Homework 2

Spring 2023

Tanner Huck

Instructions

- This homework is due in Gradescope on Wednesday April 19 by midnight PST.
- Please answer the following questions in the order in which they are posed.
- Don't forget to knit the document frequently to make sure there are no compilation errors.
- When you are done, download the PDF file as instructed in section and submit it in Gradescope.

Exercises

1. (Simulation) From problem session 1, we saw from our robustness studies that the ratio

$$T = \frac{\sqrt{n}(\bar{X} - \mu_0)}{S}$$

is not well approximated by a t-distribution for small samples taken from $Exp(\lambda_0 = 2)$. Therefore, use of the t confidence interval formula $\bar{X} \pm t_{\alpha/2,n-1}\frac{S}{\sqrt{n}}$ in small samples where the data may have come from an exponential will likely not work very well.

In this problem, you will develop an alternate confidence interval formula for μ_0 which works for small samples. Specifically, suppose $X_1, X_2, \dots, X_n \overset{i.i.d.}{\sim} Exp(\lambda_0)$. Then it can be shown that (you do not have to show this):

$$\lambda_0 \, \bar{X} \sim Gamma(n, n).$$

a. Use the distribution of $\bar{X}$ to construct an *equal tailed* $100(1-\alpha)\%$ confidence interval for the mean $\mu_0 = \frac{1}{\lambda_0}$. Show your work and the formula explicitly. (You may leave q_1 and q_2 in terms of the R function that will be used to calculate them.)

Since we know that $\lambda_0 \, \bar{X} \sim Gamma(n, n)$,

$$P(q_1 \leq \lambda_0 \, \bar{X} \leq q_2) = 1 - \alpha$$
$$P(\frac{q_1}{\bar{X}} \leq \lambda_0 \leq \frac{q_2}{\bar{X}}) = 1 - \alpha$$
$$P(\frac{\bar{X}}{q_1} \leq \frac{1}{\lambda_0} \leq \frac{\bar{X}}{q_2}) = 1 - \alpha$$
$$P(\frac{\bar{X}}{q_1} \leq \mu_0 \leq \frac{\bar{X}}{q_2}) = 1 - \alpha$$

Hence the $100(1-\alpha)\%$ confidence interval for the mean μ_0 is

$$[\frac{\bar{X}}{\text{qgamma}(1-\alpha/2, n, n)}, \frac{\bar{X}}{\text{qgamma}(\alpha/2, n, n)}]$$

```
b.  Use simulations to verify the coverage probability of your interval from part a by following the st

    i. Generate $B = 10,000$ samples each of size 6 from an $Exp(\lambda_0 = 2)$ distribution. Please

    ii. Calculate the 95% $t$ confidence interval for each sample. Also calculate your confidence int

    iii. Calculate and report the coverage rates for the two confidence intervals across the 10,000 s

    iv. Perform a large sample significance test to evaluate whether the intervals from part iii. hav
```

In generating samples from a $Exp(\lambda_0 = 2)$ distribution, we can see that the coverage rates of the confidence intervals for μ_0 from the samples is about 89 percent for the t confidence interval and about 95 percent for the confidence interval from part a. We can also see that there are 1075 samples that lie below the t confidence interval and 40 samples that lie above. Similarly for the other confidence interval, 221 below and 238 above. We can see that from the confidence interval created in part a, there is about the same number of samples above and below the interval. However, for the t confidence interval there are many more missed samples that are below the interval compared to above. Finally, computing a significance test to evaluate if the intervals have a nominal coverage rate of 95% at a 0.01 alpha level. We get a p-value of about 3.5×10^{-175} for the t confidence interval and about 0.0599 for the confidence interval from part a. Since $3.5 \times 10^{-175} < 0.01$, we have enough evidence to reject the null hypothesis that the t confidence interval has a nominal coverage rate of 95%. Whereas for the confidence interval in part a, $0.0599 > 0.01$, meaning we do have have enough evidence to reject the null hypothesis that the confidence interval has a nominal coverage rate of 95%.

