J. Theoret. Neurobiol. **1**, 105–119 (1981)
Printed in Australia

THE GENERAL LINEAR MODEL IN THE DESIGN AND ANALYSIS OF EVOKED RESPONSE EXPERIMENTS

DAVID R. BRILLINGER
Department of Statistics, University of California, Berkeley, CA 94720, U.S.A.

(Received April 15, 1981)

Abstract

Through the setting down of a linear time invariant model it is possible to formalize a wide variety of questions and issues that arise in the use and interpretation of evoked response experiments. In the model considered, it is assumed that the effects of individual stimuli combine in an additive time invariant fashion to provide the responses observed. The model allows several stimuli to be applied in one experiment and for multiple responses to be analyzed. It allows for complex experimental designs being employed and provides formal tests of significance as well as estimates of parameters of interest.

1. Introduction

Evoked response experiments play an essential role in quantitative biology. Because the experimenter is able to choose which stimuli to apply, and when to apply them, conclusions drawn can pass far beyond associations noted to inferences concerning causal mechanisms. The intention of this paper is to illustrate how, in many situations, the descriptions of such experiments may be formalized in a manner that the extensive subject matter of the fields of statistical inference and design of experiments may be drawn upon.

Evoked response experiments consist of the input of a succession of impulses to a system of interest and the recording of a corresponding succession of responses. The procedure extends the traditional one of system identification via pulse probing in that many pulses are applied over the duration of the experiment. The desire is the same—to obtain an understanding of how the system operates.

Evoked response experiments are used particularly by neuroscientists who take sensory stimuli for the impulses and record the corresponding brain electrical activity. (The terms evoked potential and event related potential are often used by these workers.) Questions that arise in their work include the following.

(1) Does a particular stimulus actually elicit a response?
(2) Do two different stimuli elicit the same average response?
(3) Is the average response the same when measured at two different locations?
(4) Has the average response changed after some passage of time?
(5) Does the order of application of different stimuli affect the average response?
(6) Are the individual effects of two different stimuli additive (superposable) when both are applied simultaneously?
(7) How does the strength (intensity) of a stimulus affect the average response?
(8) Do the average responses of each of a group of individuals relate to some other characteristics of those individuals?

It will be seen that each of the above questions is analagous to a hypothesis concerning the parameters of the traditional linear model of statistical inference. In consequence the neuroscientist can in fact draw on the extensive bank of knowledge that statisticians have developed, concerning the linear model, in addressing questions such as those above.

Workers in the evoked response field express desires to:

(*a*) collect data quickly and efficiently;
(*b*) examine for nonlinearities;
(*c*) formalize variability between and within subjects;
(*d*) have precise estimates and indications of variability.

Difficulties they have include:

(*a*) small response and large noise;
(*b*) variability in response;
(*c*) superposed effects;
(*d*) artifacts.

It will be seen that each of these issues may be addressed in a formal fashion within the context of a statistical model presented.

The traditional evoked response experiment and analysis proceeds by the application of a single stimulus periodically, at well spaced time points, followed by the averaging together of the EEG values that occur at the same time lag after the application of the stimulus. Specifically, if $Y(t)$ is the value of the response series at time t and if the stimulus is applied at the times $j\sigma$, $j=1, \ldots, M$ one computes

$$\bar{Y}(u) = \frac{1}{M}\sum_{j=1}^{M} Y(u + j\sigma) \tag{1}$$

for $0 \leqslant u \leqslant U$. This statistic is typically referred to as the average evoked response (AER). The interval width σ is to be taken large enough that neighbouring responses evoked do not interfere with each other. This desire for a large value of σ conflicts directly with the accompanying desires for efficiency of estimation and speed in collection of data. Among other things, this paper will discuss how, by

applying the stimulus at irregularly spaced times σ_j one may estimate an expected evoked response more rapidly and efficiently in certain circumstances.

In order to make decisions and test hypotheses after data have been collected one needs some indication of the level of sampling fluctuations of the statistics computed. A usual method of providing such is the $(\pm)$ reference of Schimmel (1967) wherein differences of successive pairs of evoked responses are computed. In many circumstances this method is very effective. However, it is inefficient and incapable of dealing with more complicated experimental designs (e.g. when the actual stimulus applied varies irregularly). An advantage of setting down a formal model for the situation, as is done in this paper, is that the computation of estimates of variability (even for complex statistics) becomes an algebraic exercise.

The paper begins with a description of the standard linear regression (or Gauss–Markov) model of mathematical statistics and of its uses and properties. The paper then indicates an extension of that model appropriate to the evoked response situation.

General references to the use and interpretation of evoked response experiments include Donchin and Lindsley (1969), Shagass (1972), John (1977), Thatcher and John (1977), Callaway *et al.* (1978). References to the statistical analysis of such experiments are Glaser and Ruchkin (1976), Freeman (1980) and Brillinger (1981). General references to the linear model of statistics and to the design of experiments include Cox (1958), Rao (1965) and Seber (1977).

2. The Gauss–Markov model

Consider an $s \times n$ random matrix $\boldsymbol{\epsilon}$ with $E\boldsymbol{\epsilon} = \mathbf{0}$ and $\operatorname{var}\boldsymbol{\epsilon} = \boldsymbol{\Sigma} \otimes \mathbf{I}_n$. (Here $\otimes$ denotes the Kronecker product of matrices, E denotes expectation and var denotes the covariance matrix of the column vector resulting from stacking the successive columns of $\boldsymbol{\epsilon}$ one under the other.) Suppose that the fixed $r \times n$ (design) matrix $\mathbf{X}$ is known and that the $s \times n$ matrix $\mathbf{Y}$ is known and given by

$$\mathbf{Y} = \boldsymbol{\beta}\mathbf{X} + \boldsymbol{\epsilon}. \tag{2}$$

This structure is the one of Gauss–Markov. The known results concerning it include:

(*a*) the least squares estimate of $\boldsymbol{\beta}$ is $\hat{\boldsymbol{\beta}} = \mathbf{Y}\mathbf{X}'(\mathbf{X}\mathbf{X}')^{-1}$;
(*b*) the unbiased estimate of $\boldsymbol{\Sigma}$ is $\hat{\boldsymbol{\Sigma}} = (n-r)^{-1}\mathbf{Y}(\mathbf{I} - \mathbf{X}'(\mathbf{X}\mathbf{X}')^{-1}\mathbf{X})\mathbf{Y}'$;
(*c*) $\operatorname{var}\hat{\boldsymbol{\beta}} = \boldsymbol{\Sigma} \otimes (\mathbf{X}\mathbf{X}')^{-1}$;
(*d*) the best linear unbiased estimate (BLUE) of $\boldsymbol{\Gamma} = \mathbf{A}\boldsymbol{\beta}\mathbf{B}$ is $\hat{\boldsymbol{\Gamma}} = \mathbf{A}\hat{\boldsymbol{\beta}}\mathbf{B}$;
(*e*) if $\boldsymbol{\epsilon}$ has a multivariate normal distribution so does $\hat{\boldsymbol{\beta}}$;
(*f*) confidence regions for the entries of $\boldsymbol{\beta}$ may be constructed directly;
(*g*) significance tests can be set up directly for hypotheses of the form $\mathbf{A}\boldsymbol{\beta}\mathbf{B} = \mathbf{C}$. The tests are based on $\hat{\boldsymbol{\Sigma}}$ and $\mathbf{S} = (\mathbf{A}\hat{\boldsymbol{\beta}}\mathbf{B} - \mathbf{C})(\mathbf{B}'(\mathbf{X}\mathbf{X}')^{-1}\mathbf{B})^{-1}(\mathbf{A}\hat{\boldsymbol{\beta}}\mathbf{B} - \mathbf{C})'$ and are F tests in the case that $\mathbf{B}$ and $\mathbf{C}$ have one column.

Further the adequacy of the model may be checked by examination of the residual values $\hat{\boldsymbol{\epsilon}} = \mathbf{Y} - \hat{\boldsymbol{\beta}}\mathbf{X}$ or by bringing further values into the $\mathbf{X}$ matrix corresponding to possibly relevant variates.

The above results are standard in the multivariate analysis of variance and may be found in many places; for example, Rao (1965), Roy *et al.* (1971), Brillinger (1975), Seber (1977), Timm (1980). The case with $s = 1$ is often referred to as the linear regression model. In its case the model may be written

$$Y_j = \sum_{i=1}^{r} \beta_i X_{ij} + \varepsilon_j \tag{3}$$

$j = 1, \ldots, n$.

In experimental design situations the entries of the matrix $\mathbf{X}$ typically take on the values 0 or 1. For example in the one-way classification, corresponding to a single factor with A levels, one writes

$$Y_{ak} = \mu_a + \varepsilon_{ak} \tag{4}$$

$k = 1, \ldots, n_a$; $a = 1, \ldots, A$. (Here $n = n_1 + \ldots + n_A$.) This may be brought into the form (3) by defining indicator variables $X_{ij} = 1$ if unit j has the factor at level i and $X_{ij} = 0$ otherwise. The hypothesis of interest is often $\mu_1 = \mu_2 = \ldots = \mu_A$; that is, that the factor (or treatments) has no effect.

In the two-way classification, with two factors having A, B levels respectively, one writes

$$Y_{abk} = \mu + \alpha_a + \beta_b + \gamma_{ab} + \varepsilon_{abk} \tag{5}$$

and introduces the constraints $\sum_a \alpha_a, \sum_b \beta_b, \sum_a \gamma_{ab}, \sum_b \gamma_{ab} = 0$ to make the parameters well defined. The model (5) is linear in the parameters and so may be written in the form (3) through the definition of certain indicator variables. The γ_{ab} terms are called interactions. Hypotheses of interest include: all the interactions are 0 and, the interactions are 0 and so are the β_b.

One especially important design is the randomized block. Here there are B blocks of A, approximately, homogeneous units. A factor of A levels is applied at each level within each block, levels being assigned to units randomly. One model for this experiment takes the form (5) with no interactions, namely

$$Y_{ab} = \mu + \alpha_a + \beta_b + \varepsilon_{ab} \tag{6}$$

$a = 1, \ldots, A$; $b = 1, \ldots, B$. The α_a are here called treatment effects and the β_b, block effects. Generally the block effects are non-zero and the hypothesis of interest is: $\alpha_1, \ldots, \alpha_A = 0$. This model and hypothesis may be dealt with as above; however, the randomization employed allows another, broader, analysis. The randomization, (i) provides a protection against systematic biases due to extraneous factors that might enter, (ii) allows experimental error to be estimated and (iii) leads to exact significance tests based on the randomization introduced rather than some assumption of normality. Randomization changes a study from the status of an observational one to that of an experiment. Nonrandomized studies often have no inferential value.

The testing and variance estimation involved in models, such as those above is, conveniently carried through by means of analysis of variance tables. These tables, the above designs and issues and many related matters are discussed in detail in the books by Kempthorne (1952) and Cochran and Cox (1957). In this paper the ideas of such designs and their uses in inference will be discussed in the particular context of evoked response experiments.

3. A linear model for evoked responses

In order to proceed further it is necessary to set down some notation and assumptions specific to the evoked response situation. In general it will be assumed that the experiment is not evolving in time, that functional transformations involved are time invariant and that the noise processes present are stationary stochastic processes.

To begin, consider the case of a single stimulus being applied with constant intensity at times σ_j, $j = 0, \pm 1, \pm 2, \ldots$. It is convenient to view this sequence of times as a point process, $M(\cdot)$, with $M(t)$ denoting the number of times the stimulus was applied in the time interval $(0, t]$ (that is the number of σ_j satisfying $0 < \sigma_j \leqslant t$). (Point processes are discussed in the book by Cox and Lewis (1966), for example.) Let the response at time t be denoted by $Y(t)$. Suppose that the time period of observation is $(0, T]$. Denote the total number of times the stimulus was applied by $M = M(T)$. Then the average evoked response may be written

$$\overline{Y}(u) = \frac{1}{M} \sum_{j=1}^{M} Y(u + \sigma_j) = \frac{1}{M} \int_0^{T-u} Y(u + t)\, dM(t). \tag{7}$$

An examination of expression (7) indicates that the AER may be viewed as an estimate of

$$E\{Y(u + t) \mid \text{stimulus at time } t\}. \tag{8}$$

As such it is not a system invariant. It has the distribution of the σ_j confounded therein. In practice one is interested in aspects of the system, not the input. Suppose one defines

$$E\{Y(u + t) \mid \text{single stimulus at time } t\} = \mu + a(u). \tag{9}$$

That is, (9) provides the expected response of an experiment in which the stimulus is applied just once, at time t. This function $a(\cdot)$ might be viewed as a system invariant.

