1

C.R. Math. Rep. Acad. Sci. Canada - Vol. VIII, No. 1, Feb. 1986 fév.

WHAT DO SEISMOLOGY AND NEUROPHYSIOLOGY HAVE IN COMMON? - STATISTICS!

David R. Brillinger

F.R.S.C.

Dedicated to my teachers at the University of Toronto and the University of Toronto Schools

1. Introduction

Seismology is the branch of science concerned with the investigation of earthquakes and related phenomena. Its goals include: learning about the Earth's and planets' interior composition and predicting the time, size and location of future earthquakes. In contrast, *neurophysiology* is the branch of science concerned with how the elements of the nervous system develop, function and work together. Its goals include: explaining notions like memory, emotion, learning, sleep, expectation and, less heroically, how individual neurons respond to stimuli, transmit information and change with environment. The definitions may make these two fields seem remote from each other. In point of fact, however, they are intimately tied together through use of a common methodology - statistics.

Statistics is the science concerned with the collection and analysis of numerical information (data) in order to answer questions wisely. It is characterised by an interplay between axioms and data and, in particular, is concerned: with making statements that go beyond the data collected (inferences), with explanation and understanding, with prediction and control, with discovery and application, with justification and classification. These concerns are patently common to seismology and neurophysiology - hence a connection.

Throughout much of my career I have collaborated with seismologists and neurophysiologists. In this article I would like to present some examples of statistical concepts and techniques that I have found myself making use of in problems arising from *both* seismology and neurophysiology.

2. Some Statistical Background

"What's the use of their having names," the Gnat said, "if they won't answer to them? "No use to them," said Alice, "but its useful to the people that name them, I suppose. If not, why do they have names at all?"

Alice's Adventures in Wonderland

The statistician approaches a substantive problem with knowledge and experience concerning a particular collection of concepts and techniques. These tools often incorporate a notion of randomness and have proven

pertinent to problems arising in a broad range of scientific, technological and social fields. Brief descriptions of some of those pertinent to our examples follow. In these examples, dynamics and time will be central. This has affected the choice of the statistical apparatus highlighted.

It will be taken that the notion of a random element, ω, is given. A *time series* is a random real-valued function, $Y(.,\omega)$, of a real or integer-valued random variable t, usually referred to as *time*. A *point process* is a random, nonnegative, integer-valued measure, $N(.,\omega)$. The values of these quantities, for a particular realisation of ω, are typically denoted by $Y(t)$, $Y(x,y,t)$, $N(I)$ respectively with I referring to a measureable set in the last case and with dependence on ω suppressed.

If $X(\omega)$ denotes a particular random variable and if $P(.)$ denotes the (probability) measure of ω, then in what follows $E\{X\}$ will denote $\int X(\omega)P(d\omega)$. *Moment functions* and *product densities* are of substantial use in discussing time series and random processes. When they exist, these take the forms

$$E\{Y(t_1)...Y(t_K)\} = m_{Y...Y}(t_1,...,t_K)$$

and

$$E\{N(dt_1)...N(dt_K)\} = p_{N...N}(t_1,...,t_K)dt_1...dt_K \qquad (1)$$
$$= Prob\{N(dt_1)=1,...,N(dt_K)=1\}$$

respectively with the dt_k distinct. Given data, typically assumed to be part of a realisation of a random process, useful estimates may be constructed for these quantities in a broad class of instances, particularly when the process involved is *stationary*, that is when its probabilistic properties are invariant under simple translations of time. In the stationary case, the process has a spectral representation, eg.

$$Y(t) = \int \exp\{i\lambda t\}Z(d\lambda)$$

or

$$N(I) = \int[\int_I \exp(i\lambda t)dt]Z(d\lambda)$$

$Z(.)$ being a random function with orthogonal increments. A further useful parameter, the *power spectrum*, is now given by

$$E\{Z(d\lambda)\overline{Z(d\mu)}\} = \delta(\lambda-\mu)f(\lambda)d\lambda d\mu$$

$\lambda\neq0$, when it exists. (Here $\delta(.)$ denotes the Dirac delta 'function'.)

