

1. Suponha que um analista retira uma amostra de 10 carregamentos feitos por camião de uma companhia e anota a distância em quilómetros e o tempo de entrega em dias. Os dados obtidos encontram-se na tabela seguinte:

Carregamento	1	2	3	4	5	6	7	8	9	10
Distancia, X	825	215	1070	550	480	920	1350	325	670	1215
Tempo de entrega, Y	3.5	1.0	4.0	2.0	1.0	3.0	4.5	1.5	3.0	5.0

- (a) Calcule o coeficiente de correlação.
  - (b) Determine a equação de regressão dos mínimos quadrados.
  - (c) Estime o tempo de entrega de um carregamento a 1000km de distância, com base na alínea anterior.
  - (d) Construa um intervalo de previsão de 95% para o tempo de entrega, envolvendo um carregamento para 1000km.
  - (e) Determine um intervalo de confiança a 95% para o parâmetro  $b$  ( $y=a+bx$ ).
  - (f) Teste a hipótese  $H_0: b=0$  para os dados de distância e tempo de entrega a um nível de significância de 5%.
2. A tabela seguinte apresenta dados relacionados com o número de semanas de experiência a colocar fios em pequenos componentes eletrónicos bem como o número de tais componentes que foram rejeitados durante a última semana. Estes dados referem-se a 10 trabalhadores selecionados aleatoriamente.

Trabalhador	1	2	3	4	5	6	7	8	9	10
Semanas de experiência, X	7	9	6	14	8	10	4	2	11	1
Componentes rejeitados, Y	26	20	28	16	23	24	26	38	22	32

- (a) Calcule o coeficiente de correlação .
- (b) Determine a equação de regressão para prever o número de componentes rejeitados, dado o número de semanas de experiência.
- (c) Estime o número de componentes rejeitados para um empregado com 3 semanas de experiência.
- (d) Construa um intervalo de confiança de 95% para o parâmetro ordenada na origem da reta de regressão.
- (e) Construa um intervalo de previsão de 95% para o número de componentes rejeitados para um empregado com 3 semanas de experiência.

(f) Construa um intervalo de confiança de 95% para estimar o declive da reta de regressão.

3. Na tabela seguinte apresentam-se os dados de uma amostra referentes ao número de horas de estudo e respetivas notas de exame obtidas na disciplina de MATCP:

Estudante	1	2	3	4	5	6	7	8
Horas de estudo, X	20	16	34	23	27	32	18	22
Nota de exame, Y	128	122	168	140	176	184	144	154

- (a) Calcule o coeficiente de correlação.
- (b) Determine a equação de regressão dos mínimos quadrados.
- (c) Utilizando a expressão obtida na alínea anterior, estime a nota que poderá obter no exame um estudante que dedicou 30 horas de estudo.
- (d) Construa um intervalo de confiança de 90% para a nota de exame, dado que um estudante dedicou 30 horas de estudo.
- (e) Teste a hipótese nula de que o declive da reta de regressão é zero, usando um nível de significância de 1%.
- (f) A um nível de significância de 5% a ordenada na origem poderá ser nula?
4. Submeteram-se 7 engenheiros informáticos escolhidos ao acaso a um regime alimentar para reduzir peso e obtiveram-se os seguintes resultados:

Peso	1	2	3	4	5	6	7
Antes	65	71	69	75	78	66	74
Depois	63	69	64	76	75	65	70

- (a) Calcule o coeficiente de correlação.
- (b) Determine a equação da reta de regressão dos mínimos quadrados.
- (c) Determine o diagrama de dispersão para os dados da tabela e represente a reta de regressão obtida na alínea anterior.
- (d) Construa um intervalo de confiança de 95% para estimar o peso médio de um engenheiro informático que pesa 77Kg antes do regime alimentar.
- (e) Teste a hipótese que o declive da reta de regressão é diferente de zero com um nível de significância de 5%.
5. No âmbito de um estudo do tráfego rodoviário registaram-se em 20 ocasiões diferentes, os valores observados das duas variáveis seguintes:

X: densidade do tráfego, expressa em número de veículos por km

Y: velocidade média dos veículos, expressa em km/hora

As medições efetuadas produziram as seguintes estatísticas:

$$\bar{x} = 60, \quad \bar{y} = 25, \quad S_{xx} = 5000, \quad S_{xy} = -1500 \text{ e } S_{yy} = 630$$

