The Annals of Statistics
1978, Vol. 6, No. 1, 101-111

MAXIMUM LIKELIHOOD ESTIMATION OF DOSE-RESPONSE FUNCTIONS SUBJECT TO ABSOLUTELY MONOTONIC CONSTRAINTS

BY H. A. GUESS AND K. S. CRUMP

National Institutes of Health and Louisiana Tech University

Statistical properties are derived for maximum likelihood estimates of dose-response functions in which the response probability is related to the dose by means of a polynomial of unknown degree with nonnegative coefficients. Dose-response functions of this form are predicted by the multistage model of carcinogenesis. We first establish necessary and sufficient conditions for strong consistency of the estimates. For these results no assumptions are made about the polynomial degree, so the number of coefficients to be estimated is effectively infinite. Under some additional assumptions, which do involve restrictions on the polynomial degree, we obtain the asymptotic distribution of the vector of maximum likelihood estimates about the true vector of polynomial coefficients. Because the coefficients are constrained to be nonnegative, the limiting distribution will generally not be normal.

1. Introduction. The procedures currently used to assess the human cancer risk associated with a potentially carcinogenic chemical often involve estimating the cancer risk to laboratory animals at very low doses of the chemical on the basis of data from experiments conducted at much higher doses. This problem of interpolating a dose-response relation in the low-dose range between the background (usually zero) dose and the other test doses is frequently referred to in the cancer literature as the *low-dose extrapolation problem*. The qualitative outcome of this extrapolation depends strongly on the statistical model used for estimating dose-response relations.

Several standard dose-response models are capable of providing a reasonably good fit to the data in the high dose range (e.g., where the cancer risk at the given dose exceeds the risk at background dose by about 10^{-1} or more) but yield risk estimates differing by several orders of magnitude in the very low dose range, where the increased risk over background is perhaps 10^{-6} or less. These large differences in estimated low-dose risks are due to differences in the assumed shapes of the dose-response curves in the low-dose range. Some models (e.g., the one-hit model) assume that the increased risk over background is roughly linear in dose in the low-dose range. Other models (e.g., the probit model and the multi-hit models) assume the low-dose part of the dose-response curve to be

Received July 1976; revised January 1977.

AMS 1970 *subject classifications*. Primary 62P10; Secondary 60F15, 62E20.

Key words and phrases. Absolutely monotonic functions, cancer risk estimation, constrained maximum likelihood estimates, convergence in distribution, dose-response functions, strong consistency, Tchebycheff systems.

extremely flat, so that the increased risk over background decreases with decreasing dose at a much faster than linear rate.

To help resolve questions regarding the shapes of dose-response curves for chemical carcinogens, we have developed an estimation technique which assumes a parametric dose-response curve form which is biologically plausible and is general enough to include both linear dose-response curves and curves that are much flatter than linear in the low-dose range. Our dose-response relation is taken from a general multistage carcinogenesis model developed by Armitage and Doll (1961), extended by Peto (1974) and by Crump, et al. (1976), and investigated by Brown (1976). In this multistage model it is assumed that some unknown number of different random events (e.g., mutations) must occur in a single cell before cancer is initiated. The occurrence times of the events are assumed to be exponentially distributed. The time from cancer initiation in a single cell until an observable tumor develops in a tissue is assumed not to depend on the dose rate. These assumptions can be used to show that the excess rate of cancer incidence over background should be given by the product of a function of dose and a function of the exposure duration. The function of dose is a polynomial with nonnegative coefficients, representing sums of products of the nonnegative rate coefficients for the exponential event time distributions. Expressions of this form for the excess rate of incidence over background lead to dichotomous response probabilities of the form

$$P\{Q(d)\} = 1 - \exp\{-Q(d)\}, \tag{1}$$

where $Q(d)$ is an absolutely monotonic polynomial of unknown degree in the dose rate d, that is

$$Q(d) = \sum_{i=0}^{\infty} q_i d^i, \tag{2}$$

where $q_i \geqq 0$ for all i and only a finite number of the q_i are nonzero. We call $P\{Q(d)\}$ the response probability at dose d or the cancer risk at dose d. When d_0 is a background dose level we call $P\{Q(d)\} - P\{Q(d_0)\}$ the increased risk over background at dose $d > d_0$.

