

Laboratory for Repeated Measures ANOVA

[My notes for Ch 16 show how to analyze a single-factor repeated measures design using the example from K&W p. 355. Thus, your notes provide a prototype for the analyses discussed below using SPSS.]

1. First of all, let's look at the notion of error in a repeated measures design. As you know, it is the interaction of Treatment and Subject. For the data below, you should be able to make a prediction as to the size of the treatment effect, the size of the individual difference (subject) effect, and the size of the error term. Can you tell me where you'd look for each of those effects?

	a ₁	a ₂	a ₃
p ₁	1	3	5
p ₂	2	4	6
p ₃	3	5	7
p ₄	4	6	8
p ₅	5	7	9
p ₆	6	8	10

Here is the ANOVA for the above data. As you can see, SPSS blows up (as would any self-respecting ANOVA program) because it's being asked to divide by zero. Note that you do have *SS* for Subject and Treatment. Can you articulate the source of those *SS*'s?

Tests of Within-Subjects Effects

Measure: MEASURE_1

Source		Type III Sum of Squares	df	Mean Square	F	Sig.
A	Sphericity Assumed	48.000	2	24.000	.	.
	Greenhouse-Geisser	48.000
	Huynh-Feldt	48.000
	Lower-bound	48.000	1.000	48.000	.	.
Error(A)	Sphericity Assumed	.000	10	.000		
	Greenhouse-Geisser	.000	.	.		
	Huynh-Feldt	.000	.	.		
	Lower-bound	.000	5.000	.000		

Now, let's consider two different data sets. Look at the two data sets below and predict what will change in any ANOVA computed on the data. (Don't peek ahead!)

a ₁	a ₂	a ₃		a ₁	a ₂	a ₃
1	5	3		1	8	5
2	4	6		2	7	6
3	7	5		3	6	7
4	6	8		4	5	8
5	9	7		5	4	9
6	8	10		6	3	10

Tests of Within-Subjects Effects

Measure: MEASURE_1

Source		Type III Sum of Squares	df	Mean Square	F	Sig.
A	Sphericity Assumed	36.000	2	18.000	15.000	.001
	Greenhouse-Geisser	36.000	1.000	36.000	15.000	.012
	Huynh-Feldt	36.000	1.000	36.000	15.000	.012
	Lower-bound	36.000	1.000	36.000	15.000	.012
Error(A)	Sphericity Assumed	12.000	10	1.200		
	Greenhouse-Geisser	12.000	5.000	2.400		
	Huynh-Feldt	12.000	5.000	2.400		
	Lower-bound	12.000	5.000	2.400		

Tests of Within-Subjects Effects

Measure: MEASURE_1

Source		Type III Sum of Squares	df	Mean Square	F	Sig.
A	Sphericity Assumed	48.000	2	24.000	5.143	.029
	Greenhouse-Geisser	48.000	1.000	48.000	5.143	.073
	Huynh-Feldt	48.000	1.000	48.000	5.143	.073
	Lower-bound	48.000	1.000	48.000	5.143	.073
Error(A)	Sphericity Assumed	46.667	10	4.667		
	Greenhouse-Geisser	46.667	5.000	9.333		
	Huynh-Feldt	46.667	5.000	9.333		
	Lower-bound	46.667	5.000	9.333		

Can you tell which of these ANOVA's goes with which of the two preceding data sets?

Note that the source table in SPSS reports a number of different F s, etc. The topmost F represents an uncorrected F (sphericity assumed). The next two rows produce corrected estimates of probability using either the Greenhouse-Geisser correction or the Huynh-Feldt procedure. When there is a lack of sphericity, the significance-values will be higher than the uncorrected significance-value. When there is sphericity, the corrected significance-values will be similar to the uncorrected significance-value. Note, for example, that in the lower ANOVA, you would have a significant effect with the uncorrected significance-value (Sig. = .029) and an insignificant effect with the corrected significance-value (Sig. = .073).

2. For the next part of this laboratory, you'll be working with three data sets. Here's how I created these data sets:

- a) to model order effects, $p_1 = +0$, $p_2 = +4$, $p_3 = +8$, and $p_4 = +10$ (i.e., practice effects)
- b) to model treatment effects, $a_1 = +0$, $a_2 = +2$, $a_3 = +4$, $a_4 = +6$ (i.e., increasing treatment effect)
- c) to model individual differences, $s_1 = +10$, $s_2 = +5$, $s_3 = +8$, $s_4 = +2$, etc. (i.e., variable effects)
- d) to model random effects, I went through and randomly added +0, +1, +2, and +3 to scores

So that you can see these effects separately, I've created one data set with no individual differences or random effects, one data set with individual differences but no random effects, and one data set with individual differences and random effects.

a. First, compute the repeated measures ANOVA on *RM.Exercise.No.Ind.Diff*. Can you tell from the analysis that no individual differences are present?

SOURCE	SS	df	MS	F
Treatment				
Error				

b. Now compute a repeated measures ANOVA on *RM.Exercise.Ind.Diff*. Can you tell that individual differences are present? Has the error term changed? Can you explain?

SOURCE	SS	df	MS	F
Treatment				
Error				

c. As K&W point out in the text, the introduction of counterbalancing leads to variability within the data (ultimately lowering the *F*-ratio by increasing the error term). However, it is possible to estimate the extent to which there is an effect of counterbalancing (position) by analyzing for the effects of position. All that you need to do is to re-enter your data so that they are lined up by position rather than by treatment. (I've provided the position information in *RM.Exercise.Ind.Diff* so that you don't need to input the data, but you should be able to do so yourself when needed.) Compute a position effect ANOVA on these data.

SOURCE	SS	df	MS	F
Position				
Error				

Note that if you remove the effects of position from your original ANOVA, you'll be left with no error. That is, if you subtracted the position effect from the original error term, there would be no residual error variability. That's because I've created the data set to include only three sources of variability (a. order effects, b. treatment effects, and c. individual differences). As a result the only *SS* present in the data are due to Subject, Treatment, and Position.

d. Now compute the repeated measures ANOVA on *RM.Exercise.All*. How does this ANOVA differ from the prior ANOVA?

SOURCE	SS	df	MS	F
Treatment				
Error				

e. Compute the position effects and then remove them from the error term in this ANOVA. Can you see what's different now that you've removed these position effects?

SOURCE	SS	df	MS	F
Treatment				
Error				
Position				
Residual Error				

f. Comparisons for repeated measures designs are both easier and more difficult. They are easier because they are always computed in the same fashion and don't require a different error term (always using the error term that's found in the comparison). They are more difficult because you can't just compute a critical mean difference and use it to detect differences. Instead, you must compute a separate ANOVA for each comparison of interest. For the last data set (with random effects) compute a simple comparison for a_2 vs. a_4 , and a_3 vs. a_4 .

Comparison	$MS_{Comparison}$	MS_{Error}	$F_{Comparison}$	$F_{Critical}$
a_2 vs. a_4				
a_3 vs. a_4				

g. Complex comparisons get a bit tricky. The logic is the same as that for an independent groups design. That is, you need to combine information before computing the ANOVA. However, for repeated measures designs, the way that you combine information is by taking means, which involves using Transform->Compute. Compute a complex comparison for $(a_1 + a_3)$ vs. a_4 .

Comparison	$MS_{Comparison}$	MS_{Error}	$F_{Comparison}$	$F_{Critical}$
$(a_1 + a_3)$ vs. a_4				