BST 676 – Spring 2011 – Dr. Charnigo

Unit VI: Techniques for Interval Estimation

a. Motivating case study #1: Can an interval for a cancer rate contain a negative number?

Suppose that, over some fixed period of time, we observe 6 cancer cases in a sample of 500 individuals and that we want a 99% confidence interval for the proportion of individuals developing cancer in the population to which the sample generalizes. Noting that $z_{0.995} = 2.576$, we may consider using the familiar formula

$$\hat{p} \pm 2.576\sqrt{\hat{p}(1-\hat{p})/n}.$$

First we calculate that

$$n\hat{p}(1-\hat{p}) = 500(6/500)(494/500) = 5.928 > 5,$$

so that we in fact decide to proceed with the familiar formula.

We have $\hat{p} = 6/500 = 0.0120$, $n = 500$, and $2.576\sqrt{\hat{p}(1-\hat{p})/n} = 0.0125$, so that the familiar formula yields -0.0005 to 0.0245 as a 99% confidence interval for p.

Since a proportion cannot be a negative number, this confidence interval does not make sense!

To get some idea about what has gone wrong, consider what happens mathematically when the familiar formula yields a confidence interval that contains a negative number. We have

$$\hat{p} < z_{1-\alpha/2}\sqrt{\hat{p}(1-\hat{p})/n}$$

or, on squaring both sides and dividing by $\hat{p}$,

$$\hat{p} < \chi^2_{1,1-\alpha}(1-\hat{p})/n. \quad (1)$$

Assuming that $(1-\hat{p})$ is close to 1, relation (1) is approximated by

$$n\hat{p}(1-\hat{p}) < \chi^2_{1,1-\alpha}.$$

Thus, if we seek a 95% confidence interval for p, we do not have to worry about negative numbers with $n\hat{p}(1-\hat{p}) \geq 5$ since $\chi^2_{1,0.95} = 3.841$. However,

$n\hat{p}(1-\hat{p}) \geq 5$ is not good enough if we seek a 99% confidence interval since $\chi^2_{1,0.99} = 6.635$.

We will return to this case study, to identify some better methods for obtaining a 99% confidence interval in the presence of rare events, at the end of Unit VI.

b. Motivating case study #2: Constructing intervals from non-normal point estimators

Suppose that

$$X_1, X_2, X_3, X_4 \overset{iid}{\sim} f(x;\theta) := 1_{\{x>0\}}\theta^{-1}\exp[-x\theta^{-1}]$$

for $\theta \in \Theta := (0,\infty)$. Since $\bar{X}$ is the maximum likelihood estimator of θ, we may consider constructing a confidence interval of the form $\bar{X} \pm$ "something", where the "something" could be any of $1.96\bar{X}/\sqrt{n}$, $1.96S/\sqrt{n}$, $3.182\bar{X}/\sqrt{n}$, or $3.182S/\sqrt{n}$.

To be definite, let us consider the fourth proposal for "something". Supposing that $\theta = 1$, we may carry out a simulation study to see how often $\bar{X} \pm 3.182S/\sqrt{n}$ contains 1. When I ran the R code shown below, I obtained 0.88115, which is quite far from the nominal confidence level of 95%.

```
Data <- matrix(rexp(100000*4,rate=1),nrow=100000,ncol=4)
Xbar <- apply(Data,1,mean)
S <- apply(Data,1,sd)
n <- 4
CILower <- Xbar - qt(.975,3)*S/sqrt(n)
CIUpper <- Xbar + qt(.975,3)*S/sqrt(n)
mean( (CILower < 1)&(CIUpper > 1) )
```

We will return to this case study, to identify some better methods for obtaining a 95% confidence interval from a non-normal point estimator, at the end of Unit VI.

c. Inverting tests

In introductory methods courses, we tell students that θ_0 should be included in a $100(1-\alpha)\%$ confidence interval for θ if and only if $H_0 : \theta = \theta_0$ is accepted over $H_1 : \theta \neq \theta_0$ at significance level α.