In the experimental setting which we shall consider, a population of laboratory animals is partitioned into $n + 1$ treatment groups, the jth group consisting of N_j animals, $0 \leqq j \leqq n$. The animals in the jth treatment group are exposed to a dose rate d_j ($0 \leqq d_0 < d_1 < \cdots < d_n$) of a chemical carcinogen whose dose-response curve is to be estimated. Let X_j be the number of animals in the jth treatment group that are observed to have a positive response (e.g., a specific type of tumor). Each animal in the jth treatment group is assumed to have a probability $P\{Q(d_j)\}$ of responding, independently of all other animals in the experiment. Thus the likelihood of the experimental outcome is given by

$$L = \prod_{j=0}^{n} \binom{N_j}{X_j} [P\{Q(d_j)\}]^{X_j} [1 - P\{Q(d_j)\}]^{N_j - X_j}. \tag{3}$$

We wish to obtain and study maximum likelihood estimates of, first, all of

the coefficients q_i, and subsequently the function $P\{Q(d)\}$, $0 \leqq d \leqq d_n$. The present paper is devoted to the statistical properties of the estimates. Their mathematical properties (e.g., existence and uniqueness of the maximum likelihood estimates and algorithms for their computation) are described in Guess and Crump (1976). Applications of this work to the problem of obtaining risk estimates for specific carcinogens from experimental animal data are discussed in Crump, Guess and Deal (1977).

In the present paper we establish the asymptotic properties of the constrained maximum likelihood estimates as all of the sample sizes N_j increase with the test dose levels $d_0, \cdots, d_n$ remaining fixed. We first show that a necessary and sufficient condition for strong consistency of the maximum likelihood estimates of all of the q_i $(0 \leqq i < \infty)$ is that the true polynomial $Q(d)$ be determined uniquely within the class of absolutely monotonic polynomials by its values $Q(d_j)$ at the experimental doses d_j, $0 \leqq j \leqq n$. Necessary and sufficient conditions for the latter follow at once form the work of Krein and Rehtman (1959, Theorems 4.3 and 3.1) and Karlin and Studden (1966, Chapter 7). In essence, our result means that the statistical interpolation problem yields consistent estimators as long as the idealized deterministic interpolation problem has a unique absolutely monotonic solution, which it will have whenever the number of test doses is somewhat larger than the true number of stages in the multistage carcinogenic process. It is interesting to note that in a certain sense we have consistent estimators for infinitely many parameters using data from only a fixed finite set of doses. Although only finitely many q_i can be nonzero we do not know which these are and hence we admit an infinite number of maximum likelihood estimators, all but finitely many of which will converge to zero.

Once strong consistency established we obtain the asymptotic distribution of the vector of maximum likelihood estimates about the true coefficient vector. To prove this result we must assume the degree of the true polynomial Q to be known and we must limit the number of estimated coefficients to be no greater than the number of test doses. Since the true values of some of the coefficients to be estimated can be zero and since the estimates are constrained to be nonnegative, the limiting distributions will generally not be normal.

We summarize below some mathematical properties of the maximum likelihood estimates. These properties are needed for the statement and proof of the statistical results. For a more detailed discussion, the reader is referred to Guess and Crump (1976).

In many, but not all, cases there will exist a unique absolutely monotonic polynomial Q whose likelihood (3) is higher than that of any other absolutely monotonic polynomial. In some cases it is not possible to attain the supremum of the likelihood functions within the class of absolutely monotonic polynomials. An example of this is given in Guess and Crump (1976). The reason for this has to do with the shape of the dose-response curve near d_n, the highest dose tested. If the response frequency X_n/N_n at d_n lies above the general trend of responses

at the lower doses, then as absolutely monotonic polynomials of increasingly high degree are admitted into the class over which the maximum is taken, it becomes possible to fit the response at the highest dose with a high order term $c(d/d_n)^k$ which hooks up sharply at d_n and is vanishingly small at all lower doses. In the limit it is therefore possible for the maximum likelihood polynomials of increasingly high degree to approach a curve having a positive jump discontinuity at the highest dose tested and to approach a smooth curve at all doses $d < d_n$. The easiest way to accommodate any mathematical difficulties caused by this behavior in the high-dose range is to enlarge the class of functions over which the maximum is taken, so as to include functions which are absolutely monotonic polynomials on the interval $[0, d_n)$ and are permitted to have a nonnegative jump at d_n. Let

$$(4) \qquad AMJ = \{f \mid f(d) = Q(d) + q_\infty \delta_{d_n}(d),\ Q(d) = \textstyle\sum_{i=0}^{\infty} q_i d^i,\ q_i \geqq 0 \text{ for all } i,\ q_i = 0 \text{ for all but finitely many } i\}$$

