**In class Activity:**

1. **Chemical composition of rain water.** The *Journal of Agricultural, Biological, and Environmental Statistics* (March 2005) presented a study of thechemical composition of rain water. The nitrateconcentration, *y* (milligrams per liter), in a rainwater sample was modeled as a function of twoindependent variables: water source (groundwater,subsurface flow, or overground flow) and silica concentration(milligrams per liter). Identify the type(quantitative or qualitative) for each independentvariable.

Qualitative Variable : Water Source

Quantitative Variable: Silica concentration

1. **Tire wear and pressure.** Underinflated or overinflated tires can increase tire wear and decrease gas mileage. A new tire was tested for wear at different pressures with the results recorded in data set: Tires2.
2. Graph the data in a scatterplot.

A graph with numbers and points

Description automatically generated

(b) If you were given the information for x = 30, 31, 32, and 33 only, what kind of model would you suggest?

First order model/ Linear model

c) For x = 33, 34, 35, and 36?

First order model/ Linear model

d) For all the data?

Quadratic Model

1. **Earnings of Mexican street vendors.** Refer to the *World Development* (February 1998) study of street vendors in the city of Puebla, Mexico, mean annual earnings, *E(y)*,was modeled as a firstorder function of age *(x*1*)* and hours worked *(x*2*)*. The data for the study are reproduced in the table.
2. Write a complete second-order model for mean annual earnings, *E(y)*, as a function of age *(x*1*)* and hours worked *(x*2*)*.

E(y) = β0 + β1x1 + β2x2 + β3x1x2+ β4x1^2 + β5x2^2

1. Fit the model and write the least squares prediction equation.

A screenshot of a computer code

Description automatically generated

E(y) = 605.94+119.68x1-139.79x2+2.6617x1x2-1.57x1^2+8.08x2^2

1. Is the model statistically useful for predicting annual earnings? Test using *α* = *.*05.

Ho: β1= β2= β3= β4= β5=0

Ha: at least one βi !=0

F − statistic: 5.594 on 5 and 9 DF, p-value = 0.0128 < 0.05 , we reject Ho

Conclusion: At 5% level of significance there is sufficient evidence to indicate that the model is statistically useful for predicting annual earnings.

1. How would you test the hypothesis that the second-order terms in the model are not necessary for predicting annual earnings?

Ho: β4= β5=0

Ha: at least one of β4 and β5 differs from 0

Reduced model :

E(y) = β0 + β1x1 + β2x2 + β3x1x2

Complete Model:

E(y) = β0 + β1x1 + β2x2 + β3x1x2+ β4x1^2 + β5x2^2

A screenshot of a computer

Description automatically generated

(e) Use R and interpret the results in part d.

The test statistic, F = 2.644 and p-value = 0.1249>0.05, we fail to reject Ho.  
we do not have sufficient evidence to support the alternative hypotheisis that the Quadratic terms contribute to the prediction of annual earnings.

1. **Expert testimony in homicide trials of battered women.** The *Duke Journal of Gender Law and**Policy* (Summer 2003) examined the impact of examined the impact of expert testimony on the outcome of homicide trials that involve battered woman syndrome. Multiple regression was employed to model the likelihood of changing a verdict from not guilty to guilty after deliberations, *y*, as a function of juror gender and whether or not expert testimony was given.
2. Identify the independent variables in the model as quantitative or qualitative.  
     
   

b) Write a main effects model for *E*(*y*) as a function of gender and expert testimony. Interpret the *β* coefficients in the model.

|  |  |  |  |
| --- | --- | --- | --- |
|  |  | **Expert Testimony** | |
|  |  | **No** | **Yes** |
| **Gender** | **Female** | **μ11** | **μ12** |
| **Male** | **μ21** | **μ22** |

Gender : 1 : if Male. Expert Testimony: 1: If Yes

0 : if Female 0: If No

Base level : Female and No

Main effect model with 2 Qualitative independent variables, gender at 2 levels (female and male) and Expert Testimony at 2 levels (No and Yes)

E(y) = β0 + β1x1 + β2x2

Interpretation of β coefficients:

β0: μ11 (Female + No)

β1: μ2j – μ1j for any level Expert testimony (j = 1,2) (Male)

β2: μi2– μi1for any level Gender (i= 1,2) (Yes)

1. Write an interaction model for *E*(*y*) as a function of gender and expert testimony. Interpret the *β* coefficients in the model.

Interaction Model with 2 Qualitative independent variables, gender at 2 (female and male) and Expert Testimony at 2 levels (No and Yes)

E(y) = β0 + β1x1 + β2x2 + β3x1x2

Interpretation of β coefficients:

β0: μ11 (Female + No)

β1: μ2j – μ1j for any level Expert testimony (j = 1,2) (Male)

β2: μi2– μi1for any level Gender (i= 1,2) (Yes)

β3: (μ22– μ12) – (μ21– μ11) (interaction between gender and testimony)

d) Based on data collected on individual juror votes from past trials, the article reported that ‘‘when expert testimony was present, women jurors were more likely than men to change a verdict from not guilty to guilty after deliberations.’’ Assume that when no expert testimony was present, male jurors were more likely than women to change a verdict from not guilty to guilty after deliberations. Which model, part b or part c, hypothesizes the relationships reported in the article?

Part c – Interaction model

Ho : β3 = 0

Ha : β3 != 0

Pvalue <0.05, we reject Ho, β3 is significant