A particularly simple example arises when θ represents the unknown mean of a normal distribution with variance known to equal 1. We accept $H_0 : \theta = \theta_0$ over $H_1 : \theta \neq \theta_0$ at significance level 0.05 if and only if

$$|\bar{X} - \theta_0| \leq 1.96/\sqrt{n}.$$

Yet, algebraic manipulation shows that this is the same as

$$\theta_0 \in [\bar{X} - 1.96/\sqrt{n}, \bar{X} + 1.96/\sqrt{n}],$$

which is to say that θ_0 falls in the 95% confidence interval for θ.

One practical implication of this duality between hypothesis tests and confidence intervals arises when we have a confidence interval in hand but have not performed the companion hypothesis test. In this case, we can judge statistical significance using the confidence interval without ever calculating a test statistic or a p-value for the hypothesis test. For instance, if I tell you that an odds ratio has been estimated as 1.08 with accompanying 95% confidence interval 1.03 to 1.13, then you need no further information to conclude that the null hypothesis of a unit odds ratio should be rejected and that the 1.08 should be declared statistically significant.

A second practical implication arises when we have a hypothesis testing procedure available but are not sure how to construct a confidence interval. We can then "invert" the testing procedure to obtain a confidence interval. To illustrate, suppose that $X_1, \ldots, X_n$ are independent Bernoulli random variables with success probability $p \in (0, 1)$.

If we are comfortable with Wald testing, we may note that $H_0 : p = p_0$ is accepted over $H_1 : p \neq p_0$ at approximate significance level α if and only if

$$|\hat{p} - p_0| \leq z_{1-\alpha/2}\sqrt{\hat{p}(1-\hat{p})/n},$$

which is the same as

$$p_0 \in [\hat{p} - z_{1-\alpha/2}\sqrt{\hat{p}(1-\hat{p})/n}\ ,\ \hat{p} + z_{1-\alpha/2}\sqrt{\hat{p}(1-\hat{p})/n}].$$

Hence,

$$[\hat{p} - z_{1-\alpha/2}\sqrt{\hat{p}(1-\hat{p})/n}\ ,\ \hat{p} + z_{1-\alpha/2}\sqrt{\hat{p}(1-\hat{p})/n}]$$

defines an approximate $100(1-\alpha)\%$ confidence interval. This "Wald interval" is, of course, the familiar formula that we teach in introductory methods courses. However, as we have seen in the first motivating case study, the Wald interval is not a good choice when $n\hat{p}(1-\hat{p}) < \chi^2_{1,1-\alpha}$.

If we are comfortable with score testing, we may note that $H_0 : p = p_0$ is accepted over $H_1 : p \neq p_0$ at approximate significance level α if and only if

$$|\hat{p} - p_0| \leq z_{1-\alpha/2}\sqrt{p_0(1-p_0)/n},$$

which is equivalent to

$$(\hat{p} - p_0)^2 \leq \chi^2_{1,1-\alpha} p_0(1-p_0)/n. \tag{2}$$

We rearrange (2) as

$$p_0^2(1 + \chi^2_{1,1-\alpha}/n) + p_0(-2\hat{p} - \chi^2_{1,1-\alpha}/n) + \hat{p}^2 \leq 0. \tag{3}$$

The "$\leq$" in (3) becomes "$=$" when

$$p_0 = \frac{(2\hat{p} + \chi^2_{1,1-\alpha}/n) \pm \sqrt{(2\hat{p} + \chi^2_{1,1-\alpha}/n)^2 - 4\hat{p}^2(1 + \chi^2_{1,1-\alpha}/n)}}{2(1 + \chi^2_{1,1-\alpha}/n)}, \tag{4}$$

and so the right member of (4) defines the endpoints of an approximate $100(1-\alpha)\%$ confidence interval. This "score interval" never contains a negative number and is not centered exactly at $\hat{p}$ but rather shifted slightly rightward if $\hat{p} < 1/2$ and slightly leftward if $\hat{p} > 1/2$.

