MATH3911 Assignment 2

Victor Chan

This assignment is my own work. I have read and understood the University Rules in respect to Student Academic Misconduct.

1. a) Now $X_1, ..., X_{10}$ are i.i.d. $N(\mu, 4)$ random variables with common density

$$f(x;\mu) = \frac{1}{2\sqrt{2\pi}} e^{-\frac{1}{8}(x-\mu)^2}.$$

Then the likelihood function is

$$\begin{aligned}
L(\mathbf{x};\mu) &= \prod_{i=1}^{10} f(x_i;\mu) \qquad \text{as } X_1, ..., X_{10} \text{ are i.i.d.} \\
&= \left(\frac{1}{2\sqrt{2\pi}}\right)^{10} e^{-\frac{1}{8}\sum_{i=1}^{10}(x_i-\mu)^2} \\
&= \left(\frac{1}{2\sqrt{2\pi}}\right)^{10} e^{-\frac{1}{8}\left(\sum_{i=1}^{10} x_i^2 - 2\mu\sum_{i=1}^{10} x_i + 10\mu^2\right)} \\
&= \left(\frac{1}{2\sqrt{2\pi}}\right)^{10} \left(e^{-\frac{1}{8}\sum_{i=1}^{10} x_i^2}\right)\left(e^{\frac{1}{4}\mu\sum_{i=1}^{10} x_i}\right)\left(e^{-\frac{5}{4}\mu^2}\right).
\end{aligned}$$

Fix $\mu_1, \mu_2 \in \mathbb{R}$ with $\mu_1 < \mu_2$. Then the ratio of the likelihoods at μ_2 and at μ_1 is

$$\begin{aligned}
\frac{L(\mathbf{x};\mu_2)}{L(\mathbf{x};\mu_1)} &= \frac{\left(e^{-\frac{1}{8}\sum_{i=1}^{10} x_i^2}\right)\left(e^{\frac{1}{4}\mu_2\sum_{i=1}^{10} x_i}\right)\left(e^{-\frac{5}{4}\mu_2^2}\right)}{\left(e^{-\frac{1}{8}\sum_{i=1}^{10} x_i^2}\right)\left(e^{\frac{1}{4}\mu_1\sum_{i=1}^{10} x_i}\right)\left(e^{-\frac{5}{4}\mu_1^2}\right)} \\
&= \left(e^{\frac{5}{4}(\mu_1^2-\mu_2^2)}\right)\left(e^{\frac{1}{4}(\mu_2-\mu_1)\sum_{i=1}^{10} x_i}\right)
\end{aligned}$$

which is a non-decreasing function of $\sum_{i=1}^{10} x_i$, as $\mu_2 - \mu_1 > 0$.

Hence the joint density of $X_1, ..., X_{10}$ has monotone likelihood ratio in the statistic $T(\mathbf{X}) = \sum_{i=1}^{10} X_i$.

b) Now the likelihood function of $X_1, ..., X_{10}$ is

$$\begin{aligned}
L(\mathbf{x};\mu) &= \left(\frac{1}{2\sqrt{2\pi}}\right)^{10} \left(e^{-\frac{1}{8}\sum_{i=1}^{10} x_i^2}\right)\left(e^{\frac{1}{4}\mu\sum_{i=1}^{10} x_i}\right)\left(e^{-\frac{5}{4}\mu^2}\right) \\
&= \left(\frac{1}{2\sqrt{2\pi}} e^{-\frac{1}{8}\mu^2}\right)^{10} \left(\prod_{i=1}^{10} e^{-\frac{1}{8}x_i^2}\right) e^{\frac{1}{4}\mu\sum_{i=1}^{10} x_i} \\
&= (a(\mu))^n \left(\prod_{i=1}^{n} b(x_i)\right) e^{c(\mu)\sum_{i=1}^{n} d(x_i)}
\end{aligned}$$

where

$$\begin{aligned} n &= 10 \\ a(\mu) &= \frac{1}{2\sqrt{2\pi}} e^{-\frac{1}{8}\mu^2} \\ b(x) &= e^{-\frac{1}{8}x_i^2} \\ c(\mu) &= \frac{1}{4}\mu \\ d(x) &= x. \end{aligned}$$

We have

$$T(\mathbf{X}) = \sum_{i=1}^{10} X_i = \sum_{i=1}^{10} d(X_i),$$

so by Theorem 2 in Lecture 6, the UMP unbiased size $\alpha = 0.05$ test of $H_0 : \ \mu = 1$ versus $H_1 : \ \mu \neq 1$ is

$$\varphi^*(\mathbf{X}) = \begin{cases} 1, & T(\mathbf{X}) < c_1 \text{ or } T(\mathbf{X}) > c_2 \\ 0, & c_1 \leq T(\mathbf{X}) \leq c_2 \end{cases}$$

where $c_1, c_2 \in \mathbb{R}$ are constants satisfying

$$\begin{aligned} \mathrm{E}_{\mu=1}\left[\varphi^*(\mathbf{X})\right] &= \alpha && (1) \\ \left[\frac{\partial}{\partial\mu}\mathrm{E}_{\mu}\left[\varphi^*(\mathbf{X})\right]\right]_{\mu=1} &= 0. && (2) \end{aligned}$$

We can use a deterministic test, as $X_1, ..., X_{10}$ are continuous.

