Longitudinal data: simple univariate methods of analysis

Danish version by Henrik Stryhn, June 1996
Department of Mathematcis and Physics, KVL
Translation (and rewriting) by Ib Skovgaard, March 1998
(the English version is the sole responsibility of the translator)

These notes describe various simple models and methods for analysis of longitudinal data and show how the analyses can be carried out using SAS. The descriptions require some knowledge of analysis of variance, also with random effects as in split-plot experiments, in particular. Knowledge of multivariate analysis is not required, and the plain multivariate analysis of variance, treating each series as a single multivariate measurement, is not accouted for here. Other relevant methods that are not included are random coefficients regression and the so-called ante-dependence methods. General references for statistical analyses of this kind of data are [1] and [2].

Below is first given a brief introduction to the problems and an example of multivariate data which will be used to illustrate the methods. Section 2 contains the description of the methods, while SAS-programs and output are collected in Section 3.

1. Introduction

A longitudinal study can be characterized by having several consequtive measurements on the same individuals, or experimental units, as opposed to investigations where only one measurement is made per individual. Also the phrase "repeated measurements" is used to describe such studies. The consequtive measurements are typically measurements taken at various time points, but time may be replaced by one-dimensional space, for example observations on a row of trees. Although the "individuals" in some applications might be experimental units of a different type, for example a sample plot in a forest, we use the word "individuals" throughout; this also mathces the example used here.

The reason that an analysis of longitudinal data requires special considerations is that measurements on the same individual cannot be considered independent, as if they were from different individuals. Thus it is sometimes simply seen that some individuals give persistently higher observations than others, but sometimes the pattern is more complicated. In any case it is wrong to use time as a factor in an analysis of variance and regard observations as independent, without effect of the individual in the model. Such an analysis may lead to gross misinterpretations of data.

The following example is a typical and relatively simple example of results from an experiment with longitudinal data.

Example 1. Growth of guinea pigs

The following data are taken from [1]. In an investigation of the effect of vitamine E on the growth of guinea pigs 15 animals were observed for 7 weeks. In week 1 they were given a growth-inhibiting substance. In the beginning of week 5 they received different amounts of vitamine E (dosage 0, 1 or 2 in appropriate units). There were 5 animals for each treatment group, and each animal was weighted at the end of week 1, 3, 4, 5, 6, and 7. Figure 1 shows the growth curves for the animals in the three groups. Let

$$\begin{aligned} Y_{ij} &= \text{weight of } i\text{th animal at } j\text{th time point} \\ \mathtt{gr}(i) &= \text{treatment group for } i\text{th animal} \\ t_j &= \text{time for } j\text{th measurement}\,, \end{aligned}$$

where $i = 1, \ldots, m = 15$ and $j = 1, \ldots, n = 6$ with $(t_1, t_2, t_3, t_4, t_5, t_6) = (1, 3, 4, 5, 6, 7)$.