If neither Wald testing nor score testing is palatable, another option is to invert an exact test. Actually, we have already seen an example of how to do this, in the first motivating case study of Unit IV. The general operating procedure is concisely summarized as follows. Let

$$f(y;p) := \binom{n}{y} p^y (1-p)^{n-y}$$

for $y \in \{0, 1, \ldots, n\}$ and $p \in (0, 1)$. Suppose that the actual number of successes observed is k. We include $p_0 \in (0, 1)$ in the $100(1-\alpha)\%$ confidence interval if and only if

$$\sum_{y \in \{0,1,\ldots,n\} : f(y;p_0) \leq f(k;p_0)} f(y;p_0) \geq \alpha. \tag{5}$$

We may also be curious to see what happens when we invert a one-sided test, so let us consider testing $H_0 : \mu \leq \mu_0$ against $H_1 : \mu > \mu_0$ for $X_1, \ldots, X_n \overset{iid}{\sim} N(\mu, 1)$ and $\mu \in \mathbb{R}$, where μ_0 is a fixed real number.

A reasonable testing procedure (in fact, the uniformly most powerful testing procedure at significance level 0.05) is to accept H_0 if and only if

$$\bar{X} \leq \mu_0 + 1.645/\sqrt{n},$$

which is algebraically equivalent to $\mu_0 \geq \bar{X} - 1.645/\sqrt{n}$. Thus, we actually obtain a 95% confidence interval of infinite length, namely

$$\bar{X} - 1.645/\sqrt{n} \;\text{ to }\; +\infty. \tag{6}$$

We refer to (6) as a "one-sided" confidence interval since only one of the bounds is finite. Interestingly, although one-sided tests are routinely taught in introductory methods courses, very little is said about one-sided confidence intervals. In Unit VII we will discuss the practical merits of one-sided confidence intervals, but for now the key point is that inverting a one-sided test usually yields a one-sided confidence interval.

d. Pivoting

Another approach to constructing confidence intervals, and the approach that we usually teach in introductory methods courses before noting the duality between hypothesis testing and confidence intervals, is called pivoting.

Let $X_1, \ldots, X_n \overset{iid}{\sim} f(x;\theta)$ for some probability mass or density function $f(x;\theta)$ and some unknown parameter $\theta \in \Theta$. A pivot is a quantity whose probabilistic behavior is known and does not depend on θ. For instance, if $f(x;\theta)$ represents the normal probability density function with mean θ and unit variance, then $\sqrt{n}(\bar{X} - \theta)$ is a pivot. This is because, regardless of θ, the distribution of $\sqrt{n}(\bar{X} - \theta)$ is standard normal.

To see how this pivot can be used to construct a confidence interval, note that for any numbers $c_1 \in [0, 1]$ and $c_2 \in [0, 1]$ with $c_2 - c_1 = 1 - \alpha$ we have

$$P(z_{c_1} < \sqrt{n}(\bar{X} - \theta) < z_{c_2}) = 1 - \alpha.$$

Above, we interpret z_0 as $-\infty$ and z_1 as $+\infty$. Also, since the distribution of the pivot is continuous rather than discrete, the first "$<$" can be made "$\leq$" if $c_1 > 0$ and the second "$<$" can be made "$\leq$" if $c_2 < 1$.

Algebraically rearranging

$$z_{c_1} < \sqrt{n}(\bar{X} - \theta) < z_{c_2}$$

yields

$$\bar{X} - z_{c_2}/\sqrt{n} < \theta < \bar{X} - z_{c_1}/\sqrt{n}, \tag{7}$$

which defines a $100(1-\alpha)\%$ confidence interval. Note that the minus sign in the right member of (7) is correct since z_{c_1} will in general be negative. Also, either "$<$" in (7) can be made "$\leq$" as long as no infinity is in play.

Note that pivoting can yield a one-sided interval (if $c_1 = 0$ and $c_2 = 1 - \alpha$, or if $c_1 = \alpha$ and $c_2 = 1$), an "equal-tail two-sided interval" (if $c_1 = \alpha/2$ and $c_2 = 1 - \alpha/2$), or an "unequal-tail two-sided interval" (if $c_1 \in (0, \alpha/2)$ or $c_2 \in (1 - \alpha/2, 1)$).

