

S722 HW5

John Koo

To save on typing, I will denote $\frac{\partial^k}{\partial x^k} f(x) = \partial_x^k f(x)$.

9.2

The probability is greater than 0.95, since $\text{Var}(X_{n+1}) = 1$, not $n^{-1/2}$.

9.3

a

From HW9, we have the likelihood:

$$L(\alpha, \beta) = \alpha^n \beta^{-n\alpha} I(x_{(1)} \geq 0) I(x_{(n)} \leq \beta) \prod x_i^{\alpha-1}$$

Ignoring the indicator parts and taking the derivative w.r.t. β results in a negative expression, so for the MLE, we just choose the smallest possible value, i.e., $\hat{\beta} = X_{(n)}$.

Then for the upper limit, we need to find c such that

$$\begin{aligned} .05 &= P(X_{(n)}/\beta \leq c|\beta) \\ &= P(X_{(n)} \leq c\beta|\beta) \\ &= \prod P(X_i \leq c\beta|\beta) \\ &= P(X_1 \leq c\beta|\beta)^n \\ &= (c\beta/\beta)^{\alpha_0 n} \\ \implies c &= (.05)^{\frac{1}{\alpha_0 n}} \end{aligned}$$

Plugging this back in, we get

$$\begin{aligned} .05 &= P(X_{(n)}/\beta \leq c|\beta) \\ \implies .95 &= P(\beta < \frac{X_{(n)}}{c}|\beta) \end{aligned}$$

so the upper confidence limit for $\beta = \frac{X_{(n)}}{(.05)^{\frac{1}{\alpha_0 n}}}$

b

To find the MLE for α , it's easier to work with the log-likelihood:

$$\ell(\alpha, \beta) = n \log \alpha - n\alpha \log \beta + (\alpha - 1) \sum \log x_i$$

Taking the derivative w.r.t. α and setting equal to 0 yields:

$$\begin{aligned} n/\alpha - n \log \beta + \sum \log x_i &= 0 \\ \implies \hat{\alpha} &= \frac{n}{n \log \hat{\beta} - \sum \log X_i} \end{aligned}$$

Plugging this into α_0 in part (a) yields:

```
import:from(magrittr, `%>%`, `%<>%`)

alpha <- .05
```

```

# data from 7.10c
x <- c(22, 23.9, 20.9, 23.8, 25, 24, 21.7, 23.8, 22.8, 23.1, 23.1, 23.5, 23, 23)
n <- length(x)

# MLE for beta
beta.hat <- max(x)

# MLE for alpha
alpha.hat <- n / (n * log(beta.hat) - sum(log(x)))

# upper confidence limit for beta
beta.hat / alpha ** (1 / alpha.hat / n)

[1] 25.42837

```

9.4

a

First, the MLEs:

Under the alternative hypothesis, we can just deal with each sample separately, and we should obtain $\hat{\sigma}_X^2 = \frac{1}{n} \sum X_i^2$ and $\hat{\sigma}_Y^2 = \frac{1}{m} \sum Y_i^2$.

Under the null, we have:

$$L(\sigma_X^2) = (2\pi\sigma_X^2)^{-n/2} (2\pi\lambda_0\sigma_X^2)^{-m/2} \exp\left(-\frac{\sum x_i^2/\sigma_X^2 + \sum y_i^2/(\lambda_0\sigma_X^2)}{2}\right)$$

It's easier to work with the log-likelihood:

$$\ell(\sigma_X^2) = -\frac{n}{2} \log 2\pi\sigma_X^2 - \frac{m}{2} \log 2\pi\lambda_0\sigma_X^2 - \frac{1}{2\sigma_X^2} \sum x_i^2 - \frac{1}{2\lambda_0\sigma_X^2} \sum y_i^2$$

Differentiating w.r.t. σ_X^2 and setting to 0 yields:

$$\begin{aligned}
0 &= -\frac{n}{2\sigma_X^2} - \frac{m}{2\sigma_X^2} + \frac{\sum x_i^2}{2(\sigma_X^2)^2} + \frac{\sum y_i^2}{2\lambda_0(\sigma_X^2)^2} \\
\implies (n+m)\sigma_X^2 &= \sum x_i^2 + \frac{1}{\lambda_0} \sum y_i^2 \\
\implies \hat{\sigma}_X^2 &= \frac{\sum X_i^2 + \sum Y_i^2/\lambda_0}{n+m}
\end{aligned}$$

To make it possible to distinguish the null from the alternative, I will set $\sigma_X^2 = \sigma_0^2$ for the null MLE.

