A STOCHASTIC MODEL FOR THE INTERPRETATION OF CLINICAL TRIALS

BY GEORGE H. WEISS† AND MARVIN ZELEN‡

UNIVERSITY OF MARYLAND AND MATHEMATICS RESEARCH CENTER, U.S. ARMY

Communicated by J. Neyman, September 12, 1963

1. Introduction.—There are several diseases in which possibly recurrent phases may be distinguished. Different investigators have proposed Markovian models to describe data dealing with the time-dependent phenomena associated with these diseases. We mention in particular the work of Fix and Neyman on cancer,¹ the

⁵ Bhaduri, A., and P. A. Srere, unpublished results.
¹² Abbreviation: Tris for tris(hydroxymethyl)aminomethane.
²⁶ P. F. V. Ward and R. A. Peters, [Biochem. J., 75, 661 (1961)] have shown that a single fluorocitric acid accumulates in vitro in rabbit kidney poisoned with monofluoroacetate. Two diastereoisomers are theoretically derivable through the action of citrate synthase. Such an accumulation would arise if the enzyme is able to catalyze the replacement of only one of the two nonequivalent α-hydrogens of fluoroacetate-CoA in the condensation with oxaloacetate.
³² Cahn, R. S., C. K. Ingold, and V. Prelog, Experientia, 12, 81 (1956).
work of Marshall and Goldhamer on the epidemiology of mental disease,\textsuperscript{2} and the
work of Alling on tuberculosis.\textsuperscript{4} All of these assume that the distribution of time
spent in an occurrence of a particular phase is negative exponential. In the following
paper we present a semi-Markov model for data analysis, in which it is possible
to consider any distribution for stay in a given phase. The theory has been applied
to data on victims of both acute myelocytic and acute lymphocytic leukemia who
have received experimental drug therapy. In our application we have found that
gamma distributions give a convenient representation of the relevant probability
densities. Hence, any Markov theory will not be sufficient for the study of the
statistics relating to leukemia. We feel that the semi-Markov model presented
here has wide applicability to many clinical situations.

The advantage of having a model is that it guides the investigator in the types
and kinds of data to collect. Further, a model serves as a convenient frame of
reference for posing questions and suggesting further experiments. Too often the
"effectiveness" of a treatment is measured by the "success ratio." It is clear that
other factors are also important; e.g. (using the terminology of the acute leukemia
study), time in a remissive state, time to reach a remissive state, degree of toxicity,
time to failure (if ailment is a fatal one), time to complete cure, etc.

2. Formulation of the Model.—The health of a patient can be characterized at
any instant of time by being in one of a finite number of states. In the clinical
terminology, the patient may be in a relapse state, a remissive state, toxic state, etc.
These remissive and relapse states may be further classified by the degree of re-
mission or relapse and also by how many and what kinds of relapse or remissions
preceded the present state. In addition to these transient states, the patient may
have entered a terminal (or absorbing, in the language of Markov processes) state
such as failure, failure from other causes, cure, or patient lost.

We shall let $T$ denote the set of transient states and $A$ the set of absorbing states.
Also we define:

$$U_i(t) = \text{probability of being in state } i \text{ at time } t;$$

$$q_i(t) = \text{probability density function for a single time in state } i \epsilon T;$$

$$Q_i(t) = \int_{t^-}^t q_i(x)dx;$$

$$\omega_i(t)dt = \text{probability of leaving state } i (i \epsilon T) \text{ during the time interval } (t, t + dt);$$

$$p_{ij} = \text{probability of passing from state } i \text{ to } j, \text{ conditional upon leaving } i \text{ (by}
\text{ convention we set } p_{ii} = 0 \text{ for } i \epsilon T).$$

It will be convenient to define the vector and matrix analogues of the above
quantities. For this purpose we let

$$\mathbf{U}_T(t) = \text{column vector of } U_i(t), i \epsilon T;$$

$$\mathbf{U}_A(t) = \text{column vector of } U_i(t), i \epsilon A;$$

$$\mathbf{q}(t) = (\delta_{ij} q_i(t)), \text{ where } \delta_{ij} \text{ is the Kronecker delta;}$$

$$\mathbf{Q}(t) = (\delta_{ij} Q_i(t));$$

$$\mathbf{\omega}(t) = \text{column vector of } \omega_i(t); \mathbf{U}^0 = (U_i(0)\delta_{ij}), i \epsilon T;$$
\[ \mathbf{P} = \begin{bmatrix} A & T \\ I & 0 \end{bmatrix} \begin{bmatrix} R & \mathbf{P} \end{bmatrix}. \]

As a further convention, we denote the Laplace transform of any time-dependent function by the same function with argument \( s \) and an asterisk; i.e., \( \omega^*(s) = \int_0^\infty e^{-st} \omega(t) dt, \omega^*(s) = \int_0^\infty e^{-st} \omega(t) dt \).