In their work statisticians make continual use of *stochastic models*. These are analytic idealizations of real-world circumstances containing some random element. They tie the observables to the phenomenon of concern and are designed to lead to a broad variety of inferences concerning the phenomenon. Stochastic models often take the form of *systems*, that is mappings carrying functions, measures and the like over into other functions or measures. Stochastic models and systems usually involve unknown *parameters* (these may be finite or infinite dimensional) and the *estimation* (or identification) *problem* is to attach reasonable empirical values to these unknowns given observational or experimental data. A central role in this endeavour is played by the *likelihood function*. In simple terms, if θ denotes the unknown parameter the likelihood is the Radon-Nikodyn derivative of the probability measure of the data relative to some known measure, viewed as a function of θ. In many circumstances one has to work with an approximation to the likelihood, perhaps derived via an asymptotic method. One seeks a θ that is physically interpretable whenever possible.

3

D.R. Brillinger

schnological and
ur examples fol-
his has affected

is given. A *time*
eger-valued ran-
a random, non-
quantities, for a
y,t), $N(I)$ respec-
and with depen-

enotes the (pro-
note $\int X(\omega)P(d\omega)$.
se in discussing
the forms

(1)

ed to be part of
constructed for
hen the process
es are invariant
e process has a

A further useful

.)

chastic models.
ontaining some
of concern and
concerning the
ns, that is map-
ier functions or
known *parame-
estimation* (or
es to these unk-
al role in this
ms, if θ denotes
lerivative of the
ire, viewed as a
approximation
One seeks a θ

3. Example I - the Autointensity Function

A central entity in the method by which nerve cells communicate is the spike train. If a microelectrode is inserted into the axon (that is the output component) of a neuron, a changing voltage is recorded. This time series is made up of essentially identical spikes, or pulses, repeating at generally irregularly spaced times. Supposing these spikes to occur at times τ_k, $k=0,\pm1,\pm2,...$ one may define a counting measure via $N(I)$ = the number of τ_k in the interval I (of the real line). In various circumstances it seems reasonable to talk of probabilities of events such as: there is a spike (or point) in the small interval $(t,t+dt)$ or there is a spike in the interal $(t,t+dt)$ *and* in the interval $(t+u,t+u+du)$. If one is willing to view a given spike train as part of a realisation of a stochastic point process, then these probabilities correspond to product densities as defined by (1) above. In the stationary case it is convenient to define the *rate*

$$h_N = Prob\{point\ in\,(t,t+dt)\}/dt$$

and the *autointensity function*

$$h_{NN}(u) = Prob\{point\ in\,(t+u,t+u+du)\,|\,point\ at\ t\}/du$$

Given a stretch of data, these two parameters may be estimated by n/T and

$$\#\{|\tau_k-\tau_j-u|\,<b/2\}/nb$$

respectively where b is a small bin width and where n is the number of points τ_k observed in the time period, T, of observation. The autointensity function is an important descriptor of the behaviour of a firing neuron. For example in the case of a pacemaker cell, the autointensity is essentially 0 except when u is near some multiple of the (constant) interval between spikes. If the neuron is firing completely at random, the autointensity will be essentially constant. If bursting of firing is occurring, then $h_{NN}(u)$ will be essentially high for small to moderate $|u|$ and drop down to h_N as $|u|$ increases. If bursting is taking place at regular intervals with, for example, an accelerando pattern within bursts, then $h_{NN}(u)$ will show mass broadly near 0 and also for u near multiples of the interval between bursts. From an estimate of the autointensity of a spike train, the behaviour of a nerve cell may be described and classified. A broad variety of experimental examples may be found in Bryant *et al.* (1973) and Brillinger *et al.* (1976).

Earth scientists, engineers, government officials and the like are interested in the *seismicity* of the habitat, that is the timing, location and strength of earthquakes occuring in their region of interest. They are further interested in earthquake prediction and corresponding risk assessment. The sequence of times of earthquake occurrence in a given region may be viewed as corresponding to part of a realisation of a stochastic point process. The rate of the process tells how many earthquakes may be expected in a unit time interval. The autointensity provides a means of describing future probabilities of earthquakes given the past record. For example, if earthquakes tend to recur periodically, then the autointensity will have the pacemaker shape described above. If earthquakes tend to occur in clusters, the shape will be as for a nerve cell firing in bursts. If the times of earthquakes are totally random, the $h_{NN}(u)$ will be essentially constant. Various empirical examples are given in Vere-Jones (1970). Data of China for the period 1000 A.D. to the present, is studied in Lee and Brillinger (1979) by means of a technique developed to handle the incompleteness of the early records.

