ESTATÍSTICA I – 2.º Ano/Gestão – 1.º Semestre – Época Normal – Duração: 2 horas

1.ª Parte – Teórica **N.º de Exame: RESOLUÇÃO** **09.01.2015**

Este exame é composto por duas partes. **Esta é a 1ª Parte — Teórica (Cotação: 8 valores). As respostas às questões de escolha múltipla são efectuadas na correspondente folha de resposta anexa, que será recolhida 40 minutos após o início da prova. As outras questões devem ser respondidas no próprio enunciado, no espaço disponibilizado para o efeito.** No decorrer da prova não serão prestados quaisquer esclarecimentos. BOA SORTE!

Nome: ______________________________ N.º: __________

Cada um dos cinco grupos de perguntas de escolha múltipla vale 10 pontos (1 valor). Cada resposta certa vale 2,5; cada resposta errada vale -2,5. A classificação de cada grupo variará entre um mínimo de zero e um máximo de 10 pontos.

Classifique as seguintes afirmações como verdadeiras (**V**) ou falsas (**F**).

1. Sejam A, B e C acontecimentos de um espaço de resultados Ω.

a) Se A e B são mutuamente exclusivos, então também são independentes (**F**)

b) Se $P(A-B) = P(A)$, então $P(A \cup B) = P(A) + P(B)$ (**V**)

c) Se $P(B) > 0$, então $P(A \cap B \cap C) = P(B) \times P(A \cap C \mid B)$ (**V**)

d) Se $A \subset B$ e $P(A) > 0$, então tem-se que $P(B) > 0$ (**V**)

2. Considere uma v.a. X e a respectiva função de distribuição $F(x)$.

a) Sejam a e b números reais, $a < b$. Se X é contínua, então $P(a < X < b) = F(b) - F(a)$ (**V**)

b) Se X for discreta, quaisquer que sejam $h > 0$ e $x \in \mathbb{R}$ tem-se que $F(x+h) - F(x) > 0$ (**F**)

c) Sendo X contínua com função densidade $f(x)$, tem-se $F(x) = \int_{-\infty}^{x} f(u)\, du$ para todo o x (**V**)

d) Se $F(x)$ tiver pontos de descontinuidade, então X é discreta (**F**)

3. Sejam X e Y variáveis aleatórias com função distribuição conjunta $F(x, y) = P(X \leq x, Y \leq y)$, representando $F_X(x)$ e $F_Y(y)$ as respectivas funções de distribuição marginais.

a) Se o coeficiente de correlação é nulo, então X e Y são independentes (**F**)

b) $F(x, y) = F_X(x) \times F_Y(y)$ para todo o (x, y) (**F**)

c) Para todo o x, tem-se $F_X(x) = F(x, +\infty)$ (**V**)

d) $\text{Var}(X+Y) \geq \text{Var}(X) + \text{Var}(Y)$ (**F**)

4. Seja X uma variável aleatória e sejam $E(X) = \mu$ e $\text{Var}(X) = \sigma^2 < \infty$.

a) Se $X \sim \text{N}(\mu, \sigma^2)$, então $W = (X-\mu)^2/\sigma^2 \sim \chi^2(1)$ (**V**)

b) Seja $X \sim \text{B}(n, \theta)$. Então $Y = \theta - X \sim \text{B}(n, 1-\theta)$ (**F**)

c) Se $X \sim \text{U}(a, a+1)$, onde a é um número real, então a probabilidade de X pertencer a qualquer intervalo contido em $(a, a+1)$ é igual ao comprimento desse intervalo (**V**)

d) Considere a variável $Z = (X-\mu)/\sigma$. Então $E[Z] = 0$ e $\text{Var}(Z) = 1$ (**V**)

5. Seja $X_1, \ldots, X_n$ uma amostra casual de tamanho $n > 2$ proveniente de uma população X possuindo média μ e variância σ^2 finita.

a) $P(X_1 \leq x_1, \ldots, X_n \leq x_n) = \sum_{i=1}^{n} P(X \leq x_i)$ (**F**)

b) Se X tem distribuição Normal, então $\bar{X}$ é uma estatística com distribuição Normal(**V**)

c) A correcção de continuidade só deve ser usada se a distribuição de X for Binomial ou Poisson (**F**)

d) O teorema do limite central determina a distribuição aproximada da estatística $\sum_{i=1}^{n} X_i$ para n suficientemente grande (**V**)

Responda às perguntas que se seguem no espaço disponibilizado para o efeito. Justifique cuidadosamente todos os passos. Cotação de cada pergunta: 15 pontos.

