

UNIVERSIDADE DE SÃO PAULO
 FACULDADE DE ECONOMIA, ADMINISTRAÇÃO E CONTABILIDADE DE RIBEIRÃO PRETO
 DEPARTAMENTO DE ADMINISTRAÇÃO

RAD1507 – Estatística Aplicada à Administração I

Lista 09

Nome: _____

Correlação e regressão linear simples. 12 exercícios:
 13.4 13.5 13.7 13.9 13.49 13.51
 13.52 Tabela 13.7 (entender) 13.74 13.78 13.80 13.84

13.4 O gerente de marketing de uma grande cadeia de supermercados gostaria de determinar o efeito do espaço

disponível em prateleira sobre as vendas de rações para animais de estimação. Uma amostra aleatória de 12 lojas de igual tamanho foi selecionada, com os seguintes resultados (armazenados no arquivo **ração.xls**):

| Loja | Espaço de Prateleira (X) (Pés) | Vendas Semanais (Y) (\$) |
|------|-----------------------------------|-----------------------------|
| 1 | 5 | 160 |
| 2 | 5 | 220 |
| 3 | 5 | 140 |
| 4 | 10 | 190 |
| 5 | 10 | 240 |
| 6 | 10 | 260 |
| 7 | 15 | 230 |
| 8 | 15 | 270 |
| 9 | 15 | 280 |
| 10 | 20 | 260 |
| 11 | 20 | 290 |
| 12 | 20 | 310 |

- a. Construa um gráfico de dispersão.
 Para esses dados, $b_0 = 145$ e $b_1 = 7,4$.
- b. Interprete o significado da inclinação, b_1 , neste problema.
- c. Faça a previsão da média aritmética das vendas semanais (em centenas de dólares) de rações para animais de estimação para lojas com 8 pés de espaço na prateleira destinado ao produto.

13.5 A circulação é a força vital do negócio de venda de revistas. Quanto maior o volume de vendas de uma revista, mais ela consegue cobrar de seus anunciantes. Recentemente, foi detectada uma defasagem entre os dados informados pelos editores de revistas das vendas em bancas e subseqüentes auditorias feitas pelo Audit Bureau of Circulations (Departamento de Auditoria de Circulação), nos Estados Unidos. Os dados no arquivo *circulação.xls* representam o volume de vendas de revistas informado e o auditado (em milhares), em 2001, para as 10 revistas a seguir:

| Revista | Informado (X) | Auditado (Y) |
|-------------------|---------------|--------------|
| <i>YM</i> | 621,0 | 299,6 |
| <i>CosmoGirl</i> | 359,7 | 207,7 |
| <i>Rosie</i> | 530,0 | 325,0 |
| <i>Playboy</i> | 492,1 | 336,3 |
| <i>Esquire</i> | 70,5 | 48,6 |
| <i>TeenPeople</i> | 567,0 | 400,3 |
| <i>More</i> | 125,5 | 91,2 |
| <i>Spin</i> | 50,6 | 39,1 |
| <i>Vogue</i> | 353,3 | 268,6 |
| <i>Elle</i> | 263,6 | 214,3 |

Fonte: Extraído de M. Rose, "In Fight for Ads, Publishers Often Overstate Their Sales," *The Wall Street Journal*, 6 de agosto de 2003, pp. A1, A10.

- Construa um gráfico de dispersão.
Para esses dados, $b_0 = 26,724$ e $b_1 = 0,5719$.
- Interprete o significado da inclinação, b_1 , neste problema.
- Faça a previsão da média aritmética para o volume de vendas de revistas auditado para um editor de revistas que informa suas vendas em bancas em 400.000.

13.7 Uma grande empresa de venda por remessa postal acredita que existe uma relação linear entre o peso da remessa dos pedidos que recebe e o número de formulários de remessa a serem preenchidos. Ela gostaria de investigar a relação para poder prever o número de formulários de remessa com base no peso da remessa. De uma perspectiva operacional, o fato de o número de formulários de remessa ser conhecido ajudará a planejar o processo de preenchimento desses formulários. Uma amostra de 25 remessas postais foi selecionada, com pedidos de itens com peso de 200 a 700 libras. Os resultados (armazenados no arquivo encomenda.xls) são os apresentados na tabela a seguir.