2. (Reproducing chi squares) Suppose $X \sim \chi^2_m$ independently of $Y \sim \chi^2_n$. That is, X has PDF

$$f_1(x) = \frac{(1/2)^{m/2}}{\Gamma(m/2)} x^{\frac{m}{2}-1} e^{-x/2}, \quad x > 0$$

and Y has PDF

$$f_2(y) = \frac{(1/2)^{n/2}}{\Gamma(n/2)} y^{\frac{n}{2}-1} e^{-y/2}, \quad x > 0$$

Let $S = X + Y$ be their sum. Show, using the method of convolution, that

$$S = X + Y \sim \chi^2_{m+n}.$$

That is, show that S has PDF

$$f(s) = \frac{(1/2)^{(m+n)/2}}{\Gamma((m+n)/2)} s^{(m+n)/2-1} e^{-s/2} \quad s > 0.$$

You may use without proof that

$$\frac{\Gamma(m)\,\Gamma(n)}{\Gamma(m+n)} = \int_0^1 x^{m-1}\,(1-x)^{n-1}\,dx$$

Hint: Denote the PDF of S by f Recall from chapter 16 that

$$f(s) = \int_0^s f_1(x)\,f_2(s-x)\,dx$$

Plug in for f_1 and f_2 and simplify.

If we let $S = X + Y$ where $X \sim \chi^2_m$ independently of $Y \sim \chi^2_n$, then using convolution we want to show that $S = X + Y \sim \chi^2_{m+n}$. Solving for the PDF of S,

$$f(s) = \int\limits_0^s f_1(x) f_2(s-x) dx$$

substituting in $f_1(x)$ and $f_2(y)$ with $s-x$ instead of y in f_2

$$= \int\limits_0^s \frac{(1/2)^{m/2}}{\Gamma(m/2)} x^{\frac{m}{2}-1} e^{-x/2} \times \frac{(1/2)^{n/2}}{\Gamma(n/2)} (s-x)^{\frac{n}{2}-1} e^{-(s-x)/2} dx$$

then simpliying

$$= \int\limits_0^s \frac{(1/2)^{(m+n)/2}}{\Gamma(m/2)\Gamma(n/2)} \times x^{\frac{m}{2}-1} e^{-x/2} \times (s-x)^{\frac{n}{2}-1} e^{-(s-x)/2} dx$$

$$= \frac{(1/2)^{(m+n)/2}}{\Gamma(m/2)\Gamma(n/2)} \int\limits_0^s x^{\frac{m}{2}-1} (s-x)^{\frac{n}{2}-1} e^{-s/2} dx$$

then making the substitution $x = s \cdot u$ and $dx = s \cdot du$

note when $x = s$ then $u = 1$ giving us our top bound

$$= \frac{(1/2)^{(m+n)/2}}{\Gamma(m/2)\Gamma(n/2)} \int_0^1 (us)^{\frac{m}{2}-1} (s-us)^{\frac{n}{2}-1} e^{-s/2} s du$$

$$= \frac{(1/2)^{(m+n)/2}}{\Gamma(m/2)\Gamma(n/2)} \int_0^1 u^{\frac{m}{2}-1} s^{\frac{m}{2}-1} (1-u)^{\frac{n}{2}-1} (s)^{\frac{n}{2}-1} e^{-s/2} s du$$

pulling out the constants

$$= \frac{(1/2)^{(m+n)/2}}{\Gamma(m/2)\Gamma(n/2)} \times s^{\frac{m}{2}-1} \times (s)^{\frac{n}{2}-1} \times e^{-s/2} \times s \int_0^1 u^{\frac{m}{2}-1} (1-u)^{\frac{n}{2}-1} du$$

then using the provided fact of $\dfrac{\Gamma(m)\Gamma(n)}{\Gamma(m+n)} = \displaystyle\int \ldots$

