|  |  |  |
| --- | --- | --- |
| **STA 672** | **Fall 2015** | **Test** |

**DO ONLY 2 OF THE 4 PROBLEMS** – The first two columns must be *filled in\** and this page *submitted* with any work you wish to have considered for credit.

|  |  |  |
| --- | --- | --- |
| Name \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ | Problem  # | Points |
| First problem (number)  to be graded: |  | /20 |
| Second problem (number)  to be graded: |  | /20 |
| Total: |  | / 40 |

This test is open-book, open-note—you may use any resource you wish other than the help of another human being. All cell phones must be ***turned off*** (not in vibration-only mode,) or the person not following this instruction will not be allowed to turn his/her exam for credit.

\*Submission with more or fewer than two numbers to be graded will be returned for no credit.

#1. A statistics instructor has noticed, while blowing up balloons for his granddaughter’s birthday party, that balloons of different colors seem to vary in the amount of time to blow up. He decides this is worthy of research, and with a package of 100 balloons (20 of each color: red, pink, yellow, blue, and orange) and the help of a team of poor starving teaching assistants (TAs) working for pizza, he has each of the five TAs pull a random color out of a hat. Whichever color they draw is the color of balloon they will be inflating in the experiment—blowing up each balloon to its target diameter of 7 inches, with the instructor using a stopwatch to record each time to the nearest tenth. The TAs proceed in order—red, pink, yellow, blue, orange—giving them time to recover between balloons.

(a) (5 points) Identify the following elements of their experiment:

i) Research question:

ii) Factor(s):

iii) Treatment(s):

(b) (7 points) Is this a well-designed experiment? If not, explain what is wrong and how it could be improved.

(c) (8 points) Some suggestions for investigating the instructor’s research question:

* Suggestion 1: Average the inflation times across each color, comparing the averages for each color to determine any variations in the inflation times.
* Suggestion 2: Randomly assign a new color balloon to each TA on each round, at the end comparing the averages for each color to determine any variations in the inflation times.
* Suggestion 3: Randomly assign a color to the 5 TAs, having them blow up five balloons of that color—each successive five balloons, randomly assign each TA to a color he/she hasn’t inflated yet. At the end, compare the averages for each color, using the TAs as a blocking variable.

Comment on the three suggestions. Explain which solution is best approach to answer the instructor's research question and also explain what is wrong with the other two.

#2. An experiment on the efficacy of nail varnish solvent in removing nail varnish from cloth was run using two sorts of varnish and two sorts of varnish solvent and is excerpted here:

|  |  |  |
| --- | --- | --- |
| **Varnish** | **Varnish Solvent** | |
| 1 | 2 |
| 1 | 27.58, 28.75, 25.83, 29.25, 22.08 | 29.17, 31.75, 32.08, 28.50, 31.25 |
| 2 | 22.50, 25.17, 27.75, 21.50, 25.00 | 34.42, 26.00, 31.08, 29.17, 24.75 |

The response variable is time (in minutes) for the varnish to dissolve from the cloth.

(a) (6 points) Write down the cell means model for the data and identify the components.

(b) (4 points) Use the JMP output as necessary to decide whether the efficacy of the two varnish solvents in removing nail varnish depended on the type of varnish.

(c) (4 points) The researcher’s decision was that the interaction between nail varnish and nail varnish solvent was negligible, so that term has been removed from subsequent JMP analyses for this problem. Using these, would you decide there is a significant effect of the different solvents in efficacy in removing varnish from cloth?

(d) (6 points) Consistent with your answers to the previous parts, use the appropriate JMP analysis (some additional are included) for the significant effect in this model. Tell me which effect you are reporting, why, and what the *comparison*-wise error rate would have been.

#3. The data from problem #2 are repeated here (an experiment on the efficacy of nail varnish solvent in removing nail varnish from cloth was run using two sorts of varnish and two sorts of varnish solvent):

|  |  |  |
| --- | --- | --- |
| **Varnish** | **Varnish Solvent** | |
| 1 | 2 |
| 1 | 27.58, 28.75, 25.83, 29.25, 22.08 | 29.17, 31.75, 32.08, 28.50, 31.25 |
| 2 | 22.50, 25.17, 27.75, 21.50, 25.00 | 34.42, 26.00, 31.08, 29.17, 24.75 |

a) (3 points) What assumption about a blocking variable and the factor of interest can be justified using the first JMP output? Explain your answer.

b) (6 points) Fit the analysis of variance model using the 2 varnishes as a blocking factor, testing for an effect of different solvents on the time at the ** = 0.05 level of significance. State the hypotheses in words, identify the appropriate test statistic and *p*-value, and state your conclusion.

c) (6 points) Fit the analysis of variance model ignoring the effect of the two varnishes, testing for an effect of different solvents on the time at the ** = 0.05 level of significance. State the hypotheses in words, identify the appropriate test statistic and *p*-value, and state your conclusion.

d) (5 points) Comment on the difference in the results in b) and c), citing specific outcomes and explain this difference.

