Design and Analysis of Experiments, HW 2

Due Tuesday, January 6, beginning of class.

Instructions: (Late hw's/e-mailed hw's are not accepted.) Be sure to follow the formatting instructions (see instructions for hw 1 which is posted on course website). In particular, be sure to put all output/work for a given problem **together in one place** and **include a copy of all relevant Minitab output - text and graphics - in your hw.**

1. In this problem you will analyze data for comparing two methods of measuring mercury in fish, selective reduction and what is referred to as "the permanganate method." The mercury in each of 25 juvenile black marlin was measured using both techniques. Note that since each fish was measured by both methods, we have paired data. The 25 measurements of mercury by the selective reduction and permanganate method (in ppm of mercury) are given in columns Hg1 and Hg2, respectively.

i. What hypotheses should be tested in order to determine if there is a systematic difference between the levels of mercury measured by the two tests?

ii. Download the data from the following URL

`www.rose-hulman.edu/~inlow/fish.MTW`

and test the hypotheses in part i using a t-test. What do you conclude at $\alpha = 0.05$?

iii. Are the assumptions/requirements of this paired t-test met? Recall that the assumptions/requirements on your flow chart are that the differences are IID and from a normal population. An additional assumption (not typically discussed in intro stat courses) is that the differences d_i should be IID with respect to the magnitude of the measurements. Check this by making a scatterplot of d_i versus the Hg2 measurements. (This is an alternative to d_i versus $(Hg1_i + Hg2_i)/2$ plot.) What do you conclude about the assumption and how the two measurements compare to each other? Be specific.

2. Bubba wants a 308 deer rifle. A good buddy of his has three 308 rifles, each from a different manufacturer, and is willing to sell one of them to Bubba for \$200. Bubba borrows all three to see which is most accurate. Since Bubba always uses the cheapest ammo he can get (whatever is on sale), he decides to compare the accuracy of the rifles across several different randomly selected ammo brands. He randomly selects 4 boxes of 308 shells, each from a different brand. Each box contains 12 shells. Design an experiment so that Bubba can optimally compare the three rifles (using the 4 boxes of shells) to see which is most accurate. Describe in gory detail how Bubba should shoot the 4 boxes of shells using the 3 rifles. Discuss the applicability, if any, of each of the three basic principles of experimental design to the experiment you devised for Bubba.

3. On the first problem on the quiz many of you proposed using a complete randomized block design to control the drift of the test instrument over time, a nuisance factor. Each set of three contigous measurements constituted a block - hence there are $15/3 = 5$ blocks. Within each block a bolt of each of the 3 alloys is measured once so the design is a balanced complete block. The order of the three bolts within each block is random so this design is thus a balanced randomized complete block design. In this problem you will finish specifying the experimental design by determining the order of each alloy within each block using Minitab as follows:

i. In column c1 enter the block number for each row/trial by putting "1" in the first three rows, "2" in the next three rows and so on until the first 15 rows are filled. In column c3 enter "A" in the first cell, "B" in the second cell, and "C" in the third cell. Repeat this pattern until the first 15 rows are filled.

ii. Fill the first 15 rows of column c2 with pseudorandom numbers using `Calc -> Random Data -> Uniform` and specifying 15 for `number of rows to generate`.

iii. Sort column c3 by column c1 then column c2 using `Data -> Sort` as follows:

1. Select c3 as the sort column.
2. Select c1 as the first column to sort by and column c2 as the second column.
3. Choose settings so that the sorted version of c3 will be stored in column c4 in your current worksheet then click OK.

iv. By sorting on column c1 first - which is already sorted - you "trick" Minitab into sorting on the three pseudorandom numbers corresponding to each block. Since these numbers are in random order, when placed in ascending order they randomize the order of the alloys **within each block** - see the ordering of alloy type in column c4. Provide a printout of the four columns in your hw. This printout, in particular column c4, constitutes your experimental design. To run the experiment a technician would determine the type of bolt alloy for each trial/run then randomly select a bolt of that alloy and measure it. By using this experimental design, bias due to instrument drift in your alloy strength comparisons is minimized.