$$= \frac{(1/2)^{(m+n)/2}}{\Gamma(m/2)\Gamma(n/2)} \times s^{\frac{m}{2}-1} \times (s)^{\frac{n}{2}-1} \times e^{-s/2} \times s \times \frac{\Gamma(\frac{m}{2})\Gamma(\frac{n}{2})}{\Gamma(\frac{m}{2}+\frac{n}{2})}$$

combining like terms

$$= \frac{(1/2)^{(m+n)/2}}{\Gamma(\frac{m}{2}+\frac{n}{2})} s^{\frac{m}{2}+\frac{n}{2}-1} e^{-s/2}$$

$$= \frac{(1/2)^{(m+n)/2}}{\Gamma((m+n)/2)} s^{(m+n)/2-1} e^{-s/2}$$

Hence S has PDF $f(s) = \frac{(1/2)^{(m+n)/2}}{\Gamma((m+n)/2)} s^{(m+n)/2-1} e^{-s/2}$.

3. (Dark Matter) Two independent research teams claim to have discovered the elusive dark matter. They have used completely independent methods, and completely different statistical tests (although in both cases, rejecting the null hypothesis implies the discovery of dark matter). However, neither group has obtained a significant P-value, achieving 0.06 and 0.08, respectively. They want to combine their results somehow. Here are two facts:

   - When a null hypothesis is true, P-values follow a $Unif(0,1)$ distribution.
   - If $U \sim Unif(0,1)$ then $-2\ln(U) \sim \chi^2_2$.

   Knowing that the 95th percentile of a χ^2_4 distribution is 9.49, how would you suggest they combine their results?

Hint: you will need to also use what you learned from problem 2 to define a combined test statistic whose distribution you know under the null hypothesis. Then see whether the observed value of the test statistic provides a contradiction to this distribution.

We know that when the null hypothesis is true, the P-values follow a $Unif(0,1)$ distribution and if we let U be a random variable from this distribution, $U \sim Unif(0,1)$, than $-2\ln(U) \sim \chi^2_2$. When the null hypothesis is true, there is not sufficient evidence to prove the existence of dark matter. Whereas the alternative hypothesis is that we have enough evidence to prove the existence of dark matter. Let U_1 be the P-value obtained by team one and U_2 the P-value obtained by team 2. Since the two teams and results are completely independent of one another, we can combine their respective test statistics by adding them together, by what we showed in problem 2. More specifically, $S = U_1 + U_2 \sim \chi^2_{n+m}$. Hence the combine test statistic is,

$$T = -2\ln(U) - 2\ln(U) = -2\ln(0.06) - 2\ln(0.08) = 10.6783$$

Furthermore, we know that the 95th percentile of a χ^2_4 distribution (a chi squared distribution with the trials from team 1 and team 2 added together) is 9.49. Since $10.68 > 9.49$, at the 0.05 significance level, we have sufficient evidence to reject the null hypothesis. Hence, we can conclude that dark matter has been discovered.

4. (Blood pressure) An *Arterisonde machine* prints blood-pressure readings on a tape so that the measurement can be read rather than heard. A major argument for using such a machine is that the variability of measurements obtained by different observers on the same person will be lower than the variability with a standard blood-pressure cuff. From previously published work, the variance with a standard blood pressure cuff is $\sigma_0^2 = 35$.

Suppose we have data consisting of systolic blood pressure (SBP) measurements obtained on 10 people and read by two observers. We use the difference between the first and second observers to assess inter-observer variability. In particular, if we assume the underlying distribution of these **differences** is normal with mean μ_0 and variance σ_0^2, then it is of primary interest to make inference about σ_0^2.

The data is in the file `systolic.csv`. Calculate and interpret (in context) a 95% confidence interval for σ_0^2. (Even though investigators think the variability of the new method will be lower, we calculate a two sided confidence interval as the observers are less experienced in using it and this might result in an increase in the variability.)

Create a brief report (of sorts) where you include a description of the data and scientific problem, model/assumptions, the confidence interval, and a conclusion. (Put R code in the appendix.)