Suppose further that the situation is causal, and so $a(u) = 0$ for $u < 0$, and that $a(u)$ is negligible for sufficiently large u. If the stimulus is applied at times that are sufficiently far apart that the individual effects do not overlap, then from (9)

(10) $$E\{Y(t) \mid M(u), u \leqslant t\} = \mu + \sum_{\sigma_j \leqslant t} a(t - \sigma_j)$$
$$= \mu + \int a(t-u)dM(u).$$

Expression (10) suggests the model on which much of the analysis to come will be based, namely,

(11) $$Y(t) = \mu + \int a(t-u)dM(u) + \varepsilon(t)$$

with $M(\cdot)$ denoting the input, $Y(\cdot)$ the output, $a(\cdot)$ the system response and $\varepsilon(\cdot)$ the noise. In the case that the function $a(\cdot)$ vanishes outside the interval $[0, V]$ and that the stimulus is applied at times farther than V time units apart, one has from (11)

(12) $$\bar{Y}(u) = \mu + a(u) + \bar{\varepsilon}(u)$$

for $0 \leqslant u \leqslant V$ and the function estimated by the AER is $a(\cdot)$. Experiments have been carried out (see for example Bieddenbach and Freeman, 1965) indicating that the model (11) is reasonable in evoked response experiments even in situations where the stimulus is applied at times so close together that the individual responses do overlap—provided that the minimal interstimulus interval is not exceedingly small. It may be remarked that even when the system is non-linear, the linear model (11) may prove a useful approximation.

In the case that a single stimulus of constant intensity is applied and that a single response series is recorded, the above discussion suggests consideration of the stochastic model,

(13) $$Y(t) = \mu + \int a(t-u)dM(u) + \varepsilon(t)$$

$-\infty < t < \infty$, with $\varepsilon(\cdot)$ a noise series. The hypothesis that the stimulus has no effect may now be formalized as: is $a(\cdot) \equiv 0$? Through assumptions concerning the noise series, and using statistical techniques, one can proceed to construct efficient estimates $\hat{a}(\cdot)$ of the parameter $a(\cdot)$, to the computation of measures of the sampling fluctuations of $\hat{a}(\cdot)$ and to the construction of estimates of $a(\cdot)$ that are resistant to artifacts.

Rather than investigating the model (13) in real detail (it is considered in Brillinger (1978)), it seems appropriate to proceed to the general case of stimuli of several types and of possibly varying intensities. Suppose that there are stimuli of r types. Set $dM_i(t) =$ the relative intensity with which the i-th stimulus is applied at time t. If the i-th stimulus is not applied at time t, then $dM_i(t) = 0$. In particular, if the i-th stimulus has constant intensity throughout the experiment set

(14) $$dM_i(t) = \begin{cases} 1, & \text{if the } i\text{-th stimulus is applied at time } t, \\ 0, & \text{otherwise.} \end{cases}$$

Collect the $M_i(\cdot)$ into the column vector $\mathbf{M}(t) = [M_i(t)]$. Next, suppose that response series are observed concurrently at s different sites. Let $Y_j(\cdot)$ denote the series observed at the jth site and set up the column vector $\mathbf{Y}(t) = [Y_j(t)]$.

Model I

The response series relate to the stimuli applied via

(15) $$\mathbf{Y}(t) = \boldsymbol{\mu} + \int \mathbf{a}(t-u)d\mathbf{M}(u) + \boldsymbol{\epsilon}(t)$$

where $\boldsymbol{\mu}$ is an s-vector, where $\mathbf{a}(\cdot)$ is an $s \times r$ matrix-valued function and where $\boldsymbol{\epsilon}(\cdot)$ is an s-vector-valued stationary noise process having mean $\mathbf{0}$ and spectral density matrix $\mathbf{f}(\lambda)$.

The entry $a_{ji}(t)$, of the matrix $\mathbf{a}(t)$, provides the effect on the response observed at the site j of applying the i-th stimulus at time 0 with intensity 1. (If the stimulus was applied then with intensity I_j, the effect would be $I_j a_{ji}(t)$.) A variety of hypotheses of interest may now be formalized in terms of the matrix $\mathbf{a}(\cdot)$: for example, (i) are stimuli i, i' having the same effects?—is $a_{ji}(\cdot) \equiv a_{ji'}(\cdot), j = 1, \ldots, s$?; (ii) are the average responses at symmetrically placed sites the same?—is $a_{ji}(\cdot) \equiv a_{j'i}(\cdot)$ for j, j' corresponding to symmetrically placed sites?; (iii) does the ith stimulus have any effect?—is $a_{ji}(\cdot) \equiv 0, j = 1, \ldots, s$? In order to test such hypotheses it is first necessary to estimate the parameter $\mathbf{a}(\cdot)$ of (15).

Given data $\mathbf{Y}(t)$, $\mathbf{M}(t), 0 \leqslant t \leqslant T$, with T sufficiently large, the parameters of Model I may be estimated in a reasonable fashion. It turns out to be easier to proceed in the frequency domain. To this end define

(16) $$\mathbf{A}(\lambda) = \int \exp(-i\lambda t)\mathbf{a}(t)\, dt$$

(17) $$\mathbf{d}_Y^T(\lambda) = \int_0^T \exp(-i\lambda t)\mathbf{Y}(t)dt$$

(18) $$\mathbf{d}_M^T(\lambda) = \int_0^T \exp(-i\lambda t)d\mathbf{M}(t).$$

With these definitions the relationship (15) yields the relationships

(19) $$\mathbf{d}_Y^T\left(\frac{2\pi k}{T}\right) \doteq \mathbf{A}(\lambda)\mathbf{d}_M^T\left(\frac{2\pi k}{T}\right) + \mathbf{d}_\varepsilon^T\left(\frac{2\pi k}{T}\right)$$

for integers k, not equal to 0, such that $2\pi k/T$ is near λ. Now the central limit theorem for the finite Fourier transform of stationary mixing processes suggests that the distribution of the values $\mathbf{d}_\varepsilon^T(2\pi k/T)$ may be approximated by that of independent complex normals with mean $\mathbf{0}$ and covariance matrix $2\pi T\mathbf{f}(\lambda)$. (See, for example, Brillinger (1974) for such central limit theorems.) Suppose that there are n such distinct frequencies $2\pi k/T$ and set up the $s \times n$, $r \times n$ and $s \times n$ matrices.

$$\text{(20)} \qquad \mathbf{Y} = \left[\mathbf{d}_Y^T\left(\frac{2\pi k}{T}\right)\right], \qquad \mathbf{X} = \left[\mathbf{d}_M^T\left(\frac{2\pi k}{T}\right)\right], \qquad \boldsymbol{\epsilon} = \left[\mathbf{d}_\varepsilon^T\left(\frac{2\pi k}{T}\right)\right].$$

The relationship (19) for the given k's is seen to have the approximate form of the relationship (2) with $\boldsymbol{\beta} = \mathbf{A}(\lambda)$, except that in (20) the entries are complex-valued.

