More ANOVA

Tests to determine equivalence of variances and means between two or more samples is termed ANOVA or one way ANalysis Of VAriance

Includes:

- F-test on group variances or Kruskal-Wallis on group medians
- Partitions variance *within* and *between* group variance
- Used to determine if observed difference in means can be attributed to natural variation in population

One-way ANOVA (One variable)

- Groups need to be independent
- Each group is random sample from a normal population
- In the population, the variances of the groups are equal (should test first)

- Partition variation into parts:

  1. Between group variation
  2. Within group variation
  3. Total variation

k = number of groups, n = objects in group, N = total objects , i = row, j = column, $\bar{x}$ = group mean, $\bar{\bar{x}}$ = grand mean

Total Variation = $SS_T = \sum_{j=i}^{k} \sum_{i=1}^{n_j} (x_{ij} - \bar{\bar{x}})^2$ variation = $S_T^2 = SS_T / N-1$

Within Groups = $SS_W = \sum_{j=i}^{k} \sum_{i=1}^{n_i} (x_{ij} - \bar{x}_j)^2$ variation = $S_W^2 = SS_W / N-K$

Between Groups = $SS_B = \sum_{j=1}^{k} n_j (\bar{x}_j - \bar{\bar{x}})^2$ variation = $S_B^2 = SS_B / K-1$

$SS_B + SS_W = SS_T$

To test hypothesis that population means from the groups are equal (and any differences are due to natural variation), use F test where

$$F = \frac{SS_B / K-1}{SS_W / N-K}$$

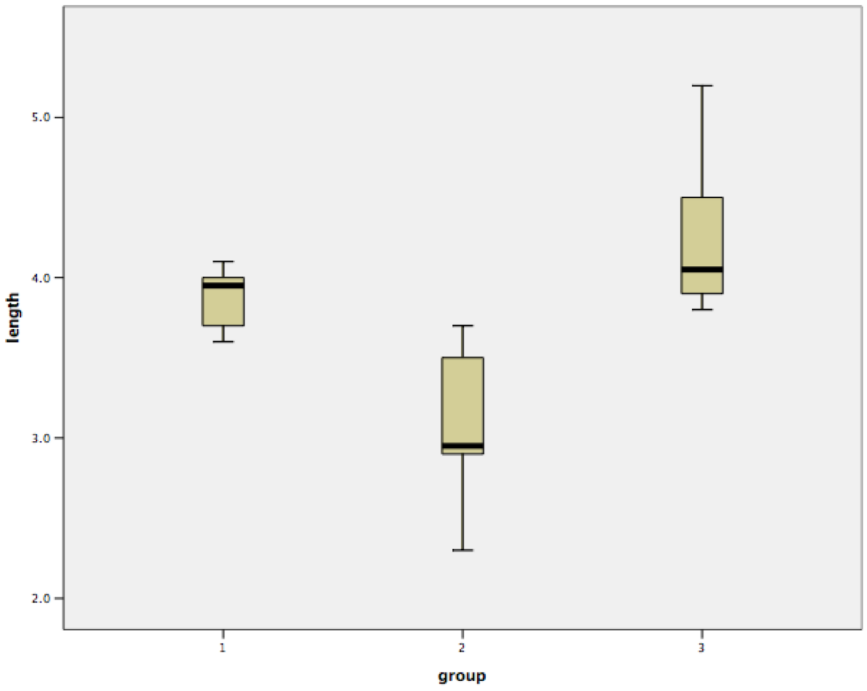
F is distributed with K-1 and N-K degrees of freedom

Example

Data are wing lengths (in m) of adult pterosaurs

Group 1	Group 2	Group 3
4.0	2.9	4.5
3.6	2.3	3.8
3.7	2.9	4.0
4.1	3.5	5.2
3.9	3.7	3.9
4.0	3.0	4.1

Box plot of data.



Step 1. Formulate hypothesis and set α

H_o: population means from the groups are equal
H_1: population means from the groups are unequal

$\alpha = 0.05$

Step 2. Calculate SS_B, SS_W, etc... and F value

Output from SPSS (Analyze -> Compare means->ANOVA

Descriptives

length

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
1	6	3.883	.1941	.0792	3.680	4.087	3.6	4.1
2	6	3.050	.4970	.2029	2.528	3.572	2.3	3.7
3	6	4.250	.5244	.2141	3.700	4.800	3.8	5.2
Total	18	3.728	.6569	.1548	3.401	4.054	2.3	5.2

ANOVA

length

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	4.538	2	2.269	12.162	.001
Within Groups	2.798	15	.187		
Total	7.336	17			

Step 3. Compare with critical value from table

With N-1 (18-1=17) and N-K (18-3=15) degrees of freedom and a two-tailed α of 0.5 the critical value of F = 2.37

Because a F_{obs} of 12.162 is outside (greater than) a F_{crit} of 2.37, we must reject the null hypothesis that the population means from the groups are equal.

ANOVA

length

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	4.538	2	2.269	12.162	.001
Within Groups	2.798	15	.187		
Total	7.336	17			