STAT 371 Discussion 10

1. Basic concepts

1.1 Data Structure

Assume we have k treatment groups

Group 1: Y_{11}, Y_{12}, Y_{13}, Y_{14}, …..

.

.

Group k: Y_{k1}, Y_{k2}, Y_{k3}, Y_{k4}, …..

1.2 Some Notations

n_t	number of cases in treatment group t
N	number of cases (overall)
Y_{ti}	response i in treatment group t
$\bar{Y}_{t\cdot}$	average response in treatment group t
$\bar{Y}_{\cdot\cdot}$	average response (overall)

1.3 Sum of Squares

1) Sum of squares for each group

$$S_t = \sum_{i=1}^{n_t} (Y_{ti} - \bar{Y}_{t\cdot})^2.$$

2) Within group sum of squares, df_W = K-1

$$S_W = S_1 + \cdots + S_k = \sum_t \sum_i (Y_{ti} - \bar{Y}_{t\cdot})^2$$

3) Between group sum of squares, df_B = N-k

$$S_B = \sum_{t=1}^{k} n_t(\bar{Y}_{t\cdot} - \bar{Y}_{\cdot\cdot})^2$$

4) Total group sum of squares: $\mathbf{S_T = S_W + S_B}$, df = N-1 = $df_W + df_B$

1.4 ANOVA Table

source	sum of squares	df	mean square
between treatments	$S_B = \sum_t n_t(\bar{Y}_{t\cdot} - \bar{Y}_{\cdot\cdot})^2$	$k-1$	$M_B = S_B/(k-1)$
within treatments	$S_W = \sum_t \sum_i (Y_{ti} - \bar{Y}_{t\cdot})^2$	$N-k$	$M_W = S_W/(N-k)$
total	$S_T = \sum_t \sum_i (Y_{ti} - \bar{Y}_{\cdot\cdot})^2$	$N-1$	

Note: book uses SS_B, SS_W, SS_T, MS_B, MS_W

2. Standard ANOVA Model

2.1 Model

$Y_{ti} = \mu_t + \varepsilon_{ti}$

$Y_{ti} = \mu + \tau_t + \varepsilon_{ti}$,

where $Y_{ti} \sim$ iid $N(\mu_t, \sigma^2)$ and $\varepsilon_{ti} \sim$ iid $N(0,\sigma^2)$

2.2 Hypothesis Test

1) State the hypothesis

   $H_0 : \mu_1 = \cdot \cdot \cdot = \mu_k$ vs. Ha : H_0 is false

2) Test statistic

   $F = M_B/M_W$

3) Distribution of F under H_0

   $F \sim F_{k-1,\, N-k}$

   where k-1 is the numerator degrees of freedom, N-k is the denominator degrees of freedom.

4) Use one-sided F test to find the p-value

   p-value = $Pr(F>F_{obs})$

   Method I: Table 10

   Method II. P-value = **1-pf(F_{obs}, k-1, N-k)**

3. Example

11. 13 A researcher studied the flexibility of 10 women in an aerobic exercise class, 10 women in a modern dance class and a control group of 9 women. One measurement she made on each woman was spinal extension, which was a measure of how far the woman could bend her back. Measurements were made before and after a 16-week training period. The change in spinal extension was recorded for each woman. The ANOVA SS_B is 7.04 and the SS_W is 15.08.

(a) State the null hypothesis

(b) Construct the ANOVA table and test the null hypothesis. Let $\alpha = 0.01$