Plugging these into $\lambda(X, Y)$, we can note that as usual, the terms in the exponentials cancel out, as do the 2π terms, so we are left with:

$$\lambda(X, Y) = \frac{(\hat{\sigma}_X^2)^{n/2} (\hat{\sigma}_Y^2)^{m/2}}{\lambda_0^{m/2} (\hat{\sigma}_0^2)^{\frac{n+m}{2}}}$$

And we reject when this value is small.

b

Manipulating the expression for $\lambda(X, Y)$, we get:

$$\begin{aligned}
\lambda(X, Y) &= \left(\frac{\hat{\sigma}_X^2}{\hat{\sigma}_0^2}\right)^{n/2} \left(\frac{\hat{\sigma}_Y^2}{\lambda_0 \hat{\sigma}_0^2}\right)^{m/2} \\
&= \left(\frac{\sum_i X_i^2}{\sum_i X_i^2 + \sum_i Y_i^2 / \lambda_0}\right)^{n/2} \left(\frac{\sum_i Y_i^2 / \lambda_0}{\sum_i X_i^2 + \sum_i Y_i^2 / \lambda_0}\right)^{m/2} \\
&\propto \left(\frac{\chi_n^2 / n}{\chi_{n+m}^2 / (n+m)}\right)^{n/2} \left(\frac{\chi_m^2 / m}{\chi_{n+m}^2 / (n+m)}\right)^{m/2} \\
&\sim (F_{n,n+m})^{n/2} (F_{m,n+m})^{m/2}
\end{aligned}$$

And we reject if this value is too small.

c

We have some c such that $P((F_{n,n+m})^{n/2} (F_{m,n+m})^{m/2} > c | H_0) = 1 - \alpha$, where c is chosen according to the F distribution terms. So the $1 - \alpha$ confidence set is $\{\lambda : (F_{n,n+m})^{n/2} (F_{m,n+m})^{m/2} > c\}$.

We need to put this in terms of λ_0 , so we should substitute back in the original terms to get:

$$\left(\frac{\sum_i X_i^2}{\sum_i X_i^2 + \sum_i Y_i^2 / \lambda_0}\right)^{n/2} \left(\frac{\sum_i Y_i^2 / \lambda_0}{\sum_i X_i^2 + \sum_i Y_i^2 / \lambda_0}\right)^{m/2}$$

We can pull out some constants and absorb them into c and we get

$$\{\lambda : \left(\frac{\lambda \sum_i X_i^2}{\lambda \sum_i X_i^2 + \sum_i Y_i^2}\right)^{n/2} \left(\frac{\sum_i Y_i^2}{\lambda \sum_i X_i^2 + \sum_i Y_i^2}\right)^{m/2} > c'\}$$

To show that this set is an interval, we can note that as $\lambda \rightarrow \pm\infty$, the left term goes to ± 1 while the right term goes to 0, so there must be both upper and lower limits. We can also note that the left term is monotone increasing to 1 while the right term is monotone decreasing to 0 for $\lambda > 0$, so their product must have one maximum. Therefore, we are guaranteed that this set is an interval.

9.11

We know that $F_T(T|\theta) \sim Unif(0, 1)$
 $\implies P(\alpha_1 \leq F_T(T|\theta) \leq 1 - \alpha_2 | \theta = \theta_0) = 1 - \alpha_2 - \alpha_1 = 1 - \alpha$
under H_0 .

