STK3100 Exercises, Week 2

Vinnie Ko, Jonas Moss

September 14, 2018

Exercise 2.22

Consider the null model $\mathrm{E}[Y_i] = \beta$, $i = 1, 2$.

(a)

$$\boldsymbol{X} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \boldsymbol{\beta} = \begin{bmatrix} \beta_0 \end{bmatrix}$$

The model space is $C(\boldsymbol{X}) = \left\{ \boldsymbol{\eta} = \boldsymbol{X}\boldsymbol{\beta} = \begin{bmatrix} \beta_0 \\ \beta_0 \end{bmatrix} \mid \beta_0 \in \mathbb{R} \right\}$.
Its orthogonal complement is $C(\boldsymbol{X})^{\perp} = \left\{ \boldsymbol{\tau} = \begin{bmatrix} -a \\ a \end{bmatrix} \mid a \in \mathbb{R} \right\}$ because we have $\boldsymbol{\eta}^{\mathrm{T}}\boldsymbol{\tau} = 0$.

$$\begin{aligned} \boldsymbol{P}_X &= \boldsymbol{X}(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X})^{-1}\boldsymbol{X}^{\mathrm{T}} = \begin{bmatrix} 0.5 & 0.5 \\ 0.5 & 0.5 \end{bmatrix} \\ \boldsymbol{I} - \boldsymbol{P}_X &= \begin{bmatrix} 0.5 & -0.5 \\ -0.5 & 0.5 \end{bmatrix} \end{aligned}$$

(b)

$$\boldsymbol{y} = \begin{bmatrix} 5 \\ 10 \end{bmatrix}, \quad \boldsymbol{X} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \widehat{\boldsymbol{\beta}} = (\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X})^{-1}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{y} = \begin{bmatrix} 7.5 \end{bmatrix}$$

$$\widehat{\boldsymbol{\mu}} = \boldsymbol{H}\boldsymbol{y} = \boldsymbol{X}(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X})^{-1}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{y} = \begin{bmatrix} 7.5 \\ 7.5 \end{bmatrix}$$

Sum of squares decomposition

$$\sum_{i=1}^{n} y_i^2 = n\overline{y}^2 + \sum_{i=1}^{n} (y_i - \overline{y})^2 \qquad \text{(p.42 of the book)}$$

$$s = \widehat{\sigma} = \sqrt{\frac{(\boldsymbol{y} - \widehat{\boldsymbol{\mu}})^{\mathrm{T}}(\boldsymbol{y} - \widehat{\boldsymbol{\mu}})}{n - p}} = \sqrt{\frac{\sum_{i=1}^{n} (y_i - \overline{y})^2}{n - p}} = 3.5355$$

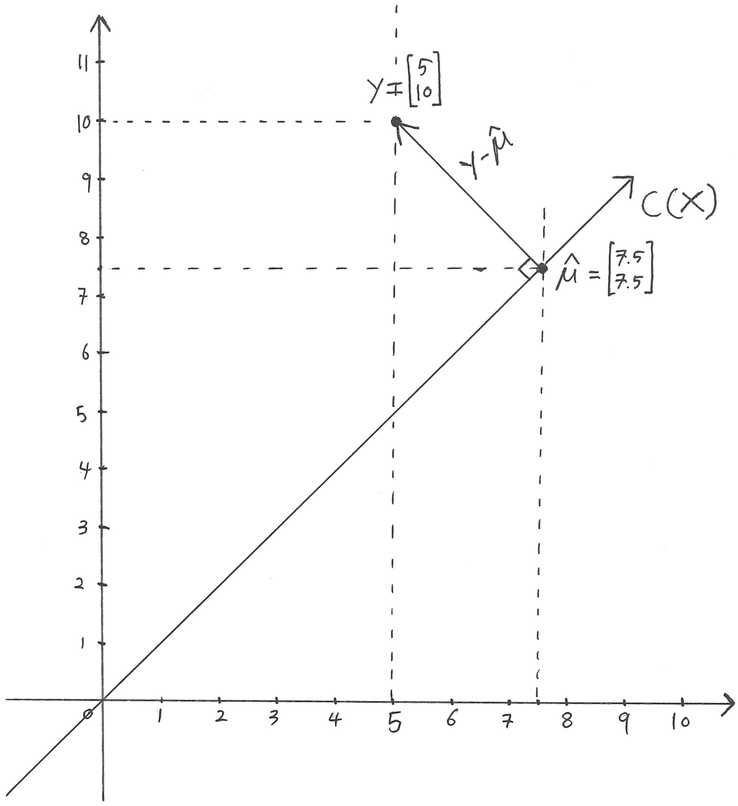

Figure 1: Visualization of $\boldsymbol{y}, \widehat{\boldsymbol{\mu}}, C(\boldsymbol{X})$

Exercise 2.23

Consider the saturated model $\mathrm{E}[Y_i] = \beta_i$, $i = 1, \cdots, n$.

