ST4241: Design and Analysis of Clinical Trials

2009/2010: Semester I

Tutorial 2

1. Two methods are applied to train patients with senile dementia to care for themselves. After the completion of the training, patients are asked to take 20 tests involving activities of daily living. The response from each patient is the proportion of his or her tests that are successful. The data for the two groups are given below:

```
Group 1:  0.05, 0.15, 0.35, 0.25, 0.20, 0.05, 0.10, 0.05, 0.30, 0.05, 0.25

Group 2:  0, 0.15, 0, 0.05, 0, 0, 0.05, 0.10.
```

(i) Conduct a t test for the difference between the two groups based on the data. Comment on the appropriateness of this t-test.

R-code:

```
x1=c(0.05, 0.15, 0.35, 0.25, 0.20, 0.05, 0.10, 0.05, 0.30, 0.05, 0.25 )
x2=c(0, 0.15, 0, 0.05, 0, 0, 0.05, 0.10)
t.test(x1,x2,alternative = "two.sided",  var.equal = T, onf.level = 0.95)
```

Results:

```
t = 2.7686, df = 17, p-value = 0.01315
```

The difference between the two group means is significant with p-value 0.01315. But the data is far from normally distributed. The p-value does not provide an trustful evidence for the significance.

(ii) Transform the data to $Y = \arcsin\sqrt{X}$. Conduct a t test for the difference between the two groups based on the transformed data. Compare the result with that in part (i).

R-code:

```
y1=asin(sqrt(x1))
y2=asin(sqrt(x2))
t.test(y1,y2,alternative = "two.sided",  var.equal = T, onf.level = 0.95)
```

Results:

```
t = 3.3231, df = 17, p-value = 0.004024
```

The test provides a much stronger evidence for the significance of the difference.

(iii) Conduct the Mann-Whitney-Wilcoxon test based on the data: (a) without adjustment on ties, and (b) with adjustment on ties. Compare the results.

The ordered observations and their ranks are given in the following table:

Group	x_{ij}	*Rank without ties*	*Rank with ties*
2	0.00	1	2.5
2	0.00	2	2.5
2	0.00	3	2.5
2	0.00	4	2.5
1	0.05	5	7.5
1	0.05	6	7.5
1	0.05	7	7.5
1	0.05	8	7.5
2	0.05	9	7.5
2	0.05	10	7.5
1	0.10	11	11.5
2	0.10	12	11.5
1	0.15	13	13.5
2	0.15	14	13.5
1	0.20	15	15.0
1	0.25	16	16.5
1	0.25	17	16.5
1	0.30	18	18.0
1	0.35	19	19.0

From the table, we obtain

$$\bar{R}_1 = 12.73, \bar{R}_2 = 6.25.$$

$$\chi^2 = \frac{12 n_1 n_2 (\bar{R}_1 - \bar{R}_2)^2}{n_.^2 (n_. + 1)} = \frac{12 \times 11 \times 8 \times (12.73 - 6.25)^2}{19^2 \times 20} = 6.1415.$$

(a) The p-value of the test with adjustment for ties is $Pr(\chi_1^2 > 6.1415) = 0.013.$

(b) To adjust for ties, compute f as follows:

$$T = 5,\ t_1 = 4, t_2 = 6, t_3 = t_4 = t_5 = 2.$$
$$\sum_{i=1}^{T} t_i(t_i - 1)(t_i + 1) = 4 \times 3 \times 5 + 6 \times 5 \times 7 + 3 \times (2 \times 1 \times 3) = 288.$$
$$n.(n. - 1)(n. + 1) = 6840.$$
$$f = 1 - 288/6840 = 0.9579.$$

The adjusted statistic is $\chi^2/0.9579 = 6.4114$. The corresponding p-value is 0.011.