9.12

We know $\frac{\bar{X} - \theta}{\sqrt{\theta/n}} \sim \mathcal{N}(0, 1)$, so the $1 - \alpha$ confidence interval for θ is characterized by $|\frac{\bar{X} - \theta}{\sqrt{\theta/n}}| \leq z_{\alpha/2}$

$$\begin{aligned}
&\implies (\bar{X} - \theta)^2 \leq z_{\alpha/2}^2 \theta / n \\
&\implies n\theta^2 - (2\bar{X}n - z_{\alpha/2}^2)\theta + n\bar{X}^2 \leq 0 \\
&\implies \theta \in \frac{2\bar{X}n + z_{\alpha/2}^2 \pm \sqrt{4\bar{X}^2 n^2 + z_{\alpha/2}^4 + 4\bar{X}n z_{\alpha/2}^2 - 4n^2 \bar{X}^2}}{2n} \\
&\implies \theta \in \frac{2\bar{X}n + z_{\alpha/2}^2 \pm \sqrt{z_{\alpha/2}^4 + 4\bar{X}n z_{\alpha/2}^2}}{2n}
\end{aligned}$$

9.13

a

$$X \sim \text{Beta}(\theta, 1) \implies f_X(x) = \theta x^{\theta-1}$$

$$Y = -(\log X)^{-1} \implies X = \exp(-Y^{-1}) \implies X' = Y^{-2} \exp(-Y^{-1})$$

$$\begin{aligned} \text{So } f_Y(y) &= \theta (\exp(-y^{-1}))^{\theta-1} \exp(-y^{-1}) / y^2 \\ &= \frac{\theta}{y^2} \exp(-\theta/y) \end{aligned}$$

$$\text{Then } P(Y/2 \leq \theta \leq Y) = P(Y \geq \theta) - P(Y/2 \geq \theta)$$

$$= P(Y \leq 2\theta) - P(Y \leq \theta)$$

$$= \int_{\theta}^{2\theta} \frac{\theta}{y^2} e^{-\theta/y} dy$$

$$u = -\theta/y \implies du = \frac{\theta}{y^2} dy, \text{ so we have}$$

$$= \int_{-1}^{-1/2} e^u du = e^{-1/2} - e^{-1} \approx 0.239$$

b

$$F_Y(y) = \int_0^y \frac{\theta}{t^2} \exp(-\theta/t) dt$$

$$= \int_{-\infty}^{-\theta/y} e^u du$$

$$= e^{-\theta/y}$$

$$\text{So } \exp(-\theta/Y) \sim \text{Unif}(0, 1)$$

$$\implies .239 = P(a \leq \exp(-\theta/Y) \leq b) \text{ for some } a \text{ and } b \text{ s.t. } b - a = .239$$

$$\implies \text{the .239-confidence interval is } [-Y \log b, -Y \log a].$$

c

The length of the interval in part (a) is just $Y/2$, while the length of the interval in part (b) depends on how a and b are set and is proportional to $\log b - \log a$. We can try to minimize this quantity under the constraint $b - a = 1 - \alpha$ where α was found in part (a).

We can first rewrite the constraint as $a = b - 1 + \alpha$, and substituting that into the quantity we wish to minimize, we get $\log b - \log(b - 1 + \alpha) = \log \frac{b}{b-1-\alpha}$. Since the term inside the logarithm is decreasing for positive b , the entire term must also be decreasing w.r.t. b . So we take the largest possible value of b . Since $b \in [0, 1]$, we just set $b = 1$ and so $a = \alpha$ to obtain the interval $[0, -Y \log \alpha]$.

This interval's length is just $-Y \log \alpha$, so this interval is shorter for $\alpha \geq \exp(-1/2)$, which it is in this case (and of course, α would change if we were to change the interval in part (a)).

9.25

The confidence interval from example 9.2.13 is $[Y + \frac{1}{n} \log \frac{\alpha}{2}, Y + \frac{1}{n} \log(1 - \frac{\alpha}{2})]$

LRT inversion method

The likelihood for μ is

$$L(\mu) = \prod e^{-(x_i - \mu)}$$

$$\implies \ell(\mu) = -\sum x_i + n\mu$$

which is an increasing function in μ . So we just choose the smallest possible value. $\hat{m}u = X_{(n)} = Y$.

