

Regressão Linear Simples

16.1 Introdução

No Capítulo 8 introduzimos o conceito de regressão para duas v.a. quantitativas, X e Y . Vimos que a esperança condicional de Y , dado que $X = x$, por exemplo, denotada por $E(Y|x)$, é uma função de x , ou seja,

$$E(Y|x) = \mu(x). \quad (16.1)$$

Em (8.27) definimos precisamente essa função. Uma definição similar vale para $E(X|y)$, que será uma função de y . Estamos considerando aqui o caso em que X e Y são definidas sobre uma mesma população P . Por exemplo, X pode ser a idade e Y o tempo de reação ao estímulo, no Exemplo 15.1. Nesse exemplo, a análise sugeriu a existência de uma relação mais forte entre as duas variáveis, e a modelamos por

$$y_{ij} = \mu_i + e_{ij}, \quad i = 1, \dots, 5, \quad j = 1, \dots, 4, \quad (16.2)$$

onde μ_i é a média do grupo de idade i . Podemos pensar que o fator idade determina cinco subpopulações (ou estratos) em P e de lá escolhemos cinco amostras aleatórias de tamanhos $n_i = 4$, $i = 1, \dots, 5$.

Em (16.1), $\mu(x)$ pode ser qualquer função de x ; veja o Exemplo 8.21. Um caso simples de interesse é aquele em que X e Y têm distribuição conjunta normal bidimensional. Nesse caso, $\mu(x)$ e $\mu(y)$ são, de fato, funções lineares. Veja a seção 8.8.

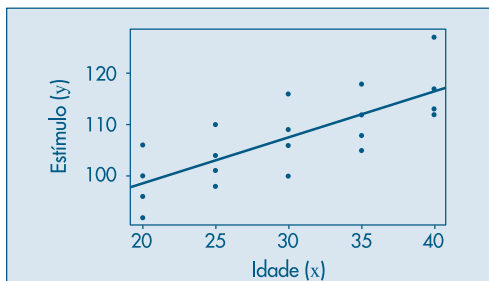
Continuando com o Exemplo 15.1, tanto X (idade) como Y (tempo de resposta ao estímulo) são v.a. contínuas, e podemos pensar em introduzir um modelo alternativo para y_{ij} , dada a relação entre X e Y . Observando as médias de Y , segundo os grupos de idades, ou seja, $E(Y|x)$, percebemos que estas aumentam conforme as pessoas envelhecem. A Figura 16.1 mostra os dados observados, onde notamos uma tendência crescente, bem como os valores repetidos de Y para cada nível de idade x .

Um modelo razoável para $E(Y|x)$ pode ser

$$E(Y|x) = \mu(x) = \alpha + \beta x, \quad (16.3)$$

ou seja, o tempo médio de reação é uma função linear da idade.

Figura 16.1: Gráfico de dispersão de idade e reação ao estímulo, com reta ajustada.



A forma da função $\mu(x)$ deve ser definida pelo pesquisador, em função do grau de conhecimento teórico que ele tem do fenômeno sob estudo. Um modelo alternativo a (16.2) seria, então,

$$y_{ij} = \mu(x_i) + e_{ij}, \quad (16.4)$$

com $E(Y|x_i) = \mu(x_i) = \alpha + \beta x_i$, $i = 1, 2, \dots, 5$. Entretanto, a forma usual de escrever o modelo é

$$y_i = \mu(x_i) + e_i, \quad (16.5)$$

onde y_i indica o tempo de reação do i -ésimo indivíduo com x_i anos de idade, $i = 1, 2, \dots, n$ e n é o número total de observações. Teremos, então, com essa notação, valores repetidos para X , por exemplo, $x_1 = \dots = x_4 = 20$. Convém reforçar a idéia que estamos propondo um modelo de comportamento para as médias das subpopulações, logo teremos de estimar os parâmetros envolvidos na função $\mu(x)$, baseados numa amostra de $n = 20$ observações, no exemplo.

