INSTITUTO POLITÉCNICO DE BRAGANÇA — Escola Superior de Tecnologia e de Gestão	Estatística Engenharia Informática 7 julho 2023 Prova Época Recurso

Consulta de Tabelas e Formulário autorizada. É obrigatória a devolução do enunciado, devidamente identificado. Utilização de telemóveis e internet não autorizada. Duração: 120 min. Justifique convenientemente todas as suas respostas.

Nome:____ Número:_______________

1. As empresas Caracol Expressos e a Sempressas Bus monopolizam o transporte de passageiros em autocarro entre o Porto e Bragança. As duas empresas competem relativamente à pontualidade que anunciam aos seus clientes. Recolheram-se alguns dados referentes à pontualidade das viagens realizadas pelas duas empresas entre estas duas cidades (apenas sentido Bragança), tendo-se verificado que a Caracol Expressos representa 65% do total das viagens. Os dados recolhidos mostram que as viagens da Caracol Expressos possuem atrasos de tempo em 66% das ocasiões enquanto esse valor é de 40% para a Sempressas Bus.

1.1 Apresente um diagrama em árvore que traduza a informação do enunciado, definindo os acontecimentos correspondentes.

1.2 Se selecionar aleatoriamente um autocarro que realizou a referida rota e pertence a uma destas duas empresas, qual a probabilidade de estar atrasado?

1.3 Se um autocarro proveniente de uma qualquer destas empresas, chegar atrasado a Bragança, qual é a probabilidade de pertencer à Caracol Expressos?

1.4 Se um autocarro chegar a Bragança e não estiver atrasado, qual é a probabilidade de pertencer à Sempressas Bus?

2. Três cartas são retiradas aleatoriamente e sem reposição de um baralho normal contendo 52 cartas. Considere que Y representa o número de copas obtido. Nestas condições, responda às seguintes questões.

2.1 A variável Y é discreta ou contínua?

2.2 Determine a função de probabilidade e a função distribuição de probabilidade da variável Y. Apresente um esboço gráfico de ambas as funções.

2.3 Determine os parâmetros de localização (média, mediana e moda) e de dispersão (variância), para a variável Y.

2.4 Qual a probabilidade de se obterem pelo menos 2 copas?

3. Admita que a quantidade de proteína no hambúrguer "Diabólico" comercializado por uma cadeia internacional de fast-food segue aproximadamente uma distribuição normal com um valor médio de 21.0 g e um desvio padrão de 0.8 g.

3.1 Se um hambúrguer for escolhido aleatoriamente de entre os "diabólicos", qual é a probabilidade de conter menos do que 20.0 g?

3.2 Qual é a distribuição da quantidade total de proteína em 2 hambúrgueres escolhidos aleatoriamente de entre os "diabólicos"?

3.3 Qual é a distribuição da quantidade média de proteína em 2 hambúrgueres escolhidos aleatoriamente de entre o referido tipo de hambúrgueres? Defina o seu valor médio e variância)

3.4 Se forem escolhidos aleatoriamente 2 hambúrgueres de entre os "diabólicos", qual é a probabilidade de a quantidade média de proteína ser pelo menos 20.5 g?

4. Pretende-se calcular um intervalo de confiança a 99% para a proporção de economistas que discorda com a subida das taxas de juros bancários como forma de conter a subida da inflação.

4.1 Numa sondagem de opinião envolvendo 400 economistas, 150 manifestaram concordância com a medida económica referida. Nestas condições, determine o intervalo de confiança para a proporção de economistas que concorda com esta medida.

4.2 Qual o tamanho mínimo da amostra de economistas a inquirir de modo que o intervalo a calcular tenha uma amplitude inferior a 5 pontos percentuais.

