

# Two Factor Analysis with Missing Cells

In this example, we consider 3 levels of Drug and 4 levels of Dose for a total of 12 possible treatment combinations. However two treatment combinations are not represented. We refer to this situation as “missing cells”.

```
library(car)
library(emmeans)
DrugTrial2 <- read.csv("C:/hess/STAT512/RNotes/ExpDesign3/ED3_MissingCells2Factor.csv")
str(DrugTrial2)
```

```
## 'data.frame': 56 obs. of 3 variables:
## $ Drug: int 1 1 1 1 1 1 1 1 1 1 ...
## $ Dose: int 1 1 1 2 2 2 2 2 2 4 ...
## $ Y : int 45 44 43 42 44 45 43 44 45 22 ...
```

*#Important: Need to define Drug and Dose as factors!!!!*

```
DrugTrial2$Drug<-as.factor(DrugTrial2$Drug)
DrugTrial2$Dose<-as.factor(DrugTrial2$Dose)
str(DrugTrial2)
```

```
## 'data.frame': 56 obs. of 3 variables:
## $ Drug: Factor w/ 3 levels "1","2","3": 1 1 1 1 1 1 1 1 1 1 ...
## $ Dose: Factor w/ 4 levels "1","2","3","4": 1 1 1 2 2 2 2 2 2 4 ...
## $ Y : int 45 44 43 42 44 45 43 44 45 22 ...
```

*#Use table to check the number of observations for each Drug/Dose combination.*

```
table(DrugTrial2$Drug, DrugTrial2$Dose)
```

```
##
##      1 2 3 4
##    1 3 6 0 7
##    2 6 8 4 0
##    3 3 7 7 5
```

## Two-way analysis with missing cells

Due to the missing cells, Anova() gives an error and does not return any results. We can still look at the emmeans, but note the NA values.

```
options(contrasts=c("contr.sum", "contr.poly"))
Model1 <- lm(Y ~ Drug*Dose, data = DrugTrial2)
#Rmarkdown won't let me run this code because it generates an error.
#Error in Anova.III.lm(mod, error, singular.ok = singular.ok, ...) :
#there are aliased coefficients in the model
#Anova(Model1, type = 3)
summary(Model1)
```

```
##
## Call:
## lm(formula = Y ~ Drug * Dose, data = DrugTrial2)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
```

```
## -2.3333 -0.8750 0.0000 0.7143 1.7143
##
## Coefficients: (2 not defined because of singularities)
##           Estimate Std. Error t value Pr(>|t|)
## (Intercept) 40.0847    0.2475 161.939 < 2e-16 ***
## Drug1       -3.1307    0.3660  -8.553 4.59e-11 ***
## Drug2        5.6725    0.4849  11.699 2.20e-15 ***
## Dose1       -3.9735    0.3269 -12.154 5.79e-16 ***
## Dose2       -2.0450    0.2987  -6.846 1.55e-08 ***
## Dose3       -0.2571    0.3886  -0.662 0.511
## Drug1:Dose1 11.0196    0.4499  24.494 < 2e-16 ***
## Drug2:Dose1 -22.4503    0.6462 -34.743 < 2e-16 ***
## Drug1:Dose2  8.9243    0.4183  21.334 < 2e-16 ***
## Drug2:Dose2 -5.7122    0.5863  -9.743 9.24e-13 ***
## Drug1:Dose3      NA         NA      NA      NA
## Drug2:Dose3      NA         NA      NA      NA
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.11 on 46 degrees of freedom
## Multiple R-squared:  0.9862, Adjusted R-squared:  0.9835
## F-statistic: 365.1 on 9 and 46 DF,  p-value: < 2.2e-16
```

```
emmeans(Model1, ~Dose*Drug)
```

```
## Dose Drug emmean    SE df lower.CL upper.CL
## 1 1 44.0 0.641 46 42.7 45.3
## 2 1 43.8 0.453 46 42.9 44.7
## 3 1 nonEst NA NA NA NA
## 4 1 23.3 0.419 46 22.4 24.1
## 1 2 19.3 0.453 46 18.4 20.2
## 2 2 38.0 0.392 46 37.2 38.8
## 3 2 45.5 0.555 46 44.4 46.6
## 4 2 nonEst NA NA NA NA
## 1 3 45.0 0.641 46 43.7 46.3
## 2 3 32.3 0.419 46 31.4 33.1
## 3 3 37.3 0.419 46 36.4 38.1
## 4 3 35.6 0.496 46 34.6 36.6
##
## Confidence level used: 0.95
```

```
#For Illustration
```

```
emmeans(Model1, ~ Dose)
```

```
## NOTE: Results may be misleading due to involvement in interactions
```

```
## Dose emmean    SE df lower.CL upper.CL
## 1 36.1 0.338 46 35.4 36.8
## 2 38.0 0.244 46 37.5 38.5
## 3 nonEst NA NA NA NA
## 4 nonEst NA NA NA NA
##
```

```
## Results are averaged over the levels of: Drug
```

```
## Confidence level used: 0.95
```

## One-way Analysis

We start by creating a new variable called “Trt” which just combines the Drug, Dose information. Then run a one way ANOVA analysis with Trt (10 levels). Note that the emmeans are the same as the previous analysis.

```
DrugTrial2$Trt <- paste("Drug", DrugTrial2$Drug, "Dose", DrugTrial2$Dose)
#Use table to check the number of observations for each Trt
table(DrugTrial2$Trt)
```

```
##
## Drug 1 Dose 1 Drug 1 Dose 2 Drug 1 Dose 4 Drug 2 Dose 1 Drug 2 Dose 2
##           3           6           7           6           8
## Drug 2 Dose 3 Drug 3 Dose 1 Drug 3 Dose 2 Drug 3 Dose 3 Drug 3 Dose 4
##           4           3           7           7           5
```

```
Model2 <- lm(Y ~ Trt, data = DrugTrial2)
anova(Model2)
```

```
## Analysis of Variance Table
##
## Response: Y
##           Df Sum Sq Mean Sq F value    Pr(>F)
## Trt         9 4047.3  449.71   365.15 < 2.2e-16 ***
## Residuals 46   56.7    1.23
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
emmeans(Model2, ~ Trt)
```

```
## Trt           emmean    SE df lower.CL upper.CL
## Drug 1 Dose 1    44.0 0.641 46     42.7     45.3
## Drug 1 Dose 2    43.8 0.453 46     42.9     44.7
## Drug 1 Dose 4    23.3 0.419 46     22.4     24.1
## Drug 2 Dose 1    19.3 0.453 46     18.4     20.2
## Drug 2 Dose 2    38.0 0.392 46     37.2     38.8
## Drug 2 Dose 3    45.5 0.555 46     44.4     46.6
## Drug 3 Dose 1    45.0 0.641 46     43.7     46.3
## Drug 3 Dose 2    32.3 0.419 46     31.4     33.1
## Drug 3 Dose 3    37.3 0.419 46     36.4     38.1
## Drug 3 Dose 4    35.6 0.496 46     34.6     36.6
##
## Confidence level used: 0.95
```