where

$$\delta_{d_n}(d) = 0 \quad d < d_n$$
$$= 1 \quad d \geqq d_n .$$

Henceforth we shall assume the response probabilities to have the form

$$(5) \qquad P\{f(d)\} \equiv 1 - \exp\{-f(d)\}, \qquad f \in AMJ .$$

2. Strong consistency of the risk estimates. We establish necessary and sufficient conditions for strong consistency of the risk estimates for the case of biological interest, where the true function f in (5) is an absolutely monotonic polynomial, i.e., $q_\infty = 0$. Of course we must allow the maximum likelihood estimates of q_∞ to take on positive values in order to guarantee that the likelihood function will attain its supremum over the class of absolute monotonic polynomials. It will follow from the theorem that the maximum likelihood estimates of q_∞ will converge to zero with probability one. Thus we assume that the true response probabilities P are of the form (5) with $q_\infty = 0$ and we use data from tests conducted at a fixed finite set of doses $\{d_j\}_{j=0}^n$ to estimate $P\{f(d)\}$ for all d, $0 \leqq d \leqq d_n$. The type of limit theorem appropriate to this experimental setting is one in which the sample sizes N_j at the doses d_j are increased but the set of doses is fixed and finite.

It is easy to state a necessary condition for strong consistency: The finite set of true response probabilities, $[P\{f(d_j)\}]_{j=0}^n$, at the experimental doses must determine the true response function f uniquely within the class of absolutely monotonic polynomials. What is remarkable is that this condition is also sufficient. To prove this result we will need to introduce a definition and some notation.

DEFINITION. Let $f \in AMJ$ have the form (4). The *index* of f, written $I(f)$, is defined as follows. Let $k_1 < k_2 < \cdots < k_m$ be the subscripts of the coefficients q_i in the definition of f ($0 \leqq i \leqq \infty$) such that $q_i > 0$. Divide the numbers k_j beginning with k_m into groups in the following manner. If $k_m - k_{m-1} > 1$ or if

$k_m = \infty$ then the first group consists only of k_m; otherwise the first group consists of k_m and k_{m-1}. Remove the first group from the sequence $k_1 < k_2 < \cdots < k_m$ and repeat the above operation on the sequence which remains. Once all subscripts in the sequence have been grouped in this manner we compute $I(f)$ by counting one for each group consisting of ∞ alone or 0 alone and two for each other group.

This definition is adopted from Krein and Rehtman (1959, page 127). What we call $I(f)$ is what they call the index of the sequence of subscripts k_i such that $q_{k_i} > 0$. It is easy to see that $I(f)$ is greater than or equal to the number of positive coefficients in the representation of f in the form (4).

A function $g(d) = \sum_{i=0}^{\infty} a_i d^i$ will be said to be absolutely monotonic on the closed interval $[0, D]$ if $a_i \geqq 0$ for all i and $g(D) < \infty$. This definition is equivalent to the definition of Krein and Rehtman (1959, page 130).

LEMMA 1. *Let $f \in AMJ$ have the form* (4). *If $d_0 > 0$ then $I(f) \leqq n$ is a necessary and sufficient condition that Q be the only absolutely monotonic function on $[0, d_n]$ satisfying the conditions*

$$Q(d_j) = f(d_j) \qquad 0 \leqq j \leqq n-1$$
$$Q(d_n) \leqq f(d_n)\,. \tag{6}$$

If $d_0 = 0$ then a necessary and sufficient condition is that $I(f_R) \leqq n - 1$, where $f_R(d) = Q'(d) + q_\infty \delta_{d_n}(d)$.

PROOF. First take the case $d_0 > 0$. By Theorems 4.3 and 3.1 of Krein and Rehtman (1959) a necessary and sufficient condition that Q be the only absolutely monotonic function on $[0, d_n]$ satisfying (6) is that the index of the canonical representation of the set $\{f(d_j)\}_{j=0}^{n}$ in the form of their equation 3.4 be less than or equal to n where, translating their notation into ours, $c_j = f(d_j)$, $u_j(i) = d_j^{\,i}$ for $i < \infty$ and $u_j(\infty) = \delta_{jn}$. This index is the same as our $I(f)$.