Using familiar arguments in the theory of semi-Markov processes, one can obtain the following equations for \( \omega^*(s) \) and \( U_T^*(s) \):

\[
\begin{align*}
\omega^*(s) &= q^*(s)U^0 + q^*(s)\mathbf{P}'\omega^*(s) \\
U_T^*(s) &= Q^*(s)U^0 + Q^*(s)\mathbf{P}'U_T^*(s)
\end{align*}
\]

which yield the solutions

\[
\begin{align*}
\omega^*(s) &= [I - q^*(s)\mathbf{P}']^{-1}q^*(s)U^0 \\
U_T^*(s) &= Q^*(s)[I - \mathbf{P}'q^*(s)]^{-1}U^0.
\end{align*}
\]

Let \( U_A(t) \) denote the probability of being in an absorbing state at time \( t \). Then we have

\[ U_A^*(s) = \frac{1}{s} - \sum_{i \in T} U_i^*(s) \]

from which the moments of the time to reach absorption can easily be calculated by noting that the Laplace transform of the survivorship function, i.e., \( G(t) = \text{Pr.} \{ \text{failure time} > t \} \) is

\[ G^*(s) = \sum_{i \in T} U_i^*(s) = \mathbf{1}'U_T^*(s) \]

where \( \mathbf{1} \) is a column vector of appropriate dimension having all elements unity.

For the purpose of writing the moments associated with \( G(t) \), define

\[ m_i(k) = \int_0^\infty t^k q_i(t) dt, \quad M(k) = (m_i(k) \delta_{ij}) \]

\[ \tau = \lim_{s \to 0^+} \omega^*(s) = (I - \mathbf{P}^{-1})^{-1}U^0. \]

Note that \( \tau_i \) is the expected number of times state \( i \) is visited. Then the mean and variance of the time to reach absorption are

\[ E(t) = \lim_{s \to 0^+} G^*(s) = \mathbf{1}'M(1)(I - \mathbf{P}^{-1})^{-1}U^0 = \mathbf{1}'M(1)\tau \]

\[ \sigma^2 = -\lim_{s \to 0^+} \left\{ 2 \frac{dG^*(s)}{ds} + [G^*(s)]^2 \right\} = \mathbf{1}'\Sigma \tau \]

\[ + \mathbf{1}'M(1)\left\{ [2(I - \mathbf{P}^{-1})^{-1} - I - \mathbf{1}'D(\tau)M(1)] \right. \]

where

\[ \Sigma = M(2) - M^2(1) \]

\[ D(\tau) = \begin{pmatrix} \tau_1 & 0 \\ \tau_2 & \tau_1 \\ \vdots & \ddots \\ 0 & \cdots & \tau_n \end{pmatrix} \]
Relatively simple recursive relations can be derived for the higher moments. In addition to the results given so far, it is possible to develop a theory in which the p.d.f. \( q_i(t) \) is replaced by \( q_{ij}(t) \), i.e., in which the time spent in any state depends either on the succeeding or the following state. However, the results do not have the simplicity of those of equations (7) and (8). They will appear in a forthcoming publication. Another statistic of some interest in drug evaluation is the total time spent in a given transient state. This is important in the leukemia study, since it is desirable to prolong a patient’s life in a condition of remission rather than in a condition of active illness. Let us consider a single remissive state \( i \) and partition the transition matrix \( P \) as

\[
P = \begin{bmatrix}
A & i & T - i \\
I & O & O^T A \\
\delta & O & \beta \\
\gamma & \alpha & \Pi \\
\end{bmatrix}.
\]

Then the probability of entering state \( i \) at least once starting from state \( j \) is

\[
h_{ji} = [(I - \Pi)^{-1}\alpha]_j, \quad j \neq i
\]

\[
h_{ii} = \beta(I - \Pi)^{-1}\alpha.
\]

The first two moments of the total sojourn time in state \( i \) conditional on starting from state \( j \) are

\[
\mu_{ji}(1) = \frac{m_i(1)}{1 - h_{ii}} \delta_{ij} + \frac{m_i(1)h_{ji}}{1 - h_{ii}} (1 - \delta_{ij})
\]

\[
\mu_{ji}(2) = \left[ \frac{m_i(2)}{1 - h_{ii}} + \frac{2m_i^2(1)h_{ji}}{(1 - h_{ii})^2} \right] [\delta_{ij} + h_{ji}(1 - \delta_{ij})].
\]

3. Applications to Acute Leukemia Clinical Data.—A large number of clinical trials were conducted by the Acute Leukemia Group B (Frei) on patients having acute leukemia. In this section we will illustrate the application of our model to a portion of these data. A more complete discussion of the data will be given in another publication.