An unequal-tail two-sided interval for the mean parameter of a symmetric distribution does not make much sense, but an unequal-tail two-sided interval may make sense for a parameter describing a non-symmetric distribution since in that case an unequal-tail two-sided interval may be of shorter length than an equal-tail two-sided interval.

For more examples of pivoting, suppose that $f(x;\theta)$ represents the normal probability density function with $\theta = (\theta_1, \theta_2)^T$ a vector whose components represent the mean and variance respectively.

Regardless of θ_1 and θ_2, the quantity $\sqrt{n/\theta_2}(\bar{X} - \theta_1)$ has the standard normal distribution. Unfortunately, this quantity is not of practical use as a pivot since

$$z_{c_1} < \sqrt{n/\theta_2}(\bar{X} - \theta_1) < z_{c_2}$$

results in

$$\bar{X} - z_{c_2}\sqrt{\theta_2/n} < \theta_1 < \bar{X} - z_{c_1}\sqrt{\theta_2/n},$$

an interval that cannot be calculated since θ_2 is unknown.

So, instead we will work with $\sqrt{n}(\bar{X} - \theta_1)/S$, which has the T distribution on $(n-1)$ degrees of freedom. This quantity is of practical use as a pivot since

$$t_{n-1,c_1} < \sqrt{n}(\bar{X} - \theta_1)/S < t_{n-1,c_2}$$

results in

$$\bar{X} - t_{n-1,c_2} S/\sqrt{n} < \theta_1 < \bar{X} - t_{n-1,c_1} S/\sqrt{n},$$

an interval that can be calculated.

Approximate pivots are frequently used to generate approximate confidence intervals. Continuing from the above example, one can show (using methods beyond the scope of BST 676) that

$$\sqrt{n}(S^2 - \theta_2) \xrightarrow{L} N(0, 2\theta_2^2).$$

Noting the consistency of S^2 for θ_2, we apply Slutsky's Theorem #3 to obtain

$$\sqrt{n/2}(S^2 - \theta_2)/S^2 \xrightarrow{L} N(0, 1).$$

Thus, for large n, $\sqrt{n/2}(S^2 - \theta_2)/S^2$ is an approximate pivot, and we obtain an approximate confidence interval

$$S^2(1 - z_{c_2}\sqrt{2/n}) < \theta_2 < S^2(1 - z_{c_1}\sqrt{2/n}). \tag{8}$$

Actually, we can obtain an exact confidence interval using a different pivot, namely $(n-1)S^2/\theta_2$, which has the chi-square distribution on $(n-1)$ degrees of freedom. Rearranging

$$\chi^2_{n-1,c_1} < (n-1)S^2/\theta_2 < \chi^2_{n-1,c_2}$$

yields

$$(n-1)S^2/\chi^2_{n-1,c_2} < \theta_2 < (n-1)S^2/\chi^2_{n-1,c_1}. \tag{9}$$

Sometimes (9) is presented in introductory methods courses.

Incidentally, comparing (8) to (9) suggests that, for large n,

$$\chi^2_{n-1,c_1} \approx (n-1)/(1 - z_{c_1}\sqrt{2/n}) \quad \text{and} \quad \chi^2_{n-1,c_2} \approx (n-1)/(1 - z_{c_2}\sqrt{2/n}).$$

You may wonder why (8) would be considered at all when (9) is available. The reason is that (8) can be modified to accommodate non-normal data with finite fourth moment. More specifically, the "2" is replaced by

$$\frac{E[(X - E[X])^4]}{(Var[X])^2} - 1. \tag{10}$$

If quantity (10) depends on an unknown parameter (not the case for normal data), we can replace it by a consistent estimator and then invoke Slutsky's Theorem #3.

e. Bayesian credible intervals

We have already spoken of Bayesian inference in the context of point estimation. Recall that, given a prior distribution $p(\theta)$ reflecting our beliefs about θ before observing the data and a likelihood function $L(\theta; \mathbf{x})$ describing the probabilistic structure of the data, we can calculate a posterior distribution $p(\theta; \mathbf{x})$ reflecting our beliefs about θ after observing the data,

$$p(\theta; \mathbf{x}) = \frac{p(\theta)L(\theta; \mathbf{x})}{\int_{\eta \in \Theta} p(\eta)L(\eta; \mathbf{x})\ d\eta} \propto p(\theta)L(\theta; \mathbf{x}).$$

