Instituto Superior de Economia e Gestão

Universidade de Lisboa

Licenciaturas em Economia e em Finanças

Estatística II – Teste intercalar – 27/03/2017 – Duração: 1 hora

Nome:____ Número:___________

Notas: Certifique-se de que o seu telemóvel está desligado. Se não estiver, é motivo suficiente para anulação da prova. As perguntas de escolha múltipla valem 1 valor; respostas erradas são penalizadas em 0.25. Caso nada seja dito em contrário, utilize um nível de significância de 5% nos testes de hipóteses que efetuar. Fundamente e formalize devidamente todas as respostas. Pode usar a última página para continuar qualquer questão.

Espaço reservado para classificações

Tópicos de resolução

1. Na adega *Bacchus* procede-se a medições regulares do conteúdo das garrafas de vinho. Numa amostra de 16 garrafas escolhidas aleatoriamente obteve-se uma quantidade média de 0.745 litros por garrafa. Admita que a quantidade de vinho (litros) por garrafa tem distribuição normal de média μ e desvio padrão igual a 0.02.

   a) Construa um intervalo de confiança a 95% para a quantidade média de vinho por garrafa e interprete o resultado obtido. [**1.5**]

Seja X a variável aleatória que representa a quantidade de vinho, em litros, por garrafa:
$X \sim N(\mu, \sigma^2)$, com $\sigma = 0.02$.

Amostra: $n = 16$, $\bar{x} = 0.745$.

O intervalo de confiança a 95% para μ é

$$\left(\bar{x} - z_{0.025} \times \frac{\sigma}{\sqrt{n}}, \bar{x} + z_{0.025} \times \frac{\sigma}{\sqrt{n}}\right)$$

obtido com base na variável fulcral

$$Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim N(0,1)$$

Como

$$P(-1.96 < Z < 1.96) = 0.95$$

Obtém-se:

$$\left(0.745 - 1.96 \times \frac{\sigma}{\sqrt{n}}, 0.745 + 1.96 \times \frac{\sigma}{\sqrt{n}}\right) = (0.745 - 1.96 \times 0.005\,, 0.745 + 1.96 \times 0.005)$$

O intervalo de confiança para μ é $(0.7352, 0.7548)$.
Pode afirmar-se, com uma confiança de 95% que a quantidade média de vinho por garrafa da adega Bacchus se situa entre 0.7352 e 0.7548 litros.

b) Determine, justificando, a estimativa da máxima verosimilhança para a probabilidade de uma garrafa da adega *Bacchus* ter menos que 0.75 litros. [**1.5**]

Pretende-se a estimativa da máxima verosimilhança para $P(X < 0.75)$, ou seja:

$$\hat{P}(X < 0.75) \ = \ \hat{P}(\frac{X-\mu}{\sigma} < \frac{0.75-\mu}{\sigma})$$

Utilizando a propriedade de invariância dos estimadores da máxima verosimilhança e sabendo que $\hat{\mu} = \bar{X}$, e $Z = \frac{X-\mu}{\sigma} \sim N(0,1)$, vem

$$\hat{P}(X < 0.75) = \ \hat{P}\left(Z < \frac{0.75-\mu}{\sigma}\right) = P\left(Z < \frac{0.75-\hat{\mu}}{0.02}\right) = P\left(Z < \frac{0.75-\bar{x}}{0.02}\right)$$

Como $\bar{x} = 0.745$, obtém-se

$$\hat{P}(X < 0.75) = \ P\left(Z < \frac{0.75-0.745}{0.02}\right) = P(Z < 0.25) = 0.5987$$

A estimativa de máxima verosimilhança para a probabilidade de uma garrafa da adega *Bacchus* ter menos de 0.75 litros é igual a 0.5987.

c) Admita que, com base na mesma amostra, realizou o teste H_0: $\mu = 0.75$ contra H_1: $\mu < 0.75$ ao nível de 5%. O valor da função potência quando $\mu = 0.73$, $\beta(0.73)$ obtém-se calculando:

■ $P(\bar{X} < 0.75 - 1.645 \times 0.005 | \mu = 0.73)$.

□ $P(\bar{X} > 0.75 + 1.645 \times 0.005 | \mu = 0.73)$.

□ $P(Z < 1.645)$, onde Z é a estatística de teste.

□ $P(\bar{X} > 0.73 + 1.645 \times 0.005 | \mu = 0.75)$.

2. Seja *X* uma variável aleatória com função densidade $f(x|\theta)$, de média $E(X) = \frac{3}{2}\theta + 1$. O estimador para o parâmetro θ pelo método dos momentos, obtido com base numa amostra casual de dimensão *n*, é:

□ $\tilde{\theta} = \bar{X}$.

■ $\tilde{\theta} = \frac{2\bar{X}-2}{3}$.

□ $\tilde{\theta} = \frac{2\bar{X}}{3} - 1$.

□ não se pode obter porque a função densidade não é dada.

