EX5

roi hezkiyahu

24 3 2022

Q1

שאלה 1

בהינתן מודל ניתוח שונות מעורב עם אפקט מקרי אחד a_i ואפקט קבוע אחד β_j, עבור מערך מאוזן,

$$Y_{ijk} = \mu + a_i + \beta_j + \epsilon_{ijk} \quad \epsilon_{ijk} \sim N(0, \sigma^2_\epsilon) \quad (\epsilon_{ijk}\ iid) \quad a_i \sim N(0, \sigma^2_a)$$

עבור $i = 1, ..., I, j = 1, ..., J, k = 1, ..., n$,

א. חשבו כיצד מתפלג הממוצע הכללי ($\bar{Y}$).

ב. חשבו כיצד מתפלג הממוצע של כל קבוצה $i = 1, ..., I$ של האפקט המקרי ($\overline{Y}_{i..}$).

ג. חשבו כיצד מתפלג הממוצע של כל קבוצה $j = 1, ..., J$ של האפקט הקבוע ($\overline{Y}_{.j.}$).

a

$$\bar{Y} \sim N(\mu_{\bar{Y}}, \sigma^2_{\bar{Y}}) \text{ as a sum of normally distributed random variables}$$
$$E(\bar{Y}) = E[\frac{1}{N}\sum_i\sum_j\sum_k(\mu + \alpha_i + \beta_j + \varepsilon_{ijk})] = \mu + \frac{1}{J}\sum_j \beta_j$$
$$V(\bar{Y}) = \frac{1}{N^2}V[\sum_i\sum_j\sum_k(\mu + \alpha_i + \beta_j + \varepsilon_{ijk})] = \frac{1}{N^2}V[\sum_i\sum_j\sum_k(\alpha_i + \varepsilon_{ijk})] = \frac{1}{N^2}[V(\sum_i\sum_j\sum_k\alpha_i) + V(\sum_i\sum_j\sum_k\varepsilon_{ijk})] =$$
$$= \frac{1}{N^2}[n^2J^2\sum_i V(\alpha_i) + n^2J^2 2Cov(\sum_i\alpha_i, \sum_i\alpha_i) + N\sigma^2_\varepsilon] =$$
$$\frac{1}{N^2}[NnJ\sigma^2_\alpha + N\sigma^2_\varepsilon] = \frac{\sigma^2_\varepsilon + nJ\sigma^2_\alpha}{N}$$
$$\text{thus we get } \bar{Y} \sim N(\mu + \frac{1}{J}\sum_j\beta_j, \frac{\sigma^2_\varepsilon + nJ\sigma^2_\alpha}{N})$$

b

$$\bar{Y}_{i..} \sim N(\mu_{\bar{Y}_{i..}}, \sigma^2_{\bar{Y}_{i..}}) \text{ as a sum of normally distributed random variables}$$
$$E(\bar{Y}_{i..}) = E[\frac{1}{nJ}\sum_j\sum_k(\mu + \alpha_i + \beta_j + \varepsilon_{ijk})] = \mu + \frac{1}{J}\sum_j\beta_j$$
$$V(\bar{Y}_{i..}) = V[\frac{1}{nJ}\sum_j\sum_k(\mu + \alpha_i + \beta_j + \varepsilon_{ijk})] = V[\frac{1}{n^2J^2}\sum_j\sum_k(\alpha_i + \varepsilon_{ijk})] = \frac{\sigma^2_\alpha + \sigma^2_\varepsilon}{nJ}$$
$$\text{thus we get } \bar{Y}_{i..} \sim N(\mu + \frac{1}{J}\sum_j\beta_j, \frac{\sigma^2_\alpha + \sigma^2_\varepsilon}{nJ})$$

c

$$\bar{Y}_{.j.} \sim N(\mu_{\bar{Y}_{.j.}}, \sigma^2_{\bar{Y}_{.j.}}) \text{ as a sum of normally distributed random variables}$$

$$E(\bar{Y}_{.j.}) = E[\frac{1}{nI}\sum_i\sum_k(\mu+\alpha_i+\beta_j+\varepsilon_{ijk})] = \mu+\beta_j$$

$$V(\bar{Y}_{.j.}) = V[\frac{1}{nI}\sum_i\sum_k(\mu+\alpha_i+\beta_j+\varepsilon_{ijk})] = \frac{1}{n^2I^2}V[\sum_i\sum_k(\alpha_i+\varepsilon_{ijk})] = \frac{\sigma^2_\varepsilon}{nI} + \frac{V(\sum_i\alpha_i)}{I^2} = \frac{\sigma^2_\varepsilon}{nI} + \frac{\sigma^2_\alpha}{I}$$

$$\text{thus we get } \bar{Y}_{.j.} \sim N(\mu+\beta_j, \frac{\sigma^2_\varepsilon}{nI} + \frac{\sigma^2_\alpha}{I})$$