Description of Data Our Data consists of systolic blood pressure (SBP) measurements obtained on 10 people which is then read by two observers. From previous publications the variance with a standard blood pressure cuff is $\sigma_0^2 = 35$. In our new experiment, using an Arterisonde machine, the observer can take observations via printings rather than sound. We aim to assess the inter-observer variability, to guess the variance of the measurements obtained by the different observers.

Model/Assumptions We first assume that the underlying distribution of these differences is normal with a mean of μ_0 and variance of σ_0^2 (meaning the difference in observers 1 and 2's measurements in norm. dist.) Furthermore, if we let X_i represent the blood pressure reading from observer 1 and Y_i be the reading from observer 2, than $X_i - Y_i \sim Norm(\mu_0, \sigma_0)$. We are also assuming that the observations made by the observers are independent per subject and independent of the other observer. Meaning that any observation made by the other observer or any other observation about another subject will not effect any other observation. Finally, our null hypothesis is that the inter-observer variability is 35, $H_0 : \sigma^2 = 35$, which is telling us that our null hypothesis is that the variance of the difference in ratings with the machine is the same as the standard cuff. the alternative hypothesis is $H_1 : \sigma^2 \neq 35$, meaning the variance of the difference in ratings with the machine is different than the standard cuff.

Confidence Interval The 95 percent confidence interval for σ_0^2 is about $[3.87, 27.26]$ (see Appendix for code).

Conclusion This confidence interval tells us that we are 95 percent confident that the true variance of difference of the measurements obtained by the observers is in the range $[3.87, 27.26]$. This tells us that inter-observer variability may be as low as 3.87 with the new machine, but it may be as high as 27.26 as well. However, in comparison to the variance with a standard blood pressure cuff, the variability with the machine is lower. This means that it may be beneficial to use this new machine over a standard blood pressure cuff when the inter-observer variability is in question. More precisely, since 35 is not in our confidence interval, we have significant evidence to reject the null hypothesis that the variance is equal to 35, meaning that the variance with the machine is different than a standard cuff. Thus, the new machine has a inter-observer variability that is lower than a standard blood pressure cuff and may be better to use in the future.

Appendix

Problem 1 code

```
set.seed(544)

## i and ii code to generate samples of size 6 from Exp(lambda=2) and calculate the 95% confidence inte

B <- 10000
n <- 6
df <- lapply(X = 1:B, FUN = function(x){
  data = rexp(n, 2)
  data.frame(t_lower = mean(data) - qt(0.975, 5) * sd(data)/sqrt(n),
             t_upper = mean(data) + qt(0.975, 5) * sd(data)/sqrt(n),
             q_lower = mean(data) / qgamma(0.975, n, n),
             q_upper = mean(data) / qgamma(0.025, n, n))
})
df <- do.call(rbind, df)

## iii code to calculate coverage rates of the two intervals
t_coverage_rate <- length(which(df$t_lower <= 0.5 & df$t_upper >= 0.5))/B
t_above_interval <- length(which(df$t_low > 0.5))
t_below_interval <- length(which(df$t_upper < 0.5))

q_coverage_rate <- length(which(df$q_lower <= 0.5 & 0.5 <= df$q_upper))/B
q_above_interval <- length(which(df$q_low > 0.5))
q_below_interval <- length(which(df$q_upper < 0.5))

## code to calculate P-value for part iv.
t_pval <- 2*pnorm(t_coverage_rate, 0.95, sqrt(0.95*0.05/B))
q_pval <- 2*pnorm(q_coverage_rate, 0.95, sqrt(0.95*0.05/B), lower.tail = F)
```

Problem 4 code

```
data <- read.csv("systolic.csv")
observation_diff <- data$observer.1 - data$observer2
n <- 10
sample_vars <- var(observation_diff)
alpha <- 0.05
lower_bound <- (n-1) * sample_vars / qchisq(1-alpha/2, n-1)
upper_bound <- (n-1) * sample_vars / qchisq(alpha/2, n-1)
```

Hence the confidence interval is about [3.87,27.26].