Review of linear models

$$y = X\beta + e$$

y = dependent variable (N x 1), ~ $\mathbf{N}(\boldsymbol{X\beta}, \sigma^2 I_N)$
e = residual error
X = fixed effect design matrix (N x p)
$\boldsymbol{\beta}$ = fixed effects parameters (p x 1)
e ~$\mathbf{N}(0, \sigma^2 I_N)$

Linear Regression

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + e$$

β_0 = intercept
β_1= slope in the x_1 direction holding x_2 fixed
β_2= slope in the x_2 direction holding x_1 fixed

$$y = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \\ y_6 \\ y_7 \\ y_8 \\ y_9 \end{bmatrix} \quad x = \begin{bmatrix} 1 & x_{11} & x_{21} \\ 1 & x_{21} & x_{22} \\ 1 & x_{31} & x_{32} \\ 1 & x_{41} & x_{42} \\ 1 & x_{51} & x_{52} \\ 1 & x_{61} & x_{62} \\ 1 & x_{71} & x_{72} \\ 1 & x_{81} & x_{82} \\ 1 & x_{91} & x_{92} \end{bmatrix} \quad \beta = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{bmatrix} \quad e = \begin{bmatrix} e_1 \\ e_2 \\ e_3 \\ e_4 \\ e_5 \\ e_6 \\ e_7 \\ e_8 \\ e_9 \end{bmatrix}$$

or

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + e_i$$

Prediction equation

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 x_2$$

Independent variable type: direct
no CLASS statement

```
SAS code:

PROC MIXED;
MODEL Y=X1 X2;
RUN;
```

Analysis of Variance (ANOVA)

Completely Randomized Design (CRD) - treatments randomly assigned to experimental units.

One-way classification model

$$y = X\beta + e$$

3 levels of Treatment, 3 experimental units each treatment

$$y = \begin{bmatrix} y_{11} \\ y_{12} \\ y_{13} \\ y_{21} \\ y_{22} \\ y_{23} \\ y_{31} \\ y_{32} \\ y_{33} \end{bmatrix} \qquad X = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \end{bmatrix} \qquad \beta = \begin{bmatrix} \mu \\ \tau_1 \\ \tau_2 \\ \tau_3 \end{bmatrix}$$

or

$y_{ij} = \mu + \tau_i + e_{ij}$

SAS Code

```
PROC MIXED;
     CLASS B;
     MODEL Y=B;
RUN;
```

Randomized block design –treatments randomly assigned to experimental units within homogeneous blocks (naturally occuring differences not related to treatments). Randomized **complete** block design when all treatments are assigned to all blocks. Blocks can be treated either as fixed or random.

3 blocks and 2 levels of Treatment B, 2 e.u. per treatment within block. Block as fixed effect.

$$y = \begin{bmatrix} y_{111} \\ y_{112} \\ y_{121} \\ y_{122} \\ y_{211} \\ y_{212} \\ y_{221} \\ y_{222} \\ y_{311} \\ y_{312} \\ y_{321} \\ y_{322} \end{bmatrix} \quad X = \begin{bmatrix} 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 & 0 & 1 \end{bmatrix} \quad B = \begin{bmatrix} \mu \\ blk_1 \\ blk_2 \\ blk_3 \\ \tau_1 \\ \tau_2 \end{bmatrix}$$

or

$y_{ijk} = \mu + blk_i + \tau_j + e_{ijk}$

SAS Code

```
PROC MIXED;
     CLASS BLK B;
     MODEL Y=BLK B;
RUN;
```

Row-column designs-two or more identifiable sources of variation that are independent.

Latin Square design –Special case of row-column design where number of treatments=number of rows=number of columns.

Run	Position			
	1	2	3	4
1	A	C	D	B
2	B	D	C	A
3	D	B	A	C
4	C	A	B	D

SAS Code

```
PROC MIXED;
     CLASS RUN POS B;
     MODEL Y=RUN POS B;
RUN;
```

Factorial Treatment Designs – two or more factors. The response is modeled as a function of the treatment.

- look at every combination of levels.

- Goal is to describe “as simply as possible” how the factors affect the response variable.

2 levels of Treatment A, 2 levels of Treatment B, 3 experimental units each treatment combination.

$$y = \begin{bmatrix} y_{111} \\ y_{112} \\ y_{113} \\ y_{121} \\ y_{122} \\ y_{123} \\ y_{211} \\ y_{212} \\ y_{213} \\ y_{221} \\ y_{222} \\ y_{223} \end{bmatrix} \quad X = \begin{bmatrix} 1 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 \end{bmatrix} \quad B = \begin{bmatrix} \mu \\ \alpha_1 \\ \alpha_2 \\ \beta_1 \\ \beta_2 \\ \alpha\beta_{11} \\ \alpha\beta_{12} \\ \alpha\beta_{21} \\ \alpha\beta_{22} \end{bmatrix}$$

or

$y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + e_{ijk}$

SAS Code

```
PROC MIXED;
      CLASS A B;
      MODEL Y=A B A*B;
RUN;
```