Q2

שאלה 2

בהינתן מודל רגרסיה לינארית עם חותך מקרי,

$$Y_{ij} = \beta_0 + b_{0i} + x_{ij}^T\beta + \epsilon_{ij} \quad \epsilon_{ij} \sim N(0,\sigma^2_\epsilon) \quad (\epsilon_{ij}\ iid) \quad b_{0i} \sim N(0,\sigma^2_b)$$

עבור $i=1,...,I, j=1,...,n_i$,

הוכיחו כי ה ICC המוגדר להיות $\rho = corr(Y_{ij}, Y_{ij'})$ שווה ל $\frac{\sigma^2_b}{\sigma^2_b+\sigma^2_\epsilon}$ עבור $j \neq j'$.

$$V(Y_{ij}) = \sigma^2_\varepsilon + \sigma^2_b$$

$$corr(Y_{ij}, Y_{ij'}) = \frac{Cov(Y_{ij}, Y_{ij'})}{\sqrt{\sigma^2_\varepsilon+\sigma^2_b}\sqrt{\sigma^2_\varepsilon+\sigma^2_b}}$$

$$Cov(Y_{ij}, Y_{ij'}) = Cov(\beta_0+b_{0i}+x_{ij}^T\beta+\varepsilon_{ij}, \beta_0+b_{0i}+x_{ij'}^T\beta+\varepsilon_{ij'}) = Cov(b_{0i}+\varepsilon_{ij}, b_{0i}+\varepsilon_{ij'}) = Cov(b_{0i}+\varepsilon_{ij}, b_{0i}) + Cov(b_{0i}+\varepsilon_{ij}, \varepsilon_{ij'}) =$$

$$= Cov(b_{0i}+\varepsilon_{ij}, b_{0i}) = V(b_{0i}) = \sigma^2_b \Rightarrow corr(Y_{ij}, Y_{ij'}) = \frac{\sigma^2_b}{\sigma^2_\varepsilon+\sigma^2_b}$$

Q3

שאלה 3

בשיעור הזכרתם את ההשפעה של מולטיקולינאריות על אומדי מקדמי הרגרסיה. בשאלה זו נבצע סימולציות על מנת להדגים זאת.

עליכם לבצע את הסימולציה המתוארת מטה ארבע פעמים, בכל פעם עם פרמטרים שונים.

א. בצעו את הסימולציה עם $n = 100$ ועם $\sigma^2 = 1$.

ב. בצעו את הסימולציה עם $n = 100$ ועם $\sigma^2 = 3^2$.

ג. בצעו את הסימולציה עם $n = 1000$ ועם $\sigma^2 = 1$.

ד. בצעו את הסימולציה עם $n = 1000$ ועם $\sigma^2 = 3^2$.

ה. השוו בין ארבע הסימולציות:

I. האם הממוצעים, סטיות התקן, הקורלציה והגרף השתנו כאשר גודל המדגם גדל? הסבירו.
II. האם הממוצעים, סטיות התקן, הקורלציה והגרף השתנו כאשר סטיית התקן גדלה? הסבירו.

<u>הסימולציה:</u>
בצעו תהליך של יצירת מדגם 1,000 פעמים. בכל איטרציה צרו מדגם בגודל n עם שני משתנים מסבירים (x_1, x_2) ומשתנה תלוי (y). התאימו מודל רגרסיה לינארית ושימרו את המקדמים המתקבלים. כך תעשו זאת:

- דגמו וקטור x_1 של משתנים נורמליים סטנדרטיים $x_1 \sim N(0,1)$
- צרו וקטור x_2 באופן הבא: $\epsilon \sim N(0,\sigma^2) \quad x_2 = x_1 + \epsilon$
כלומר לכל תצפית, x_2 שווה לערך המקביל של x_1 בתוספת רעש ייחודי לתצפית זו, שהוא דגימה של משתנה נורמלי.
- צרו וקטור y באופן הבא: $\epsilon_2 \sim N(0,1) \quad y = x_1 + x_2 + \epsilon_2$
כלומר לכל תצפית, y שווה לערך המקביל של x_1 ועוד x_2 בתוספת רעש ייחודי לתצפית זו, שהוא דגימה של משתנה נורמלי.
- התאימו מודל רגרסיה לינארית פשוטה: $y \sim x_1 + x_2$
- חלצו את $\hat\beta_1$ המקדם של x_1, ואת $\hat\beta_2$ המקדם של x_2 ממודל הרגרסיה ושימרו אותם

מכל אחד מ1000 האיטרציות של הסימולציה קיבלתם שני מקדמים, $\hat\beta_1$ ו $\hat\beta_2$. נרצה לבחון את המקדמים ואת הקשר בניהם.

- דווחו את הממוצע ואת סטיית התקן של כל אחד מהמקדמים $\hat\beta_1$ ו $\hat\beta_2$.
- דווחו את הקורלציה בין המקדמים $\hat\beta_1$ ו $\hat\beta_2$.
- צרו גרף נקודות של $\hat\beta_1$ לעומת $\hat\beta_2$
- מה ניתן ללמוד מכך? הסבירו.

```
#imports
library(tidyverse)
```

```
## -- Attaching packages ------------------------------------------------------------------------------------------------------
------- tidyverse 1.3.0 --
```

```
## <U+221A> ggplot2 3.3.2     <U+221A> purrr   0.3.4
## <U+221A> tibble  3.0.3     <U+221A> dplyr   1.0.2
## <U+221A> tidyr   1.1.2     <U+221A> stringr 1.4.0
## <U+221A> readr   1.3.1     <U+221A> forcats 0.5.0
```

```
## -- Conflicts ---------------------------------------------------------------------------------------------------------------
- tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

```
library(glue)
```

```
##
## Attaching package: 'glue'
```

```
## The following object is masked from 'package:dplyr':
##
##     collapse
```

```
library(patchwork)
library(lme4)
```

```
## Loading required package: Matrix
```

```
##
## Attaching package: 'Matrix'
```

```
## The following objects are masked from 'package:tidyr':
##
##     expand, pack, unpack
```

