Homework 6

In this homework you will use the Wilcoxon sum rank test, Wilcoxon sign rank test, paired t-test and ANOVA as well as getting some JMP practice.

(1) What is wrong with each of these statements?

(i) A researcher wants to test $H_0 : \bar{x}_1 = \bar{x}_2$ against $H_A : \bar{x}_1 \neq \bar{x}_2$.

(ii) A study recorded the IS scores of 100 college students. The scores of 56 males in the study were compared with all 100 students in the study using the independent two-sample t-test.

(iii) A two-sample t statistic gave a p-value of 0.94. From this we can reject the null hypothesis with 90% confidence.

(iv) A researcher is interested in testing $H_0 : \mu_1 - \mu_2 = 0$ against $H_A : \mu_1 - \mu_2 < 0$. The test gave $t = 2.15$. Since the p-value for the two-sided alternative ($H_A : \mu_1 - \mu_2 \neq 0$) gave p-value equal to 3.6%, the researcher concludes that the p-value for the one-sided test is 1.8%.

(2) Our objective is to see if there is a difference in weights of eight week old calves given different treatments, in particular treatment B and treatment C. Two random samples (each of size eleven) of new born calves are given the different treatments. At 8 weeks, those on treatment B had a sample mean of 139.54 pounds and those on treatment C had a sample mean of 144.45 pounds. This is difference of almost 5 pounds, to a layman this may suggest there is a difference in the treatments.

However, by now you should realise that this difference could be due to sampling differences. Your objective is to analyse the data to see whether this difference is statistically significant (ie, the difference is too large and the standard error is too small so that probability that this difference is due to random is small).

The (Statcrunch) output for the result is given the right figure, and the steps for answering this question are given below.

(i) Using the output evaluate the t-transform under the null hypothesis (that there is no difference).

(ii) How many degrees of freedom do we use for the t-transform?

(iii) Go to Statcrunch – > Stat – > Calculation – > T, and calculate the probability associated with the t-transform.

(iv) Based on your result in (iii), what is the p-value for the test and is there evidence to reject the null and suggest there is a difference in the mean weight of calves on treatment B and C?
(3) Nutritionists are studying the effect a calorie controlled diet may have on the life expectancy of mice. They randomly sampled 12 mice and put 6 on a calorie controlled diet and the other 6 on a typical mouse diet. The life time of each mouse was monitored and the results are given below.

Let 1 indicate the mice was on a calorie controlled diet and 0 indicate it was put on a usual diet.

(a) State the null and alternative that the nutritionist wants to investigate.
(b) State the test you would do and why?
(c) Do the test at the 5% level.

(4) A group of ornithologists want to investigate whether cell phone masts have increased the number of eggs laid by birds nesting near by. To see whether there is an increase, the ornithologists locate 8 permanent nests and count the number eggs laid in the spring before the masts were constructed and then count the number of eggs after the masts were constructed. The data is collected below (you can do the calculations below in either JMP or by hand).

(i) State the null and alternative that the ornithologists want to investigate.
(ii) Make a scatter plot of the Before masts against the After masts, from the plots do you see a association between the location and the number of eggs (are the points 'random' or do they appear to have a pattern?).

(iii) Do you think the assumptions to do a paired t-test are satisfied?

(iv) Explain why it is best to use a Wilcoxon sign rank test.

(v) Do the test at the 10% level.

(vi) What happens if I do the test at the 5% level?

(vii) Do the paired t-test at the 10% level. Explain any differences in the results between the paired t-test and the Wilcoxon sign rank test.

(5) Import the calf weight data into JMP. It is believed that the weight of a newborn animal in general drops immediately after they are born and only after a few weeks does the weight get back to the birth weight and above. We want to investigate whether the calf weight data suggests this to be true. Look at this data in JMP. The column with Wt 0 contains all the weights at birth. Wt 0.5 contains the weights at week 0.5, Wt 1 the weights at week 1 etc.

(a) Make a scatter plot of the week 0 weights against the week 1 weights and from this explain why a paired t-test should be used (rather than an independent sample t-test) to compare the mean weight of a calf at different weeks.

(b) Do the following tests at the 5% level.

   (i) Do a paired t-test to test whether the weight has dropped between week 0 and week 1.

   (ii) Do a paired t-test to test whether the weight has dropped between week 0 and week 2.

   (iii) Do a paired t-test to test whether the weight has dropped between week 0 and week 3.

   (iv) Do a paired t-test to test whether the weight has dropped between week 0 and week 4.

   (v) Do a paired t-test to test whether the weight has increased between week 0 and week 3.

   (vi) Do a paired t-test to test whether the weight has increased between week 0 and week 4.

(c) Based on your results in (b) summarise the mean behaviour of calf weights from week 0 to week 4.

   Write this as a mini report (say about 5 lines) summarising how you came to your conclusions.

(d) Why should we be a little cautious about doing multiple tests (in the way that we have done above).
(6) Import the calf weight data into JMP. Investigate whether treatment (denoted as TRT) has an impact on the mean weight of a calf at week 8. Write this as a mini report (say about 7 lines), stating your null and alternative, the ANOVA table, the distribution used (including the number of degrees of freedom) and the p-value (do the test at the 5% level). Also comment about the confidence intervals of the mean weight for the 4 treatments.