Now

$$T(\mathbf{X}) = \sum_{i=1}^{10} X_i \sim N\left(10\mu, 40\right)$$

so

$$\frac{T(\mathbf{X}) - 10\mu}{2\sqrt{10}} \sim N\left(0, 1\right).$$

Also note that if $c_1 \geq c_2$, $\mathrm{Pr}_{\mu}\left[\varphi^*(\mathbf{X}) = 1\right] = 1$, so $\mathrm{E}_{\mu=1}\left[\varphi^*(\mathbf{X})\right] = 1 \neq 0.05 = \alpha$. Hence we need $c_1 < c_2$.

Then

$$\begin{aligned} \mathrm{E}_{\mu}[\varphi^*(\mathbf{X})] &= \Pr_{\mu}\left[\varphi^*(\mathbf{X}) = 1\right] \\ &= \Pr_{\mu}\left[T(\mathbf{X}) < c_1\right] + \Pr_{\mu}\left[T(\mathbf{X}) > c_2\right] \qquad \text{as } c_1 < c_2 \\ &= \Pr_{\mu}\left[\frac{T(\mathbf{X}) - 10\mu}{2\sqrt{10}} < \frac{c_1 - 10\mu}{2\sqrt{10}}\right] + \Pr_{\mu}\left[\frac{T(\mathbf{X}) - 10\mu}{2\sqrt{10}} > \frac{c_2 - 10\mu}{2\sqrt{10}}\right] \\ &= \Phi\left(\frac{c_1 - 10\mu}{2\sqrt{10}}\right) + 1 - \Phi\left(\frac{c_2 - 10\mu}{2\sqrt{10}}\right) \end{aligned}$$

where Φ is the standard normal cumulative distribution function.

Consequently

$$\frac{\partial}{\partial \mu} \mathrm{E}_\mu \left[\varphi^*(\mathbf{X})\right] = -\frac{\sqrt{10}}{2} \phi\left(\frac{c_1 - 10\mu}{2\sqrt{10}}\right) + \frac{\sqrt{10}}{2} \phi\left(\frac{c_2 - 10\mu}{2\sqrt{10}}\right)$$

where

$$\phi(z) = \frac{d}{dz}\Phi(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2}$$

is the standard normal probability density function.

From equation (2), we have

$$\begin{aligned}
\left[\frac{\partial}{\partial \mu} \mathrm{E}_\mu \left[\varphi^*(\mathbf{X})\right]\right]_{\mu=1} &= 0 \\
-\frac{\sqrt{10}}{2} \phi\left(\frac{c_1 - 10}{2\sqrt{10}}\right) + \frac{\sqrt{10}}{2} \phi\left(\frac{c_2 - 10}{2\sqrt{10}}\right) &= 0 \\
\phi\left(\frac{c_1 - 10}{2\sqrt{10}}\right) &= \phi\left(\frac{c_2 - 10}{2\sqrt{10}}\right).
\end{aligned}$$

Now if $\phi(z_1) = \phi(z_2)$,

$$\begin{aligned}
\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z_1^2} &= \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z_2^2} \\
z_1^2 &= z_2^2 \\
z_1 &= \pm z_2
\end{aligned}$$

so

$$\begin{aligned}
\frac{c_1 - 10}{2\sqrt{10}} &= \pm\frac{c_2 - 10}{2\sqrt{10}} \\
(c_1 - 10) &= \pm(c_2 - 10)
\end{aligned}$$

$$c_1 = c_2 \text{ or } c_1 = 20 - c_2.$$

As $c_1 > c_2$, the only possibility is $c_1 = 20 - c_2$.

Then from equation (1), we have

$$\begin{aligned}
\mathrm{E}_{\mu=1}\left[\varphi^*(\mathbf{X})\right] &= \alpha = 0.05 \\
\Phi\left(\frac{c_1-10}{2\sqrt{10}}\right) + 1 - \Phi\left(\frac{c_2-10}{2\sqrt{10}}\right) &= 0.05 \\
\Phi\left(\frac{-(c_2-10)}{2\sqrt{10}}\right) + 1 - \Phi\left(\frac{c_2-10}{2\sqrt{10}}\right) &= 0.05 \qquad \text{as } (c_1-10) = -(c_2-10) \\
1-\Phi\left(\frac{c_2-10}{2\sqrt{10}}\right) + 1 - \Phi\left(\frac{c_2-10}{2\sqrt{10}}\right) &= 0.05 \qquad \text{as } \Phi(-x) = 1-\Phi(x) \\
1 - \Phi\left(\frac{c_2-10}{2\sqrt{10}}\right) &= 0.025 \\
\Phi\left(\frac{c_2-10}{2\sqrt{10}}\right) &= 0.975 \\
\frac{c_2-10}{2\sqrt{10}} &= \Phi^{-1}(0.975) \approx 1.96 \\
c_2 &= 10 + 2\sqrt{10}\Phi^{-1}(0.975) \approx 22.396 \\
c_1 &= 20 - c_2 \\
&= 10 - 2\sqrt{10}\Phi^{-1}(0.975) \approx -2.396
\end{aligned}$$