Since Y is a sufficient statistic, the LRT ratio can be written in terms of Y . Plugging in $\hat{\mu} = Y$ into the density for Y results in the exponential term canceling out, so we are just left with n , which cancels out in the LRT ratio. So we get

$$\lambda(Y) = \exp(-n(Y - \mu_0))$$

and we reject H_0 if this is too small. Since this is decreasing w.r.t. Y , this is equivalent to rejecting H_0 when Y is large.

To find the rejection region:

$$\begin{aligned} \alpha &= P(Y > c | \mu_0) = \int_c^\infty n \exp(-n(y - \mu_0)) dy = \exp(-n(c - \mu_0)) \\ \implies c &= -\frac{\log \alpha}{n} + \mu_0 \end{aligned}$$

So the acceptance region is $\{\mu : Y \leq -\frac{\log \alpha}{n} + \mu\}$ which is equivalent to $\{\mu : \mu \geq Y + \frac{\log \alpha}{n}\}$. But we should also remember that $\mu \leq Y$, so we get a proper interval for μ : $[Y + \frac{\log \alpha}{n}, Y]$

Pivot method

We need an “equivalent” random variable that does not depend on the parameter of interest, μ . We can accomplish this by just shifting the random variable:

$$Z = Y - \mu \sim n \exp(-nz) I(z > 0)$$

$$\begin{aligned} \text{Then we set } P(a \leq Z \leq b) &= 1 - \alpha \\ &= \int_a^b n e^{-nz} dz = e^{-na} - e^{-nb} \end{aligned}$$

and our interval is $[Y - a, Y - b]$ where $e^{-na} - e^{-nb} = 1 - \alpha$.

In order to minimize this interval, we minimize the quantity $b - a$ under the constraint $e^{-na} - e^{-nb} = 1 - \alpha$
 $\implies a = -\frac{1}{n} \log(1 - \alpha + e^{-nb})$.

Plugging this constraint into the quantity we wish to minimize, we get $b + \frac{1}{n} \log(1 - \alpha + e^{-nb})$ which is increasing w.r.t. b . So we merely choose the smallest possible b , so we have to shift the interval to the left as much as possible. Since a and b are positive and $b > a$, we should set $a = 0$. Then the constraint becomes:

$$1 - e^{-nb} = 1 - \alpha \implies b = -\frac{1}{n} \log \alpha$$

So our interval for Z is $[0, -\frac{1}{n} \log \alpha]$, and shifting back to Y , we get the same interval as we did with the LRT method.

9.36

The joint density is

$$\begin{aligned} f(x|\theta) &= \prod e^{i\theta - x_i} I(x_i > i\theta) \\ &= e^{\sum i\theta - x_i} I(\min(x_i/i) > \theta) \\ &= e^{-\sum x_i} e^{n\theta} I(\min(x_i/i) > \theta) \end{aligned}$$

So $h(x) = e^{-\sum x_i}$, $T(x) = \min(x_i/i)$, and $g(t) = e^{n\theta} I(t > \theta)$. Therefore, $T(X) = \min(X_i/i)$ is a sufficient statistic.

To find the density of T :

$$\begin{aligned} P(T > t) &= \prod P(X_i > it) = \prod \int_{it}^\infty e^{i\theta - x} dx \\ &= \prod e^{i(\theta - t)} = e^{-\frac{n(n-1)}{2}(t-\theta)} \end{aligned}$$

So $F_T(t) = 1 - e^{-\frac{n(n-1)}{2}(t-\theta)}$, and differentiating once, $f_T(t) = \frac{n(n-1)}{2} e^{-\frac{n(n-1)}{2}(t-\theta)}$

$$\begin{aligned}
\text{Then let } Y = T - \theta &\sim \frac{n(n-1)}{2} e^{-\frac{n(n-1)}{2}y} \text{ and we can set} \\
1 - \alpha &= P(a \leq Y \leq b) \\
&= \int_a^b \frac{n(n-1)}{2} e^{-\frac{n(n-1)}{2}y} dy \\
&= e^{-\frac{n(n-1)}{2}a} - e^{-\frac{n(n-1)}{2}b}
\end{aligned}$$