When $d_0 = 0$ we have $q_0 = f(0)$ and the conditions (6) are equivalent to

$$Q'(d_j) = d_j^{-1}\{f(d_j) - f(0)\} \qquad 1 \leqq j \leqq n-1$$
$$Q'(d_n) \leqq d_n^{-1}\{f(d_n) - f(0)\}\,. \tag{7}$$

By the argument for the case $d_0 > 0$ a necessary and sufficient condition that Q' be determined uniquely by (7) within the class of functions absolutely monotonic on $[0, d_n]$ is that $I(f_R) \leqq n - 1$. Since $q_0 = f(0)$ holds automatically, uniquely determining Q' is equivalent to uniquely determining Q. This completes the proof.

REMARK. It follows from Karlin and Studden (1966, page 230) that if g is an absolutely monotonic *function* on $[0, d_n]$ satisfying (6) then there exists an absolutely monotonic *polynomial* satisfying (6). Thus Lemma 1 also provides a necessary and sufficient condition for the polynomial Q to be determined uniquely within the class of absolutely monotonic polynomials.

Let $(\Omega, \mathscr{B}, \mathscr{P})$ be a probability space on which all of the random variables

in the subsequent discussion are defined. For each vector $\mathbf{p} = (p_0, \cdots, p_n)$ with $0 \leqq p_j \leqq 1$, we define the likelihood function of $\mathbf{X} = (X_0, \cdots, X_n)$ with respect to $\mathbf{p}$ by

$$L(\mathbf{X}, \mathbf{p}) = \prod_{j=0}^{n} \binom{N_j}{X_j} p_j^{X_j} (1 - p_j)^{N_j - X_j} . \tag{8}$$

Let

$$S = [\mathbf{p} \mid p_j = P\{g(d_j)\},\ 0 \leqq j \leqq n;\ g \in AMJ] . \tag{9}$$

Then, for each $\omega \in \Omega$, a maximum likelihood estimate of f is defined to be any function f_ω satisfying

$$f_\omega \in AMJ \tag{10}$$

and

$$L[\mathbf{X}(\omega), \mathbf{P}\{f_\omega(\cdot)\}] = \sup [L(\mathbf{X}(\omega), \mathbf{p}) \mid \mathbf{p} \in S] . \tag{11}$$

Let $0 \leqq d_0 < d_1 < \cdots < d_n$ be a fixed finite set of doses and let $\{N_0^{(m)}, \cdots, N_n^{(m)}\}_{m=1}^{\infty}$ be a sequence of sample sizes such that

$$\lim_{m \to \infty} N_j^{(m)} = \infty \tag{12}$$

and

$$0 < \limsup_m \{N_j^{(m)} / N_i^{(m)}\} < \infty \tag{13}$$

for all $0 \leqq i, j \leqq n$. We are now ready to state and prove a strong consistency theorem for the maximum likelihood estimates.

THEOREM 1. *Let f be an absolutely monotonic polynomial satisfying the applicable uniqueness condition of Lemma* 1. *If the sample sizes $N_1^{(m)}, \cdots, N_n^{(m)}$ satisfy* (12) *and* (13) *and if f_m satisfies* (10) *and* (11) *for each m,*

$$f_m(d) = \sum_{i=0}^{\infty} q_{im} d^i + q_{i\infty} \delta_{d_n}(d) ,$$

then

$$\lim_{m \to \infty} q_{im} = q_i \quad \text{for} \quad 0 \leqq i \leqq \infty \quad \text{w.p. 1} , \tag{14}$$

$$\lim_{m \to \infty} \sup_{d \leqq d_n} |f_m(d) - f(d)| = 0 \quad \text{w.p. 1} , \tag{15}$$

$$\lim_{m \to \infty} \sup_{d \leqq d_n} |P\{f_m(d)\} - P\{f(d)\}| = 0 \quad \text{w.p. 1} , \tag{16}$$

and

$$\lim_{m \to \infty} E[P\{f_m(d)\}] = P\{f(d)\} \quad d \leqq d_n . \tag{17}$$

PROOF. By the result (i) of Rao (1973, page 356) it follows that

$$\lim_{m \to \infty} f_m(d_j) = f(d_j) \qquad 0 \leqq j \leqq n , \ \text{w.p. 1} . \tag{18}$$

Hence to prove (14) it suffices to show that $q_{im\omega} \to q_i$ for all $0 \leqq i \leqq \infty$ whenever $\omega \in \Omega$ is such that $f_{m\omega}(d_j) \to f(d_j)$ for $0 \leqq j \leqq n$.