The results of Section 2 now suggest the estimation of the transfer function $\mathbf{A}(\lambda)$ by

$$\text{(21)} \qquad \hat{\mathbf{A}}(\lambda) = \mathbf{Y}\overline{\mathbf{X}}'(\mathbf{X}\overline{\mathbf{X}}')^{-1}$$
$$= \left(\sum_k \mathbf{d}_Y^T\left(\frac{2\pi k}{T}\right)\overline{\mathbf{d}_M^T\left(\frac{2\pi k}{T}\right)}'\right)\left(\sum_k \mathbf{d}_M^T\left(\frac{2\pi k}{T}\right)\overline{\mathbf{d}_M^T\left(\frac{2\pi k}{T}\right)}'\right)^{-1}$$

with the sums in (21) over n distinct frequencies $2\pi k/T$ near λ. Further, the distribution of $\hat{\mathbf{A}}(\lambda)$ may be approximated by a complex normal with mean $\mathbf{A}(\lambda)$ and with vec $\hat{\mathbf{A}}(\lambda)$, the column vector resulting from stringing out the columns of $\hat{\mathbf{A}}(\lambda)$ each underneath the previous, having covariance matrix

$$\text{(22)} \qquad 2\pi T\mathbf{f}(\lambda) \otimes \left(\sum_k \mathbf{d}_M^T\left(\frac{2\pi k}{T}\right)\overline{\mathbf{d}_M^T\left(\frac{2\pi k}{T}\right)}'\right)^{-1}.$$

Once one notes that the spectral density matrix $\mathbf{f}(\lambda)$ may be estimated by

$$\text{(23)} \qquad \hat{\mathbf{f}}(\lambda) = (2\pi T)^{-1}(n-r)^{-1}\mathbf{Y}(\mathbf{I} - \overline{\mathbf{X}}'(\mathbf{X}\overline{\mathbf{X}}')^{-1}\mathbf{X})\overline{\mathbf{Y}}'$$

whose distribution is approximately $(n-r)^{-1}$ times a complex Wishart, degrees of freedom $(n-r)$, parameter $\mathbf{f}(\lambda)$ independent of $\hat{\mathbf{A}}(\lambda)$, the means are at hand for testing hypotheses of interest and for constructing confidence regions for the entries of $\mathbf{A}(\lambda)$. Important cases of particular interest will be considered in the next section.

If desired, the average evoked response matrix $\mathbf{a}(t)$ may be estimated by

$$\text{(24)} \qquad \hat{\mathbf{a}}(t) = Q^{-1}\sum_{q=0}^{Q-1}\exp\left(i\frac{2\pi kq}{Q}\right)\hat{\mathbf{A}}\left(\frac{2\pi q}{Q}\right),$$

but for many purposes it is satisfactory and even advantageous to stick to the frequency domain.

The advantages of the approach of this section are now seen to include: (1) multiple stimuli and varying intensities are handled directly, (2) the individual effects may overlap (and still be deconvolved), (3) approximate sampling properties may be stated succinctly. In the next section it will be seen that substantially different and complex experimental designs may be dealt with within the framework of the model. It is important to remark, however, that for the estimates to be reasonably stable the matrix $(\mathbf{X}\overline{\mathbf{X}}')$ should not be near singular at frequencies where estimation is being carried out. Indeed, as expression (22) shows, one wants the matrix $(\mathbf{X}\overline{\mathbf{X}}')$ as large as possible, in a certain sense, for precise estimates. Beyond these comments, the timing of stimuli application may be arbitrary.

4. Some particular cases

This section presents details concerning certain specific experimental designs and concerning the significance testing of certain specific hypotheses.

4.1 Single stimulus—single response

Consider the case of a single stimulus, applied at times σ_j, with constant intensity. Suppose that a single response series is recorded. In this case the general model (15) becomes

$$\text{(25)} \qquad Y(t) = \mu + \sum_j a(t - \sigma_j) + \varepsilon(t).$$

Issues at hand include: the estimation of $a(t)$ and $A(\lambda)$, the testing of $a(\cdot) \equiv 0$ and of $A(\cdot) \equiv 0$ and the testing of $A(\lambda) = 0$. The estimate of $A(\lambda)$ is given by (21),

$$\text{(26)} \qquad \hat{A}(\lambda) = \sum_k d_Y^T\left(\frac{2\pi k}{T}\right)\overline{d_M^T\left(\frac{2\pi k}{T}\right)} \Big/ \sum_k \left|d_M^T\left(\frac{2\pi k}{T}\right)\right|^2$$

with

$$d_M^T(\lambda) = \sum_{j=1}^{M}\exp(-i\lambda\sigma_j)$$

and M the number of times the stimulus is applied. The variance of $\hat{A}(\lambda)$ is approximately

$$\text{(27)} \qquad 2\pi T f(\lambda) \Big/ \sum_k \left|d_M^T\left(\frac{2\pi k}{T}\right)\right|^2$$

and the error spectrum may be estimated by

$$\text{(28)} \qquad \hat{f}(\lambda) = (2\pi T)^{-1}(n-1)^{-1}\left[\sum_k \left|d_Y^T\left(\frac{2\pi k}{T}\right)\right|^2 - \right.$$
$$\left. - \left|\sum_k d_Y^T\left(\frac{2\pi k}{T}\right)\overline{d_M^T\left(\frac{2\pi k}{T}\right)}\right|^2 \Big/ \sum_k \left|d_M^T\left(\frac{2\pi k}{T}\right)\right|^2\right]$$

Expressions (27), (28) may be combined with the approximate complex normal distribution of $\hat{A}(\lambda)$ to construct confidence regions for $A(\cdot)$.

