Chapter 11
Stuff for Students

11.1 R

R is available from the **CRAN** website (https://cran.
r-project.org/). As of January 2020, the author's personal computer has Version 3.3.1 (June 21, 2016) of *R*. *R* is similar to *Splus*, but is free. *R* is very versatile since many people have contributed useful code, often as packages.

Downloading the book's files into R

Many of the homework problems use *R* functions contained in the book's website (http://parker.ad.siu.edu/Olive/linmodbk.htm) under the file name *linmodpack.txt*. The following two *R* commands can be copied and pasted into *R* from near the top of the file (http://parker.ad.siu.edu/Olive/
linmodrhw.txt).

Downloading the book's R functions *linmodpack.txt* and data files *linmoddata.txt* into *R*: the commands

```
source("http://parker.ad.siu.edu/Olive/linmodpack.txt")
source("http://parker.ad.siu.edu/Olive/linmoddata.txt")
```

can be used to download the *R* functions and data sets into *R*. Type *ls()*. Nearly 10 *R* functions from *linmodpack.txt* should appear. In *R*, enter the command *q()*. A window asking "*Save workspace image?*" will appear. Click on *No* to remove the functions from the computer (clicking on *Yes* saves the functions in *R*, but the functions and data are easily obtained with the source commands).

Citing packages

We will use *R* packages often in this book. The following *R* command is useful for citing the Mevik et al. (2015) `pls` package.

```
citation("pls")
```

Other packages cited in this book include `MASS` and `class`: both from Venables and Ripley (2010), `glmnet`: Friedman et al. (2015), and `leaps`: Lumley (2009).

This section gives tips on using *R*, but is no replacement for books such as Becker et al. (1988), Crawley (2005, 2013), Fox and Weisberg (2010), or Venables and Ripley (2010). Also see Mathsoft (1999ab) and use the website (www.google.com) to search for useful websites. For example enter the search words *R documentation*.

The command *q()* gets you out of *R*.

Least squares regression can be done with the function *lsfit* or *lm*.

The commands *help(fn)* and *args(fn)* give information about the function fn, e.g. if fn = lsfit.

Type the following commands.

```
x <- matrix(rnorm(300),nrow=100,ncol=3)
y <- x%*%1:3 + rnorm(100)
out<- lsfit(x,y)
out$coef
ls.print(out)
```

The first line makes a 100 by 3 matrix x with N(0,1) entries. The second line makes $y[i] = 0+1*x[i,1]+2*x[i,2]+3*x[i,2]+e$ where e is N(0,1). The term 1:3 creates the vector $(1,2,3)^T$ and the matrix multiplication operator is $\%*\%$. The function `lsfit` will automatically add the constant to the model. Typing "out" will give you a lot of irrelevant information, but *out$coef* and *out$resid* give the OLS coefficients and residuals respectively.

To make a residual plot, type the following commands.

```
fit <- y - out$resid
plot(fit,out$resid)
title("residual plot")
```

The first term in the plot command is always the horizontal axis while the second is on the vertical axis.

To put a graph in *Word,* hold down the *Ctrl* and *c* buttons simultaneously. Then select "Paste" from the *Word* menu, or hit *Ctrl* and *v* at the same time.

To enter data, open a data set in *Notepad* or *Word*. You need to know the number of rows and the number of columns. Assume that each case is entered in a row. For example, assuming that the file *cyp.lsp* has been saved on your flash drive from the webpage for this book, open *cyp.lsp* in *Word.* It has 76 rows and 8 columns. In *R* , write the following command.

```
cyp <- matrix(scan(),nrow=76,ncol=8,byrow=T)
```

Then copy the data lines from *Word* and paste them in *R*. If a cursor does not appear, hit *enter.* The command *dim(cyp)* will show if you have entered the data correctly.

Enter the following commands

```
cypy <- cyp[,2]
cypx<- cyp[,-c(1,2)]
lsfit(cypx,cypy)$coef
```

to produce the output below.

```
   Intercept           X1           X2           X3
205.40825985   0.94653718   0.17514405   0.23415181
          X4           X5           X6
  0.75927197  -0.05318671  -0.30944144
```

Making functions in R is easy.

For example, type the following commands.

```
mysquare <- function(x){
# this function squares x
r <- x^2
r }
```

The second line in the function shows how to put comments into functions.

Modifying your function is easy.

Use the fix command.
fix(mysquare)
This will open an editor such as *Notepad* and allow you to make changes. (In *Splus*, the command *Edit(mysquare)* may also be used to modify the function *mysquare*.)

To save data or a function in *R*, when you exit, click on *Yes* when the "*Save worksheet image?*" window appears. When you reenter *R*, type *ls()*. This will show you what is saved. You should rarely need to save anything for this book. To remove unwanted items from the worksheet, e.g. *x*, type *rm(x)*,
pairs(x) makes a scatterplot matrix of the columns of *x*,
hist(y) makes a histogram of *y*,
boxplot(y) makes a boxplot of *y*,
stem(y) makes a stem and leaf plot of y,
scan(), source(), and *sink()* are useful on a *Unix* workstation.
To type a simple list, use *y <− c(1,2,3.5)*.
The commands *mean(y), median(y), var(y)* are self explanatory.

The following commands are useful for a scatterplot created by the command *plot(x,y)*.
lines(x,y), lines(lowess(x,y,f=.2))

identify(x,y)
abline(out$coef), abline(0,1)

The usual arithmetic operators are $2 + 4$, $3 - 7$, $8 * 4$, $8/4$, and

```
2^{10}.
```

The ith element of vector y is $y[i]$ while the ij element of matrix x is $x[i, j]$. The second row of x is $x[2,]$ while the 4th column of x is $x[, 4]$. The transpose of x is *t(x)*.

The command *apply(x,1,fn)* will compute the row means if fn = mean. The command *apply(x,2,fn)* will compute the column variances if fn = var. The commands *cbind* and *rbind* combine column vectors or row vectors with an existing matrix or vector of the appropriate dimension.

Getting information about a library in R

In *R*, a *library* is an add–on package of *R* code. The command *library()* lists all available libraries, and information about a specific library, such as `leaps` for variable selection, can be found, e.g., with the command *library(help=leaps)*.

Downloading a library into R

Many researchers have contributed a *library* or *package* of *R* code that can be downloaded for use. To see what is available, go to the website (http://cran.us.r-project.org/) and click on the Packages icon.

Following Crawley (2013, p. 8), you may need to "Run as administrator" before you can install packages (right click on the *R* icon to find this). Then use the following command to install the *glmnet* package.

```
install.packages("glmnet")
```

Open *R* and type the following command.
library(glmnet)
Next type *help(glmnet)* to make sure that the library is available for use.

Warning: *R* is free but not fool proof. If you have an old version of *R* and want to download a library, you may need to update your version of *R*. The libraries for robust statistics may be useful for outlier detection, but the methods have not been shown to be consistent or high breakdown. All software has some bugs. For example, Version 1.1.1 (August 15, 2000) of *R* had a random generator for the Poisson distribution that produced variates with too small of a mean θ for $\theta \geq 10$. Hence simulated 95% confidence intervals might contain θ 0% of the time. This bug seems to have been fixed in Versions 2.4.1 and later. Also, some functions in *lregpack* may no longer work in new versions of *R*.

11.2 Hints for Selected Problems

Chapter 1

1.1 a) Sort each column, then find the median of each column. Then $\text{MED}(\boldsymbol{W}) = (1430, 180, 120)^T$.

b) The sample mean of $(X_1, X_2, X_3)^T$ is found by finding the sample mean of each column. Hence $\overline{\boldsymbol{x}} = (1232.8571, 168.00, 112.00)^T$.

1.2 a) $7 + \beta X_i$

b) $\hat{\beta} = \sum (Y_i - 7)X_i / \sum X_i^2$

1.3 See Section 1.3.5.

1.5 a) $\hat{\beta}_3 = \sum X_{3i}(Y_i - 10 - 2X_{2i}) / \sum X_{3i}^2$. The second partial derivative $= \sum X_{3i}^2 > 0$.

1.10 a) $X_2 \sim N(100, 6)$.

b)

$$\begin{pmatrix} X_1 \\ X_3 \end{pmatrix} \sim N_2 \left(\begin{pmatrix} 49 \\ 17 \end{pmatrix}, \begin{pmatrix} 3 & -1 \\ -1 & 4 \end{pmatrix} \right).$$

c) $X_1 \perp\!\!\!\perp X_4$ and $X_3 \perp\!\!\!\perp X_4$.

d)

$$\rho(X_1, X_2) = \frac{Cov(X_1, X_3)}{\sqrt{\text{VAR}(X_1)\text{VAR}(X_3)}} = \frac{-1}{\sqrt{3}\sqrt{4}} = -0.2887.$$

1.11 a) $Y|X \sim N(49, 16)$ since $Y \perp\!\!\!\perp X$. (Or use $E(Y|X) = \mu_Y + \Sigma_{12}\Sigma_{22}^{-1}(X - \mu_x) = 49 + 0(1/25)(X - 100) = 49$ and $\text{VAR}(Y|X) = \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21} = 16 - 0(1/25)0 = 16$.)

b) $E(Y|X) = \mu_Y + \Sigma_{12}\Sigma_{22}^{-1}(X - \mu_x) = 49 + 10(1/25)(X - 100) = 9 + 0.4X$.

c) $\text{VAR}(Y|X) = \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21} = 16 - 10(1/25)10 = 16 - 4 = 12$.