2. Prove that the MWW test statistic is identical to the expression given below:

$$\chi^2 = \frac{12\left(S_i - \frac{n_i(n.+1)}{2}\right)^2}{n_1 n_2 (n. + 1)},$$

where $S_i = n_i \bar{R}_i$ is the sum of the ranks of the n_i measurements in Group i, $i = 1, 2$. Note the MWW test statistic is originally given by

$$\chi\text{mww} = \frac{12 n_1 n_2 (\bar{R}_1 - \bar{R}_2)^2}{n_\cdot^2 (n. + 1)}.$$

Note that $\bar{R}_i = S_i/n_i$ and $S_1 + S_2 = \frac{n.(n.+1)}{2}$ or $S_2 = \frac{n.(n.+1)}{2} - S_1$. Thus

$$\bar{R}_1 - \bar{R}_2$$
$$= \frac{S_1}{n_1} + \frac{S_1}{n_2} - \frac{n.(n. + 1)}{2n_2}$$
$$= \frac{n.}{n_1 n_2}[S_1 - \frac{n_1(n. + 1)}{2}].$$

Substituting the expression for $\bar{R}_1 - \bar{R}_2$ in χmww, the result follows.

3. The following table summarizes the data from a trial for the comparison of two drugs to a control. The response variable is a blood count (in millions of cells per cubic millimeter).

Group	n_i	$\bar{X}_i$	s_i
Control	6	8.25	0.94
Drug A	4	8.90	0.90
Drug B	5	10.88	1.56

(i) Compute s^2 and the two test statistics L_1 and L_2 for testing whether the drugs are significantly different from the control.

$$s^2 = (5 \times 0.94^2 + 3 \times 0.9^2 + 4 \times 1.56^2)/(6 + 4 + 5 - 3) = 1.3819.$$
$$L_1 = (8.9 - 8.25)/\sqrt{1.3819} \times \sqrt{6 \times 4/(4+6)} = 0.8566.$$
$$L_2 = (10.88 - 8.25)/\sqrt{1.3819} \times \sqrt{6 \times 5/(5+6)} = 3.6948.$$

(ii) Test whether the mean for Drug B is significantly greater than the control mean at level 0.01. Should a two-sided test or a one-sided test be used? What is the appropriate critical value to be used for the test?

A one-sided test should be conducted. The critical value for overall significance level 0.01 with one-sided test in Table A.6 of Fleiss should be used. It is found that $d_{2,12,0.01} = 3.01$. *Since* $|L_2| = 3.6948 > 3.01$. *The mean of Drug B is significantly greater than the control mean at level 0.01.*

4. Suppose that a total of $n.$ experimental subjects are to be assigned randomly to a control group or to one of p experimental groups. Let n_0 denote the number to receive the control treatment and n_t the number to receive each experimental treatment, so that $n. = n_0 + pn_t$. Show that the variance of $\bar{X}_i - \bar{X}_0$ is minimized if

$$n_0 = \frac{n.}{1+\sqrt{p}}, \quad n_t = \frac{n.}{p+\sqrt{p}}.$$

The variance is proportional to $\frac{n_0+n_t}{n_0 n_t}$. *Substituting* $n_0 = n. - pn_t$ *into it, we have*

$$h(n_t) = \frac{n. - (1-p)n_t}{n_t(n. - pn_t)}$$

Differentiating $h(n_t)$ *and setting the derivative to zero yields the equation*

$$p(1-p)n_t^2 + 2n.pn_t - n_.^2 = 0.$$

The two roots of the equation are given by

$$\frac{n.(p+\sqrt{p})}{(p+\sqrt{p})(p-\sqrt{p})} = \frac{n.}{(p-\sqrt{p})},$$

and

$$\frac{n.(p-\sqrt{p})}{(p+\sqrt{p})(p-\sqrt{p})} = \frac{n.}{(p+\sqrt{p})}.$$

Since pn_t *must be smaller than* $n.$, *the right root is* $n_t = \frac{n.}{(p+\sqrt{p})}$.