FIGURE 11.3
Plot of $E(Y) = \beta_0 + \beta_1 x_1 + \beta_2 x_2$

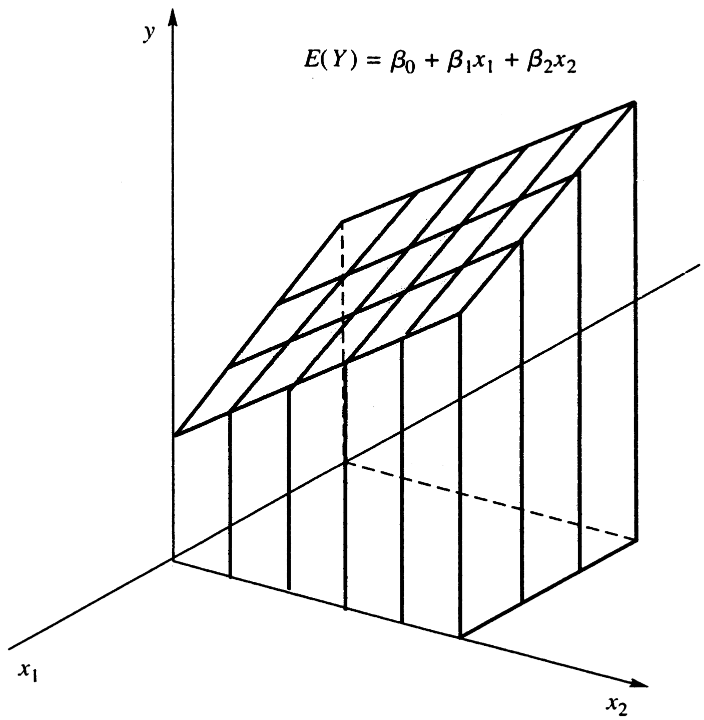
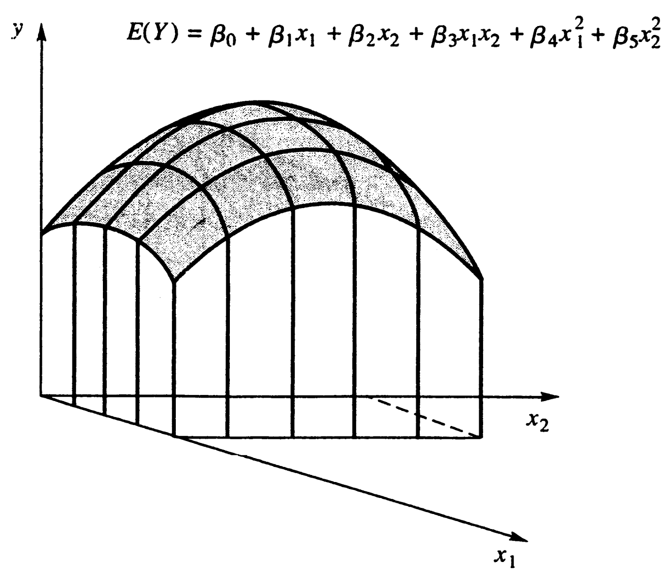


FIGURE 11.4
Plot of $E(Y) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_1 x_2 + \beta_4 x_1^2 + \beta_5 x_2^2$



Least-Squares Estimators for the Simple Linear Regression Model

1. $\hat{\beta}_1 = \frac{S_{xy}}{S_{xx}}$, where $S_{xy} = \sum_{i=1}^{n}(x_i - \overline{x})(y_i - \overline{y})$ and $S_{xx} = \sum_{i=1}^{n}(x_i - \overline{x})^2$.
2. $\hat{\beta}_0 = \overline{y} - \hat{\beta}_1\overline{x}$.

Properties of the Least-Squares Estimators; Simple Linear Regression

1. The estimators $\hat{\beta}_0$ and $\hat{\beta}_1$ are unbiased—that is, $E(\hat{\beta}_i) = \beta_i$, for $i = 0, 1$.
2. $V(\hat{\beta}_0) = c_{00}\sigma^2$, where $c_{00} = \sum x_i^2/(nS_{xx})$.
3. $V(\hat{\beta}_1) = c_{11}\sigma^2$, where $c_{11} = \frac{1}{S_{xx}}$.
4. $\text{Cov}(\hat{\beta}_0, \hat{\beta}_1) = c_{01}\sigma^2$, where $c_{01} = \frac{-\overline{x}}{S_{xx}}$.
5. An unbiased estimator of σ^2 is $S^2 = \text{SSE}/(n-2)$, where $\text{SSE} = S_{yy} - \hat{\beta}_1 S_{xy}$ and $S_{yy} = \sum(y_i - \overline{y})^2$.