On the other hand, we were curiously silent about Bayesian inference in the context of hypothesis testing. To understand this, consider testing $H_0 : \theta = \theta_0$ against $H_1 : \theta \neq \theta_0$. An ostensibly reasonable hypothesis testing procedure would be to reject H_0 if its posterior probability were smaller than some threshold $c \in (0, 1)$. Yet, regardless of the data, the posterior probability of H_0 is 0 because

$$\int_{\theta_0}^{\theta_0} p(\theta; \mathbf{x})\ d\theta = 0.$$

Indeed, the prior probability of H_0 is also 0 since

$$\int_{\theta_0}^{\theta_0} p(\theta)\ d\theta = 0. \tag{11}$$

Thus, an attempt at two-sided testing from a Bayesian perspective basically dooms H_0 from the start, unless one is willing to adopt a prior distribution that has both discrete and continuous parts; such a prior distribution is decidedly contrived.

Incidentally, you may have heard someone make a statement like, "In practice, the null hypothesis is always false; we just need to collect enough data to prove it with a hypothesis test." Whether or not the speaker explicitly recognizes his/her own Bayesian thinking, such a statement finds a clear expression in result (11).

However, Bayesian inference does allow for meaningful interval estimation. To distinguish Bayesian interval estimates from frequentist confidence intervals, we refer to the Bayesian interval estimates as credible intervals.

As an illustration, let $X_1, \ldots, X_n \stackrel{iid}{\sim} 1_{\{x>0\}}\theta \exp[-x\theta]$ for $\theta \in \Theta := (0, \infty)$. Let a and b be positive constants, and impose prior distribution

$$p(\theta) = \frac{b^a}{\Gamma[a]} \theta^{a-1} \exp[-b\theta]$$

for $\theta \in \Theta$. As shown in Unit II, this yields posterior distribution

$$p(\theta; \mathbf{x}) = \frac{(b + n\bar{x})^{a+n}}{\Gamma[a+n]} \theta^{a+n-1} \exp[-(b + n\bar{x})\theta]. \tag{12}$$

We define a $100(1-\alpha)\%$ credible interval as $[d_1, d_2]$ for any numbers $d_1 < d_2$ with

$$\int_{d_1}^{d_2} p(\theta; \mathbf{x}) \, d\theta = 1 - \alpha. \tag{13}$$

Note that requirement (13) does not uniquely determine a $100(1-\alpha)\%$ credible interval. One option, called the "equal-tail" credible interval, is to let d_1 be the $\alpha/2$ quantile of the posterior distribution and d_2 the $1-\alpha/2$ quantile of the posterior distribution. A second option entails minimizing the length of the credible interval, which is accomplished by requiring that

$$p(d_1; \mathbf{x}) = p(d_2; \mathbf{x}) \tag{14}$$

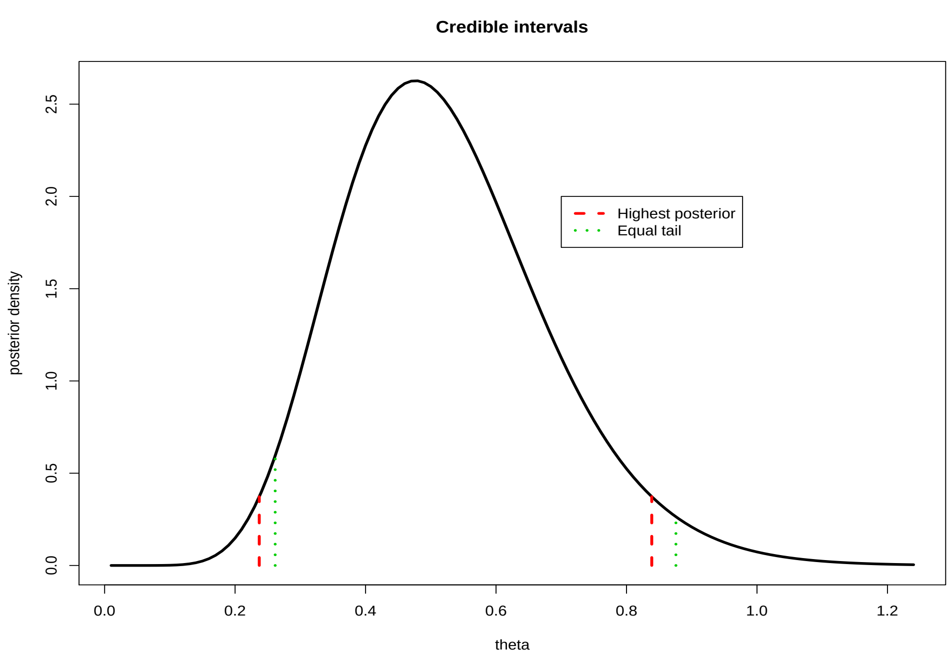
in addition to (13).