(a)

$$\boldsymbol{X} = \boldsymbol{I}_n, \quad C(\boldsymbol{X}) = \left\{ \boldsymbol{\eta} = \boldsymbol{X}\boldsymbol{\beta} = \boldsymbol{\beta} = \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_n \end{bmatrix} \mid \boldsymbol{\beta} \in \mathbb{R}^n \right\} = \mathbb{R}^n$$

Its orthogonal complement is $C(\boldsymbol{X})^\perp = \{\boldsymbol{\tau} = \boldsymbol{0}_n\}$ because we have $\boldsymbol{\eta}^\mathrm{T}\boldsymbol{\tau} = 0$.

$$\begin{aligned} \boldsymbol{P}_X &= \boldsymbol{X}(\boldsymbol{X}^\mathrm{T}\boldsymbol{X})^{-1}\boldsymbol{X}^\mathrm{T} = \boldsymbol{I}_n \\ \boldsymbol{I} - \boldsymbol{P}_X &= \boldsymbol{0}_{n\times n} \end{aligned}$$

(b)

$$\widehat{\boldsymbol{\beta}} = \boldsymbol{y}, \quad \widehat{\boldsymbol{\mu}} = \widehat{\boldsymbol{\beta}} = \boldsymbol{y}$$

$$s = \sqrt{\frac{\sum_{i=1}^{n}(y_i - y_i)^2}{n-n}} = \frac{0}{0}$$

This saturated model is not sensible for practical use because it's not processing anything. It merely returns the raw information from data as output, namely $\widehat{\mu}_i = y_i$.

Exercise 3.2

Note that we solve this exercise for more general case where T has noncentral t-distribution.

It's given that $T \sim t_{p,\mu}$. We can rewrite T as

$$T = \frac{Z}{\sqrt{X/p}}$$

where $Z \sim N(\mu, 1)$, $X \sim \chi_p^2$ and $Z \perp\!\!\!\perp X$. (The notation $A \perp\!\!\!\perp B$ means that A and B are independent.) Let $W = Z^2$, then $W \sim \chi^2_{1,\mu^2}$ (noncentral chi-squared distribution).

Now we look at T^2:

$$T^2 = \frac{Z^2}{X/p} = \frac{W/1}{X/p}$$

where $W \sim \chi_{1,\mu^2}$, $X \sim \chi_p^2$ and $W \perp\!\!\!\perp X$.

Then, by definition, $T^2 \sim F_{1,p,\mu^2}$ (noncentral F-distribution).

If we let $\mu = 0$, then we have the (central) F-distribution.

The relationship between noncentral t-distribution and noncentral F-distribution:

$$T \sim t_{p,\mu} \;\Rightarrow\; F = T^2 \sim F_{1,p,\mu^2}$$

Exercise 3.4

Sum-of-squares decomposition (p.42 of the book): $\boldsymbol{Y}^{\mathrm{T}}\boldsymbol{Y} = \boldsymbol{Y}^{\mathrm{T}}\boldsymbol{P}_X\boldsymbol{Y} + \boldsymbol{Y}^{\mathrm{T}}(\boldsymbol{I} - \boldsymbol{P}_X)\boldsymbol{Y}$.

Since $\boldsymbol{P}_X + (\boldsymbol{I} - \boldsymbol{P}_X) = \boldsymbol{I}$, $\boldsymbol{Y}^{\mathrm{T}}\boldsymbol{P}_X\boldsymbol{Y} \perp\!\!\!\perp \boldsymbol{Y}^{\mathrm{T}}(\boldsymbol{I} - \boldsymbol{P}_X)\boldsymbol{Y}$ by Cochran's theorem.

In case of null model, $\boldsymbol{P}_X = \dfrac{1}{n}\mathbf{1}_n\mathbf{1}_n^{\mathrm{T}}$ and we have

$$\boldsymbol{Y}^{\mathrm{T}}\boldsymbol{P}_X\boldsymbol{Y} = \boldsymbol{Y}^{\mathrm{T}}\overline{Y}\mathbf{1}_n = \overline{Y}\boldsymbol{Y}^{\mathrm{T}}\mathbf{1}_n = n\overline{Y}^2$$

$$\boldsymbol{Y}^{\mathrm{T}}(\boldsymbol{I} - \boldsymbol{P}_X)\boldsymbol{Y} = \boldsymbol{Y}^{\mathrm{T}}(\boldsymbol{I} - \boldsymbol{P}_X)^{\mathrm{T}}(\boldsymbol{I} - \boldsymbol{P}_X)\boldsymbol{Y} = \left(\boldsymbol{Y} - \overline{Y}\mathbf{1}_n\right)^{\mathrm{T}}\left(\boldsymbol{Y} - \overline{Y}\mathbf{1}_n\right) = \sum_{i=1}^{n}(Y_i - \overline{Y})^2$$

So, $n\overline{Y}^2 \perp\!\!\!\perp \displaystyle\sum_{i=1}^{n}(Y_i - \overline{Y})^2$ which implies $\overline{Y}^2 \perp\!\!\!\perp \dfrac{\sum_{i=1}^{n}(Y_i - \overline{Y})^2}{n-1} = S^2$.

