Simple Repeated Measures Primer

With regular analysis of variance, remember that we partition the total variance into two quantities, a between-treatment variation quantity and a within-treatment variation quantity (sometimes referred to as ERROR -- see below):

Regular Oneway ANOVA

Total Variation (N - 1)

Variation Between Treatments (k - 1)

Variation Within Treatments (ERROR) (N - k)

Within Treatment variation, or ERROR, is really just how much subjects vary about the means in their respective groups. The "closer" subjects are to their respective group means, the less error. Error consists of two entities; simple measurement error and error associated with how far our subjects vary from their own group mean. It is this last quantity that repeated measures is concerned with.

In repeated measures, we can set the error associated with SUBJECT DIFFERENCES aside. In essence, we can partition individual subject variability from the overall model, therefore reducing ERROR and increasing the sensitivity of our hypothesis test. Below we see the variance partitioning for a repeated measures ANOVA (with degrees of freedom as well).

Repeated Measures Oneway ANOVA

Total Variation (N - 1)   N = #obs; k = #levels of IV; s = #subs

Variation Between Subjects (s - 1)

Variation Within s(k-1)

Treatment Differences (k - 1)

*Subject * Treatment Interaction (ERROR)* (s - 1)*(k - 1)

Note that the *subject*treatment interaction* reflects measurement error and the possibility that some subjects will react differently to different treatments.

How is subject variability removed? In simple terms, we would need to get multiple measurements of our subjects. This would allow us to see which subjects naturally score high, middle, or low on our measurements. Also, since we are using the same subjects across multiple measurements, we get inherent correlations across measures which we need to remove, and this will be removed if we isolate the subject effect.

Let's look at some data from Howell (1999). Say we did a study where we wanted to increase the amount of HIV knowledge (e.g., how does someone contract AIDS, available drugs, etc.) in a group of subjects. We get access to 4 students who are taking an AIDS Awareness class to participate in the study, and decide to measure them once a week for three (3) weeks. Hence, we will have a baseline measurement (T1), and measurements at weeks two and three (T2, T3). Given that these students are in an AIDS Awareness class, it is expected that they will become more knowledgeable about HIV-risk factors over time (so there should be a linear trend effect over time AFTER baseline). Data are reflected below:

Subject	T1	T2	T3	TOTAL
1	2	4	7	13
2	10	12	13	35
3	23	29	30	34
4	30	31	34	95
Totals	65	76	84	225

Notice first that the COLUMN totals increase over time (65 to 76 to 84). That suggests the AIDS Awareness class may be having some effect (because knowledge is increasing). However, look at the amount of variability in just the T1 assessment (ranges from 2 to 30). It seems rather large. The same goes for T2 and T3. The "largeness" of the variability comes from the individual differences of subjects. Subject 1 started with a low level of knowledge at T1 and has increased a bit to T3. Subject 4 started with a higher level of knowledge and has also increased a bit to T3. However, the RELATIVE increase for Subject 1 is much more substantial than for Subject 4 (Subject 1's knowledge doubled from T1 to T2, and almost doubled again from T2 to T3). It would be nice if we could somehow reduce the variability in each assessment, yet retain the relative increases for each subject. Repeated measures analysis allows us to do this.

[repeat.wpd]

What if you have times points such as above (T1 T2 T3), but also a nonrepeated IV (such as group membership)? This is sometimes called a mixed design. Let's say your data looks like this:

Group	Subject	T1	T2	T3	TOTAL
G1	1	2	4	7	13
G1	2	10	12	13	35
G2	3	23	29	30	34
G2	4	30	31	34	95
	Total	65	76	84	225

How do we partition the variance for that kind of design (referred to as a Mixed Design)? Below I represent the partitioning, which takes a slightly different form from those presented earlier. In this representation, we partition the total variance into *between-subject variation*, and *within-subject variation*. Both of these can be further partitioned into systematic and error components.

Mixed Repeated Measures ANOVA

Total Variation (k * s - 1)

Variation Between Subjects *Variation Within Subjects*

Variation Between Groups *Subject Differences* *Subject * Treatment Interaction (ERROR)*

(s - 1) *(s - 1)*(k - 1)*