The hypothesis $A(\lambda) = 0$ may be tested via the statistic

$$\text{(29)} \qquad |\hat{A}(\lambda)|^2\hat{f}(\lambda)^{-1}\sum_k \left|d_M^T\left(\frac{2\pi k}{T}\right)\right|^2(2\pi T)^{-1}$$

which is distributed approximately as $F_{2;2(n-1)}$ in the case that $A(\lambda) = 0$ and is stochastically larger otherwise.

The fact that the $\hat{A}(\lambda)$ are asymptotically independent at distinct frequencies λ, allows the construction of a composite test statistic for the hypothesis $A(\cdot) \equiv 0$ if desired. It seems, however, that the use of such an overall statistic constitutes too brutal a summary of the data collected.

4.2 Several stimuli—single response

Suppose that there are r distinct stimuli (or treatments) of interest. Suppose that each individual stimulus is applied with constant intensity, then the treatment times may be assimilated to an r-variate point process $\{M_1(\cdot), \ldots, M_r(\cdot)\}$. If the treatments do not interact with each other, then a useful model for the response series is provided by

$$Y(t) = \mu + \sum_{j=1}^{r} \int a_j(t-u)dM_j(u) + \varepsilon(t) \tag{30}$$

with the function $a_j(\cdot)$ providing the effect of the jth treatment.

Hypotheses of interest in this situation include: (i) $a_j(\cdot) \equiv 0$, $j = 1, \ldots, r$ (i.e. none of the treatments are having an effect) and (ii) $a_j(\cdot) \equiv a_1(\cdot), j = 2, \ldots, r$ (i.e. all of the treatments are having the same effect).

Supposing that $\hat{A}(\lambda)$ denotes the estimate (derived above) of the r-vector $\mathbf{A}(\lambda)$ the hypothesis (i) may be tested at frequency λ via the statistic

$$\hat{\mathbf{A}}(\lambda)\mathbf{X}\overline{\mathbf{X}}'\overline{\hat{\mathbf{A}}(\lambda)}'/r2\pi T\hat{f}(\lambda) \tag{31}$$

whose null distribution may be approximated by $F_{2r;2(n-r)}$. The distribution will be stochastically larger if the hypothesis does not hold. (The procedure of 4.1 above is a particular case of this.)

The hypothesis (ii) is undoubtedly of greater scientific interest. Following the discussion of Section 2, this hypothesis may be examined via the statistic

$$\hat{\mathbf{A}}(\lambda)\mathbf{B}(\mathbf{B}'(\mathbf{XX}')^{-1}\mathbf{B})^{-1}\mathbf{B}'\overline{\hat{A}(\lambda)}'/(r-1)2\pi T\hat{f}(\lambda) \tag{32}$$

with $\mathbf{B}$ an $r \times r-1$ matrix whose first row is all -1, with $B_{j+1,j} = 1, j = 1, \ldots, r-1$ and with entries 0 otherwise. The distribution of the statistic (32) may be approximated by $F_{2(r-1);2(n-r)}$ in the null case. In some circumstances it may be helpful to reparameterize the model in this case via $a_j(\cdot) = b(\cdot) + b_j(\cdot)$ with the constraint $\sum_j b_j(\cdot) \equiv 0$. The hypothesis (ii) then becomes $b_j(\cdot) \equiv 0$, $j = 1, \ldots, r$.

4.3 Factorial experiment—single response

The model (30) assumed that the treatments did not interact. In some situations the question of interest is whether two stimuli, applied simultaneously, do interact. Suppose that one has two basic stimuli A, B and that an experiment is carried out in which sometimes A alone is applied, sometimes B alone is applied and sometimes both are applied. Let the point process $M_1(\cdot)$ record the times of application of A alone, $M_2(\cdot)$ record the times of B alone and $M_3(\cdot)$ the times of simultaneous application. A relevant model for this situation is provided by

$$Y(t) = \mu + \sum_{j=1}^{3} \int a_j(t-u)dM_j(u) + \varepsilon(t). \tag{33}$$

The hypothesis of interest may be formalized as: is $a_3(\cdot) \equiv a_1(\cdot) + a_2(\cdot)$? Again following the discussion of Section 2, this hypothesis may be examined via

$$\hat{\mathbf{A}}(\lambda)\mathbf{B}(\mathbf{B}'(\mathbf{XX}')^{-1}\mathbf{B})^{-1}\mathbf{B}'\overline{\hat{\mathbf{A}}(\lambda)}'/2\pi T\hat{f}(\lambda) \tag{34}$$

where $\mathbf{B}' = [-1\ -1\ 1]$. The null distribution of the statistic (34) will be approximately $F_{2;2(n-1)}$.

This type of experiment, with periodic application of the treatments is considered in Diamond (1964). There the stimuli are two nearly identical light pulses in one case and two lights of different colours in another. His results suggest consideration of a further hypothesis; namely, is $a_3(\cdot) \equiv k(a_1(\cdot) + a_2(\cdot))$? Such a hypothesis may also be considered within the framework developed in this paper.

4.4 Trends or covariates present

The model (30) may be expanded to deal with the case that trends or concomitant series are present. Specifically consider the set-up

$$Y(t) = \sum_{i=1}^{q} \mu_i g_i(t) + \sum_{j=1}^{r} \int a_j(t-u)dM_j(u) + \varepsilon(t) \tag{35}$$

with the $g_i(\cdot)$ known functions. The case of linear trend corresponds to $q = 2$, $g_1(t) = 1$, $g_2(t) = t$. The case of blocks corresponds to $g_i(t) = 1$ for t in the i-th block and 0 otherwise. The case of the response depending on covarying series as well as the treatments applied, corresponds to $g_i(\cdot)$ being the i-th concomitant series. As in the case of ordinary time series (see Section 5.11 of Brillinger (1975)), this model may be dealt with by first estimating the μ_i by least squares, then correcting the $Y(t)$ by the fitted $g_i(t)$ and proceeding as in Section 3. An experiment may be made considerably more sensitive by the employment of blocks and/or the recording of relevant covariate series, in many cases.

4.5 Vector response

This section considers the case that the response observed is an s-vector of time series $\mathbf{Y}(t) = [Y_j(t)]$, $j = 1, \ldots, s$. In the case of a single stimulus, the linear model takes the form

$$\mathbf{Y}(t) = \boldsymbol{\mu} + \int \mathbf{a}(t-u)dM(u) + \boldsymbol{\epsilon}(t) \tag{36}$$

with $\boldsymbol{\mu}$, $\mathbf{a}(t)$, $\boldsymbol{\epsilon}(t)$, s-vectors. Hypotheses of interest include: (i) $\mathbf{a}(\cdot) \equiv \mathbf{0}$ (i.e. the stimulus is not having an effect on any of the series), and (ii) $a_1(\cdot) \equiv a_2(\cdot)$ in the case $s = 2$ (i.e. the average responses at say two symmetrically placed leads are the same).