In order that hypotheses and models may be checked, some indication of the sampling uncertainty of the estimates is needed. Also, a parameter that

proves to be even more useful than the autointensity defined above is the crossintensity. It gives the probability of a point of one type occurring given that a point of another type has occurred, say u time units, earlier. Various results related to these last ideas may be found in Brillinger (1975).

4. Example II - Probit Analysis

A conceptual model for the firing of a neuron is the following: input to the nerve cell leads to (postsynaptic) electric current genesis. This current flows to a trigger zone, being filtered in the course of its passage. When the voltage level at the trigger zone exceeds a threshold value, the nerve cell fires. This process may be specified analytically as follows. Let $U(t)$ denote the voltage (membrane potential) at the trigger zone at time t. Let $B(t)$ denote the time elapsed since the last firing. Let $X(t)$ denote the (measured) input to the cell. Then, assuming linearity and time invariance

$$U(t) = \int_0^{B(t)} a(u)X(t-u)du$$

for some response function $a(.)$. Suppose the threshold level at time t has the form $\alpha + \epsilon(t)$, with $\epsilon(t)$ a normal variate of mean 0 and variance 1. Then, given $U(t)$, the probability the neuron fires at time t is $\Phi(U(t)-\alpha)$, with $\Phi(.)$ denoting the standard normal cumulative function. Supposing the data to be recorded at times $t=0,1,2,...,T-1$ and $Y(t)$ to be observed and defined to be 1 if a spike occurred in the immediately preceding interval and to be 0 otherwise, the likelihood of the data may be written

$$\prod_{t=0}^{T-1} \Phi(V_t - \alpha)^{Y(t)}[1 - \Phi(V_t - \alpha)]^{1-Y(t)}$$

with

$$V_t = \sum_{u=0}^{B(t)-1} a(u)X(t-u)$$

The unknowns, $a(.)$, α, may be estimated by maximising the likelihood. Once the estimates have been obtained, the model may be checked to an extent by comparing the empirical firing probability with the fitted. This is done for a variety of inputs and neurons in Brillinger and Segundo (1979).

The generally agreed description of earthquake genesis is the following: earthquakes are due to faulting. Specifically a crack initiates at a point and spreads out to form a fault plane. As the crack passes a given point, slip takes place on the fault plane resulting in a stress drop and the radiation of seismic waves. The ground is initially compressed and dilated around the focus of the earthquake. The pattern of compressions and dilations is preserved in the seismic waves radiated out and may be observed in the seismograms of stations detecting the event. Further, given the orientation of the fault plane (usually expressed by three angles) there is a formula for the theoretical relative amplitude of the signal arriving at a given station. Data then consists of the following: the estimated focus of an earthquake, the locations of the stations recording it, and whether each station recorded compression or dilation (corresponding to whether the first motion noted was positive or negative). The problem is to estimate the orientation of the fault plane. This information is useful for understanding the physical nature of the Earth in general and for seismic risk assessment in particular.

Let A_j denote the theoretical amplitude for station j. Let t_j denote the arrival time of the seismic signal, $s_j(.)$. Then $s_j(t_j)=\alpha A_j$ for some α. Further the seismogram may be written $Y_j(t)=s_j(t)+\epsilon_j(t)$ with $\epsilon_j(t)$ a noise series. Now the

probability that the first motion is positive, assuming $\epsilon_j(t_j)$ normal mean 0 variance σ^2, is

$$Prob\{Y_j(t_j) > 0\} = Prob\{\epsilon_j(t_j) > -s_j(t_j)\} = \Phi(\rho A_j)$$

with $\rho = \alpha/\sigma$. Assuming the noises independent at different stations the likelihood function is given by

$$\prod_{j=1}^{J} \Phi(\rho A_j)^{Z_j}[1 - \Phi(\rho A_j)]^{1-Z_j}$$

with $Z_j = 1$ if the first motion is positive and $= 0$ if it is negative. The parameters may be estimated by maximising this likelihood. Once again the model may be checked by comparing an empirical with a fitted probability. The details of all this are given in Brillinger et al. (1980) and illustrated by computations with the great 1964 Alaskan earthquake and for some California events.

