Analysis of Variance (ANOVA) – Prof. Richard B. Goldstein

Assumptions: k populations are *independent* and *normally distributed* with means $\mu_1, \mu_2, \ldots, \mu_k$ and *common variance* σ^2

Hypothesis:
H_0: $\mu_1 = \mu_2 = \ldots = \mu_k$
H_1: at least two of the means are not equal

Notation:
x_{ij} = j^{th} observation of i^{th} treatment or group

$\sum_{j=1}^{n_i} x_{ij}$ the sum of observations of the i^{th} treatment

$\bar{x}_i = \dfrac{\sum_{j=1}^{n_i} x_{ij}}{n_i}$ the mean of the i^{th} group

$\sum_{i=1}^{k}\sum_{j=1}^{n_i} x_{ij}$ the total of all observations

$\bar{x} = \dfrac{\sum_{i=1}^{k}\sum_{j=1}^{n_i} x_{ij}}{N}$ the mean of all $N = n_1 + n_2 + \ldots + n_k$ observations

Sum of Sqs:

$$\sum_{i=1}^{k}\sum_{j=1}^{n_i}\left(x_{ij}-\bar{x}\right)^2 = \sum_{i=1}^{k} n_i\left(\bar{x}_i-\bar{x}\right)^2 + \sum_{i=1}^{k}\sum_{j=1}^{n_i}\left(x_{ij}-\bar{x}_i\right)^2$$

$SS_{TOT} = SS_{BET} + SS_W$

$$SS_{TOT} = \sum x_{TOT}^2 - \frac{\left(\sum x_{TOT}\right)^2}{N} = \sum_{i=1}^{k}\sum_{j=1}^{n_i} x_{ij}^2 - \frac{\left(\sum_{i=1}^{k}\sum_{j=1}^{n_i} x_{ij}\right)^2}{N}$$

$$SS_{BET} = \sum_{\text{all groups}}\left(\frac{\left(\sum x_i\right)^2}{n_i}\right) - \frac{\left(\sum x_{TOT}\right)^2}{N} = \sum_{i=1}^{k}\frac{\sum_{j=1}^{n_i} x_{ij}^2}{n_i} - \frac{\left(\sum_{i=1}^{k}\sum_{j=1}^{n_i} x_{ij}\right)^2}{N} \qquad s_1^2 = MS_{BET} = \frac{SS_{BET}}{k-1}$$

$$SS_W = \sum_{\text{all groups}}\left(\sum x_i^2 - \frac{\left(\sum x_i\right)^2}{n_i}\right) = \sum_{i=1}^{k}\left(\sum_{j=1}^{n_i} x_{ij}^2 - \frac{\left(\sum_{j=1}^{n_i} x_{ij}\right)^2}{n_i}\right) \qquad s_2^2 = MS_W = \frac{SS_W}{N-k}$$

$F = \dfrac{MS_{BET}}{MS_W}$ has $k-1$ d.f. in numerator and $N-k$ d.f. in denominator