Licenciatura em Engenharia Informática

Matcp

Exame Modelo

Obs: Justifique todos os cálculos que efetuar

Nome: Nº:

1. Um artigo é comercializado a um preço de venda de 15€, sendo o seu custo 10€.
O abastecimento desse artigo é feito no início de cada semana e torna-se irrecuperável se não for vendido na própria semana. No início de cada semana o stock é reposto com 3 unidades. A tabela seguinte apresenta a função de probabilidade para o número de artigos procurados semanalmente

x	1	2	3	4	5
$f(x)$	0.2	0.3	0.15	0.25	0.1

(a) Qual a probabilidade de haver falhas de stock?

X-v.a.que representa o nº de artigos procurados semanalmente

Falhas de stock, significa que a procura é superior a 3 unidades

Probabilidade pedida

$P(X > 3) = P(X = 4) + P(X = 5) = 0.35$

(b) Se numa semana foram vendidos menos de 5 artigos, qual a probabilidade de se terem vendido no mínimo 2? Probabilidade pedida

$P(X \geq 2|X < 5) = \frac{P(X<5 \wedge X \geq 2)}{P(X<5)} = \frac{P(2 \leq X \leq 4)}{P(X \leq 4)} = \frac{0.7}{0.9} = \frac{7}{9} \approx 0.7778$

(c) Determine a função de probabilidade do lucro semanal obtido.

Y- v.a. que representa o lucro semanal que depende do número de itens vendidos

1 artigo vendido , lucro= $15 - 30 = -15$ euros

2 artigos vendidos, lucro $= 30 - 30 = 0$ euros

3 artigos vendidos (só são vendidos no máximo 3 artigos), lucro= $45 - 30 = 15$euros

x	1	2	3	4	5
y	-15	0	15	15	15

A função de probabilidade do lucro é

y	-15	0	15
$g(y)$	0.2	0.3	0.5

(d) Qual o lucro médio semanal e o desvio-padrão do lucro? $E(Y) = -15 \times 0.2 + 0 \times 0.3 + 15 \times 0.2 = 4.5$ euros

$\sigma(Y) = \sqrt{(-15 - 4.5)^2 \times 0.2 + (0 - 4.5)^2 \times 0.3 + (15 - 4.5)^2 \times 0.5} \approx 11.72$euros

(e) Admita agora que não há stock inicial e que número de artigos procurados semanalmente segue uma distribuição de Poisson de média 3 artigos.

i. Qual a probabilidade da procura semanal exceder a procura média?

X - v.a. que representa o número de artigos procurados semanalmente
$X \sim P_o(3)$

Probabilidade pedida

$$P(X > 3) = 1 - P(X \leq 3) = 1 - \sum_{i=1}^{3} \frac{e^{-i} \times 3^i}{i!} \left\{ \begin{array}{l} \text{Tabela Poisson} \\ \text{Função distribuição} \\ \lambda = 3 \\ x = 3 \end{array} \right. = 1 - 0.6472 = 0.3528$$

ii. Num mês (4 semanas), qual a probabilidade da procura exceder 9 unidades e não ultrapassar 15?

Y - v.a. que representa o número de artigos procurados num mês

$\lambda_y = 4 \times 3 = 12$ - n$^{\text{o}}$ médio de artigos por mês

Probabilidade pedida

$P(9 < Y \leq 15) = P(Y \leq 15) - P(Y \leq 9)$

$$= 0.8444 \left\{ \begin{array}{l} \text{Tabela Poisson} \\ \text{Função distribuição} \\ \lambda = 12 \\ x = 15 \end{array} \right. - 0.2424 \left\{ \begin{array}{l} \text{Tabela Poisson} \\ \text{Função distribuição} \\ \lambda = 12 \\ x = 9 \end{array} \right. = 0.6020$$

2. Após um grande número de registos concluiu-se que a extensão (em km) dos serviços das empresas de distribuição de encomendas, é uma variável aleatória de média 20 e desvio-padrão 9. Foi registada uma amostra aleatória das extensões de 34 serviços.

(a) Qual a probabilidade da média da amostra ser inferior a 19km?

Seja X_i - v.a. que representa a extensão (em km) do serviço i de distribuição de encomendas.