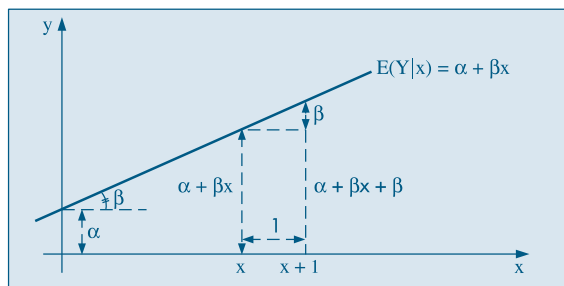
No caso de (16.3) o modelo pode ser escrito como

$$y_i = E(Y|x_i) + e_i = \alpha + \beta x_i + e_i, \quad i = 1, 2, \dots, n, \quad (16.6)$$

devendo-se encontrar os valores mais prováveis para α e β , segundo algum critério, a partir de n observações de pares de valores de (X, Y) .

Antes de prosseguirmos, seria conveniente interpretar os parâmetros envolvidos no modelo (16.5). Sabemos que α , o intercepto, representa o ponto onde a reta corta o eixo das ordenadas, e β , o coeficiente angular, representa o quanto varia a média de Y para um aumento de uma unidade da variável X . Esses parâmetros estão representados na Figura 16.2.

Figura 16.2: Representação do modelo $E(Y|x) = \alpha + \beta x$.



Voltando ao nosso exemplo, onde X é a idade e Y o tempo de reação, β representa o acréscimo no tempo médio de reação para cada ano de envelhecimento das pessoas. Aqui α representa o tempo de reação para a idade zero (recém-nascido), o que é uma inadequação do modelo.

Observação. Chamamos (16.3) de modelo linear, pois este representa uma reta. Todavia, em casos mais gerais, o termo *linear* refere-se ao modo como os *parâmetros entram no modelo*, ou seja, de forma linear. Por exemplo, o modelo

$$E(Y|x) = \alpha + \beta x + \gamma x^2,$$

embora graficamente represente uma parábola, é *modelo linear em α , β e γ* . Por outro lado,

$$E(Y|x) = \alpha e^{\beta x} \tag{16.7}$$

não é um modelo linear em α e β .

Determinados modelos não-lineares podem ser transformados em lineares, por meio de transformações das variáveis. Assim, tomando-se o logaritmo (de base e) em (16.7) obtemos

$$\ell n E(Y|x) = \ell n(\alpha) + \beta x = \alpha' + \beta x,$$

que é linear em α' e β .

Ao lado de um tratamento formal para estudar o modelo (16.6), devemos usar as técnicas de análise de dados que estudamos na Parte 1 do livro. Em particular, podemos fazer diversos tipos de gráficos *antes* que o modelo seja ajustado, *durante* o processo de ajuste e, finalmente, *depois* que o modelo foi ajustado.

A Figura 16.1 é um exemplo de um gráfico que deve ser feito antes de selecionar o modelo. Ou seja, temos um gráfico de dispersão entre as variáveis X (idade) e Y (tempo de reação ao estímulo). Esse tipo de diagrama permite ver qual o tipo de relação existente entre as variáveis, se há valores atípicos, se há valores repetidos (como no Exemplo 15.1), se a variabilidade de Y está aumentando ou não com X etc. Nesse mesmo exemplo, se decidirmos incluir a variável “acuidade visual” no modelo, teríamos duas variáveis explicativas e poderíamos fazer, por exemplo, gráficos de dispersão entre a resposta e cada variável explicativa e entre as duas variáveis explicativas. Este último nos daria uma idéia do *planejamento* envolvido, ou seja, se os pares de valores das variáveis explicativas estão cobrindo o plano (x_1, x_2) , se há grupos de pontos etc.

Exemplos de gráficos depois do ajuste serão vistos na seção 16.5, quando fizermos uma análise dos resíduos, para avaliar a adequação do modelo aos dados. Gráficos durante o ajuste são utilizados quando estudarmos a possibilidade de considerar vários modelos alternativos para o problema em questão. Esse tópico não será explorado com detalhes no livro.