1.13 The proof is identical to that given in Example 3.2. (In addition, it is fairly simple to show that $M_1 = M_2 \equiv M$. That is, M depends on $\boldsymbol{\Sigma}$ but not on c or g.)

1.19 $\boldsymbol{\Sigma B} = E[E(\boldsymbol{X}|\boldsymbol{B}^T\boldsymbol{X})\boldsymbol{X}^T\boldsymbol{B})] = E(\boldsymbol{M}_B\boldsymbol{B}^T\boldsymbol{X}\boldsymbol{X}^T\boldsymbol{B}) = \boldsymbol{M}_B\boldsymbol{B}^T\boldsymbol{\Sigma B}$. Hence $\boldsymbol{M}_B = \boldsymbol{\Sigma B}(\boldsymbol{B}^T\boldsymbol{\Sigma B})^{-1}$.

1.26 a)

$$N_2 \left(\begin{pmatrix} 3 \\ 2 \end{pmatrix}, \begin{pmatrix} 3 & 1 \\ 1 & 2 \end{pmatrix} \right).$$

b) $X_2 \perp\!\!\!\perp X_4$ and $X_3 \perp\!\!\!\perp X_4$.

c) $\dfrac{\sigma_{12}}{\sqrt{\sigma_{11}\sigma_{33}}} = \dfrac{1}{\sqrt{2}\sqrt{3}} = 1/\sqrt{6} = 0.4082$.

1.31 See Section 1.3.6.

1.32 a) Model I:

$$\hat{\beta}_1 = \frac{\sum_{i=1}^n (X_i - \overline{X})Y_i}{\sum_{j=1}^n (x_j - \overline{x})^2} = \sum_{i=1}^n k_i Y_i \quad \text{with} \quad \mathrm{k_i} = \frac{\mathrm{x_i - \overline{x}}}{\sum_{\mathrm{j=1}}^{\mathrm{n}} (\mathrm{x_j - \overline{x}})^2}.$$

Model II:

$$\hat{\beta}_1 = \frac{\sum_{i=1}^n x_i Y_i}{\sum_{j=1}^n x_j^2} = \sum_{i=1}^n k_i Y_i \quad \text{with} \quad \mathrm{k_i} = \frac{\mathrm{x_i}}{\sum_{\mathrm{j=1}}^{\mathrm{n}} \mathrm{x_j^2}}.$$

b) Model I:

$$V(\hat{\beta}_1) = \sum_{i=1}^n k_i^2 V(Y_i) = \sigma^2 \sum_{i=1}^n k_i^2 = \sigma^2 \frac{\sum_{i=1}^n (x_i - \overline{x})^2}{[\sum_{j=1}^n (x_j - \overline{x})^2]^2} = \sigma^2 / \sum_{i=1}^n (x_i - \overline{x})^2.$$

Model II:

$$V(\hat{\beta}_1) = \sum_{i=1}^n k_i^2 V(Y_i) = \sigma^2 \sum_{i=1}^n k_i^2 = \sigma^2 \frac{\sum_{i=1}^n x_i^2}{[\sum_{j=1}^n x_j^2]^2} = \sigma^2 / \sum_{i=1}^n x_i^2.$$

c) The result follows if $\sum_{i=1} x_i^2 \geq \sum_{i=1} (x_i - \overline{x})^2$, but $\sum_{i=1}^n (x_i - \mu)^2$ is the least squares criterion for the model $x_i = \mu + e_i$, and the criterion is minimized by the least squares estimator $\hat{\mu} = \overline{x}$. Hence using $\tilde{\mu} = 0$ gives a least squares criterion at least as large as that using $\hat{\mu}$, and the result holds.

1.33 a) $E(\boldsymbol{r}) = E[(\boldsymbol{I} - \boldsymbol{P})\boldsymbol{Y}] = (\boldsymbol{I} - \boldsymbol{P})\boldsymbol{X}\boldsymbol{\beta} = \boldsymbol{0}$. $Cov(\boldsymbol{r}) = Cov[(\boldsymbol{I} - \boldsymbol{P})\boldsymbol{Y}] = (\boldsymbol{I} - \boldsymbol{P})Cov(\boldsymbol{Y})(\boldsymbol{I} - \boldsymbol{P})^T = \sigma^2(\boldsymbol{I} - \boldsymbol{P})$.

b) $\text{Cov}(\boldsymbol{r}, \boldsymbol{Y}) = E([\boldsymbol{r} - E(\boldsymbol{r})][\boldsymbol{Y} - E(\boldsymbol{Y})]^T) =$

$$E([(\boldsymbol{I} - \boldsymbol{P})\boldsymbol{Y} - (\boldsymbol{I} - \boldsymbol{P})E(\boldsymbol{Y})][\boldsymbol{Y} - E(\boldsymbol{Y})]^T) =$$

$$E[(\boldsymbol{I}-\boldsymbol{P})[\boldsymbol{Y}-E(\boldsymbol{Y})][\boldsymbol{Y}-E(\boldsymbol{Y})]^T] = (\boldsymbol{I}-\boldsymbol{P})Cov(\boldsymbol{Y}) = (\boldsymbol{I}-\boldsymbol{P})\sigma^2\boldsymbol{I} = \sigma^2(\boldsymbol{I}-\boldsymbol{P}).$$

c) $\text{Cov}(\boldsymbol{r}, \hat{\boldsymbol{Y}}) = E([\boldsymbol{r} - E(\boldsymbol{r})][\hat{\boldsymbol{Y}} - E(\hat{\boldsymbol{Y}})]^T) =$

$$E([(\boldsymbol{I} - \boldsymbol{P})\boldsymbol{Y} - (\boldsymbol{I} - \boldsymbol{P})E(\boldsymbol{Y})][\boldsymbol{P}\boldsymbol{Y} - \boldsymbol{P}E(\boldsymbol{Y})]^T) =$$

$$E[(\boldsymbol{I} - \boldsymbol{P})[\boldsymbol{Y} - E(\boldsymbol{Y})][\boldsymbol{Y} - E(\boldsymbol{Y})]^T\boldsymbol{P}] = (\boldsymbol{I} - \boldsymbol{P})\sigma^2\boldsymbol{I}\boldsymbol{P} = \sigma^2(\boldsymbol{I} - \boldsymbol{P})\boldsymbol{P} = \boldsymbol{0}.$$

Chapter 2

2.1 See the proof of Theorem 2.18.

2.14 For fixed $\sigma > 0$, $L(\boldsymbol{\beta}, \sigma^2)$ is maximized by minimizing $Q(\boldsymbol{\beta}) \geq 0$. So $\hat{\boldsymbol{\beta}}_Q$ maximizes $L(\boldsymbol{\beta}, \sigma^2)$ regardless of the value of $\sigma^2 > 0$. So $\hat{\boldsymbol{\beta}}_Q$ is the MLE.

b) Let $Q = Q(\hat{\boldsymbol{\beta}}_Q)$. Then the MLE $\hat{\sigma}^2$ is found by maximizing the profile likelihood, $L_p(\sigma^2) = L(\hat{\boldsymbol{\beta}}_Q, \sigma^2) = c_n \; \frac{1}{\sigma^n} \exp\left(\frac{-1}{2\sigma^2} Q\right)$. Let $\tau = \sigma^2$. The $L_p(\tau) = c_n \; \frac{1}{\tau^{n/2}} \exp\left(\frac{-1}{2\tau} Q\right)$, and the log profile likelihood $\log L_p(\tau) = d - \frac{n}{2}\log(\tau) - \frac{Q}{2\tau}$. Thus

$$\frac{d \; \log L_p(\tau)}{d\tau} = \frac{-n}{2\tau} + \frac{Q}{2\tau^2} \stackrel{set}{=} 0$$

or $-n\tau + Q = 0$ or $\hat{\tau} = \hat{\sigma}^2 = Q/n$, unique. Then

$$\frac{d^2 \; \log L_p(\tau)}{d\tau^2} = \frac{n}{2\tau^2} - \frac{2Q}{2\tau^3}\Bigg|_{\hat{\tau}} = \frac{n}{2\hat{\tau}^2} - \frac{2n\hat{\tau}}{2\hat{\tau}^3} = \frac{-n}{2\hat{\tau}^2} < 0$$

which proves that $\hat{\sigma}^2$ is the MLE of σ^2.

2.32 a) If λ is an eigenvalue of $\boldsymbol{P}$, then for some $\boldsymbol{x} \neq \boldsymbol{0}$, $\lambda \boldsymbol{x} = \boldsymbol{P}\boldsymbol{x} = \boldsymbol{P}^2\boldsymbol{x} = \lambda^2 \boldsymbol{x}$. So $\lambda(\lambda - 1) = 0$, which only has possible solutions $\lambda = 0$ or $\lambda = 1$.

b) Thus $rank(\boldsymbol{P})$ = number of nonzero eigenvalues of $\boldsymbol{P} = tr(\boldsymbol{P})$ by a).