- Construa um gráfico de dispersão.
- Pressupondo uma relação linear, utilize o método dos mínimos quadrados para encontrar os coeficientes de regressão, b_0 e b_1 .
- Interprete o significado da inclinação, b_1 , neste problema.
- Faça a previsão para a média aritmética do número de formulários para remessa quando o peso da remessa corresponde a 500 libras.

| Peso da Remessa (Libras) | Pedidos (Milhares) | Peso da Remessa (Libras) | Pedidos (Milhares) |
|--------------------------|--------------------|--------------------------|--------------------|
| 216 | 6,1 | 432 | 13,6 |
| 283 | 9,1 | 409 | 12,8 |
| 237 | 7,2 | 553 | 16,5 |
| 203 | 7,5 | 572 | 17,1 |
| 259 | 6,9 | 506 | 15,0 |
| 374 | 11,5 | 528 | 16,2 |
| 342 | 10,3 | 501 | 15,8 |
| 301 | 9,5 | 628 | 19,0 |
| 365 | 9,2 | 677 | 19,4 |
| 384 | 10,6 | 602 | 19,1 |
| 404 | 12,5 | 630 | 18,0 |
| 426 | 12,9 | 652 | 20,2 |
| 482 | 14,5 | | |

13.9 Um corretor de uma imobiliária em uma grande cidade gostaria poder prever o custo mensal de aluguéis de apartamentos, com base no tamanho do apartamento, em pés quadrados. Foi selecionada uma amostra com 25 apartamentos (armazenada no arquivo ~~aluguel.xls~~), em um determinado bairro residencial, e as informações coletadas revelam o seguinte:

| Apartamento | Aluguel Mensal (\$) | Tamanho (Pés Quadrados) | Apartamento | Aluguel Mensal (\$) | Tamanho (Pés Quadrados) |
|-------------|---------------------|-------------------------|-------------|---------------------|-------------------------|
| 1 | 950 | 850 | 14 | 1.800 | 1.369 |
| 2 | 1.600 | 1.450 | 15 | 1.400 | 1.175 |
| 3 | 1.200 | 1.085 | 16 | 1.450 | 1.225 |
| 4 | 1.500 | 1.232 | 17 | 1.100 | 1.245 |
| 5 | 950 | 718 | 18 | 1.700 | 1.259 |
| 6 | 1.700 | 1.485 | 19 | 1.200 | 1.150 |
| 7 | 1.650 | 1.136 | 20 | 1.150 | 896 |
| 8 | 935 | 726 | 21 | 1.600 | 1.361 |
| 9 | 875 | 700 | 22 | 1.650 | 1.040 |
| 10 | 1.150 | 956 | 23 | 1.200 | 755 |
| 11 | 1.400 | 1.100 | 24 | 800 | 1.000 |
| 12 | 1.650 | 1.285 | 25 | 1.750 | 1.200 |
| 13 | 2.300 | 1.985 | | | |

- Construa um gráfico de dispersão.
- Utilize o método dos mínimos quadrados para encontrar os coeficientes de regressão, b_0 e b_1 .
- Interprete o significado de b_0 e b_1 neste problema.
- Faça a previsão da média aritmética do custo mensal do aluguel para um apartamento de 1.000 pés quadrados.
- Por que não seria apropriado utilizar o modelo para prever o aluguel mensal correspondente a apartamentos de 500 pés quadrados?
- Seus amigos Jim e Jennifer estão avaliando a possibilidade de assinar um contrato de aluguel de um apartamento nesse bairro residencial. Eles estão tentando decidir entre dois apartamentos, um deles com 1.000 pés quadrados e aluguel mensal de \$1.275 e outro com 1.200 pés quadrados e aluguel mensal de \$1.425. O que você recomendaria a eles, tendo como base os itens (a) até (d)?