Seja $\overline{X} = \frac{1}{34} \sum_{i=1}^{34} X_i$- v.a que representa a média das extensões de 34 serviços.

Como $n > 30$, pelo Teorema do Limite Central, tem-se que

$\overline{X} \sim N\left(20, \frac{9^2}{34}\right)$

Probabilidade pedida: $P\left(\overline{X} < 19\right) = P\left(Z \leq \frac{19-20}{\sqrt{\frac{9^2}{34}}}\right) = \Phi(-0.65) = 0.2585$

Obs: Pode ser utilizada a função NORMCDF.

(b) Qual a probabilidade dos 34 serviços excederem um total de 600km?

Seja $T = \sum_{i=1}^{34} X_i$- v.a que representa a extensão total de 34 serviços.

Como $n > 30$, pelo Teorema do Limite Central, tem-se que

$T \sim N\left(34 \times 20, 34 \times 9^2\right)$

Probabilidade pedida:

$P(T > 600) = 1 - P(T \leq 600) =$

$= 1 - P\left(Z \leq \frac{600-680}{\sqrt{34\times 9^2}}\right) = 1 - \Phi(-1.52) = 1 - 0.0637 = 0.9363$

3. Numa indústria extratora de pedras preciosas, sabe-se que o salário médio de um trabalhador é 5 euros por hora, com variancia 0,36 euros^2 por hora. Considerando que os salários dos trabalhadores desta indústria são variáveis aleatórias independentes e identicamente distribuídas, calcule:

   (a) a probabilidade de, num grupo de 47 trabalhadores desta indústria, se observar um salário médio inferior ou igual a 4.90 euros por hora.

   Seja X_i - v.a. que representa o salário (em euros) do trabalhador i
   Seja $\overline{X} = \frac{1}{47}\sum\limits_{i=1}^{47} X_i$- v.a que representa o salário médio 47 trabalhadores.

   Como $n > 30$, pelo Teorema do Limite Central, tem-se que

   $\overline{X} \sim N\left(5, \frac{0.36}{47}\right)$

   Probabilidade pedida: $P\left(\overline{X} \leq 4.9\right) = P\left(Z \leq \frac{4.9-5}{\sqrt{\frac{0.36}{47}}}\right) = \Phi(-1.14) = 0.1266$

   (b) a probabilidade de um orçamento de 405 euros por hora não ser suficiente para pagar a um grupo de 81 trabalhadores desta indústria.

   Seja $T = \sum\limits_{i=1}^{81} X_i$- v.a que representa o pagamento total de 81 trabalhadores.

   Como $n > 30$, pelo Teorema do Limite Central, tem-se que
   $T \sim N(81 \times 5, 81 \times 0.36)$

   Probabilidade pedida:

   $P(T > 405) = 1 - P(T \leq 405) =$

   $= 1 - P\left(Z \leq \frac{405-405}{\sqrt{81\times 0.36}}\right) = 0.5000$

4. Uma empresa de prestação de serviços ambientais publicita que o tempo de resposta ao cliente é, em média, de 11,5 horas. Um cliente habitual, desconfiando que tempo médio de resposta é superior, verificou aleatória e independentemente 38 serviços à sua empresa, obtendo um tempo médio de 11,54h e um desvio-padrão de 1,2h.

   (a) Construa um intervalo de confiança a 95% para o verdadeiro tempo de resposta médio por parte da empresa.

X_i- v.a. que representa o tempo de resposta do serviço i, em horas.

Estamos em presença de uma "população"desconhecida e variância desconhecida.

Pretende-se determinar um I.C. a 95% para a média $E()X_i) = \mu$

Variável aleatória fulcral

$\overline{X}$ - v.a. que representa a média do tempo de resposta quando considerada uma amostra aleatória de 38 serviços.

$$\overline{X} = \frac{1}{38}\sum_{i=1}^{38} X_i$$

Como $n > 30$, pelo Teorema do Limite Central, tem-se que :

$$\overline{X} = \frac{1}{38}\sum_{i=1}^{38} X_i \sim N\left(\mu, \frac{\sigma^2}{38}\right) \Leftrightarrow Z = \frac{\overline{X} - \mu}{\sigma/\sqrt{38}} \sim N(0,1)$$

