Homework 1

Instructions: This homework is due in class on Friday Apr 12.

Please read and follow the guidelines for presenting your work.

- Write your full name clearly on the top right of the first page. **Staple** pages on the left hand corner. Write neatly in complete sentences.
- You are required to work all the problems, however, only 5 will be graded. The page numbers below refer to the fifth edition of the text.
- Answer the questions in the order in which they are posed. Clearly number the questions as I have.
- You must work independently on the homework. Please post questions on the discussion board or come to office hours.
- Be sure to show/explain your work thoughtfully. How you write your answers is important.
- If you use R to make plots or as a calculator, it is enough to simply include the output (e.g., appropriately labeled plot) in the main part of your homework without the R code.

---

1. A machine stamps out a certain type of automobile part. When working properly the part has an average weight of $\mu = 1.6$ pounds and standard deviation $\sigma = 0.22$ pounds. To test the proper working of the machine, quality control staff take forty of the parts and weigh them. They will reject the hypothesis that the machine is working properly (in other words they will test $H_0 : \mu \leq 1.6$ versus $H_1 : \mu > 1.6$) if the average weight $\bar{y} \geq 1.67$.
   (a) What is the significance level of the test?
   (b) What is β, the probability of a Type II error of this test when $\mu = 1.68$?
   (c) How many parts should they sample to have a 90% power of detecting a significant result with a significance level of $\alpha = 0.05$?
2. Larsen & Marx 378: 6.4.10
3. Larsen & Marx Page 379: 6.4.20
4. For the one-sided test of the mean μ of a normal with known variance σ^2:

$$\begin{aligned} H_0 : &\quad \mu \leq 100, \\ H_1 : &\quad \mu > 100. \end{aligned}$$

As we discussed in class, the power curve for a size $\alpha = 0.05$ test is given by the expression:

$$P_{H_1}\left(\bar{X} \geq 100 + 1.645 \times \frac{\sigma}{\sqrt{n}}\right).$$

(a) Draw the power curve for the test when

i. $n = 25$ and $\sigma = 14$.

ii. $n = 100$ and $\sigma = 14$

iii. $n = 25$ and $sigma = 28$

Sample code is given below. Please tinker with it – figure out how to add legends for what the colors represent; also if you are not printing in color, use different line types to show the three power curves.

(b) Write a couple of sentences describing what you notice about the effect of n and σ on the power.

5. Suppose a sample of size 3 is taken from the PDF

$$f_Y(y;\theta) \;=\; 2y/\theta^2, \quad 0 \leq y \leq \theta.$$

The largest order statistic Y_3' will be used to conduct a test of

$$H_0 : \theta \leq 5$$

versus

$$H_1 : \theta > 5.$$

(a) Find the critical value to give a test of significance level $\alpha = 0.05$.

(b) Suppose $\theta = 7$. What is the Type II error rate at this value for the test in part (a)?

6. (From the "Fundamentals of Biostatistics by Rosner, B.) An investigator wants to test a new eye drop that is supposed to prevent ocular itching. To study the drug, she uses a contralateral design, whereby for each participant one eye is randomized to get active drug (A), while the other eye gets a placebo (P). The participants use eye drops three times a day for 1 week and report their degree of itching on a four point scale. Ten subjects are randomized into the study.

(a) What is the principal advantage of the contralateral design?

(b) What is the principal advantage of comparing the active drug with a placebo (as opposed to no drug)?

(c) Suppose the randomization assignment is as given in the table.

	Eye			Eye	
Subject	L	R	Subject	L	R
1	A	P	6	A	P
2	P	A	7	A	P
3	A	P	8	P	A
4	A	P	9	A	P
5	P	A	10	A	P

More left (L) eyes seem to be assigned to A than to P and the investigator wonders if the assignment is truly random. Perform a significance test to assess how well the randomization is working. Choose an appropriate significance level and you may use R to calculate binomial probabilities.

```r
##drawing the power curve for problem 4.

install.packages("ggplot2")
library(ggplot2)

mu<-seq(100,110,0.01)      #pick mu values in H_1 to calculate power at
n=25
sigma=14
k=100+1.645*sigma/sqrt(n)      #critical value for size alpha=0.05 test
powerfn1 <- 1-pnorm(sqrt(n)*(k-mu)/sigma)

n=100                                #recalculate for n=100, sigma=14
sigma=14
k=100+1.645*sigma/sqrt(n)
powerfn2<-1-pnorm(sqrt(n)*(k-mu)/sigma)

n=25                                     #recalculate for n=25,sigma=28
sigma=28
k=100+1.645*sigma/sqrt(n)
powerfn3<-1-pnorm(sqrt(n)*(k-mu)/sigma)

powercomp <- data.frame(mu,powerfn1,powerfn2,powerfn3)

powercurve <- ggplot() + geom_line(data=powercomp, aes(x=mu, y=powerfn1,
                                                       color="red")) +
    geom_line(data=powercomp, aes(x=mu, y=powerfn2, color="blue")) +
      geom_line(data=powercomp, aes(x=mu, y=powerfn3, color="green")) +
    ggtitle("Power curve for normal mean
            with known variance") + xlab("Mean mu") + ylab("Power")

print(powercurve)
```