

Ch. 11 Practice

Key

1. A forester wishes to predict the volume (in cubic feet) of usable lumber in a certain species of tree using the height (in feet) and diameter (in inches) of the trees. The height and diameter of 31 trees of a certain species were measured, the trees were cut down, and the volume of usable lumber was determined.

(a) What are the explanatory variables and response variables in this study?

$Y = \text{response} = \text{usable volume}$

$X_i = \text{predictors} = \text{height \& diameter}$

(b) A multiple regression was performed on data obtained from 31 trees to predict usable volume from the height (in feet) and the diameter (in inches) of the tree. Report the regression model and complete the result table below (highlighted cells).

	Predictor	Coef	SE of Coef	t	p-value
	Intercept	-57.988	8.638	-6.71	---
$X_1 =$	Diameter	4.7082	0.2643	17.81	$\approx 0$
$X_2 =$	Height	0.3393	0.1302	2.606	0.014

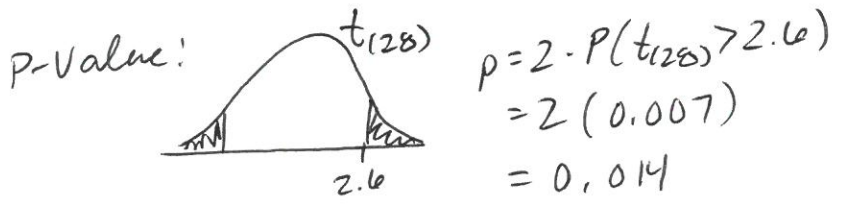
Model:  $\hat{Y}_i = -57.988 + 4.7082 X_{1i} + 0.3393 X_{2i}$

T-stat & p-value for  $X_2 = \text{Height}$ :

$H_0: \beta_2 = 0$   
 $H_A: \beta_2 \neq 0$

$t^* = \frac{0.3393 - 0}{0.1302} = 2.606$

with  $df = n - (p+1) = 31 - 3 = 28$



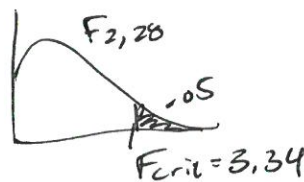
(c) Use F test to examine if this regression model is significant using  $\alpha = 0.05$ .

Analysis of Variance

Source	SS	df	MS	F
Model	$SSM = 8106.1 - 421.9$ $= 7684.2$	$p = 2$	$\frac{7684.2}{2}$ $= 3842.1$	$\frac{3842.1}{15.07}$
Error	421.9	$n - (p + 1)$ $= 28$	$421.9 / 28$ $= 15.07$	$= 254.95$
Total	8106.1	$n - 1 = 30$		

$H_0: \beta_1 = \beta_2 = 0$   
 $H_A: \text{At least one } \beta_k \neq 0$   
 $\alpha = .05$

$F^* = 254.95$   
 $df_1 = 2$   
 $df_2 = 28$



$$RR = \{F^* > 3.34\}$$

Reject  $H_0$  and conclude model is significant.

(d) Let  $\beta_2$  be the slope coefficient for height. What is the conclusion for testing the hypotheses  $H_0: \beta_2 = 0$  vs.  $H_A: \beta_2 \neq 0$  using  $\alpha = 0.05$

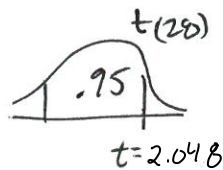
From (b),  $p\text{-value} = 0.014 < \alpha$

Reject  $H_0$ , we can conclude that height is significantly associated with usable volume.

(e) What is a 95% confidence interval for the slope coefficient of diameter (denoted by  $\beta_1$ )?

$$95\% \text{ CI for } \beta_1: b_1 \pm |t_{\alpha/2, df}| \cdot \hat{\sigma}_{b_1} = 4.7082 \pm 2.048 (0.2643)$$

$b_1 = 4.7082$   
 $\alpha_2 = 0.025$   
 $df = 28$   
 $\hat{\sigma}_{b_1} = 0.2643$



$$= (4.1669, 5.2495)$$

\*Note! Does not contain 0

(f) What is the amount of variation in the usable volume explained by the model?

$$R^2 = 1 - \frac{SSE}{SST} = 1 - \frac{421.9}{8106.1} = 0.948$$

(g) What is the adjusted  $R^2$

$$R_{adj}^2 = 1 - \frac{421.9/28}{8106.1/30} = 0.944$$

(h) What is the residual SD?

$$S_e = \sqrt{MSE} = \sqrt{15.07} = 3.9 \text{ (ft}^3\text{)}$$

(i) What is the **predicted** usable volume for a tree with a height of 10 feet and a diameter of 24 inches?

$$\hat{y}_i = -57.988 + 4.7082(24) + 0.3393(10) = 58.3814 \text{ ft}^3$$

(j) If an interval estimate should be provided for (i), should one report the confidence interval or the prediction interval?

Since (i) asks "for a tree", meaning a single future value, we should use a prediction interval.