

## Math 566 - Homework 4

1. We want to test a hypothesis involving a population proportion. The unknown population proportion is  $p$ . The null hypothesis is  $p = 1/2$  and the alternative hypothesis is  $p \neq 1/2$ . We want the level of the test to be 0.01, and we want the power of the test to be at least 0.9 when  $p > 0.55$ . Determine how large a sample we need and specify what the test is, i.e., when we accept the null hypothesis and when we reject it. Note that we worked out the power function for the case of the alternative hypothesis being  $p > 1/2$  in class. (It is also worked out in the book.) You'll have to work it out in the case of this two sided alternative.
2. Problem 1 in chapter 5 in the book.
3. Problem 2 in chapter 5 in the book.
4. Problem 3 in chapter 5 in the book. If you are using R you will find the following functions useful. `pnorm()` computes the c.d.f. of a normal distribution. `sort()` puts a sample in increasing order. `(1:n-0.5)/n` will create an array with the numbers  $(i - 0.5)/n$  for  $i = 1, 2, \dots, n$ .
5. Suppose  $X$  (the population) is a continuous random variable with probability density function

$$f(x) = \frac{\lambda}{2} \exp(-\lambda|x - \mu|), \quad -\infty < x < \infty \quad (1)$$

where  $\lambda$  and  $\mu$  are unknown parameters. The mean of  $X$  is  $\mu$  and the density is symmetric about  $x = \mu$ , so the median is also  $\mu$ . We are given a random sample  $X_1, X_2, \dots, X_n$  with  $n$  large. Determine whether the sample mean or the sample median is a better estimator for  $\mu$ . By "better" I mean "has smaller variance."

6. The random variable  $X$  is uniformly distributed on the interval  $[0, \theta]$ . (This is the population.)  $\theta$  is an unknown parameter. We have a random sample  $X_1$  of size 1. We want to use it to estimate the unknown parameter  $\theta$ . Consider estimators of the form  $T = cX_1$  where  $c$  is a constant.
  - (a) Find the value of  $c$  which makes this an unbiased estimator.
  - (b) Find the value of  $c$  which minimizes the mean square error. This is the risk when we take the loss function to be  $(T - \theta)^2$ .

Now suppose we have a random sample  $X_1, \dots, X_n$  of size  $n$ . Recall that the order statistic  $X_{(n)}$  is just the maximum of all the  $X_1, \dots, X_n$ . We consider estimators of the form  $T = cX_{(n)}$ .

- (c) Find the value of  $c$  which makes this an unbiased estimator.
- (d) Find the value of  $c$  which minimizes the mean square error. This is the risk when we take the loss function to be  $(T - \theta)^2$ .
- Hint: A previous problem gives the mean of  $X_{(n)}$  in the case of  $\theta = 1$ . Section 5.6 of the book gives the variance. Feel free to use these results.