```
library(lmerTest)
```

```
##
## Attaching package: 'lmerTest'
```

```
## The following object is masked from 'package:lme4':
##
##     lmer
```

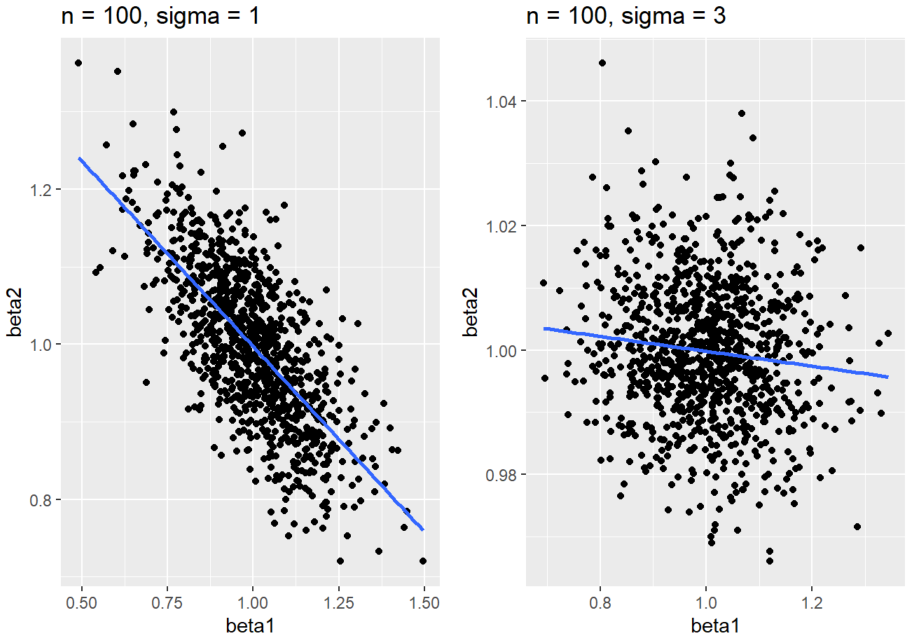
```
## The following object is masked from 'package:stats':
##
##     step
```

a-d

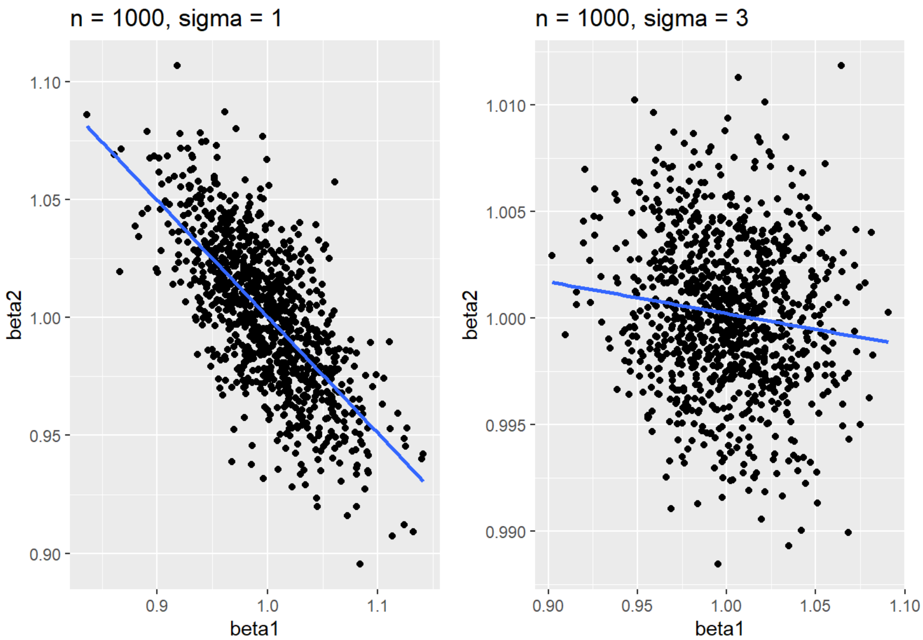
```
N <-  1000
simulation <- function(n,sigma){
  beta1 <- c()
  beta2 <- c()
  for (i in 1:N){
    x1 <- rnorm(n)
    x2 <- x1 + rnorm(n,0,sigma^2)
    y <- x1+x2+rnorm(n)
    lm_model <- lm(y~x1+x2)
    beta1[i] <- lm_model$coefficients[2]
    beta2[i] <- lm_model$coefficients[3]
  }
  tbl <- tibble(beta1 = beta1, beta2 = beta2)
  summary_tbl <- tbl%>%
    summarise("mean beta1" = mean(beta1),
              "mean beta2" = mean(beta2),
              "sd beta1" = sd(beta1),
              "sd beta2" = sd(beta2),
              "cor beta1,beta2" = cor(beta1,beta2))
  plt <- tbl %>%
    ggplot(aes(y=beta2,x=beta1)) +
    geom_point()+
    geom_smooth(method='lm', formula= y~x,se=F)+
    ggtitle(glue("n = {n}, sigma = {sigma}"))
  return(list(summary_tbl,plt))
}
s1 <- simulation(100,1)
s2 <- simulation(100,3)
s3 <- simulation(1000,1)
s4 <- simulation(1000,3)
bind_rows(s1[[1]],s2[[1]],s3[[1]],s4[[1]]) %>%
  mutate(n = c(100,100,1000,1000), sd = c(1,3,1,3))
```

```
## # A tibble: 4 x 7
##   `mean beta1` `mean beta2` `sd beta1` `sd beta2` `cor beta1,beta2`     n    sd
##          <dbl>        <dbl>      <dbl>      <dbl>             <dbl> <dbl> <dbl>
## 1        0.999        0.998     0.146     0.0998            -0.700   100     1
## 2        1.00         1.00      0.107     0.0112            -0.115   100     3
## 3        0.999        1.00      0.0447    0.0316            -0.699  1000     1
## 4        1.00         1.00      0.0308    0.00354           -0.129  1000     3
```