Since this is decreasing in a , we can already conclude that the optimal $a = 0$. Then $\alpha = e^{-\frac{n(n-1)}{2}b} \implies b = -\frac{2 \log \alpha}{n(n-1)}$. So the interval is $[T, T - \frac{2 \log \alpha}{n(n-1)}]$.

9.37

This is similar to the example from class. The sufficient statistic is $T = X_{(n)}$, and we can see that $T/\theta \sim f(t) = nt^{n-1}$ for $t \in (0, 1)$.

$$\begin{aligned}
\text{Our interval now is } [a, b] \text{ where } 1 - \alpha &= P(a \leq T \leq b) \\
&= \int_a^b nt^{n-1} dt = b^n - a^n \\
\implies b &= (1 - \alpha + a^n)^{1/n}
\end{aligned}$$

Then the quantity we wish to minimize becomes $(1 - \alpha + a^n)^{1/n} - a$, which is decreasing in a , so we set it to the largest possible value. This would mean shifting the interval as far as possible to the right, so $b = 1$ and $a = \alpha^{1/n}$.

Plugging this back in, we get:

$$\begin{aligned}
1 - \alpha &= P(a \leq T \leq b) \\
&= P(a \leq Y/\theta \leq b) \\
&= P(\alpha^{1/n} \leq Y/\theta \leq 1) \\
&= P(Y \leq \theta \leq Y/\alpha^{1/n})
\end{aligned}$$

So the confidence interval is $[Y, Y/\alpha^{1/n}]$.

9.52

a

From an example in class, we saw that $\lambda(X) = \lambda(W(X))$ where $W(X)$ is the sufficient statistic $W(X) = \frac{\sum (X_i - \bar{X})^2}{\sigma_0^2}$. In particular, we saw that $\lambda(W) \propto W^{n/2} \exp(-\frac{W}{2})$ and we required that $\lambda(a) = \lambda(b)$. Then we can see that $\lambda(w)$ is proportional to the density of a χ^2 random variable, in particular, $\lambda(w) \propto w^{n/2-1+2/2} \exp(-w/2) = w^{(n+2)/2-1} \exp(-w/2)$, so if we require $\lambda(a) = \lambda(b)$, then $f_{n+2}(a) = f_{n+2}(b)$.

b

From an example in class:

We have the condition $1 - \alpha = P(a \leq \chi_{n-1}^2 \leq b)$ and we want to minimize $1/a - 1/b$.

$$\begin{aligned}
\text{The condition is equivalent to } F_{n-1}(b) - F_{n-1}(a) &= 1 - \alpha \\
\implies b &= F_{n-1}^{-1}(1 - \alpha + F_{n-1}(a)).
\end{aligned}$$

Plugging this into the expression we wish to minimize, we get $a^{-1} - (F_{n-1}^{-1}(1 - \alpha + F_{n-1}(a)))^{-1}$. Then the derivative of this w.r.t. a is $a^{-2} + \frac{f_{n-1}(1 - \alpha + F_{n-1}(a))}{(F_{n-1}^{-1}(1 - \alpha + F_{n-1}(a)))^2}$

$$= -a^{-2} + \frac{f_{n-1}(a)}{b^2 f_{n-1}(b)}$$

Setting this to 0, we get $a^2 f_{n-1}(a) = b^2 f_{n-1}(b)$
 $\implies f_{n+3}(a) = f_{n+3}(b)$

c

The condition is $1 - \alpha = P(\sigma^2 \in I(S^2) | \sigma^2) \leq P((\sigma^2)' \in I(S^2) | \sigma^2)$, i.e., the interval generated has a higher probability of containing the true value than it containing any other value.

