Homework Assignment 5
Solutions

1. (a) The p-value is larger than 5%, so we cannot reject the null hypothesis that the difference in treatment responses is zero. However, this does not mean that the population means are necessarily the same: maybe we did not have enough power to detect this difference. In fact, a p-value of 0.057 means that there is a bit of evidence in the data that the null might not be true. We should get some more data!

(b) P-values are derived from data, and therefore are random quantities. That’s why my friend should not get exactly the same p-value. Whether or not my friend gets roughly the same p-value depends on a few things. For example, if there was no difference in the population means in truth, we both should have gotten draws from a Uniform(0,1) distribution, and therefore my friend might get a totally different p-value. If the null was false, then I would expect my friend to get a somewhat small p-value as well. However, a total of 20 samples is not an awful lot, so we expect that the standard deviation of the distribution of the test statistic, even under the alternative, is somewhat large (unless we have very precise measurements). Therefore, my friend may or may not get a similar p-value.

[3 points]

2. Please also see the code.

(a) We assume that the variance of the observations in the case group is the same as in the control group of healthy individuals.

(b) The standard deviation of the measurements in the 10 healthy individuals is 3.4, and is used in the power calculations. With `power.t.test(n=40, sd=3.4, power=0.8, sig.level=0.01)` you can find in the output that `delta = 2.655605`, so about 2.66. This is the true difference in population mean lipid levels between cases and controls you can detect with 80% power at a 1% significance level in the hypothesis test.

(c) The detectable difference in means is 3.00.

(d) The detectable difference in means is 2.16.

(e) The detectable difference in means is 2.50.

(f) When your desired power increases, the detectable difference in means has to increase, all other parameters being the same. When your desired type I error rate decreases for a fixed power level, the detectable difference in means has to increase. You are being more stringent about making a type I error, thus, you need a stronger signal to surpass this stringency.

[4 points]

3. Please also see the code.

(a) Only one hypothesis test would be considered significant when controlling the FWER at 5% using the Bonferroni correction (the one with `p=0.003`).
(b) When controlling the FDR at 5% with Benjamini-Hochberg, four brain regions would be called significant (the ones with \( p < 0.016 \)).

[3 points]

4. (a) The standard error for the difference in sample means will be \( \sqrt{\frac{\sigma^2}{n} + \frac{\sigma^2}{n}} = \sigma \sqrt{\frac{2}{n}} \). Thus, the 95% confidence interval will have a margin of error of \( 1.96 \times \sigma \sqrt{\frac{2}{n}} \). For this margin of error to be less than one, we have

\[
1.96 \times \sigma \sqrt{\frac{2}{n}} < 1 \iff \sqrt{n} > 1.96 \times \sigma \sqrt{2} \iff n > (1.96)^2 \times 2\sigma^2 = 3.84 \times 2 \times 100 = 768.32.
\]

Thus, we need 769 subjects per group!

(b) Using the formula from class, the power is \( \approx P\{Z > 1.96 - \frac{(\Delta \sqrt{n})}{(\sigma \sqrt{2})}\} \) where \( Z \) is a standard Gaussian distribution. For 50% power, we thus need \( 1.96 - \frac{(\Delta \sqrt{n})}{(\sigma \sqrt{2})} = 0 \) since \( P\{Z > 0\} = 0.5 \). Solving this equation for \( n \) yields

\[
\sqrt{n} > 1.96 \times \sigma \sqrt{2}/\Delta \iff n > (1.96)^2 \times 2 \times (\sigma/\Delta)^2 = 3.84 \times 2 \times 25 = 192.08.
\]

Thus, we need at least 193 subjects per group.

[6 points]