```
(s1[[2]] + s2[[2]])
```




```
(s3[[2]] + s4[[2]])
```



e

for changing n

we can see: the means,correlation and plots didn't change much but the beta_sd decreased dramatically

this is due to the fact that the correlation depends on the sd, the mean becomes closer to the expected value, and the plots don't change because the distribution is the same. beta_sd changes because as n increases beta_sd decreases

for changing the sd

we can see: the means didn't change much but the correlation and beta_sd decreased dramatically and the plots changed

the mean does not change because the sd does not effect the expected value

beta_sd changes because we can better estimate it without correlation between x1,x2

correlation changes because beta_2 has much more weight in the formula as it increases the correlation decreases

plots changes because the correlation changes

Q4

שאלה 4

בשאלה זו נשתמש בנתונים בשם feedback_bi הזמינים כקובץ csv במודל.

אלו נתונים מאותו מחקר בו השתמשתם בתרגיל בית 4. המשתנה התלוי performance מכיל מדד בשם emg_bi. המדד הזה מכיל תוצאות מדידה של פוטנציאל חשמלי הנמדד על ידי מכשיר אלקטרומיוגרף (EMG) בשריר הזרוע הדו ראשי (biceps brachii muscles). הנבדקים בקבוצת הביקורת לא קיבלו עידוד לפני מדידת הפוטנציאל החשמלי שלהם, ואילו אלו שבקבוצות הטיפול קיבלו חיזוק חיובי ("You are looking great.") או שלילי ("You are not trying.") .

העמודות הנדרשות לנו בשאלה זו הן -

- id - מספר מזהה ייחודי למשתתף
- performance - (נומרי) emg bi הציון המתקבל במדד
- feedback- (פידבק חיובי / פידבק שלילי / ביקורת) האם קיבל פידבק
- gender - (גבר/אישה) מגדר

א. נרצה לבחון את האפקט של המגדר (gender) על התיפקוד (performance), תוך כדי שניקח בחשבון את השונות בין המשתתפים השונים (id) .

I. איזה סוג ניתוח מתאים ביותר למקרה הזה? הסבירו.

II. דווחו את התוצאה שקיבלתם. מה המסקנות עבור $\alpha = 0.05$?

III. חשבו אומדים לכל אחד מהפרמטרים של המודל.

IV. בנו רווחי סמך ברמת סמך של 95% עבור כל אחד מהפרמטרים של המודל שהתאמתם בR.

ב. נרצה לבחון את האפקט של הפידבק (feedback) על התיפקוד (performance), תוך כדי שניקח בחשבון את השונות בין המשתתפים השונים (id) .

I. איזה סוג ניתוח מתאים ביותר למקרה הזה? הסבירו.

II. דווחו את התוצאה שקיבלתם. מה המסקנות עבור $\alpha = 0.05$?

III. חשבו אומדים לכל אחד מהפרמטרים של המודל.

IV. בנו רווחי סמך ברמת סמך של 95% עבור כל אחד מהפרמטרים של המודל שהתאמתם בR.

ג. נרצה לבחון את האפקט של המגדר (gender) והאפקט של הפידבק (feedback) על התיפקוד (performance), תוך כדי שניקח בחשבון את השונות בין המשתתפים השונים (id) .

I. איזה סוג ניתוח מתאים ביותר למקרה הזה? הסבירו.

II. דווחו את התוצאה שקיבלתם. מה המסקנות עבור $\alpha = 0.05$?

III. חשבו אומדים לכל אחד מהפרמטרים של המודל.

IV. בנו רווחי סמך ברמת סמך של 95% עבור כל אחד מהפרמטרים של המודל שהתאמתם בR.

ד. נרצה לבחון את האפקט של המגדר (gender) והאפקט של הפידבק (feedback) והאינטראקציה בניהם על התיפקוד (performance), תוך כדי שניקח בחשבון את השונות בין המשתתפים השונים (id) .

I. איזה סוג ניתוח מתאים ביותר למקרה הזה? הסבירו.

II. דווחו את התוצאה שקיבלתם. מה המסקנות עבור $\alpha = 0.05$?

III. חשבו אומדים לכל אחד מהפרמטרים של המודל.

IV. בנו רווחי סמך ברמת סמך של 95% עבור כל אחד מהפרמטרים של המודל שהתאמתם בR.

ה. השוו בין ארבעת הניתוחים שביצעתם. הסבירו איך הפרמטרים, רווחי הסמך ואחוזי השונות המוסברת השתנו. מה ניתן ללמוד מזה? איזה מודל תעדיפו להתאים?

a

the proper analysis would be a mixed model because each participant has it's own variance and we expect a fixed effect for gender