16.2 Estimação dos Parâmetros

Como no capítulo anterior, iremos encontrar os estimadores de mínimos quadrados para os parâmetros do modelo linear (16.6), mas o mesmo desenvolvimento pode ser aplicado em modelos mais complexos. Será necessário ainda introduzir algumas suposições para as v.a. envolvidas. A primeira delas é que a variável X é por hipótese controlada e não está sujeita a variações aleatórias. Dizemos que X é uma variável fixa (ou sem erro ou determinística). Segundo, para dado valor x de X , os erros distribuem-se ao redor da média $\alpha + \beta x$ com média zero, isto é,

$$E(e_i | x) = 0. \quad (16.8)$$

Em terceiro lugar, e pela mesma razão apresentada no capítulo anterior, devemos supor que os erros tenham a mesma variabilidade em torno dos níveis de X , ou seja,

$$\text{Var}(e_i | x) = \sigma_e^2. \quad (16.9)$$

E em quarto lugar, introduziremos a restrição de que os erros sejam não-correlacionados.

Colhida uma amostra de n indivíduos, teremos n pares de valores (x_i, y_i) , $i = 1, \dots, n$, que devem satisfazer ao modelo (16.6), isto é,

$$y_i = \alpha + \beta x_i + e_i, \quad i = 1, \dots, n. \quad (16.10)$$

Temos, então, n equações e $n + 2$ incógnitas (α , β , e_1 , e_2 , ..., e_n). Precisamos introduzir um critério que permita encontrar α e β . Como no capítulo anterior, vamos adotar o critério que consiste em encontrar os valores de α e β que minimizam a soma dos quadrados dos erros, dados por

$$e_i = y_i - (\alpha + \beta x_i), \quad i = 1, \dots, n. \quad (16.11)$$

Obtemos, então, a quantidade de informação perdida pelo modelo ou soma dos quadrados dos erros (ou desvios)

$$SQ(\alpha, \beta) = \sum_{i=1}^n e_i^2 = \sum_{i=1}^n \{y_i - (\alpha + \beta x_i)\}^2. \quad (16.12)$$

Para cada valor de α e β teremos um resultado para essa soma de quadrados, e a solução de mínimos quadrados (MQ) é aquela que torna essa soma mínima. Temos, então, o problema de encontrar o mínimo de uma função de duas variáveis, α e β , no caso (ver Morettin et al., 2005). Derivando em relação a α e β e igualando a zero, observamos que as soluções $\hat{\alpha}$ e $\hat{\beta}$ devem satisfazer

$$\begin{aligned} n\hat{\alpha} + \hat{\beta} \sum_{i=1}^n x_i &= \sum_{i=1}^n y_i, \\ \hat{\alpha} \sum_{i=1}^n x_i + \hat{\beta} \sum_{i=1}^n x_i^2 &= \sum_{i=1}^n x_i y_i, \end{aligned} \quad (16.13)$$

as quais produzem as soluções

$$\begin{aligned}\hat{\alpha} &= \bar{y} - \hat{\beta}\bar{x}, \\ \hat{\beta} &= \frac{\sum_{i=1}^n x_i y_i - n\bar{x}\bar{y}}{\sum_{i=1}^n x_i^2 - n\bar{x}^2}.\end{aligned}\quad (16.14)$$

Substituindo em (16.3), teremos o estimador para a média $\mu(x)$, dado por

$$\hat{\mu}(x_i) = \hat{\alpha} + \hat{\beta}x_i, \quad i = 1, \dots, n, \quad (16.15)$$

que iremos indicar por

$$\hat{y}_i = \hat{\alpha} + \hat{\beta}x_i, \quad (16.16)$$

ou, ainda, por

$$\hat{y}_i = \bar{y} - \hat{\beta}\bar{x} + \hat{\beta}x_i = \bar{y} + \hat{\beta}(x_i - \bar{x}). \quad (16.17)$$

Exemplo 16.1. Voltemos ao Exemplo 15.1 e vamos ajustar o modelo (16.10), com:

y_i : tempo de reação do i -ésimo indivíduo,

x_i : idade do i -ésimo indivíduo,

e_i : desvio, $i = 1, 2, \dots, 20$.

