

Choosing Mutually Orthogonal Coefficients

1. Select a comparison that tests the relations among all groups (i.e. Coefficients should carry the sign + or -). There should be no zeros as coefficients). It is handy if this comparison is the experimenter's hypothesized a priori comparison.
2. Make sure the sum of those coefficients is zero.
3. Make sure all the plus signs are adjacent to one another and all the minus signs are adjacent to one another.
4. Draw a line between the + and - signs of the first contrast.
5. Choose to create your comparison on either side of the line (this means all your + and - signs go on one side).
6. Create another set of coefficients that sum to zero. If you follow the rules carefully, they will be mutually orthogonal to the previous and subsequent comparisons.
7. Draw a line between the + and - signs of the second contrast. Repeat steps 5 and 6 as necessary.
8. Never "cross a line" in a subsequent contrast. This means that the coefficients that have plus and minus signs must always be between the vertical lines you draw. You can't have a plus on one side of a vertical line, and a minus on the other.

An example with 5 groups

	V	W	X	Y	Z	Hypothesis tested
C1	-1	-1	-1	-1	+4	-is Z is different?
C2	-1	-1	+1	+1	0	-are X&Y different than V&W?
C3	-1	+1	0	0	0	-is W different than V?
C4	0	0	-1	+1	0	-is Y different than X?

Another example

	V	W	X	Y	Z	Hypothesis tested
C1	-2	-2	-2	+3	+3	-are Y&Z different than V,W&X?
C2	-1	-1	+2	0	0	-is X different than V&W?
C3	-1	+1	0	0	0	-is W different than V?
C4	0	0	0	+1	-1	-is Z different than Y?

Note: To **prove** a set of coefficients is mutually orthogonal, sum the cross-products of each pair of coefficients. For example, in the immediately preceding example, C2 and C3 are proven orthogonal by showing that product of the coefficients for V are $-1 \cdot -1 = +1$, for W are $-1 \cdot +1 = -1$, for X are $+2 \cdot 0 = 0$, for Y are $0 \cdot 0 = 0$, and for Z are $0 \cdot 0 = 0$. The sum of $+1 -1 0 0$ and 0 is 0 .

.....0

Summary of Tests Subsequent to ANOVA

Test	Statistic	Formula	df	a priori?	Type	alpha
Orthogonal Comparison	F	$\frac{L^2/n\sum a^2}{MS_{within}}$	1 dfw	yes	$\sum a=0$ df=k	each .05
Bonferroni Comparison	F	$\frac{L^2/n\sum a^2}{MS_{within}}$	1 dfw	yes	$\sum a \neq 0$ df \neq k	each .05/k
Scheffé Comparison	F	$\frac{(L^2/n\sum a^2)/g-1}{MS_{within}}$	g-1 dfw	no	$\sum a \neq 0$ df \neq k	family .05
Tukey	q	$\frac{M - M}{\sqrt{MS_{within}/n}}$	g dfw	yes	all pairs	family .05
Newman-Keuls	q	$\frac{M - M}{\sqrt{MS_{within}/n}}$	g* dfw	yes	all pairs	family .05

g* number of groups within the comparison range

M = the sample mean

Which kind of ANOVA is it?

When designing an experiment, researchers choose among various types of experimental designs. Certain kinds of questions are best tested with between-subjects designs (i.e., one-way ANOVA or a between-subjects factorial design, such as a $2 \times 2 \times 3$). Other questions require purely within-subjects designs (e.g., $S \times A$, $S \times A \times B$) or a mixture of between and within factors (e.g., $S/A \times B$).

In making their decision about which design to use, researchers consider a variety of issues; these issues include theory, experimental paradigm, and state-of-the-art technology. We haven't considered these issues in Psych 212; you will probably consider them in your methodology courses. We have discussed some reasons for the choice of one design over another, however. For example, if there are likely to be carry-over effects from one level of an independent variable to another level (no matter how long the interval between levels), one might consider making that IV between-subjects. As another example, when the cost of data collection is high, a within-subjects or mixed design might be preferred.