```
feed_back <- read.csv("feedback_df_bi.csv") %>%
  select(id,performance,feedback,gender) %>%
  mutate(across(c(gender,feedback),as.factor))
model1 <- lmerTest::lmer(performance ~ gender + (1|id) , data = feed_back)
summary(model1)
```

```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: performance ~ gender + (1 | id)
##    Data: feed_back
##
## REML criterion at convergence: 6170.7
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -4.4407 -0.4590 -0.0056  0.3728  4.6852
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  id       (Intercept)  84.59    9.197
##  Residual             131.28   11.458
## Number of obs: 792, groups:  id, 22
##
## Fixed effects:
##             Estimate Std. Error     df t value Pr(>|t|)
## (Intercept)   92.121      2.832 20.000  32.526   <2e-16 ***
## gender2       -1.064      4.005 20.000  -0.266    0.793
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##         (Intr)
## gender2 -0.707
```

```
mu <- 92.121
sigma_a <- 84.59
sigma_e <- 131.28
glue("the estimators are:
     mu = {mu}
     sigma alpha = {sigma_a}
     sigma epsilon = {sigma_e}")
```

```
## the estimators are:
## mu = 92.121
## sigma alpha = 84.59
## sigma epsilon = 131.28
```

```
confint(lme4::lmer(performance ~ gender + (1|id) , data = feed_back))
```

```
## Computing profile confidence intervals ...
```

```
##                 2.5 %    97.5 %
## .sig01       6.571680 12.281426
## .sigma      10.908391 12.054634
## (Intercept) 86.588734 97.653534
## gender2     -8.887814  6.760177
```

we do not reject the null with pvlaue of 0.793 therefor we conclude that gender has no contribution (which makes sense beacuse results are normalized based on participant)

b

```
model2 <- lmerTest::lmer(performance ~ feedback + (1|id) , data = feed_back)
summary(model2)
```

```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: performance ~ feedback + (1 | id)
##    Data: feed_back
##
## REML criterion at convergence: 6119.1
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -4.4058 -0.4546 -0.0217  0.4058  5.0225
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  id       (Intercept)  80.92    8.995
##  Residual             122.87   11.085
## Number of obs: 792, groups:  id, 22
##
## Fixed effects:
##                      Estimate Std. Error       df t value Pr(>|t|)
## (Intercept)           95.7091     2.0356  24.5328  47.019  < 2e-16 ***
## feedbackno feedback  -6.2064     0.9648 768.0000  -6.433  2.2e-10 ***
## feedbackpositive     -6.1532     0.9648 768.0000  -6.378  3.1e-10 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr) fdbckf
## fdbcknfdbck -0.237
## feedbckpstv -0.237  0.500
```

```
mu <- 95.7
sigma_a <- 80.92
sigma_e <- 122.87
glue("the estimators are:
     mu = {mu}
     sigma alpha = {sigma_a}
     sigma epsilon = {sigma_e}")
```

```
## the estimators are:
## mu = 95.7
## sigma alpha = 80.92
## sigma epsilon = 122.87
```

```
confint(lme4::lmer(performance ~ feedback + (1|id) , data = feed_back))
```

```
## Computing profile confidence intervals ...
```

```
##                          2.5 %    97.5 %
## .sig01               6.602662 12.313431
## .sigma              10.539466 11.646941
## (Intercept)         91.652670 99.765507
## feedbackno feedback -8.097255 -4.315556
## feedbackpositive    -8.044036 -4.262336
```

we can see that feedback has a pv <0.05 therefore we can reject the null at 95% confidence and conclude that feedback has an effect on performance

C

```
model3 <- lmerTest::lmer(performance ~ feedback+ gender + (1|id) , data = feed_back)
summary(model3)
```