In fact the same computer program was employed to fit the nerve firing model and the first motion model, even though these two models had such totally different origins.

5. Example III - Average Evoked Response

Consider the linear time invariant system with input $X(t)$ and output $Y(t)$

$$Y(t) = \int a(t-u)X(u)du$$

Here $a(.)$ is referred to as the *impulse response*, because if the input $X(t)$ is taken as the Dirac delta function , then the output is $a(t)$. A broad variety of naturally occurring systems seem to be linear and time invariant in the above manner, to a good approximation. Prominent among these is the Earth's transmission of seismic (acoustic) waves, be they generated by earthquakes, explosions or other vibratory sources. This effect is highly useful in seismic exploration. Suppose an impulse of energy is input to the Earth in a region of interest. Part of this energy will be reflected back to the surface by subsurface geologic structures after time delays 'proportionate' to the depth of the structure (wherever there is a difference in acoustic impedance). If $Y(t)$ denotes the signal recorded by a sensor on the surface, then its peaks (really peaks of the impulse response $a(.)$) may be interpreted in terms of subsurface layering. From estimates of such 'reflectivity functions' along lines of shots, geologically interesting structures at depth may be inferred. In practice a single pulse at a location rarely proves incisive. Hence prospectors are led to replicate the pulses at times σ_m, $m = 1,...,M$. One can then form the *average evoked response* or stacked estimate

$$\frac{1}{M} \sum_{m=1}^{M} Y(u + \sigma_m)$$

as an improved estimate of $a(u)$, provided the σ_m are sufficiently far apart that the corresponding individual responses do not overlap. Neitzel (1958) presents the results for some early experiments of this type.

It has long been traditional to average numbers. The novelty in the present circumstance is that it is curves that are being averaged. Such averaging has proved to be crucial in studies of brain waves because of the fact that signals evoked by various sensory stimuli are much smaller than the ongoing noise. The stimulus may be auditory, visual, olfactory, somatosensory or gustatory in character. The data available for analysis consists of the ongoing electrencephalogram (EEG) observed at an array of locations on the skull and

the times of application of the stimuli. Quite a variety of questions arise concerning the evoked response phenomena. These include: Does a given stimulus in fact evoke a response? Do different stimuli elicit the same response? Does the same response occur at different sensors? Are the responses repeatable? If stimuli are reordered, is the response the same? Are the effects of different stimuli additive? How does the response depend on the stimulus intensity? Answers to these questions are complicated by the phenomena of: weak response, variability of response, occurrence of artifacts, among other things. The papers Brillinger (1981a,b) review the history of evoked response experiments and their analysis, describe a number of success stories concerning the technique and provide some formal answers to the preceding questions. In particular, the following class of procedures is proposed for dealing with the complications caused by the presence of artifacts.

Nowadays considerable statistical research effort is directed towards the construction of robust/resistant techniques, that is procedures that remain effective in the presence of bad data values or of long-tailed error distributions. The traditional average value (or sample mean) is a prime example of a nonresistant sample quantity. Its value may be shifted an arbitrarily large amount by merely shifting a single sample value. By contrast, the interquartile or mid- mean, that is the mean of the central 50% of the sample values, does not even involve the 50% most extreme sample values in explicit fashion and hence is highly unshiftable. In Brillinger (1981a) the following class of estimates was proposed for the evoked response case, with a discussion of computational procedures for both the live and dead time cases. These estimates may be computed automatically. Set $Y_m(u) = Y(u + \sigma_m)$ and

$$| Y - \theta |^2 = \int_0^V | Y(u) - \theta(u) |^2 du$$

where, in this last, it is assumed that the evoked response dies off after V time units and that $\sigma_{m+1} - \sigma_m > V$. As a resistant estimate consider $\hat{\theta}(.)$ satisfying

$$\hat{\theta}(u) = \sum_m W_m Y_m(u) / \sum_m W_m$$

with $W_m = W(| Y_m - \hat{\theta} | / \hat{\rho})$, $W(.)$ being a weight function having most of its mass near 0 and $\hat{\rho}$ an estimate of scale. As a generalization of the mid-mean above, one can consider

$$\hat{\theta}(u) = \sum' Y_j(u) / \beta M$$

with $\sum'$ denoting the summation over the βM smallest $| Y_m - \hat{\theta} |$. The statistical properties of this last estimate are studied in the Berkeley Ph.D. thesis of Folledo (1983).