All patients entered these clinical trials while in a state of relapse. The data examined in this paper were for patients who were initially given Methotrexate (MTX). The patients either have responded and expired, or eventually reached a remissive state. A distinction was made between a partial and complete remission. If the patients did not show remission within the first 6 weeks, the MTX therapy was stopped, a period of two weeks was allowed to elapse, at the conclusion of which the surviving patients who were still in a state of relapse were given 6-mercaptopurine (6-MP). Further, if a patient who was in remission entered a relapsed state, the therapy was changed to the alternate drug (MTX or 6-MP). In this paper we will apply the model to data from 26 patients who achieved a state of remission within 8 weeks of entering these clinical trials. In our application we will be mainly interested in the distribution of the time to failure. Since, by definition, a remissive state is always followed by a relapse state, we will combine the sojourn time in a relapse state with the sojourn time of the remissive state which immediately preceded it. Another characteristic of this application is that the data
An investigation of the distribution of sojourn times within the various transient states showed that with the exception of the initial state, these distributions can be well approximated by gamma distributions, i.e.,

\[ \varphi(t) = \lambda e^{-\lambda t} \frac{t^{a-1}}{\Gamma(a)}, \quad t \geq 0, \quad a > 0. \]

The parameters \((a, \lambda)\) were estimated from the data by the method of maximum likelihood with the aid of the convenient tables of Wilk, Gnanadesikan, and Huyett. Table 1 summarizes the results of these calculations.

<table>
<thead>
<tr>
<th>State</th>
<th>(a)</th>
<th>(\lambda)</th>
<th>(m(1))</th>
<th>(\varphi)</th>
<th>Sample size</th>
</tr>
</thead>
<tbody>
<tr>
<td>(S_1) (first P. R.)</td>
<td>2.3</td>
<td>0.101</td>
<td>23.1</td>
<td>189</td>
<td>14</td>
</tr>
<tr>
<td>(S_2) (second P. R.)</td>
<td>2.9</td>
<td>0.122</td>
<td>23.6</td>
<td>215</td>
<td>7</td>
</tr>
<tr>
<td>(S_3) (first C. R.)</td>
<td>4.7</td>
<td>0.132</td>
<td>35.6</td>
<td>310</td>
<td>27</td>
</tr>
<tr>
<td>(S_4) (second C. R.)</td>
<td>15.5</td>
<td>0.463</td>
<td>33.6</td>
<td>90</td>
<td>7</td>
</tr>
</tbody>
</table>

* These estimates are based on all patients who went into a remission state whether they left the initial relapse state within eight weeks or after.

It remains to obtain estimates of the transition probabilities \(\mathbf{P}\) before the formulas of the preceding section can be applied. The observed relative frequencies (conditional on going to a remission state) observed in these clinical trials were used as estimates of the \((p_{ij})\). The numerical results are:

\[
\mathbf{P} = \begin{bmatrix}
S_0 & S_1 & S_2 & S_3 & S_4 & S_5 \\
S_0 & 1 & 0 & 10/26 & 0 & 0 & 16/26 & 0 \\
S_1 & 5/10 & 0 & 0 & 3/10 & 0 & 0 & 2/10 \\
S_2 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
S_4 & 9/16 & 0 & 0 & 4/16 & 0 & 0 & 3/16 \\
S_5 & 1 & 0 & 0 & 0 & 0 & 0 & 0
\end{bmatrix}
\]

The estimate of the transition matrix \(\mathbf{P}\) and the parameters of the gamma distribution summarize the relevant information with the exception of the sojourn time distribution for the initial relapse state \((S_1)\). Here the distribution is complicated by being truncated at the end of eight weeks. Since the data are summarized in units of a week, this distribution (conditional on a patient reaching a remissive state) was assumed to be a discrete distribution where \(p_n\) denotes the
probability of entering a remissive state after $n$ weeks in the initial relapse state. The observed relative frequencies were taken as the estimates for $p_n$ and were:

$$n = 2 \quad 3 \quad 4 \quad 5 \quad 6 \quad 7$$

$$p_n = [\frac{2}{26} \quad \frac{6}{26} \quad \frac{8}{26} \quad \frac{5}{26} \quad \frac{4}{26} \quad \frac{1}{26}]$$

A check on the model can be obtained by comparing the sample mean and variance of the time to reach failure with the theoretical formulas given in (7) and (8). The numerical results are:

<table>
<thead>
<tr>
<th>Data Mean</th>
<th>Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>48.4</td>
<td>49.7</td>
</tr>
<tr>
<td>Variance</td>
<td>515.</td>
</tr>
</tbody>
</table>

We now turn our attention to estimating the distribution of the time to reach the failure state. The probability density function for those patients reaching a remissive state can be calculated using (6). After some reduction, the result is:

$$g_0(t) = p_{20}p_{12} [\varphi_1 \ast \varphi_2] + p_{20} \left[ p_3 p_{15} (\varphi_1 \ast \varphi_2 \varphi_3) + p_{14} p_{43} (\varphi_1 \ast \varphi_3 \varphi_4) \right] + p_{50} p_{14} [\varphi_1 \ast \varphi_4] + p_{50} \left[ p_3 p_{15} (\varphi_1 \ast \varphi_5) + p_{14} p_{45} (\varphi_1 \ast \varphi_4) \right].$$

(The notation $\varphi_1 \ast \varphi_2$ denotes the convolution of $\varphi_1(t)$ and $\varphi_2(t)$, where $\varphi_1(t)$ denotes the frequency function of the initial relapse state.)

Note that $g_0(t)$ is made up of a mixture of distributions which involve convolutions of gamma distributions. For our purposes, the approximation of convolutions of gamma distributions by a gamma distribution with the same first two moments was deemed a sufficient approximation. Figure 2 contains a plot of $G_0(t) = \int_0^t g_0(x) dx$, along with the sample data.

We would like to thank Dr. Edmund Gehan of the National Institutes of Health for helpful discussions of this problem and his cooperation in providing us with records of the leukemia study.

* Sponsored in part by the Mathematics Research Center, U.S. Army, Madison, Wisconsin, under contract DA-11-022-ORD-2059, and in part by contract NONR 595(17).

† Present address: Rockefeller Institute, New York, New York.

‡ Present address: National Institutes of Health, Bethesda, Maryland.


ON DIFFERENTIAL SYSTEMS, II

BY VICTOR GUILLEMIN AND SHLOMO STERNBERG

DEPARTMENT OF MATHEMATICS, HARVARD UNIVERSITY

Communicated by Lars V. Ahlfors, September 30, 1963

In the study of the "general equivalence problem," one of the principal tools is the use of the structure function to reduce the group. (For a discussion of this technique for the case of G-structures, cf. ref. 3; the general setting, where the structure groups are allowed to vary from point to point, is the notion of differentiable groupoid of Ehresmann.1 We hope to discuss the foundations of the general theory in the near future.) Here we shall present some techniques applicable in the case that our structure is subordinate to a nonintegrable differential system. More precisely, we shall assume that the Lie algebra of the structural group at each point has the form

\[
\begin{pmatrix}
\hat{g}o0 \\
nh \end{pmatrix}
\]

where the lower-right-hand block corresponds to the invariant subspace \( V_1 \) of the tangent space \( V \) (and \( g \) is the restriction of the structure algebra to \( V_1 \)) where \( \hat{g} \) is the induced action of the structure algebra on \( V/V_1 \) and where we assume that the structure function \( \rho : V_1 \wedge V_1 \rightarrow V/V_1 \) of the differential system is surjective. (This is our nonintegrability assumption.) In particular, \( \hat{g} \) is a homomorphic image of \( g \) and we have \( P \circ \rho = \rho \circ \wedge^2(Y) \) for any \( Y \in g \). We wish to describe how to cut down \( h \).

Setting the structure function constant has the following effect on \( h \):

1. For any \( X \in h \), \( X \circ \rho \) lies in \( V_1 \otimes \wedge^2(V_1^*) \). Setting the structure function constant implies that

\[
X \circ \rho \text{ lies in } \partial(g \otimes V_1^*)
\]

(in the Spencer complex). Notice that if \( g^{(1)} = 0 \) (i.e., \( H^{5,0} = 0 \)) then we have a unique \( S \) with

\[
X \circ \rho = \partial S.
\]

(This is useful when \( g \) is a "small" subalgebra of \( \text{Hom} (V_1, V_1) \), in particular if \( V_1 \) is of small codimension \( >2 \).)

2. To each \( v \in V_1 \) and \( X \in h \) associate \((v \underline{\rho}) \circ X \in \text{Hom} (V/V_1, V/V_1) \). Then setting the structure function constant implies that

\[
(v \underline{\rho}) \circ X \in \hat{g}
\]

so that \( X \) gives \( T \in \hat{g} \otimes V_1^* \). (This is useful when \( V_1 \) is of large codimension.)

3. If \( g^{(1)} = 0 \) (so that (2) holds) then we have

\[
\]