Chapter 11: Comparison of Means from Many Independent Samples

Part II

Eric D. Nordmoe

Math 261
Department of Mathematics and Computer Science
Kalamazoo College

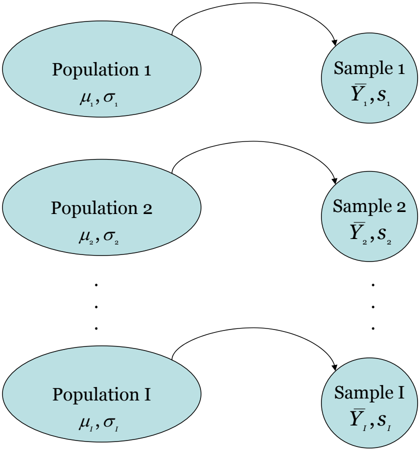
Spring 2009

Outline

- Checking Conditions
- Multiple Comparisons

Analysis of Variance Sampling Model

Draw samples from I independent populations to compare population means μ_1, μ_2, . . ., and μ_I:



Condition: $\sigma_1 = \sigma_2 = \cdots = \sigma_I$

Standard Conditions for ANOVA

- Design conditions:
  - Random samples: reasonable to consider observations a random sample from respective populations
  - Independent samples: the I samples are independent of each other
- Population:
  - Normal: Population distributions are normal (not crucial if n_i are large and similar)
  - Equal standard deviations:

$$\sigma_1 = \sigma_2 = \cdots = \sigma_I$$

Rules of Thumb for Checking ANOVA Conditions

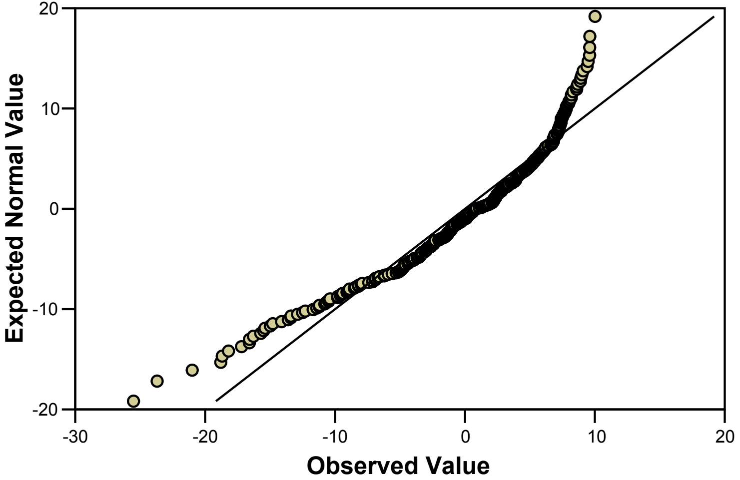
- Inference requires random samples.
- Outliers are always problematic! Plot your data.
- ANOVA methods are robust if group samples sizes are similar and not too small.
- Ratio of largest sample SD to smallest should not be much greater than 2.
  - Biggest problem is if sample sizes are unequal and SD from a small sample is much larger than others.
- Normality is not critical if sample sizes n_i are large and approximately equal.

Checking ANOVA Conditions

- Checking normality
  - Normal probability plots for each group
  - Normal probability plots of deviations from group means $(y_{ij} - \bar{y}_{i\cdot})$
- Checking equal standard deviations
  - Compare SDs
  - Plot deviations $(y_{ij} - \bar{y}_{i\cdot})$ against means $\bar{y}_{i\cdot}$ and check for trend