6. Example IV - Decaying Cosines

After a great earthquake the whole earth rings like a bell, with the vibrations lasting for days sometimes. Because the Earth is a finite body, it can only resonate as a whole at certain discrete frequencies. Because the medium is dissipative, the vibrations eventually damp away. These phenomena are in accord with the equations of motion being linear with constant coefficients and in consequence having solutions

$$s(t) = \sum_k \alpha_k \exp\{-\beta_k t\} \cos(\gamma_k t + \delta_k)$$

$t > 0$, assuming initial condition of a Dirac delta function at 0. An observed seismogram will have the form $Y(t) = s(t) + \epsilon(t)$ with $\epsilon(.)$ denoting a noise series. The problem arising is how to estimate the unknown parameters, particularly

ons arise con-
)oes a given
cit the same
)rs? Are the
ie same? Are
se depend on
icated by the
:e of artifacts,
he history of
ber of success
iswers to the
:dures is pro-
of artifacts.
1 towards the
; that remain
:rror distribu-
: example of a
bitrarily large
the interquar-
ample values,
xplicit fashion
)wing class of
discussion of
s. These esti-

ff after V time
itisfying

)st of its mass
l-mean above,

The statistical
thesis of Fol-

vith the vibra-
: body, it can
:e the medium
nomena are in
.nt coefficients

An observed
a noise series.
rs, particularly

the β_k, γ_k. In Bolt and Brillinger (1979) the following solution is developed. Suppose the noise series $\epsilon(.)$ is stationary and *mixing*, (that is well-separated in time values are at most weakly dependent), then the Fourier transform values of lengthy time segments satisfy a central limit theorem, that is are asymptotically normal. In particular if

$$\epsilon_j = \sum_{t=0}^{T-1} \epsilon(t) \exp\{-i 2\pi j t / T\}$$

the values ϵ_j for $2\pi j/T$ near λ will be approximately independent complex normal variates with mean 0 and variance $2\pi T f_{\epsilon\epsilon}(\lambda)$, $f_{\epsilon\epsilon}(.)$ being the power spectrum of the series $\epsilon(.)$. Now one has $Y_j = s_j + \epsilon_j$, with s_j depending on the unknown parameters of interest. If one sets down the approximate likelihood function for the data here and works in a neighborhood of λ containing only one of the γ_k, then obtaining (approximate) maximum likelihood estimates comes down to minimizing

$$\sum_j |Y_j - s_j|^2$$

as a function of the unknown parameters. The details of this, as well as a procedure for checking the validity of the assumed form for $s(t)$, may be found in Bolt and Brillinger (1979). For example, it is found there that if a limiting process with $\beta_k = \phi_k/T$ as $T \to \infty$ is employed then

$$var \beta_k, var \gamma_k \approx T^{-3} 4\pi f_{\epsilon\epsilon}(\gamma_k) \alpha_k^{-2} I_0(\phi_k) J(\phi_k)^{-1}$$

where

$$I_l(\phi) = \int_0^1 u^l \exp(-2\phi u) du$$

and $J(\phi) = I_0(\phi) I_2(\phi) - I_1(\phi)^2$. The T^{-3} decrease of variance is initially surprising.

The decaying cosine model has also proven useful in neurophysiology. In the work of Freeman (1972, 75, 79) the olfactory system of rabbits has been studied via evoked response experiments. Freeman found that the averaged response could be well-fitted by the sum of a few decaying cosine terms. He developed a model involving spike to wave conversion, involving collections of constant coefficient second-order differential equations, involving feed forward and feedback and involving wave to pulse conversion. Various types of neurons and connections were postulated. He employed nonlinear regression, in the time domain, to estimate the unknowns. In one case involving two cosines he was led to view the larger wave as representing intracortical negative feedback and the smaller as representing another feedback loop. Of some interest in this type of work is what happens to the frequencies and the decay rates when the experimental conditions are altered.

