$y_{ij} = \mu_i + \epsilon_{ij}$, $j = 1, 2, \ldots, n_i$; $i = 1, 2, \ldots, k$, $\epsilon_{ij} \sim N(0, \sigma^2)$ i.i.d. Are the group means different?

$\begin{pmatrix} \hat{\mu}_1 \\ \vdots \\ \hat{\mu}_k \end{pmatrix} = \begin{pmatrix} \bar{y}_1 \\ \vdots \\ \bar{y}_k \end{pmatrix}$ so that RSS $= \sum_{i=1}^k \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_i)^2$.

To test $H_0 : \mu_1 = \mu_2 = \cdots = \mu_k$, consider
$A_{(k-1)\times k} = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 & -1 \\ 0 & 1 & 0 & \cdots & 0 & -1 \\ \vdots & \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & -1 \end{pmatrix}$. Then we test $H_0 : A\mu = 0$ where A has rank $k-1$. To test H_0, we obtain $\hat{\mu}_{H_0}$, RSS_{H_0} and consider

$$F = \frac{(\text{RSS}_{H_0} - \text{RSS})/(k-1)}{\text{RSS}/(\sum_{i=1}^k n_i - k)}, \text{ which} \sim F_{k-1, \sum_{i=1}^k n_i - k} \text{ under } H_0.$$

To find $\hat{\mu}_{H_0}$, RSS_{H_0}, note that, under $H_0 : \mu_1 = \mu_2 = \cdots = \mu_k$, these means are equal, and so it is enough to find

$$\min_{\mu_1 = \mu_2 = \cdots = \mu_k} \sum_{i=1}^k \sum_{j=1}^{n_i} (y_{ij} - \mu_i)^2 = \min_{\mu} \sum_{i=1}^k \sum_{j=1}^{n_i} (y_{ij} - \mu)^2.$$

Therefore,

$$\hat{\mu}_{H_0} = \frac{1}{\sum_{i=1}^k n_i} \sum_{i=1}^k \sum_{j=1}^{n_i} y_{ij} \equiv \bar{y}_{..}, \text{ and hence } \text{RSS}_{H_0} = \sum_{i=1}^k \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_{..})^2.$$

Introduce further notation: $\bar{y}_{i.} = \bar{y}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} y_{ij}$, $i = 1, 2, \ldots, k$. Note, further, that

$$\begin{aligned} &\text{RSS}_{H_0} \\ &= \sum_{i=1}^k \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_{..})^2 = \sum_{i=1}^k \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_{i.} + \bar{y}_{i.} - \bar{y}_{..})^2 \\ &= \sum_{i=1}^k \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_{i.})^2 + \sum_{i=1}^k n_i (\bar{y}_{i.} - \bar{y}_{..})^2 + 2 \sum_{i=1}^k \left\{ (\bar{y}_{i.} - \bar{y}_{..}) \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_{i.}) \right\} \\ &= \text{RSS} + \sum_{i=1}^k n_i (\bar{y}_{i.} - \bar{y}_{..})^2, \end{aligned}$$

since $\sum_{j=1}^{n_i}(y_{ij} - \bar{y}_{i.}) = 0$ for all i. Therefore,

$$\text{RSS}_{H_0} - \text{RSS} = \sum_{i=1}^{k} n_i(\bar{y}_{i.} - \bar{y}_{..})^2$$

and therefore,

$$F = \frac{\sum_{i=1}^{k} n_i(\bar{y}_{i.} - \bar{y}_{..})^2/(k-1)}{\sum_{i=1}^{k}\sum_{j=1}^{n_i}(y_{ij} - \bar{y}_{i.})^2/(\sum_{i=1}^{k} n_i - k)} \sim F_{k-1, \sum_{i=1}^{k} n_i - k} \text{ under } H_0.$$

It is instructive to consider these sum of squares.
$\text{RSS} = \sum_{i=1}^{k}\sum_{j=1}^{n_i}(y_{ij} - \bar{y}_{i.})^2$
= the sum total of all the sum of squares of deviations from the sample means
= within groups or within treatments sum of squares, SS_W.

$\text{RSS}_{H_0} = \sum_{i=1}^{k}\sum_{j=1}^{n_i}(y_{ij} - \bar{y}_{..})^2$
= total sum of squares of deviations assuming no treatment effect
= total variability (corrected) in the k samples, SS_T.