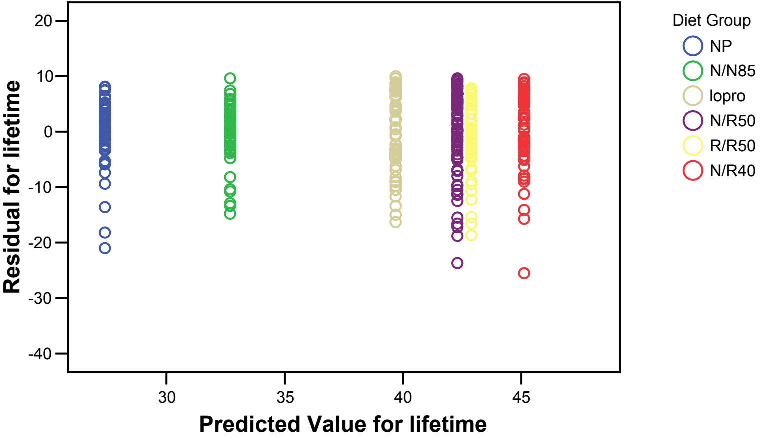
Checking ANOVA Conditions

Normal plot of $y_{ij} - \bar{y}_{i\cdot}$ for Diet Restriction Data



Checking ANOVA Conditions

Scatterplot of $y_{ij} - \bar{y}_{i\cdot}$ versus $\bar{y}_{i\cdot}$ for Diet Restriction Data



Multiple Comparisons

Follow-up to Global *F* Test

- If the global *F*-test rejects equality of means, use multiple comparison (post hoc) methods to identify significant differences.
  - Simple *t* tests have α probability of Type I Errors
  - With many tests, the familywise error rate (FWER) grows.
- Multiple comparison procedures attempt to control the FWER.
- Two methods are discussed in the Samuels-Witmer text:
  - The Bonferroni method
  - The Newman-Keuls or (Student-Newman-Keuls) method

Bonferroni Method

- The Bonferroni method is a general method for multiple comparisons.
  - Can be applied whenever multiple significance tests are being carried out.
  - Very conservative.
- Basic idea:
  - Goal: Test k hypotheses at overall α error rate.
  - Procedure: Carry out each of k tests at the α/k significance level.
- For a one-way ANOVA with I groups:
  - Total possible comparison: $k = I(I-1)/2$.
  - Test each at α/k level of significance.
- For confidence intervals, use multiplier $t_{\alpha/(2k)}$ to obtain overall $100(1-\alpha)\%$ confidence.

Bonferroni Example

Diet Restriction Study

- In the diet restriction study, there are $I = 6$ groups and $6(6-1)/2 = 15$ possible comparisons but ...
  - Only 5 comparisons are of interest so let $k = 5$.
- Desired overall level of significance is $\alpha = 0.05$.
- Set pairwise error rate equal to $.05/5 = .01$.
- Test $H_0 : \mu_1 = \mu_2$ by test statistic

$$t = \frac{\bar{y}_1 - \bar{y}_2}{s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

where

$$t \sim t(\text{df}_w) \text{ and } s_p = \sqrt{\text{MS(within)}}$$

Bonferroni Example

Confidence Intervals

- Bonferroni confidence interval:

$$(\bar{y}_1 - \bar{y}_2) \pm t_{\alpha/(2k)} s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

  where k is the number of intervals
  - Each interval has confidence level $100(1 - \alpha/k)\%$.
  - Overall confidence level is $100(1 - \alpha)\%$.

Bonferroni Example

SPSS Output for Diet Restriction Study

SPSS will only compute Bonferroni procedures for all possible comparisons:

Multiple Comparisons

Dependent Variable: Months Survived
Bonferroni

(I) Diet Group	(J) Diet Group	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
NP	N/N85	-5.28919*	1.30101	.001	-9.1348	-1.4435
	lopro	-12.28367*	1.30637	.000	-16.1452	-8.4222
	N/R50	-14.89514*	1.24030	.000	-18.5613	-11.2289
	R/R50	-15.48367*	1.30637	.000	-19.3452	-11.6222
	N/R40	-17.71463*	1.28588	.000	-21.5156	-13.9137
N/N85	NP	5.28919*	1.30101	.001	1.4435	9.1348
	lopro	-6.99449*	1.25652	.000	-10.7086	-3.2803
	N/R50	-9.60596*	1.18768	.000	-13.1166	-6.0953
	R/R50	-10.19449*	1.25652	.000	-13.9086	-6.4803
	N/R40	-12.42544*	1.23521	.000	-16.0766	-8.7743
. . .						
	N/R50	2.81948	1.17110	.249	-.6422	6.2811
	R/R50	2.23095	1.24086	1.000	-1.4369	5.8988

*. The mean difference is significant at the .05 level.