First consider the hypothesis (i). The discussion of Section 2 suggests the computation of the $s \times s$ matrix

(37) $$\mathbf{S} = \hat{\mathbf{A}}(\lambda)(\mathbf{X}\overline{\mathbf{X}}')^{-1}\overline{\hat{\mathbf{A}}(\lambda)}'.$$

Its large sample null distribution is complex Wishart with 1 degree of freedom and parameter $2\pi T\mathbf{f}(\lambda)$. At the same time the large sample distribution of $(n-1)\hat{\mathbf{f}}(\lambda)$ is complex Wishart with $n-1$ degrees of freedom and parameter $\mathbf{f}(\lambda)$. Because of the matrix values of the statistics $\mathbf{S}, \hat{\mathbf{f}}(\lambda)$, there is no most sensible test statistic. Tests may be based upon the matrix $\mathbf{S}\hat{\mathbf{f}}(\lambda)^{-1}/(n-1)2\pi T$. By analogy with the ordinary case, test statistics to consider include: its trace (cp. Lawley–Hotelling), its determinant (cp. Wilks, likelihood ratio test) and its largest root (cp. Roy's test). Approximations to the distributions of these test statistics may be developed in the manner of Krishnaiah (1980). Part of the motivation for considering such multivariate test procedures is that they may be more sensitive than univariate procedures to the presence of small effects in the individual response series.

In the case of hypothesis (ii) above, the test statistic

(38) $$\left|\hat{A}_1(\lambda) - \hat{A}_2(\lambda)\right|^2 \sum_k \left|d_M^T\left(\frac{2\pi k}{T}\right)\right|^2 \Big/ 4\pi T\hat{f}(\lambda)$$

is appropriate. Its large sample null distribution is $F_{2:2(n-1)}$. These remarks apply to the case of $s = 2$. If desired a similar multivariate test statistic may be developed for example for the hypothesis that $a_1(\cdot) = a_2(\cdot)$ and $a_3(\cdot) = a_4(\cdot)$ and ....

Vector responses also occur in situations of quite different character, namely, when a basic experiment is replicated possibly with the same subject, possibly with different subjects. The individual response series in this situation may be viewed as statistically independent. This simplifies the analysis greatly. For example a model might be

(39) $$Y_j(t) = \mu_j + \int a_j(t-u)dM_j(u) + \varepsilon_j(t)$$

$j = 1, \ldots, s$ and hypothesis of interest $a_j(\cdot) = a(\cdot), j = 1, \ldots, s$. It may be that the index 'j' has a further structure, for example corresponding to a factorial experiment or a randomized block design. It appears that the model and general approach of Section 3 is general enough to handle many situations of this character (Brillinger, 1980).

In summary, the basic idea throughout this Section 4 is that with the model (15) the hypothesis $\mathbf{AA}(\lambda)\mathbf{B} = \mathbf{C}$ may be tested by noting that the large sample distribution of

(40) $$(\mathbf{AA}(\lambda)\mathbf{B} - \mathbf{C})(\mathbf{B}'(\mathbf{X}\overline{\mathbf{X}}')^{-1}\mathbf{B})^{-1}(\overline{\mathbf{AA}(\lambda)\mathbf{B} - \mathbf{C}})'$$

is complex Wishart with degrees of freedom rank($\mathbf{B}$) and parameter $\mathbf{Af}(\lambda)\mathbf{A}'$, while $(n-r)\mathbf{A}\hat{\mathbf{f}}(\lambda)\mathbf{A}'$ is independently complex Wishart d.f. $n-r$ and parameter $\mathbf{Af}(\lambda)\mathbf{A}'$.

5. Optimal design

So far in this work, the input process $\mathbf{M}(\cdot)$, providing the times of stimuli application has been taken as given. On many occasions it is possible, however, to select the times at which treatments are applied. The expressions of Section 3 give some indication of how times might be selected in an optimal fashion.

Specifically, the large sample covariance matrix of the entries of $\hat{\mathbf{A}}(\lambda)$ is proportional to $(\mathbf{X}\overline{\mathbf{X}}')^{-1}$. Now, it may be shown (e.g. exercise 6.14.18 in Brillinger (1975)) that for given diagonal entries of $\mathbf{X}\overline{\mathbf{X}}'$, the (large sample) variances of the entries of $\hat{\mathbf{A}}(\lambda)$ are minimized by taking the rows of $\mathbf{X}$ to be orthogonal. This orthogonality will occur, approximately, if the various $M_j(\cdot)$ are taken to be statistically independent of each other. Further, approximate orthogonality will occur simultaneously at all λ.

Consider next the choice of $\mathbf{X}$ as a function of λ. Suppose that it is essentially diagonal. The j-th diagonal entry is proportional to $\hat{f}_{jj}(\lambda)$ an estimate of the power spectrum of the j-th input at frequency λ. Suppose $s = 1$ and consider the problem of estimating $A_j(\cdot)$ simultaneously at a number of frequencies λ_k. Noting that the $\hat{A}_j(\lambda_k)$ are asymptotically independent of each other for distinct λ_k, take

(41) $$\sum_k \operatorname{var} \hat{A}_j(\lambda_k)$$

as a measure of the overall precision of the estimate. In the large sample case expression (41) is approximately proportional to

(42) $$\sum_k f(\lambda_k)/\hat{f}_{jj}(\lambda_k).$$

For given (total input power), $\sum \hat{f}_{jj}(\lambda_k)$, the Cauchy–Schwarz inequality shows that the criterion (42) is minimized by choosing $\hat{f}_{jj}(\lambda_k)$ to be proportional to $f(\lambda_k)^{1/2}$. In the case of white noise error (with constant $f(\lambda)$), one is led to choose $\hat{f}_{jj}(\lambda)$ to be constant. This will happen, approximately, if the $M_j(\cdot)$ are taken to be realizations of Poisson processes.

Finally, expression (42) will be smaller, the larger $\hat{f}_{jj}(\cdot)$ is. It will be larger, the greater the rate the process $M_j(\cdot)$ has.