2.35 a) Note that $E(\boldsymbol{Y}\boldsymbol{Y}^T) = \boldsymbol{\Sigma} + \boldsymbol{\theta}\boldsymbol{\theta}^T$. Since the quadratic form is a scalar and the trace is a linear operator, $E[\boldsymbol{Y}^T\boldsymbol{A}\boldsymbol{Y}] = E[tr(\boldsymbol{Y}^T\boldsymbol{A}\boldsymbol{Y})] = E[tr(\boldsymbol{A}\boldsymbol{Y}\boldsymbol{Y}^T)] = tr(E[\boldsymbol{A}\boldsymbol{Y}\boldsymbol{Y}^T]) = tr(\boldsymbol{A}\boldsymbol{\Sigma} + \boldsymbol{A}\boldsymbol{\theta}\boldsymbol{\theta}^T) = tr(\boldsymbol{A}\boldsymbol{\Sigma}) + tr(\boldsymbol{A}\boldsymbol{\theta}\boldsymbol{\theta}^T) = tr(\boldsymbol{A}\boldsymbol{\Sigma}) + \boldsymbol{\theta}^T\boldsymbol{A}\boldsymbol{\theta}$.

b) Note that $\sum_i (Y_i - \overline{Y})^2$ is the residual sum of squares for the linear model $\boldsymbol{Y} = \boldsymbol{1} + \boldsymbol{e}$. Hence $\sum_i (Y_i - \overline{Y})^2 = \boldsymbol{Y}^T(\boldsymbol{I} - \boldsymbol{H})\boldsymbol{Y} = \boldsymbol{Y}^T(\boldsymbol{I} - \frac{1}{n}\boldsymbol{1}\boldsymbol{1}^T)\boldsymbol{Y}$ where $\boldsymbol{H} = \boldsymbol{1}(\boldsymbol{1}^T\boldsymbol{1})^{-1}\boldsymbol{1}^T$. Now $tr(\boldsymbol{A}\boldsymbol{\Sigma}) = tr(\boldsymbol{\Sigma}) - tr(\frac{1}{n}\boldsymbol{1}\boldsymbol{1}^T\boldsymbol{\Sigma})$. Now $\boldsymbol{1}^T\boldsymbol{\Sigma} = (\sigma^2[1 + (n-1)\rho], ..., \sigma^2[1 + (n-1)\rho]$, $\boldsymbol{1}\boldsymbol{1}^T\boldsymbol{\Sigma} = (\sigma^2[1 + (n-1)\rho])$, and $tr(\frac{1}{n}\boldsymbol{1}\boldsymbol{1}^T\boldsymbol{\Sigma}) = \sigma^2[1 + (n-1)\rho]$. So $tr(\boldsymbol{A}\boldsymbol{\Sigma}) = n\sigma^2 - \sigma^2[1 + (n-1)\rho] = \sigma^2[n - 1 - (n-1)\rho] = \sigma^2(n-1)(1-\rho)$. Now $\boldsymbol{\theta}^T\boldsymbol{A}\boldsymbol{\theta} = \theta\boldsymbol{1}^T(\boldsymbol{I} - \frac{1}{n}\boldsymbol{1}\boldsymbol{1}^T)\boldsymbol{1} = \theta^2(n - n^2/n) = 0$. Hence the result follows by a).

c) Assume $\boldsymbol{Y} \sim N_n(\boldsymbol{\theta}, \sigma^2\boldsymbol{I})$. Then $\overline{\boldsymbol{Y}} = \boldsymbol{B}\boldsymbol{Y}$ where $\boldsymbol{B} = \frac{1}{n}\boldsymbol{1}^T$. Now $\boldsymbol{Y}^T\boldsymbol{A}\boldsymbol{Y} = \boldsymbol{Y}^T\boldsymbol{A}^T\boldsymbol{A}\boldsymbol{Y}$. Hence the two terms are independent if $\boldsymbol{A}\boldsymbol{Y} \perp\!\!\!\perp \boldsymbol{B}\boldsymbol{Y}$ iff $\boldsymbol{A}\boldsymbol{B}^T = \boldsymbol{0}$, but $\boldsymbol{A}\boldsymbol{B}^T = \frac{1}{n}(\boldsymbol{I} - \frac{1}{n}\boldsymbol{1}\boldsymbol{1}^T)\boldsymbol{1} = \frac{1}{n}(\boldsymbol{1} - \boldsymbol{1}) = \boldsymbol{0}$.

2.37 a) Use either proof of Theorem 2.5. Normality is not necessary.

b) i)

Source	df	SS	MS	F	p-value
Regression	p-1	$SSR = \boldsymbol{Y}^T(\boldsymbol{P} - \frac{1}{n}\boldsymbol{1}\boldsymbol{1}^T)\boldsymbol{Y}$	MSR	$F_0 = \frac{MSR}{MSE}$	for H_0:
Residual	n-p	$SSE = \boldsymbol{Y}^T(\boldsymbol{I} - \boldsymbol{P})\boldsymbol{Y}$	MSE		$\beta_2 = \cdots = \beta_p = 0$

ii) $E(MSE) = \sigma^2$, so $E(SSE) = (n-p)\sigma^2$. By a)

$$E(SSR) = \boldsymbol{\beta}^T\boldsymbol{X}^T(\boldsymbol{P} - \frac{\boldsymbol{1}\boldsymbol{1}^T}{n})\boldsymbol{X}\boldsymbol{\beta} + tr[\sigma^2(\boldsymbol{P} - \frac{\boldsymbol{1}\boldsymbol{1}^T}{n})] = \boldsymbol{\beta}^T\boldsymbol{X}^T(\boldsymbol{P} - \frac{\boldsymbol{1}\boldsymbol{1}^T}{n})\boldsymbol{X}\boldsymbol{\beta} + \sigma^2(p-1).$$

When H_0 is true $\boldsymbol{X}\boldsymbol{\beta} = \boldsymbol{1}\beta_1$ and $E(SSR) = \sigma^2(p-1)$.

iii) By Theorem 2.14 g), if $\boldsymbol{Y} \sim N_n(\boldsymbol{\mu}, \sigma^2\boldsymbol{I})$ then $\dfrac{\boldsymbol{Y}^T\boldsymbol{A}\boldsymbol{Y}}{\sigma^2} \sim \chi^2\left(r, \dfrac{\boldsymbol{\mu}^T\boldsymbol{A}\boldsymbol{\mu}}{2\sigma^2}\right)$ iff $\boldsymbol{A}$ is idempotent with rank$(\boldsymbol{A}) = tr(\boldsymbol{A}) = r$.

This theorem applies to SSE/σ^2 with $\boldsymbol{A} = \boldsymbol{I} - \boldsymbol{P}$, $r = n-p$, and $\boldsymbol{\mu} = \boldsymbol{X}\boldsymbol{\beta}$. Then $\boldsymbol{\mu}^T(\boldsymbol{I}-\boldsymbol{P})\boldsymbol{\mu} = \boldsymbol{0}$ since $\boldsymbol{P}\boldsymbol{X} = \boldsymbol{X}$. Hence $SSE/\sigma^2 \sim \chi^2(n-p, 0) \sim \chi^2_{n-p}$.

2.38 a) $\boldsymbol{A}^-$ is a generalized inverse of $\boldsymbol{A}$ if $\boldsymbol{A}\boldsymbol{A}^-\boldsymbol{A} = \boldsymbol{A}$.

b) i) $\boldsymbol{P} = \boldsymbol{X}(\boldsymbol{X}^T\boldsymbol{X})^-\boldsymbol{X}^T$.

ii) $C(\boldsymbol{X}) = C(\boldsymbol{1})$. Hence $\boldsymbol{P} = \boldsymbol{1}(\boldsymbol{1}^T\boldsymbol{1})^{-1}\boldsymbol{1}^T = \frac{1}{3}\boldsymbol{1}\boldsymbol{1}^T$.

iii) $SSE = \boldsymbol{Y}^T(\boldsymbol{I}-\boldsymbol{P})\boldsymbol{Y} = \boldsymbol{Y}^T\boldsymbol{Y} - \frac{1}{3}(\sum Y_i)^2 = 1+4+9-(1+2+3)^2/3 = 14 - 36/3 = 2$.

