Homework 7
STA4322/STA5328
Introduction to Statistics Theory
Spring 2020, MWF 3:00pm
Due date: Friday, April 3rd, 2020

All work must be shown for complete credit. W.M.S. denotes the course textbook, *Mathematical Statistics with Applications*.

1. We want to test whether a coin is fair based on the number of heads in 40 tosses, $Y_1, \ldots, Y_{40}$, where $Y_i = 1$ if toss i was heads and $Y_i = 0$ otherwise. Specifically, we want to test $H_0 : p = 0.50$ vs $H_A : p \neq 0.50$. Suppose we use the rejection region $|\sum_{i=1}^{40} Y_i - 20| \geq K$. You may leave both answers in terms of sums

   (a) What is the type I error probability α if $K = 6$?

   (b) What is the value of β if $p = 0.7$ and $K = 6$?

2. Let X_1 and X_2 be independent and identically distributed with a uniform distribution on the interval $(\theta, \theta + 1)$. If we wanted to test $H_0 : \theta = 0$ vs $H_A : \theta > 0,$ we could consider hypothesis tests based on two rejection regions:

   Rejection region 1: $\{X_1 > 0.95\}$

   Rejection region 2: $\{X_1 + X_2 > K\}$

   Note that when H_0 is true, $W = X_1 + X_2$ has a special distribution known as the triangular distribution. The density of the triangular distribution is

$$f_W(w) = \begin{cases} w & 0 \leq w \leq 1 \\ 2 - w & 1 \leq w \leq 2 \end{cases}$$

   (a) What value of K would ensure that a test based on rejection region 1 has the same level α as a test based on rejection region 2?

   (b) For the K you computed in part (a), what is the power (i.e., $1 - \beta$) of each of the two tests if $\theta = 0.5$? Which test would you prefer? (Hint: recall that if $\theta = 0.50$, $X_1 - 0.5 \sim \text{Uniform}(0, 1)$ and $X_2 - 0.5 \sim \text{Uniform}(0, 1)$)

3. The wages (hourly) in a particular industry has are mandated to be normally distributed with mean \$13.20 and standard deviation \$2.50. One company in this industry employs 40 workers, and has an average hourly wage of \$12.20.

   (a) Is there sufficient evidence to suggest that this company is paying substandard wages at the $.01$ level?

(b) What is the probability that if we were to randomly select 40 workers from this industry, who were not being underpaid, that we would observe an average hourly wage of $12.20 or less? Leave your answer in terms of the distribution function of the standard normal distribution.

4. A manufacturer creates three products denoted A, B, and C. Of the first 1000 products sold (assume sales are independent and identically distributed), 400 were product A. Would you conclude that customers have a preference for product A? Justify your answer via a hypothesis test with level $.01$.

5. A large-sample α-level hypothesis test for $H_0 : \theta = \theta_0$ vs $H_A : \theta < \theta_0$ rejects H_0 if

$$\frac{\hat{\theta} - \theta_0}{\hat{\mathrm{SE}}(\hat{\theta})} < -z_{1-\alpha}.$$

Show that this is equivalent to rejecting H_0 if θ_0 is greater than the large sample $100(1-\alpha)$% upper confidence bound for θ.

6. Suppose that $U_1, \ldots, U_n$ are a random sample distributed uniformly on the interval $(0, \theta)$. Let $U_{(n)} = \max(U_1, \ldots, U_n)$ be the test statistic. We want to test the hypotheses $H_0 : \theta = 2$ vs $H_A : \theta > 2$.

   (a) If we use a rejection region of the form

   $$\left\{U_{(n)} > K\right\},$$

   determine K such that the hypothesis test has exactly level-α. It may be helpful to recall that $U_{(n)}/\theta$ is a pivot.

   (b) Based on the K you have chosen in (a), compute the power of your test $(1-\beta)$ if $\theta = 3$.

7. A federal agency collected a random sample of 500 measurements of the length of "long term" hospital stays. The random sample had mean 5.4 days and standard devation 3.1 days. Beforehand, the agency had hypothesized that the average length of stay was greater than the reported average of 5 days.

   (a) Test whether there is sufficient evidence to support the alternative hypothesis at least $.01$.

   (b) Calculate β, the type II error probability, if $\mu = 5.5$

Computing standard normal quantiles

To compute standard normal quantiles or probabilities which may be needed in this homework assignment, you can use the statistical software R, which can be run through an internet browser here: https://rdrr.io/snippets/. Once here, to compute the, say, 0.975th quantile of the standard normal distribution, we would type

```
qnorm(0.975)
```

into the console and hit "Run". Then, we see the output

```
[1] 1.959964
```

which we already knew since $P(-1.96 < Z < 1.96) \approx 0.95$. For more exotic quantiles which we do not know off the topic of our heads, this may be helpful.

Conversely, we may want to evaluate the standard normal distribution at some particular quantile. Suppose we want to compute $P(Z < 1.224) \equiv \Phi(1.224)$: we would enter

```
pnorm(1.224)
```

into the console and hit "Run". We would then get the output

```
[1] 0.8895239
```

which tells us that

$$P(Z < 1.224) \equiv \Phi(1.224) = 0.8895239.$$