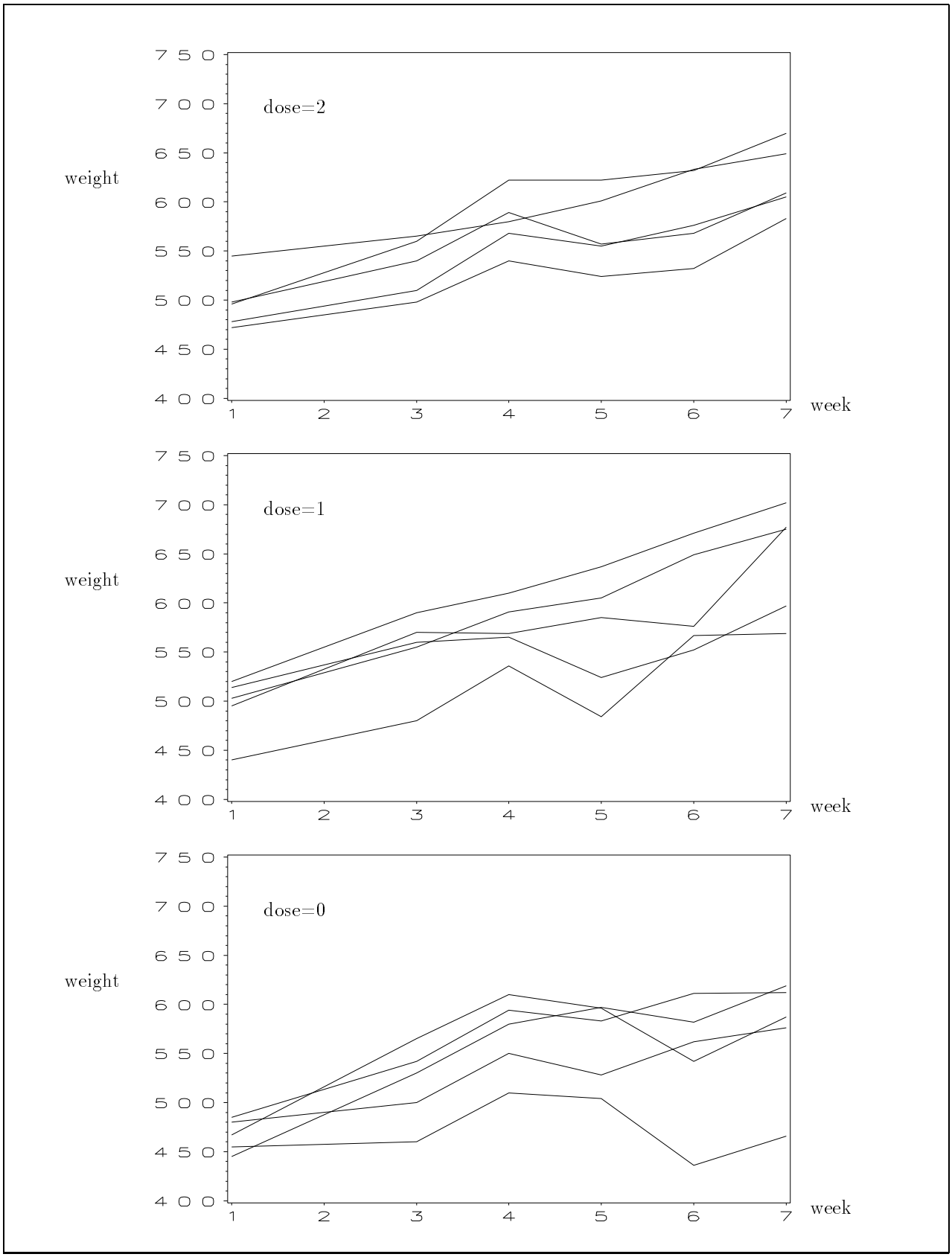

Figure 1: Weight of guinea pigs from Example 1.

Dosage	Animal	Week 1	3	4	5	6	7
0	1	455	460	510	504	436	466
	2	467	565	610	596	542	587
	3	445	530	580	597	582	619
	4	485	542	594	583	611	612
	5	480	500	550	528	562	576
1	6	514	560	565	524	552	597
	7	440	480	536	484	567	569
	8	495	570	569	585	576	677
	9	520	590	610	637	671	702
	10	503	555	591	605	649	675
2	11	496	560	622	622	632	670
	12	498	540	589	557	568	609
	13	478	510	568	555	576	605
	14	545	565	580	601	633	649
	15	472	498	540	524	532	583

Table 1: Data from Example 1: Weights of guinea pigs (g).

The structure of the design is like in a one-way analysis of variance comparing the three groups (dosages), but for each animal we have a series of measurements instead of just one. To describe the dependence between the measurements on the same animal, $(Y_1, \ldots, Y_n)$, we introduce the *covariance matrix* (or the variance-covariance matrix) $V(Y)$, an $n \times n$-matrix in which the (j,k)th element $v_{j,k}$ is defined as

$$v_{j,k} = \begin{cases} \mathrm{Var}(Y_j) & j = k \\ \mathrm{Cov}(Y_j, Y_k) & j \neq k \end{cases}.$$

The corresponding correlations are $\rho_{jk} = \mathrm{corr}(Y_j, Y_k) = \mathrm{Cov}(Y_j, Y_k)/\sqrt{v_{j,j} v_{k,k}}$, and the correlation matrix with these elements is also sometimes considered. Thus, the covariance matrix contains the variances in the main diagonal, the covariances between all pairs outside the diagonal. The simplest structure of the covariance matrix is a diagonal matrix with identical diagonal elements,

$$V(Y) = \begin{pmatrix} \sigma^2 & 0 & \cdots & 0 \\ 0 & \sigma^2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \sigma^2 \end{pmatrix}.$$

which means that $Y_1, \ldots, Y_n$ are independent with the same variance σ^2. Other types of covariance structures are needed to describe longitudinal data. In multivariate analysis of variance the model allows for any covariance matrix, so the covariance matrix is estimated in the analysis. However, this requires estimation of many parameters (all variances and covariances) and may therefore be ineffective, especially with long series of measurements. Furthermore the multivariate analysis of variance does not use the order of measurements. Other methods attempt to take advantage of this information.