#4. (Due to Larry Lesher) An experiment was run to determine whether pushing a certain pedestrian light button had an effect on how long the person had to wait before the light showed “walk.” The factor of interest was the number of pushes of the button. The data are reproduced here:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Time in seconds for “Walk” sign to appear** | | | | |
| **Number of Pushes** | **0** | **1** | **2** | **3** |
| Time | 38.14 | 38.28 | 38.17 | 38.14 |
| 38.20 | 38.17 | 38.13 | 38.30 |
| 38.31 | 38.08 | 38.16 | 38.21 |
| 38.14 | 38.25 | 38.30 | 38.04 |
| 38.29 | 38.18 | 38.34 | 38.37 |

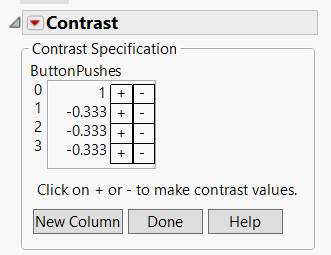
Use the JMP output on the handout where necessary to answer the following questions.

(a) (6 points) Write out the cell means model for this experiment and explain the components.

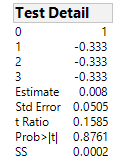
(b) (6 points) Test for an effect of pushing the pedestrian light button on the waiting time at the level of significance. Clearly state your hypotheses, the appropriate test statistic and *p*-value, and your conclusion.

(c) (2 points) Comment on the appropriateness (utility) of any further analyses on these data (considering number of pushes on the button as a factor.)

(d) (6 points) Regardless of your answer to part (c), explain what is being tested with the following option:



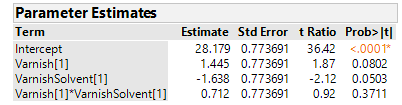
and output:



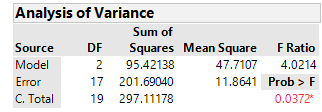
| **SS** | **NumDF** | **DenDF** | **F Ratio** | **Prob > F** |
| --- | --- | --- | --- | --- |
| 2e-4 | 1 | 16 | 0.0251 | 0.8761 |

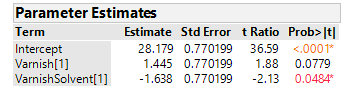
summarizing the results of this procedure using an error rate of , and explain how these results are consistent with those of part (b).

**Problem #2**

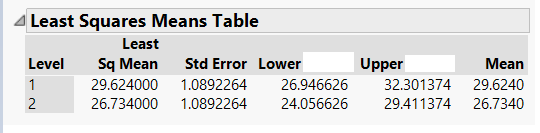
****

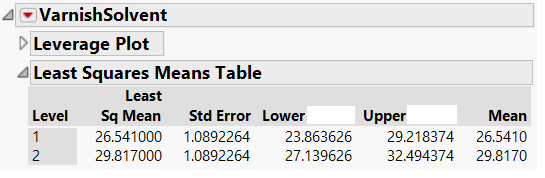
***Subsequently:***

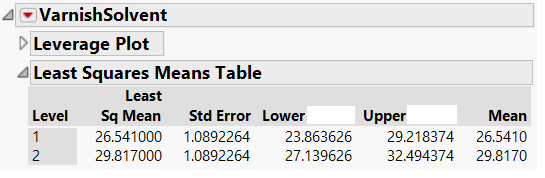
****

****



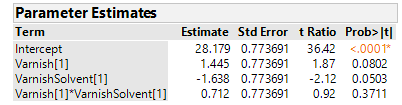






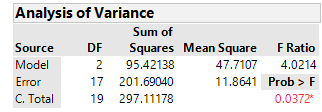
**Problem #3**

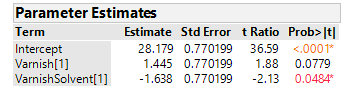
For part a)

****

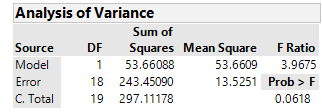
**Problem #3, (cont’d)**

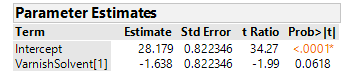
For part b)

****

****

For part c)

****

****

**Problem #4**