7. Example V - System Identification / Deconvolution

Regression analysis is one of the more long standing and potent tools in the statistician's kit. There are variants, for time series and point process data, that have proved useful in studying both neurophysiological and seismological data. Let $M(.)$ and $N(.)$ be two stochastic point processes whose realisations are imagined to correspond to the spike trains of two given neurons. Consider modelling the rate of firing of one neuron as it is affected by the other, as follows

$$Prob\{N(dt) = 1 \mid M(.)\} = [\alpha + \int a(t-u) M(du)] dt$$

for some constant α and function $a(.)$. Supposing $\{M(.), N(.)\}$ to be a stationary point process, the above relationship leads to

$$f_{NM}(\lambda) = A(\lambda) f_{MM}(\lambda)$$

where $A(.)$ denotes the Fourier transform of $a(.)$ and $f_{NM}(.)$ denotes the *cross-spectrum* of $N(.)$ with $M(.)$. (In terms of the spectral representations of the two processes it is defined via $E\{dZ_N(\lambda)\overline{dZ_M(\mu)} = \delta(\lambda-\mu)f_{NM}(\lambda)d\lambda d\mu$, $\lambda \neq 0$.) Once estimates of the spectra $f_{NM}(.)$ and $f_{MM}(.)$ are at hand, the function $A(.)$ may be estimated and after it $a(.)$. There are various ways that such spectral estimates may be formed, see Brillinger (1975). A variety of examples for neuro-physiological data are presented in Brillinger *et al.* (1976). The strength of the postulated 'linear' relationship may be measured by the *coherence* function, $|R_{NM}(\lambda)|^2 = |f_{NM}(\lambda)|^2/f_{NN}(\lambda)f_{MM}(\lambda)$. It lies between 0 and 1. In this last reference a variant of the coherence is employed to untangle the issue of how some given triples of nerve cells are causally connected.

Calculations similar to the above are also extremely useful in the seismic exploration case. We indicate how they may be employed to design an effective probing signal there. Consider again the system

$$Y(t) = \int a(t-u)X(u)du \qquad (2)$$

for the time series $Y(.),X(.)$. Let $m_{YX}(.)$ denote the convolution of Y with X, in some sense, and let $m_{XX}(.)$ denote the convolution of X with X. Then from (2) one has

$$m_{YX}(u) = \int a(u-v)m_{XX}(v)dv$$

Suppose one wishes an $X(.)$ such that $m_{YX}(u)$ is approximately $a(u)$. Let $f_{XX}(.)$ denote the Fourier transform of $m_{XX}(.)$, then the right-hand side above is

$$\int \exp\{i\lambda u\}A(\lambda)f_{XX}(\lambda)d\lambda$$

and one sees that for this to be $a(u)$ what is needed is that $f_{XX}(\lambda)$ be approximately 1 on the support of $A(.)$. Supposing $A(\lambda)$ to be approximately 1 on $\lambda_0 < \lambda < \lambda_1$ and 0 elsewhere, one is seeking $X(.)$ with $f_{XX}(\lambda)$ of the same character. An example of such an $X(.)$ is the chirp function

$$X(t) = \cos([\lambda_0 + (\lambda_1 - \lambda_0)\frac{t}{\tau}]t)$$

for $0 < t < \tau$. (This signal was first introduced formally by researchers in radar and is employed by bats in natural flight as well). In the seismic case it is input to the ground (for which λ_0 and λ_1 are known) repeatedly and the responses averaged. It should be remarked that in actual applications, substantial further processing is carried out to handle further physical effects present, such as wavefront spreading.

The tools of system identification are extremely powerful and may often be used to obtain indications of the mechanisms and states underlying some structure of interest. The above examples provide but a glimpse of the strength of the systems approach.

8. Example VI - the Analysis of Array Data

Array data is collected in both the earth and neuro- sciences. By array data is meant a collection of measurements of the form $Y(x_j,y_j,t)$, $j=1,...,J$ and $t=0,...,T-1$ with the (x_j,y_j), $j=1,...,J$ the coordinates of J sensors, and given j the measurements $Y(x_j,y_j,t)$, $t=0,...,T-1$ a segment of a time series. In the seismological case, the (x_j,y_j) refer to the locations of seismometers. In the

ie a station-

:s the *cross-*
s of the two
) Once esti-
$A(.)$ may be
pectral esti-
s for neuro-
strength of
'rence func-
In this last
ssue of how

the seismic
) design an

(2)

' with X, in
ien from (2)

) . Let $f_{XX}(.)$
ove is

be approxi-
mately 1 on
ie character.

iers in radar
tic case it is
dly and the
:ations, sub-
/sical effects

d may often
erlying some
npse of the

es. By array
$j=1,...,J$ and
and given j
ries. In the
:ters. In the

neurophysiological case, they refer to the grid locations of electrodes on the skull. Such data may often be reasonably viewed as part of a realisation of a planar-temporal (or spatial-temporal) random process $Y(x,y,t)$.

