$\boxed{\text{Chapter 16}}$

Balanced Completely Random Design: $Y_{ij} = \beta_0 + \beta_1 X_{ij1} + \beta_2 X_{ij2} + \dots + \beta_{r-1} X_{ij,r-1} + \epsilon_{ij}$

- Instead of the predictors having their own distribution they are qualitative
- Their predicted coefficients are the mean responses.

Cell Means Model — denote r the number of levels of the factor with index i and n_i = # of cases for the i^{th} factor level w/ n_T = total # of cases ($n_T = \sum_i n_i$)

$Y_{ij} = \mu_i + \epsilon_{ij}$ w/
- $E(Y_{ij}) = \mu_i$
- $\sigma^2(Y_{ij}) = \sigma^2(\mu_i + \epsilon_{ij}) = \sigma^2(\epsilon_{ij}) = \sigma^2$
- independent errors $\rightarrow Y_{ij} \overset{ind}{\sim} N(\mu_i, \sigma^2)$

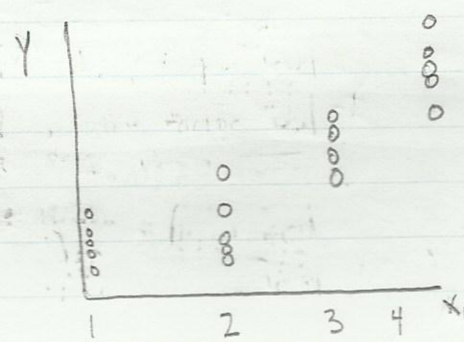
Can be expressed in Matrix Terms $\underline{Y} = \underline{X}\,\underline{\beta} + \underline{\epsilon}$

$$\underline{Y} = \begin{bmatrix} Y_{11} \\ Y_{12} \\ Y_{21} \\ Y_{22} \\ Y_{31} \\ Y_{32} \end{bmatrix}, \quad \underline{X} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix}, \quad \underline{\beta} = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \end{bmatrix}, \quad \underline{\epsilon} = \begin{bmatrix} \epsilon_{11} \\ \epsilon_{12} \\ \epsilon_{21} \\ \epsilon_{22} \\ \epsilon_{31} \\ \epsilon_{32} \end{bmatrix}$$

$$E(\underline{Y}) = [\mu_1, \mu_1, \mu_2, \mu_2, \mu_3, \mu_3]' \qquad \sigma^2(\underline{\epsilon}) = \sigma^2 \underline{I}$$

$\boxed{\text{Sas}}$

```
proc sgscatter;
    plot Y * X1;
run;
```



SAS

```
proc glm;
    class x1;
    model y=x1;
    lsmeans x1;
run;
proc ANOVA;
    class x1;
    Model y=x1;
    means x1;
run;
```

Levels of x	x LSMEAN
1	μ_1
2	μ_2
3	μ_3
⋮	⋮

Notation

$\boxed{Y_{i\cdot}} = \sum_{j}^{n_i} Y_{ij}$

$\boxed{\bar{Y}_{i\cdot}} = \frac{\sum_{j}^{n_i} Y_{ij}}{n_i} = \frac{Y_{i\cdot}}{n_i}$

$\boxed{\bar{Y}_{\cdot\cdot}} = \frac{\sum_{i=1}^{r}\sum_{j=1}^{n_i} Y_{ij}}{n_T} = \frac{Y_{\cdot\cdot}}{n_T}$

$\boxed{Y_{\cdot\cdot}} = \sum_{i=1}^{r}\sum_{j=1}^{n_i} Y_{ij}$

$\boxed{\hat{\mu}_i} = \bar{Y}_{i\cdot}$

$\boxed{\hat{Y}_{ij}} = \bar{Y}_{i\cdot}$

$\boxed{\sum_j e_{ij}} = 0 \quad i = 1, 2, \ldots r$ errors for each factor level sum to 0

Given by both proc GLM and ANOVA above

$\boxed{SSTO} = \sum_i \sum_j (Y_{ij} - \bar{Y}_{\cdot\cdot})^2$

$\boxed{SSTreatment} = \sum_i n_i (\bar{Y}_{i\cdot} - \bar{Y}_{\cdot\cdot})$

$\boxed{SSE} = \sum_i \sum_j (Y_{ij} - \bar{Y}_{i\cdot})^2 = \sum_i \sum_j e_{ij}^2$

$\boxed{MStreatment} = \frac{SSTreatment}{r-1}$

$\boxed{MSE} = \frac{SSE}{n_T - r}$

degrees of freedom

$\underset{n_T-1}{SSTO} = \underset{r-1}{SSTR} + \underset{n_T-r}{SSE}$

$\boxed{E(MSTreatment)} = \sigma^2 + \frac{\sum n_i (\mu_i - \mu_\cdot)^2}{r-1}$

$\boxed{E(MSE)} = \sigma^2$

F Test for equality of factor level means

$H_0: \mu_1 = \mu_2 = \dots \mu_r$

H_a: not all μ_i are equal

$F^* = \dfrac{MSTreatment}{MSE}$

Reject when $F^* > F(1-\alpha;\ r-1,\ n_T - r)$

*p-value for this given by proc anova and GLM on p 69

Factor Effects Model

$Y_{ij} = \underset{\substack{\uparrow \\ \text{mean response}}}{\mu_\cdot} + \underset{\substack{\uparrow \\ \text{factor}}}{\tau_i} + \epsilon_{ij}$

$\tau_i = \mu_i - \mu$

$i = 1, \dots r$

$j = 1, \dots, n_i$

$\sum_i \tau_i = 0$

Weighted Mean

$\mu_\cdot = \sum_{i=1}^{r} w_i \mu_i \qquad \sum_{i=1}^{r} w_i = 1$

$\sum_{i=1}^{r} w_i \tau_i = 0$

Test for equality of factor level means

$H_0: \tau_1 = \tau_2 = \dots \tau_r = 0$

H_a: not all $= 0$

use same F test above.

Factor Effects Model

$\sum_1^r \tau_i = 0 \rightarrow \tau_r = -\tau_1 - \tau_2 \cdots \tau_{r-1}$

We can fit the linear model w/ $\mu, \tau_1, \tau_2, \dots \tau_{r-1}$

So: $Y_{ij} = \mu_\cdot + \tau_1 X_{ij1} + \cdots \tau_{r-1} X_{ijr-1} + \epsilon_{ij}$

regular $X_{ijk} = \begin{cases} 1 & \text{if case is from factor level } k \\ -1 & \text{if case is from factor level } r \\ 0 & \text{ow} \end{cases}$

weighted $X_{ijk} \begin{cases} 1 & \text{if case is from factor level } k \\ -\frac{n_k}{n_r} & \text{if case is from factor level } r \\ 0 & \text{ow} \end{cases}$

* Follow same linear regression from 704

$\boxed{\text{SAS}}$

```
/* set up taus in data step */
proc glm;
  model y= tau1 tau2 ...;
  estimate 'tau r' tau1 -1 tau2 -1 ... taur-1 -1;
run;
```