2. Models and methods

Methods discussed further below are:

1) separate analyses for different time-points: weak analysis, awkward for longer series,

2) analysis of a summary statistic: good and simple method, relies on a good choice of summary measure,

3) split-plot model with individual as "mainplot": good for short series, dubious for longer series,

4) (ϵ-)correction of split-plot analysis: good for short as well as long series.

Before discussing these methods let us briefly consider the effect of ignoring that the measurements within a series are from the same individual. Thus, pretending that the weights in Example 1 were from 90 guinea pigs instead of 15, one would disregard a source of variation which is almost invariably of importance: the variation between individuals. This results in wrong estimates of other types of variation, for example the residual variance. In the present case it would lead to a comparison of the groups with the variation between measurements rather than between individuals which would be more correct. For illustration and comparison the table of analysis of variance is given as Table 2.

Source	SS	DF	MS	F	P
Groups	18548	2	9274	4.84	0.01
Weeks	142555	5	28510	14.9	<0.001
G × W	9763	10	976	0.51	>0.5
Residual	137987	72	1916		
Total	308852	89			

Table 2: Wrong analysis of data from Example 1, neglecting variation between individuals.

The method and hence the resulting significant difference between groups is incorrect.

Method 1): Separate analyses for different time-points
Suppose that in Example 1 we only use the last measurement for each animal. We would then perform a one-way analysis of variance based on the model

$$Y_{in} = \alpha(\mathtt{gr}_i) + \varepsilon_{in}\,. \tag{1}$$

This would not be wrong, but it would be inefficient because all the remaining measurements are wasted. We could then make a similar analysis for each of the other time-points, but it might then become difficult to combine the conclusions from the different analyses. Since the analysis would not tell how strong correlations there are between different measurements from the same series, we cannot tell whether a few significances at different time-points stregthen the evidence of group effects, or whether we essentially just see the same evidence several times. One might then settle for a few time points (far apart) for analysis, but should then resist the temptation of choosing the time points where there seems to be group differences. This would cause a selection effect that is difficult to take into account when interpreting the significances. The difficulty can be partly resolved by using the Bonferroni correction for performing n tests (one for each time point); this would require a P-value as low as $0.05/n$ to be interpreted as a 0.05 significance, for example.

Thus it is possible to make correct analyses time-by-time, but it is a weak analysis, because it does not use the combined information. Even more important is that this type analysis does not describe or analyse the development over time. A one-way analysis of variance as (1) has individuals as replicates and these are represented by single measurements. The pattern in Figure 1 suggest that more precise information about the individual can be obtained using the entire curve.

As an illustration of the problems arising from separate analyses group means and F-tests are shown in Table 3. No significances can be obtained even though it seems that the control group (dosage 0) has

Week j	Average $\bar{Y}_{\cdot j}$ Dosage 0	Dosage 1	Dosage 2	F (groups)
1	466	494	498	2.10
3	519	551	535	0.87
4	568	574	580	0.14
5	562	567	572	0.05
6	546	603	588	1.39
7	572	644	623	2.46

Table 3: Separate analyses of data from Example 1 for each week. F-test statistics should be compared with $F_{.95,2,12} = 3.89$ or with the Bonferroni-correction: $F_{.992,2,12} = 7.33$.

lower values than the other groups after week 5 when the treatment started.

Method 2): Analysis of a summary statistic
The analysis is performed in two steps. In the first step you choose a single quantity to calculate from each individual curve, for example the increase from first two last measurement. This results in a single measurement for each individual which is then analysed by usual methods in the second step, typically by an analysis of variance, and in Example 1 just like the analysis of a single time-point. In fact, the analysis of a single time-point can be considered a special case of a summary measure, but rarely a very good choice. The use of a summary statistic avoids the problem of dependence between several measurements from the same individual, but it is important to make good choices of summary statistics, and one will often use a figure like Figure 1 as guidance in this choice. Common choices of summary statistics are

- average over time $\bar{Y}_{i\cdot}$,
- slope, $\hat{\beta}_i$, in regression on time,
- total increase $Y_{in} - Y_{i1}$,
- area under curve,
- single time-point Y_{ij} with j fixed.

Note that although the analysis is carried out on a single summary statistic nothing prevents us from analysing several summary statistics separately. Thus, it is quite common to make separate analyses of the average and the slope, for example.