Without loss of generality we will assume throughout the remainder of the proof that $d_n = 1$ so that the $q_{im\omega}$ are uniformly bounded by $\sup_m f_{m\omega}(1) < \infty$.

To prove (14) it suffices to show that the limit $\mathbf{q}_\omega = (q_{0\omega}, \cdots, q_{\infty\omega})$ of any componentwise convergent subsequence $\mathbf{q}_{m'\omega}$ is the vector $\mathbf{q}$. Since $0 \leqq d_j \leqq 1$ for $0 \leqq j \leqq n-1$, we have by the dominated convergence theorem

$$f_{m'\omega}(d_j) = \sum_{i=0}^{\infty} q_{im'\omega} d_j^{\,i} \to \sum_{i=0}^{\infty} q_{i\omega} d_j^{\,i}$$

and so by (18)

$$\sum_{i=0}^{\infty} q_{i\omega} d_j^{\,i} = f(d_j) \qquad 0 \leqq j \leqq n-1 . \tag{19}$$

By Fatou's lemma we have

$$\sum_{i=0}^{\infty} q_{i\omega} + q_{\infty\omega} \leqq \liminf \left(\sum_{i=0}^{\infty} q_{im'\omega} + q_{\infty m'\omega}\right) = \lim f_{m'\omega}(1) = f(1) \tag{20}$$

and hence

$$\sum_{i=0}^{\infty} q_{i\omega} \leqq f(1) . \tag{21}$$

By Lemma 1 there exists only one absolutely monotonic function on [0, 1] satisfying (19) and (21). This implies $q_{i\omega} = q_i$ for $0 \leqq i < \infty$. Using this fact in (20) and recalling that $q_{\infty\omega} \geqq 0$ and that $q_\infty = 0$ (because f is an absolutely monotonic polynomial by hypothesis) yields $q_{\infty\omega} = 0 = q_\infty$. This proves (14).

Since the $q_{im\omega}$ are uniformly bounded by $\sup_m f_m(1)$, the dominated convergence theorem and (14) together imply that

$$\lim_{m\to\infty} f_m(d) = f(d) \qquad \text{w.p. } 1 \tag{22}$$

for each $d \in [0, 1]$. Since f_m and f are bounded right continuous functions on [0, 1], (15) and (16) follow from (22) by the standard argument used to prove the Glivenko–Cantelli theorem. (See Chung (1968), pages 124–125.) The reasoning there applies to our case even though f_m and f are not probability distribution functions. Equation (17) follows from (16) since the terms in (16) are uniformly bounded. This completes the proof.

3. Asymptotic distributions. In this section we obtain the limiting distribution of the maximum likelihood estimates for the true vector of polynomial coefficients. Using this result it is straightforward to compute asymptotic distributions for other quantities of interest, such as the increased risk over background. To derive the limiting distribution we must restrict the number of coefficients that are allowed to be nonzero in maximizing the likelihood to belong to a fixed finite set whose cardinality is no greater than the fixed finite number of experimental doses.

The main idea of the proof is to use the Kuhn–Tucker conditions for the likelihood maximization problem to express the difference between the maximum likelihood estimate of the coefficient vector and the true coefficient vector as a function of an asymptotically normal vector. We can then deduce the limit theorem via a continuous mapping argument. For the proof we will need some additional notation.

Let T be a nonrandom fixed finite subset of $\{0, 1, \cdots, \infty\}$, let card (T) denote

the number of elements in T, and let

$$AMJ(T) = \{f \in AMJ \mid f(d) = \sum_{i=1}^{\infty} q_i d^i + q_\infty \delta_{d_n}(d),\ q_i \geqq 0 \text{ for all } i = 0, 1, \cdots, \infty \text{ and } q_i = 0 \text{ for } i \notin T\}. \tag{23}$$

In this section we will assume that the true function f is in $AMJ(T)$ and we will maximize the likelihood function over $AMJ(T)$ instead of over AMJ as in the previous section. For all results except Theorem 3 we will require that card $(T) \leqq n$ if both $d_0 = 0$ and $0 \notin T$ and that card $(T) \leqq n + 1$ otherwise. Using Descartes' rule of signs it is simple to show that this condition guarantees that a function $f \in AMJ(T)$ is determined uniquely within $AMJ(T)$ by the values $f(d_0), \cdots, f(d_n)$. Since we can have card $(T) = n + 1 = I(f)$, it can happen that these values determine f uniquely in $AMJ(T)$ but not uniquely in the larger set AMJ.