Then the length of interval is:

$$\begin{aligned} P((\sigma^2)' \in I(S^2) | \sigma^2) &= P\left(\frac{(n-1)S^2}{b\sigma^2} \leq (\sigma^2)'/\sigma^2 \leq \frac{(n-1)S^2}{a\sigma^2} | \sigma^2\right) \\ &= P(\chi_{n-1}^2/b \leq (\sigma^2)'/\sigma^2 \leq \chi_{n-1}^2/a) \\ &= \int_{a(\sigma^2)'/\sigma^2}^{b(\sigma^2)'/\sigma^2} f_{n-1}(t) dt \end{aligned}$$

Taking the derivative w.r.t. $(\sigma^2)'/\sigma^2$ yields $b f_{n-1}(b(\sigma^2)'/\sigma^2) - a f_{n-1}(a(\sigma^2)'/\sigma^2)$. This is 0 if $(\sigma^2)'/\sigma^2 = 1 \implies \sigma^2 = (\sigma^2)'$ and $b f_{n-1}(b) = a f_{n-1}(a) \implies f_{n+1}(b) = f_{n+1}(a)$.

d

If the probability within the interval is $1 - \alpha$ and the probability in each tail is equal, then each tail has probability $\alpha/2$, which is equivalent to the quantities stated.

e

First, note that we always have the constraint $F_{n-1}(b) - F_{n-1}(a) = 1 - \alpha$, so it must be that $b = F_{n-1}^{-1}(1 - \alpha + F_{n-1}(a))$.

We will use Newton-Raphson optimization to find a and b for the above constraint as well as the constraints specified individually in parts (a) and (b).

```
# specified in the problem
alpha <- .1
n <- 3

# tolerance
eps <- 1e-6

# starting guesses
a <- 0
b <- 5

# required for newton-raphson optimization
pdf.deriv <- function(x, k) {
  (k / 2 - 1) * x ** (k / 2 - 1) * exp(-x / 2) -
  x / 2 * x ** (k / 2 - 1) * exp(-x / 2)
}

# condition specified above
cond.1 <- function(a, b) {
  pchisq(b, n - 1) - pchisq(a, n - 1) - 1 + alpha
}

# --- part a --- #
```

```

# condition specified in part a
cond.a <- function(a, b) {
  dchisq(b, n + 2) - dchisq(a, n + 2)
}

# newton-raphson
while (abs(cond.1(a, b)) > eps | abs(cond.a(a, b)) > eps) {
  # compute jacobian
  J <- rbind(c(-dchisq(a, n - 1), dchisq(b, n - 1)),
             c(-pdf.deriv(a, n + 2), pdf.deriv(b, n + 2)))
  # newton-raphson step
  x <- c(a, b) - solve(J) %*% c(cond.1(a, b), cond.a(a, b))
  # update
  a <- x[1]
  b <- x[2]
}

# save results
out.df <- dplyr::data_frame(method = 'LRT', a = a, b = b,
                             `relative length` = 1 / a - 1 / b)

# --- part b --- #

# condition specified in part b
cond.b <- function(a, b) {
  dchisq(b, n + 3) - dchisq(a, n + 3)
}

# newton-raphson
while (abs(cond.1(a, b)) > eps | abs(cond.b(a, b)) > eps) {
  # compute jacobian
  J <- rbind(c(-dchisq(a, n - 1), dchisq(b, n - 1)),
             c(-pdf.deriv(a, n + 3), pdf.deriv(b, n + 3)))
  # newton-raphson step
  x <- c(a, b) - solve(J) %*% c(cond.1(a, b), cond.b(a, b))
  # update
  a <- x[1]
  b <- x[2]
}

# save results
out.df %<>% dplyr::bind_rows(
  dplyr::data_frame(method = 'minimum length', a = a, b = b,
                     `relative length` = 1 / a - 1 / b)
)

# print results
out.df %>%
  xtable::xtable() %>%
  print(include.rownames = FALSE)

```


method	a	b	relative length
LRT	0.21	12.52	4.76
minimum length	0.21	18.01	4.70