Summary statistics should be chosen to have specific interpretations related to the problem under investigation. They are not (primarily) based on statistical considerations — for example the use of a slope from a regression does not require that the curve is modelled well by a straight line. The slope can simply be used as a measure of average increase even if there is some curvature. If the curvature is very clear it might, however, suggest the curvature itself as a summary statistic.

Usually an analysis based on summary statistics should be limited to 2 or 3 summary statistics, otherwise the same type of problems arise as with analyses of each time-point separately, namely the combination of many, possibly dependent, results. The summary statistics are most usefully chosen to represent different aspects of the curve, so that the individual analyses supplement each other rather than tell the same story. For example, it is usually better to analyse average value and slope, rather than last value and slope, because a high slope usually reults in a high last value. With several summary statistics it is also possible to use a multivariate analysis of these.

Some loss of information may be involved in the reduction from the curve to the summary statistic(s). Thus, consider a regression of the measurements for each individual on time. The linear regression model can be extended by a quadratic term, a cubic term, and so on until the term of degree $n-1$, because a polynomial of degree $n-1$ may be chosen to fit the n measurements perfectly. This model for a particular individual can be written

$$\begin{aligned} Y_j &= \alpha + \beta_1 t_j + \beta_2 t_j^2 + \ldots + \beta_{n-1} t_j^{n-1} \\ &= \gamma_0 + \gamma_1 p_1(t_j) + \gamma_2 p_2(t_j) + \ldots + \gamma_{n-1} p_{n-1}(t_j)\,, \end{aligned}$$

where p_k, $k = 1, \ldots, n-1$ are so-called orthogonal polynomials of degree k. The advantage of using orthogonal polynomials are first of all that the estimate of a particular coefficient, $\hat{\gamma}_k$, does not depend on the degree of the polynomial fitted, and second that these estimates are statistically independent. The model does not change by using orthogonal polynomials, it is just rewritten in terms of other parameters. The expressions for orthogonal polynomials are somewhat awkward and the easiest way to obtain the estimate $\hat{\gamma}_k$ using statistical programs is as the estimate of the coefficient of degree k in a polynomial regression of degree k. Thus, for each k we run a polynomial regression analysis of degree k:

$$Y_j = \alpha + \beta_1 t_j + \beta_2 t_j^2 + \ldots + \beta_k t_j^k + \varepsilon_j\,.$$

and use only the coefficient of highest degree, β_k. Note, by the way, that this has to be done for each individual.

Use of orthogonal polynomials in this connection is described in [5], among others. For each individual you compute the estimates for different degrees: intercept (average), first degree (slope), second degree (curvature), etc., then you analyse each of these as a summary measure. It is easiest to interpret the coefficients of low degree, and usually at most the three mentioned are used.

In the analysis of the coefficient of each degree with respect to treatment effects, etc., there is still the problem of combining the information from several analyses. The problem is not as prominent here if only a few degrees are used. However, to see what further information is obtained by analysing a term of higher degree, [3] suggest to use the coefficents of lower degrees as covariates in each analysis. Thus, first the intercepts ($\hat{\gamma}_{0i}$) are analysed in the factorial model corresponding to the design. Then the slopes ($\hat{\gamma}_{1i}$) are analysed in the same model *adding* ($\hat{\gamma}_{0i}$) as a covariate. In this analysis the test for treatment effect is adjusted for the information already provided by the intercept coefficients ($\hat{\gamma}_{0i}$). Next step is to analyse the curvatures $\hat{\gamma}_{2i}$ using ($\hat{\gamma}_{0i}$) and ($\hat{\gamma}_{1i}$) as covariates, and so on. In this way it is sometimes seen that all treatment effect is contained in the first one or two terms, which makes it easier, for example in pairwise comparisons, to explain which differences between treatments are found. To present the different developments over time one should, however, not forget if also higher order coefficients differ — even if this is explained by the lower order coefficients.

Example 2. Growth of guinea pigs (continued)
We show two examples of the use of summary statistics for these data. Since the treatment is not applied until week 5 it seems natural to attempt to measure the treatment effect by the following variable

$$D_i = Y_{i7} - Y_{i5}\,, \quad i = 1, \ldots, m$$

and to use the "baseline" measurement Y_{i5} as a covariate:

$$D_i = \alpha(\mathtt{gr}_i) + \beta\, Y_{i5} + \varepsilon_i\,, \tag{2}$$

As seen in the analysis of variance, Table 4, the covariate is unimportant, but we can establish a treatment effect with reasonable certainty.

