Quiz #1

Dustin Leatherman

February 4, 2019

1

With respect to the meadowfoam data set in case0901 (see Section 9.1.1) in The Statistical Sleuth): a. Why were the numbers of flowers from 10 plants averaged to make a response, rather than representing them as 10 different response?

By setting a single response as the mean, it automatically means that the distribution of sample means will be normal. This makes data analysis and modeling easier since normality is a requirement for many statistical tools. Additionally, it helps cut out other confounding factors like plants who don't flower.

b. What assumption is assisted by averaging the numbers from the 10 plants?

The assumption of normality across samples is assisted.

2

With respect to the insulating data set in case0802 (see Section 8.1.2) in The Statistical Sleuth), would it be possible to test for lack of fit to the straight line model for the regression of log breakdown time on voltage by including a voltage-squared term in the model, and testing whether the coefficient of the squared term is zero?

It is not possible because by introducing a squared-term in the model, the regression line will no longer be straight.

3

Describe what σ measures in the meadowfoam problem and in the brain weigh problem (case0902, see Section 9.1.2 in The Statistical Sleuth).

In the meadowfoam problem, σ measures the standard deviation in number of flowers for all meadowfoam flowers for a given Light Intensity and PFI indicator.

In the Brain Weight problem, σ measures the standard deviation of Brain Weight for all mammals for a given gestation period, Litter Size, and Body Weight.

4

The weekly gas consumption (in 1000 cubic feet) and the average outside temperature (in degrees Celsius) were recorded for 26 weeks before and 30 weeks after cavity-wall insulation had been installed in a house in southeast England in the 1960's. The house thermostat was set at 20°C throughout.

The following model is proposed:

$\mu\{\text{Gas Consumption}|\text{Before, Temperature}\} = \beta_0 + \beta_1\text{before} + \beta_2\text{Temperature}$ where before is an indicator variable for before insulation was installed.

a. What is the model for the mean gas consumption for a week before insulation was installed, in terms of the parameters?

$\beta_0 + \beta_1 + \beta_2$Temperature

b. What is the model for the mean gas consumption for a week after insulation was installed, in terms of the parameters?

$\beta_0 + \beta_2$Temperature

c. Which parameter describes the change in mean gas consumption before and after insulation installation, holding temperature constant?

β_1

5

FEV (forced expiratory volume) is an index of pulmonary function that measures the volume of air expelled after one second of constant effort. It is of interest whether being a smoker affects FEV, but it is also known the gender and height also affect FEV.

The following model is found to fit well:

$\mu\{\text{FEV}|\text{Height, Female, Smoker}\} = \beta_0 + \beta_1\text{height} + \beta_2\text{smokerCurrent} + \beta_3\text{female} + \beta_4(\text{smokerCurrent} \times \text{female})$

where height is height in inches, smokerCurrent is an indicator variable for currently being a smoker, and female is an indicator for being female.

a. What is the effect of height, in terms of the parameters?

$\beta_0 + \beta_1$height + $\beta_2$smokerCurrent + $\beta_3$female + β_4(smokerCurrent $\times$ female) - ($\beta_0 + \beta_2$smokerCurrent + $\beta_3$female + β_4(smokerCurrent $\times$ female))

$\rightarrow \beta_1$

The effect of height is β_1.

b. What is the difference in mean FEV between a female smoker and a female non-smoker of the same height, in terms of the parameters?

$\beta_0 + \beta_1$height + $\beta_2 + \beta_3 + \beta_4$ - ($\beta_0 + \beta_1$height + β_3)

$\rightarrow \beta_2 + \beta_4$

c. The model fitted to 654 people. Using the fitted model, what estimate the fitted mean FEV for a male non-smoker 60 inches tall?

$\beta_0 + \beta_1$height

$\hat{\beta}_0$ = -5.2284

$\hat{\beta}_1$ = 0.1294

= 2.5356

6

Consider the following regression model

$\mu\{\text{Y}|\text{X,Z}\} = \beta_0 + \beta_1\text{X} + \beta_2\text{Z} + \beta_3(\text{X} \times \text{Z})$,

where X is a continuous variable and Z is also a continuous variable. What is the effect of X?

$\beta_0 + \beta_1\text{X} + \beta_2\text{Z} + \beta_3(\text{X} \times \text{Z})$ - $(\beta_0 + \beta_2\text{Z})$

$\rightarrow \beta_1\text{X} + \beta_3(\text{X} \times \text{Z})$

$\rightarrow (\beta_1 + \beta_3\text{Z})\text{X}$

The effect of X is $\beta_1 + \beta_3\text{Z}$.