13.49 A volatilidade de uma ação é geralmente medida por seu valor beta. Você pode estimar o valor beta de uma ação desenvolvendo um modelo de regressão linear simples, utilizando a variação percentual semanal do valor da ação como a variável dependente e a variação percentual semanal em um índice de mercado como a variável independente. O Índice S&P 500 é um índice empregado usualmente. Por exemplo, se desejasse estimar beta para a IBM, você poderia utilizar o modelo a seguir, que algumas vezes é conhecido como *modelo de mercado*:

$$(\text{variação \% semanal na IBM}) = \beta_0 + \beta_1 (\text{variação \% semanal no Índice S\&P 500}) + \varepsilon$$

A estimativa da regressão dos mínimos quadrados para a inclinação, b_1 , é a estimativa do valor de beta para a IBM. Uma ação com um valor de beta igual a 1,0 tende a variar da mesma maneira que o mercado global. Uma ação com um valor de beta igual a 1,5 tende a variar 50% a mais do que o mercado global, e uma ação com um valor de beta igual a 0,6 tende a variar somente 60% em relação ao mercado global. Ações com valores de beta negativos tendem a variar de maneira oposta à do mercado global. A tabela a seguir fornece alguns valores de beta para algumas ações amplamente negociadas:

| Empresa | Sigla | Beta |
|----------------|-------|------|
| AT&T | T | 0,80 |
| IBM | IBM | 1,20 |
| Disney Company | DIS | 1,40 |
| Alcoa | AA | 2,26 |
| LSI Logic | LSI | 3,61 |

Fonte: Extraída de finance.yahoo.com, 31 de maio de 2006.

- Para cada uma dessas cinco empresas, interprete o valor de beta.
- De que modo os investidores utilizam o valor de beta como um guia para investimentos?

13.51 Os dados no arquivo `bebidascafe.xls` representam as calorias e a gordura (em gramas) para 16 onças de bebidas geladas feitas com café na Dunkin' Donuts e na Starbucks:

| Produto | Calorias | Gordura |
|---|----------|---------|
| Latte Iced Mocha Swirl (leite integral), da Dunkin' Donuts | 240 | 8,0 |
| Café Frappuccino blended, da Starbucks | 260 | 3,5 |
| Café Coolatta (creme), da Dunkin' Donuts | 350 | 22,0 |
| Café Espresso Gelado Mocha (leite integral e creme chantilly), da Starbucks | 350 | 20,0 |
| Café Mocha Frappuccino blended (creme chantilly), da Starbucks | 420 | 16,0 |
| Café Chocolate Brownie Frappuccino blended (creme chantilly), da Starbucks | 510 | 22,0 |
| Chocolate Frappuccino Blended Crème (creme chantilly), da Starbucks | 530 | 19,0 |

Fonte: Extraída de "Coffee as Candy at Dunkin' Donuts and Starbucks," Consumer Reports, junho de 2004, p. 9.

- Calcule e interprete o coeficiente de correlação, r .
- No nível de significância de 0,05, existe uma relação linear significativa entre calorias e gordura?

13.52 Existem diversos métodos para se calcular a economia de combustível. A tabela a seguir (contida no arquivo [milhagem.xls](#)) indica a milhagem calculada por proprietários e pelos padrões correntes do governo:

| Veículo | Proprietário | Padrões Governamentais |
|--------------------------|--------------|------------------------|
| Ford F-150 2005 | 14,3 | 16,8 |
| Chevrolet Silverado 2005 | 15,0 | 17,8 |
| Honda Accord LX 2002 | 27,8 | 26,2 |
| Honda Civic 2002 | 27,9 | 34,2 |
| Honda Civic Hybrid 2004 | 48,8 | 47,6 |
| Ford Explorer 2002 | 16,8 | 18,3 |
| Toyota Camry 2005 | 23,7 | 28,5 |
| Toyota Corolla 2003 | 32,8 | 33,1 |
| Toyota Prius 2005 | 37,3 | 56,0 |

- Calcule e interprete o coeficiente de correlação, r .
- No nível de significância de 0,05, existe uma relação linear significativa entre a milhagem calculada pelos proprietários e a dos padrões governamentais correntes?