An important use of array data is the detection of propagating waves and the consequent estimation of their number, directions and velocities. For example in the seismic case one might have an array of (strong-motion) instruments located close to an earthquake fault. These instruments would be triggered by a sufficiently large event. In this case the source of the energy would be moving as the fault ripped. The seismologist would like to estimate the orientation of the fault and the rupture velocity. Bolt *et al.* (1982) discuss this problem and provide some elementary estimates based on data collected during an earthquake in Taiwan. They proceed by estimating the frequency-wavenumber spectrum of separate time segments of the data. The frequency-wavenumber spectrum has also been employed in the analysis of visual evoked response data, see Childers (1977). This last researcher first notes an apparent high velocity wave. After this wave has been 'removed', in a number of experiments he notes the occurrence of a pair of waves moving in opposite directions. His research is directed at developing a diagnostic procedure for various visual disorders and at obtaining insight concerning how the visual system functions.

Consider a planar-temporal wave of the form

$$Y(x,y,t) = \rho \cos(\alpha x + \beta y + \gamma t + \delta) + \epsilon(x,y,t)$$

with ρ,α,β,δ unknown constants and $\epsilon(x,y,t)$ a stationary noise process. The signal here is a plane wave propagating in direction θ given by $\tan \theta = \beta/\alpha$ with speed $\gamma/\sqrt{\alpha^2+\beta^2}$. In Brillinger (1985) the following maximum likelihood procedure is developed for detecting the presence of such a wave and for estimating its parameters.

Collect the J time series into a vector $Y(t) = [Y(x_j,y_j,t)]$. Let

$$Y_k = \frac{1}{T} \sum_{t=0}^{T-1} Y(t)\exp\{i2\pi k/T\}$$

and

$$M = \sum_k Y_k \bar{Y}_k^T$$

with the sum over $2\pi k/T$ near $\gamma = 2\pi k'/T$ say. Further set

$$S = \sum_k Y_k \bar{Y}_k^T - Y_{k'}\bar{Y}_{k'}^T$$

and let $B = [\exp\{i(\alpha x_j + \beta y_j)\}]$. The value $|BY_{k'}|^2$ is referred to as the *conventional statistic*. It may be expected to be large when $\rho \neq 0$ and it is evaluated at the 'correct' (α,β). The matrix S provides an estimate of the spectral density matrix of the series $\epsilon(.)$. Invoking a central limit theorem for the ϵ_k , an approximation may be set down for the likelihood function based on the Y_k (with $2\pi k/T$ near γ). It is found that the maximum likelihood detection statistic, given (α,β), is

$$\bar{B}^T S^{-1} B / \bar{B}^T M^{-1} B - 1 \qquad (3)$$

with null distribution $(K-J)^{-1}$ times an F distribution with degrees of freedom 2 and $2(K-J)$. It is further found that the maximum likelihood estimates of α and β are the coordinates of the maximum of the detection statistic. It is often convenient to prepare contour plots of the statistic (3) as a function of (α,β). An example of this is given in Brillinger (1975).

9. Discussion

In this article we have presented a number of examples, drawn mostly from our own experience, showing the use of the same statistical technique in the rather separate sciences of seismology and neurophysiology. It now seems appropriate to ask what, if anything, have the *three* sciences - statistics, seismology, neurophysiology - gained from each other as a result of connections albethey indirect? Having in mind a broader class of examples then those discussed in this paper, one can say that: i) statistics is richer for having been led to develop and study various novel methods to handle specific problems arising in seismology or neurophysiology, ii) both seismology and neurophysiology are the richer for the other's field having generated a problem for the statistician to abstract sufficiently that the result's applicability to their field became apparent, iii) either seismology or neurophysiology benefit from a statistical formulation because various of their problems seem necessarily to need to be stated in terms of probabilities (eg. neither neuron firings nor earthquakes seem predictable) and because these fields need procedures to validate results and to fit conceptual models. The methods of statistics often lead to important insight and understanding in substantive problems.