6. Mostre que se $X \sim \text{Ex}(\lambda)$, onde $\lambda > 0$ e $c > 0$, então $Y = cX$ é tal que $Y \sim \text{Ex}(\lambda/c)$.

RESPOSTA:

$$\begin{aligned} F_Y(y) =& P(Y \leq y) \\ =& P(cX \leq y) \\ =& P(X \leq y/c) \\ =& F_X(y/c) \\ =& 1 - \exp(-\lambda\ y/c) \\ =& 1 - \exp[-(\lambda/c)y],\ \ y > 0\ . \end{aligned}$$

7. Seja $X_1, \ldots, X_n$ uma amostra casual proveniente de uma população com função de distribuição $F(x)$. Mostre que para todo x se tem $P(X_1 > x, \ldots, X_n > x) = [1 - F(x)]^n$.

RESPOSTA:

$$\begin{aligned} P(X_1 > x, \ldots, X_n > x) =& P(X_1 > x)\ \cdots P(X_n > n) \quad \text{por independência} \\ =& [1 - F_{X_1}(x)] \cdots [1 - F_{X_n}(x)] \\ =& [1 - F_X(x)] \cdots [1 - F_X(x)] \quad X_i \text{ identicamente distribuído a } X \\ =& [1 - F_X(x)]^n, \quad x \in \mathbb{R}\ . \end{aligned}$$

FIM DA 1ª PARTE **VSFF**

Este exame é composto por duas partes. **Esta é a 2.ª Parte – Prática (Cotação: 12 valores). Esta parte é composta por 4 questões, cada uma na sua folha. As questões devem ser respondidas no espaço disponibilizado para o efeito.** No decorrer da prova não serão prestados quaisquer esclarecimentos. BOA SORTE!

Atenção: Nas perguntas com alternativas, uma resposta certa vale 10 pontos, uma resposta errada vale -2.5 pontos.

Cotação:

1.a)	b)	2.a)	b)	3.a)	b)	4.a)	b)
10	20	10	20	10	20	10	20

1. Determinada empresa produz garrafas de vidro usando 3 máquinas diferentes: M1, M2 e M3. A probabilidade de produzir uma garrafa defeituosa é 0.01 para as máquinas M1 e M2; 0.05 para a máquina M3.

a) Em 20 garrafas seleccionadas ao acaso da produção da máquina M3, qual a probabilidade de se encontrarem pelo menos 3 defeituosas?

☐ **i)** 0.0596 ☐ **ii)** 0.05 ☐ **iii)** 0.0159 ☒ **iv)** 0.0755

b) Num lote de 15000 garrafas, sabemos que 7000 foram produzidas pela máquina M3. Seleccionada ao acaso uma garrafa desse lote, verifica-se que é defeituosa. Qual a probabilidade de esta ter sido produzida pela máquina M3?

RESPOSTA 1.b) Definam-se os acontecimentos M_i = "garrafa é produzida pela máquina Mi", $i = 1, 2, 3$; D = "garrafa é defeituosa". Queremos calcular $P(M_3 \mid D)$. Tem-se que $P(D \mid M_1) = P(D \mid M_2) = 0.01$ e que $P(D \mid M_3) = 0.05$. Como $\{D_i\}$ forma uma partição do espaço de resultados tem-se pelo teorema de Bayes que

$$\begin{aligned} P(M_3 \mid D) &= \frac{P(D \mid M_3)\ P(M_3)}{P(D \mid M_1)\ P(M_1) + P(D \mid M_2)\ P(M_2) + P(D \mid M_3)\ P(M_3)} \\ &= \frac{0.05\ P(M_3)}{0.01[P(M_1) + P(M_2)] + 0.05\ P(M_3)}\ . \end{aligned}$$

Por outro lado, $P(M_3) = 7000/15000$ e $P(M_1) + P(M_2) = P(M_1 \cup M_2) = 8000/15000$. Substituindo na fórmula acima, otém-se $P(M_3 \mid D) \approx 0.814$

2. A função densidade conjunta do par (X, Y) é dada pela expressão

$$f(x, y) = \begin{cases} kxy & \text{se } 0 < x < 4, 1 < y < 5 \\ 0 & \text{outros } x \ . \end{cases}$$

a) O valor da constante k é dado por

☒ **i)** 1/96 ☐ **ii)** 1 ☐ **iii)** 1/100 ☐ **iv)** 1/20

b) Determine $E[X \mid Y = y]$ para $1 < y < 5$. O que se pode dizer acerca da independência entre X e Y tendo em conta apenas este resultado?

