475.101/102/107/108 Semester 1 2000
Assignment 5 Solutions

Question 1.

(a) **(i)** Paired, observational study. **(ii)** Paired, experiment.
(iii) Paired, observational study.

(b) Let Y be the number of positive signs. Assuming null hypothesis is true and ignoring hypothesized values of 0 in the data, we get $Y \sim$ Binomial(n=8, p=0.5). We observed 1 positive sign and 7 negative signs.
P-value = 2×pr($Y \le 1$) = 2×0.035 = 0.07.

Note: We can calculate the *P*-value for the test in a number of different ways, all of which are equally valid. (You need to know how to use at least one of these approaches.)
Let y = minimum of the number of +'s and the number of -'s = minimum of 1 and 7 = 1.
P-value = 2×pr($Y \le$ minimum) = 2×pr($Y \le 1$) = 2×0.035 = 0.07.
or: Let y = maximum of the number of +'s and the number of -'s = maximum of 1 and 7 = 7.
P-value = 2×pr($Y \ge$ maximum) = 2×pr($Y \ge 7$) = 2×(1–pr($Y \le 6$)) = 2×0.035 = 0.07.
or: Let y = number of +'s = 1.
P-value = 2×minimum(pr($Y \le 1$), pr($Y \ge 1$))
= 2×minimum(0.035, 0.965) = 2×0.035 = 0.07.

(c) $df_1 = k - 1 = 5 - 1 = 4$. $df_2 = n - k = (8+11+9+10+12) - 5 = 50 - 5 = 45$.
$df_{tot} = df_1 + df_2 = 4 + 45 = 49$. $f_0 = \dfrac{215.57}{68.49} = 3.147$.

Question 2.

(a) The parameter of interest here is μ_{Diff}, the underlying mean of the differences in the reaction times of male 100 metre sprinters. We wish to test $H_0 : \mu_{Diff} = 0$ vs $H_1 : \mu_{Diff} \neq 0$.
The *P-value* of 0.006 provides very strong evidence against H_0. Thus we have very strong evidence that there is a difference in the average reaction time of male 100 metre sprinters between round 3 and round 1. With 95% confidence the mean reaction time of male 100 metre sprinters in round 3 is between 0.005 and 0.024 seconds faster than in round 1.

(b) The null hypothesis being tested is that differences calculated between rounds 3 and 1 are Normally distributed versus the alternative hypothesis that they are not Normally distributed. The Normal probability plot and the *W*-test (*P-value* > 0.100) provide no evidence against the null hypothesis, thus indicating that the assumption of Normality is valid.

(c) The purpose of this study was to investigate the reaction times of male 100 metre elite sprinters, using data measuring the reaction times from sprinters running in round 1 and round 3 of the men's 100 metres at the 1996 Atlanta Summer Olympics.
We can conclude from our results that as they advance in the competition the reaction times of male 100 metre elite sprinters are on average faster. There is a significant difference in the average reaction time of 100 metre sprinters between round 3 and round 1, with the mean reaction time of 100 metre sprinters in round 3 being between 0.005 and 0.024 seconds faster than in round 1.
The assumptions required for the analysis of the data to be applicable were checked and found to be satisfied.

Question 3.

(a) **(i)** The parameter of interest here is $\mu_E - \mu_Y$ where μ_E is the mean amount of forward/backward sway experienced by an "elderly" group and μ_Y is the mean amount of forward/backward sway experienced by a "young" group. We wish to test $H_0 : \mu_E - \mu_Y = 0$ vs $H_1 : \mu_E - \mu_Y \neq 0$.

We have some evidence (*P-value* ≈ 4%) of a difference in the mean amount of forward/backward sway experienced by "elderly" and "young" groups. With 95% confidence the mean amount of forward/backward sway experienced by an "elderly" group is between 0.3 and 16.1 millimetres more than that experienced by a "young" group.

Minitab output

Two Sample T-Test and Confidence Interval

Two sample T for elderly -fwd.bkwd vs young - fwd.bkwd

	N	Mean	StDev	SE Mean
elderly	9	26.33	9.77	3.3
young -	8	18.13	4.09	1.4

95% CI for mu elderly - mu young -: (0.3, 16.1)
T-Test mu elderly = mu young - (vs not =): T = 2.30 **P = 0.044** DF = 10

Excel output

t-Test: Two-Sample Assuming Unequal Variances

	elderly	*young*
Mean	26.33333333	18.125
Variance	95.5	16.69642857
Observations	9	8
Hypothesized Mean Difference	0	
df	11	
t Stat	2.303480324	
P(T<=t) one-tail	0.020887664	
t Critical one-tail	1.795883691	
P(T<=t) two-tail	**0.041775329**	
t Critical two-tail	2.200986273	

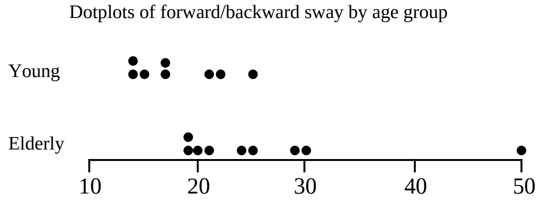
(ii) The parameter of interest here is $\tilde{\mu}_E - \tilde{\mu}_Y$ where $\tilde{\mu}_E$ is the median amount of forward/backward sway experienced by an "elderly" group and $\tilde{\mu}_Y$ is the median amount of forward/backward sway experienced by a "young" group. We wish to test $H_0 : \tilde{\mu}_E - \tilde{\mu}_Y = 0$ vs $H_1 : \tilde{\mu}_E - \tilde{\mu}_Y \neq 0$.

