

Chapter 11

Simple Linear Regression

Probabilistic Models

5 steps of Simple Linear Regression

1. Hypothesize the deterministic component
2. Use sample data to estimate unknown model parameters
3. Specify probability distribution of ε , estimate standard deviation of the distribution
4. Statistically evaluate model usefulness
5. Use for prediction, estimation, once model is useful

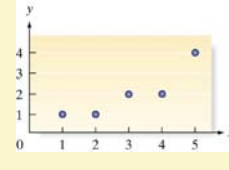
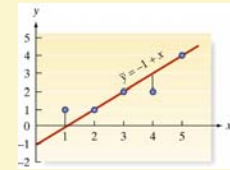
Probabilistic Models

General form of Probabilistic Models

Y = Deterministic Component + Random Error
 where
 E(y) = Deterministic Component

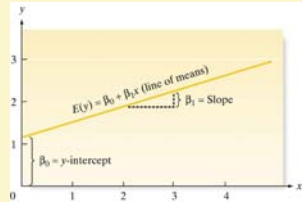
Fitting the Model: The Least Squares Approach

Reaction Time versus Drug Percentage		
Subject	Amount of Drug x (%)	Reaction Time y (seconds)
1	1	1
2	2	1
3	3	2
4	4	2
5	5	4

Probabilistic Models

First Order (Straight-Line) Probabilistic Model

$$y = \beta_0 + \beta_1 x + \varepsilon$$


Fitting the Model: The Least Squares Approach

Least Squares Line $\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x$ has:

- Sum of errors (SE) = 0
- Sum of Squared errors (SSE) is smallest of all straight line models

Formulas:

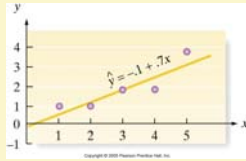
Slope: $\hat{\beta}_1 = \frac{SS_{xy}}{SS_{xx}}$ y-intercept $\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$

$$SS_{xx} = \sum (x_i - \bar{x})^2 = \sum x_i^2 - \frac{(\sum x_i)^2}{n}$$

$$SS_{xy} = \sum (x_i - \bar{x})(y_i - \bar{y}) = \sum x_i y_i - \frac{(\sum x_i)(\sum y_i)}{n}$$

Fitting the Model: The Least Squares Approach

Preliminary Computations				
x_i	y_i	x_i^2	$x_i y_i$	
1	1	1	1	
2	1	4	2	
3	2	9	6	
4	2	16	8	
5	4	25	20	
Totals	$\sum x_i = 15$	$\sum y_i = 10$	$\sum x_i^2 = 55$	$\sum x_i y_i = 37$



Comparing Observed and Predicted Values for the Least Squares Prediction Equation				
x	y	$\hat{y} = -1 + .7x$	$(y - \hat{y})$	$(y - \hat{y})^2$
1	1	.6	.4	.16
2	1	1.3	-.3	.09
3	2	2.0	0.0	.00
4	2	2.7	-.7	.49
5	4	3.4	.6	.36
Sum of Errors = 0			SSE = 1.10	