so the UMP unbiased size $\alpha = 0.05$ test is

$$\varphi^*(\mathbf{X}) = \begin{cases} 1, & T(\mathbf{X}) < -2.396 \text{ or } T(\mathbf{X}) > 22.396 \\ 0, & -2.396 \le T(\mathbf{X}) \le 22.396. \end{cases}$$

c) From part b), the power function is

$$\begin{aligned}
\mathrm{E}_{\mu}\left[\varphi^*(\mathbf{X})\right] &= \Phi\left(\frac{c_1-10\mu}{2\sqrt{10}}\right) + 1 - \Phi\left(\frac{c_2-10\mu}{2\sqrt{10}}\right) \\
&= \Phi\left(\frac{-2.396-10\mu}{2\sqrt{10}}\right) + 1 - \Phi\left(\frac{22.396-10\mu}{2\sqrt{10}}\right).
\end{aligned}$$

Hence we have

$$\begin{aligned}
\mathrm{E}_{\mu=0}\left[\varphi^*(\mathbf{X})\right] &= \Phi\left(\frac{-2.396}{2\sqrt{10}}\right) + 1 - \Phi\left(\frac{22.396}{2\sqrt{10}}\right) \approx 0.3526 \\
\mathrm{E}_{\mu=0.5}\left[\varphi^*(\mathbf{X})\right] &= \Phi\left(\frac{-7.396}{2\sqrt{10}}\right) + 1 - \Phi\left(\frac{17.396}{2\sqrt{10}}\right) \approx 0.1241 \\
\mathrm{E}_{\mu=1.5}\left[\varphi^*(\mathbf{X})\right] &= \Phi\left(\frac{-17.396}{2\sqrt{10}}\right) + 1 - \Phi\left(\frac{7.396}{2\sqrt{10}}\right) \approx 0.1241 \\
\mathrm{E}_{\mu=2}\left[\varphi^*(\mathbf{X})\right] &= \Phi\left(\frac{-22.396}{2\sqrt{10}}\right) + 1 - \Phi\left(\frac{2.396}{2\sqrt{10}}\right) \approx 0.3526 \\
\mathrm{E}_{\mu=3}\left[\varphi^*(\mathbf{X})\right] &= \Phi\left(\frac{-32.396}{2\sqrt{10}}\right) + 1 - \Phi\left(\frac{-7.604}{2\sqrt{10}}\right) \approx 0.8854.
\end{aligned}$$

d) The density of the second order statistic $X_{(2)}$ is

$$\begin{aligned} f_{X_{(2)}}(x;\mu) &= \frac{10!}{1!8!} f(x;\mu)\ [F(x;\mu)]\ [1-F(x;\mu)]^8 \\ &= 90 \left(\frac{1}{2\sqrt{2\pi}} e^{-\frac{1}{8}(x-\mu)^2}\right) [F(x;\mu)]\ [1-F(x;\mu)]^8 \end{aligned}$$

where

$$\begin{aligned} F(x;\mu) &= \Pr_{\mu}\left[X_i \leq x\right] \\ &= \Pr_{\mu}\left[\frac{X_i-\mu}{2} \leq \frac{x-\mu}{2}\right] \\ &= \Phi\left(\frac{x-\mu}{2}\right), \end{aligned}$$

as

$$\frac{X_i-\mu}{2} \sim N(0,1) \qquad \text{for all } i \in \{1,...,10\}.$$

Hence

$$f_{X_{(2)}}(x;\mu) = \left(\frac{45}{\sqrt{2\pi}} e^{-\frac{1}{8}(x-\mu)^2}\right) \left[\Phi\left(\frac{x-\mu}{2}\right)\right] \left[1-\Phi\left(\frac{x-\mu}{2}\right)\right]^8,$$

so

$$f_{X_{(2)}}(x;1) = \left(\frac{45}{\sqrt{2\pi}} e^{-\frac{1}{8}(x-1)^2}\right) \left[\Phi\left(\frac{x-1}{2}\right)\right] \left[1-\Phi\left(\frac{x-1}{2}\right)\right]^8.$$

The density is plotted for μ=1.

2. a) Now $X_1, ..., X_n$ are i.i.d. random variables with common conditional density

$$f_{X|\Theta}(x\,|\,\theta) = \theta^x(1-\theta) \qquad \text{for } x \in \mathbb{N} = \{0, 1, ...\},\ \theta \in (0,1)$$

and the prior distribution of the random parameter Θ is

$$\tau(\theta) = I_{(0,1)}(\theta).$$

Then the conditional joint density of $\mathbf{X}$ is

$$\begin{aligned} f_{\mathbf{X}|\Theta}(\mathbf{x}\,|\,\theta) &= \prod_{i=1}^{n} f_{X|\Theta}(x\,|\,\theta) \qquad \text{as } X_1, ..., X_n \text{ are i.i.d., conditional on } \Theta \\ &= \theta^{\sum_{i=1}^n x_i}\,(1-\theta)^n \end{aligned}$$

so the marginal density of $\mathbf{X}$ is

$$\begin{aligned} g(\mathbf{x}) &= \int_{-\infty}^{\infty} f_{\mathbf{X}|\Theta}(\mathbf{x}\,|\,\theta)\,\tau(\theta)\,d\theta \\ &= \int_{-\infty}^{\infty} \theta^{\sum_{i=1}^n x_i}\,(1-\theta)^n\,I_{(0,1)}(\theta)\,d\theta \\ &= \int_0^1 \theta^{\sum_{i=1}^n x_i}\,(1-\theta)^n\,d\theta \\ &= B\left(\sum_{i=1}^n x_i + 1,\ n+1\right) \end{aligned}$$

where

$$B(\alpha, \beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$$

is the beta function.