2.39 a)

Source	df	SS	MS	E(MS)	F
Reduced	$n-p_1$	$SSE(R) = \boldsymbol{Y}^T(\boldsymbol{I} - \boldsymbol{P}_1)\boldsymbol{Y}$	MSE(R)	E(MSE(R))	$F_R = \frac{SSE(R)-SSE}{p_2 MSE} =$
Full	$n-p$	$SSE = \boldsymbol{Y}^T(\boldsymbol{I} - \boldsymbol{P})\boldsymbol{Y}$	MSE	σ^2	$\dfrac{\boldsymbol{Y}^T(\boldsymbol{P} - \boldsymbol{P}_1)\boldsymbol{Y}/p_2}{\boldsymbol{Y}^T(\boldsymbol{I} - \boldsymbol{P})\boldsymbol{Y}/(n-p)}$

where

$$E(MSE(R)) = \frac{1}{n-p_1}[\sigma^2 tr(\boldsymbol{I}-\boldsymbol{P}_1) + \boldsymbol{\beta}^T\boldsymbol{X}^T(\boldsymbol{I}-\boldsymbol{P}_1)\boldsymbol{X}\boldsymbol{\beta}] = \frac{1}{n-p_1}[\sigma^2(n-p_1) + \boldsymbol{\beta}^T\boldsymbol{X}^T(\boldsymbol{I}-\boldsymbol{P}_1)\boldsymbol{X}\boldsymbol{\beta}].$$

If H_0 is true, then $\boldsymbol{Y} \sim N_n(\boldsymbol{X}_1\boldsymbol{\beta}_1, \sigma^2\boldsymbol{I})$, and $E(MSE(R)) = \sigma^2$.

b) Need to show that $SSE(R) - SSE = \boldsymbol{Y}^T(\boldsymbol{P} - \boldsymbol{P}_1)\boldsymbol{Y}$ and $SSE = \boldsymbol{Y}^T(\boldsymbol{I} - \boldsymbol{P})\boldsymbol{Y}$ are independent. This result follows from Craig's Theorem since $(\boldsymbol{P} - \boldsymbol{P}_1)(\boldsymbol{I} - \boldsymbol{P}) = \boldsymbol{P} - \boldsymbol{P}_1 - \boldsymbol{P} + \boldsymbol{P}_1 = \boldsymbol{0}$.

c) By Theorem 2.14 g), if $\boldsymbol{Y} \sim N_n(\boldsymbol{\mu}, \sigma^2\boldsymbol{I})$ then $\dfrac{\boldsymbol{Y}^T\boldsymbol{A}\boldsymbol{Y}}{\sigma^2} \sim \chi^2\left(r, \dfrac{\boldsymbol{\mu}^T\boldsymbol{A}\boldsymbol{\mu}}{2\sigma^2}\right)$ iff $\boldsymbol{A}$ is idempotent with rank$(\boldsymbol{A}) = tr(\boldsymbol{A}) = r$.

This theorem applies to SSE/σ^2 with $\boldsymbol{A} = \boldsymbol{I} - \boldsymbol{P}$ and $r = n-p$. Then $\boldsymbol{\mu} = \boldsymbol{X}\boldsymbol{\beta}$, and $\boldsymbol{\mu}^T(\boldsymbol{I} - \boldsymbol{P})\boldsymbol{\mu} = \boldsymbol{0}$ since $\boldsymbol{P}\boldsymbol{X} = \boldsymbol{X}$. Hence $SSE/\sigma^2 \sim \chi^2(n-p, 0) \sim$

χ^2_{n-p}. Similarly, when H_0 is true, the theorem applies to $\boldsymbol{Y}^T(\boldsymbol{P}-\boldsymbol{P}_1)\boldsymbol{Y}/\sigma^2$ with $\boldsymbol{A} = \boldsymbol{P}-\boldsymbol{P}_1$ and $r = p - p_1 = p_2$. Then $\boldsymbol{\mu} = \boldsymbol{X}_1\boldsymbol{\beta}_1$, and $\boldsymbol{\mu}^T(\boldsymbol{P}-\boldsymbol{P}_1)\boldsymbol{\mu} = \boldsymbol{0}$ since $\boldsymbol{P}\boldsymbol{X}_1 = \boldsymbol{P}_1\boldsymbol{X}_1 = \boldsymbol{X}_1$. Hence $\boldsymbol{Y}^T(\boldsymbol{P}-\boldsymbol{P}_1)\boldsymbol{Y}/\sigma^2 \sim \chi^2(p_2, 0) \sim \chi^2_{p_2}$. Thus

$$F_R = \frac{\boldsymbol{Y}^T(\boldsymbol{P}-\boldsymbol{P}_1)\boldsymbol{Y}/p_2}{\boldsymbol{Y}^T(\boldsymbol{I}-\boldsymbol{P})\boldsymbol{Y}/(n-p)} \sim F_{p_2,n-p}.$$

2.40 a) $\boldsymbol{Y}^T\boldsymbol{A}\boldsymbol{Y} \sim \chi^2(rank(\boldsymbol{A}))$ iff $\boldsymbol{A}\boldsymbol{\Sigma}$ is idempotent and $\boldsymbol{\mu}^T\boldsymbol{A}\boldsymbol{\mu} = 0$ by Theorem 2.13.

b) This proof similar to the proof of Theorem 2.8. Let $\boldsymbol{u} = \boldsymbol{A}\boldsymbol{Y}$ and $\boldsymbol{w} = \boldsymbol{B}\boldsymbol{Y}$. Then $\boldsymbol{A}\boldsymbol{Y} \perp\!\!\!\perp \boldsymbol{B}\boldsymbol{Y}$ iff $\text{Cov}(\boldsymbol{w}, \boldsymbol{u}) = \boldsymbol{B}\boldsymbol{\Sigma}\boldsymbol{A} = \boldsymbol{0}$. Thus $\boldsymbol{A}\boldsymbol{Y} \perp\!\!\!\perp \boldsymbol{B}\boldsymbol{Y}$.

Let $g(\boldsymbol{A}\boldsymbol{Y}) = \boldsymbol{Y}^T\boldsymbol{A}^T\boldsymbol{A}^-\boldsymbol{A}\boldsymbol{Y} = \boldsymbol{Y}^T\boldsymbol{A}\boldsymbol{A}^-\boldsymbol{A}\boldsymbol{Y} = \boldsymbol{Y}^T\boldsymbol{A}\boldsymbol{Y}$. Then $g(\boldsymbol{A}\boldsymbol{Y}) = \boldsymbol{Y}^T\boldsymbol{A}\boldsymbol{Y} \perp\!\!\!\perp \boldsymbol{B}\boldsymbol{Y}$ since $\boldsymbol{A}\boldsymbol{Y} \perp\!\!\!\perp \boldsymbol{B}\boldsymbol{Y}$.

c) $\overline{Y} = \boldsymbol{1}^T\boldsymbol{Y}/n$ and $\sum_{i=1}^n (Y_i - \overline{Y})^2 = \boldsymbol{Y}^T(\boldsymbol{I}-\boldsymbol{P}_1)\boldsymbol{Y}$ where $\boldsymbol{P}_1 = \boldsymbol{1}\boldsymbol{1}^T/n$ is the projection matrix on $C(\boldsymbol{1})$ since $\sum_{i=1}^n (Y_i - \overline{Y})^2$ is the residual sum of squares for the model $\boldsymbol{Y} = \boldsymbol{1}\mu + \boldsymbol{e}$ with least squares estimator $\hat{\mu} = \overline{Y}$. Hence the quantities are independent if $\boldsymbol{B}\boldsymbol{Y} = \boldsymbol{1}^T\boldsymbol{Y}$ and $\boldsymbol{Y}^T\boldsymbol{A}\boldsymbol{Y} = \boldsymbol{Y}^T(\boldsymbol{I}-\boldsymbol{P}_1)\boldsymbol{Y}$ are independent, or if $\boldsymbol{1}^T\boldsymbol{I}(\boldsymbol{I}-\boldsymbol{P}_1) = 0$ by b). This result holds since $\boldsymbol{1}^T\boldsymbol{P}_1 = \boldsymbol{1}^T$ since $\boldsymbol{P}_1$ is the projection matrix on $C(\boldsymbol{1})$ means $\boldsymbol{P}_1\boldsymbol{1} = \boldsymbol{1}$.

2.41 a) $\hat{\boldsymbol{\beta}} = (\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T\boldsymbol{Y}$ and $\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^n r_i^2 = \frac{1}{n}SSE = \frac{1}{n}\boldsymbol{Y}^T(\boldsymbol{I} - \boldsymbol{P})\boldsymbol{Y}$.

b) By Theorem 2.14 g), if $\boldsymbol{Y} \sim N_n(\boldsymbol{\mu}, \sigma^2\boldsymbol{I})$ then $\dfrac{\boldsymbol{Y}^T\boldsymbol{A}\boldsymbol{Y}}{\sigma^2} \sim \chi^2\left(r, \dfrac{\boldsymbol{\mu}^T\boldsymbol{A}\boldsymbol{\mu}}{2\sigma^2}\right)$ iff $\boldsymbol{A}$ is idempotent with $\text{rank}(\boldsymbol{A}) = tr(\boldsymbol{A}) = r$.

This theorem applies to SSE/σ^2 with $\boldsymbol{A} = \boldsymbol{I}-\boldsymbol{P}$, $r = n-p$, and $\boldsymbol{\mu} = \boldsymbol{X}\boldsymbol{\beta}$. Then $\boldsymbol{\mu}^T(\boldsymbol{I}-\boldsymbol{P})\boldsymbol{\mu} = \boldsymbol{0}$ since $\boldsymbol{P}\boldsymbol{X} = \boldsymbol{X}$. Hence $SSE/\sigma^2 \sim \chi^2(n-p, 0) \sim \chi^2_{n-p}$. Thus

$$(n-p)\hat{\sigma}^2/\sigma^2 = \frac{n-p}{n}\,\frac{SSE}{\sigma^2} \sim \frac{n-p}{n}\chi^2_{n-p}.$$

c) $\boldsymbol{B}\boldsymbol{Y} \perp\!\!\!\perp \boldsymbol{Y}^T\boldsymbol{A}\boldsymbol{Y}$ if $\boldsymbol{B}\boldsymbol{A} = \boldsymbol{0}$ by Theorem 2.8 b). Here $\boldsymbol{B}\boldsymbol{A} = (\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T(\boldsymbol{I} - \boldsymbol{P}) = \boldsymbol{0}$ since $\boldsymbol{X}^T\boldsymbol{P} = \boldsymbol{X}^T$. Thus the MLEs are independent.

d) The MLE is the generalized least squares estimator $\hat{\boldsymbol{\beta}} = (\boldsymbol{X}^T\boldsymbol{V}^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}^T\boldsymbol{V}^{-1}\boldsymbol{Y}$.

