Research Methods 2

Week 10: Exercise Sheet 2

Solution sheet

Question 1

The results of applying the paired *t*-test to compare the admission and one-hour means is reproduced from the study document below.

Paired T-Test and CI: Admission, one_hour

```
Paired T for Admission - one_hour

                 N      Mean     StDev   SE Mean
Admission       10     129.0      63.8      20.2
one_hour        10     120.0      72.0      22.8
Difference      10      9.00     26.01      8.23

95% CI for mean difference: (-9.61, 27.61)
T-Test of mean difference = 0 (vs not = 0): T-Value = 1.09  P-Value = 0.302
```

Repeating the exercise but using the re-arranged one-hour PEFR values gives the following output

Paired T-Test and CI: Admission, one_hour (re-arr.)

```
Paired T for Admission - one_hour (re-arr.)

                 N      Mean     StDev   SE Mean
Admission       10     129.0      63.8      20.2
one_hour (re    10     120.0      72.0      22.8
Difference      10       9.0      79.8      25.2

95% CI for mean difference: (-48.1, 66.1)
T-Test of mean difference = 0 (vs not = 0): T-Value = 0.36  P-Value = 0.730
```

Some aspects of the output are the same. The means, SDs and SEs of the individual admission and one hour PEFRs are the same, regardless of the ordering applied to the latter. However, these quantities for the differences are the same only for the mean. The SD of the differences is much smaller for the genuine difference than for that based on rearranged data. A consequence of this is that the confidence interval based on the genuine data is narrower than that based on the re-arranged data. Also the P-value is different in the two cases.

This illustrates that when the correct analysis is applied it does make a difference if you ignore the structure of the data.

It was noted above that the SD of the differences is smaller for the genuine differences. A possible reason for this, which often obtains in practice, is that people will differ between one another in their PEFR (perhaps because of varying pain tolerances, body sizes and intrinsic quality of their lung function) much more than they differ between different measurement occasions. The act of taking the (correct) difference eliminates this source of variation.

Question 2

Because the each patient gives a score on each drug, the analysis should be based on a paired *t*-test. This assertion assumes that the differences (score on drug A – score on drug B) have a Normal distribution.

A problem with this design is that the order of the treatments is always the same. A consequence is that changes in the nausea scores from week 4 to week 6 may be due to other factors than the change in the treatment, such as the cumulative effects of the chemotherapy. Consequently, while the analysis will be able to tell you if there is a difference or not, you may be unable to infer what has caused the difference. A more satisfactory design (which it self has some problems) is the crossover design, in which some patients receive the anti-emetics in one order and the other patients receive them in the other order. See section 2.6 of Bland.

End of solution sheet