Then the posterior distribution of Θ is

$$\begin{aligned} h(\theta\,|\,\mathbf{X} = \mathbf{x}) &= \frac{f_{\mathbf{X}|\Theta}(\mathbf{x}\,|\,\theta)\,\tau(\theta)}{g(\mathbf{x})} \\ &= \frac{\theta^{\sum_{i=1}^n x_i}\,(1-\theta)^n}{B\left(\sum_{i=1}^n x_i + 1,\ n+1\right)}\,I_{(0,1)}(\theta) \end{aligned}$$

so

$$\Theta|(\mathbf{X} = \mathbf{x}) \sim \text{Beta}\left(\sum_{i=1}^n x_i + 1,\ n+1\right).$$

b) The Bayes estimator of θ with respect to quadratic loss is

$$
\begin{aligned}
\mathrm{E}[\Theta \mid \mathbf{X}] &= \int_{-\infty}^{\infty} \theta \, h(\theta \mid \mathbf{X}) \, d\theta \\
&= \left[B\left(\sum_{i=1}^{n} X_i + 1, \ n+1 \right) \right]^{-1} \int_{-\infty}^{\infty} \theta \, \left(\theta^{\sum_{i=1}^n X_i} (1-\theta)^n \right) \, I_{(0,1)}(\theta) \, d\theta \\
&= \left[\frac{\Gamma(\sum_{i=1}^n X_i + 1)\, \Gamma(n+1)}{\Gamma(\sum_{i=1}^n X_i + n + 2)} \right]^{-1} \int_0^1 \theta^{\sum_{i=1}^n X_i + 1} (1-\theta)^n \, d\theta \\
&= \frac{\Gamma(\sum_{i=1}^n X_i + n + 2)}{\Gamma(\sum_{i=1}^n X_i + 1)\, \Gamma(n+1)} \, B\left(\sum_{i=1}^{n} X_i + 2, \ n+1 \right) \\
&= \frac{\Gamma(\sum_{i=1}^n X_i + n + 2)}{\Gamma(\sum_{i=1}^n X_i + 1)\, \Gamma(n+1)} \, \frac{\Gamma(\sum_{i=1}^n X_i + 2)\, \Gamma(n+1)}{\Gamma(\sum_{i=1}^n X_i + n + 3)} \\
&= \frac{\Gamma(\sum_{i=1}^n X_i + n + 2)}{\Gamma(\sum_{i=1}^n X_i + 1)} \, \frac{(\sum_{i=1}^n X_i + 1)\, \Gamma(\sum_{i=1}^n X_i + 1)}{(\sum_{i=1}^n X_i + n + 2)\, \Gamma(\sum_{i=1}^n X_i + n + 2)} \\
&= \frac{\sum_{i=1}^n X_i + 1}{\sum_{i=1}^n X_i + n + 2} \\
&= \frac{\bar{X} + \frac{1}{n}}{\bar{X} + 1 + \frac{2}{n}}.
\end{aligned}
$$

3. S-PLUS Code

```
# Question 3
# part a)
# function to estimate skewness

myskewness <- function (x) {
n <- length (x)
resid <- x - mean (x)

numer <- sqrt (n) * sum (resid^3)
denom <- (sum (resid^2))^(3/2)
skew <- numer / denom
return (skew)
}


# function to estimate kurtosis
mykurtosis <- function (x) {
n <- length (x)
resid <- x - mean (x)

numer <- n * sum (resid^4)
denom <- (sum (resid^2))^2
kurt <- numer / denom - 3
return (kurt)
```

```
}

# part b)
# estimate skewness and kurtosis for the vector 'time' in the 'lung'
data set
# using the functions in a) and built in S-PLUS functions

attach (lung)
myskewness (time)
skewness (time, method="moment")
skewness (time)
mykurtosis (time)
kurtosis (time, method="moment")
kurtosis (time)

# part c)
# bootstrap the estimators of skewness and kurtosis from part a)

boot.objs <- bootstrap (lung, myskewness (time), B=3000)
boot.objk <- bootstrap (lung, mykurtosis (time), B=3000)
summary (boot.objs)
summary (boot.objk)

# part d)
# produce a normal quantile-quantile plot for the vector 'time'
qqnorm (time, ylab="Time", main="Quantile-Quantile Plot for the Lung
Data")
```

```
> myskewness(time)
[1] 1.086788
> skewness(time, method = "moment")
[1] 1.086788
> skewness(time)
[1] 1.093999


> mykurtosis(time)
[1] 0.8934417
> kurtosis(time, method = "moment")
[1] 0.8934417
> kurtosis(time)
[1] 0.9401333
```

Frm the S-PLUS output, when called on the vector 'time', the functions 'myskewness' and 'mykurtosis' give different values to 'skewness' and 'kurtosis', when the latter are called without specifying the method, but they give the same values when the latter are called with 'method = moment'.

