Single-Factor ANOVA

Hypothesis $H_0 : \mu_1 = \ldots = \mu_I$ vs H_1 : at least two $\mu_i's$ differ.

Assumptions: (1)Each sample has equal sample size (J) (2) Normal populations (3) Equal variance: $\sigma_1^2 = \ldots = \sigma_I^2$.

Idea: In the pooled two-sample t test with $m = n$, $T = \frac{\bar{X}_{1.} - \bar{X}_{2.} - 0}{\sqrt{S_p^2 \times 2/n}} \sim t_{2n-2} \rightarrow T^2 = \frac{(\bar{X}_{1.} - \bar{X}_{..})^2 + (\bar{X}_{2.} - \bar{X}_{..})^2}{S_p^2/n} \sim F_{1,2n-2}$

Raw data	Sample 1	...	Sample I		Treat all the data from one sample
	X_{11}	...	X_{I1}		$\bar{X}_{..} = \frac{1}{IJ}(X_{11} + \ldots + X_{1J} + \ldots + X_{I1} + \ldots + X_{IJ})$
	X_{12}	...	X_{I2}		$= \frac{1}{I}(\bar{X}_{1.} + \ldots + \bar{X}_{I.})$
	$\vdots$	$\vdots$	$\vdots$	Total sum of square (SST)=	$\sum_{i=1}^{I} \sum_{j=1}^{J} (X_{ij} - \bar{X}_{..})^2$
	X_{1J}	...	X_{IJ}		
Sample mean	$\bar{X}_{1.} = \frac{1}{J}\sum_{j=1}^{J} X_{1j}$	...	$\bar{X}_{I.} = \frac{1}{J}\sum_{j=1}^{J} X_{Ij}$	Sum of square of treatment (SSTr)=	$J\sum_{i=1}^{i}(\bar{X}_{i.} - \bar{X}_{..})^2$
Sample variance	$S_1^2 = \frac{1}{J-1}\sum_{j=1}^{J}(X_{1j} - \bar{X}_{1.})^2$	...	$S_I^2 = \frac{1}{J-1}\sum_{j=1}^{J}(X_{Ij} - \bar{X}_{I.})^2$	Sum of square of error (SSE)=	$\sum_{i=1}^{I}\sum_{j=1}^{J}(X_{ij} - \bar{X}_{i.}) = (J-1)\sum_{i=1}^{I} S_i^2$
					This one is used to calculate the pool sample variance.

Theorem: Under all the assumptions (balanced data, normal distribution, equal variance), we have

1. $SSE/\sigma^2 \sim \chi^2_{I(J-1)}$ no matter H_0 is true or not;
2. $SSTr/\sigma^2 \sim \chi^2_{I-1}$ if and only if H_0 is true;
3. $SSTr$ and SSE are independent random variables, where $SSE = \sum_{i=1}^{I}\sum_{j=1}^{J}(\bar{X}_{ij} - \bar{X}_{i.})^2 = (J-1)\sum_{i=1}^{I} S_i^2$ and $SSTr = J\sum_{i=1}^{I}(\bar{X}_{i.} - \bar{X}_{..})^2$

Test: The F test: $F = \frac{MSTr}{MSE} = \frac{SSTr/[\sigma^2(I-1)]}{SSE/[\sigma^2 I(J-1)]} \overset{H_0}{\sim} F_{I-1,I(J-1)}$. We reject H_0 at level α whenever $F > F_{\alpha,I-1,I(J-1)}$.

Table 1: ANOVA TABLE

Source of variation	df	Sum of Squares	Mean Squares	F
Treatments	$I-1$	$SSTr$	$MSTr$	$MSTr/MSE$
Error	$I(J-1)$	SSE	MSE	
Total	$IJ-1$	SST		