TABELA 13.7

Quatro Conjuntos de Dados Artificiais

| Conjunto de Dados A | | Conjunto de Dados B | | Conjunto de Dados C | | Conjunto de Dados D | |
|---------------------|-------|---------------------|-------|---------------------|-------|---------------------|-------|
| X_i | Y_i | X_i | Y_i | X_i | Y_i | X_i | Y_i |
| 10 | 8,04 | 10 | 9,14 | 10 | 7,46 | 8 | 6,58 |
| 14 | 9,96 | 14 | 8,10 | 14 | 8,84 | 8 | 5,76 |
| 5 | 5,68 | 5 | 4,74 | 5 | 5,73 | 8 | 7,71 |
| 8 | 6,95 | 8 | 8,14 | 8 | 6,77 | 8 | 8,84 |
| 9 | 8,81 | 9 | 8,77 | 9 | 7,11 | 8 | 8,47 |
| 12 | 10,84 | 12 | 9,13 | 12 | 8,15 | 8 | 7,04 |
| 4 | 4,26 | 4 | 3,10 | 4 | 5,39 | 8 | 5,25 |
| 7 | 4,82 | 7 | 7,26 | 7 | 6,42 | 19 | 12,50 |
| 11 | 8,33 | 11 | 9,26 | 11 | 7,81 | 8 | 5,56 |
| 13 | 7,58 | 13 | 8,74 | 13 | 12,74 | 8 | 7,91 |
| 6 | 7,24 | 6 | 6,13 | 6 | 6,08 | 8 | 6,89 |

Fonte: Extraída de F. J. Anscombe, "Graphs in Statistical Analysis," *American Statistician*, Vol. 27 (1973), pp. 17-21.

13.74 A gerência de uma empresa de envasamento de refrigerantes quer desenvolver um método para transferir os custos de entrega aos clientes. Embora um dos custos esteja claramente relacionado ao tempo de transporte dentro de um determinado roteiro, uma outra variável de custo reflete o tempo necessário para descarregar as caixas de refrigerante no ponto de entrega. Foi selecionada uma amostra de 20 entregas dentro de um determinado roteiro. Os tempos de entrega e o número de caixas entregues foram registrados no arquivo `entrega.xls`.

| Cliente | Número de Caixas | Tempo de Entrega (Minutos) | Cliente | Número de Caixas | Tempo de Entrega (Minutos) |
|---------|------------------|----------------------------|---------|------------------|----------------------------|
| 1 | 52 | 32,1 | 11 | 161 | 43,0 |
| 2 | 64 | 34,8 | 12 | 184 | 49,4 |
| 3 | 73 | 36,2 | 13 | 202 | 57,2 |
| 4 | 85 | 37,8 | 14 | 218 | 56,8 |
| 5 | 95 | 37,8 | 15 | 243 | 60,6 |
| 6 | 103 | 39,7 | 16 | 254 | 61,2 |
| 7 | 116 | 38,5 | 17 | 267 | 58,2 |
| 8 | 121 | 41,9 | 18 | 275 | 63,1 |
| 9 | 143 | 44,2 | 19 | 287 | 65,6 |
| 10 | 157 | 47,1 | 20 | 298 | 67,3 |

Desenvolva um modelo de regressão para prever o tempo de entrega, com base no número de caixas entregues.

- Utilize o método dos mínimos quadrados para calcular os coeficientes de regressão, b_0 e b_1 .
- Interprete o significado de b_0 e b_1 neste problema.
- Faça a previsão para o tempo de entrega de 150 caixas de refrigerantes.
- Você deveria utilizar o modelo para prever o tempo de entrega para um cliente que esteja recebendo 500 caixas de refrigerantes? Por que sim ou por que não?
- Determine o coeficiente de determinação, r^2 , e explique seu significado neste problema.
- Realize uma análise dos resíduos. Existe alguma evidência de um padrão nos resíduos? Explique.
- No nível de significância de 0,05, existem evidências de uma relação linear entre o tempo de entrega e o número de caixas entregues?
- Construa uma estimativa para o intervalo de confiança de 95% para a média aritmética do tempo de entrega de 150 caixas de refrigerantes.
- Construa um intervalo de previsão de 95% para o tempo de entrega correspondente a uma única entrega de 150 caixas de refrigerantes.
- Construa uma estimativa para o intervalo de confiança de 95% para a inclinação da população.
- Explique como os resultados obtidos em (a) até (j) podem ajudar a transferir custos de entrega aos clientes.