Exercise 3.5

Vinnie's solution

We apply the sum-of-squares decomposition to $(\boldsymbol{Y}-\boldsymbol{\mu}_0)$ where $\boldsymbol{Y} \sim N(\boldsymbol{\mu}, \sigma^2 \boldsymbol{I})$, $\boldsymbol{\mu_0} = \mu_0 \mathbf{1}_n$, $\boldsymbol{\mu} = \mu \mathbf{1}_n$:

$$(\boldsymbol{Y}-\boldsymbol{\mu}_0)^{\mathrm{T}}(\boldsymbol{Y}-\boldsymbol{\mu}_0) = (\boldsymbol{Y}-\boldsymbol{\mu}_0)^{\mathrm{T}}\boldsymbol{P}_X(\boldsymbol{Y}-\boldsymbol{\mu}_0) + (\boldsymbol{Y}-\boldsymbol{\mu}_0)^{\mathrm{T}}(\boldsymbol{I}-\boldsymbol{P}_X)(\boldsymbol{Y}-\boldsymbol{\mu}_0).$$

Since $\boldsymbol{P}_X + (\boldsymbol{I}-\boldsymbol{P}_X) = \boldsymbol{I}$, Cochran's theorem gives $(\boldsymbol{Y}-\boldsymbol{\mu}_0)^{\mathrm{T}}\boldsymbol{P}_X(\boldsymbol{Y}-\boldsymbol{\mu}_0) \perp\!\!\!\perp (\boldsymbol{Y}-\boldsymbol{\mu}_0)^{\mathrm{T}}(\boldsymbol{I}-\boldsymbol{P}_X)(\boldsymbol{Y}-\boldsymbol{\mu}_0)$ and

$$\frac{(\boldsymbol{Y}-\boldsymbol{\mu}_0)^{\mathrm{T}}\boldsymbol{P}_X(\boldsymbol{Y}-\boldsymbol{\mu}_0)}{\sigma^2} \sim \chi^2_{r_1,\lambda_1}$$

$$\frac{(\boldsymbol{Y}-\boldsymbol{\mu}_0)^{\mathrm{T}}(\boldsymbol{I}-\boldsymbol{P}_X)(\boldsymbol{Y}-\boldsymbol{\mu}_0)}{\sigma^2} \sim \chi^2_{r_2,\lambda_2}$$

where $r_1 = \operatorname{rank}(\boldsymbol{P}_X)$, $\lambda_1 = \dfrac{(\boldsymbol{\mu}-\boldsymbol{\mu}_0)^{\mathrm{T}}\boldsymbol{P}_X(\boldsymbol{\mu}-\boldsymbol{\mu}_0)}{\sigma^2}$ and $r_2 = \operatorname{rank}(\boldsymbol{I}-\boldsymbol{P}_X)$, $\lambda_2 = \dfrac{(\boldsymbol{\mu}-\boldsymbol{\mu}_0)^{\mathrm{T}}(\boldsymbol{I}-\boldsymbol{P}_X)(\boldsymbol{\mu}-\boldsymbol{\mu}_0)}{\sigma^2}$.

In case of null model, $\boldsymbol{P}_X = \dfrac{1}{n}\mathbf{1}_n\mathbf{1}_n^{\mathrm{T}}$ and we have $\boldsymbol{P}_X\boldsymbol{Y} = \dfrac{1}{n}\mathbf{1}_n\mathbf{1}_n^{\mathrm{T}}\boldsymbol{Y} = \overline{Y}\mathbf{1}_n$, $\boldsymbol{P}_X\boldsymbol{\mu} = \dfrac{1}{n}\mathbf{1}_n\mathbf{1}_n^{\mathrm{T}}\mathbf{1}_n\mu = \mu\mathbf{1}_n = \boldsymbol{\mu}$.

Further,

$$\begin{aligned}
(\boldsymbol{Y}-\boldsymbol{\mu}_0)^{\mathrm{T}}\boldsymbol{P}_X(\boldsymbol{Y}-\boldsymbol{\mu}_0) &= (\boldsymbol{Y}-\boldsymbol{\mu}_0)^{\mathrm{T}}\boldsymbol{P}_X^{\mathrm{T}}\boldsymbol{P}_X(\boldsymbol{Y}-\boldsymbol{\mu}_0) \\
&= (\overline{Y}\mathbf{1}_n - \mu_0\mathbf{1}_n)^{\mathrm{T}}(\overline{Y}\mathbf{1}_n - \mu_0\mathbf{1}_n) = (\overline{Y}-\mu_0)^2\mathbf{1}_n^{\mathrm{T}}\mathbf{1}_n \\
&= n(\overline{Y}-\mu_0)^2 \\
(\boldsymbol{Y}-\boldsymbol{\mu}_0)^{\mathrm{T}}(\boldsymbol{I}-\boldsymbol{P}_X)(\boldsymbol{Y}-\boldsymbol{\mu}_0) &= (\boldsymbol{Y}-\boldsymbol{\mu}_0)^{\mathrm{T}}(\boldsymbol{I}-\boldsymbol{P}_X)^{\mathrm{T}}(\boldsymbol{I}-\boldsymbol{P}_X)(\boldsymbol{Y}-\boldsymbol{\mu}_0) \\
&= (\boldsymbol{Y}-\overline{Y}\mathbf{1}_n)^{\mathrm{T}}(\boldsymbol{Y}-\overline{Y}\mathbf{1}_n) \\
&= \sum_{i=1}^n (Y_i-\overline{Y})^2
\end{aligned}$$

and

$$\begin{aligned}
r_1 &= \operatorname{rank}(\boldsymbol{P}_X) = \operatorname{rank}(\frac{1}{n}\mathbf{1}_n\mathbf{1}_n^{\mathrm{T}}) = 1 \\
r_2 &= \operatorname{rank}(\boldsymbol{I}-\boldsymbol{P}_X) = n-1 \\
\lambda_1 &= \frac{(\boldsymbol{\mu}-\boldsymbol{\mu}_0)^{\mathrm{T}}\boldsymbol{P}_X(\boldsymbol{\mu}-\boldsymbol{\mu}_0)}{\sigma^2} = \frac{n(\mu-\mu_0)^2}{\sigma^2} \\
\lambda_2 &= \frac{(\boldsymbol{\mu}-\boldsymbol{\mu}_0)^{\mathrm{T}}(\boldsymbol{I}-\boldsymbol{P}_X)(\boldsymbol{\mu}-\boldsymbol{\mu}_0)}{\sigma^2} = \frac{\sum_{i=1}^n(\mu-\mu)^2}{\sigma^2} = 0.
\end{aligned}$$