It may be remarked that the applicability of statistical procedures to these two substantive fields has further grown in direct consequence of their move to greater quantification and digital data collection.

Acknowledgement

References

[1] Bolt, B. A. and D. R. Brillinger (1979). Estimation of the uncertainties in eigenspectral estimates from decaying geophysical time series. *Geophys. J. R. astr. Soc.* **59** , 593-603.

[2] Bolt, B. A., Y. B. Tsai, K. Yeh and M. K. Hsu (1982). Earthquake strong motion recorded by a large near-source array of digital seimographs. *Earthquake Eng. Structural Dynam.* **10** , 561-573.

[3] Brillinger, D. R. (1975). Statistical inference for stationary point processes. pp. 55-99 in *Stochastic Processes and Related Topics* (Ed. M. L. Puri). Academic, New York.

[4] Brillinger, D. R. (1981a). Some aspects of the analysis of evoked response experiments. pp. 155-168 in *Statistics and Related Topics* (Eds. M. Csorgo *et al.*) North-Holland, Amsterdam.

[5] Brillinger, D. R. (1981b). The general linear model in the design and analysis of evoked response experiments. *J. Theoret. Neurobiol.* **1** , 105-119.

[6] Brillinger, D. R. (1985). A maximum likelihood approach to frequency-wavenumber analysis. *IEEE Trans. Acoust. Speech Signal Proc.* **ASSP-33.**

[7] Brillinger, D. R., H. L. Bryant and J. P. Segundo (1976). Identification of synaptic interactions. *Biol. Cybernetics* **22** , 213-228.

[8] Brillinger, D. R. and J. P. Segundo (1979). Empirical examination of the threshold model of neuron firing. *Biol. Cybernetics* **35** , 213-228.

[9] Brillinger, D.R., A. Udias and B. A. Bolt (1980). A probability model for regional focal mechanism solutions. *Bull. Seismol. Soc. America* **70** ,

drawn mostly
l technique in
It now seems
s - statistics,
ilt of connec-
xamples then
ner for having
specific prob-
gy and neuro-
a problem for
bility to their
benefit from a
necessarily to
on firings nor
procedures to
statistics often
ems.

dures to these
of their move

tional Science

ncertainties in
ries. *Geophys.*

hquake strong
seimographs.

oint processes.
. M. L. Puri).

oked response
ics (Eds. M.

ne design and
obiol. **1** , 105-

to frequency-
Proc. **ASSP-**

lentification of

lination of the
3-228.
ility model for
America **70** ,

149-170.

[10] Bryant, H. L., A. Ruiz Marcos and J. P. Segundo (1973). Correlations of neuronal spike discharges produced by monosynaptic connections and by common inputs. *J. Neurophysiol.* **36** , 205-225.

[11] Childers, D. G. (1977). Evoked responses: electrogenesis, models, methodology, and wavefront reconstruction and tracking analysis. *Proc. IEEE* **65** , 611-626.

[12] Folledo, M. (1983). *Robust/Resistant Mthods in the Estimation of the Evoked Response Curve*. Ph.D. Thesis, University of California, Berkeley.

[13] Freeman, W. J. (1972). Linear analysis of the dynamics of neural masses. *Ann. Rev. Biophysics and Bioeng.* **1** , 225-256.

[14] Freeman, W. J. (1975). *Mass Action in the Nervous System.* Academic, New York.

[15] Freeman, W. J. (1979). Nonlinear dynamics of paleocortex manifested in the olfactory EEG. *Biol. Cybernetics* **35** , 21-37.

[16] Lee, W. H. K. and D. R. Brillinger (1979). On Chineese earthquake history - an attempt to model an incomplete data set by point process analysis. *Pageoph.* **117**, 1229-1257.

[17] Neitzel, E. B. (1958). Seismic reflection records obtained by dropping a weight. *Geophysics* **23** , 58-80.

[18] Vere-Jones, D. (1970). Stochastic models for earthquake occurrence (with Discussion). *J. Royal. Statist. Soc.* **B 32** , 1-62.

Department of Statistics
University of California
Berkeley, Ca. 94720