It is not the purpose of this course to teach you how to create an experiment. But, once you are ready to create an experiment, you will have some tools with which to analyze data. For the purposes of this course, you should be able to recognize different experimental designs so that you can analyze them appropriately. Following the steps below will help you decide which ANOVA to use.

1. Identify the dependent variable (i.e., what is being measured). Now ask yourself: are subjects completing that measure more than once? If the answer is no, then you have a between subjects design. If the answer is yes, you have at least one within subjects variable.
2. Identify the independent variable(s). These are the factors that are responsible for systematic variability among subjects' responses to the DV (as opposed to error variability). Some Ivs will be manipulated by the researcher and as such are true causes. Other IVs will be "Subject Variables" which the researcher expects contribute variability but they aren't really manipulated (e.g., gender, IQ). (Remember that if you answered "yes" to question 1, subjects is a random independent variable).
3. For each of the independent variables (except for the random IV subjects), count the number of levels
4. Draw the design as a series of boxes.
5. For the first independent variable (except for the random IV subjects), ask yourself whether any given subject receives only one level of it or if a subject receives all levels of that IV. If you answer the former, then this IV is between subjects. On your drawing, write "between" in parenthesis next to the factor name. If you answer the latter, this IV is within subjects. On your drawing, write "within" in parenthesis next to the factor name.
6. Repeat step 5 for as many IVs as you have identified.
7. If all of your IVs are between, describe the design in terms of how many IVs there are (and the number of levels in each IV). For example, if there are three independent variables, each of which have 2 levels, you have a $2 \times 2 \times 2$ between subjects design (in which case you do a 3-way between subjects ANOVA). For purposes of this course, you know 1-way, 2-way, and 3-way between-subjects ANOVAs. (There are 4-ways, 5-ways, etc).
- 7a. If all your IVs are within, you have an S by "something else" design. For purposes of this course, S by "something else" has been $S \times A$ and $S \times A \times B$. (There are $S \times A \times B \times C$, in which there are three within-subjects variables, etc).
- 7b. For the purposes of this course, if one IV is within and the other is between, you have an $S/A \times B$ design. As you might imagine, there are more complex designs like $S/A \times S/B \times C$, in which A and B are between subjects while C is within.

Exam Review

*Need to know for all subsequent noncumulative examinations as well as current examination; the final is cumulative, so anything covered in the course is “fair game”

Exam 1

Memorize or Distinguish among:

Variables and constants
 Parameter, statistic
 range, standard deviation, variance
 normal distribution characteristics
 sample vs sampling distribution
 Why $N-1$ for sample standard deviation
 Why expected value of F is 1 under the null*
 Determinants of power
 Type I and Type II errors

Population, sample
 Mean, median, mode
 skew, kurtosis, modality
 the standard normal (z) distribution
 the Central Limit Theorem
 null and alternative hypotheses*
 similarities and differences among the tests subsequent to ANOVA
 independent and dependent variables*
 assumptions of b/n groups ANOVA
 determine which procedure is appropriate given described study*

Know how to:

Recognize formulas*
 Compute degrees of freedom*
 Find percentile rank
 Use statistical tables (e.g., z , F)*
 State the null and alternative hypotheses*
 determine orthogonality
 “split up” variance in the various designs*
 compute chi-square tests

compute 1-way b/n groups ANOVA using heuristic formula and raw score formulae
 conduct orthogonal, scheffe, and bonferroni comparisons
 conduct tukey and neuman-keuls tests
 compute standard deviation and variance for populations and samples
 find confidence interval around the mean
 Decide whether to retain or reject the null hypothesis*
 find N required for $x\%$ power in one-way ANOVA

Exam 2

Memorize or Distinguish among:

problems with unequal n , how to deal with
 fixed effect vs random effects independent variables
 assumptions of repeated measures ANOVA
 Box's Epsilon
 2-way interaction

why Ss are an IV in repeated measures designs
 between vs within independent variables
 problems with difference scores
 appropriate error terms in the repeated measures designs
 main effect
 simple effect

