STAT 422: Homework #7

Elliott Pryor

Problem 1

Statement

Section 8.9 in your text describes how to create a 100(1-α)% CI for σ^2.

a. Construct a 100(1-α)% CI for the sample standard deviation, σ.

b. Now, suppose that you are interested in the variability in credit card debt among MSU students, and in a RS of 40 students, you observe a sample standard devaition of $s = 200$\$. Construct a 95% CI for σ and *interpret the interval in the context of the problem.*

c. Will a 90% CI for σ be wider or smaller than a 95% CI for σ? Explain. Assume the same sample size and sample standard deviation as in part b.

Solution

a. First we need an unbiased estimator for σ. We know that S^2 is an unbiased estimator or σ^2, so its square root S as unbiased estimator of σ. We know from Theorem 7.3 that $\frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1}$, we take the square root of both sides to find that $\frac{\sqrt{n-1}S}{\sigma} \sim \chi_{n-1}$. We then need to follow the pivotal method to find two numbers χ_L, χ_U such that

$$P\left[\chi_L \leq \frac{\sqrt{n-1}S}{\sigma} \leq \chi_U\right] = 1 - \alpha$$

There are many combinations for χ_L, χ_U, we choose the one that removes equal areas on left and right ie. $P(X \leq \chi_L) = P(X \geq \chi_U) = \alpha/2$ if $X \sim \chi_{n-1}$. This corresponds to picking χ_L to be the α-quantile of χ_{n-1} and χ_U to be the $(1-\alpha)$-quantile of χ_{n-1}.These selections for χ_L, χ_U are denoted $\chi_\alpha, \chi_{1-\alpha}$ respectively. Then we re-arrange the formula above to solve for σ.

$$P\left[\chi_\alpha \leq \frac{\sqrt{n-1}S}{\sigma} \leq \chi_{1-\alpha}\right] = 1 - \alpha$$

$$P\left[\frac{\sqrt{n-1}S}{\chi_{1-\alpha}} \leq \sigma \leq \frac{\sqrt{n-1}S}{\chi_\alpha}\right] = 1 - \alpha$$

Is a $100(1-\alpha)$% CI for σ.

```r
# Example of how to compute Chi_l and Chi_u
library(chi)  # for the chi distribution
n = 5  # for example
alpha = 0.05
result <- qchi(p=c(alpha, 1-alpha), df = n)
chi_l = result[1]
chi_u = result[2]
print(chi_l)
```

```
## [1] 1.070269
```

```r
print(chi_u)
```

```
## [1] 3.327236
```

b. So first we need to compute χ_L, χ_U with the following r script

```r
# Example of how to compute Chi_l and Chi_u
n = 40  # 40 students
alpha = 0.05
result <- qchi(p=c(alpha, 1-alpha), df = n)
chi_l = result[1]
chi_u = result[2]
print(chi_l)
```

```
## [1] 5.148719
```

```r
print(chi_u)
```

```
## [1] 7.46716
```

Then we compute the bounds:

```r
S = 200
lower_bound = (sqrt(n - 1) * S) / chi_u
upper_bound = (sqrt(n - 1) * S) / chi_l
print(paste0("The 95% confidence interval for sigma is ", lower_bound, " <= sigma <= ",
```

```
## [1] "The 95% confidence interval for sigma is 167.265679397627 <= sigma <= 242.584553
```

c. The interval for 90% will be more narrow than 95%. If we think about it, we cut off $\alpha/2$ probability on both ends. If α is bigger, then we cut out more of a region and thus the region we keep is smaller.

Problems related to power and power functions

1. **Graded for accuracy.** An experimenter has prepared a drug dosage level that she claims will induce sleep for 80% of people suffering from insomnia. After examining the dosage, we think that her claims regarding the effectiveness of the dosage are inflated.