Source	SS	DF	MS	F	P
Groups	11287	2	5643	11.96	0.002
Initial wt. (week 5)	3	1	3	0.01	>0.5
Residual	5192	11	472		

Table 4: Analysis of variance table (partial tests) for the model (2) of weight gains from week 5 to week 7 for the data from Example 1.

The average weight gains for the treatment groups are

	Dosage 0	Dosage 1	Dosage 2
Estimate	10	77	55
$\text{LSD}_{0.95}$		29	

so the analysis establishes a positive effect of vitamine E on weight gain for the guinea pigs. In Table 5 an analysis is shown using orthogonal polynomials of degrees 0, 1 and 2. Thus, for each individual the intercept, slope and curvature are estimated, and a one-way analysis of variance is carried out for each degree. The table shows estimates from each group, corresponding LSD-values, F-tests for treatment effects from the one-way ANOVA, and in the last column the corrected F-tests using the method from [3] with lower order coefficients as covariates as described above.

	Estimate						
Variable	Dosage 0	Dosage 1	Dosage 2	F(groups)	P	$\text{LSD}_{0.95}$	F_{corr}
intercept	539	572	566	1.06	0.38	—	1.06
slope	16.0	22.6	19.6	0.83	0.46	—	0.15
curvature	-3.88	0.49	-0.40	7.45	0.008	2.61	6.05

Table 5: Analysis of coefficients from orthogonal polynomials up to degree 2 for data from Example 1.

It is seen from the table that there is a significant difference between the curvatures from the three groups, due to a negative curvature for dosage 0 (see also Figure 1). The effect is the same as observed earlier with the smaller weight gain for dosage 0 after week 5. In the present case the covariate adjustments are not of any use because there are no significant differences between intercepts and slopes.

Method 3): The split-plot model

It is possible to view an experiment with longitudinal data as a kind of split-plot experiment, with individuals as the 'main-plots' to which the treatments are applied. The 'sub-plots' are then the single measurements (or occasions) for each individual. The main-plot factor in Example 1 is the dosage applied to the guinea pig, and the sub-plot factor is time. The statistical model is

$$Y_{ij} = \alpha(\mathtt{gr}_i, j) + A_i + \varepsilon_{ij}\,, \quad A_i \sim N(0, \sigma_A^2),\ \varepsilon_{ij} \sim N(0, \sigma^2)\,, \tag{3}$$

where the term $\alpha(\mathtt{gr}_i, j)$ contains main effects of group and time as well as their interaction, and where σ_A^2 is the variation between individuals, and σ^2 is the residual variance between single measurements.

The model can be critisized for various reasons and is by now considered inadequate for most longitudinal data analyses. One critical point is that times cannot be randomized within individuals; this may also be seen as an indication that the natural order of time should be taken into account in the analysis. The consequence of the lack of randomisation is that there may be systematic, but uncontrolled, experimental variation with time, so that the variation over time cannot be reproduced in other experiments. The issue is further discussed in [4] and [5].

Another critique arises from inspection of the covariance structure in the split-plot model. For a series, $Y(i) = (Y_{i1}, \ldots, Y_{in})$, of measurements for one individual the covariance matrix is

$$V(Y(i)) = \begin{pmatrix} \sigma^2 + \sigma_A^2 & \sigma_A^2 & \cdots & \sigma_A^2 \\ \sigma_A^2 & \sigma^2 + \sigma_A^2 & \cdots & \sigma_A^2 \\ \vdots & \vdots & \vdots & \vdots \\ \sigma_A^2 & \sigma_A^2 & \cdots & \sigma^2 + \sigma_A^2 \end{pmatrix}.$$

where we have used the computation

$$\mathrm{Cov}(Y_{i1}, Y_{i2}) = \mathrm{Cov}(A_i + \varepsilon_{i1}, A_i + \varepsilon_{i2}) = \mathrm{Var}(A_i) = \sigma_A^2.$$

Thus all measurements on the same individual are positively correlated, and *all pairwise correlations are the same.* This is hardly a reasonable assumption since a pair of measurements is more likely to be highly correlated if they are close in time than if they are further apart. Thus, in the example it would mean that dependence from week 1 to week 7 is as strong as the dependence from week 6 to week 7, for example. This type of covariance structure is called 'compound symmetry'. More reasonable covariance structures are available for analysis of repeated measures, but these will not be discussed here. The procedure `PROC MIXED` in SAS has several built-in facilities for such analyses.