2.42 Note that $\boldsymbol{H} = \boldsymbol{P}$ and that $\boldsymbol{Z} = \boldsymbol{Y} - \boldsymbol{\mu} \sim N_n(\boldsymbol{0}, \boldsymbol{\Sigma})$.

a) i) $E[(\boldsymbol{Y}-\boldsymbol{\mu})^T\boldsymbol{A}(\boldsymbol{Y}-\boldsymbol{\mu})] = E[\boldsymbol{Z}^T\boldsymbol{A}\boldsymbol{Z}] = tr(\boldsymbol{A}\boldsymbol{\Sigma}) + \boldsymbol{0}^T\boldsymbol{A}\boldsymbol{0} = tr(\boldsymbol{A}\boldsymbol{\Sigma})$ by Theorem 2.5 using $E(\boldsymbol{Z}) = \boldsymbol{0}$.

Alternatively, $E(\boldsymbol{Z}\boldsymbol{Z}^T) = \boldsymbol{\Sigma}$ since $E(\boldsymbol{Z}) = \boldsymbol{0}$. Since the quadratic form is a scalar and the trace is a linear operator, $E[\boldsymbol{Z}^T\boldsymbol{A}\boldsymbol{Z}] = E[tr(\boldsymbol{Z}^T\boldsymbol{A}\boldsymbol{Z})] = E[tr(\boldsymbol{A}\boldsymbol{Z}\boldsymbol{Z}^T)] = tr(E[\boldsymbol{A}\boldsymbol{Z}\boldsymbol{Z}^T]) = tr(\boldsymbol{A}\boldsymbol{\Sigma})$.

Normality is not needed for this result.

ii) $\boldsymbol{A}\boldsymbol{\Sigma}$ is idempotent by Theorem 2.13.

iii) $\boldsymbol{B}\boldsymbol{\Sigma}\boldsymbol{A} = \boldsymbol{0}$ (or $\boldsymbol{A}\boldsymbol{\Sigma}\boldsymbol{B}^T = \boldsymbol{0}$) by Theorem 2.8.

b) i) $\frac{1}{\sigma}(\boldsymbol{I}-\boldsymbol{H})\boldsymbol{Y} \sim N_n(\frac{1}{\sigma}(\boldsymbol{I}-\boldsymbol{H})\boldsymbol{X}\boldsymbol{\beta}, \frac{1}{\sigma}(\boldsymbol{I}-\boldsymbol{H})\sigma^2\boldsymbol{I}\frac{1}{\sigma}(\boldsymbol{I}-\boldsymbol{H}) \sim N_n(\boldsymbol{0}, \boldsymbol{I}-\boldsymbol{H})$ since $\boldsymbol{HX} = \boldsymbol{X}$.

ii) By Theorem 2.14 g), if $\boldsymbol{Y} \sim N_n(\boldsymbol{\mu}, \sigma^2\boldsymbol{I})$ then $\frac{\boldsymbol{Y}^T\boldsymbol{AY}}{\sigma^2} \sim \chi^2\left(r, \frac{\boldsymbol{\mu}^T\boldsymbol{A\mu}}{2\sigma^2}\right)$ iff $\boldsymbol{A}$ is idempotent with $\text{rank}(\boldsymbol{A}) = tr(\boldsymbol{A}) = r$.

This theorem applies to $u = \frac{\boldsymbol{Y}^T(\boldsymbol{I}-\boldsymbol{H})\boldsymbol{Y}}{\sigma^2} = SSE/\sigma^2$ with $\boldsymbol{A} = \boldsymbol{I} - \boldsymbol{H}$, $r = n - p$, and $\boldsymbol{\mu} = \boldsymbol{X\beta}$. Then $\boldsymbol{\mu}^T(\boldsymbol{I}-\boldsymbol{H})\boldsymbol{\mu} = \boldsymbol{0}$ since $\boldsymbol{HX} = \boldsymbol{X}$. Hence $SSE/\sigma^2 \sim \chi^2(n-p, 0) \sim \chi^2_{n-p}$.

iii) By Theorem 2.8 b), independence follows since $\boldsymbol{H}(\boldsymbol{I}-\boldsymbol{H}) = \boldsymbol{0}$.

2.43 a) $Q(\beta) = \sum_{i=1}^n (y_i - \beta x_i)^2$. By the chain rule,

$$\frac{dQ(\beta)}{d\beta} = -2\sum_{i=1}^n (y_i - \beta x_i)x_i.$$

Setting the derivative equal to 0 and calling the unique solution $\hat{\beta}$ gives $\sum_{i=1}^n x_i y_i = \hat{\beta}\sum_{i=1}^n x_i^2$ or

$$\hat{\beta} = \frac{\sum_{i=1}^n x_i y_i}{\sum_{i=1}^n x_i^2}.$$

b) $MSE = \frac{1}{n-1}\sum_{i=1}^n r_i^2$ since $p = 1$.

c) Since $y_i \sim N(x_i\beta, \sigma^2)$, the likelihood function

$$L(\beta, \sigma^2) = \prod_{i=1}^n f_{y_i}(y_i) = \prod_{i=1}^n \frac{1}{\sqrt{2\pi}}\frac{1}{\sigma}\exp[\frac{-1}{2\sigma^2}(y_i - x_i\beta)^2] = c_n\frac{1}{\sigma^n}\exp[\frac{-1}{2\sigma^2}\sum_{i=1}^n (y_i - x_i\beta)^2] =$$

$$c_n\frac{1}{\sigma^n}\exp[\frac{-1}{2\sigma^2}Q(\beta)]$$

where $Q(\beta)$ is the least squares criterion. For fixed $\sigma > 0$, maiximizing $L(\beta, \sigma)$ is equivalent to minimizing the least squares criterion $Q(\boldsymbol{\beta})$. Thus $\hat{\beta}$ from a) is the MLE of β. To find the MLE of σ^2, use the profile likelihood function

$$L_p(\sigma^2) = L_p(\tau) = c_n\frac{1}{\sigma^n}\exp[\frac{-1}{2\sigma^2}Q] = c_n\frac{1}{\tau^{n/2}}\exp[\frac{-1}{2\tau}Q]$$

where $Q = Q(\hat{\beta})$. Then the log profile likelihood function

$$\log(L_p(\tau)) = d_n - \frac{n}{2}\log(\tau) - \frac{Q}{2\tau},$$

$$\text{and } \frac{\mathrm{d}}{\mathrm{d}\tau}\log(\mathrm{L_p}(\tau)) = \frac{-\mathrm{n}}{2\tau} + \frac{\mathrm{Q}}{2\tau^2} \overset{\text{set}}{=} 0.$$

Thus $n\tau = Q$ or $\hat{\tau} = \hat{\sigma}^2 = Q/n = \sum_{i=1} r_i^2/n$, which is a unique solution. Now

$$\frac{d^2}{d\tau^2}\log(L_p(\tau)) = \frac{n}{2\tau^2} - \frac{2Q}{2\tau^3}\Bigg|_{\hat{\tau}} = \frac{n}{2\hat{\tau}^2} - \frac{2n\hat{\tau}}{2\hat{\tau}^3} = \frac{-n}{2\hat{\tau}^2} < 0.$$

Thus $\hat{\sigma}^2$ is the MLE of σ^2.

2.44 Let Y_1 and Y_2 be iindependent random variables with mean θ and 2θ respectively. Find the least squares estimate of θ and the residual sum of squares.

Solution:

$$\boldsymbol{Y} = \boldsymbol{X\beta} + \boldsymbol{e} = \begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}\theta + \begin{pmatrix} e_1 \\ e_2 \end{pmatrix}.$$

Then

$$\hat{\theta} = (\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T\boldsymbol{Y} = \left[(1\ 2)\begin{pmatrix} 1 \\ 2 \end{pmatrix}\right]^{-1}(1\ 2)\begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix} = \frac{Y_1 + 2Y_2}{5}.$$

$$\text{Now}\ \ \hat{\boldsymbol{Y}} = \boldsymbol{X}\hat{\theta} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}\frac{\mathrm{Y}_1 + 2\mathrm{Y}_2}{5} = \begin{pmatrix} \frac{Y_1+2Y_2}{5} \\ \\ \frac{2Y_1+4Y_2}{5} \end{pmatrix}.$$

Thus

$$RSS = \left(Y_1 - \frac{Y_1 + 2Y_2}{5}\right)^2 + \left(Y_2 - \frac{2Y_1 + 4Y_2}{5}\right)^2.$$

2.45 a) $\sqrt{n}\boldsymbol{A}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}) \xrightarrow{D} N_r(\boldsymbol{0}, \sigma^2\boldsymbol{AWA}^T)$.

b) $\boldsymbol{A}(\boldsymbol{Z}_n - \boldsymbol{\mu}) \xrightarrow{D} N_r(\boldsymbol{0}, \boldsymbol{AA}^T)$.