RESPOSTA 2.b) Por definição, para $1 < y < 5$, $E[X \mid Y = y] = \int_{-\infty}^{+\infty} x f_{X|Y=y}(x)\ dx$, onde

$$f_{X|Y=y}(x) = \frac{f(x, y)}{f_Y(y)} \ .$$

Como

$$\begin{aligned} f_Y(y) &= \int_{-\infty}^{+\infty} f(x, y)\ dx \\ &= \int_0^4 kxy\ dx \quad 1 < y < 5 \\ &= 8ky \quad 1 < y < 5 \ . \end{aligned}$$

Assim,

$$E[X \mid Y = y] = \int_0^4 x\ \frac{kxy}{8ky}\ dx = \frac{8}{3} \ .$$

Pelo facto de $E[X \mid Y = y]$ não depender de y podemos concluir que X é independente de Y em média, mas isto não implica que X e Y sejam independentes. Nada se pode concluir sobre a independência entre X e Y.

3. O número de pessoas que entra em determinada superfície comercial segue um processo de Poisson. Em média, numa hora entram 20 pessoas neste estabelecimento. Admita que o estabelecimento está aberto diariamente entre as 10:00 e as 18:00.

a) Determine a probabilidade de ter que se esperar mais de 5 minutos até entrar a primeira pessoa no estabelecimento.

☐ **i)** 0.0067 ☐ **ii)** 0.5332 ☒ **iii)** 0.1889 ☐ **iv)** 1/5

b) No final de cada dia, a equipa de marketing classifica cada uma das 8 horas de funcionamento diário como "produtiva" se durante essa hora entrarem no estabelecimento mais de 25 pessoas. Qual a probabilidade de em dois dias de funcionamento se verificarem no máximo 4 horas "produtivas"?

RESPOSTA 3.b) Seja $X =$ "número de horas produtivas entre as 16". Sabemos que $X \sim \mathrm{B}(16, \theta)$, onde, com $Y =$ "número de clientes que entra no estabelecimento numa hora", se tem $\theta = P(Y > 25) = 1 - F_Y(25)$. Como $Y \sim \mathrm{Po}(20)$, segue-se que $\theta = 0.1122$. Queremos calcular $P(X \leq 4)$, o que usando uma máquina de calcular resulta em 0.9731. Para usar as tabelas da binomial há que fazer $\theta \approx 0.10$, o que origina 0.9830 como resposta aproximada.

VSFF

4. Um empresa produz baterias cuja duração, em horas, segue uma distribuição de probabilidades cuja média é 15 horas e o desvio padrão é 2 horas.

a) Um lote de baterias é considerado defeituoso se a média da duração das baterias nesse lote for inferior a 14.8 horas. Qual a probabilidade (aproximada) de um lote de 100 baterias ser considerado defeituoso?

☐ **i)** 0 ☒ **ii)** 0.159 ☐ **iii)** 0.242 ☐ **iv)** 0.308

b) Um novo tipo de baterias foi desenvolvido pela mesma empresa, e tem uma duração (em horas) cuja média é 16 horas e o desvio padrão 2.5 horas. Comparando duas amostras (uma de cada tipo) de 5 baterias cada, qual a probabilidade de se observar que a média amostral relativa à bateria antiga é superior à média amostral relativa à bateria nova? Enuncie as suposições efectuadas.

RESPOSTA 4.b) Queremos calcular $P(\bar{X}_1 > \bar{X}_2) = P(\bar{X}_1 - \bar{X}_2 > 0)$, onde $\bar{X}_1$ corresponde à média da amostra de baterias antigas e $\bar{X}_2$ corresponde à média da amostra das baterias novas. Assumindo que

- duração de ambos os tipos de bateria segue uma distribuição normal;
- amostras são casuais e independentes e entre si,

temos que

$$\bar{X}_1 - \bar{X}_2 \sim \mathrm{N}(\mu_1 - \mu_2, \sigma_1^2/m + \sigma_2^2/n)$$

com $\mu_1 = 15$, $\mu_2 = 16$, $\sigma_1^2 = 2^2$, $\sigma_2^2 = 2.5^2$, $m = n = 5$. Assim, a probabilidade pedida é

$$P(\bar{X}_1 - \bar{X}_2 > 0) = 1 - \Phi\left(\frac{0 - (15 - 16)}{\sqrt{2^2/5 + 2.5^2/5}}\right) = 0.24245\ .$$

FIM