Take a sequence of sample sizes satisfying (12), let $N_m = (n + 1)^{-1} \sum_{j=0}^{n} N_j^{(m)}$ and let $\mathbf{N}_m$ be the vector $(N_0^{(m)}, \cdots, N_n^{(m)})$. In place of (13) we impose the stronger condition that

$$\lim_{m\to\infty} \left(\frac{N_j^{(m)}}{N_0^{(m)}} - r_j \right) (N_0^{(m)})^{\frac{1}{2}} = 0 \tag{24}$$

for positive constants r_j and for all $0 \leqq j \leqq n$. Let $\mathbf{X}_m = (X_0^{(m)}, \cdots, X_n^{(m)})$ be the response vector associated with the sample size vector $\mathbf{N}_m$ and let $\mathbf{B}_m(\mathbf{u}) = [b_{k,l}^{(m)}(\mathbf{u})]_{k,l \in T}$ be the matrix whose elements are defined by

$$b_{k,l}^{(m)}(\mathbf{u}) = \frac{\partial^2 \ln L}{\partial u_k\, \partial u_l}(\mathbf{X}_m, P\{g(\cdot)\}),$$

where $g \in AMJ(T)$ and u is its coefficient vector. Further let $\mathbf{\Sigma} = [\sigma_{k,l}]_{k,l \in T}$ be defined by

$$\sigma_{k,l} = \frac{(n + 1) \sum_{j=0}^{n} r_j P^{-1}\{f(d_j)\}[1 - P\{f(d_j)\}] d_j^{k+l}}{\sum_{j=0}^{n} r_j}. \tag{25}$$

In (25) it is to be understood that $d_0^0 \equiv 1$ (even if $d_0 = 0$) and that $P^{-1}\{f(d_0)\}d^{k+l} \equiv 0$ if $0 = d_0 = f(0)$ and $k + l > 0$. Since we shall not consider the case for which $d_0 = 0$, $f(0) = 0$, $0 \in T$ all hold simultaneously, $\mathbf{\Sigma}$ will be well defined. In order that $\mathbf{\Sigma}$ be strictly positive definite it suffices that $\Sigma_{k \in T} x_k d_j^k = 0$ for $j = 0, \cdots, n$ implies $x_k = 0$ for all $k \in T$. Under the hypotheses of Theorem 2 below, this follows easily from Descartes' rule of signs.

Let $\mathbf{q} = (q_i)_{i \in T}$ be the coefficient vector associated with the true function $f \in AMJ(T)$ and let $\mathbf{q}_m$ be the coefficient vector associated with a maximum likelihood estimate f_m of f in $AMJ(T)$.

THEOREM 2. *Suppose* $q_i > 0$ *for at least one* $i \in T/\{\infty\}$ *and* (12) *and* (24) *hold. Assume* card $(T) \leqq n + 1$ *except when* $d_0 = 0$ *and* $f(0) = 0$. *If both* $d_0 = 0$ *and* $f(0) = 0$ *assume* card $(T) \leqq n$ *and* $0 \notin T$. *Then*

$$N_m^{\frac{1}{2}}(\mathbf{q}_m - \mathbf{q}) \Rightarrow \mathbf{Y},$$

where $\Rightarrow$ *denotes convergence in distribution and* $\mathbf{Y} = (Y_i)_{i \in T}$ *is a random vector whose distribution is uniquely determined by the conditions*

$$\mathbf{\Sigma Y} = \mathbf{Z} + \mathbf{C}\,, \tag{26}$$

$$C_i \geqq 0\,, \quad Y_i \geqq 0 \quad \textit{if} \quad q_i = 0\,, \quad C_i Y_i = 0\,, \quad \textit{and} \quad C_i q_i = 0\,, \quad \textit{for all} \quad i \in T\,, \tag{27}$$

and $\mathbf{Z} = N(\mathbf{0}, \mathbf{\Sigma})$. *For a given value of* $\mathbf{Z}$ (26) *and* (27) *uniquely determine* $\mathbf{Y}$ *and* $\mathbf{C}$.