Know how to:

interpret graphs*
 recognize simple and main effects*
 use Epsilon to adjust df

recognize 2-way interactions
 describe designs in terms of levels, type of IVs*

Exam 3

Memorize or Distinguish among:

assumptions of pearson product moment correlation
what confidence intervals around rho means
r as extent to which x and y have same absolute value percentile rank
scatter plot
equation for multiple r

z-test for independent r s
what r means
regression line, slope, intercept
standard vs hierarchical MR
what R^2 means

Know how to:

Find a confidence interval around rho
compute pearson r, slope, intercept, regression line
test significance of unstandardized regression coefficients
F-test for multiple
interpret output of multiple regression analyses f

interpret and test the significance of a r and R
use a simple regression line
find standard error of estimate
F test for change in R-squared
ind N required for x% power in simple regression

FormulasThe Mean: $\sum x/N$ Population Variance: $\frac{\sum x^2 - (\sum x)^2/N}{N}$ z-score = $(x-\mu)/\sigma$ t-test for sample: $(M-\mu)/s/\sqrt{N}$
vs. populationvariance
of the mean: σ^2/N standard error
the mean: s/\sqrt{N} linear contrast: $L=\sum aT$ SSContrast= $\frac{L^2}{n\sum a^2}$ Power: $1-\beta$ $\phi' = \sqrt{[SSb/N]/[MSw]}$ $\phi^2/\phi'^2 = n$ per group β for ncF $\phi = \phi' \sqrt{n}$ CI around population mean: $M \pm t(s/\sqrt{N})$ control for family-
wise error rate: α/k Tukey/Newman-Keuls: $(M_1 - M_2)/\sqrt{MS_{Within}/n}$ Chi-square: $\sum(O-E)^2/E$ Heuristic F= $\frac{nS_x^2}{\sum S^2/g}$ GrandTotal = $\sum x$ CF = G^2/N

One-way between groups

 $\sum x^2$ - CF
 $\sum T^2/n$ - CF
SSTotal-SSBetween

Two-way between groups

 $\sum x^2$ - CF
 $\sum \text{Cells}^2/n$ - CF
 $\sum R^2/cn$ - CF
 $\sum C^2/rn$ - CF
BetweenCells-Row-Column
Total-BetweenCellsNOTE: M refers to the sample mean

SxA One-way Repeated Measures:

$$\text{Total} = \sum x^2 - CF$$

$$\sum S^2/a - CF$$

$$\text{Within} = \text{Total} - \text{Between}$$

$$\sum A^2/s - CF$$

$$sxa (\text{error}) = \text{Within} - A$$

S/A x B: Two-way mixed design

$$\sum x^2 - (\sum x)^2/N$$

$$\sum S^2/b - (\sum x)^2/N$$

$$\sum A^2/bn - (\sum x)^2/N$$

$$SS_{\text{Between}} - SSA$$

$$SST_{\text{Total}} - SS_{\text{Between}}$$

$$\sum B^2/an - (\sum x)^2/N$$

$$\sum (AB)^2/n - (\sum x)^2/N - SSB - SSA$$

$$SS_{\text{Within}} - SSB - SSAxB$$

$$\text{intercept: } (\sum y/N) - b(\sum x/N)$$

$$r = \frac{N\sum xy - \sum x\sum y}{\sqrt{[N\sum x^2 - (\sum x)^2][N\sum y^2 - (\sum y)^2]}}$$

$$\frac{R^2/k}{[(1-R^2)/(N-k-1)]}$$

z test for

$$\text{independent } r_s: (Z_1' - Z_2') / \sqrt{[1/(N_1-3)] + [1/(N_2-3)]}$$

$$4/r^2$$

$$t = b/s_e$$

$$\frac{\Delta R^2 / k_\Delta}{[(1-R_0^2)/(N-k_0-1)]}$$

$$\text{regression line: } y' = a + bx$$

$$\text{slope: } \frac{N\sum xy - \sum x\sum y}{[N\sum x^2 - (\sum x)^2]}$$

$$\text{slope: } r(S_y/S_x)$$

$$r^2 / [(1-r^2)/(N-2)]$$

$$\text{CI around rho: } Z' \pm z(1/\sqrt{N-3})$$