Let

$$F = \frac{1}{2\hat{\sigma}^2}\left[na^2 + 2a(b-1)\sum_{i=1}^{n} X_i + (b-1)^2 \sum_{i=1}^{n} X_i^2\right], \qquad (3.3.7)$$

where a and b are the least-squares estimates of the intercept and the slope as given in (2.3.9), and $\hat{\sigma}^2$ is the mean-squared error in the ANOVA table given in (2.3.16). The null hypothesis of (3.3.6) is rejected at the α level of significance if

$$F > F(\alpha, 2, n-2),$$

where $F(\alpha, 2, n-2)$ is the αth upper quantile of a central F distribution with 2 and $n-2$ degrees of freedom. The $(1-\alpha) \times 100\%$ simultaneous confidence region, which is an ellipse, for the intercept and slope can be obtained as follows:

$$n(\alpha - a)^2 + 2(\alpha - a)(\beta - b)\sum_{i=1}^{n} X_i + (\beta - b)^2 \sum_{i=1}^{n} X_i^2 = 2\hat{\sigma}^2 F(\alpha, 2, n-2). \qquad (3.3.8)$$

We fail to reject the null hypothesis of (3.3.6) at the α level of significance if the $(1-\alpha) \times 100\%$ simultaneous confidence ellipse in (3.3.8) contains the point (0,1).

The advantage of the simple linear regression approach for the evaluation of accuracy is that the predicted recovered amount at a particular known added amount of active ingredient within the range of the concentrations used in the recovery studies can be obtained. In addition, the $(1-\alpha) \times 100\%$ confidence interval for the amount recovered can also be obtained. Similarly, the $(1-\alpha) \times 100\%$ confidence interval for the bias over the range can also be obtained. The predicted percent recovery and its corresponding $(1-\alpha) \times 100\%$ confidence interval at a particular known added amount X_i are given, respectively, as

$$\hat{Z}_i = \frac{\hat{Y}_i}{X_i} = 100\left(\frac{a + bX_i}{X_i}\right)\% = 100\left(\frac{a}{X_i} + b\right)\%, \qquad (3.3.9)$$

$$U_Z(L_Z) = \frac{100}{X_i}\left\{\hat{Y}_i \pm t\left(\tfrac{1}{2}\alpha, n-2\right)\hat{\sigma}\left[1 + \frac{1}{n} + \frac{(X_i - \bar{X})^2}{S_{XX}}\right]\right\}, \qquad (3.3.10)$$

$$\widehat{PB}_i = \hat{Z}_i - 100, \qquad (3.3.11)$$

$$U_{PB}(L_{PB}) = U_Z - 100(L_Z - 100). \qquad (3.3.12)$$

Note that it is recommended that statistical methods described in Sec. 2.6.1 be applied to test the validity of the simple regression model for validation of the assay method. These tests include the test for normality assumptions and the lack-of-fit test. An alternative approach to testing the validity of the simple linear regression model is to fit the following quadratic model:

$$Y_i = \alpha + \beta_1 X_i + \beta_2 X_i^2 + e_i, \qquad i = 1, \ldots, n.$$

Let b_2 and $SE(b_2)$ be the least-squares estimator of β_2 and its estimated standard error, respectively. Then the null hypothesis of $\beta_2 = 0$ is rejected at the α level of significance if

$$|T_2| = \left|\frac{b_2}{SE(b_2)}\right| > t\left(\tfrac{1}{2}\alpha, n - 3\right).$$

If the null hypothesis H_{01} in (3.3.4) is rejected, the null hypothesis of $\beta_2 = 0$ is not rejected, R^2 is high, and no evidence exists for the departure from normality assumption and the lack of fit, the simple linear regression model may be an adequate model to describe the relationship between the recovered and known added amounts of an active ingredient for evaluation of the accuracy of the assay method.

Example 3.3.1

To illustrate the statistical methods for evaluation of the accuracy of an assay method, consider the data given in Table 3.3.1 from a recovery study that was conducted to quantify the accuracy of the proposed assay method. The data consist of nine determinations of recovery and added amounts of an active ingredient. Table 3.3.1 also provides the percent recovery, absolute bias, and per-

TABLE 3.3.1 Recovered Amounts and Added Amounts of an Active Ingredient from a Recovery Study

Day	Added amount	Recovered amount	Percent recovery	Absolute bias	Percent bias
1	0.128	0.139	108.594	0.011	8.5938
2	0.127	0.132	103.937	0.005	3.9370
3	0.126	0.136	107.937	0.010	7.9365
1	0.383	0.365	95.300	−0.018	−4.7000
2	0.372	0.372	100.000	0.000	0.0000
3	0.376	0.395	105.053	0.019	5.0320
1	0.624	0.625	100.160	0.001	0.1603
2	0.640	0.622	97.188	−0.018	−2.8125
3	0.602	0.628	104.319	0.026	4.3189

cent bias. It can easily be verified that

$$\overline{Z} = 102.45\%, \qquad \overline{B} = 0.004, \qquad \overline{PB} = 2.499\%,$$
$$s_Z^2 = 21.296, \qquad s_B^2 = 2.235 \times 10^{-4}, \qquad s_{PB}^2 = s_Z^2 = 21.296.$$

Hence the 95% confidence interval for percent recovery and percent bias are given by

$$(L_Z, U_Z) = (98.951\%, 106.046\%),$$
$$(L_{PB}, U_{PB}) = (-1.049\%, 6.046\%),$$

respectively. Since the 95% confidence interval for percent recovery contains 100, or equivalently, the 95% confidence interval for percent bias contains 0, the assay method is considered accurate and validated.