So, we have

$$\frac{n(\overline{Y}-\mu_0)^2}{\sigma^2} \sim \chi^2_{1,\lambda_1} \quad \text{and} \quad \frac{(n-1)S^2}{\sigma^2} = \frac{\sum_{i=1}^n(Y_i-\overline{Y})^2}{\sigma^2} \sim \chi^2_{n-1}$$

and they are independent.

Then, by the definition of noncentral F-distribution,

$$F = \frac{n(\overline{Y}-\mu_0)^2}{S^2} = \frac{n(\overline{Y}-\mu_0)^2/1}{\sum_{i=1}^n(Y_i-\overline{Y})^2/(n-1)} \sim F_{1,n-1,\lambda_1}.$$

F is our test statistic.

Null hypothesis $H_0 : \mu = \mu_0$. Under this H_0, $\lambda_1 = \dfrac{n(\mu_0 - \mu_0)^2}{\sigma^2} = 0$. So, our test statistic follows (central) F-distribution: $F \sim F_{1,n-1}$.

A level α F-test rejects the null hypothesis if $F > f_{1-\alpha,1,n-1}$.

Now, we construct a t-test.

Since $Y_i, \cdots Y_n \overset{i.i.d.}{\sim} N(\mu, \sigma^2)$, $\overline{Y} \sim N\left(\mu, \dfrac{\sigma^2}{n}\right)$ and by transformation $\dfrac{\sqrt{n}(\overline{Y} - \mu_0)}{\sigma} \sim N\left(\dfrac{\sqrt{n}(\mu - \mu_0)}{\sigma}, 1\right)$.

It's well known that $\dfrac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1}$. (You can easily show this by using the result from p.82 of the book.) By combining those results, we have

$$T = \frac{\frac{\sqrt{n}(\overline{Y}-\mu_0)}{\sigma}}{\sqrt{\frac{(n-1)S^2}{\sigma^2}/(n-1)}} = \frac{\sqrt{n}(\overline{Y} - \mu_0)}{S} \sim t_{n-1,\sqrt{\lambda_1}}.$$

Under $H_0 : \mu = \mu_0$, $\lambda_1 = \dfrac{n(\mu_0 - \mu_0)^2}{\sigma^2} = 0$. So, our test statistic follows (central) t-distribution: $T \sim t_{n-1}$.

A level α two-sided t-test rejects the null hypothesis if $|T| > t_{1-\frac{\alpha}{2},n-1}$.

Jonas' solution

Assume μ is the true mean and assume $\mu_0 = 0$ without loss of generality. Then

$$\begin{aligned} \sigma^{-2} Y^T P Y &\sim \chi^2_{1,\sigma^{-2}\mu^T P_0 \mu}, \\ \sigma^{-2} Y^T (I - P) Y &\sim \chi^2_{n-1}, \end{aligned}$$

independently by Cochran's theorem, as the rank of $I-P$ is $n-1$ and the rank of P is 1. The non-centraliy parameter of $\sigma^{-2} Y^T (I-P) Y$ is $\sigma^{-2}\mu^T (I-P)\mu = 0$. As for the non-centrality of $\sigma^{-2} Y^T P Y$, define $\theta = \dfrac{\mu}{\sigma}$. This is the *effect size*, which is used in power analysis and meta analysis. Typically, the non-central distributions are functions of $n\theta^2$. And now we observe that $\sigma^{-2}\mu^T P \mu = n\theta^2$, as $P = n^{-1} 1_n 1_n^T$. Then

$$\frac{Y^T P_X Y}{Y^T (I - P_X) Y} \sim F_{1,n-1,n\left(\frac{\mu}{\sigma}\right)^2}.$$

Under the null $\mu = 0$, this is the usual F-test. The t-test is arrived at similarily. Since $n^{-1/2}\sigma^{-1}\sum_{i=1}^n Y_i \sim N\left(n^{1/2}\theta, 1\right)$ and $\sigma^{-2} Y^T (I-P) Y = \sigma^{-2}(n-1) S^2 \sim \chi^2_{n-1}$, the t-statistic $t = n^{-1/2}\sum_{i=1}^n Y_i / S$ is distributed as $t_{n-1,\sqrt{n}\theta}$.

Exercise 3.7

(a)

In one-way ANOVA, the following sum of squares decomposition is used:

$$\boldsymbol{Y}^{\mathrm{T}}\boldsymbol{Y} = \boldsymbol{Y}^{\mathrm{T}}\boldsymbol{P}_0\boldsymbol{Y} + \boldsymbol{Y}^{\mathrm{T}}(\boldsymbol{P}_X - \boldsymbol{P}_0)\boldsymbol{Y} + \boldsymbol{Y}^{\mathrm{T}}(\boldsymbol{I} - \boldsymbol{P}_X)\boldsymbol{Y}.$$

(See p.46 of the book for details.)