PROOF. By a minor adaptation of Guess and Crump (1976, Theorem 2) there exists a unique $f_m \in AMJ(T)$ for m sufficiently large maximizing the likelihood function L over $AMJ(T)$. Let $\mathbf{G}_m(\mathbf{q}) = (\partial \ln L/\partial q_i)_{i \in T}$. Then the coefficient vector $\mathbf{q}_m$ associated with f_m satisfies the Kuhn–Tucker conditions

$$\mathbf{G}_m(\mathbf{q}_m)_i \overset{(=)}{\underset{(>)}{\leqq}} 0 \quad \text{if} \quad q_{im} = 0\,, \tag{28}$$

for all $i \in T$. Expanding $\mathbf{G}_m$ in a Taylor series about $\mathbf{q}$ and using (28) we get

$$\mathbf{G}_m(\mathbf{q})_i \overset{(=)}{\underset{(>)}{\leqq}} -\{\mathbf{B}_m(\mathbf{q}_m') \cdot (\mathbf{q}_m - \mathbf{q})\}_i \quad \text{if} \quad q_{im} = 0\,, \tag{29}$$

where $\mathbf{q}_m'$ lies on the line joining $\mathbf{q}$ and $\mathbf{q}_m$. Adapting the argument in Zacks (1971, pages 246–247) and using standard asymptotic normal theory it is not difficult to show that $N_m^{\frac{1}{2}}\mathbf{G}_m(\mathbf{q}) \Rightarrow \mathbf{Z}$, where $\mathbf{Z} = N(\mathbf{0}, \mathbf{\Sigma})$. Next define

$$\mathbf{Z}_m = N_m^{-\frac{1}{2}}\mathbf{G}_m(\mathbf{q}) + \{\mathbf{\Sigma} + N_m^{-1}\mathbf{B}_m(\mathbf{q}_m')\}\mathbf{Y}_m\,, \tag{30}$$

where $\mathbf{Y}_m = N_m^{\frac{1}{2}}(\mathbf{q}_m - \mathbf{q})$. As in the proof of Theorem 1 it can be shown that $\mathbf{q}_m \to \mathbf{q}$ w.p. 1. It follows easily that $N_m^{-1}\mathbf{B}_m(\mathbf{q}_m') \to -\mathbf{\Sigma}$ w.p. 1. Thus the second term in (30) converges to zero in probability and hence we have

$$\mathbf{Z}_m \Rightarrow \mathbf{Z}\,. \tag{31}$$

It now follows from (29) that $\mathbf{Y}_m$ satisfies the Kuhn–Tucker conditions for the quadratic programming (QP) problem

$$\begin{aligned} &\text{minimize} \quad -\mathbf{Z}_m{}^T\mathbf{y} + \tfrac{1}{2}\mathbf{y}^T\mathbf{\Sigma}\mathbf{y} \\ &\text{subject to} \quad y_i \geqq -N_m^{\frac{1}{2}}q_i\,, \quad i \in T\,. \end{aligned} \tag{32}$$

Since $\mathbf{\Sigma}$ is strictly positive definite this QP problem has a unique solution for each value of the vector $\mathbf{Z}_m$. Thus we may write $\mathbf{Y}_m = \mathbf{F}_m(\mathbf{Z}_m)$ where $\mathbf{F}_m(\mathbf{z})$ denotes the unique solution to (32) when $\mathbf{Z}_m$ is replaced by $\mathbf{z}$. If $\{\mathbf{z}_m\}_{m=1}^{\infty}$ is a sequence of vectors converging to a vector $\mathbf{z}$ it is readily seen that $\mathbf{F}_m(\mathbf{z}_m) \to \mathbf{F}(\mathbf{z}) \equiv \mathbf{y}$, where $\mathbf{y}$ is the unique solution to the QP problem

$$\begin{aligned} &\text{minimize} \quad -\mathbf{z}^T\mathbf{y} + \tfrac{1}{2}\mathbf{y}^T\mathbf{\Sigma}\mathbf{y} \\ &\text{subject to} \quad y_i \geqq 0 \quad \text{if} \quad q_i = 0\,, \quad i \in T\,. \end{aligned} \tag{33}$$

Hence by the continuous mapping theorem of Billingsley (1968, Theorem 5.5) it follows from (31) that $\mathbf{F}_m(\mathbf{Z}_m) \Rightarrow \mathbf{F}(\mathbf{Z})$. In other words $\mathbf{Y}_m \Rightarrow \mathbf{Y}$ where $\mathbf{Y}$ is a

random vector whose distribution is uniquely determined by the requirement that $\mathbf{Y}$ solve (33) when $\mathbf{z}$ is replaced by $\mathbf{Z}$. Since (26) and (27) are the Kuhn–Tucker conditions for this problem they determine the distribution of $\mathbf{Y}$ uniquely. This completes the proof.

