Biol. Cybernetics 31, 141—144 (1978)

Biological Cybernetics

A Note on the Estimation of Evoked Response*

D. R. Brillinger

Department of Statistics, University of California, Berkeley, CA, USA

Abstract. The traditional average evoked response is compared and contrasted with several alternate estimates, derived from frequency domain considerations, for a model of evoked responses superimposed upon a stationary noise series. Further, a means of constructing approximate confidence intervals for the values of an evoked response is indicated. The case of several simultaneously recorded series is also considered.

I. Introduction

Consider a situation in which stimuli are applied at times $\sigma_1 < \sigma_2 < \sigma_3 < \ldots$ and simultaneously a continuously varying response function $Y(t)$ is recorded. Suppose that it is reasonable to model the series $Y(t)$ by

$$Y(t) = \sum_j s(t - \sigma_j) + \varepsilon(t) \tag{1.1}$$

with $s(t)$ an evoked response function and with $\varepsilon(t)$ a stationary, zero mean, error series. In many cases the function $s(t)$ will vanish for $t<0$ and will be of finite duration. In the classical evoked response experiment, see for example Donchin and Lindsley (1969), the successive σ_j are taken farther apart than the length of the interval where $s(t)$ is non-zero and $s(t)$ is estimated by the average evoked response

$$s_0^T(t) = \sum_{j=1}^{M} Y(t + \sigma_j)/M \tag{1.2}$$

where $\sigma_1, \sigma_2, \ldots, \sigma_M$ are the times at which stimuli were applied during the time interval $[0, T]$. It is apparent that if the interval between the σ_j is not sufficiently large, then the estimate (1.2) will be biased. It is plausible that with autocorrelation in the error series $\varepsilon(t)$ the equal weighting of the M terms of (1.2) will not be efficient. It is not obvious how to set confidence intervals for the estimate, in the presence of autocorrelation. This paper is concerned with certain statistical properties of the estimate (1.2), with developing estimates of $s(t)$ that place no a priori restriction on how close the σ_j may be to each other, and with developing efficient estimates of $s(t)$ in the presence of autocorrelation.

We assume that $s(t)$ is a fixed, unknown function defined for $-\infty < t < \infty$. We assume that $\varepsilon(t)$ is a stationary time series with mean $E\varepsilon(t) = 0$, with autocovariance function

$$\operatorname{cov}\{\varepsilon(t+u), \varepsilon(t)\} = c_{\varepsilon\varepsilon}(u), \tag{1.3}$$

with power spectrum

$$f_{\varepsilon\varepsilon}(\lambda) = (2\pi)^{-1} \int_{-\infty}^{\infty} c_{\varepsilon\varepsilon}(u) \exp\{-i\lambda u\}\, du \tag{1.4}$$

and satisfying the asymptotic independence or mixing Assumption I of the Appendix. Let $M(t)$ count the number of σ_j in the interval $[0, t]$. Then $M(t)$ is a step function increasing by 1 each time a σ_j occurs. We assume that the following limits exist for $M(t)$

$$p_M = \lim_{T\to\infty} M(T)/T \tag{1.5}$$

and

$$P_{MM}(u) = \lim_{T\to\infty} \int_0^T [\mathrm{M}(\mathrm{t}+\mathrm{u}) - \mathrm{M}(\mathrm{t})]\, \mathrm{dM}(\mathrm{t})/\mathrm{T}\,. \tag{1.6}$$

The parameter p_M specifies the (asymptotic) rate with which stimuli are being applied. The parameter (1.6) provides information concerning the relative spacings of the σ_j. We further require $M(t)$ to satisfy Assumption II of the Appendix.

* Prepared with the support of the National Science Foundation Grant MCS76-06117

0340-1200/78/0031/0141/$01.00

With the point process $M(t)$ defined in the above fashion, the model (1.1) may be written

$$Y(t)=\int s(t-u)dM(u)+\varepsilon(t) \quad (1.7)$$

and the estimate (1.2) may be written

$$s_0^T(t)=\int_0^{T-t} Y(t+u)dM(u)/M(T). \quad (1.8)$$

We turn to a statistical investigation of the average evoked response, (1.8).

II. The Average Evoked Response

Because the noise series has 0 mean,

$$EY(t)=\int s(t-u)dM(u)=\sum_j s(t-\sigma_j)$$

and

$$Es_0^T(t)=\int_0^{T-t}\int s(t-u+v)dM(u)dM(v)$$
$$=\sum_j\sum_k s(t-\sigma_j+\sigma_k)/M(T). \quad (2.1)$$

From the last expression here it is apparent that $Es_0^T(t)=s(t)$ provided $t<T-\sigma_{M(T)}$ and provided $|\sigma_j-\sigma_k|$, $j\neq k$, is greater than the interval length of non-zero $s(t)$ for all j,k. In the general case

$$Es_0^T(t)\sim\int s(t-u)dP_{MM}(u)/p_M \quad (2.2)$$

as $T\to\infty$, using expressions (1.5), (1.6), and this is not generally $s(t)$. Clearly an alternate estimate is called for.