Newman-Keuls Procedure

- Procedure is based on the distribution of the differences of order statistics of sample means when all population means are equal and sample sizes are equal.
- The procedure is as follows:
  - Sort the sample means in increasing order.
  - Compute critical values $R_i = q_i s_p/\sqrt{n}$.
  - Make pairwise comparisons.

Newman-Keuls Procedure

Samuels-Witmer Example 11.28

- Blood urea concentrations (mg/d/Li) of rats on 5 different diets, $n_i = 4$, $I = 5$.
- ANOVA table

Source	df	SS	MS	F
Between diets	4	894.8	223.7	10.51
Within diets	15	319.35	21.29	
Total	19	1214.15		

- Test for differences between means at the $\alpha = 0.05$ level.

Newman-Keuls Procedure

Samuels-Witmer Example 11.28

- The unsorted group sample means are:

Diet	A	B	C	D	E
$\bar{y}_{i\cdot}$	40.0	40.7	32.9	29.6	48.8

- The sorted group sample means are:

Diet	D	C	A	B	E
$\bar{y}_{i\cdot}$	29.6	32.9	40.0	40.7	48.8

- To compute the critical values R_i, first find the scale factor $s_p/\sqrt{n} = \sqrt{21.29}/\sqrt{4} = 2.307$.
- Using the q_i from Table 11 with df $=$ df(within) $= 15$, construct the table of critical values:

i	2	3	4	5
q_i	3.01	3.67	4.08	4.37
R_i	6.9	8.5	9.4	10.1

Newman-Keuls Procedure

Samuels-Witmer Example 11.28

- The value $q_i = 4.37$ means that the test statistic

$$\frac{\bar{y}_{\text{max}} - \bar{y}_{\text{min}}}{s_p/\sqrt{n}}$$

  has a 5% chance of exceeding 4.37 if the population means are equal and standard conditions hold.
- Use the critical values R_i to make pairwise comparisons among means starting with the most extreme comparisons.
- At each step, use underlining to indicate groups that are not significantly different from each other.

Newman-Keuls Procedure

Samuels-Witmer Example 11.28

Sorted group sample means:

Diet	D	C	A	B	E
$\bar{y}_{i\cdot}$	29.6	32.9	40.0	40.7	48.8

Newman-Keuls Procedure

Samuels-Witmer Example 11.28

- Groups connected by lines are not significantly different at the $\alpha = 0.05$.
- Groups not connected by lines are significantly different.
- Newman-Keuls does not produce *P*-values or confidence intervals.

Newman-Keuls Procedure: SPSS Output

Diet Restriction Study

Instead of underlines, SPSS puts means that are not significantly different in the same subset:

Months Survived

Student-Newman-Keuls[a,b]

Diet Group	N	Subset for alpha = .05			
		1	2	3	4
NP	49	27.4020			
N/N85	57		32.6912		
lopro	56			39.6857	
N/R50	71				42.2972
R/R50	56				42.8857
N/R40	60				45.1167
Sig.		1.000	1.000	1.000	.062

Means for groups in homogeneous subsets are displayed.

a. Uses Harmonic Mean Sample Size = 57.462.

b. The group sizes are unequal. The harmonic mean of the group sizes is used. Type I error levels are not guaranteed.

- Results are only approximate since sample sizes are not equal.

The One-Way ANOVA Model

- For experiments, the one-way ANOVA model can be expressed in terms of treatment effects:

y_{ij}	$=$	μ	$+$	τ_i	$+$	ϵ_{ij}
observation		overall average		group effect		random error

where

  - μ is the grand population mean
  - τ_i is the *i*th treatment effect, $\tau_i = \mu_i - \mu$
  - ϵ_{ij} is the random error term.

- The mean of group *i* is $\mu_i = \mu + \tau_i$
- Two equivalent ANOVA hypothesis statements:
  - $H_0 : \mu_1 = \mu_2 = \cdots = \mu_I$
  - $H_0 : \tau_1 = \tau_2 = \cdots = \tau_I = 0$