This is because the functions 'myskewness' and 'mykurtosis' calculate the moment estimates of skewness and kurtosis, while the built-in S-PLUS functions 'skewness' and 'kurtosis' calculate Fisher's g1 estimate for the skewness and Fisher's g2 estimate for the kurtosis, by default. However, they can be made to calculate the moment estimates by specifying 'method = moment'.

b)
```
> summary(boot.objs)
Call:  bootstrap(data = lung, statistic = myskewness(time), B = 3000)


Number of Replications:  3000

Summary Statistics:
        Observed    Bias     Mean     SE
 Param    1.087    -0.01345  1.073  0.1441

Empirical Percentiles:
        2.5%    5%      95%    97.5%
 Param   0.8   0.8418  1.316  1.362

BCa Confidence Limits:
         2.5%     5%      95%    97.5%
 Param  0.8301  0.8753   1.343   1.397

> summary(boot.objk)
Call:  bootstrap(data = lung, statistic = mykurtosis(time), B = 3000)


Number of Replications:  3000


Summary Statistics:
        Observed    Bias      Mean      SE
 Param    0.8934   -0.01379  0.8797  0.4855

Empirical Percentiles:
          2.5%      5%      95%   97.5%
 Param  0.03808  0.1393  1.696  1.896

BCa Confidence Limits:
          2.5%      5%      95%   97.5%
 Param  0.08429  0.2001  1.793  1.983
```

From the S-PLUS output, 95% BCa confidence intervals for the skewness and kurtosis are (0.8301, 1.397) and (0.08429, 1.983), respectively.

d) As neither of the 95% BCa confidence intervals include zero, we conclude that the 'time' variable is not normal. The normal quantile-quantile plot is not close to a straight-line, which confirms that the 'time' variable is not normally distributed. The plot is concave up, which suggests that 'time' is skewed to the right. This is what we would expect, as the estimated skewness is positive.

4. (a) S-PLUS Code

```
set.seed (round (log (3287921)))

# generate observations from the 'idea' distribution
x70 <- runif (70, 0.5, 4)
e70 <- rnorm (70, mean=0, sd=0.2)
y70 <- 2 + x70 + e70

# generate observations from the 'contaminating' distribution (outliers)
x30 <- rnorm (30, mean=6, sd=0.5)
y30 <- rnorm (30, mean=3, sd=0.5)

x <- c (x70, x30)
y <- c (y70, y30)
simuldata <- data.frame (x, y)

# plot the observations, with the least squares, default M-estimate and
default LTS regression lines
plot (x, y, main="Simulated Data with Regression Lines 1")
abline (lm (y~x, simuldata))
text (5.0, 3.9, "LS")

abline (rreg (x,y))
text (6.3, 3.6, "M")
abline (ltsreg.formula (y~x, simuldata))
text (4.0, 3.6, "LTS")
# redraw the plot with the new LTS regression line
plot (x, y, main="Simulated Data with Regression Lines 2")
abline (lm (y~x, simuldata))
text (5.0, 3.9, "LS")

abline (rreg (x,y))
text (6.3, 3.6, "M")
abline (ltsreg.formula (y~x, simuldata, quan=70))
text (3.0, 4.8, "LTS")

# find the slope and intercept of the old and new LTS regression lines
summary (ltsreg.formula (y~x, simuldata))
summary (ltsreg.formula (y~x, simuldata, quan=70))
```

(Partial) S-PLUS Output

```
summary(ltsreg.formula(y ~ x, simuldata))
Method:  [1] "Least Trimmed Squares Robust Regression."


Call:  ltsreg.formula(formula = y ~ x, data = simuldata)


Coefficients:
 Intercept     x
  4.1035   -0.1019

Scale estimate of residuals:  1.176

Robust Multiple R-Squared:  0.02574

Total number of observations:  100

Number of observations that determine the LTS estimate:  90


> summary(ltsreg.formula(y ~ x, simuldata, quan = 70))
Method:  [1] "Least Trimmed Squares Robust Regression."

Call:  ltsreg.formula(formula = y ~ x, data = simuldata, quan = 70)

Coefficients:
 Intercept    x
  2.0517   0.9865

Scale estimate of residuals:  0.216

Robust Multiple R-Squared:  0.8912

Total number of observations:  100

Number of observations that determine the LTS estimate:  70
```

The intercept and slope for the second LTS regression line are 2.0517 and 0.9865, respectively. These estimates are quite close to the true values for the intercept and the slope in the 'ideal' model (2 and 1, respectively). The robust multiple R-squared is 89.12%, so the fitted regression line explains most of the variation in the response. In contrast, the first LTS regression line had an intercept of 4.1035 and a slope of -0.1019, which are quite far away from the true values. The robust multiple R-squared was only 2.574%, so very little of the variation in the response was explained.

In least squares (LS) regression, the regression line is fitted by minimising the sum of squares of all residuals:

$$\text{SSRes} = \sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n} (y_i - b_0 - b_1 x_i)^2$$

where b_0 and b_1 are the intercept and slope of the fitted regression line. It is not robust to outliers, as any large residuals will increase SSRes.