Grau de confiança:

$$1 - \alpha = 0.95 \Leftrightarrow \alpha = 0.05 \Rightarrow z_c = z_{0.975} = \Phi^{-1}(1 - 0.025) = \Phi^{-1}(.975) = 1.96$$

Intervalo de confiança para μ:

$$IC_{(1-\alpha)*100\%}(\mu) = \left]\overline{X} - z_c\frac{\sigma}{\sqrt{n}}, \overline{X} + z_c\frac{\sigma}{\sqrt{n}}\right[$$

$$IC_{(1-\alpha)*100\%}(\mu) = \left]\overline{X} - z_c\frac{1.2}{\sqrt{38}}, \overline{X} + z_c\frac{1.2}{\sqrt{38}}\right[\quad \text{onde } \sigma^2 \approx 1.2^2$$

Depois de realizada a amostragem, considerando os dados da amostra.

$n = 38$, $\overline{x} = 11.54$ e $s = 1.2$,

$$IC_{(1-\alpha)*100\%}(\mu) = \left]\overline{x} - z_c\frac{\sigma}{\sqrt{n}}, \overline{x} + z_c\frac{\sigma}{\sqrt{n}}\right[$$

obtemos,

$$IC_{95\%}(\mu) = \left]11.54 - 1.96 \times \frac{1.2}{\sqrt{38}}, 11.54 + 1.96 \times \frac{1.2}{\sqrt{38}}\right[=]11.15, 11.93[$$

.

(b) Com base nos resultados obtidos, haverá publicidade enganosa?

Resposta: O IC obtido contem o tempo de resposta ao cliente (11,5 horas). Não se pode concluir que haja publicidade enganosa.

5. Foi largamente divulgado na imprensa das últimas semanas um ciberataque à escala mundial de um tipo de *malware*, conhecido como *ransomware*, que infeta os sistemas informáticos bloqueando o acesso ao sistema e aos ficheiros para u m posterior pedido de resgate. Num determinado país A, e com o objetivo de analisar a extensão do ataque, recolheu-se uma amostra aleatória

de 2500 computadores tendo-se verificado que 627 tinham sido infetados. O mesmo estudo foi realizado no país B, onde em 1300 computadores analisados verificou-se que 343 estavam infetados.
Determine um intervalo de confiança a 95% para a diferença de proporções de computadores infetados entre os dois países. Podemos concluir que a infeção teve o mesmo impacto em ambos os países? Justifique.

Seja X_A - v.a. que representa o número de computadores infetados no país A .
Seja X_B - v.a. que representa o número de computadores infetados no país B .

Amostras: $n_A = 2500$, $\hat{p}_A = \frac{627}{2500} = 0.2508$, $n_B = 1300$, $\hat{p}_B = \frac{343}{1300} = 0.2638$

Seja $\hat{P}_A$ - v.a. que representa a proporção de computadores infetados no país A quando considerada uma amostra aleatória de 2500 clientes.
Seja $\hat{P}_B$ - v.a. que representa a proporção de computadores infetados no país B quando considerada uma amostra aleatória de 1300 clientes.
$\hat{P}_A \sim N\left(p_A, \frac{p_A q_A}{n_A}\right)$ sendo p_A a proporção de computadores infetados no país A.

$\hat{P}_B \sim N\left(p_B, \frac{p_A q_A}{n_A}\right)$ sendo p_B a proporção de computadores infetados no país B.

Para compararmos as proporções de computadores infetados temos de considerar a variável aleatória representativa da sua diferença: $\hat{P}_A - \hat{P}_B \sim N\left(p_A - p_B, \sqrt{\frac{p_A q_A}{n_A} + \frac{p_B q_B}{n_B}}\right)$
Pretendemos estimar $p_A - p_B$ através de um intervalo de confiança a 95%

Para um nível de confiança de 95% temos $\alpha = 0.05$ pelo que $1 - \frac{\alpha}{2} = 0.975$. Logo, $z_c = \Phi^{-1}\left(1 - \frac{\alpha}{2}\right) = \Phi^{-1}(0.975) = 1.96$

$$IC_{95\%}(p_A - p_B) = \left](\hat{p}_A - \hat{p}_B) - z_{0.975}\sqrt{\frac{\hat{p}_A\hat{q}_A}{n_A} + \frac{\hat{p}_B\hat{q}_B}{n_B}}, (\hat{p}_A - \hat{p}_B) + z_{0.975}\sqrt{\frac{\hat{p}_A\hat{q}_A}{n_A} + \frac{\hat{p}_B\hat{q}_B}{n_B}}\right[$$