In summary, the above crude arguments suggest strongly, that when possible, one should take the times of application of stimuli to correspond to the times of events of Poisson processes with high rates and with different stimuli corresponding to statistically independent Poisson processes. (In practice the rates of the Poisson processes may not be taken to be arbitrarily high because the assumed linearity of the model will break down.) The argument was given for the case $s = 1$. It applies equally to the case of general s when the components of $\boldsymbol{\epsilon}(\cdot)$ are independent white noises.

So far the discussion has centered around the model (15). When trends or covariates are present, as in the model (35), other designs suggest themselves as they do in the ordinary experimental design case. For example, suppose that $\boldsymbol{\mu}$ of (15) is not constant throughout the whole course of the experiment, but may be reasonably viewed as constant during blocks of time. (This corresponds to defining $g_i(\cdot)$ in (35) to be 1 for the i-th time block and 0 otherwise.) Classical experimental design, specifically the work concerning randomized blocks, suggests that one should ensure that the r treatments are applied the same number of times and in random order, throughout each of the blocks. (Undoubtedly analytic arguments could be generated to demonstrate the 'optimality' of such a design for the model (35) as specified.)

6. Randomization

Turning to another statistical aspect of evoked response experiments, consider a situation in which a single stimulus (of constant intensity) may be applied to a subject and in which it is desired to test whether the stimulus has any effect in fact. The test statistic (39) of section 4 was derived under the assumptions of Model I. The derivation of the indicated null distribution for (29) made use of assumptions of stationarity and mixing for the error series $\varepsilon(\cdot)$.

In fact, an experimental design and inferential procedure exist that require much less in the way of assumptions. Let $\tau_1, \ldots, \tau_N$ denote a collection of possible times of stimulation. Let stimulation take place at time τ_j with probability 1/2 independently of what takes place at the other τ's. Let the point process $M_1(\cdot)$ correspond to the times at which stimulation does take place and $M_2(\cdot)$ correspond to the times at which it does not. As a test statistic one may compute (32), with $r = 2$. Its null distribution may be evaluated by computer rerandomization (over choices of τ's in M_1 as opposed to M_2.) it turns out that the null distribution may still be reasonably approximated by $F_{2:2(n-2)}$. The difference is that the distribution results from the randomness physically introduced by the experimenter. He does not have to hope that an error series happens to exist with certain properties.

A similar randomization procedure exists for the case of r treatments. At each of the τ_j one applies one of the r treatments, with the one applied selected randomly from the r available, independently of the treatments applied at the other τ's. To check, for example, that the treatments each have the same average effect the statistic (32) may be computed and its null distribution approximated by $F_{2(r-1):2(n-r)}$.

This technique may be similarly extended to other experimental designs, for example, those involving covariates or blocks. The advantage of the approach is its allowing the experimenter to take advantage of the physical act of randomization to justify tests of significance rather than having to introduce assumptions of stationarity and mixing concerning the error series.

7. Concluding remarks

Before ending the paper it seems appropriate to summarize what has been gained by the approach advocated, beyond the traditional technique of periodic stimulation. The principal contribution is the presentation of a general framework encompassing the traditional experiments and those with stimuli applied at irregularly spaced times. The framework allows the use of standard results concerning linear models to design experiments and to suggest test statistics. A secondary advantage is the result that through irregular stimulation at high rate one may be able to determine average stimulus effects more rapidly or more precisely.

References

Bieddenbach, M. A. and Freeman, W. J. (1965). Linear domain of potentials for the prepyriform cortex with respect to stimulus parameters. *Exp. Neurology* **11**, 400–417.

Brillinger, D. R. (1974). The Fourier analysis of stationary processes. *Proc. IEEE* **62**, 1628–1643.

Brillinger, D. R. (1975). *Time Series: Data Analysis and Theory.* Holt–Rinehart, New York. (Expanded Edition, 1980, Holden–Day, San Francisco.)

Brillinger, D. R. (1978). A note on the estimation of evoked response. *Biol. Cybernetics* **31**, 141–144.

Brillinger, D. R. (1980). Analysis of variance and problems under time series models. *In: Handbook of Statistics* **1** (Krishnaiah, P. R., Ed.), 237–277. North–Holland, Amsterdam.

Brillinger, D. R. (1981). Some aspects of the analysis of evoked response experiments. *In: Proceedings of the International Symposium on Statistics and Related Topics* (Saleh, A. K. Md. E., Csorgo, M., Dawson, D. A. and Rao, J. N. K., Eds.). North–Holland, Amsterdam.

Callaway, E., Tueting, P. and Koslow, S. H. (1978). *Event-Related Brain Potentials in Man.* Academic Press, New York.

Cochran, W. G. and Cox, G. M. (1957). *Experimental Designs.* Wiley, New York.

Cox, D. R. (1958). *Planning of Experiments.* Wiley, New York.

Cox, D. R. and Lewis, P. A. W. (1966). *The Statistical Analysis of Series of Events.* Methuen, London.

Diamond, S. P. (1964). Input–output relations. *Ann. New York Acad. Sci.* **112**, 160–171.

Donchin, E. and Lindsley, D. B. (1969). *Average Evoked Potentials.* NASA, Washington.

Freeman, W. J. (1980). Measurement of cortical evoked potentials by decomposition of their wave forms. *J. Cybernetics and Inf. Sci.* **2**, 44–56.

Glaser, E. M. and Ruchkin, D. S. (1976). *Principles of Neurobiological Signal Analysis.* Academic Press, New York.

John, E. R. (1977). *Neurometrics.* Wiley, New York.

Kempthorne, O. (1952). *The Design and Analysis of Experiments.* Wiley, New York.

Krishnaiah, P. R. (1980). Computations of some multivariate distributions. *In: Handbook o Statistics* **1** (Krishnaiah,P. R., Ed.), pp. 745–971. North–Holland, Amsterdam.

Rao, C. R. (1965). *Linear Statistical Inference and Its Applications.* Wiley, New York.

Roy, S. N., Gnanadesikan, R. and Srivastava, J. N. (1971). *Analysis and Design of Certain Quantitative Multiresponse Experiments.* Pergamon, Oxford.

Schimmel, H. (1967). The (±) reference: accuracy of estimated mean components in average response studies. *Science* **157**, 92–93.

Seber, G. A. F. (1977). *Linear Regression Analysis.* Wiley, New York.

Shagass, C. (1972). *Evoked Brain Potentials in Psychiatry.* Plenum, New York.