Da Tabela 16.1 obtemos as informações:

$$\begin{aligned}n &= 20, \quad \sum y_i = 2.150, \quad \sum x_i = 600, \quad \sum x_i y_i = 65.400, \\ \bar{y} &= 107,50, \quad \bar{x} = 30, \quad \sum x_i^2 = 19.000.\end{aligned}$$

Substituindo em (16.14) obtemos

$$\begin{aligned}\hat{\beta} &= \frac{65.400 - (20)(30)(107,50)}{19.000 - (20)(30)^2} = 0,90, \\ \hat{\alpha} &= 107,50 - (0,90)(30) = 80,50,\end{aligned}$$

o que nos dá o *modelo ajustado*

$$\hat{y}_i = 80,50 + 0,90x_i, \quad i = 1, 2, \dots, 20. \quad (16.18)$$

Com esse modelo podemos prever, por exemplo, o tempo médio de reação para pessoas de 20 anos, que será indicado por $\hat{y}(20)$ e determinado por

$$\hat{y}(20) = 80,50 + (0,90)(20) = 98,50.$$

De modo análogo, os tempos médios para as idades 25, 30, 35 e 40 serão, respectivamente, estimados por: 103,00, 107,50, 112,00, e 116,50. Esses valores são muito próximos daqueles encontrados na seção 15.3, e a vantagem desse modelo sobre aquele é a possibilidade de estimar o tempo de reação médio para um grupo de idades não observado. Suponhamos, por exemplo, que se deseja estimar o tempo médio para o grupo de pessoas com 33 anos; este será dado por

$$\hat{y}(33) = 80,50 + (0,90)(33) = 110,20.$$

Na Figura 16.1 aparecem representados os dados observados, bem como a reta ajustada. Podemos observar que o modelo parece ser adequado, não apresentando nenhum ponto com desvio exagerado.

Problemas

- Usando os dados do Exemplo 15.1:
 - Encontre a reta de mínimos quadrados $\hat{z}_i = \alpha + \beta x_i$, onde z mede a acuidade visual e x , a idade.
 - Interprete o significado de α e β nesse problema.
 - Para cada indivíduo, encontre o desvio $\hat{e}_i = z_i - \hat{z}_i$; existe algum com valor muito exagerado?
- A tabela abaixo indica o valor y do aluguel e a idade x de cinco casas.
 - Encontre a reta de MQ, supondo a relação $E(y|x) = \alpha + \beta x$.
 - Faça o gráfico dos pontos e da reta ajustada. Você acha que o modelo adotado é razoável?
 - Qual o significado do coeficiente angular nesse caso?
 - E do coeficiente linear?

x	10	13	5	7	20
y	4	3	6	5	2

- Um laboratório está interessado em medir o efeito da temperatura sobre a potência de um antibiótico. Dez amostras de 50 gramas cada foram guardadas a diferentes temperaturas, e após 15 dias mediu-se a potência. Os resultados estão no quadro abaixo.
 - Faça a representação gráfica dos dados.
 - Ajuste a reta de MQ, da potência como função da temperatura.
 - O que você acha desse modelo?
 - A que temperatura a potência média seria nula?

Temperatura	30°			50°			70°			90°	
Potência	38	43	32	26	33	19	27	23	14	21	

- Ainda usando os dados do exemplo numérico 15.1, investigue o ajuste da reta de MQ na variável tempo de reação como função da acuidade visual.

16.3 Avaliação do Modelo

Nesta seção e nas seguintes estudaremos várias formas de avaliar se o modelo linear postulado é adequado ou não, dadas as suposições que fizemos sobre ele.

16.3.1 Estimador de σ_e^2

Como no capítulo anterior, para julgar a vantagem da adoção de um modelo mais complexo (linear ou outro qualquer), vamos usar a estratégia de compará-lo com o modelo mais simples, que é aquele discutido na seção 15.2, ou seja,

$$y_i = \mu + e_i. \quad (16.19)$$

A vantagem será sempre medida por meio da diminuição dos erros de previsão, ou ainda, da variância residual S_e^2 . Para o modelo ajustado (16.16), cada *resíduo* é dado por

$$\hat{e}_i = y_i - \hat{y}_i = y_i - \hat{\alpha} - \hat{\beta}x_i. \quad (16.20)$$

Como vimos na seção 16.1, vários gráficos envolvendo esses resíduos podem ser feitos para avaliar se eles são “bons representantes” dos verdadeiros e_i desconhecidos, no sentido de que as suposições feitas sobre estes estão satisfeitas. Esses gráficos serão estudados na seção 16.5.