In an attempt to provide statistical evidence against her claim, we administer her prescribed dosage to 20 insomniacs, and we observe Y, the number for which the drug dose induces sleep. We wish to test the hypothesis $H_0 : p = 0.8$ versus the alternative $H_a : p > 0.8$ Assume that the rejection region $\{y \leq 12\}$ is used.

a. In terms of this problem, what is a Type I error?

b. Find the significance level,α of the test.

c. In terms of this problem, what is a Type II error?

d. Find the probability of a Type II error when $p = 0.6$

e. Find the probability of a Type II error when $p = 0.4$

f. Find the power for the following values of p

- $p = 0.4$
- $p = 0.5$
- $p = 0.6$
- $p = 0.7$

g. Graph the power function in R. Provide the graph as well as the code used to produce it.

h. What is the power when $p = 0.8$? Why is this the same value as your answer to part (b)

i. What would happen to the power at $p = 0.8$ if we change the rejection region to $\{y \leq 13\}$? What would happen to the power at $p = 0.8$ if we change the rejection region to $\{y \leq 11\}$? Briefly explain why each of these changes occurrs.

2. A sample of size 1 is taken from an exponential(θ) distribution. To test $H_0 : \theta = 1$ against $H_a : \theta > 1$, the test to be used is $\zeta(X) = \begin{cases} 1 & X > 2 \\ 0 & X \leq 2 \end{cases}$

a. Find the significance level of the test.
b. Derive the power function.

Problems related to hypothesis tests

1. Let $X_1, ..., X_n$ be a random sample from the Gamma(α,β) distribution.

a. Assuming that α (the *shape* parameter in the distriubtion, not the alpha that defines the Type I error rate) is known, derive the likelihood ratio test (LRT) for testing $H_0 : \beta = \beta_0$ versus $H_a : \beta \neq \beta_0$. Use the large-sample approximation for the LRT to find the critical value for the test.
b. The Kew Observatory in West London has been collecting rainfall data since around 1700. For 1700-1950, the July rainfall (in mm) can be modeled using a Gamma distribution with $\alpha = 5$ and $\beta = 12.2$. For 1991-2008, the data are assumed to follow a Gamma distribution with $\alpha = 5$, but researchers suspect that

β may have changed. For 1991-2008 (18 years), the average July rainfall was 44 mm. Assuming that $\alpha = 5$, do these data indicate that the β parameter is no longer 12.2? Use the test you derived in part (a) to test these hypotheses at the 0.01 significance level. Be sure to also find the corresponding $p - value$ for the study and write a conclusion beyond reject H_0 or fail to reject H_0.

2. Suppose that X is one observation from the distribution with pdf $f(x;\lambda) = \lambda e^{-\lambda x} I_{(0,\infty)}(x)$ and $\lambda > 0$. Consider testing $H_0 : \lambda = 1$ versus $H_a : \lambda \neq 1$. Let $\zeta(X)$ be the critical function

$$\zeta(X) = \begin{cases} 1 & x < 1/2 \text{ or } x > 2 \\ 0 & else \end{cases}$$

   a. Find the power function of the test corresponding to ζ.
   b. What is the significance level of the test determined by ζ?

3. In a given city it is assumed that the number of automobile accidents in a given year follows a Poisson distribution. In past years the average number of accidents per year was 15, and this year the number of accidents was 10.

   a. Is it justified to claim that the accident rate has dropped? Justify your answer using sound statistical evidence and reasoning (i.e., find a *p-value* for the test and interpret the *p-value* in the context of the problem).
   b. What else should you provide to complete your summary of statistical evidence? Briefly describe how you could obtain this information.

4. A sample of size 1 is taken from a Poisson(λ) distribution. To test $H_0 : \lambda = 1$ against $H_a : \lambda = 2$, consider the test:

$$\phi(X) = \begin{cases} 1 & X > 3 \\ 0 & X \leq 3 \end{cases}$$

   a. What is the probability of making a Type I error?
   b. What is the probability of making a Type II error?