Chapter 3

3.7 Note that $\boldsymbol{Z}_A^T\boldsymbol{Z}_A = \boldsymbol{Z}^T\boldsymbol{Z}$,

$$\boldsymbol{G}_A\ \boldsymbol{\eta}_A = \begin{pmatrix} \boldsymbol{G\eta} \\ \sqrt{\lambda_2^*}\ \ \boldsymbol{\eta} \end{pmatrix},$$

and $\boldsymbol{Z}_A^T\boldsymbol{G}_A\boldsymbol{\eta}_A = \boldsymbol{Z}^T\boldsymbol{G\eta}$. Then

$$RSS(\boldsymbol{\eta}_A) = \|\boldsymbol{Z}_A - \boldsymbol{G}_A\boldsymbol{\eta}_A\|_2^2 = (\boldsymbol{Z}_A - \boldsymbol{G}_A\boldsymbol{\eta}_A)^T(\boldsymbol{Z}_A - \boldsymbol{G}_A\boldsymbol{\eta}_A) =$$

$$\boldsymbol{Z}_A^T\boldsymbol{Z}_A - \boldsymbol{Z}_A^T\boldsymbol{G}_A\boldsymbol{\eta}_A - \boldsymbol{\eta}_A^T\boldsymbol{G}_A^T\boldsymbol{Z}_A + \boldsymbol{\eta}_A^T\boldsymbol{G}_A^T\boldsymbol{G}_A\boldsymbol{\eta}_A =$$

$$\boldsymbol{Z}^T\boldsymbol{Z} - \boldsymbol{Z}^T\boldsymbol{G\eta} - \boldsymbol{\eta}^T\boldsymbol{G}^T\boldsymbol{Z} + \begin{pmatrix}\boldsymbol{\eta}^T\boldsymbol{G}^T & \sqrt{\lambda_2}\ \ \boldsymbol{\eta}^T\end{pmatrix}\begin{pmatrix} \boldsymbol{G\eta} \\ \sqrt{\lambda_2^*}\ \ \boldsymbol{\eta} \end{pmatrix}.$$

Thus

$$Q_N(\boldsymbol{\eta}_A) = \boldsymbol{Z}^T\boldsymbol{Z} - \boldsymbol{Z}^T\boldsymbol{G\eta} - \boldsymbol{\eta}^T\boldsymbol{G}^T\boldsymbol{Z} + \boldsymbol{\eta}^T\boldsymbol{G}^T\boldsymbol{G\eta} + \lambda_2^*\boldsymbol{\eta}^T\boldsymbol{\eta} + \gamma\|\boldsymbol{\eta}_A\|_1 =$$

$$\|\boldsymbol{Z} - \boldsymbol{G}\boldsymbol{\eta}\|_2^2 + \lambda_2^* \|\boldsymbol{\eta}\|_2^2 + \frac{\lambda_1^*}{\sqrt{1+\lambda_2^*}} \|\boldsymbol{\eta}_A\|_1 =$$

$$RSS(\boldsymbol{\eta}) + \lambda_2^* \|\boldsymbol{\eta}\|_2^2 + \lambda_1^* \|\boldsymbol{\eta}\|_1 = Q(\boldsymbol{\eta}). \quad \square$$

3.12 a) $SSE = \boldsymbol{Y}^T(\boldsymbol{I} - \boldsymbol{P})\boldsymbol{Y}$ and $SSR = \boldsymbol{Y}^T(\boldsymbol{P} - \frac{1}{n}\boldsymbol{1}\boldsymbol{1}^T)\boldsymbol{Y} = \boldsymbol{Y}^T(\boldsymbol{P} - \boldsymbol{P}_1)\boldsymbol{Y}$ where $\boldsymbol{P}_1 = \frac{1}{n}\boldsymbol{1}\boldsymbol{1}^T = \boldsymbol{1}(\boldsymbol{1}^T\boldsymbol{1})^{-1}\boldsymbol{1}^T$ is the projection matrix on $C(\boldsymbol{1})$.

b) $E(MSE) = \sigma^2$, so $E(SSE) = (n-r)\sigma^2$. By a) and Theorem 2.5,

$$E(SSR) = \boldsymbol{\beta}^T\boldsymbol{X}^T(\boldsymbol{P} - \frac{\boldsymbol{1}\boldsymbol{1}^T}{n})\boldsymbol{X}\boldsymbol{\beta} + tr[\sigma^2(\boldsymbol{P} - \frac{\boldsymbol{1}\boldsymbol{1}^T}{n})] = \boldsymbol{\beta}^T\boldsymbol{X}^T(\boldsymbol{P} - \frac{\boldsymbol{1}\boldsymbol{1}^T}{n})\boldsymbol{X}\boldsymbol{\beta} + \sigma^2(r-1).$$

When H_0 is true $\boldsymbol{X}\boldsymbol{\beta} = \boldsymbol{1}\beta_1$ and $E(SSR) = \sigma^2(r-1)$.

c) By Theorem 2.14 g), if $\boldsymbol{Y} \sim N_n(\boldsymbol{\mu}, \sigma^2\boldsymbol{I})$ then $\frac{\boldsymbol{Y}^T\boldsymbol{A}\boldsymbol{Y}}{\sigma^2} \sim \chi^2\left(a, \frac{\boldsymbol{\mu}^T\boldsymbol{A}\boldsymbol{\mu}}{2\sigma^2}\right)$ iff $\boldsymbol{A}$ is idempotent with rank$(\boldsymbol{A}) = tr(\boldsymbol{A}) = a$.

i) Theorem 2.14 g) applies to SSE/σ^2 with $\boldsymbol{A} = \boldsymbol{I} - \boldsymbol{P}$ and $a = n - r$. Since $\boldsymbol{\mu} = \boldsymbol{X}\boldsymbol{\beta}$, and $\boldsymbol{\mu}^T(\boldsymbol{I} - \boldsymbol{P})\boldsymbol{\mu} = \boldsymbol{0}$ since $\boldsymbol{P}\boldsymbol{X} = \boldsymbol{X}$. Hence $SSE/\sigma^2 \sim \chi^2(n-r, 0) \sim \chi^2_{n-r}$. Thus $SSE \sim \sigma^2\chi^2_{n-r}$ regardless of whether H_0 is true or false.

ii) Theorem 2.14 g) applies to SSR/σ^2 with $\boldsymbol{A} = \boldsymbol{P} - \boldsymbol{P}_1$ and $a = r - 1$. If H_0 is true, then $\boldsymbol{\mu} = \boldsymbol{1}\beta_1$ and and $\boldsymbol{\mu}^T(\boldsymbol{P} - \boldsymbol{P}_1)\boldsymbol{\mu} = \boldsymbol{0}$ since $\boldsymbol{1}$ is the first column of $\boldsymbol{X}$ and $\boldsymbol{P}_1$ is the projection matrix on $C(\boldsymbol{1})$. Thus $\boldsymbol{P}\boldsymbol{1} = \boldsymbol{P}_1\boldsymbol{1} = \boldsymbol{1}$. Hence $SSR/\sigma^2 \sim \chi^2(r-1, 0) \sim \chi^2_{r-1}$. Thus $SSR \sim \sigma^2\chi^2_{r-1}$.

iii) SSE and SSR are independent by Craig's theorem since $(\boldsymbol{I} - \boldsymbol{P})(\boldsymbol{P} - \boldsymbol{P}_1) = \boldsymbol{P} - \boldsymbol{P}_1 - \boldsymbol{P} + \boldsymbol{P}_1 = \boldsymbol{0}$. MSE = SSE/(n-r) and MSR = SSR/(r-1). Thus

$$MSE/MSR = \frac{SSR/[\sigma^2(r-1)]}{SSE/[\sigma^2(n-r)]} \sim F_{r-1,n-r}.$$

3.13 a) i) Let $\boldsymbol{a}$ and $\boldsymbol{b}$ be constant vectors. Then $\boldsymbol{a}^T\boldsymbol{\beta}$ is estimable if there exists a linear unbiased estimator $\boldsymbol{b}^T\boldsymbol{Y}$ so $E(\boldsymbol{b}^T\boldsymbol{Y}) = \boldsymbol{a}^T\boldsymbol{\beta}$. Also, the quantity $\boldsymbol{a}^T\boldsymbol{\beta}$ is estimable iff $\boldsymbol{a}^T = \boldsymbol{b}^T\boldsymbol{X}$ iff $\boldsymbol{a} = \boldsymbol{X}^T\boldsymbol{b}$ iff $\boldsymbol{a} \in C(\boldsymbol{X}^T)$.

ii) Let a least squares estimator $\hat{\boldsymbol{\beta}}$ be any solution to the normal equations $\boldsymbol{X}^T\boldsymbol{X}\hat{\boldsymbol{\beta}} = \boldsymbol{X}^T\boldsymbol{Y}$. Then the least squares estimator of $\boldsymbol{a}^T\boldsymbol{\beta}$ is $\boldsymbol{a}^T\hat{\boldsymbol{\beta}} = \boldsymbol{b}^T\boldsymbol{X}\hat{\boldsymbol{\beta}} = \boldsymbol{b}^T\boldsymbol{P}\boldsymbol{Y}$.

iii) $MSE = \boldsymbol{Y}^T(\boldsymbol{I} - \boldsymbol{P})\boldsymbol{Y}/(n-r) = SSE/(n-r)$.

b) ii) $E(\boldsymbol{b}^T\boldsymbol{P}\boldsymbol{Y}) = \boldsymbol{b}^T\boldsymbol{P}\boldsymbol{X}\boldsymbol{\beta} = \boldsymbol{b}^T\boldsymbol{X}\boldsymbol{\beta} = \boldsymbol{a}^T\boldsymbol{\beta}$.