In least trimmed squares (LTS) regression, the regression line is fitted by minimising the sum of squares of the m residuals with least magnitude:

$$\text{SSRes}^* = \sum_{i=1}^{m} |e_{(i)}|^2.$$

This method is more robust to outliers, as the $n - m$ largest residuals will not affect SSRes* at all.

Initially, all three regression lines were highly influenced by the outliers in the simulated data, and none of them accurately represented the 'ideal' model. The LTS regression was not sufficiently robust because the default value of m is `max(floor((n+p+1)/2), floor(0.9*n))`, which is 90 in this case (n=100, p=2). This was too large, as 20 of the residuals corresponding to outliers would have been included in SSRes*.

Decreasing m to 70 caused the LTS regression to ignore the 30 largest residuals, which would have come from the 30 outliers. The remaining residuals corresponded to only the points that came from the ideal model, so the fitted regression line was an accurate representation of this model.

5. (a) Now $X_1, ..., X_n$ are i.i.d. uniform $[0, \theta)$ random variables, with common density function

$$f(x;\theta) = \begin{cases} \frac{1}{\theta}, & 0 \le x < \theta \\ 0, & \text{otherwise} \end{cases}$$

and common distribution function

$$F(x;\theta) = \begin{cases} 0, & x < 0 \\ \frac{x}{\theta}, & 0 \le x < \theta \\ 1, & x \ge \theta. \end{cases}$$

Consider the estimator $T_n = X_{(n)} = \max\{X_1, ..., X_n\}$ of θ. The bias-corrected jackknife estimator is

$$JK(T_n) = nT_n - \frac{n-1}{n} \sum_{i=1}^{n} T_n^{(i)},$$

where $T_n^{(i)} = \max\{X_1, ..., X_{i-1}, X_{i+1}, ..., X_n\}$ is the i^{th} jackknife replication of T_n.

Now if $X_i \neq X_{(n)}$, the i^{th} jackknife sample $\{X_1, ..., X_{i-1}, X_{i+1}, ..., X_n\}$ contains $X_{(n)}$. As $X_{(n)}$ was the largest value in the original sample, it must also be the greatest value in the jackknife sample, so $T_n^{(i)} = X_{(n)}$.

If $X_i = X_{(n)}$, the i^{th} jackknife sample does not contain $X_{(n)}$, so the largest value in the jackknife sample must be the second largest value in the original sample. Hence $T_n^{(i)} = X_{(n-1)}$.

Consequently, we have

$$
\begin{aligned}
JK(T_n) &= nX_{(n)} - \frac{n-1}{n}\left[(n-1)X_{(n)} + X_{(n-1)}\right] \\
&= nX_{(n)} - \frac{(n-1)^2}{n}X_{(n)} - \frac{n-1}{n}X_{(n-1)} \\
&= \frac{1}{n}\left[n^2 - (n^2 - 2n + 1)\right] X_{(n)} - \frac{n-1}{n}X_{(n-1)} \\
&= \frac{2n-1}{n}X_{(n)} - \frac{n-1}{n}X_{(n-1)}.
\end{aligned}
$$

(b) Now the density of $X_{(n-1)}$ is

$$
\begin{aligned}
f_{X_{(n-1)}}(u;\theta) &= \frac{n!}{(n-2)!1!}\ f(u;\theta)\ \left[F(u;\theta)\right]^{n-2}\ \left[1 - F(u;\theta)\right] \\
&= \begin{cases} n(n-1)\left(\frac{1}{\theta}\right)\left(\frac{u}{\theta}\right)^{n-2}\left(1 - \frac{u}{\theta}\right) & \\ = \frac{n(n-1)}{\theta^n}\ u^{n-2}\ (\theta - u), & 0 \le u < \theta \\ 0, & \text{otherwise,} \end{cases}
\end{aligned}
$$

so we have

$$
\begin{aligned}
\mathrm{E}_\theta[X_{(n-1)}] &= \int_0^\theta u\ f_{X_{(n-1)}}(u;\theta)\ du \\
&= \int_0^\theta u\ \frac{n(n-1)}{\theta^n}\ u^{n-2}\ (\theta - u)\ du \\
&= \frac{n(n-1)}{\theta^n}\left[\theta\int_0^\theta u^{n-1}du - \int_0^\theta u^n du\right] \\
&= \frac{n(n-1)}{\theta^n}\left[\theta\left[\frac{1}{n}u^n\right]_{u=0}^\theta - \left[\frac{1}{n+1}u^{n+1}\right]_{u=0}^\theta\right] \\
&= \frac{n(n-1)}{\theta^n}\left[\frac{\theta^{n+1}}{n} - \frac{\theta^{n+1}}{n+1}\right] \\
&= n(n-1)\theta\left[\frac{1}{n} - \frac{1}{n+1}\right] \\
&= n(n-1)\theta\left[\frac{1}{n(n+1)}\right] = \frac{n-1}{n+1}\theta.
\end{aligned}
$$