The between-groups sum of squares is

$$\boldsymbol{Y}^{\mathrm{T}}\left(\boldsymbol{P}_X-\boldsymbol{P}_0\right)\boldsymbol{Y}=\sum_{i=1}^{c} n_i\left(\overline{Y}_i-\overline{Y}\right)^2.$$

By Cochran's theorem, $\dfrac{\boldsymbol{Y}^{\mathrm{T}}\left(\boldsymbol{P}_X-\boldsymbol{P}_0\right)\boldsymbol{Y}}{\sigma^2}\sim\chi^2_{r,\lambda}$ with $r=\operatorname{rank}(\boldsymbol{P}_X-\boldsymbol{P}_0)=c-1, \lambda=\dfrac{\boldsymbol{\mu}^{\mathrm{T}}\left(\boldsymbol{P}_X-\boldsymbol{P}_0\right)\boldsymbol{\mu}}{\sigma^2}$.

By replacing $\boldsymbol{Y}$ by $\boldsymbol{\mu}$ in the expression of between-groups sum of squares, we directly obtain

$$\lambda=\frac{\boldsymbol{\mu}^{\mathrm{T}}\left(\boldsymbol{P}_X-\boldsymbol{P}_0\right)\boldsymbol{\mu}}{\sigma^2}=\frac{1}{\sigma^2}\sum_{i=1}^{c} n_i\left(\mu_i-\overline{\mu}\right)^2$$

where $\overline{\mu}=\dfrac{1}{n}\displaystyle\sum_{i=1}^{c} n_i\mu_i$.

(b)

To perform F-test, we need to evaluate $\boldsymbol{Y}^{\mathrm{T}}\left(\boldsymbol{I}-\boldsymbol{P}_X\right)\boldsymbol{Y}$:

$$\boldsymbol{Y}^{\mathrm{T}}\left(\boldsymbol{I}-\boldsymbol{P}_X\right)\boldsymbol{Y}=\sum_{i=1}^{c}\sum_{j=1}^{n_i}\left(Y_{i,j}-\overline{Y}_i\right)^2.$$

By Cochran's theorem, $\dfrac{\boldsymbol{Y}^{\mathrm{T}}\left(\boldsymbol{I}-\boldsymbol{P}_X\right)\boldsymbol{Y}}{\sigma^2}\sim\chi^2_{r_*,\lambda_*}$ where $r_*=\operatorname{rank}(\boldsymbol{I}-\boldsymbol{P}_X)=n-c, \lambda_*=\dfrac{\boldsymbol{\mu}^{\mathrm{T}}\left(\boldsymbol{I}-\boldsymbol{P}_X\right)\boldsymbol{\mu}}{\sigma^2}=\dfrac{1}{\sigma^2}\displaystyle\sum_{i=1}^{c} n_i\left(\mu_i-\mu_i\right)^2=0$.

By the definition of noncentral F-distribution, we have the test statistic

$$F=\frac{\frac{\sum_{i=1}^{c} n_i\left(\overline{Y}_i-\overline{Y}\right)^2}{\sigma^2}/(c-1)}{\frac{\sum_{i=1}^{c}\sum_{j=1}^{n_i}\left(Y_{i,j}-\overline{Y}_i\right)^2}{\sigma^2}/(n-c)}\sim\frac{\chi^2_{c-1,\lambda}/(c-1)}{\chi^2_{n-c}/(n-c)}=F_{c-1,n-c,\lambda}$$

Now, we evaluate the noncentrality parameter λ by using given conditions:

$$\begin{aligned}
\lambda&=\frac{1}{\sigma^2}\sum_{i=1}^{c} n_i\left(\mu_i-\overline{\mu}\right)^2\\
&=\frac{n}{\sigma^2}\left[\left(\mu_1-\overline{\mu}\right)^2+\left(\mu_2-\overline{\mu}\right)^2+\left(\mu_3-\overline{\mu}\right)^2\right]\\
&=\frac{n}{9\sigma^2}\left[\left(3\mu_1-\left(\mu_1+\mu_2+\mu_3\right)\right)^2+\left(3\mu_2-\left(\mu_1+\mu_2+\mu_3\right)\right)^2+\left(3\mu_3-\left(\mu_1+\mu_2+\mu_3\right)\right)^2\right]\\
&=\frac{n}{9\sigma^2}\left[\left(\left(\mu_1-\mu_2\right)+\left(\mu_1-\mu_3\right)\right)^2+\left(-\left(\mu_1-\mu_2\right)+\left(\mu_2-\mu_3\right)\right)^2+\left(-\left(\mu_1-\mu_3\right)-\left(\mu_2-\mu_3\right)\right)^2\right]\\
&=\frac{n}{9\sigma^2}\left[\left(\left(\mu_1-\mu_2\right)+\left(\mu_1-\mu_3\right)\right)^2+\left(-\left(\mu_1-\mu_2\right)+\left(\mu_2-\mu_3\right)\right)^2+\left(-\left(\mu_1-\mu_3\right)-\left(\mu_2-\mu_3\right)\right)^2\right]\\
&=\frac{n}{9\sigma^2}\left[\left(\frac{\sigma}{2}+\sigma\right)^2+\left(-\frac{\sigma}{2}+\frac{\sigma}{2}\right)^2+\left(-\sigma-\frac{\sigma}{2}\right)^2\right]\\
&=\frac{n}{9\sigma^2}\left(\frac{9\sigma^2}{2}\right)\\
&=\frac{n}{2}.
\end{aligned}$$

So, our test statistic becomes

$$F_{2,n-3,\frac{n}{2}}$$

and level 0.05 F-test rejects the null hypothesis if $F > f_{0.95,2,n-3}$.

