

Homework 4
STAT 334, Spring 2019, Prof. Walker
Due Thursday, May 9 at the beginning of class

1

- **Problem 9 requires R. R commands are given in the lecture notes. All other problems may be done in JMP or R. (Use the .jmp files for JMP and the .csv files for R.)**
 - **Interpret the results of all tests and confidence intervals in the context of the data.**
 - **You may switch HW groups if you want. If you make changes to your HW group, please let me know by Monday, May 6.**
1. Problem 6.6, p. 248-249. Use the data in **BrandPref.jmp** (or **BrandPref.csv** for R). (See Problem 6.5 for a description of the data.)
 2. (a) Check the regression assumptions for the **BrandPref** data using residual plots. For each assumption, list what plots you looked at (or “none” if you didn’t need plots), your conclusion, and why you made that decision.
(b) Do a lack-of-fit test at $\alpha = 0.05$. State the hypotheses and p-value, then interpret the result of the test in the context of the data.
 3. Problem 6.8, p. 249. Use the **BrandPref** data.
 4. Problem 6.19, p. 252. Use the data in **CommProp.jmp** or **CommProp.csv**. (See Problem 6.18 for a description of the data.)
In part (c), interpret the R^2 adjusted—not the R^2 .
 5. Check the regression assumptions for the **CommProp** data using residual plots. For each assumption, list what plots you looked at (or “none” if you didn’t need plots), your conclusion, and why you made that decision.
 6. Problem 10.8, p. 415. Use the **CommProp** data.
Make partial regression plots (also known as added variable plots or JMP leverage plots) for each of the 4 predictor variables. Answer the questions below.
(a) Explain how you would compute the values plotted on both the vertical and horizontal axes of the partial regression plot for the Vacancy variable.
(b) Do the partial regression plots suggest that any of the relationships between the response and predictor variables are non-linear? If so, which one(s)?
 7. Problem 6.22, p. 253. Hint: Some of these may be transformable into linear regression models.
 8. Problem 6.25, p. 253
 9. Problem 6.27, p. 254.
Import the data in file **HW4.csv** into R. Then, run the code in the R script **HW4.r** to format the data.
Type in commands given in the R scripts for our previous lectures to compute the answers. Remember that matrix formulas for regression don’t change from simple to multiple regression.
 10. Problem 7.7, p. 289-290. Use the **CommProp** data.
 11. Problem 7.8, p. 290. Use the **CommProp** data.
 12. Problem 7.10, p. 290. Use the **CommProp** data.

JMP Instructions for this homework are on the next page.

Many commands needed for this homework were covered in earlier homework instructions. Here is what's new.

Global F-test

- The **Analyze>Fit Model** command produces the ANOVA table as part of the basic regression output.

Residual plots for multiple regression

- From the regression output triangle menu, choose **Save Columns>Studentized Residuals** and **Save Columns>Predicted Values**.
- Use the **Analyze>Fit Y-by-X** command. Select the studentized residuals as **Y**, then select the predicted values as well as every predictor variable (x_1, x_2, \dots) as **X**.
- Once the plots have been created, use the triangle menu above each plot to choose **Fit Mean** to put a horizontal line across the graph.
- Normal probability plots of residuals and index plots of residuals (if necessary) are created the same way as in simple regression.

Confidence interval for $E(y_v)$ and prediction interval for y_v

- This is the same as in simple regression, you just need to enter a plug-in value for every predictor in the first blank line at the bottom of the JMP data table. **Do not enter a new value for y .** In the regression output, use the triangle menu to **Save Columns** for the kind of interval that you want.

Partial regression plot/added variable plot/JMP leverage plot

- Run a regression using the **Fit Model** command using the default **Effect Leverage** emphasis. The leverage plots are created automatically.

Sequential Sums of Squares

- Add the predictor variables to the regression in the order that you want when you specify the variables for the analysis.
- Once you have the regression output, use the triangle menu to select **Estimates>Sequential Tests**.

Partial F-test

- There is no automatic way to do a partial F-test in JMP. Run regressions for both the reduced model and the full model. Use the results from the ANOVA tables of both models to compute the F-ratio by hand.
- To get the p-value for the partial F-test, follow these steps:
 - Use the JMP command **Help>Sample Data>Teaching Scripts>Interactive Teaching Modules>Distribution Calculator**
 - Select the F-distribution, enter d.f., select Input values, click probability $X > q$, and enter F-statistic as Value.

For *R* commands, go back through the *R* scripts for each lecture and modify them for the data files in HW4.