The visual interpretation of (14) is that we are including in the credible interval all elements of the parameter space for which the posterior density is sufficiently large. Thus, the credible interval determined from (14) is sometimes called a highest posterior density region. Figure 1 on the next page, obtained with the R code below, illustrates both the equal-tail credible interval $[0.261, 0.876]$ and the highest posterior density region $[0.237, 0.839]$ based on (12) supposing that $a = b = 1$, $n = 10$, $\bar{x} = 2$, and $\alpha = 0.05$.

```
theta<-(1:124)/100
postdensity<-dgamma(theta,shape=11,rate=21)
plot(theta,postdensity,type="l", lwd=3, ylab="posterior density")
c1<- (1:499)/10000
c2<-  0.95 + c1
hpdcoord <- which.min( abs( dgamma( qgamma(c1,shape=11,rate=21),
```

```
  shape=11,rate=21) - dgamma( qgamma(c2,shape=11,rate=21),
  shape=11,rate=21)) )
hpdleft <- qgamma(c1[hpdcoord],shape=11,rate=21)
hpdright <- qgamma(c2[hpdcoord],shape=11,rate=21)
eqtleft <- qgamma(c1[250],shape=11,rate=21)
eqtright <- qgamma(c2[250],shape=11,rate=21)
lines( c(hpdleft,hpdleft), c(0,dgamma(hpdleft,shape=11,rate=21)),
  col=2, lty=2, lwd=3)
lines( c(hpdright,hpdright), c(0,dgamma(hpdright,shape=11,rate=21)),
  col=2, lty=2, lwd=3)
lines( c(eqtleft,eqtleft), c(0,dgamma(eqtleft,shape=11,rate=21)),
  col=3, lty=3, lwd=3)
lines( c(eqtright,eqtright), c(0,dgamma(eqtright,shape=11,rate=21)),
  col=3, lty=3, lwd=3)
title("Credible intervals")
legend(0.7,2, c("Highest posterior", "Equal tail"), col=c(2,3),
  lty=c(2,3), lwd=c(3,3))
```

Figure 1:



f. Resolution of motivating case studies

Our first motivating case study pointed out the potential inadequacy of the Wald interval for a population proportion, despite having $n\hat{p}(1-\hat{p}) > 5$, if the desired confidence level were 99%.

One may consider using the score interval (4) instead, which yields 0.0184 ± 0.0140, or $[0.0044, 0.0324]$. Besides not containing any negative numbers, the score interval is also centered to the right of $\hat{p}$. In fact, the center of the score interval

$$\frac{2\hat{p} + \chi^2_{1,1-\alpha}/n}{2(1 + \chi^2_{1,1-\alpha}/n)} = \frac{\sum_{i=1}^n x_i + \chi^2_{1,1-\alpha}/2}{n + \chi^2_{1,1-\alpha}},$$

basically reflects an artificial addition of three successes and three failures to the Bernoull trials if $\alpha = 0.01$ or an artificial addition of two successes and two failures if $\alpha = 0.05$, since $\chi^2_{1,0.99} = 6.635 \approx 6$ and $\chi^2_{1,0.95} = 3.841 \approx 4$.