In practice, however, the split-plot model may give a reasonable analysis for short series ($n = 2, 3, 4$) since compound symmetry may not be that far off in such cases. Split-plot method is the simplest kind of analysis of variance which uses the full data set, and the choice of summary statistics is avoided. Furthermore it is, in fact, possible to test the hypothesis of compound symmetry to see whether the split-plot method is reasonable. This will be discussed further in the next section which also describes ways of correcting the split-plot method when it is not reasonable.

Example 3. Growth of guinea pigs (continued)
Table 6 shows the analysis of variance for the split-plot method First it is seen from the table that

Source	SS	DF	MS	F	P
Groups	18548	2	9274	1.06	0.38
Weeks	142555	5	28511	52.6	<0.001
G × U	9763	10	976	1.80	0.080
Individuals (main-plots)	105434	12	8786	16.2	<0.001
Residual	32553	60	543		
Total	308852	89			

Table 6: Split-plot analysis, model (2), for data from Example 1.

there is a considerable variation between individuals, thus confirming the importance of a method that takes this into account. Further it is seen that the interaction between groups and time is close to being significant, that (obviously) there is an overwhelmingly significant difference between weeks, and that aside from the almost significant interaction there is no evidence of difference between the groups. The test for main effect of groups is, in fact, the same as the test for difference between the intercepts in the analysis using orthogonal polynomials (Table 5). The present analysis uses time as a factor rather than as a covariate. For the covariate (regression) approach the method of summary statistics may be preferable ([4],[5]).

Method 4): ϵ-correction of the split-plot method
In light of the criticism of the split-plot method it is natural to ask when the method seems reasonable,

and what can be done if it is not. This has been investigated over some time in the statistical literature (see for example [1]) and by now reasonably clear recommendations can be made. These are summarized below. The method described here is together with its variants often refered to as 'Repeated measures ANOVA'.

It may be shown that the F-tests obtained by the split-plot method are correct if the covariance matrix satisfies a certain condition called the condition of "sphericity" of which compound symmetry is a special case. We shall not describe this condition precisely here, only note that there exists a test for the hypothesis of sphericity, the so-called Mauchly's test, and that this test is provided by `PROC GLM` in SAS. This test is not very efficient and may be quite sensitive to non-normality of the data. Thus, the test should be considered but only partly be trusted for choice of method.

In this connection it is important to note that an inappropriate application of the split-plot method may lead to false significances, which is much more dangerous than a conservative test which gives too few significances. However, there exist corrections of the F-tests which (approximately) account for the possible lack of sphericity. These are the so-called ϵ-corrections, by which a quantity, ϵ, measures the departure from sphericity. An $\epsilon = 1$ corresponds to sphericity while smaller values correspond to deviations from sphericity. The ϵ-corrected test uses *the same* F-statistic as the split-plot analysis, but replaces the corresponding degrees of freedom $(\mathrm{DF}_1, \mathrm{DF}_2)$ by $(\epsilon\mathrm{DF}_1, \epsilon\mathrm{DF}_2)$. This correction should be used for sub-plot tests, in the example these are the tests involving time, while the main-plot tests should not be corrected.

There are two ways to estimate ϵ, named Greenhouse-Geisser and Huynh-Feldt after two pairs of authors. It is, unfortunately, not easy to say which is best. There seems to be some consensus that the Greenhouse-Geisser estimate may lead to conservative tests in small samples. Both are given in SAS as well as the correspondingly corrected tests (see Section 3). If the discrepancy from sphericity is very large (ϵ near zero) the method of ϵ-correction is not to be recommended.

Let us illustrate the method on the guinea pig example.

Example 4. Growth of guinea pigs (continued)
The first step is to compute Mauchly's test for sphericity. In SAS a χ^2-square distribution with $(n-2)(n+1)/2$ degrees of freedom is used to approximate the distribution of the test statistic. In the present case the test statistic is 29.4 on 14 degrees of freedom, and the corresponding P-value is 0.0093. Thus, the test is clearly significant implying that at least some of the F-tests in Table 6 are dubious. Next, the following estimates of ϵ are given in the SAS output:

Greenhouse-Geisser:	$\hat{\epsilon} = 0.49$,
Huynh-Feldt:	$\tilde{\epsilon} = 0.72$.

This is the typical pattern, that $\hat{\epsilon} < \tilde{\epsilon}$. The two ϵ-values are used in Table 7 to correct the tests from the analysis of variance.