Note: these hypotheses are a consequence of testing the hypothesis that the distribution of forward/backward sway for the elderly is the same as that for the young versus the distributions are the same shape, but centred at different locations.

We have some to strong evidence (*P-value* ≈ 3%) of a difference in the median amount of forward/backward sway experienced by "elderly" and "young" groups. With 95% confidence the median amount of forward/backward sway experienced by an "elderly" group is between 2 and 13 millimetres more than that experienced by a "young" group.

(b) The plots show that "elderly" data is centred higher than the "young" data. Both data sets appear to be slightly skewed to the right. Also the "elderly" group has an outlier at 50 units. Due to the outlier and the small size of both samples a 2 independent sample *t*-test would not be ideal here, so the Mann-Whitney test is the most appropriate test for this data.

By hand:



Back to back stem and leaf plot
of Forwad/Backward sway by Age Group

```
Young      Elderly
  4  4 | 1 |
7  7  5 | 1 | 9  9
  2  1 | 2 | 0  1  4
     5 | 2 | 5  9
       | 3 | 0
       | 3 |
       | 4 |
       | 4 |
       | 5 | 0
```

Units: 2| 0 =20

Minitab output:

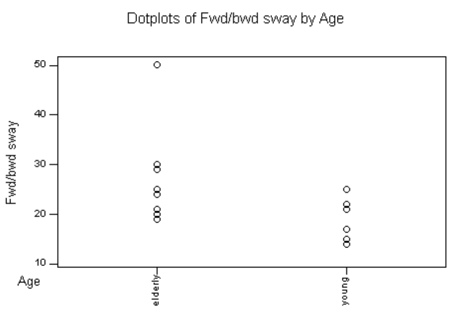
Character Stem-and-Leaf Display

Stem-and-leaf of elderly N = 9
Leaf Unit = 1.0

```
  2    1 99
(3)    2 014
  4    2 59
  2    3 0
  1    3
  1    4
  1    4
  1    5 0
```

Stem-and-leaf of young - N = 8
Leaf Unit = 1.0

```
  3    1 445
(2)    1 77
  3    1
  3    2 1
  2    2 2
  1    2 5
```



Question 4.

(a) From the dotplots we can see that fabric 1 appears to have longer burn times than the other fabrics. There does not seem to be large differences between the burn times for fabrics 2, 3 and 4. There is a possible high outlier for fabric 2. Except for the effect of the outlier, there does not appear to be much difference in variability between the burn times for the different fabrics.

(b) Boxplots would not be appropriate here as each sample only has 5 observations.

(c) The outcome of one test should not have any effect on the outcome of another test. They are independent of each other. This satisfies the independence assumption for the *F*-test.

(d) The highest standard deviation is 2.33 for fabric 2. The lowest is 1.144 for fabric 3. The ratio of the highest standard deviation to the lowest is 2.33/1.144 = 2.04.

(e) There are some doubts about the validity of the *F*-test. The outlier in the Fabric 2 group shows problems with the Normality assumption. The difference in variability between the groups is also at the extreme end of what is acceptable, even with identical sample sizes. (Though this difference is also due to the outlier in the fabric 2 group.) There are no problems with the independence assumption.

(f) The non-parametric alternative to the *f*-test is the Kruskal-Wallis test.

(g) **(i)** Let μ_1 be the mean burn time for fabric 1, similarly define μ_2, μ_3 and μ_4 for fabrics 2, 3 and 4. $H_0 : \mu_1 = \mu_4 = \mu_3 = \mu_4$

H_0 : The mean burn time is the same for the four different fabrics.

(ii) H_1 : The mean burn time is different for at least two of the four different fabrics.

(iii) As the P-value from the F-test is 0.000, we have extremely strong evidence against the null hypothesis. Thus, we have extremely strong evidence that the mean burn time is different for at least two of the four different fabrics.

(h) **(i)** Fabric 1 has the highest average burn time. (Looking at the intervals, we see that there is evidence that the burn time for fabric 1 is higher than that for each of the other three fabrics.)

(ii) Fabrics 2, 3 and 4 have the lowest average burn times. (Note: as all the intervals for pairwise differences between fabrics 2, 3 and 4 contain 0 we do not have evidence that any one of the fabrics leads to a lower average burn time than any other, so we must state all three fabrics.

(iii) We estimate, with 95% confidence, that the mean burn time for fabric 2 is between 3.3 seconds lower and 2.86 seconds higher than the mean burn time for fabric 4.

Question 5.

(a) The most appropriate design for this experiment is to use two independent samples. Randomly allocate half of the households to be surveyed with the promise of one $500 prize. The remaining households get the same survey but with the promise of five $100 prizes. The hypotheses we are interested in testing are H_0 : there is no difference in the proportion of respondents to surveys between surveys promising a chance at one $500 prize and surveys promising five $100 prizes. The alternative hypothesis would be that there is a difference in the two proportions.

(b) The most appropriate design for this experiment is a paired design in order to attempt to cut down on the variability (which might otherwise overshadow any difference in means) between the drivers. We have 20 drivers and 20 cars available. The cars should be randomly allocated to drivers. Half of the cars (chosen randomly) should be fitted with the first type of carburettor and the remaining cars fitted with the other type. After two weeks, the fuel consumption should be measured and then the carburettors changed in each car to the other type. After the second two weeks the fuel consumption should be measured again. The hypotheses we are interested in testing are $H_0 : \mu_{\text{diff}} = 0$ versus $H_1 : \mu_{\text{diff}} \neq 0$, where μ_{diff} is the mean difference in fuel consumption when using the different carburettors.