iii) $E(SSE) = E(\boldsymbol{Y}^T(\boldsymbol{I} - \boldsymbol{P})\boldsymbol{Y}) = tr[\sigma^2(\boldsymbol{I} - \boldsymbol{P})\boldsymbol{I}] + \boldsymbol{\mu}^T(\boldsymbol{I} - \boldsymbol{P})\boldsymbol{\mu}$ by Theorem 2.5 where $\boldsymbol{\mu} = \boldsymbol{X}\boldsymbol{\beta}$. Hence $E(SSE) = \sigma^2(tr(\boldsymbol{I} - \boldsymbol{P}) = \sigma^2(n-r)$. Hence $E(MSE) = E(SSE)/(n-r) = \sigma^2$.

c) If $\boldsymbol{a}^T\boldsymbol{\beta}$ is estimable and a least squares estimator $\hat{\boldsymbol{\beta}}$ is any solution to the normal equations $\boldsymbol{X}^T\boldsymbol{X}\hat{\boldsymbol{\beta}} = \boldsymbol{X}^T\boldsymbol{Y}$, then $\boldsymbol{a}^T\hat{\boldsymbol{\beta}}$ is the unique BLUE of $\boldsymbol{a}^T\boldsymbol{\beta}$.

d) $SSE = \boldsymbol{Y}^T(\boldsymbol{I} - \boldsymbol{P})\boldsymbol{Y}$ and $SSR = \boldsymbol{Y}^T(\boldsymbol{P} - \frac{1}{n}\boldsymbol{1}\boldsymbol{1}^T)\boldsymbol{Y} = \boldsymbol{Y}^T(\boldsymbol{P} - \boldsymbol{P}_1)\boldsymbol{Y}$ where $\boldsymbol{P}_1 = \frac{1}{n}\boldsymbol{1}\boldsymbol{1}^T = \boldsymbol{1}(\boldsymbol{1}^T\boldsymbol{1})^{-1}\boldsymbol{1}^T$ is the projection matrix on $C(\boldsymbol{1})$.

3.14 a) Note that $\boldsymbol{\beta}$ is estimable for i) since $\boldsymbol{X}$ for i) has full rank 2. Note that $\boldsymbol{\beta}$ is not estimable for ii) since $\boldsymbol{X}$ for ii) does not have full rank ($rank(\boldsymbol{X}) = 1$).

b)

$$\boldsymbol{B} = (\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T = \left(\begin{bmatrix} 2 & 1 & 0 \\ 0 & 1 & 2 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 1 & 1 \\ 0 & 2 \end{bmatrix} \right)^{-1} \boldsymbol{X}^T = \begin{bmatrix} 5 & 1 \\ 1 & 5 \end{bmatrix}^{-1} \boldsymbol{X}^T.$$

If

$$\boldsymbol{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

and $d = a_{11}a_{22} - a_{21}a_{12} \neq 0$, then

$$\boldsymbol{A}^{-1} = \frac{1}{d} \begin{bmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{bmatrix}.$$

Thus

$$\boldsymbol{B} = \frac{1}{24} \begin{bmatrix} 5 & -1 \\ -1 & 5 \end{bmatrix} \begin{bmatrix} 2 & 1 & 0 \\ 0 & 1 & 2 \end{bmatrix} = \frac{1}{24} \begin{bmatrix} 10 & 4 & -2 \\ -2 & 4 & 10 \end{bmatrix}.$$

c) Note that $\boldsymbol{b}^T\boldsymbol{Y}$ is an unbiased estimator of $\boldsymbol{b}^T\boldsymbol{X}\boldsymbol{\beta} = \boldsymbol{a}^T\boldsymbol{\beta}$ with $\boldsymbol{a}^T = \boldsymbol{b}^T\boldsymbol{X}$. If $\boldsymbol{b} = \boldsymbol{1}$, then

$$\boldsymbol{a}^T = \boldsymbol{1}^T\boldsymbol{X} = (1 \;\; 1 \;\; 1) \begin{bmatrix} 3 & 6 \\ 2 & 4 \\ 1 & 2 \end{bmatrix} = (6 \;\; 12).$$

Thus the estimable function $\boldsymbol{a}^T\boldsymbol{\beta} = 6\beta_1 + 12\beta_2$ has unbiased estimator $\boldsymbol{b}^T\boldsymbol{Y} = \boldsymbol{1}^T\boldsymbol{Y} = Y_1 + Y_2 + Y_3$.

Alternatively, let $\boldsymbol{b} = \boldsymbol{1}$ and $\boldsymbol{a}$ be as above. Then the unbiased least squares estimator $\boldsymbol{a}^T\hat{\boldsymbol{\beta}} = \boldsymbol{b}^T\boldsymbol{P}\boldsymbol{Y}$ where

$$\boldsymbol{P} = \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix} \left[(3 \;\; 2 \;\; 1) \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix} \right]^{-1} (3 \;\; 2 \;\; 1) = \frac{1}{14} \begin{bmatrix} 9 & 6 & 3 \\ 6 & 4 & 2 \\ 3 & 2 & 1 \end{bmatrix}.$$

Since $\boldsymbol{b} = \boldsymbol{1}$, the unbiased least squares estimator is

$$\frac{1}{14}(18 \;\; 12 \;\; 6) \begin{bmatrix} Y_1 \\ Y_2 \\ Y_3 \end{bmatrix} = \frac{18}{14}Y_1 + \frac{12}{14}Y_2 + \frac{6}{14}Y_3.$$

Since $E(\boldsymbol{Y}) = \boldsymbol{X\beta}$, note that $E(\boldsymbol{a}^T\hat{\boldsymbol{\beta}}) =$

$$\frac{18}{14}(3\beta_1+6\beta_2)+\frac{12}{14}(2\beta_1+4\beta_2)+\frac{6}{14}(\beta_1+2\beta_2) = (84/14)\beta_1+(168/14)\beta_2 = 6\beta_1+12\beta_2.$$

Chapter 4

4.11 a) $(\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T E(\boldsymbol{Y}^*) = (\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T\boldsymbol{X}\hat{\boldsymbol{\beta}} = \hat{\boldsymbol{\beta}}$.

b) $\boldsymbol{A}Cov(\boldsymbol{Y}^*)\boldsymbol{A}^T = (\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T diag(r_i^2)\boldsymbol{X}(\boldsymbol{X}^T\boldsymbol{X})^{-1}$.

c) We will use $\boldsymbol{X}\hat{\boldsymbol{\beta}} = \boldsymbol{PY}$ and $\boldsymbol{PX}_I = \boldsymbol{X}_I$. Then $E(\hat{\boldsymbol{\beta}}_I^*) = (\boldsymbol{X}_I^T\boldsymbol{X}_I)^{-1}\boldsymbol{X}_I^T E(\boldsymbol{Y}^*) = (\boldsymbol{X}_I^T\boldsymbol{X}_I)^{-1}\boldsymbol{X}_I^T\boldsymbol{X}\hat{\boldsymbol{\beta}} = (\boldsymbol{X}_I^T\boldsymbol{X}_I)^{-1}\boldsymbol{X}_I^T\boldsymbol{PY} = (\boldsymbol{X}_I^T\boldsymbol{X}_I)^{-1}\boldsymbol{X}_I^T\boldsymbol{Y} = \hat{\boldsymbol{\beta}}_I$.

d) $\boldsymbol{A}Cov(\boldsymbol{Y}^*)\boldsymbol{A}^T = (\boldsymbol{X}_I^T\boldsymbol{X}_I)^{-1}\boldsymbol{X}_I^T diag(r_i^2)\boldsymbol{X}_I(\boldsymbol{X}_I^T\boldsymbol{X}_I)^{-1}$.

Chapter 10

10.1

a) Since Y is a (random) scalar and $E(\boldsymbol{w}) = \boldsymbol{0}$, $\boldsymbol{\Sigma}_{\boldsymbol{u},Y} = E[(\boldsymbol{u}-E(\boldsymbol{u}))(Y - E(Y))^T] = E[\boldsymbol{w}(Y-E(Y))] = E(\boldsymbol{w}Y) - E(\boldsymbol{w})E(Y) = E(\boldsymbol{w}Y)$.

b) Using the definition of z and $\boldsymbol{r}$, note that $Y = m(z) + e$ and $\boldsymbol{w} = \boldsymbol{r} + (\boldsymbol{\Sigma}_{\boldsymbol{u}}\boldsymbol{\eta})\boldsymbol{\eta}^T\boldsymbol{w}$. Hence $E(\boldsymbol{w}Y) = E[(\boldsymbol{r} + (\boldsymbol{\Sigma}_{\boldsymbol{u}}\boldsymbol{\eta})\boldsymbol{\eta}^T\boldsymbol{w})(m(z)+e)] = E[(\boldsymbol{r} + (\boldsymbol{\Sigma}_{\boldsymbol{u}}\boldsymbol{\eta})\boldsymbol{\eta}^T\boldsymbol{w})m(z)] + E[\boldsymbol{r} + (\boldsymbol{\Sigma}_{\boldsymbol{u}}\boldsymbol{\eta})\boldsymbol{\eta}^T\boldsymbol{w}]E(e)$ since e is independent of $\boldsymbol{x}$. Since $E(e) = 0$, the latter term drops out. Since $m(z)$ and $\boldsymbol{\eta}^T\boldsymbol{w}m(z)$ are (random) scalars, $E(\boldsymbol{w}Y) = E[m(z)\boldsymbol{r}] + E[\boldsymbol{\eta}^T\boldsymbol{w}\ m(z)]\boldsymbol{\Sigma}_{\boldsymbol{u}}\boldsymbol{\eta}$.