Logo,

$$IC_{96\%}(p_A - p_B) = \left]0.2508 - 0.2638 - 1.96\sqrt{\frac{0.1879}{2500} + \frac{0.1942}{1300}}, 0.2508 - 0.2638 + 1.96\sqrt{\frac{0.1879}{2500} + \frac{0.1942}{1300}}\right[$$
$$\approx \left] -0.1686, -0.0914\right[$$

R:Como o IC é ?negativo?, conclui-se com 95% de confiança que o impacto da infeção não é igual nos dois países (o IC não contém o valor 0).

6. Realizou-se um inquérito a 125 alunos da universidade A selecionados aleatoriamente e verificou-se que 25 utilizam o smartphone durante o seu tempo de estudo. Realizou-se igual inquérito na universidade B e em 110 alunos inquiridos aleatoriamente, 37 afirmaram usar o smartphone. A um nível de significância de 1%, poder-se-á afirmar que existe diferença significativa entre as proporções de alunos que usam o smartphone durante o seu estudo, em ambas as universidades.

X_1- v.a. que representa o nº de alunos da universidade A que utilizam o smartphone durante o seu tempo de estudo, em 125 selecionados aleatoriamente.

$\hat{P}_1$-v.a. que representa a proporção de alunos da universidade A que utilizam o smartphone durante o seu tempo de estudo, quando considerada uma amostra aleatória de 125.

$\hat{P}_1 \sim N\left(p_1, \frac{p_1 q_1}{n1}\right)$

X_2- v.a. que representa o nº de alunos da universidade B que utilizam o smartphone durante o seu tempo de estudo, em 110 selecionados aleatoriamente.

$\hat{P}_2$-v.a. que representa a proporção de de alunos da universidade B que utilizam o smartphone durante o seu tempo de estudo, quando considerada uma amostra aleatória de 110.

$\hat{P}_2 \sim N\left(p_2, \frac{p_2 q_2}{n2}\right)$

Para compararmos as proporções de alunos das duas universidade temos de considerar a variável aleatória representativa da sua diferença: $\hat{P}_1 - \hat{P}_2 \sim N\left(p_1 - p_2, \sqrt{\frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}}\right)$ segundo o TAN.

Pretendemos aplicar um teste de hipóteses para a diferença de proporções com um nível de significância de 1

Teste de hipóteses bilateral para a diferença de proporções

$H_0 : p_1 = p_2 \quad v.s. \quad H_1 : p_1 - p_2 \neq 0$

Amostras: $n_1 = 125$, $\hat{p}_1 = 25/125 = 0.20$, $n_2 = 110$, $\hat{p}_2 = 37/110 = 0.34$,

Estatística de teste: $Z = \frac{(\hat{P}_1 - \hat{P}_2) - 0}{\sqrt{\hat{p}\hat{q}(\frac{1}{n_1} + \frac{1}{n_2})}} \sim N(0,1)$

Região crítica

Como o nível de significância é $\alpha = 0.01$,logo $P(Z > z_c) = \frac{0.01}{2} \Leftrightarrow P(Z \leq z_c) = 0.995 \Leftrightarrow z_c = 2.58$

$R_c =]-\infty, -2.58[\cup]2.58, +\infty[$

Valor observado da estatística de teste:

$\hat{p} = \frac{125 \times 0.20 + 110 \times 0.34}{125 + 110} = 0.2655 \quad \text{e} \quad \hat{q} = 0.7345$

$z_{obs} = \frac{0.20 - 0.34}{\sqrt{0.2655 \times 0.7345 \times (\frac{1}{125} + \frac{1}{110})}} = -2.43$

Decisão:

$-2.58 < z_{obs} = -2.43 < 2.58$, $z_{obs} \notin R_c$

Logo, não se deve rejeitar H_0. Ao nível de significância de 1%, não existe evidência estatística de que a proporção de alunos que utilizam o smartphone durante o seu tempo de estudos é significativamente diferente nas duas universidades.