

Frequência de Estatística

Duração: 2h 00m

09-11-2011

Observação: A resolução completa das questões apresentadas inclui a justificação do raciocínio utilizado e a apresentação dos cálculos efectuados.

1. Considere uma variável estatística (v.e.) bidimensional (X, Y) e sejam $(x_1, y_1), \dots, (x_n, y_n)$ ($n \geq 2$) n observações de tal variável entre as quais existem pelo menos dois valores distintos de X e de Y . Sendo r o coeficiente de correlação da v.e. (X, Y) , prove que $|r| = 1$ se e só se os pontos observados $(x_i, y_i), \forall i \in \{1, \dots, n\}$, pertencem à recta de equação

$$y = \bar{y} + r \frac{s_Y}{s_X} (x - \bar{x}),$$

onde \bar{x}, \bar{y}, s_X e s_Y designam, respectivamente, as médias e os desvios padrão das v.e. X e Y .

(**Sugestão:** Tenha em conta que $\sum_{i=1}^n \left[(y_i - \bar{y}) - r \frac{s_Y}{s_X} (x_i - \bar{x}) \right]^2 \geq 0$).

2. A fim de analisar o tipo de dependência existente entre a quantidade de cerveja ingerida, X , e a taxa de álcool no sangue em gr/l, Y , foi realizado um estudo com 50 indivíduos de uma determinada região a quem foi medida a taxa de alcoolémia 30 minutos após terem bebido um número variável de canecas de cerveja de certa dimensão. Para garantir a aleatoriedade da amostra os indivíduos foram igualmente repartidos por sexo, diferente constituição física e diferentes hábitos de consumo de álcool. Designa-se por Y_n o valor da referida taxa de álcool associada ao número x_n de canecas de cerveja bebidas pelo n -ésimo indivíduo observado ($n = 1, 2, \dots, 50$).

- a) Considere a classificação da amostra da variável estatística marginal Y presente na seguinte tabela de frequências:

Y	$[0, 0.25]$	$]0.25, 0.5]$	$]0.5, 0.75]$	$]0.75, 1]$	$]1, 1.25]$	$]1.25, 1.5]$	$]1.5, 2]$
Frequências	11	11	6	8	7	4	3

- Obtenha a função cumulativa desta distribuição estatística e deduza a ordem do percentil correspondente a uma taxa de álcool no sangue de 0.4gr/l.
- Obtenha a média e a variância da amostra classificada.

- b) Nas tabelas seguintes apresentam-se alguns resumos estatísticos, fornecidos pelo *software* estatístico SPSS, associados à distribuição da variável estatística bidimensional (X, Y) :

	Mean	Std. Deviation
X	4,86	2,680

	Y	X
Pearson Correlation Y	1,000	,932
X	,932	1,000

- Indique a média e o coeficiente de correlação da amostra observada. Com base nesta amostra, que pode afirmar sobre o tipo de dependência existente entre as variáveis em estudo?

- ii. Obtenha a relação funcional de dependência que estima existir entre as variáveis em estudo e deduza o valor previsto para a taxa de álcool no sangue de um indivíduo, daquela população, 30 minutos após ter ingerido 5 canecas de cerveja da dimensão considerada neste estudo.
3. O volume de crédito (em milhões de euros) concedido semanalmente por determinada entidade bancária é uma variável aleatória real (v.a.r.) X seguindo uma lei exponencial de parâmetro θ , ($\theta > 0$, desconhecido). Seja (X_1, \dots, X_n) uma n -amostra de X de média \bar{X}_n .
- a) Indique o modelo estatístico associado a (X_1, \dots, X_n) .
 - b) Mostre que \bar{X}_n^2 é um estimador quase certamente convergente da variância da população.
 - c) Determine a lei limite da sucessão de v.a.r. $Y_n = \frac{n\theta}{\sqrt{n+1}} (\theta\bar{X}_n - 1)$, $n \in \mathbb{N}$.
 - d) Determine um estimador da máxima verosimilhança do parâmetro da população em estudo.
 - e) Sabe-se que o volume de crédito concedido semanalmente por tal entidade bancária atingiu nos últimos dois meses os seguintes valores:

$$(40, 50, 85, 35, 70, 55, 60, 45, 40).$$

Com base nesta amostra, calcule a estimativa da máxima verosimilhança da probabilidade do referido volume de crédito ultrapassar 50 milhões de euros.

Cotação:

1.	2
2.	7
3.	11