COROLLARY. *If $q_i > 0$ for all $i \in T$ then, under the hypotheses of Theorem 2, $N_m^{\frac{1}{2}}(\mathbf{q}_m - \mathbf{q}) \Rightarrow N(\mathbf{0}, \mathbf{\Sigma}^{-1})$.*

This follows from the fact that $\mathbf{C} = \mathbf{0}$ in (26) when all the q_i are positive.

In conclusion we state without proof a theorem about the asymptotic distribution of the vector $(P\{f_m(d_0)\}, \cdots, P\{f_m(d_n)\})$ when the number of coefficients q_k which are positive is greater than the number of test doses at which the response probability is positive. In this case the coefficients are not uniquely determined by the response probabilities at the doses and an approach along the lines of Theorem 2 cannot be used. In such a situation the theorem below could be used in conjunction with linear, or in some cases nonlinear, programming to compute conservative confidence intervals for the increased risk over background. The particularly simple form of the limiting distribution in the following theorem does not, of course, hold under the hypotheses of Theorem 2.

Let

$$V_j^{(m)} = \frac{(N_j^{(m)})^{\frac{1}{2}}[P\{f_m(d_j)\} - P\{f(d_j)\}]}{\{P\{f_m(d_j)\}[1 - P\{f_m(d_j)\}]\}^{\frac{1}{2}}}$$

and let $\mathbf{I}_k$ denote the $k \times k$ identity matrix.

THEOREM 3. *Suppose f, f_m are defined as in Theorem 2 and $N_j^{(m)} \to \infty$ for each j.*

(i) *If $d_0 > 0$ or $f(0) > 0$ then $q_k > 0$ for at least $n + 2$ elements $k \in T$ implies $(V_0^{(m)}, \cdots, V_n^{(m)}) \Rightarrow N(0, I_{n+1})$.*

(ii) *If both $d_0 = 0$ and $f(0) = 0$ then $q_k > 0$ for at least $n + 1$ elements $k \in T$ implies $(V_1^{(m)}, \cdots, V_n^{(m)}) \Rightarrow N(0, I_n)$.*

Acknowledgment. We thank an Associate Editor and referees for useful comments which substantially improved an earlier version of this paper. In particular we thank a referee for greatly simplifying the proof of Theorem 1.

REFERENCES

[1] ARMITAGE, P. and DOLL, R. (1961). Stochastic models for carcinogenesis. *Proc. Fourth Berkeley Symp. Math. Statist. Prob.* **4** 19–38. Univ. of California Press.

[2] BILLINGSLEY, P. (1968). *Convergence of Probability Measures.* Wiley, New York.

[3] BROWN, C. C. (1976). Statistical aspects of extrapolation of dichotomous dose-response data. Unpublished *National Cancer Institute* report.

[4] CHUNG, K. L. (1968). *A Course in Probability Theory.* Harcourt, Brace and World, New York.

[5] CRUMP, K. S., GUESS, H. A. and DEAL, K. (1977). Confidence intervals and tests of hypotheses concerning dose-response relations inferred from animal carcinogenicity data. *Biometrics* **33** 437–451.

[6] CRUMP, K. S., HOEL, D. G., LANGLEY, C. H. and PETO, R. (1976). Fundamental carcinogenic processes and their implications for low-dose risk assessment. *Cancer Research* **36** 2973–2979.

[7] GUESS, H. A. and CRUMP, K. S. (1976). Low-dose extrapolation of data from animal carcinogenicity experiments—analysis of a new statistical technique. *Math. Biosci.* **32** 15–36.

[8] KARLIN, S. and STUDDEN, W. J. (1966). *Tchebycheff Systems: With Applications in Analysis and Statistics*. Wiley, New York.

[9] KREIN, M. G. and REHTMAN, P. G. (1959). Development in a new direction of the Cebysev–Markov theory of limiting values of integrals. *Amer. Math. Soc. Transl.* 2, **12** 123–135.

[10] PETO, R. (1974). Report of the NIEHS conference on extrapolation of risks to man from environmental toxicants on the basis of animal experiments. NIEHS, Research Triangle Park.

[11] RAO, C. R. (1973). *Linear Statistical Inference and Its Applications*, 2nd ed. Wiley, New York.

[12] ZACKS, S. (1971). *The Theory of Statistical Inference*. Wiley, New York.

NIEHS
P.O. BOX 12233
RESEARCH TRIANGLE PARK
NORTH CAROLINA 27709