From expressions (1.3) and (1.8) the variance of this estimate is given by

$$\operatorname{var} s_0^T(t)=\int_0^{T-t}\int_0^{T-t} c_{\varepsilon\varepsilon}(u-v)dM(u)dM(v)/M(T)^2$$
$$=\sum_j\sum_k c_{\varepsilon\varepsilon}(\sigma_j-\sigma_k)/M(T)^2 \quad (2.3)$$
$$\sim\int_{-\infty}^{\infty} c_{\varepsilon\varepsilon}(u)dP_{MM}(u)/Tp_M^2. \quad (2.4)$$

In the case that the noise values have correlation zero at lag equal and beyond the minimum spacing of the σ_j, it is apparent from (2.3) that the variance is $c_{\varepsilon\varepsilon}(0)/M(T)$ and so an estimate of the variance may be constructed readily. In the general case a variance estimate is not readily available.

Finally it may be remarked that, under the conditions of the Appendix, the statistic $s_0^T(t)$ is asymptotically normal with the indicated mean and variance. (See Brillinger, 1973) Provided a consistent estimate of the variance is available, approximate confidence intervals may therefore be constructed directly for $Es_0^T(t)$.

III. Alternate Estimates

When written in the form (1.7) the model is seen to have the structure of a linear time invariant system carrying an input process, M, over into an output process Y. Such systems are generally investigated by means of cross-spectral analysis, see Brillinger (1974a) and Brillinger (1975), Chap. 6. The steps involved in such an analysis follow.

Define

$$S(\lambda)=\int_{-\infty}^{\infty} s(t)\exp\{-i\lambda t\}dt, \quad (3.1)$$

the system transfer function. Define

$$d_Y^T(\lambda)=\int_0^T \exp\{-i\lambda t\}Y(t)dt$$

with a similar definition for $d_\varepsilon^T(\lambda)$. Define

$$d_M^T(\lambda)=\int_0^T \exp\{-i\lambda t\}dM(t)=\sum_j \exp\{-i\lambda\sigma_j\}$$

with the last sum over the available σ_j. It now follows from (1.7) that

$$d_Y^T(\lambda)\doteq S(\lambda)d_M^T(\lambda)+d_\varepsilon^T(\lambda)$$

and further

$$d_Y^T\left(\frac{2\pi k}{T}\right)\doteq S\left(\frac{2\pi k}{T}\right)d_M^T\left(\frac{2\pi k}{T}\right)+d_\varepsilon^T\left(\frac{2\pi k}{T}\right)$$
$$\doteq S(\lambda)d_M^T\left(\frac{2\pi k}{T}\right)+d_\varepsilon^T\left(\frac{2\pi k}{T}\right) \quad (3.2)$$

for k an integer with $2\pi k/T\doteq\lambda$. If a number of distinct frequencies $2\pi k/T$ are considered with $2\pi k/T$ near λ, then under the assumptions of the Appendix, the corresponding Fourier transforms $d_\varepsilon^T\left(\frac{2\pi k}{T}\right)$ will be asymptotically complex normal variates with mean 0 and variance $2\pi Tf_{\varepsilon\varepsilon}(\lambda)$. (See for example Brillinger, 1974b) Expression (3.2) suggests estimating $S(\lambda)$ by

$$S^T(\lambda)=\sum_k d_Y^T\left(\frac{2\pi k}{T}\right)\overline{d_M^T\left(\frac{2\pi k}{T}\right)}\Big/\sum_k\left|d_M^T\left(\frac{2\pi k}{T}\right)\right|^2$$
$$=f_{YM}^T(\lambda)/f_{MM}^T(\lambda) \quad (3.3)$$

with the sums in (3.3) over say K distinct frequencies $2\pi k/T$ near λ and where

$$f_{YM}^T(\lambda)=(2\pi T)^{-1}\sum_k d_Y^T\left(\frac{2\pi k}{T}\right)\overline{d_M^T\left(\frac{2\pi k}{T}\right)}$$

with a similar definition of $f_{MM}^T(\lambda)$.