Thatcher, R. W. and John, E. R. (1977). *Foundations of Cognitive Processes.* Wiley, New York.

Timm, N. H. (1980). Multivariate analysis of variance of repeated measurements. *In: Handbook of Statistics* **1** (Krishnaiah, P. R., Ed.), pp. 41–88. North–Holland, Amsterdam.

ably viewed as constant during blocks of time. (This corresponds to defining $g_i(\cdot)$ in (35) to be 1 for the i-th time block and 0 otherwise.) Classical experimental design, specifically the work concerning randomized blocks, suggests that one should ensure that the r treatments are applied the same number of times and in random order, throughout each of the blocks. (Undoubtedly analytic arguments could be generated to demonstrate the 'optimality' of such a design for the model (35) as specified.)

6. Randomization

Turning to another statistical aspect of evoked response experiments, consider a situation in which a single stimulus (of constant intensity) may be applied to a subject and in which it is desired to test whether the stimulus has any effect in fact. The test statistic (39) of section 4 was derived under the assumptions of Model I. The derivation of the indicated null distribution for (29) made use of assumptions of stationarity and mixing for the error series $\varepsilon(\cdot)$.

In fact, an experimental design and inferential procedure exist that require much less in the way of assumptions. Let $\tau_1, \ldots, \tau_N$ denote a collection of possible times of stimulation. Let stimulation take place at time τ_j with probability 1/2 independently of what takes place at the other τ's. Let the point process $M_1(\cdot)$ correspond to the times at which stimulation does take place and $M_2(\cdot)$ correspond to the times at which it does not. As a test statistic one may compute (32), with $r = 2$. Its null distribution may be evaluated by computer rerandomization (over choices of τ's in M_1 as opposed to M_2.) it turns out that the null distribution may still be reasonably approximated by $F_{2:2(n-2)}$. The difference is that the distribution results from the randomness physically introduced by the experimenter. He does not have to hope that an error series happens to exist with certain properties.

A similar randomization procedure exists for the case of r treatments. At each of the τ_j one applies one of the r treatments, with the one applied selected randomly from the r available, independently of the treatments applied at the other τ's. To check, for example, that the treatments each have the same average effect the statistic (32) may be computed and its null distribution approximated by $F_{2(r-1):2(n-r)}$.

This technique may be similarly extended to other experimental designs, for example, those involving covariates or blocks. The advantage of the approach is its allowing the experimenter to take advantage of the physical act of randomization to justify tests of significance rather than having to introduce assumptions of stationarity and mixing concerning the error series.

7. Concluding remarks

Before ending the paper it seems appropriate to summarize what has been gained by the approach advocated, beyond the traditional technique of periodic stimulation. The principal contribution is the presentation of a general framework encompassing the traditional experiments and those with stimuli applied at irregularly spaced times. The framework allows the use of standard results concerning linear models to design experiments and to suggest test statistics. A secondary advantage is the result that through irregular stimulation at high rate one may be able to determine average stimulus effects more rapidly or more precisely.

References

Bieddenbach, M. A. and Freeman, W. J. (1965). Linear domain of potentials for the prepyriform cortex with respect to stimulus parameters. *Exp. Neurology* **11**, 400–417.
Brillinger, D. R. (1974). The Fourier analysis of stationary processes. *Proc. IEEE* **62**, 1628–1643.
Brillinger, D. R. (1975). *Time Series: Data Analysis and Theory.* Holt–Rinehart, New York. (Expanded Edition, 1980, Holden–Day, San Francisco.)
Brillinger, D. R. (1978). A note on the estimation of evoked response. *Biol. Cybernetics* **31**, 141–144.
Brillinger, D. R. (1980). Analysis of variance and problems under time series models. *In: Handbook of Statistics* **1** (Krishnaiah, P. R., Ed.), 237–277. North–Holland, Amsterdam.
Brillinger, D. R. (1981). Some aspects of the analysis of evoked response experiments. *In: Proceedings of the International Symposium on Statistics and Related Topics* (Saleh, A. K. Md. E., Csorgo, M., Dawson, D. A. and Rao, J. N. K., Eds.). North–Holland, Amsterdam.
Callaway, E., Tueting, P. and Koslow, S. H. (1978). *Event-Related Brain Potentials in Man.* Academic Press, New York.
Cochran, W. G. and Cox, G. M. (1957). *Experimental Designs.* Wiley, New York.
Cox, D. R. (1958). *Planning of Experiments.* Wiley, New York.
Cox, D. R. and Lewis, P. A. W. (1966). *The Statistical Analysis of Series of Events.* Methuen, London.
Diamond, S. P. (1964). Input–output relations. *Ann. New York Acad. Sci.* **112**, 160–171.
Donchin, E. and Lindsley, D. B. (1969). *Average Evoked Potentials.* NASA, Washington.
Freeman, W. J. (1980). Measurement of cortical evoked potentials by decomposition of their wave forms. *J. Cybernetics and Inf. Sci.* **2**, 44–56.
Glaser, E. M. and Ruchkin, D. S. (1976). *Principles of Neurobiological Signal Analysis.* Academic Press, New York.
John, E. R. (1977). *Neurometrics.* Wiley, New York.
Kempthorne, O. (1952). *The Design and Analysis of Experiments.* Wiley, New York.
Krishnaiah, P. R. (1980). Computations of some multivariate distributions. *In: Handbook o Statistics* **1** (Krishnaiah,P. R., Ed.), pp. 745–971. North–Holland, Amsterdam.
Rao, C. R. (1965). *Linear Statistical Inference and Its Applications.* Wiley, New York.
Roy, S. N., Gnanadesikan, R. and Srivastava, J. N. (1971). *Analysis and Design of Certain Quantitative Multiresponse Experiments.* Pergamon, Oxford.
Schimmel, H. (1967). The (±) reference: accuracy of estimated mean components in average response studies. *Science* **157**, 92–93.
Seber, G. A. F. (1977). *Linear Regression Analysis.* Wiley, New York.
Shagass, C. (1972). *Evoked Brain Potentials in Psychiatry.* Plenum, New York.
Thatcher, R. W. and John, E. R. (1977). *Foundations of Cognitive Processes.* Wiley, New York.
Timm, N. H. (1980). Multivariate analysis of variance of repeated measurements. *In: Handbook of Statistics* **1** (Krishnaiah, P. R., Ed.), pp. 41–88. North–Holland, Amsterdam.