# STA305/1004 - 2017 Homework #2

**Due: Electronic submission on UofT Learning Portal course page by Friday, February 17, 2017 at 22:00. NB: e-mail submissions will NOT be accepted.**

**If you work with other students on this assignment then:**

- indicate the names of the students on your solutions;
- your solutions must be written up independently (i.e., your solutions should not be the same as another students solutions).

## Question #1

In a noninferiority trial, the new treatment is tested to investigate if it is at least similar to an existing therapy in terms of efficacy. In other words, a noninferiority trial is interested in examining whether a new treatment is not worse than the standard treatment by a prespecified noninferiority margin $\delta < 0$. The threshold $\delta$ specifies the lower bound beyond which the experimental drug is considered unacceptably inferior to the standard drug. The hypothesis test of interest is

$$H_0 : \theta \leq \delta \text{ versus } H_1 : \theta > \delta,$$

where, $\theta = \mu_1 - \mu_2$, the difference in means between the experimental treatment (treatment 1) and the standard of care (treatment 2). Rejecting $H_0$ means that the experimental treatment is claimed to be noninferior to (not worse than) the standard treatment. For example if $\theta = 0$ and $\delta = -0.5$ then the alternative hypothesis states that the experimental treatment is noninferior to the standard treatment if the mean difference of the outcome in the experimental versus the standard arm is at least -0.5.

Assume that the outcomes of a noninferiority trial are $Y_{i1} \sim N\left(\mu_1, \sigma^2\right), i = 1, \ldots, n_1$ in the group receiving the experimental treatment with, and the outcomes in the group receiving the standard treatment are $Y_{i2} \sim N\left(\mu_2, \sigma^2\right), i = 1, \ldots, n_2$. In the following questions assume that $\sigma^2$ is known.

(a) Specify the values of the test statistic for which the null hypothesis is rejected (i.e., the rejection region of the test) at level $\alpha$.

(b) Show that the power of this test is given by

$$\Phi\left(\left(\frac{\theta - \delta}{\sigma\sqrt{1/n_1 + 1/n_2}}\right) - z_\alpha\right),$$

for $\theta > \delta$, where $\Phi(\cdot)$ is the CDF, and $z_\alpha$ is the $100(1-\alpha)^{th}$ percentile of the $N(0,1)$.

(c) If the numbers of patients in the two groups are different with $n_1 = rn_2$ then show that

$$n_2 = \frac{(1 + 1/r)\sigma^2(z_\alpha + z_\beta)^2}{(\theta - \delta)^2},$$

where $\beta$ is the probability of a type II error, and $\theta > \delta$. (HINT: Use the power function from the previous part)

(d) A pharmaceutical company is interested in establishing non-inferiority of a test drug compared to the standard treatment. The primary efficacy parameter is the percent change (from the beginning of the study) in LDL - low density lipidproteins. The company considers a difference of 0.05 to be clinically important, and assumes that the true difference in mean LDL between treatment groups is 0.0, and the standard deviation is 0.1. An equal number of subjects will be recruited into both treatment groups. How many subjects should be recruited into each treatment group so that the study has 90% power at the 5% significance level? Use R to calculate the answer using the formula in the previous part of the question. Hand in your R code

## Question #2

An engineer would like to design a study to compare the average lifetime of a certain electronic component under two different conditions. Components will be randomized to the two different conditions A and B. The distribution of the component's lifetime is known to follow an exponential distribution with rate $\lambda$. The density function of this distribution is

$$f(x) = \lambda \exp\left(-\lambda x\right), x \geq 0.$$

The distribution depends on a single parameter $\lambda > 0$. The mean and standard deviation of the exponential distribution is $\lambda^{-1}$. This parameterization is in units corresponding to the reciprocal of time.

The engineer plans to compare the mean lifetimes between the two groups. She has hypothesized that the expected lifetime of the components under condition A is 10.00 months, and 6.67 months under condition B.

(a) If 100 batteries are tested in each group then what is the power of the two sample t-test to detect a difference in means at the 5% significance level? (HINT: A random sample from an exponential distribution can be generated in R using the function `rexp(n = ,rate = )`.) Hand in your R code.

(b) The engineer contacts a statistician to check if her experiment is properly designed. The engineer tells the statistician her plan. The statistician quickly points out that one of the major assumptions of the t-test is that the data are assumed to be independent samples from normal distributions. She suggests using a nonparametric method called the Mann-Whitney Test (also called the Wilcoxon Rank sum test). This test doesn't assume that the data follow any particular distribution form. This test is implemented in R via the function `wilcox.test()`. For example, the P-value of the Mann-Whitney test for comparing the means of two random samples of size 50 from exponential distributions with $\lambda_1 = 0.1$, and $\lambda_2 = 0.2$ can be calculated in R.

```
x <- rexp(n = 50,rate = 0.1)
y <- rexp(n = 50,rate = 0.2)
wilcox.test(x,y)$p.value # the P-value from the test
```

```
## [1] 0.003583888
```

Calculate the power for the engineer's study if she uses the Mann-Whitney test to compare means. Hand in your R code.

(c) Which statistical test should the engineer use to compare mean lifetimes under the two conditions? If she uses the test you recommend then will she be guaranteed that her experiment will detect a statistically significant difference between the two treatment groups? Explain your reasoning.

## Question #3

Consider the following hypothetical example. Suppose that eight patients were assigned to one of two medical operations by a physician: $Y(0)$ is the number of years lived after standard surgery; and $Y(1)$ is the number of years lived after new surgery. The hypothetical data given below shows all potential outcomes under the two different treatments.

| $Y(0)$ | $Y(1)$ |
|---|---|
| 13 | 14 |
| 6 | 0 |
| 4 | 1 |
| 5 | 2 |
| 6 | 3 |
| 6 | 1 |
| 8 | 10 |
| 8 | 9 |

The doctor assigns a patient to the surgical treatment (new or standard) such that the patient will maximize years their lived after surgery, and if there is no difference he flips a coin.

(a) Is the assignment mechanism that the doctor used ignorable? Explain your answer.

(b) What is the true average causal effect of treatment? Which treatment is better on average? Explain your answer.

(c) If we only observed the outcomes under the treatment that the doctor assigned then what is the observed treatment effect? Is it different compared to the true average causal treatment effect?

(d) The data that you used to calculate the true treatment effect in (b) can be used to calculate the mean difference in all possible treatment assignments (the randomization distribution). If the treatment assignment would have been chosen at random then would the inference based on the difference of means be correct? Compare the latter to how the doctor assigned treatment. Is it misleading to use the difference in means to compare the treatments using the doctor's treatment assignment? Explain your reasoning.