Now, we calculate power.

$$\begin{aligned}
\text{Power} &= \Pr\left(\text{reject } H_0 | H_1 \text{ is true}\right) \\
&= \Pr\left(F_{2,n-3,\frac{n}{2}} > f_{0.95,2,n-3}\right) \\
&= 1 - \Pr\left(F_{2,n-3,\frac{n}{2}} \le f_{0.95,2,n-3}\right) \\
&= 1 - G\left(f_{0.95,2,n-3}\right)
\end{aligned}$$

Calculation in R:

```
> alpha = 0.05
> c.val = 3
> n = c(10, 30, 50)
> lambda = n/2
>
> df.1 = c.val - 1
> df.2 = c.val*n - c.val
>
> critic.val = qf(1 - alpha, df.1 , df.2) # 0.95 quantile of F dist
> power.val = 1 - pf(critic.val, df.1, df.2, lambda) # right-tail prob. for noncentral
    F
>
> result.mat = data.frame(n = n, critic.val = critic.val, power.val = power.val)
> show(result.mat)
   n critic.val power.val
1 10   3.354131 0.4579923
2 30   3.101296 0.9362768
3 50   3.057621 0.9959038
```

(c)

$$\begin{aligned}
\lambda &= \frac{1}{\sigma^2}\sum_{i=1}^{c} n_i\left(\mu_i - \overline{\mu}\right)^2 \\
&= \frac{n}{\sigma^2}\left[\left(\mu_1 - \overline{\mu}\right)^2 + \left(\mu_2 - \overline{\mu}\right)^2 + \left(\mu_3 - \overline{\mu}\right)^2\right] \\
&= \frac{n}{9\sigma^2}\left[\left(3\mu_1 - (\mu_1+\mu_2+\mu_3)\right)^2 + \left(3\mu_2 - (\mu_1+\mu_2+\mu_3)\right)^2 + \left(3\mu_3 - (\mu_1+\mu_2+\mu_3)\right)^2\right] \\
&= \frac{n}{9\sigma^2}\left[\left((\mu_1-\mu_2) + (\mu_1-\mu_3)\right)^2 + \left(-(\mu_1-\mu_2) + (\mu_2-\mu_3)\right)^2 + \left(-(\mu_1-\mu_3) - (\mu_2-\mu_3)\right)^2\right] \\
&= \frac{n}{9\sigma^2}\left[\left((\mu_1-\mu_2) + (\mu_1-\mu_3)\right)^2 + \left(-(\mu_1-\mu_2) + (\mu_2-\mu_3)\right)^2 + \left(-(\mu_1-\mu_3) - (\mu_2-\mu_3)\right)^2\right] \\
&= \frac{n}{9\sigma^2}\left[\left(\Delta\sigma + 2\Delta\sigma\right)^2 + \left(-\Delta\sigma + \Delta\sigma\right)^2 + \left(-2\Delta\sigma - \Delta\sigma\right)^2\right] \\
&= \frac{n}{9\sigma^2}\cdot 2(3\Delta\sigma)^2 \\
&= 2n\Delta^2
\end{aligned}$$

Calculation in R:

```
> alpha = 0.05
> c.val = 3
> n = 10
> Delta = c(0, 0.5, 1)
> lambda = 2*n*Delta^2
>
> df.1 = c.val - 1
> df.2 = c.val*n - c.val
>
> critic.val = qf(1 - alpha, df.1 , df.2) # 0.95 quantile of F dist
> power.val = 1 - pf(critic.val, df.1, df.2, lambda) # right-tail prob. for noncentral
    F
>
> result.mat = data.frame(Delta = Delta, critic.val = critic.val, power.val = power.
   val)
> show(result.mat)
  Delta critic.val power.val
1   0.0   3.354131 0.0500000
2   0.5   3.354131 0.4579923
3   1.0   3.354131 0.9732551
```

Additional Exercise 2

Eigendecomposition of $\boldsymbol{V}$ gives us $\boldsymbol{V} = \boldsymbol{Q}\boldsymbol{\Lambda}\boldsymbol{Q}^{-1}$ with $\boldsymbol{\Lambda} = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_p \end{bmatrix}$ where λ_i's are eigenvalues.

Since $\boldsymbol{V}$ is real symmetric matrix, $\boldsymbol{Q}$ is orthogonal matrix (i.e. $\boldsymbol{Q}^{-1} = \boldsymbol{Q}^{\mathrm{T}}$, $\boldsymbol{Q}^{\mathrm{T}}\boldsymbol{Q} = \boldsymbol{Q}\boldsymbol{Q}^{\mathrm{T}} = \boldsymbol{I}$).