3. A Associação *ATP* afirma que pelo menos 90% dos residentes em Lisboa avaliam como positivo o aumento do fluxo de turistas na cidade. Num inquérito a 500 indivíduos escolhidos aleatoriamente em Lisboa, 440 deram uma resposta positiva sobre o aumento do fluxo de turistas.

   a) Teste, ao nível de 5%, a afirmação da Associação *ATP*. [**2.0**]

Seja $X = 1$, se o residente em Lisboa é favorável ao aumento do número de turistas em Lisboa; e seja $X = 0$ caso contrário; $X \sim B(1, \theta)$, com θ proporção dos indivíduos favoráveis ao aumento do número de turistas em Lisboa.

Amostra: $n = 500$, $\sum_{i=1}^{500} x_i = 440$, $\bar{x} = \frac{440}{500} = 0.88$.

O teste pretendido é $H_0: \theta \geq 0.90$ contra $H_1: \theta < 0.90$

A estatística de teste é

$$Z = \frac{\bar{X} - \theta}{\sqrt{\theta(1-\theta)/n}} \dot{\sim} N(0,1)$$

e, sob H_0,

$$Z = \frac{\bar{X} - 0.90}{\sqrt{0.9(1-0.9)/n}} \dot{\sim} N(0,1)$$

A região crítica ao nível de 5% é $W_{5\%} = \{z: z < -1.645\}$. Calculando o valor observado, obtém-se:

$$z_{obs} = \frac{\bar{x} - 0.90}{\sqrt{0.9(1-0.9)/n}} = \frac{0.88 - 0.90}{\sqrt{(0.9 \times 0.1)/500}} = -1.49$$

Como $z_{obs} \notin W_{5\%}$, não se rejeita H_0 a 5%: não há evidência contra a afirmação da Associação *ATP*: é de admitir que pelo menos 90% dos residentes em Lisboa avaliam como positivo o aumento do fluxo de turistas na cidade.

   b) Num inquérito realizado na cidade do Porto a 480 indivíduos sobre a mesma questão, a Associação *ATP* recolheu 408 respostas favoráveis ao aumento do número de turistas na sua cidade. Para testar se existem diferenças nas duas cidades no que respeita à proporção de opiniões favoráveis foi obtido o intervalo de confiança aproximado a 95% para a diferença de proporções: $(-0.0128, 0.0728)$. Então:

- [ ] nada se pode concluir porque a dimensão das amostras nas duas cidades não é igual.
- [ ] há evidência de diferenças significativas nas duas proporções, ao nível de 5%, porque as médias amostrais nas duas cidades são diferentes.
- [x] não há evidência de proporções distintas nas duas cidades, ao nível de 5%.
- [ ] nenhuma das opções anteriores é verdadeira.

4. Seja $(X_1, X_2, \dots, X_n), n \geq 2$ uma amostra casual de uma população de Bernoulli de parâmetro θ. Sejam

$$T_1 = \bar{X} \text{ e } T_2 = \frac{\sum_{i=1}^{n-1} X_i}{n-1}$$

dois estimadores para o parâmetro θ. Prove que ambos são estimadores centrados do parâmetro θ e calcule a eficiência relativa de T_2 em relação a T_1. [**2.0**]

Como $X \sim B(1,\theta)$,tem-se $E(X) = \theta$ e $var(X) = \theta(1-\theta)$ e sendo a amostra casual, $X_i \sim iid\ B(1,\theta)$ e portanto $E(X_i) = \theta$ e $var(X_i) = \theta(1-\theta), i = 1,2, \dots, n$.

- $T_i\,, i = 1, 2$ é um estimador centrado do parâmetro θ se $E(T_i) = \theta$.

Como $E(T_1) = E(\bar{X}) = E(X) = \theta$, T_1 é um estimador centrado de θ.

Calculando o valor esperado de T_2 obtém-se:

$$E(T_2) = E\left[\frac{\sum_{i=1}^{n-1} X_i}{n-1}\right] = \frac{1}{n-1} E[X_1 + X_2 + \cdots + X_{n-1}] = \frac{1}{n-1}(E[X_1] + E[X_2] + \cdots + E[X_{n-1}])$$

Como $E(X_i) = \theta$, vem:

$$E(T_2) = \frac{1}{n-1} \times (n-1)E(X) = E(X) = \theta$$

Os dois estimadores são estimadores centrados de θ.

- Como ambos são centrados, o estimador mais eficiente é o de menor variância.

$$var(T_1) = var(\bar{X}) = \frac{var(X)}{n} = \frac{\theta(1-\theta)}{n}$$

$$var(T_2) = var\left(\frac{\sum_{i=1}^{n-1} X_i}{n-1}\right) = \frac{1}{(n-1)^2} \times var\left(\sum_{i=1}^{n-1} X_i\right)$$

Como as variáveis são independentes, obtém-se:

$$var(T_2) = \frac{1}{(n-1)^2} \times \sum_{i=1}^{n-1} var(X_i) = \frac{1}{(n-1)^2} \times \sum_{i=1}^{n-1} var(X) = \frac{(n-1)\theta(1-\theta)}{(n-1)^2}$$

$$var(T_2) = \frac{\theta(1-\theta)}{n-1}$$

Assim,

$$\frac{var(T_2)}{var(T_1)} = \frac{\theta(1-\theta}{n-1} \times \frac{n}{\theta(1-\theta} = \frac{n}{n-1} > 1$$

e conclui-se que T_1é mais eficiente que T_2 porque $var(T_1) < var(T_2)$.