Therefore, $\sum_{i=1}^{k} n_i(\bar{y}_{i.} - \bar{y}_{..})^2$ = SS_T - SS_W = between groups or between treatments sum of squares = SS_B. Thus,
$\text{SS}_T = \text{SS}_W + \text{SS}_B$ is the decomposition of sum of squares along with
$\sum_{i=1}^{k} n_i - 1 = (\sum_{i=1}^{k} n_i - k) + (k-1)$, decomposition of d.f.

ANOVA for One-way classification

source	d.f.	SS	MS	F
Treatments (groups)	$k-1$	$\text{SS}_B = \sum_{i=1}^{k} n_i(\bar{y}_{i.} - \bar{y}_{..})^2$	$\text{MS}_B = \frac{\text{SS}_B}{k-1}$	$\frac{\text{MS}_B}{\text{MSE}} \sim$ (under H_0) $F_{k-1, \sum_{i=1}^{k} n_i - k}$
Error	$\sum n_i - k$	$\text{SS}_W = \sum\sum(y_{ij} - \bar{y}_{i.})^2$	$\text{MSE} = \frac{\text{SS}_W}{\sum_{i=1}^{k} n_i - k}$	
Total (corrected)	$\sum n_i - 1$	$\text{SS}_T = \sum\sum(y_{ij} - \bar{y}_{..})^2$		
Mean	1	$(\sum_{i=1}^{k} n_i)\bar{y}_{..}^2$		
Total	$\sum n_i$	$\sum_{i=1}^{k}\sum_{j=1}^{n_i} y_{ij}^2$		

Example. Tensile strength data. $k = 5$, $n_i = 5$. ANOVA is as follows.

source	d.f.	SS	MS	F
Factor levels (% cotton)	4	475.76	118.94	$14.76 >> 4.43 = F_{4,20}(.99)$
Error	20	161.20	8.06	
Total(corrected)	24	636.96		

$R^2 = \frac{475.76}{636.96} \approx 75\%$

Now that the ANOVA H_0 has been rejected, we should look at the group means (estimates) closely. Suppose we want to compare μ_r and μ_s either with $H_0 : \mu_r = \mu_s$ or using a confidence interval for $\mu_r - \mu_s$.

$$\hat{\mu}_r - \hat{\mu}_s = \bar{y}_{r.} - \bar{y}_{s.} \sim N\left(\mu_r - \mu_s, \sigma^2\left(\frac{1}{n_r} + \frac{1}{n_s}\right)\right)$$

independently of

$$\sum_{i=1}^{k}\sum_{j=1}^{n_i}(y_{ij} - \bar{y}_{i.})^2 \sim \sigma^2\chi^2_{\sum_{i=1}^k n_i - k}.$$

Therefore,

$$\frac{\{(\bar{y}_{r.} - \bar{y}_{s.}) - (\mu_r - \mu_s)\} / \sqrt{\frac{1}{n_r} + \frac{1}{n_s}}}{\sqrt{\sum_{i=1}^k \sum_{j=1}^{n_i}(y_{ij} - \bar{y}_{i.})^2 / \left(\sum_{i=1}^k n_i - k\right)}} \sim t_{\sum_{i=1}^k n_i - k}.$$

$100(1-\alpha)\%$ confidence interval for $\mu_r - \mu_s$ is

$$\bar{y}_{r.} - \bar{y}_{s.} \pm t_{\sum_{i=1}^k n_i - k}(1-\alpha/2)\sqrt{\sum_{i=1}^{k}\sum_{j=1}^{n_i}(y_{ij} - \bar{y}_{i.})^2 / \left(\sum_{i=1}^{k} n_i - k\right)}\sqrt{\frac{1}{n_r} + \frac{1}{n_s}}.$$

Further, test statistic for testing $H_0 : \mu_r = \mu_s$ is

$$T = \frac{(\bar{y}_{r.} - \bar{y}_{s.}) / \sqrt{\frac{1}{n_r} + \frac{1}{n_s}}}{\sqrt{\sum_{i=1}^k \sum_{j=1}^{n_i}(y_{ij} - \bar{y}_{i.})^2 / \left(\sum_{i=1}^k n_i - k\right)}} \sim t_{\sum_{i=1}^k n_i - k},$$

if H_0 is true.