- (a) Estime a reta de regressão que permite prever a velocidade a partir da densidade de tráfego.
- (b) Construa o intervalo de previsão a 95% para a velocidade, admitindo que a densidade de tráfego é de 70 veículos por quilómetro.
6. Para avaliar a relação existente entre o número de cafés diários que um fumador toma e o número de cigarros que fuma diariamente, foram analisados 10 fumadores, registando-se os valores das duas variáveis: X-número de cigarros e Y-número de cafés.
- A partir dos valores observados, foram calculadas as seguintes estatísticas:  
 $\bar{x} = 14.7$  ,  $\bar{y} = 5.7$  ,  $S_{xx} = 2679$  ,  $S_{xy} = 1059$  e  $S_{yy} = 439$
- (a) Determine o coeficiente de determinação.
- (b) Determine a reta de regressão que permite prever o número de cafés, a partir do número de cigarros fumados diariamente.
- (c) Construa um intervalo de confiança a 90% para o número de cafés quando são fumados 25 cigarros diariamente.
- (d) Teste a hipótese nula de que o declive da recta de regressão seja nulo, usando um nível de significância de 1%. O que poderá dizer relativamente a este parâmetro?
- (e) Construa um intervalo de confiança a 90% para o consumo médio de cafés quando são fumados 25 cigarros diariamente.
- (f) Determine um intervalo de confiança a 95% para o declive da reta de regressão.
7. A "LAVAND" é uma lavandaria situada na região do Porto. O proprietário pretende avaliar a relação entre o gasto de amaciador e consumo de sabão em pó. Na tabela seguinte apresenta-se os valores dos gastos nos últimos 10 meses.

Mês	Sabão em pó(Kg)	Amaciador(Kg)
Janeiro	27	5
Fevereiro	58	10
Março	86	15
Abril	120	20
Maio	140	25
Junho	152	30
Julho	169	35
Agosto	218	40
Setembro	226	45
Outubro	258	50

- (a) Identifique as variáveis independente e dependente.
- (b) Determine o diagrama de dispersão dos valores tabelados.
- (c) Determine o coeficiente de correlação e interprete-o.
- (d) Determine a reta de regressão.
- (e) Construa um intervalo de confiança a 95% para a quantidade de amaciador que se prevê gastar quando são consumidos 145 Kg de sabão em pó.

- (f) Teste a hipótese nula de que o declive da recta de regressão seja nulo, usando um nível de significância de 5%. O que poderá dizer relativamente a este parâmetro?
- (g) Construa um intervalo de confiança a 90% para a quantidade média de amaciador gasto quando são consumidos 130 Kg de sabão em pó.
- (h) Determine um intervalo de confiança a 95% para o parâmetro ordenada na origem da reta de regressão.
8. Uma empresa do setor informático possui 10 lojas distribuídas por outras tantas cidades. Pretende-se avaliar o nível de faturação (milhares de euros) de cada loja em função do número de habitantes (milhares) de cada cidade. Para o efeito, foi realizado um estudo que revelou os seguintes resultados estatísticos:

$$\bar{x} = 14, \quad \bar{y} = 130, \quad S_{xx} = 2528, \quad S_{xy} = 21040 \text{ e } S_{yy} = 184730$$

- (a) Identifique as variáveis dependente e independente.
- (b) Calcule o coeficiente de correlação. Interprete o valor encontrado.
- (c) Determine a equação da reta de regressão de mínimos quadrados.
- (d) Determine um intervalo de confiança a 99% para o nível médio de faturação de uma loja instalada numa cidade com 18 mil habitantes.
- (e) Para uma cidade com 21 mil habitantes, construa um intervalo de previsão a 99% para o nível de faturação.
- (f) Teste a hipótese nula de que o declive da reta de regressão seja 5.1, usando um nível de significância de 5%.

#### Soluções:

1. (a)  $\approx 0.9489$  (b)  $\hat{y} = 0.1181 + 0.0036x$  (c) 3.7 dias (d)  $]2.5168, 4.8897[$  (e)  $]0.0026, 0.0046[$
2. (a)  $\approx -0.4011$  (b)  $\hat{y} = 31.1081 - 0.6757x$  (c)  $\approx 29$  (d)  $]19.7584, 42.4578[$  (e)  $]22.1481, 36.0140[$  (f)  $] - 1.9338, 0.5824[$
3. (a)  $\approx 0.8621$  (b)  $\hat{y} = 80.1633 + 2.9332x$  (c)  $\approx 170$  (d)  $]143.1903, 196.7281[$  (f) Não
4. (a)  $\approx 0.9253$  (b)  $\hat{y} = -2.4321 + 1.0021x$  (d)  $]71.2772, 78.1754[$
5. (a)  $\hat{y} = 42 - 0.3x$  (d)  $]14.1307, 27.8693[$
6. (a)  $\approx 0.9536$  (b)  $\hat{y} = -0.1109 + 0.3953x$  (c)  $]6.6023, 12.9408[$  (e)  $]8.6731, 10.8700[$  (f)  $]0.3241, 0.4665[$
7. (c)  $\approx 0.9940$  (d)  $\hat{y} = -1.5906 + 0.2001x$  (e)  $]26.1196, 28.7203[$  (g)  $]23.3719, 25.4659[$  (h)  $]0.1821, 0.2181[$
8. (b)  $\approx 0.9736$  (c)  $\hat{y} = 13.4810 + 8.3228x$  (d)  $]142.0079, 184.5744[$  (e)  $]151.5735, 224.9455[$