c) Using result b), $\boldsymbol{\Sigma}_{\boldsymbol{u}}^{-1}\boldsymbol{\Sigma}_{\boldsymbol{u},Y} = \boldsymbol{\Sigma}_{\boldsymbol{u}}^{-1}E[m(z)\boldsymbol{r}] + \boldsymbol{\Sigma}_{\boldsymbol{u}}^{-1}E[\boldsymbol{\eta}^T\boldsymbol{w}\ m(z)]\boldsymbol{\Sigma}_{\boldsymbol{u}}\boldsymbol{\eta} = E[\boldsymbol{\eta}^T\boldsymbol{w}\ m(z)]\boldsymbol{\Sigma}_{\boldsymbol{u}}^{-1}\boldsymbol{\Sigma}_{\boldsymbol{u}}\boldsymbol{\eta} + \boldsymbol{\Sigma}_{\boldsymbol{u}}^{-1}E[m(z)\boldsymbol{r}] = E[\boldsymbol{\eta}^T\boldsymbol{w}\ m(z)]\boldsymbol{\eta} + \boldsymbol{\Sigma}_{\boldsymbol{u}}^{-1}E[m(z)\boldsymbol{r}]$ and the result follows.

d) $E(\boldsymbol{w}z) = E[(\boldsymbol{u}-E(\boldsymbol{u}))\boldsymbol{u}^T\boldsymbol{\eta}] = E[(\boldsymbol{u}-E(\boldsymbol{u}))(\boldsymbol{u}^T - E(\boldsymbol{u}^T) + E(\boldsymbol{u}^T))\boldsymbol{\eta}] = E[(\boldsymbol{u}-E(\boldsymbol{u}))(\boldsymbol{u}^T - E(\boldsymbol{u}^T))]\boldsymbol{\eta} + E[\boldsymbol{u}-E(\boldsymbol{u})]E(\boldsymbol{u}^T)\boldsymbol{\eta} = \boldsymbol{\Sigma}_{\boldsymbol{u}}\boldsymbol{\eta}$.

e) If $m(z) = z$, then $c(\boldsymbol{u}) = E(\boldsymbol{\eta}^T\boldsymbol{w}z) = \boldsymbol{\eta}^T E(\boldsymbol{w}z) = \boldsymbol{\eta}^T\boldsymbol{\Sigma}_{\boldsymbol{u}}\boldsymbol{\eta} = 1$ by result d).

f) Since z is a (random) scalar, $E(z\boldsymbol{r}) = E(\boldsymbol{r}z) = E[(\boldsymbol{w} - (\boldsymbol{\Sigma}_{\boldsymbol{u}}\boldsymbol{\eta})\boldsymbol{\eta}^T\boldsymbol{w})z] = E(\boldsymbol{w}z) - (\boldsymbol{\Sigma}_{\boldsymbol{u}}\boldsymbol{\eta})\boldsymbol{\eta}^T E(\boldsymbol{w}z)$. Using result d), $E(\boldsymbol{r}z) = \boldsymbol{\Sigma}_{\boldsymbol{u}}\boldsymbol{\eta} - \boldsymbol{\Sigma}_{\boldsymbol{u}}\boldsymbol{\eta}\boldsymbol{\eta}^T\boldsymbol{\Sigma}_{\boldsymbol{u}}\boldsymbol{\eta} = \boldsymbol{\Sigma}_{\boldsymbol{u}}\boldsymbol{\eta} - \boldsymbol{\Sigma}_{\boldsymbol{u}}\boldsymbol{\eta} = \boldsymbol{0}$.

g) Since z and $\boldsymbol{r}$ are linear combinations of $\boldsymbol{u}$, the joint distribution of z and $\boldsymbol{r}$ is multivariate normal. Since $E(\boldsymbol{r}) = \boldsymbol{0}$, z and $\boldsymbol{r}$ are uncorrelated and thus independent. Hence $m(z)$ and $\boldsymbol{r}$ are independent and $\boldsymbol{b}(\boldsymbol{u}) = \boldsymbol{\Sigma}_{\boldsymbol{u}}^{-1}E[m(z)\boldsymbol{r}] = \boldsymbol{\Sigma}_{\boldsymbol{u}}^{-1}E[m(z)]E(\boldsymbol{r}) = \boldsymbol{0}$.

11.3 Tables

Tabled values are F(k,d, 0.95) where $P(F < F(k, d, 0.95)) = 0.95$.
00 stands for ∞. Entries were produced with the `qf(.95,k,d)` command in *R*. The numerator degrees of freedom are k while the denominator degrees of freedom are d.

k d	1	2	3	4	5	6	7	8	9	00
1	161	200	216	225	230	234	237	239	241	254
2	18.5	19.0	19.2	19.3	19.3	19.3	19.4	19.4	19.4	19.5
3	10.1	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.53
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.63
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.37
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	3.67
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.23
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	2.93
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	2.71
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.54
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90	2.41
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.30
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71	2.21
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65	2.13
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59	2.07
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.01
17	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49	1.96
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46	1.92
19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42	1.88
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39	1.84
25	4.24	3.39	2.99	2.76	2.60	2.49	2.40	2.34	2.28	1.71
30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21	1.62
00	3.84	3.00	2.61	2.37	2.21	2.10	2.01	1.94	1.88	1.00

Tabled values are $t_{\alpha,d}$ where $P(t < t_{\alpha,d}) = \alpha$ where t has a t distribution with d degrees of freedom. If $d > 29$ use the $N(0,1)$ cutoffs $d = Z = \infty$.

					alpha					pvalue
d	0.005	0.01	0.025	0.05	0.5	0.95	0.975	0.99	0.995	left tail
1	-63.66	-31.82	-12.71	-6.314	0	6.314	12.71	31.82	63.66	
2	-9.925	-6.965	-4.303	-2.920	0	2.920	4.303	6.965	9.925	
3	-5.841	-4.541	-3.182	-2.353	0	2.353	3.182	4.541	5.841	
4	-4.604	-3.747	-2.776	-2.132	0	2.132	2.776	3.747	4.604	
5	-4.032	-3.365	-2.571	-2.015	0	2.015	2.571	3.365	4.032	
6	-3.707	-3.143	-2.447	-1.943	0	1.943	2.447	3.143	3.707	
7	-3.499	-2.998	-2.365	-1.895	0	1.895	2.365	2.998	3.499	
8	-3.355	-2.896	-2.306	-1.860	0	1.860	2.306	2.896	3.355	
9	-3.250	-2.821	-2.262	-1.833	0	1.833	2.262	2.821	3.250	
10	-3.169	-2.764	-2.228	-1.812	0	1.812	2.228	2.764	3.169	
11	-3.106	-2.718	-2.201	-1.796	0	1.796	2.201	2.718	3.106	
12	-3.055	-2.681	-2.179	-1.782	0	1.782	2.179	2.681	3.055	
13	-3.012	-2.650	-2.160	-1.771	0	1.771	2.160	2.650	3.012	
14	-2.977	-2.624	-2.145	-1.761	0	1.761	2.145	2.624	2.977	
15	-2.947	-2.602	-2.131	-1.753	0	1.753	2.131	2.602	2.947	
16	-2.921	-2.583	-2.120	-1.746	0	1.746	2.120	2.583	2.921	
17	-2.898	-2.567	-2.110	-1.740	0	1.740	2.110	2.567	2.898	
18	-2.878	-2.552	-2.101	-1.734	0	1.734	2.101	2.552	2.878	
19	-2.861	-2.539	-2.093	-1.729	0	1.729	2.093	2.539	2.861	
20	-2.845	-2.528	-2.086	-1.725	0	1.725	2.086	2.528	2.845	
21	-2.831	-2.518	-2.080	-1.721	0	1.721	2.080	2.518	2.831	
22	-2.819	-2.508	-2.074	-1.717	0	1.717	2.074	2.508	2.819	
23	-2.807	-2.500	-2.069	-1.714	0	1.714	2.069	2.500	2.807	
24	-2.797	-2.492	-2.064	-1.711	0	1.711	2.064	2.492	2.797	
25	-2.787	-2.485	-2.060	-1.708	0	1.708	2.060	2.485	2.787	
26	-2.779	-2.479	-2.056	-1.706	0	1.706	2.056	2.479	2.779	
27	-2.771	-2.473	-2.052	-1.703	0	1.703	2.052	2.473	2.771	
28	-2.763	-2.467	-2.048	-1.701	0	1.701	2.048	2.467	2.763	
29	-2.756	-2.462	-2.045	-1.699	0	1.699	2.045	2.462	2.756	
Z	-2.576	-2.326	-1.960	-1.645	0	1.645	1.960	2.326	2.576	
CI						90%	95%		99%	
	0.995	0.99	0.975	0.95	0.5	0.05	0.025	0.01	0.005	right tail
	0.01	0.02	0.05	0.10	1	0.10	0.05	0.02	0.01	two tail