The estimate (3.3) parallels the usual estimate of cross-spectral analysis considered for continuous series in Brillinger (1974b), for example, and for point processes in Brillinger et al. (1976). In the manner of Sect. 6.3, Brillinger (1975), the statistic (3.3) will be asymptotically complex normal with mean $S(\lambda)$ and variance $f_{\varepsilon\varepsilon}(\lambda)/Kf_{MM}^T(\lambda)$. Confidence intervals may be set in the manner of that reference.

As in Sect. 6.8, Brillinger (1975), the evoked response $s(t)$ may now be estimated by

$$s_1^T(t)=\frac{1}{P}\sum_{p=0}^{P-1} S^T\left(\frac{2\pi p}{P}\right)\exp\left\{i\frac{2\pi pt}{P}\right\} \tag{3.4}$$

for some integer P. The estimate (3.4), in contrast to the estimate (1.2), remains valid even when the individual evoked responses overlap. The variance of the estimate (3.4) is given, approximately, by

$$\frac{1}{K}\sum_{p=0}^{P-1} f_{\varepsilon\varepsilon}\left(\frac{2\pi p}{P}\right) f_{MM}^T\left(\frac{2\pi p}{P}\right)^{-1}/P^2. \tag{3.5}$$

This latter may be estimated once an estimate of the error spectrum, $f_{\varepsilon\varepsilon}(\lambda)$, is available. In the manner of Chap. 6, Brillinger (1975) an estimate of the error spectrum is provided by

$$f_{YY}^T(\lambda)-|f_{YM}^T(\lambda)|^2/f_{MM}^T(\lambda). \tag{3.6}$$

Approximate confidence intervals may now be constructed for the value $s(t)$.

From expression (2.3) it is apparent that the estimate $s_0^T(t)$ has variance proportional to

$$\int f_{\varepsilon\varepsilon}(\alpha) f_{MM}^T(\alpha)d\alpha \doteq \frac{2\pi}{P}\sum_{p=0}^{P-1} f_{\varepsilon\varepsilon}\left(\frac{2\pi p}{P}\right) f_{MM}^T\left(\frac{2\pi p}{P}\right).$$

Using the statistic (3.6) this variance may be estimated. It must be remembered however that the estimate $s_0^T(t)$ is appropriate only when the σ_j are spaced sufficiently far apart, in contrast to the estimate $s_1^T(t)$ that is valid generally.

In practice an experimenter will often be willing to assume that the function $s(t)$ vanishes for $t<0$. The estimate $s_1^T(t)$ involves no such assumption in its derivation. Wiener (1949) has developed factorization procedures for handling functions that vanish for negative arguments. These procedures may be paralleled with data in the present situation in the manner that Bhansali (1973) dealt with the case of a pair of discrete time series. Bhansali's results suggest that there is no real improvement, in terms of variability, of proceeding to the more complicated estimate however.

IV. Further Remarks

On occasion it will be of interest to ask whether any response has been evoked at all by applying the stimuli. This question may be formalized to asking whether $s(t)$ of (1.1) or whether $S(\lambda)$of (3.1) is identically zero. A useful test statistic is provided by the coherence function

$$|R_{YM}^T(\lambda)|^2=|f_{YM}^T(\lambda)|^2/[f_{YY}^T(\lambda)f_{MM}^T(\lambda)]. \tag{4.1}$$

In the case that $S(\lambda)=0$ and that the f^T are constructed as in the previous section, the upper 100α% point of the distribution of $|R^T|^2$ is given approximately by $1-(1-\alpha)^{1/(K-1)}$, and tests may be constructed.

In EEG analysis several series, $Y_j(t)$, $j=1,\ldots,J$ are recorded. An interesting question that comes up in this case is whether the evoked responses, for two symmetrically placed leads, are the same (see John and Thatcher, 1977). A model for the situation is

$$\begin{aligned} Y_1(t)&=\sum_j s(t-\sigma_j)+\varepsilon_1(t)\\ Y_2(t)&=\sum_j s(t-\sigma_j)+\varepsilon_2(t)\end{aligned} \tag{4.2}$$

leading to

$$Y(t)=Y_1(t)-Y_2(t)=\varepsilon_1(t)-\varepsilon_2(t)$$

which may be examined using the statistic of (4.1).

