

4 Multiple Linear Regression Part 2

4.1 The Linear Regression Model 140

4.2 The Least Squares Estimates, and Point Estimation and Prediction 148

4.3 The Mean Square Error and the Standard Error 154

4.4 Model Utility: R^2 , Adjusted R^2 , and the Overall F-Test 155

4.5 Testing the Significance of an Independent Variable 160

s_{b_j} = the standard error of the j'th model parameter, $j = 0, \dots, k$

We will just read s_{b_j} from the Excel SAS or SPSS printout.

We can reject $H_0: \beta_j = 0$ in favor of $H_1: \beta_j \neq 0$ if $|t| = \left| \frac{b_j}{s_{b_j}} \right| > t_{\alpha/2}^{n-k-1}$

We can reject $H_0: \beta_j = 0$ in favor of $H_1: \beta_j > 0$ if $t = \frac{b_j}{s_{b_j}} > t_{\alpha}^{n-k-1}$

We can reject $H_0: \beta_j = 0$ in favor of $H_1: \beta_j < 0$ if $t = \frac{b_j}{s_{b_j}} < -t_{\alpha}^{n-k-1}$

Important: We are testing whether the j'th variable has any effect IN THIS MODEL.

A variable may have sharply different effects on the same y depending on what other variables are in the model!

4.6 Confidence and Prediction Intervals 163

A lower bound for the Margin of Error (not in book)

When all the predictor variables are close to their sample means,
then Distance is close to $\frac{1}{n}$

We can be "almost" $1 - \alpha$ confident that $\mu_{y|[x_1, x_2, \dots, x_k]}$ is within $\left[\hat{y} \pm t_{\alpha/2}^{n-k-1} \frac{s}{\sqrt{n}} \right]$

We can be "almost" $1 - \alpha$ confident that

the "next" y whose $x_1 = x_{1_0}, x_2 = x_{2_0}, \dots, x_k = x_{k_0}$ is within $\left[\hat{y} \pm t_{\alpha/2}^{n-k-1} s \sqrt{1 + \frac{1}{n}} \right]$

When all the predictor variables are NOT close to their sample means,
then Distance is MORE THAN $\frac{1}{n}$

Thus to be even "almost" $1 - \alpha$ confident

we need wider (vaguer) confidence intervals
which can be found by matrix algebra.

Note for Excel use =TINV($\alpha, n-k-1$) for $t_{\alpha/2}^{n-k-1}$

4.7 The Quadratic Regression Model 167

y is a linear function of two variables one of which is the square of the other

This makes y a nonlinear function of one variable

Mixed quadratic models.

4.8 Interaction 175

4.9 Using Dummy Variables to Model Qualitative Independent Variables 183

4.10 The Partial F-Test: Testing the Significance of a Portion of a Regression Model 193

Exercises 197