Table 3.3.2 summarizes the estimates of the intercept and slope and their estimated standard errors from fitting a simple linear regression model of the recovered amount on the known added amount. The estimated regression line is given by

$$Y = 0.009 + 0.986X.$$

The test statistics for the null hypothesis of zero intercept and zero slope are given by

$$T_a = \frac{0.009}{0.011} = 0.846,$$
$$T_b = \frac{0.986}{0.026} = 32.232,$$

respectively. Since

$$T_a = 0.846 < t(0.025,7) = 2.365,$$

Table 3.3.2 Estimates of the Intercept and Slope for the Recovered Amount in Table 3.3.1

	Intercept	Slope
Estimate	0.009302	0.985875
Standard error	0.010995	0.025787
t statistics	0.846	38.232
p value	0.4255	<0.0001
$\hat{\sigma}^2 = 0.00024$		

we fail to reject the null hypothesis H_{03} of zero intercept at the 5% level. Furthermore, since the corresponding p value of T_b is less than 0.0001, the null hypothesis of zero slope is rejected at the 5% level of significance. However,

$$|T^*| = \left|\frac{0.986 - 1}{0.026}\right| = 0.548,$$

which is less than $t(0.025,7) = 2.365$. Hence we fail to reject the null hypothesis H_{02} of the slope being 1 at the 5% level of significance.

Figure 3.3.1 displays the scatter diagram of the recovered and added amounts, estimated regression line, and the 95% confidence band for the predicted recovered amount. It can be seen from Fig. 3.3.1 that not only a linear regression model provides an adequate fit but also the 95% confidence band is very tight. Since $R^2 = 0.995$ and the null hypothesis of zero coefficient for a quadratic term in the model was not rejected at the 5% level of significance ($T_2 = 0.176$ with a p value of 0.866), the simple linear regression model is an appropriate statistical model.

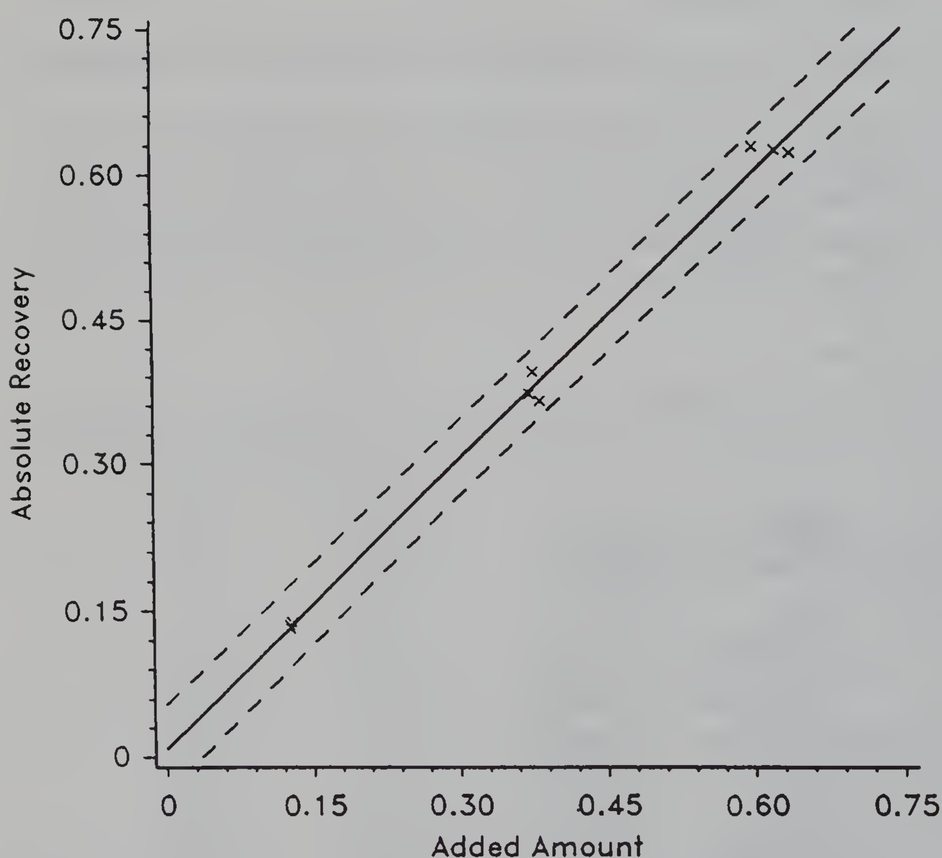

Figure 3.3.1 Scatter plot and regression line of data set in Table 3.3.1.