

# Indicator (Dummy) Variables – Part IV

Johns Hopkins Engineering

## **625.461 Statistical Models and Regression**

Module 6 – Lecture 6C



## A Case Study: Table B.18, page 571 of Textbook

$y$ : fuel consumption (g/km)

$x_1$ : vehicle (0 – bus, 1 – truck)

Objective: compare mean fuel consumption between bus and truck

Two methods: Regression, ANOVA

# The Case Study – Regression Analysis Method

Truck: mean fuel consumption  $\bar{y}_1 = 357$  g/km

Bus: mean fuel consumption  $\bar{y}_0 = 361.25$  g/km

$$y = \beta_0 + \beta_1 x_1 + \varepsilon$$

$$\begin{array}{lcl} \mu_1 = \beta_0 + \beta_1 & & \beta_1 = \mu_1 - \mu_0 \\ \mu_0 = \beta_0 & \Rightarrow & \beta_0 = \mu_0 \end{array}$$

# The Case Study – Regression Analysis Method

Truck: mean fuel consumption  $\bar{y}_1 = 357$  g/km

Bus: mean fuel consumption  $\bar{y}_0 = 361.25$  g/km

$$y = \beta_0 + \beta_1 x_1 + \varepsilon$$

$$\hat{\beta}_1 = \bar{y}_1 - \bar{y}_0 = -4.25$$

$$\hat{\beta}_0 = \bar{y}_0 = 361.25$$

# The Case Study – Regression Analysis Method

$$\hat{\sigma}^2 = 89.25$$

$$Est.Var(\hat{\beta}_1) = 22.3125 \Rightarrow Est.se(\hat{\beta}_1) = 4.72$$

To test  $H_0: \mu_1 = \mu_0$ , use the  $t$  test statistic

$$t_0 = -4.25/4.72 = -0.90 \Rightarrow |t_0| = 0.90 < t_{0.025;14}$$

Cannot conclude  $\mu_1 \neq \mu_0$       p-value=0.3835

# The Case Study – ANOVA Method

Source	Sum of Squares	DF	Mean Square	F Value	Pr > F
$SS_{\text{Reg}}$	72.25	1	72.25	0.81	0.3835
$SS_{\text{Res}}$	1249.50	14	89.25		
$SS_{\text{T}}$	1321.75	15			

To test  $H_0: \mu_1 = \mu_0$ , use the  $F$  test statistic

$$F_0 = 72.25/89.25 = 0.81 \Rightarrow \text{p-value} = 0.3835$$

Cannot conclude  $\mu_1 \neq \mu_0$



JOHNS HOPKINS  
WHITING SCHOOL  
*of* ENGINEERING