Quando estes resíduos forem pequenos, temos uma indicação de que o modelo está produzindo bons resultados. Para julgarmos se o resíduo é pequeno ou não, devemos compará-lo com os resíduos do modelo alternativo, dados por $y_i - \bar{y}$. Da dificuldade de compará-los individualmente, preferimos trabalhar com as respectivas somas de resíduos quadráticos, dadas por

$$SQTot = \sum_{i=1}^n (y_i - \bar{y})^2 \quad (16.21)$$

$$e \quad SQRes = \sum_{i=1}^n \hat{e}_i^2 = \sum_{i=1}^n (y_i - \hat{y}_i)^2. \quad (16.22)$$

Exemplo 16.1. (continuação) Na quinta coluna da Tabela 16.1 aparecem os resíduos

$$\hat{e}_i = y_i - \hat{y}_i = y_i - (80,50 + 0,90x_i)$$

que elevados ao quadrado e somados produzirão

$$SQRes = 563,00.$$

Sabemos que $SQTot = 1.373,00$, o que mostra uma sensível redução de 810 unidades. Mais ainda, a comparação da quinta coluna da Tabela 16.1 com a coluna $e(3)$ da Tabela 15.4 mostra o melhor comportamento dos resíduos do modelo de regressão (16.18).

Tabela 16.1: Resíduos para o modelo (16.18).

i	Variáveis			Resíduos
	Tempo de Reação	Sexo	Idade	$y_i - \hat{y}_i$
1	96	H	20	-2,5
2	92	M	20	-6,5
3	106	H	20	7,5
4	100	M	20	1,5
5	98	M	25	-5,0
6	104	H	25	1,0
7	110	H	25	7,0
8	101	M	25	-2,0
9	116	M	30	8,5
10	106	H	30	-1,5
11	109	H	30	1,5
12	100	M	30	-7,5
13	112	M	35	0,0
14	105	M	35	-7,0
15	118	H	35	6,0
16	108	H	35	-4,0
17	113	M	40	-4,5
18	112	M	40	-5,5
19	127	H	40	9,5
20	117	H	40	-0,5
$SQRes$				563
S_e^2				31,28
S_e				5,59
$2S_e$				11,18

No entanto, a comparação direta dessas somas de quadrados não nos parece justa, pois o modelo (16.18) tem mais parâmetros do que o modelo (16.19). Vejamos, então, como comparar as variâncias residuais. Para o modelo simples (16.19) o estimador não-viesado de σ_e^2 é

$$S^2 = \frac{1}{n-1} \sum_{i=1}^n (y_i - \bar{y})^2 = \frac{SQTot.}{n-1}. \quad (16.23)$$

Também vimos que para o modelo (16.2), com I níveis ou subpopulações, o estimado da variância residual era

$$S_e^2 = \frac{SQDen}{n-I} = \frac{SQRes}{n-I}, \quad (16.24)$$

e I também denota o número de parâmetros desconhecidos do modelo (as médias μ_i). Portanto, de modo geral, perde-se um grau de liberdade para cada parâmetro envolvido no modelo e é natural definir o estimador de σ_e^2 num modelo de regressão como sendo

$$S_e^2 = \frac{SQRes}{n-p}, \quad (16.25)$$

onde p é o número de parâmetros do modelo. No caso particular da regressão linear simples, $p = 2$ e

$$S_e^2 = \frac{SQRes}{n-2}, \quad (16.26)$$

será um estimador não-viesado de σ_e^2 , isto é, $E(S_e^2) = \sigma_e^2$. Veja o Problema 32.

Exemplo 16.2. Continuando o exemplo anterior, obteremos

$$S^2 = 1.373/19 = 72,26, \quad S = 8,50$$

e

$$S_e^2 = 563/18 = 31,28, \quad S_e = 5,59,$$

números que sugerem uma diminuição significativa nos resíduos. Observe que, passando de um modelo com um parâmetro para outro com dois, há uma redução de 813 unidades na soma de quadrados residuais. Ou seja, perdendo um grau de liberdade, reduziu-se a soma dos resíduos quadráticos em 810 unidades, o que é mais uma evidência da vantagem de adoção do segundo modelo.

