Semester 2	Statistical Tests (Advanced)	2011

Tutorial 4

1. The research department of a company that produces industrial string proposes a new technique for strengthening the string and making it able to withstand stronger forces. To find out whether the new technique really does strengthen the string, Five lengths of string are produced by the standard techniques and five by the new technique and samples are compared for their respective breaking points, in pounds of force. The breaking points are as follows:

Standard technique	144	131	155	126	134
New technique	139	154	132	143	147

   Do the data support the contention of the research department that the new technique produces stronger string? Use the Wilcoxon rank-sum test.

2. The following data give the stroke index for 10 patients before and after treatment. We wish to test the hypothesis that the treatment has no effect.

Before	109	57	53	57	68	72	51	65	52	61
After	56	44	55	40	62	46	48	41	56	49

   (a) Calculate the differences, draw a boxplot of these and comment on why a non-parametric test might be preferred.

   (b) Use the sign test to see if the treatment has no effect.

3. For the Wilcoxon rank-sum statistic based on samples of size m and n, show that the general variance formula in Lecture 9 (Week 3.3, p. 135) reduces to $\frac{mn}{12}(N+1)$ when there are no ties.

4. Calculate the exact distribution of the Wilcoxon rank-sum statistic for the case $n = 4$ and $m = 3$ when 3 values are tied on the lowest score, 2 on the next score and 2 on the largest score. Calculate $\mathrm{Var}(W)$ directly and verify that the formula given in lectures.

5. Davidson et al (1974) studied the responses to oral glucose in patients with Huntington's disease and a group of control patients. The five hour glucose responses (in mgs) to oral carbohydrate in 11 patients and 10 controls were:

Huntington's patient:	77	82	85	86	86	86	89	91	92	93	100
Control group:	65	65	73	75	77	78	83	85	90	97	

   Use the Wilcoxon test to see if both groups can be modelled by the same distribution.

Computer Exercise

1. The data in *amp* are the results of measurements made using two methods on 15 pairs of tablets to determine the dosage of ampicillin. They were analysed in week 1 using a t-test.

   (a) Use the sign test to see if there is a systematic difference between the two measure methods.

      i. Create a vector d whose elements correspond to the difference of the two columns of the data *amp*. (Note that there is a zero observation in d)

      ii. Remove the zero from d, and name the new vector as $d1$.

      iii. Calculate the size n of $d1$ and the number x of positive elements in $d1$.

      iv. Calculate the p-value.

   (b) Use the Wilcoxon sign-rank test to test the same hypothesis as in (a).

   (c) Use *wilcox.test* to confirm your findings in (b).

   (d) Perform a t-test for the same hypothesis as in (a).

   (e) Compare the results obtained by three different tests: the sign test, the Wilcoxon sign-rank test and the t-test. Comment on the differences in the three different tests including the assumptions and the p-values.

   (f) Obtain a boxplot of d. Is a transformation necessary to test the hypothesis in (a) by using the Wilcoxon signed-rank test or the t-test?