					P		
Source	SS	DF	MS	F	uncorr.	$\hat{\epsilon}$-corr. (G-G)	$\tilde{\epsilon}$-corr. (H-F)
Groups	18548	2	9274	1.06	0.38	–	–
Weeks	142555	5	28511	52.6	<0.001	<0.001	<0.001
G × W	9763	10	976	1.80	0.080	0.15	0.11
Individuals (main-plots)	105434	12	8786	16.2	<0.001	—	—
Residual	32553	60	543				
Total	308852	89					

Table 7: Corrected split-plot analysis for data from Example 1.

As noted above the corrections only affect tests involving time, including interactions with time. Note that the effect of the corrections is to increase the P-values. A simplified explanation of this is that when it is taken into account that neighbouring observations are highly correlated they should not count as completely new observations, hence the reduced number of degrees of freedom. In the present example this method of analysis does not catch the previously noted significant difference between the three groups after week 5; probably because this difference is blurred by the inclusion of several weeks with no differences. An obvious alternative was only to use the data from weeks 5, 6 and 7.

3. SAS-programs and output

The program below performs (almost) all the analyses described here. The procedure `PROC MIXED` has many facilities for analysis of repeated measures using various covariance structures, but some of the univariate methods described here are only available in `proc GLM`.

```
data guinea;
  infile 'guinea.dat';
  do grp=1 to 3;
    do guinea=1 to 5;
      input week1 week3-week7 @@;
      output;
    end;
  end;

proc glm;
  class grp;
  model week1--week7 = grp / ss3;
  repeated week 6 (1 3 4 5 6 7) polynomial / printe summary nom;
run;
```

Notes:

1. Data is input in a form especially suited for the `repeated`-statement in `proc GLM`. The factor(s) representing the "repetitions" (here time) is not represented by any variable in the SAS dataset, but indirectly as the series of variables, here `week1, week3-week7`.

2. In `PROC GLM` the model is specified by having the series of variables mentioned above on the left in the `MODEL` statement, and the right hand side then specifies the model for the individuals in the experiment (the main-plot part). A detail is the use of the dash between `week1` and `week7`, allowing us to skip writing all the variables.

3. In the `repeated`-statement a name is given to identify the repeated (sub-plot) factor. Thus `week` does *not* refer to a variable in the data set, but only gives a name used in the output. The number 6 following `week` indicates the number of measurements per individual and the numbers in the paranthesis are the "times". If the paranthesis was not included equidistant time points would have been used. The "times" given in this way are of importance for the method using polynomial fitting to construct summary statistics.

4. The option `summary` is necessary to produce the results from the orthogonal regression, `printe` produces Mauchly's test for sphericity, while the option `nom` (no multivariate) suppresses printing of results from multivariate analysis of variance.

Edited output from the program above using `proc GLM`:

General Linear Models Procedure
Class Level Information

Class	Levels	Values
GRP	3	1 2 3

Number of observations in data set = 15

Dependent Variable: WEEK1

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	2	2969.2000	1484.6000	2.10	0.1651
Error	12	8481.2000	706.7667		
Corrected Total	14	11450.4000			

R-Square	C.V.	Root MSE	WEEK1 Mean
0.259310	5.467932	26.585	486.2000

Source	DF	Type III SS	Mean Square	F Value	Pr > F
GRP	2	2969.2000	1484.6000	2.10	0.1651

Dependent Variable: WEEK3

...

General Linear Models Procedure
Repeated Measures Analysis of Variance
Repeated Measures Level Information

Dependent Variable	WEEK1	WEEK3	WEEK4	WEEK5	WEEK6	WEEK7
Level of WEEK	1	3	4	5	6	7

Partial Correlation Coefficients from the Error SS&CP Matrix /
Prob > |r|

...

E = Error SS&CP Matrix

WEEK.N represents the nth degree polynomial contrast for WEEK

...

Partial Correlation Coefficients from the Error SS&CP Matrix
of the Variables Defined by the Specified Transformation / Prob > |r|

...

Test for Sphericity: Mauchly's Criterion = 0.0544835
Chisquare Approximation = 29.389556 with 14 df
Prob > Chisquare = 0.0093

Tests of Hypotheses for Between Subjects Effects

Source	DF	Type III SS	Mean Square	F Value	Pr > F
GRP	2	18548.1	9274.0	1.06	0.3782
Error	12	105434.2	8786.2		

Univariate Tests of Hypotheses for Within Subject Effects

Source: WEEK

DF	Type III SS	Mean Square	F Value	Pr > F	Adj Pr > F G - G	Adj Pr > F H - F
5	142554.50000	28510.90000	52.55	0.0001	0.0001	0.0001