Questão	1.1	1.2	1.3	1.4	2.1	2.2	2.3	2.4	3.1	3.2	3.3	3.4	4.1	4.2
Cotação(%)	8	8	8	6	6	8	6	6	8	8	6	6	8	8

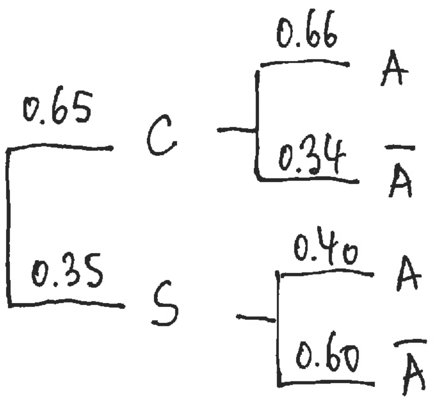
1

1.1) Acontecimentos:

C: O transporte é efetuado pela Caracol Expresso

S: O transporte é efetuado pela Sempressas Bus

A: A viagem é efetuada com atraso na chegada



1.2) $P(A) = P(A \cap C) + P(A \cap S) = P(A|C) \times P(C) + P(A|S) \times P(S)$

$= 0.66 \times 0.65 + 0.40 \times 0.35 = 0.569$

1.3) $P(C|A) = \dfrac{P(C \cap A)}{P(A)} = \dfrac{P(A|C) \times P(C)}{P(A)} = \dfrac{0.66 \times 0.65}{0.569} = 0.754$

1.4) $P(S|\bar{A}) = \dfrac{P(S \cap \bar{A})}{P(\bar{A})} = \dfrac{P(\bar{A}|S) \times P(S)}{1 - P(A)} = \dfrac{0.60 \times 0.35}{1 - 0.569} = 0.487$

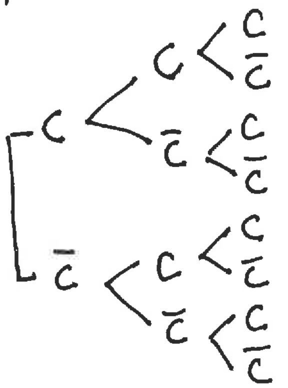
2

2.1) A variável Y é uma variável discreta já que toma apenas valores inteiros: $y = \{0, 1, 2, 3\}$

2.2) Função de probabilidade: $p(y) = P(Y = y)$

Função distribuição de probabilidade: $F(y) = P(Y \leq y)$

52 cartas: 13 cartas/naipe (39 não copas + 13 Copas)



(sugestão)

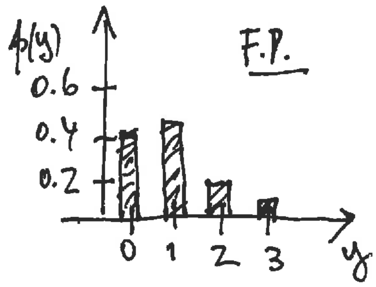
$p(0) = \frac{39}{52} \times \frac{38}{51} \times \frac{37}{50} = 0.4135$

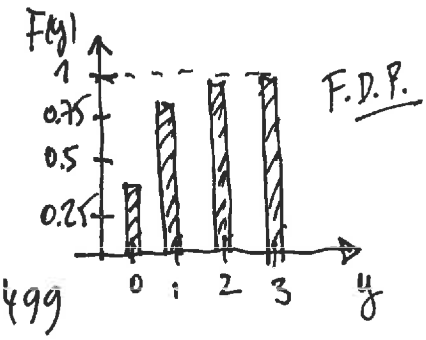
$p(1) = \frac{13}{52} \times \frac{39}{51} \times \frac{38}{50} \times 3 = 0.4360$

$p(2) = \frac{13 \times 12}{52 \quad 51} \times \frac{39}{50} \times 3 = 0.1376$

$p(3) = \frac{13}{52} \times \frac{12}{51} \times \frac{11}{50} = 0.0129$

y	p(y)	F(y)	$y \times p(y)$	$y^2 \times p(y)$
0	0.4135	0.4135	0	0
1	0.4360	0.8495	0.4360	0.4360
2	0.1376	0.9871	0.2752	0.5504
3	0.0129	1.000	0.0387	0.1161
Σ	1.000		0.7499	1.1025