If, in addition, the ε_i, for $i = 1, 2, \ldots, n$ are normally distributed,

6. Both $\hat{\beta}_0$ and $\hat{\beta}_1$ are normally distributed.
7. The random variable $\frac{(n-2)S^2}{\sigma^2}$ has a χ^2 distribution with $n - 2$ df.
8. The statistic S^2 is independent of both $\hat{\beta}_0$ and $\hat{\beta}_1$.

Test of Hypothesis for β_i

$H_0 : \beta_i = \beta_{i0}$.

$$H_a : \begin{cases} \beta_i > \beta_{i0} & \text{(upper-tail rejection region),} \\ \beta_i < \beta_{i0} & \text{(lower-tail rejection region),} \\ \beta_i \neq \beta_{i0} & \text{(two-tailed rejection region).} \end{cases}$$

Test statistic: $T = \dfrac{\hat{\beta}_i - \beta_{i0}}{S\sqrt{c_{ii}}}$.

$$\text{Rejection region: } \begin{cases} t > t_\alpha & \text{(upper-tail alternative),} \\ t < -t_\alpha & \text{(lower-tail alternative),} \\ |t| > t_{\alpha/2} & \text{(two-tailed alternative),} \end{cases}$$

where

$$c_{00} = \frac{\sum x_i^2}{nS_{xx}} \quad \text{and} \quad c_{11} = \frac{1}{S_{xx}}.$$

Notice that t_α is based on $(n - 2)$ df.

A 100(1 − α)% Confidence Interval for β_i

$$\hat{\beta}_i \pm t_{\alpha/2} S\sqrt{c_{ii}},$$

where

$$c_{00} = \frac{\sum x_i^2}{nS_{xx}} \quad \text{and} \quad c_{11} = \frac{1}{S_{xx}}.$$

A Test for $\theta = a_0\beta_0 + a_1\beta_1$

$$H_0 : \theta = \theta_0,$$

$$H_a : \begin{cases} \theta > \theta_0, \\ \theta < \theta_0, \\ \theta \neq \theta_0. \end{cases}$$

Test statistic: $$T = \frac{\hat{\theta} - \theta_0}{S\sqrt{\left(\dfrac{a_0^2 \dfrac{\sum x_i^2}{n} + a_1^2 - 2a_0a_1\overline{x}}{S_{xx}}\right)}}.$$

Rejection region: $$\begin{cases} t > t_\alpha, \\ t < -t_\alpha, \\ |t| > t_{\alpha/2}. \end{cases}$$

Here, t_α and $t_{\alpha/2}$ are based on $n - 2$ df.

The corresponding $100(1 - \alpha)\%$ confidence interval for $\theta = a_0\beta_0 + a_1\theta_1$ is as follows.

A $100(1 - \alpha)\%$ Confidence Interval for $\theta = a_0\beta_0 + a_1\beta_1$

$$\hat{\theta} \pm t_{\alpha/2}S\sqrt{\left(\frac{a_0^2 \dfrac{\sum x_i^2}{n} + a_1^2 - 2a_0a_1\overline{x}}{S_{xx}}\right)},$$

where the tabulated $t_{\alpha/2}$ is based on $n - 2$ df.

A $100(1 - \alpha)\%$ Prediction Interval for Y when $x = x^*$

$$\hat{\beta}_0 + \hat{\beta}_1 x^* \pm t_{\alpha/2}S\sqrt{1 + \frac{1}{n} + \frac{(x^* - \overline{x})^2}{S_{xx}}}.$$

In attempting to place a bound on the error of predicting Y, we would expect the error to be less in absolute value than

$$t_{\alpha/2}S\sqrt{1 + \frac{1}{n} + \frac{(x^* - \overline{x})^2}{S_{xx}}}$$

with probability equal to $(1 - \alpha)$.