Also the density of $X_{(n)}$ is

$$
\begin{aligned}
f_{X_{(n)}}(v;\theta) &= n\ f(v)\ \left[F(v;\theta)\right]^{n-1} \\
&= \begin{cases} n\left(\frac{1}{\theta}\right)\left(\frac{v}{\theta}\right)^{n-1} = \frac{n}{\theta^n}\ v^{n-1}, & 0 \le v < \theta \\ 0, & \text{otherwise,} \end{cases}
\end{aligned}
$$

so we have

$$\begin{aligned}
\mathrm{E}_\theta[X_{(n)}] &= \int_0^\theta v\ f_{X_{(n)}}(v;\theta)\ dv \\
&= \int_0^\theta v\ \frac{n}{\theta^n}\ v^{n-1}\ dv \\
&= \frac{n}{\theta^n}\int_0^\theta v^n dv \\
&= \frac{n}{\theta^n}\left[\frac{1}{n+1}v^{n+1}\right]_{v=0}^{\theta} \\
&= \frac{n}{\theta^n}\left[\frac{\theta^{n+1}}{n+1}\right] = \frac{n}{n+1}\theta.
\end{aligned}$$

Hence the bias of the original estimator $T_n = X_{(n)}$ is

$$\begin{aligned}
\mathrm{E}_\theta[X_{(n)}] - \theta &= \frac{n}{n+1}\theta - \theta \\
&= -\frac{1}{n+1}\theta.
\end{aligned}$$

Also we have

$$\begin{aligned}
\mathrm{E}_\theta[JK(T_n)] &= \frac{2n-1}{n}\ \mathrm{E}_\theta[X_{(n)}] - \frac{n-1}{n}\ \mathrm{E}_\theta[X_{(n-1)}] \\
&= \frac{2n-1}{n}\ \frac{n}{n+1}\theta - \frac{n-1}{n}\ \frac{n-1}{n+1}\theta \\
&= \frac{\theta}{n+1}\left[2n-1-\frac{(n-1)^2}{n}\right] \\
&= \frac{\theta}{n(n+1)}\left[2n^2-n-(n^2-2n+1)\right] \\
&= \frac{\theta}{n(n+1)}\left[n^2+n-1\right]
\end{aligned}$$

so the bias of the bias-corrected jackknife estimator is

$$\begin{aligned}
\mathrm{E}_\theta[JK(T_n)] - \theta &= \frac{\theta}{n(n+1)}\left[n^2+n-1\right] - \theta \\
&= \frac{\theta}{n(n+1)}\left[n^2+n-1-(n^2+n)\right] \\
&= -\frac{\theta}{n(n+1)}.
\end{aligned}$$

Thus the order of the bias has been reduced by jackknifing.

(c) Now the joint density of $X_{(n-1)}$ and $X_{(n)}$ is

$$\begin{aligned}
f_{X_{(n-1)},X_{(n)}}(u,v;\theta) &= \begin{cases} \frac{n!}{(n-2)!0!0!}\ f(u;\theta)\ f(v;\theta)\ [F(u;\theta)]^{n-2}, & u<v \\ 0, & \text{otherwise} \end{cases} \\
&= \begin{cases} \frac{n(n-1)}{\theta^n}\ u^{n-2}, & 0 \le u < v \\ 0, & \text{otherwise.} \end{cases}
\end{aligned}$$

Then we have

$$\begin{aligned}
\mathrm{E}_\theta[X_{(n-1)}\ X_{(n)}] &= \int_0^\theta \int_0^v u\ v\ f_{X_{(n-1)},X_{(n)}}(u,v;\theta)\ du\ dv \\
&= \frac{(n-1)}{\theta^n} \int_0^\theta v\ \left[\int_0^v nu^{n-1}\ du\right]\ dv \\
&= \frac{(n-1)}{\theta^n} \int_0^\theta v\ [u^n]_{u=0}^v\ dv \\
&= \frac{(n-1)}{\theta^n} \int_0^\theta v^{n+1}\ dv \\
&= \frac{(n-1)}{\theta^n} \left[\frac{1}{n+2} v^{n+2}\right]_{v=0}^\theta \\
&= \frac{n-1}{\theta^n(n+2)}\ \theta^{n+2} = \frac{n-1}{n+2}\theta^2,
\end{aligned}$$

so

$$\begin{aligned}
\mathrm{Cov}_\theta[X_{(n-1)},\ X_{(n)}] &= \mathrm{E}_\theta[X_{(n-1)}\ X_{(n)}] - \mathrm{E}_\theta[X_{(n-1)}]\mathrm{E}_\theta[X_{(n)}] \\
&= \frac{n-1}{n+2}\theta^2 - \left(\frac{n-1}{n+1}\theta\right)\left(\frac{n}{n+1}\theta\right) \\
&= (n-1)\theta^2 \left[\frac{1}{n+2} - \frac{n}{(n+1)^2}\right] \\
&= \frac{(n-1)\theta^2}{(n+1)^2(n+2)}\ \big[(n+1)^2 - n(n+2)\big] \\
&= \frac{(n-1)\theta^2}{(n+1)^2(n+2)}\ \big[(n^2+2n+1) - (n^2+2n)\big] \\
&= \frac{(n-1)\theta^2}{(n+1)^2(n+2)}.
\end{aligned}$$

