de significância $\alpha = 5\%$, de forma que $z_{\alpha/2} = 1,96$. Com $\beta = 0,1$ obtemos $z_\beta = 1,28$ e $n = 64$. Portanto, a amostra coletada não é adequada, sendo necessárias $64 - 25 = 39$ garrafas adicionais.

   Nesta solução supomos que sejam válidas as condições que permitem aplicar o teorema central do limite, justificando os resultados utilizados.

5. Denotamos por L_A, L_B e L_C as larguras das peças A, B e C (em mm), respectivamente. Conforme o enunciado, L_A, L_B e L_C são variáveis aleatórias com distribuições normal$(10; 0,1^2)$, normal$(2; 0,05^2)$ e normal$(2; 0,05^2)$.

   (a) O espaçamento entre as peças B e C (representado por d) é dado por $d = L_A - L_B - L_C$. A média do espaçamento é obtida por meio de $\mu_d = \text{E}(d) = \text{E}(L_A - L_B - L_C) = \text{E}(L_A) - \text{E}(L_B) - \text{E}(L_C) = 6$.
   Supondo independência entre as três variáveis, temos a variância de d obtida de $\sigma_d^2 = \text{var}(d) = \text{var}(L_A - L_B - L_C) = \text{var}(L_A) + (-1)^2\,\text{var}(L_B) + (-1)^2\,\text{var}(L_C) = 0,015$, de modo que o desvio padrão do espaçamento é igual a $\sigma_d = \sqrt{\sigma_d^2} = 0,1225$.

   (b) Usando o fato de que a distribuição de d também é normal, a probabilidade pedida vale

   $$\text{P}(d < 5,9) = \text{P}\left(\frac{d - \mu_d}{\sigma_d} < \frac{5,9 - \mu_d}{\sigma_d}\right) = \text{P}\left(Z < -0,82\right),$$

   sendo que a distribuição de $Z = (d - \mu_d)/\sigma_d$ é normal padrão. Consultando a tabela desta distribuição, obtemos $\text{P}(d < 5,9) = 0,2061$.