4. In the article "Review of Development and Application of CRSTER and MPTER Models" (R. Wilson, *Atmospheric Environment*, 1993:41-57), several measurements of the maximum hourly concentrations (in μgm/m^3) of SO_2 are presented for each of four power plants. The results are as follows (two outliers have been deleted):

```
Plant 1:  438   619   732   638
Plant 2:  857  1014  1153   883  1053
Plant 3:  925   786  1179   786
Plant 4:  893   891   917   695   675   595
```

Analyze this data as follows: (see next page)

i. Construct side-by-side boxplots of the SO_2 measurements for the different plants by entering the data into four columns and then using `Graph -> Boxplot -> Multipe Y's Simple`. Request that Minitab indicate the mean on the boxplots by checking `Mean symbol` under the `Data View` menu. Based on these plots, do there appear to be significant differences between the means?

ii. Test H$_o$: $\mu_1 = \cdots = \mu_4$ vs. H$_a$: $\mu_i \neq \mu_j$ *for at least one* (i, j) *pair* at $\alpha = 0.05$ using Minitab's general linear model procedure. What do you conclude?

iii. Assuming the observations are independent - you can't theck this since you don't know the order in which the data was collected - check the other ANOVA assumptions/requirements. Do they appear to be met?

5. An experiment was set up to measure the yield provided by each of three catalysts in a certain reaction. The experiment was repeated three times for each catalyst. The reactor yields, in grams are as follows:

```
Catalyst 1:  84 90 86
Catalyst 2:  88 90 87
Catalyst 3:  95 91 93
```

i. For this experiment, **MANUALLY** compute the ANOVA table; see the table at the bottom of page 3 in your ANOVA handout (Table 3-3, page 75 in your text). (Hint: Compute the easiest two of the three sums of squares then compute the third using the ANOVA Sum of Squares decomposition (page 4 ANOVA handout):

$$SS_T \quad = \quad SS_{\text{Treatments}} + SS_E$$

ii. Compute the p-value using Minitab's `Graph -> Probability Distribution Plot -> View Probability` routine.

iii. Check your results using Minitab's general linear model routine.

6. (If you have not already done so, you might want to do problem 3 prior to doing this problem.) Read the attached pages from the chapter *Statistical Analysis of Experimental Data* from the book *Mathematics Today: Twelve Informal Essays*. Answer the following questions:

i. What are the two benefits of randomization?

ii. The author, David Moore (emeritus professor, Purdue University), provides a simplified formula for F in the box on page 218 which - like the formula in the ANOVA handout - assumes the three samples have the same size, i.e., $n_1 = n_2 = n_3 = n$. However, his formula has an error. What's missing?

iii. Why is the F ratio/test statistic called "F" as opposed to G? (Incidentally, do you know why the t statistics is called "t?")

vi. See next page...

iv. On page 217, Moore writes "Fisher's approach to inference via probability was to evaluate the significance of the observed variation by comparing it to the inherent variation due to uncontrolled factors. How does our ANOVA model,

$$Y_{ij} = \mu_i + \epsilon_{ij},$$

incorporate or take into account the "inherent variation due to uncontrolled factors?"

7. Insulation shields for electrical wires are manufactured using three different types of machines. A company wants to compare the machines. A quality engineer is assigned to determine if the diameters of the shields produced by the machines differ with respect to their means. Open the data for the three machines at the following URL

`www.rose-hulman.edu/~inlow/diam.MTW`

and analyze it as follows:

i. Compare the diameters of the shields from the three machines using side-by-side boxplots. Based on the boxplots, what do you suspect concerning possible differences in the means of the shields from the three machines?

ii. Test the null hypothesis that the means are equal via ANOVA using Minitab. What do you conclude about H_0 at $\alpha = 0.05$?

iii. Assuming the independence assumption/requirement is met - you can't check this since you don't know the order - check as many of the remaining ANOVA assumptions/requirements as you can. Are they all met?

iv. At least one of the ANOVA assumptions/requirements is not met so the F-test result in part ii is not reliable. Compare the diameters with respect to stochastic order using the Kruskal-Wallis test: `Stat -> Nonparametrics -> Kruskal-Wallis`. What do you conclude at $\alpha = 0.05$? How do you reconcile this result with your conclusion in part ii?

v. You should have rejected the null of no differences with respect to stochastic ordering between the machines in part iv. But which machines differ? We need a nonparametric multiple comparison procedure. One approach is to use a two-sample nonparametric procedure to compare the machines pairwise. In order to do this, we must use a smaller α in order to control α', the *experimentwise* type I error probability, i.e, the probability of making at least one type I error. One easy way to control α' is to use the Bonferroni approach. This approach says that to do a set of r tests at α', conduct each test at α'/r. Test each of the three pairs using the Mann-Whitney procedure (`Stat -> Nonparametrics -> Mann-Whitney`) in conjunction with the Bonferroni approach so that $\alpha' = 0.05$. (NOTE: You can't use the 2-sample bootstrap/`boot2mean` because the samples are too small.) Which machines differ with respect to stochastic order based on this analysis?