Also we have

$$
\begin{aligned}
\mathrm{E}_\theta[X_{(n-1)}^2] &= \int_0^\theta u^2\ f_{X_{(n-1)}}(u;\theta)\ du \\
&= \int_0^\theta u^2\ \frac{n(n-1)}{\theta^n}\ u^{n-2}\ (\theta-u)\ du \\
&= \frac{n(n-1)}{\theta^n}\left[\theta\int_0^\theta u^n du - \int_0^\theta u^{n+1} du\right] \\
&= \frac{n(n-1)}{\theta^n}\left[\theta\left[\frac{1}{n+1}u^{n+1}\right]_{u=0}^{\theta} - \left[\frac{1}{n+2}u^{n+2}\right]_{u=0}^{\theta}\right] \\
&= \frac{n(n-1)}{\theta^n}\left[\frac{\theta^{n+2}}{n+1} - \frac{\theta^{n+2}}{n+2}\right] \\
&= n(n-1)\theta^2\left[\frac{1}{n+1} - \frac{1}{n+2}\right] \\
&= \frac{n(n-1)\theta^2}{(n+1)(n+2)},
\end{aligned}
$$

so

$$
\begin{aligned}
\mathrm{Var}_\theta[X_{(n-1)}] &= \mathrm{E}_\theta[X_{(n-1)}^2] - \mathrm{E}_\theta[X_{(n-1)}]^2 \\
&= \frac{n(n-1)\theta^2}{(n+1)(n+2)} - \left(\frac{n-1}{n+1}\theta\right)^2 \\
&= \frac{n-1}{n+1}\theta^2\ \left[\frac{n}{n+2} - \frac{n-1}{n+1}\right] \\
&= \frac{(n-1)\theta^2}{(n+1)^2(n+2)}\ [n(n+1)-(n-1)(n+2)] \\
&= \frac{(n-1)\theta^2}{(n+1)^2(n+2)}\ [n^2+n-(n^2+n-2)] \\
&= \frac{2(n-1)\theta^2}{(n+1)^2(n+2)}.
\end{aligned}
$$

Also we have

$$
\begin{aligned}
\mathrm{E}_\theta[X_{(n)}^2] &= \int_0^\theta v^2\ f_{X_{(n)}}(v;\theta)\ dv \\
&= \int_0^\theta v^2\ \frac{n}{\theta^n}\ v^{n-1}\ dv \\
&= \frac{n}{\theta^n}\int_0^\theta v^{n+1} dv \\
&= \frac{n}{\theta^n}\left[\frac{1}{n+2}v^{n+2}\right]_{v=0}^{\theta} \\
&= \frac{n}{\theta^n}\left[\frac{\theta^{n+2}}{n+2}\right] = \frac{n}{n+2}\theta^2,
\end{aligned}
$$

so

$$
\begin{aligned}
\mathrm{Var}_\theta[X_{(n)}] &= \mathrm{E}_\theta[X_{(n)}^2] - \mathrm{E}_\theta[X_{(n)}]^2 \\
&= \frac{n}{n+2}\theta^2 - \left(\frac{n}{n+1}\theta\right)^2 \\
&= n\theta^2 \left[\frac{1}{n+2} - \frac{n}{(n+1)^2}\right] \\
&= \frac{n\theta^2}{(n+1)^2(n+2)} \left[(n+1)^2 - n(n+2)\right] \\
&= \frac{n\theta^2}{(n+1)^2(n+2)} \left[(n^2+2n+1) - (n^2+2n)\right] \\
&= \frac{n\theta^2}{(n+1)^2(n+2)}.
\end{aligned}
$$

This gives

$$
\begin{aligned}
\mathrm{SD}_\theta[X_{(n-1)}]\ \mathrm{SD}_\theta[X_{(n)}] &= \sqrt{\mathrm{Var}_\theta[X_{(n-1)}]\ \mathrm{Var}_\theta[X_{(n)}]} \\
&= \sqrt{\frac{2(n-1)\theta^2}{(n+1)^2(n+2)}\ \frac{n\theta^2}{(n+1)^2(n+2)}} \\
&= \sqrt{\frac{2n(n-1)\theta^4}{(n+1)^4(n+2)^2}} \\
&= \frac{\theta^2}{(n+1)^2(n+2)}\ \sqrt{2n(n-1)},
\end{aligned}
$$

so we have

$$
\begin{aligned}
\mathrm{Corr}[X_{(n-1)},\ X_{(n)}] &= \frac{\mathrm{Cov}_\theta[X_{(n-1)},\ X_{(n)}]}{\mathrm{SD}_\theta[X_{(n-1)}]\ \mathrm{SD}_\theta[X_{(n)}]} \\
&= \frac{\frac{(n-1)\theta^2}{(n+1)^2(n+2)}}{\frac{\theta^2}{(n+1)^2(n+2)}\ \sqrt{2n(n-1)}} \\
&= \frac{n-1}{\sqrt{2n(n-1)}} \\
&= \frac{1-\frac{1}{n}}{\sqrt{2(1-\frac{1}{n})}}.
\end{aligned}
$$

Hence

$$
\mathrm{Corr}[X_{(n-1)},\ X_{(n)}] \to \frac{1}{\sqrt{2}}
$$

as $n \to \infty$.