May 2003: Think beyond simple ANOVA when a factor is time or dose—think ANCOVA. Case B: Factorial ANOVA (New Rule, 6.13).
A few corrections have been inserted in blue.

[*At times I encounter information that suggests a useful new rule—evidence that not all the rules have been covered in the book. I will number such new rules according to the chapter in which the rule fits best. So far I have not found rules for which I would create a new chapter, but that possibility is not excluded either, of course*.]

Introduction

This rule continues the discussion of the analysis of variance when one of the factors is time or dose. For this month we consider a factorial analysis of variance situation where one of the factors is time or dose—or can be ordered in some fashion.

Rule of Thumb

Think beyond simple ANOVA when a factor is time or dose—think ANCOVA.

Illustration

This example continues with the data of Table 1 from ROM for April. But now it turns out that the tablets have been stored in two types of containers: bottles and blister packs. This additional structure in the data is displayed in Table 1. Note that the observations have not changed.

Table 1. Active ingredient (in mgs) in aspirin tablets stored for four (T4), eight (T8), twelve (T12), sixteen (T16), twenty (T20), or twenty-four (T24) months. T0 is the value at baseline. Two kinds of packages.

Package	T0	T4	T8	T12	T16	T20	T24	Mean
Bottle	334	332	325	344	321	321	316	
	337	345	322	323	327	324	322	328.57
	345	325	342	334	325	317	319	
Blister	325	332	341	338	337	337	323	
	332	336	332	324	329	328	335	331.43
	328	334	335	325	331	330	328	
n	6	6	6	6	6	6	6	
Mean	333.5	334.0	332.8	331.3	328.3	326.2	323.8	
S.D.	7.06	6.54	8.18	8.66	5.47	7.08	6.79	

A factorial analysis of variance including interaction is carried out on these data producing Table 2.

Table 2. Factorial analysis of variance of amount of aspirin stored for varying lengths of time in two types of packages. Data from Table 1.

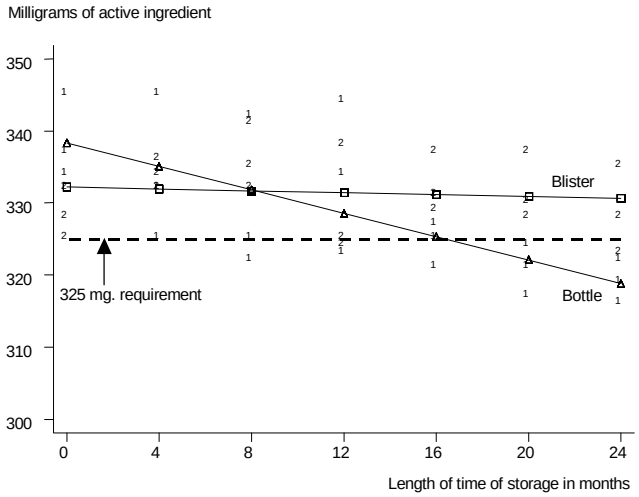
Source of variation	Degrees of freedom	Sum of Squares	Mean Square	F	Prob>F
Package (P)	1	85.71	85.714	2.12	0.156
Time (T)	6	561.33	93.556	2.31	0.061
PxT	6	584.95	97.492	2.41	0.053
Error	28	1134.00	40.500		
Total	41	2366.00			

According to this analysis there are “tantalizing” effects but nothing is significant at the o.o5 level. Again, this ignores the ordering in the time factor. We will take it into account again using an analysis of covariance. But now there are two regression patterns: one for the bottles and one for the blister pack. The linear components of the response patterns can be examined by a test for parallelism, incorporated into the analysis of covariance. This analysis is summarized in Table 3.

Table 3. Factorial analysis of covariance of data of Table 1.

Source of variation	Degrees of freedom	Sum of Squares	Mean Square	F	Prob>F
Package	1	85.71	85.71	2.12	0.156
Time	6	561.33			
Linear	1	518.01	518.01	12.79	0.0013
Remainder	5	43.32	8.66	0.21	0.96
TimexPack.	6	584.95			
Lin. x Pack.	1	375.01	375.01	9.26	0.0050
Rem. x Pack.	5	209.94	41.99	1.04	0.41
Error	28	1134.00	40.50		
Total	41	2366.00			

A graph of the data with regression lines superimposed confirms and extends the conclusions. Specifically, the material in the bottles degrades faster than the material in the blisters. The mean of the bottles crosses the 325mg line at about 16 months, suggesting that about half of the bottle-stored material will contain less than 325 mg of active ingredient by that time.