Source: WEEK*GRP

DF	Type III SS	Mean Square	F Value	Pr > F	Adj Pr > F G - G	Adj Pr > F H - F
10	9762.73333	976.27333	1.80	0.0801	0.1457	0.1103

Source: Error(WEEK)

DF	Type III SS	Mean Square
60	32552.60000	542.54333

Greenhouse-Geisser Epsilon = 0.4856
Huynh-Feldt Epsilon = 0.7191

Analysis of Variance of Contrast Variables

WEEK.N represents the nth degree polynomial contrast for WEEK

Contrast Variable: WEEK.1

Source	DF	Type III SS	Mean Square	F Value	Pr > F

```
MEAN                     1    131764.803   131764.803    87.35    0.0001
GRP                      2      2495.213     1247.607     0.83    0.4608
Error                   12     18100.874     1508.406

Contrast Variable: WEEK.2

Source                  DF   Type III SS  Mean Square  F Value    Pr > F
MEAN                     1   2011.47937   2011.47937     6.67    0.0240
GRP                      2   4489.67778   2244.83889     7.45    0.0079
Error                   12   3617.50952    301.45913

Contrast Variable: WEEK.3

...
```

Notes:

5. The first part of the output is a series of separate analyses for each week. Only the first is shown here. Printing of these can be suppressed by use of the option `nouni` in the `model`-statement.

6. The next part, 'Repeated Measures Level Information', summarizes the model as formulated by the `model`- and the `repeated`-statements, and gives some partial correlation coefficients and contrasts that are not shown and which will not be discussed here.

7. Mauchly's test and the analysis of variance tables 'Between Subjects' and 'Within Subjects' give the results for method (4): split-plot analysis of variance, ϵ-estimates and -corrections.

8. Finally, under the heading 'Analysis of Variance of Contrast Variables' we find the analyses of variance of the estimates form orthogonal regression discussed under method (2). The heading 'WEEK.1' refers to the slopes, 'WEEK.2' to the curvatures and so on up to degree 5. The F-tests are found in the `GRP` rows. There are no tests for the intercepts because these are the same as the tests in the analysis of variance 'Between Subjects'. Note that there are no significant differences between groups regarding coefficients of higher degree than 2.

The SAS program below, for which we do not show the output, saves the intercepts, slopes and curvatures from the orthogonal regression in a SAS dataset so that they can be used for further analysis. This is necessary to use the under method (2) mentioned correction for lower order terms by using the lower order coefficients as covariates. This analysis is only of interest as a supplement to the results from the `repeated`-analysis in `PROC GLM` when one or more of the polynomial degrees exhibit significant results.

```
data gp;
  infile 'guinea.dat';
  do grp=1 to 3;
    do guinea=1 to 5;
      do week=1,3,4,5,6,7;
        input wgt @@;
        weekquad=week*week;
        output;
      end;
    end;
  end;
proc sort;
  by grp guinea;

proc reg data=gp outest=e0 noprint;
  model wgt = ;
  by grp guinea;
data estint;
  set e0;
  keep grp guinea intercept;

proc reg data=gp outest=e1 noprint;
  model wgt = week;
  by grp guinea;
data estlin;
  set e1;
  weeklin=week;
  keep grp guinea weeklin;

proc reg data=gp outest=e2 noprint;
  model wgt = week weekquad;
  by grp guinea;
data estquad;
  set e2;
  keep grp guinea weekquad;

data estimpar;
  merge gp estint estlin estquad;
  by grp guinea;
  if week=1;
  keep grp intercep weeklin weekquad;
proc print;

proc glm;
  class grp;
  model intercep weeklin weekquad = grp / ss3;
  lsmeans grp / stderr;
proc glm;
  class grp;
```

```
  model weeklin = intercep grp / ss1 ss3 solution;
proc glm;
  class grp;
  model weekquad = intercep weeklin grp / ss1 ss3 solution;
run;
```

References

[1] Crowder, M. J. & Hand, D. J. (1990). *Analysis of Repeated Measures.* Chapman & Hall, London.

[2] Diggle, P. J., Liang, K.-Y. & Zeger, S. L. (1994). *Analysis of Longitudinal Data.* Oxford Science Publication, Oxford.

[3] Keuls, M. & Garretsen, F. (1982). Statistical analysis of growth curves in plant breeding. *Euphytica* **31**, 51–64.

[4] Wallenstein, S. (1982). Regression models for repeated measurements. *Biometrics* **38**, 849–850.

[5] Yates, F. (1982). Regression models for repeated measurements. *Biometrics* **38**, 850–853.