16.3.2 Decomposição da Soma de Quadrados

Ao passarmos do modelo simples para o modelo de regressão linear, vimos que a redução da soma de quadrados é dada por $SQTot - SQRes$. Esse lucro é devido à adoção do segundo modelo e será indicado por $SQReg$, significando a *soma dos quadrados devida à regressão*. Segue-se que

$$SQReg = SQTot - SQRes, \quad (16.27)$$

ou seja,

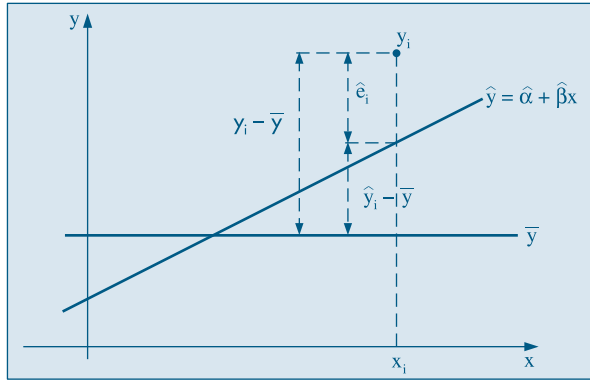
$$SQ_{Tot} = SQ_{Reg} + SQ_{Res}. \quad (16.28)$$

Observando a Figura 16.3, notamos que vale a seguinte relação:

$$y_i - \bar{y} = (y_i - \hat{y}_i) + (\hat{y}_i - \bar{y}) = \hat{e}_i + (\hat{y}_i - \bar{y}). \quad (16.29)$$

Em palavras, o desvio de uma observação em relação à média pode ser decomposto como o desvio da observação em relação ao valor ajustado pela regressão, mais o desvio do valor ajustado em relação à média.

Figura 16.3: Representação gráfica dos diversos desvios.



Elevando-se ao quadrado ambos os membros da igualdade (16.29), tomando-se a soma e observando-se que a soma do duplo produto se anula (veja o Problema 31), obtemos

$$\sum_{i=1}^n (y_i - \bar{y})^2 = \sum_{i=1}^n (\hat{y}_i - \bar{y})^2 + \sum_{i=1}^n \hat{e}_i^2, \quad (16.30)$$

ou

$$SQ_{Tot} = \sum_{i=1}^n (\hat{y}_i - \bar{y})^2 + SQ_{Res}, \quad (16.31)$$

do que deduzimos que

$$SQ_{Reg} = \sum_{i=1}^n (\hat{y}_i - \bar{y})^2. \quad (16.32)$$

De (16.17) obtemos que

$$\hat{y}_i - \bar{y} = \hat{\beta}(x_i - \bar{x}),$$

portanto, podemos escrever

$$SQ_{Reg} = \hat{\beta}^2 \sum_{i=1}^n (x_i - \bar{x})^2. \quad (16.33)$$

Daqui se pode observar que, quanto maior o valor de $\hat{\beta}$, maior será a redução da soma dos quadrados dos resíduos.

16.3.3 Tabela de Análise de Variância

Do mesmo modo como foi feito na seção 15.2, podemos resumir as informações anteriores numa única tabela ANOVA, ilustrada na Tabela 16.2.

Tabela 16.2: Tabela ANOVA para modelo de regressão.

F.V.	g.l.	SQ	QM	F
Regressão	1	SQReg	$SQReg = QMReg$	$QMReg/S_e^2$
Resíduo	$n - 2$	SQRes	$SQRes/(n - 2) = S_e^2$	
Total	$n - 1$	SQTot	$SQTot/(n - 1) = S^2$	

Também podemos medir o lucro relativo que se ganha ao introduzir o modelo, usando a estatística

$$R^2 = \frac{SQReg}{SQTot}, \quad (16.34)$$

definida anteriormente. A estatística F será discutida na seção 16.4.

Exemplo 16.3. Dos cálculos que nos levaram ao modelo (16.18), podemos construir a Tabela 16.3. Temos que

$$R^2 = \frac{810}{1.373} = 59\%.$$

Tabela 16.3: Tabela ANOVA para o modelo (16.18).

