Thursday March 12

(Tues. 10th was Test 2)

Testing $H_0: \mu_1 = \mu_2 = \cdots = \mu_k$

There are n_j observations in Group j, $j = 1, \ldots, k$

$$n = \sum_{j=1}^{k} n_j \qquad \bar{X}_{\cdot} = \sum_{j=1}^{k} \left(\frac{n_j}{n}\right) \bar{X}_j = \underbrace{\frac{1}{n} \sum_{j=1}^{k} \sum_{i=1}^{n_j} X_{ij}}$$

X_{ij} is observation i in group j, $\overset{ind}{\sim} N(\mu_j, \sigma^2)$

Found

$$\sum_{j=1}^{k} \sum_{i=1}^{n_j} (X_{ij} - \bar{X}_{\cdot})^2 = \sum_{j=1}^{k} n_j (\bar{X}_j - \bar{X}_{\cdot})^2 + \sum_{j=1}^{k} \sum_{i=1}^{n_j} (X_{ij} - \bar{X}_j)^2$$

$$SSTO = SSB + SSW$$

Total sum of squares — Between Groups sum of squares — within groups sum of squares

$$\frac{SSTO}{\sigma^2} = \frac{SSB}{\sigma^2} + \frac{SSW}{\sigma^2}$$

$$\frac{SSTO}{\sigma^2} = \frac{(n-1)S^2}{\sigma^2} = Y_1$$

$$\frac{SSB}{\sigma^2} = Y_1$$

$Y_1 \sim \chi^2(k-1)$

$$\frac{SSW}{\sigma^2} = \sum_{j=1}^{k} \frac{(n_j - 1) S_j^2}{\sigma^2} = Y_2 \sim \chi^2\left(\sum_{j=1}^{k}(n_j - 1)\right)$$

$\chi^2(n-k)$

Under H_0

$Y_1 \sim \chi^2(n-1)$

Have $Y_1 = \sum_{j=1}^{k} n_j (\bar{X}_j - \bar{X}_.)^2 \sim \chi^2(\overset{k-1}{\cancel{n-k}})$ SSB

$Y_2 \sim \sum_{j=1}^{k} \sum_{i=1}^{n_j} (X_{ij} - \bar{X}_j)^2 \sim \chi^2(n-k)$ SSW

$$F = \frac{Y_1 / k-1}{Y_2 / n-k} = \frac{SSB/(k-1)}{SSW/(n-k)} \overset{H_0}{\sim} F(k-1, n-k)$$

OFTEN arranged in an Analysis of Variance summary table.

MEAN SQUARES

Source	SS	df	MS	F
Between Groups	$\sum_{j=1}^{k} n_j(\bar{X}_j - \bar{X}_.)^2$	$k-1$	$\frac{SSB}{k-1}$	$\frac{MSB}{MSW}$
Within Groups	$\sum_{j=1}^{k}\sum_{i=1}^{n_j}(X_{ij} - \bar{X}_.)^2$	$n-k$	$\frac{SSW}{n-k}$	
Total	$\sum_{j=1}^{k}\sum_{i=1}^{n_j}(X_{ij} - \bar{X}_.)^2$	$n-1$		

$$SSTO = SSB + SSW$$

$$\sum_{j=1}^{k}\sum_{i=1}^{n_j}(x_{ij}-\bar{x}_{.})^2 = \sum_{j=1}^{k} n_j(\bar{x}_j-\bar{x}_{.})^2 \qquad \sum_{j=1}^{k}\sum_{i=1}^{n_j}(x_{ij}-\bar{x}_j)^2$$

SSTO = Total variation in X_{ij} to be "explained."

SSW = Variation still unexplained by Group

SSB Must be explained variation.

$$R^2 = \frac{SSB}{SSTO} = \frac{\text{Explained Variation}}{\text{Total variation}}$$

Proportion of variation in data explained by group membership $0 \le R^2 \le 1$

Strength of Relationship Between Group Membership & X_{ij}

$$F = \frac{\dfrac{SSB/(k-1)}{SSTO}}{\dfrac{SSW/n-k}{SSTO}} = \left(\frac{n-k}{k-1}\right)\frac{R^2}{\dfrac{SSTO-SSB}{SSTO}}$$

$$= \left(\frac{n-k}{k-1}\right)\frac{R^2}{1-R^2}$$

TWO WAYS TO Get BIG F & Reject H_0

- Strong relationship
- Big sample size

47

Example from STA 258 text: P.671

Randomly assign students to 4 different teaching methods, teach & test.

	Method			
	1	2	3	4
	65	75	59	94
	87	69	78	89
	73	83	67	80
	79	81	62	88
	81	72	83	
	69	79	76	
		90		
n_j	6	7	6	4
$\bar{x}_j$	75.67	78.43	70.83	87.75
s_j^2	66.67	50.62	91.77	33.58

$$n = \sum_{j=1}^{k} n_j = 6 + 7 + 6 + 4 = 23$$

$$\bar{x}_\circ = \sum_{j=1}^{k} \frac{n_j}{n} \bar{x}_j = \frac{1}{23}\left(6 * 75.67 + 7 * 78.43 + 6 * 70.83 + 4 * 87.75\right)$$

$$= 77.35$$

$$SSE = \sum_{j=1}^{k} \sum_{i=1}^{n_j} (x_{ij} - \bar{x}_j)^2 = 1196.6$$

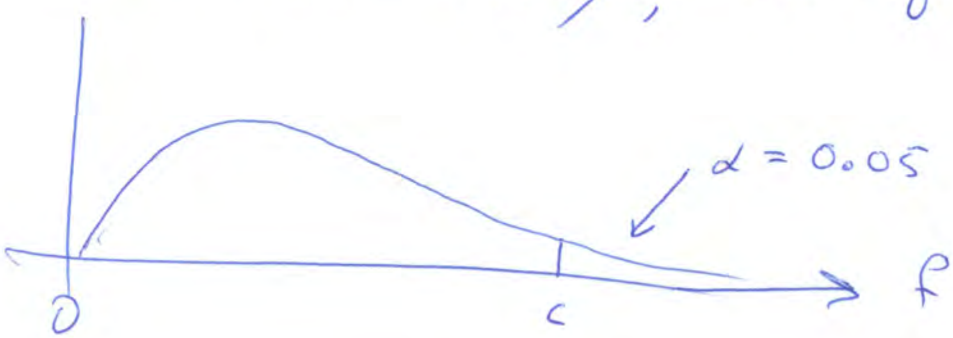
$$SSB = \sum_{j=1}^{k} n_j (\bar{x}_j - \bar{x}_\circ)^2 = 712.6$$

$$F = \frac{SSB/(k-1)}{SSW/(n-k)} = \frac{712.6/(4-3)}{1196.6/(23-4)} = 3.77$$

$$F = 3.77 > f_{0.95}(3, 19) = 3.13$$

So Reject H_0 at $\alpha = 0.05$

Conclude Not all μ_i are equal.



Follow up with 2-sample t-tests

$\binom{k}{2}$ of them. The story continues ...

$$R^2 = \frac{SSB}{SSTO} = \frac{712.6}{712.6 + 1196.6} = 0.37$$