Because $\boldsymbol{V}$ is positive definite, all eigenvalues $\lambda_1, \cdots, \lambda_p$ are positive. We define

$$\boldsymbol{\Lambda}^{\frac{1}{2}} = \begin{bmatrix} \sqrt{\lambda_1} & & \\ & \ddots & \\ & & \sqrt{\lambda_p} \end{bmatrix} \quad \text{and} \quad \boldsymbol{\Lambda}^{-\frac{1}{2}} = \begin{bmatrix} \frac{1}{\sqrt{\lambda_1}} & & \\ & \ddots & \\ & & \frac{1}{\sqrt{\lambda_p}} \end{bmatrix}.$$

Then, $\boldsymbol{\Lambda}^{\frac{1}{2}}\boldsymbol{\Lambda}^{\frac{1}{2}} = \boldsymbol{\Lambda}$ and $\boldsymbol{\Lambda}^{-\frac{1}{2}}\boldsymbol{\Lambda}^{-\frac{1}{2}} = \boldsymbol{I}$.
Now we define $\boldsymbol{V}^{\frac{1}{2}} = \boldsymbol{Q}\boldsymbol{\Lambda}^{\frac{1}{2}}\boldsymbol{Q}^{\mathrm{T}}$ and $\boldsymbol{V}^{-\frac{1}{2}} = \boldsymbol{Q}\boldsymbol{\Lambda}^{-\frac{1}{2}}\boldsymbol{Q}^{\mathrm{T}}$.

Note that $\boldsymbol{V}^{\frac{1}{2}}$ and $\boldsymbol{V}^{-\frac{1}{2}}$ are symmetric, and $\boldsymbol{V}^{\frac{1}{2}}\boldsymbol{V}^{-\frac{1}{2}} = \boldsymbol{I}$. So, $\boldsymbol{V}^{-\frac{1}{2}} = \left(\boldsymbol{V}^{\frac{1}{2}}\right)^{-1}$. Further, $\boldsymbol{V}^{\frac{1}{2}}\boldsymbol{V}^{\frac{1}{2}} = \boldsymbol{V}$.

a)

We perform variable transformation $\boldsymbol{Z} = \boldsymbol{V}^{-\frac{1}{2}}(\boldsymbol{Y} - \boldsymbol{\mu})$ which gives $\boldsymbol{Z} \sim N\left(\boldsymbol{0}, \boldsymbol{V}^{-\frac{1}{2}}\boldsymbol{V}\left(\boldsymbol{V}^{-\frac{1}{2}}\right)^{\mathrm{T}}\right) = N(\boldsymbol{0}, \boldsymbol{I})$.

Thus, we have

$$
\begin{aligned}
(\boldsymbol{Y}-\boldsymbol{\mu})^{\mathrm{T}}\boldsymbol{V}^{-1}(\boldsymbol{Y}-\boldsymbol{\mu}) &= (\boldsymbol{Y}-\boldsymbol{\mu})^{\mathrm{T}}\left(\boldsymbol{V}^{-\frac{1}{2}}\right)^{\mathrm{T}}\boldsymbol{V}^{-\frac{1}{2}}(\boldsymbol{Y}-\boldsymbol{\mu}) \\
&= \left(\boldsymbol{V}^{-\frac{1}{2}}(\boldsymbol{Y}-\boldsymbol{\mu})\right)^{\mathrm{T}}\left(\boldsymbol{V}^{-\frac{1}{2}}(\boldsymbol{Y}-\boldsymbol{\mu})\right) \\
&= \boldsymbol{Z}^{\mathrm{T}}\boldsymbol{Z} \\
&= \sum_{i=1}^{p} Z_i^2 \\
&\sim \chi_p^2.
\end{aligned}
$$

b)

We perform variable transformation $\boldsymbol{Z}=\boldsymbol{V}^{-\frac{1}{2}}\boldsymbol{Y}$ which gives $\boldsymbol{Z}\sim N\left(\boldsymbol{V}^{-\frac{1}{2}}\boldsymbol{\mu},\boldsymbol{V}^{-\frac{1}{2}}\boldsymbol{V}\left(\boldsymbol{V}^{-\frac{1}{2}}\right)^{\mathrm{T}}\right)=N\left(\boldsymbol{V}^{-\frac{1}{2}}\boldsymbol{\mu},\boldsymbol{I}\right)$.

Thus, we have

$$
\begin{aligned}
\boldsymbol{Y}^{\mathrm{T}}\boldsymbol{V}^{-1}\boldsymbol{Y} &= \boldsymbol{Y}^{\mathrm{T}}\left(\boldsymbol{V}^{-\frac{1}{2}}\right)^{\mathrm{T}}\boldsymbol{V}^{-\frac{1}{2}}\boldsymbol{Y} \\
&= \left(\boldsymbol{V}^{-\frac{1}{2}}\boldsymbol{Y}\right)^{\mathrm{T}}\left(\boldsymbol{V}^{-\frac{1}{2}}\boldsymbol{Y}\right) \\
&= \boldsymbol{Z}^{\mathrm{T}}\boldsymbol{Z} \\
&= \sum_{i=1}^{p} Z_i^2 \\
&\sim \chi_{p,\lambda}^2
\end{aligned}
$$

where $\lambda=\sum_{i=1}^{p}\mathrm{E}[Z_i]^2=\left(\boldsymbol{V}^{-\frac{1}{2}}\boldsymbol{\mu}\right)^{\mathrm{T}}\left(\boldsymbol{V}^{-\frac{1}{2}}\boldsymbol{\mu}\right)=\boldsymbol{\mu}^{\mathrm{T}}\boldsymbol{V}^{-1}\boldsymbol{\mu}$.