```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: performance ~ feedback + gender + (1 | id)
##    Data: feed_back
##
## REML criterion at convergence: 6114.5
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -4.4078 -0.4562 -0.0243  0.4025  5.0205
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  id       (Intercept)  84.82    9.21
##  Residual             122.87   11.08
## Number of obs: 792, groups:  id, 22
##
## Fixed effects:
##                      Estimate Std. Error       df t value Pr(>|t|)
## (Intercept)           96.2410     2.8865  21.5763  33.342  < 2e-16 ***
## feedbackno feedback   -6.2064     0.9648 768.0000  -6.433  2.2e-10 ***
## feedbackpositive      -6.1532     0.9648 768.0000  -6.378  3.1e-10 ***
## gender2               -1.0638     4.0054  20.0000  -0.266    0.793
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr) fdbckf fdbckp
## fdbcknfdbck -0.167
## feedbckpstv -0.167  0.500
## gender2     -0.694  0.000  0.000
```

```
mu <- 96.24
sigma_a <- 84.82
sigma_e <- 122.87
glue("the estimators are:
     mu = {mu}
     sigma alpha = {sigma_a}
     sigma epsilon = {sigma_e}")
```

```
## the estimators are:
## mu = 96.24
## sigma alpha = 84.82
## sigma epsilon = 122.87
```

```
confint(lme4::lmer(performance ~ feedback+ gender + (1|id) , data = feed_back))
```

```
## Computing profile confidence intervals ...
```

```
##                          2.5 %     97.5 %
## .sig01                6.590137  12.291286
## .sigma               10.539466  11.646941
## (Intercept)          90.610615 101.871381
## feedbackno feedback  -8.097255  -4.315556
## feedbackpositive     -8.044036  -4.262336
## gender2              -8.887795   6.760158
```

the same results as before…

d

```
model4 <- lmerTest::lmer(performance ~ feedback+ gender+ gender*feedback + (1|id) , data = feed_back)
summary(model4)
```

```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: performance ~ feedback + gender + gender * feedback + (1 | id)
##    Data: feed_back
##
## REML criterion at convergence: 6086
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -4.5657 -0.5251  0.0072  0.4470  4.9436
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  id       (Intercept)  84.91    9.215
##  Residual             119.63   10.938
## Number of obs: 792, groups:  id, 22
##
## Fixed effects:
##                              Estimate Std. Error      df t value Pr(>|t|)
## (Intercept)                    98.849      2.937  23.123  33.657  < 2e-16 ***
## feedbackno feedback            -9.868      1.346 766.000  -7.330 5.86e-13 ***
## feedbackpositive              -10.315      1.346 766.000  -7.661 5.56e-14 ***
## gender2                        -6.280      4.153  23.123  -1.512  0.14411
## feedbackno feedback:gender2     7.324      1.904 766.000   3.847  0.00013 ***
## feedbackpositive:gender2        8.323      1.904 766.000   4.371 1.40e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr) fdbckf fdbckp gendr2 ffdb:2
## fdbcknfdbck -0.229
## feedbckpstv -0.229  0.500
## gender2     -0.707  0.162  0.162
## ffdbck:gnd2  0.162 -0.707 -0.354 -0.229
## fdbckpstv:2  0.162 -0.354 -0.707 -0.229  0.500
```

```
mu <- 98.849
sigma_a <- 84.91
sigma_e <- 119.63
glue("the estimators are:
     mu = {mu}
     sigma alpha = {sigma_a}
     sigma epsilon = {sigma_e}")
```

```
## the estimators are:
## mu = 98.849
## sigma alpha = 84.91
## sigma epsilon = 119.63
```

```
confint(lme4::lmer(performance ~ feedback+ gender + gender*feedback+ (1|id) , data = feed_back))
```

```
## Computing profile confidence intervals ...
```

```
##                                  2.5 %     97.5 %
## .sig01                        6.597596  12.295277
## .sigma                       10.386109  11.477468
## (Intercept)                  93.127219 104.570493
## feedbackno feedback         -12.503534  -7.233223
## feedbackpositive            -12.949944  -7.679633
## gender2                     -14.371152   1.812081
## feedbackno feedback:gender2   3.597273  11.050618
## feedbackpositive:gender2      4.596533  12.049878
```

same results for feedback+ gender but we can see that the interaction also effects the performance

e

we can see that the parameters have not changed much

also the ICC stayed the same

the ci's for the coefficients corresponding to gender,feedback and their interaction has changed, this happens because when we add variables to the model we change the spanned space there for the projection is different leading to different coefficients

i would prefer to use the last model, this is due to the fact that the interaction is significant and when we have an interaction in the model we should keep all interacting variables (models A)