Basis of the rule

The basis of the rule is that a statistical analysis should incorporate explicit structure in data. In factorial analysis of variance with one factor involving time or dose there is the possibility of a trend, and different trends by levels of the other factors. The analysis should take this into account.

Discussion and Extensions

1. Reasons that may deter investigators from doing these trend tests may include heterogeneity of variance, anticipated non-linearity in dose-response, and unfamiliarity with these procedures and their rationale. Heterogeneity of variance can be dealt with in several ways such as transforming the data, or acknowledging the heterogeneity and working with it. Non-linearity first of all needs to be dealt with at the conceptual level. For example, in much toxicological work a dose effect is linear on a logarithmic scale, leading to a curvilinear relationship on the arithmetic scale. A determination must be made whether the logarithmic scale is a matter of computational convenience or whether it represents the biological phenomena more closely. Discomfort or unfamiliarity with these procedures must be balanced by the penalty of under-powered analyses leading to a waste of resources and the possible missing of important scientific knowledge. Given the expense of obtaining data the researcher needs only to spend a little more time and effort to come to an understanding of extracting the maximum amount of information from the data.

2. The ANCOVA and test for parallelism can be motivated as follows. First we fit separate regression lines to the bottle data and the blister data. Then we fit a single line to all of the data. Each of these three analyses leads to a sum of squares for regression. The test for parallelism involves the total of sums of squares of the separate regression and the sum of squares for the common line—the difference between these two quantities provides the means for testing whether the two slopes are equal. The sums

of squares for regression can be put into an analysis of variance table as in Table 4.

Table 4. Illustration of the interchangeability of ANCOVA and regression for the aspirin storage data in Table 1 and analyzed in Table 2.

Source of variation for slopes	**Sum of squares**	**d.f.**
SS slope for bottles	887.250	1
SS slope for blisters	5.762	1
Total SS for separate slopes	893.012	2
SS slope common slope	518.006	1
Difference SS for slopes (parallelism)	375.006	1

The difference in the SS for slopes, 375.006, is identical except for rounding error to the "Lin. x Pack." term in the analysis of covariance in Table 3.

The separate regression lines have the following equations:

Bottles,

Amount = 338.3 - 0.439*Time,

Blister packs,

Amount = 332.2 - 0.065*Time,

Combined,

Amount = 335.3 - 0.439*Time.

The last equation is identical to the one derived in last month's rule..

The error terms in the regression analyses have to take into account that there are three observations at each time point (within the package group) and thus there are two degrees of freedom per time point for a total of 14 degrees of freedom for the error term for bottles and 14 for blister packs. Thus a simple regression does not lead to the proper error term. This is one reason for preferring an analysis of covariance to just carrying out two separate regressions. Here again, as in last month's analysis, the regression approach provides the correct point estimates but it takes the equivalent of the analysis of covariance to produce the right interval estimates

3. As in last month's discussion, it needs to be emphasized that after establishing that there is a trend, or establishing that the trend patterns are not the same at levels of the other factors, it does not make sense to carry out *t*-tests to determine whether specific times are different. A difference has already been established. The interpretation of the difference should be the focus of attention. As suggested by the above example, it may be

useful to ask the question about the rate of decline of a particular preparation and the estimated time that it will reach a certain level. More sophisticated questions could be formulated in assessing when, for example, at least 5% of the preparations will fall below the required level of the active ingredient.

4. The analyses can be made more precise yet by examining trends other than linear trends. In this example there clearly is no need to examine them because the sum of squares for the residuals cannot become significant even if it were completely explained by one degree of freedom.

5. Non-linear trends in shelf-life may be particularly important. For example, a preparation declines to a certain level and then stabilizes at that level. This kind of mechanism needs biological plausibility and understanding to lead to a specific statistical model. Once you get into non-linear models there are many choices to be made—unlike linear models of which there is essentially only one.

6. Non-parametric analogues of these analyses are not easy to construct. For example, given the repeated observations at each time point it is not simple to construct a non-parametric estimate of the slope. Hollander and Wolfe [1999] contain some suggestions.

Acknowledgements
I am indebted again to Dr. Harvey Motulsky for critically and constructively reviewing earlier versions of this month's rule.

References

Hollander, M. and Wolfe, D.A. [1999]. *Nonparametric Statistics*, second edition. John Wiley and Sons, New York.

Motulsky, H. [1995]. *Intuitive Biostatistics.* Oxford University Press, New York.

Shao, J. and Chow, S-C. [1994]. Statistical inference in stability analysis. *Biometrics*, **50:** 753-763.