F.V.	g.l.	SQ	QM	F
Regressão	1	810	810	25,90
Resíduo	18	563	31,28	
Total	19	1.373	72,26	

O modelo proposto diminui a variância residual em mais da metade e explica 59% da variabilidade total. Verificamos, então, que é vantajosa a adoção do modelo linear (16.18) para explicar o tempo médio de reação ao estímulo, em função da idade.

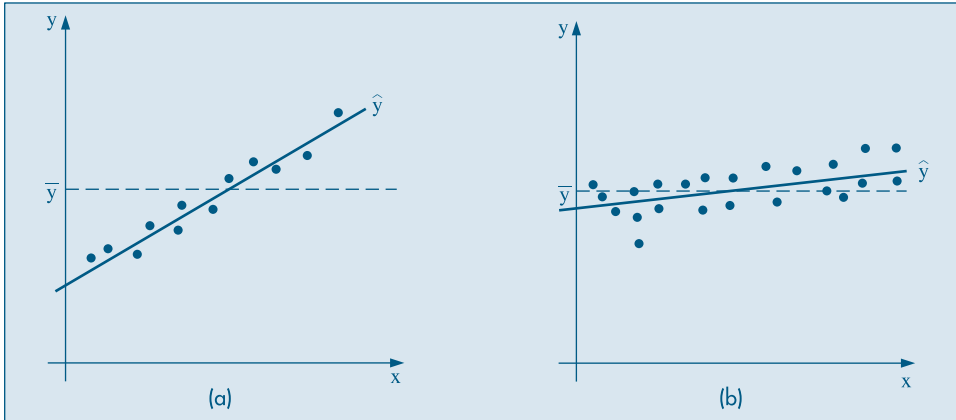
A estratégia adotada para verificar se compensa ou não utilizar o modelo $y = \alpha + \beta x + e$ é observar a redução no resíduo quando comparado com o modelo $y = \mu + e$. Se a redução for muito pequena, os dois modelos serão praticamente equivalentes, e isso ocorre quando a inclinação β for zero ou muito pequena, não compensando usar um modelo mais complexo. Estaremos, pois, interessados em testar a hipótese

$$H_0: \beta = 0, \quad (16.35)$$

o que irá exigir que se coloque uma estrutura de probabilidades sobre os erros. Esse assunto será objeto da próxima seção. A Figura 16.4 ilustra as duas situações que podem ocorrer.

Na Figura 16.4 (a) temos o caso em que claramente a variável auxiliar ajuda a prever a variável resposta. Na situação da Figura 16.4 (b) teremos dúvidas se vale a pena ou não introduzir um modelo mais complexo, ganhando muito pouco em termos de explicação.

Figura 16.4: Retas ajustadas a dois conjuntos de dados. (a) x explica y ; (b) x não explica y .



Para a avaliação final do modelo devemos investigar com mais cuidado o comportamento dos resíduos, o que será feito na seção 16.5.

Problemas

- Usando os resultados do Problema 1, construa a tabela ANOVA para o modelo $\hat{z} = \hat{\alpha} + \hat{\beta}x$, encontrado naquele problema.
 - Qual a estimativa S^2 ? E S_e^2 ?
 - Você acha que a redução nos resíduos foi grande?
 - Qual o valor de R^2 ? Interprete esse número.
- Um estudo sobre duração de certas operações está investigando o tempo requerido (em segundos) para acondicionar objetos e o volume (em dm^3) que eles ocupam. Uma amostra foi observada e obtiveram-se os seguintes resultados:

Tempo	10,8	14,4	19,6	18,0	8,4	15,2	11,0	13,3	23,1
Volume	20,39	24,92	34,84	31,72	13,59	30,87	17,84	23,22	39,65

- Faça o diagrama de dispersão dos dados.
 - Estime a reta de regressão do tempo de operação em função do volume.
 - Construa a tabela ANOVA para o modelo.
 - Qual o valor de S^2 ? É pequeno quando comparado com S_e^2 ?
 - Você acha que conhecer o volume do pacote ajuda a prever o tempo de empacotamento?
- Construa a tabela ANOVA para o Problema 2 e interprete os resultados.
 - Construa a tabela ANOVA com os dados do Problema 3.
 - Idem para o Problema 4.