13.78 O diretor da graduação de uma grande faculdade de negócios gostaria de prever a média de pontuação para graduação (GPA — Grade

Point Average) de alunos em um programa de MBA, com base nos resultados do GMAT (Graduate Management Admission Test). Foi selecionada uma amostra de 20 alunos que haviam completado 2 anos no programa. Os resultados estão armazenados no arquivo `op1gmata.xls`.

| Observação | Resultado do GMAT | GPA | Observação | Resultado do GMAT | GPA |
|------------|-------------------|------|------------|-------------------|------|
| 1 | 688 | 3,72 | 11 | 567 | 3,07 |
| 2 | 647 | 3,44 | 12 | 542 | 2,86 |
| 3 | 652 | 3,21 | 13 | 551 | 2,91 |
| 4 | 608 | 3,29 | 14 | 573 | 2,79 |
| 5 | 680 | 3,91 | 15 | 536 | 3,00 |
| 6 | 617 | 3,28 | 16 | 639 | 3,55 |
| 7 | 557 | 3,02 | 17 | 619 | 3,47 |
| 8 | 599 | 3,13 | 18 | 694 | 3,60 |
| 9 | 616 | 3,45 | 19 | 718 | 3,88 |
| 10 | 594 | 3,33 | 20 | 759 | 3,76 |

(**Dica:** Primeiramente, determine qual é a variável independente e qual é a variável dependente.)

- Construa um gráfico de dispersão e, pressupondo uma relação linear, utilize o método dos mínimos quadrados para calcular os coeficientes de regressão, b_0 e b_1 .
- Interprete o significado do intercepto de Y , b_0 , e da inclinação, b_1 , neste problema.
- Utilize a linha de previsão desenvolvida em (a) para prever o GPA para um aluno com um resultado de GMAT igual a 600.
- Determine o coeficiente de determinação, r^2 , e interprete o seu significado neste problema.
- Realize uma análise de resíduos em seus resultados e determine a adequação do ajuste do modelo.
- No nível de significância de 0,05, existem evidências de uma relação linear entre o resultado do GMAT e do GPA?
- Construa uma estimativa para o intervalo de confiança de 95% da média aritmética do GPA de alunos com um resultado de GMAT igual a 600.
- Construa um intervalo de previsão de 95% do GPA de um determinado aluno com um resultado de GMAT igual a 600.
- Construa uma estimativa para o intervalo de confiança de 95% da inclinação da população.

13.80 Em 28 de janeiro de 1986, o ônibus espacial *Challenger* explodiu, matando os sete astronautas a bordo. Antes do lançamento, a temperatura atmosférica prevista para o local de lançamento indicava congelamento. Engenheiros da Morton Thiokol (fabricante do motor do foguete) prepararam gráficos com o objetivo de defender que o lançamento não fosse realizado devido ao tempo frio. Esses argumentos foram rejeitados, e o lançamento tragicamente veio a ocorrer. Diante de investigações posteriores à tragédia, especialistas concordaram que o desastre ocorrera por causa do vazamento em um retentor de borracha que não fez a vedação apropriadamente em virtude da baixa temperatura. Dados que indicam a temperatura atmosférica de 23 lançamentos anteriores, e o índice de danos nos retentores, encontram-se armazenados no arquivo *retentor.xls*.

| Número do Vôo | Temperatura (°F) | Índice de Retentores Danificados |
|---------------|---------------------|-------------------------------------|
| 1 | 66 | 0 |
| 2 | 70 | 4 |
| 3 | 69 | 0 |
| 5 | 68 | 0 |
| 6 | 67 | 0 |
| 7 | 72 | 0 |
| 8 | 73 | 0 |
| 9 | 70 | 0 |
| 41-B | 57 | 4 |
| 41-C | 63 | 2 |
| 41-D | 70 | 4 |
| 41-G | 78 | 0 |
| 51-A | 67 | 0 |
| 51-B | 75 | 0 |
| 51-C | 53 | 11 |
| 51-D | 67 | 0 |
| 51-F | 81 | 0 |
| 51-G | 70 | 0 |
| 51-I | 67 | 0 |
| 51-J | 79 | 0 |
| 61-A | 75 | 4 |
| 61-B | 76 | 0 |
| 61-C | 58 | 4 |

Observação: Os dados do voo 4 foram omitidos devido ao desconhecimento das condições do retentor.