One may also consider using the interval (5) obtained by inverting the exact test. The following R code verifies that 0.0308 is within the 99% confidence interval (because the p-value associated with testing $H_0 : p = 0.0308$ against $H_1 : p \neq 0.0308$ exceeds 0.01) while 0.0309 is not (because the p-value associated with testing $H_0 : p = 0.0309$ against $H_1 : p \neq 0.0309$ is less than 0.01). Similarly, 0.0036 is within the 99% confidence interval while 0.0035 is not.

```
y <- (0:500)
sum ( dbinom(y[ dbinom(y,size=500,prob=.0308) <=
  dbinom(6,size=500,prob=.0308) ],size=500,prob=.0308)  )
sum ( dbinom(y[ dbinom(y,size=500,prob=.0309) <=
  dbinom(6,size=500,prob=.0309) ],size=500,prob=.0309)  )
```

Our second motivating case study pointed out the inadequacy of presuming $\bar{X}$ approximately normally distributed for small n when the data arose from an exponential distribution. We now offer two solutions to this case study.

The first solution is to use $(X_1 + X_2 + X_3 + X_4)/\theta$ as a pivot since its distribution is gamma with shape parameter 4 and scale parameter 1. Letting $g_{4,1,c_1}$ and $g_{4,1,c_2}$ denote the c_1 and c_2 quantiles of this distribution, where $c_1 \in (0,1)$ and $c_2 \in (0,1)$ satisfy $c_2 - c_1 = 0.95$, we rearrange the inequality

$$g_{4,1,c_1} < (X_1 + X_2 + X_3 + X_4)/\theta < g_{4,1,c_2}$$

as

$$(X_1 + X_2 + X_3 + X_4)/g_{4,1,c_2} < \theta < (X_1 + X_2 + X_3 + X_4)/g_{4,1,c_1}.$$

The choices of c_1 and c_2 are not unique. If we take $c_1 = 0.025$ and $c_2 = 0.975$, then we obtain

$$\begin{aligned} & 0.1141(X_1 + X_2 + X_3 + X_4) = (X_1 + X_2 + X_3 + X_4)/8.7673 \\ < \ & \theta < 0.9175(X_1 + X_2 + X_3 + X_4) = (X_1 + X_2 + X_3 + X_4)/1.0899. \end{aligned}$$

Seeking a shorter interval entails minimization of

$$1/g_{4,1,c_1} - 1/g_{4,1,c_1+0.95}$$

over $c_1 \in (0, 0.05)$. The following R code shows that the minimum is attained at $c_1 = 0.0484$.

```
c1 <- (1:499)/10000
c2<-  0.95 + c1
coord <- which.min( 1/qgamma(c1,shape=4,rate=1) -
  1/qgamma(c2,shape=4,rate=1) )
c1[coord]
```

The resulting confidence interval is

$$\begin{aligned} & 0.0802(X_1 + X_2 + X_3 + X_4) = (X_1 + X_2 + X_3 + X_4)/12.4631 \\ < \ & \theta < 0.7400(X_1 + X_2 + X_3 + X_4) = (X_1 + X_2 + X_3 + X_4)/1.3514. \end{aligned}$$

A second solution to this case study is to invert a likelihood ratio test. As shown in Example #5 of Unit IV, the likelihood ratio test statistic is

$$\lambda := (\bar{X}/\theta_0)^4 \exp[-(X_1 + X_2 + X_3 + X_4)/\theta_0 + 4].$$

Let $T := (X_1 + X_2 + X_3 + X_4)/\theta_0$. If H_0 is true, then T has the gamma distribution with shape 4 and scale 1. The appropriate critical value c is that which satisfies

$$P((T/4)^4 \exp[-T + 4] < c) = 0.05$$

and can be found using R. Once c has been found, the confidence interval contains each θ_0 such that

$$(\bar{X}/\theta_0)^4 \exp[-(X_1 + X_2 + X_3 + X_4)/\theta_0 + 4] \geq c.$$

Again, the computation can be carried out in R.