2.3) Moda = 1 ; Mediana = 1 ; Média $= \mu_Y = \sum y \times p(y) = 0.7499$

Variância $= \sum y^2 \times p(y) - \mu_Y^2 = 1.1025 - 0.7499^2 = 0.5401$

2.4) $P(Y \geq 2) = p(2) + p(3) = 0.1376 + 0.0129 = 0.1506$

③

3.1) X: Quantidade de proteína no hamburguer "diabólico" (em g)

$X \sim N(\mu_X = 21.0 \,;\, \sigma_X^2 = 0.8^2)$

$P(X < 20.0) = P\left(Z < \frac{20.0 - 21.0}{0.8}\right) = P(Z < -1.25) = P(Z > 1.25) = 0.1056$

3.2) S: Soma da quantidade de proteína em 2 hamburgueres "diabólicos" (g)

$S = X_1 + X_2$ (2 variáveis Normais)

$$\left|\begin{array}{l} \mu_S = 2 \times \mu_X = 42.0\ g \\ \sigma_S^2 = 2 \times \sigma_X^2 = 2 \times 0.8^2 = 1.28\ g^2 \end{array}\right. \rightarrow S \sim N(\mu_S = 42.0 \,;\, \sigma_S^2 = 1.28)$$

3.3) $\bar{X}$: Quantidade média de proteína em 2 hamburgueres "diabólicos" (g)

$\bar{X} = \frac{X_1 + X_2}{2}$ pelo T.L.C. $\mu_{\bar{X}} = \mu_X = 21.0\,g$

$\sigma_{\bar{X}}^2 = \frac{\sigma_X^2}{n} = \frac{0.8^2}{2} = 0.32\ g^2$

$\bar{X} \sim N(\mu_{\bar{X}} = 21.0 \,;\, \sigma_{\bar{X}}^2 = 0.32)$

3.4) $P(\bar{X} \geq 20.5) = P\left(Z \geq \frac{20.5 - 21.0}{\sqrt{0.32}}\right) = P(Z \geq -0.88)$

$= 1 - P(Z < -0.88) = 1 - P(Z > 0.88) =$

$= 1 - 0.1894 = 0.8106$

4

4.1) Trata-se de um intervalo de confiança para a proporção Binomial.
Como a amostra é de grande dimensão ($n=400$), o intervalo é dado por: $\hat{p} \pm z(\alpha/2) \times \sqrt{\frac{\hat{p}\times(1-\hat{p})}{n}}$; $\hat{p} = \frac{y}{n} = \frac{150}{400} = 0.375$

Como $1-\alpha = 0.99 \Leftrightarrow \alpha = 0.01 \rightarrow z(\alpha/2) = z(0.005) = 2.5758$
o intervalo de confiança será: $0.375 \pm 2.5758 \times \sqrt{\frac{0.375\times(1-0.375)}{400}}$

$0.375 \pm 0.062 \rightarrow [0.313 ; 0.437]$

4.2) A amplitude do intervalo de confiança é dada por:
$A = 2\times z(\alpha/2) \times \sqrt{\frac{\hat{p}\times(1-\hat{p})}{n}}$ e resolvendo em ordem a $\underline{n}$:
$$n = \frac{(2\times z(\alpha/2))^2 \times \hat{p}\times(1-\hat{p})}{A^2}$$
Para garantir que $A < 0.05$, deve-se utilizar $\hat{p} = 0.5 = \hat{q}$
$$n = \frac{(2\times 2.5758)^2 \times 0.5 \times (1-0.5)}{0.05^2} \Leftrightarrow n = 2653.9 \rightarrow \underline{n \geq 2654}$$

Se utilizarmos como estimativa $\hat{p} = 0.375$,
a dimensão mínima será ligeiramente inferior: $\underline{n \geq 2488}$