Additional Exercise 3

(a) The model matrix X is defined as the matrix satisfying $X\beta=\sum_{i=0}^{I}\beta_i X_{ni}$, hence $X=[1_n, x-\overline{x}1_n]$ as claimed. The rank is 2, since either one among $x-\overline{x}1_n$ is different from the others or $x-\overline{x}1_n=0$.

(b) We know that $P_X=X\left(X^TX\right)^{-1}X^T$. Let $Z=x-\overline{x}1_n$. First observe that

$$
X^TX=\begin{bmatrix} 1_n^T1_n & 1_n^TZ \\ 1_n^TZ & Z^TZ \end{bmatrix},
$$

but $1_n^T1_n=n$ while $1_n^TZ=\sum_{i=1}^{n}x_i-\sum_{i=1}^{n}\overline{x}=0$ and $Z^TZ=\sum_{i=1}^{n}(x_i-\overline{x})^2=M$. The inverse of this matrix is

$$
\left(X^TX\right)^{-1}=\begin{bmatrix} n^{-1} & 0 \\ 0 & M^{-1} \end{bmatrix}.
$$

Since $\begin{bmatrix} C & D \end{bmatrix}\begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}\begin{bmatrix} C & D \end{bmatrix}^T=\begin{bmatrix} CAC^T \\ BAB^T \end{bmatrix}$, we obtain

$$X\left(X^TX\right)^{-1}X^T = n^{-1}1_n1_n^T + M^{-1}\left(x-\overline{x}\right)\left(x-\overline{x}\right)^T,$$

as claimed.

(c) The fitted values are defined as $\widehat{\mu} = P_X y$. We describe them in terms of the estimated parameters $\widehat{\beta_0}$ and $\widehat{\beta_1}$. Now $P_X y = n^{-1}1_n1_n^T y + M^{-1}\left(x-\overline{x}\right)\left(x-\overline{x}\right)^T y$, where $n^{-1}1_n1_n^T y = 1_n\overline{y} = 1_n\widehat{\beta_0}$ and

$$\begin{aligned} M^{-1}\left(x-\overline{x}\right)\left(x-\overline{x}\right)^T y &= \left(x-\overline{x}\right)\frac{\left(x-\overline{x}\right)^T y}{M}, \\ &= \left(x-\overline{x}\right)\widehat{\beta_1}. \end{aligned}$$

(d) Take the the composition $Y^TY = Y^TP_0Y + Y^T\left(P_X - P_0\right)Y + Y^T\left(I-P_X\right)Y$. Then

$$\begin{aligned} P_0Y &= 1_n\overline{y}, \\ \left(P_X-P_0\right)Y &= 1_n\left(\widehat{\mu}-\overline{y}\right), \\ \left(I-P_X\right)Y &= Y - 1_n\widehat{\mu}. \end{aligned}$$

Since the matrices are projections,

$$\begin{aligned} Y^TP_0Y &= \left(P_0Y\right)^T P_0Y = n\overline{y}^2, \\ Y^T\left(P_X-P_0\right)Y &= \sum_{i=1}^n\left(\widehat{\mu_i}-\overline{y}\right)^2, \\ Y^T\left(I-P_X\right)Y &= \sum_{i=1}^n\left(y_i-\widehat{\mu}\right)^2. \end{aligned}$$

(e) They are independently χ^2-distributed since the projections sum up to I. The rank of $P_X - P_0$ is 1, since the rank of P_X is 2 and the rank of P_0 is 1. Its non-centrality parameter is $\sigma^{-2}\mu^T\left(P_X-P_0\right)\mu$. Since $\mu = \beta_0 1_n + \beta_1\left(x-\overline{x}\right)$ and $P_X - P_0 = M^{-1}\left(x-\overline{x}1_n\right)^T\left(x-\overline{x}1_n\right)$ and

$$\begin{aligned} \left(x-\overline{x}1_n\right)^T\mu &= \left(x-\overline{x}1_n\right)^T\left(\beta_0 1_n + \beta_1\left(x-\overline{x}\right)\right) \\ &= \beta_0\left(x^T1_n - \overline{x}1_n^T1_n\right) + \beta_1 M \\ &= \beta_1 M, \end{aligned}$$

we obtain the non-centrality$M^{-1}\sigma^{-2}\widehat{\mu}^T\left(P_X-P_0\right)\widehat{\mu} = \dfrac{M}{\sigma^2}\beta_1^2$. As for $Y^T\left(I-P_X\right)Y\sigma^{-2}$, the rank is $n-2$ while $\mu^T\left(\mu - P_X\mu\right) = 0$, hence its non-centrality is 0.

(f) We should use

$$\frac{Y^T\left(P_X-P_0\right)Y/1}{Y^T\left(I-P_X\right)Y/\left(n-2\right)} \sim F_{1,n-2,M\left(\beta_1^2\sigma^{-2}\right)}.$$

Under H_0, $M\left(\beta_1^2\sigma^{-2}\right) = 0$.