

**MAE0217 - Estatística Descritiva - Primeiro semestre de 2017**  
**Professora: Márcia D'Elia Branco**

**LISTA 3**

1) Na tabela a seguir apresentamos 20 observações sobre a Colheita (em milhares de hectolitros) e o Preço (escudos/litro) de uma determinada variedade de vinho tinto em Portugal, durante o período de 1942 a 1961.

Ano	Colheita	Preço	Ano	Colheita	Preço
1942	6409	2.26	1952	4233	3.38
1943	19835	1.64	1953	8790	2.02
1944	10939	1.52	1954	8959	1.68
1945	7826	2.29	1955	8289	1.78
1946	7165	3.54	1956	7910	1.99
1947	7807	2.09	1957	6775	3.02
1948	6028	2.46	1958	6088	3.23
1949	6037	2.50	1959	6381	2.98
1950	6458	2.62	1960	8600	2.74
1951	6981	2.57	1961	4805	3.38

(a) Faça o diagrama de dispersão para Preço  $\times$  Colheita, determine o coeficiente de correlação linear e interprete.

(b) Obtenha a reta de regressão, adicione a reta ao diagrama de dispersão e interprete o valor obtido para o coeficiente angular ("b"). Considere o Preço como a variável resposta.

(c) Faça um gráfico de resíduos considerando no eixo horizontal a variável explicativa (Colheita). Comente.

(d) Faça um gráfico considerando no eixo horizontal a variável "Ano" e no eixo Y a variável "Colheita", uma os pontos usando segmentos de retas. Na mesma janela gráfica adicione o gráfico "Preço" por "Ano", também usando segmentos de reta. Compare as duas séries de tempo.

(e) Obtenha e analise o gráfico de probabilidades normais dos resíduos (qqnorm).

2) A planilha "dadosmalariaCEA15P14.csv" contém dados de uma investigação sobre a incidência da malária durante a gravidez e os efeitos da infecção na mãe e no recém-nascido. O estudo foi realizado nos municípios de

Cruzeiro do Sul, Mâncio Lima, Rodrigues Alves, Porto Walter e Marechal Thaumaturgo, no estado do Acre. Ao todo, 600 mulheres — não infectadas, infectadas com o parasita *Plasmodium vivax*, com o *Plasmodium falciparum* ou com ambos — foram acompanhadas durante a gestação e tiveram informações sobre o desfecho gestacional e sobre o recém-nascido registradas. Utilize os dados para responder às seguintes questões:

(a) Determine os diagramas de dispersão e os coeficientes de correlação entre *Perimetro Cefálico*  $\times$  *Peso* e *Perimetro Cefálico*  $\times$  *Estatura*.

(b) Escolha a variável com maior correlação e determine a reta de regressão. Para um recém-nascido com peso de 3000 g e estatura de 50 cm, qual o valor predito de *Perimetro Cefálico*?

(c) Vamos separar os recém-nascidos em dois grupos conforme a condição da sua mãe: Não infectada (Controle) e Infectada com algum tipo de parasita (Malária). Determine o diagrama de dispersão para *Perimetro Cefálico*  $\times$  *Peso*, diferenciando os pontos conforme o grupo que pertence o recém-nascido.

(d) Ajuste um modelo de regressão considerando como variável resposta o *Perimetro Cefálico* e as variáveis explicativas associadas a mãe: *Grupo* (0=Controle; 1=Malária) e *Idade* (0=até 35 anos; 1=mais de 35 anos). Interprete os parâmetros desse modelo. Obtenha a estimativa do desvio padrão dos erros e faça uma análise de resíduos para o modelo.

3) A probabilidade de um teste diagnóstico produzir um resultado positivo, dado que o indivíduo é portador da doença é chamada sensibilidade do teste; e a probabilidade do teste produzir um resultado negativo, dado que o indivíduo não é portador é chamado especificidade do teste. Um gráfico importante em testes clínicos, relacionado com os dois conceitos anteriores é a denominada Curva ROC. Explique como é construída e interpretada a Curva ROC.

4) Para o modelo de regressão com variável explicativa binária (0-1), prove que os estimadores de MQ para os coeficientes linear ( $\alpha$ ) e angular ( $\beta$ ) são dados por:

$$\hat{\alpha} = \bar{y}_2 \quad e \quad \hat{\beta} = \bar{y}_1 - \bar{y}_2.$$