Fonte: Extraído de Report of the Presidential Commission on the Space Shuttle Challenger Accident, Washington, DC, 1986, Vol. II (H1-H3) e Vol. IV (664), e Post Challenger Evaluation of Space Shuttle Risk Assessment and Management, Washington, DC, 1988, pp. 135-136.

- a. Construa um gráfico de dispersão para os sete vôos nos quais havia um dano no retentor (índice de dano no retentor $\neq 0$). A que conclusões, se houver alguma, você consegue chegar sobre a relação entre a temperatura atmosférica e os danos nos retentores?
- b. Construa um diagrama de dispersão para todos os 23 vôos.
- c. Explique quaisquer diferenças na interpretação da relação entre temperatura atmosférica e danos no retentor em (a) e (b).
- d. Com base no gráfico de dispersão em (b), apresente as razões pelas quais não pôde ser feita uma previsão para uma temperatura atmosférica de 31°F , a temperatura na manhã do lançamento da *Challenger*.
- e. Embora o pressuposto de uma relação linear com a temperatura atmosférica possa não ser válido, ajuste um modelo de regressão linear simples para prever os danos nos retentores, com base na temperatura atmosférica.
- f. Inclua a linha de previsão encontrada em (e) no diagrama de dispersão desenvolvido em (b).
- g. Com base nos resultados de (f), você acredita que um modelo linear é apropriado para esses dados? Explique.
- h. Realize uma análise nos resíduos. A que conclusões você chega?

13.84 Durante a safra de outono nos Estados Unidos, abóboras são vendidas em grande quantidade em barracas nas fazendas. Frequentemente, em vez de pesar as abóboras antes da venda, o fazendeiro da barraca simplesmente coloca a abóbora no cortador circular apropriado em cima do balcão. Indagado sobre o porquê disso, um fazendeiro respondeu: “Eu posso afirmar o peso da abóbora com base em sua circunferência.” Para determinar se isso é realmente verdadeiro, uma amostra com 23 abóboras teve suas circunferências e pesos aferidos, com os seguintes resultados, armazenados no arquivo `abobora.xls`:

| Circunferência (cm) | Peso (gramas) | Circunferência (cm) | Peso (gramas) |
|------------------------|------------------|------------------------|------------------|
| 50 | 1.200 | 57 | 2.000 |
| 55 | 2.000 | 66 | 2.500 |
| 54 | 1.500 | 82 | 4.600 |
| 52 | 1.700 | 83 | 4.600 |
| 37 | 500 | 70 | 3.100 |
| 52 | 1.000 | 34 | 600 |
| 53 | 1.500 | 51 | 1.500 |
| 47 | 1.400 | 50 | 1.500 |
| 51 | 1.500 | 49 | 1.600 |
| 63 | 2.500 | 60 | 2.300 |
| 33 | 500 | 59 | 2.100 |
| 43 | 1.000 | | |

- Pressupondo uma relação linear, utilize o método dos mínimos quadrados para encontrar os coeficientes de regressão, b_0 e b_1 .
- Interprete o significado da inclinação, b_1 , neste problema.
- Faça uma previsão da média aritmética do peso de uma abóbora de 60 centímetros de circunferência.
- Você acredita que seja uma boa idéia o fazendeiro vender as abóboras com base na circunferência e não no peso? Explique.
- Determine o coeficiente de determinação, r^2 , e interprete seu significado.
- Realize uma análise dos resíduos para esses dados e determine a adequação do ajuste do modelo.
- No nível de significância de 0,05, existem evidências de uma relação linear entre a circunferência e o peso de uma abóbora?
- Construa uma estimativa para o intervalo de confiança de 95% da inclinação da população, β_1 .
- Construa uma estimativa para o intervalo de confiança de 95% da média aritmética do peso da população de abóboras que tenham uma circunferência de 60 centímetros.
- Construa um intervalo de previsão de 95% do peso de uma determinada abóbora que tenha uma circunferência de 60 centímetros.