A question that comes up is how the stimulation times σ_j should be chosen in practice. We have seen that if the average evoked response is to be taken as the estimate, then the σ_j should be taken sufficiently far apart. Examination of expression (2.2), [or alternatively of expression (3.3)], indicates that the average evoked response will be satisfactory when the point process $M(t)$ is such that $dP_{MM}(u)\alpha\delta(u)du$, with $\delta(u)$ the Dirac delta function. This will be the case if $M(t)$ is a realization of a stationary Poisson process. The expression for the variance of the estimate, (2.4), indicates that the rate of the Poisson, p_M, should be taken as large as possible. In the case that the shape of the error spectrum, $f_{\varepsilon\varepsilon}(\lambda)$, is known one can consider choosing M to minimize the variance (3.5) subject to

$$\sum_{p=0}^{P-1} f_{MM}^T\left(\frac{2\pi p}{P}\right)$$

given (as essentially corresponds to given total number of points, $M(T)$.) The solution to this problem is to take $M(t)$ such that $f_{MM}^T(\lambda)\alpha[f_{\varepsilon\varepsilon}(\lambda)]^{1/2}$. One would further take the overall number of points, $M(T)$, to be as large as possible.

Of course, the above remarks are based upon the assumption that the model (1.1) is reasonable. It will often be the case that if the σ_j are too close to each other, then the response is no longer linear, in the manner of (1.1). The system is thrown into a non-linear domain of operation.

Albrecht and Radil-Weiss (1976), Indra et al. (1976), Albrecht et al. (1977) consider a different model for the investigation of evoked responses. It may be viewed as of the form of (1.1), but with $s(t)$ a stationary process uncorrelated with the noise process $\varepsilon(t)$. The stylized forms of evoked response obtained in practice suggest that stationarity will not often prove a reasonable assumption.

Appendix

The degree of dependence of a collection of random variables is conveniently measured by their joint cumulants, see Brillinger (1975) for example. Given the random variate $\{Y_1, \ldots, Y_I\}$ denote its joint cumulant of order I by $\mathrm{cum}\{Y_1, \ldots, Y_I\}$. (In the case of $I=2$, this is the covariance of Y_1 with Y_2.)

In connection with the noise process $\varepsilon(t)$ set

$$c_{\varepsilon\ldots\varepsilon}(u_1, \ldots, u_I) = \mathrm{cum}\{\varepsilon(t+u_1), \ldots, \varepsilon(t+u_I), \varepsilon(t)\} .$$

We can now state,

Assumption I. $\varepsilon(t)$, $-\infty < t < \infty$, is a stationary process, continuous in mean, whose cumulant functions exist and satisfy

$$|c_{\varepsilon\ldots\varepsilon}(u_1, \ldots, u_I)| < L_I(1+u_1^2)^{-1} \ldots (1+u_I^2)^{-1}$$

for some finite L_I, $I=1, 2, \ldots$.

This assumption has the implication that values of the process at some distance from each other are only weakly dependent, statistically.

In connection with the point process we require.

Assumption II. $\sigma_1, \sigma_2, \ldots$ is an increasing sequence of positive numbers with the properties

$$|M(s)-M(t)| < A + B|s-t|$$

$0 \leq s$, $t < \infty$ for some finite A, B. Also the limits (1.5), (1.6) exist for almost all u.

References

Albrecht, V., Radil-Weiss, T.: Some comments on the derivation of the Wiener filter for average evoked potentials. Biol. Cybernetics **24**, 43–46 (1976)

Albrecht, V., Lansky, P., Indra, M., Radil-Weiss, T.: Wiener filtration versus averaging of evoked responses. Biol. Cybernetics **27**, 147–154 (1977)

Bhansali, R.J.: Estimation of the Wiener filter. Bul. Internat. Statist. Inst. **39**, 82–88 (1973)

Brillinger, D.R.: Estimation of the mean of a stationary time series by sampling. J. Appl. Prob. **10**, 419–431 (1973)

Brillinger, D.R.: Cross-spectral analysis of processes with stationary increments including the stationary $G/G/\infty$ queue. Ann. Probab. **2**, 815–827 (1974a)

Brillinger, D.R.: Fourier analysis of stationary processes. Proc. IEEE **62**, 628–1643 (1974b)

Brillinger, D.R.: Time Series: Data Analysis and Theory. New York: Holt, Rinehart and Winston 1975

Brillinger, D.R., Bryant, H.L., Segundo, J.P.: Identification of synaptic interactions. Biol. Cybernetics **33**, 213–228 (1976)

Donchin, E., Lindsley, D. (eds.): Average evoked potentials: Methods, results and evaluations. Washington: National Aeronautics and Space Administration 1969

Indra, M., Albrecht, V., Lansky, P., Radil-Weiss, T.: Average evoked brain potential comparison on the basis of spectral and coherency functions. Biol. Cybernetics **24**, 103–110 (1976)

John, E.R., Thatcher, R.W.: Functional neuroscience, Vol. 2. New York: Halsted 1977

Wiener, N.: The extrapolation, interpolation and smoothing of stationary time series with engineering applications. New York: Wiley 1949

Received: August 7, 1978

David R. Brillinger
Department of Statistics
University of California
Berkeley, CA 94720, USA