RERF TR 25-81

TECHNICAL REPORT
RERF TR 25-81
業 績 報 告 書

A TAYLOR SERIES APPROACH TO SURVIVAL ANALYSIS

テイラー級数法による生存の解析

JOEL B. BRODSKY, Ph.D.
PETER G. GROER, Ph.D.*

放影研
RERF

RADIATION EFFECTS RESEARCH FOUNDATION
財団法人 放射線影響研究所

A Cooperative Japan – United States Research Organization
日 米 共 同 研 究 機 関

ACKNOWLEDGMENT

謝　辞

We would like to thank Miss Michiko Takagi for her excellent typing and secretarial assistance and Dr. Arthur V. Peterson for his comments on earlier drafts of manuscript.

タイプ及び原稿作成上の援助をされた高木迪子氏と草稿に対する助言をされた Arthur V. Peterson 博士に謝意を表する．

RERF TECHNICAL REPORT SERIES

放影研業績報告書集

The RERF Technical Reports provide the official bilingual statements required to meet the needs of Japanese and American staff members, consultants, and advisory groups. The Technical Report Series is not intended to supplant regular journal publication.

放影研業績報告書は，日米専門職員，顧問，諮問機関の要求に応えるための日英両語による公式報告記録である．業績報告書は通例の誌上発表論文に代わるものではない．

The Radiation Effects Research Foundation (formerly ABCC) was established in April 1975 as a private nonprofit Japanese Foundation, supported equally by the Government of Japan through the Ministry of Health and Welfare, and the Government of the United States through the National Academy of Sciences under contract with the Department of Energy.

放射線影響研究所（元ABCC）は，昭和50年4月1日に公益法人として発足したもので，その経費は日米両政府の平等分担により，日本は厚生省の補助金，米国はエネルギー省との契約に基づく米国学士院の補助金とをもって運営されている．

RERF TR 25-81

放影研 RERF
RADIATION EFFECTS RESEARCH FOUNDATION
財団法人 放射線影響研究所

A TAYLOR SERIES APPROACH TO SURVIVAL ANALYSIS

テイラー級数法による生存の解析

JOEL B. BRODSKY, Ph.D.; PETER G. GROER, Ph.D*

Department of Epidemiology & Statistics
疫学統計部

SUMMARY

A method of survival analysis using hazard functions is developed. The method uses the well known mathematical theory for Taylor Series. Hypothesis tests of the adequacy of many statistical models, including proportional hazards and linear and/or quadratic dose responses, are obtained. A partial analysis of leukemia mortality in the Life Span Study cohort is used as an example. Furthermore, a relatively robust estimation procedure for the proportional hazards model is proposed.

要約

ハザード関数を用いた生存解析法を開発する．この方法はよく知られたテイラー級数理論を用いる．比例ハザードモデルや線型あるいは2次線量反応モデルを含む多数のモデルの適合性についての仮説検定を行う．寿命調査集団中の白血病死亡率の部分解析を一つの例として使用する．更に比例ハザードモデルのロバスト推定法を提案する．

INTRODUCTION

Currently the analysis of survival data is a very important topic in the scientific literature. Assessment of radiation effects and cancer research are two examples where survival methods are used. The ability to incorporate covariate information with survival history was greatly enhanced by the introduction of the concept of proportional hazards functions.[1] The enormous amount of resulting statistical work is testimonial to the need for such methods. Unfortunately, there are many unanswered questions about how these methods behave when the modeling assumptions are violated. We posed such questions while attempting to analyze leukemia mortality of atomic bomb surviovrs. During a preliminary analysis the proportional hazard assumption was seriously questioned and we felt that the usual method of stratifying was unsatisfactory. Also, we felt the need to develop a

緒言

今日，学術文献において生存資料の解析は非常に重要な課題となっている．放射線影響と癌調査の二つは，その評価に生存法が用いられている例である．比例ハザード関数の概念が導入されたことによって，共変量情報と生存歴とを混合できる可能性が大きく増大した．[1] この結果生じる膨大な量の統計解析がこれらの方法の必要性を物語っている．しかし残念なことに，モデルに関する仮定が崩れた場合，これらの方法がどうなるかについては不明な点が多い．本調査では原爆被爆者の白血病死亡率を解析するに当たり，そうした問題点を提起した．予備解析では比例ハザード仮説が大きく疑問視され，通常の層化法では不十分であると考えられた．また，資料

**Institute for Energy Analysis, Oak Ridge Associated Universities, RERF visiting research associate*
Oak Ridge 連合大学エネルギー分析研究所，放影研来所研究員

Approved 承認 9 November 1981

Printed 印刷 September 1984

formal test of whether data are consistent with some type of a proportional hazards assumption. A reviewer brought to our attention a different approach to this question.[2]

が何らかの種類のハザード仮説と一致するか否かを調べるための本式な検定を開発する必要性も感じられた．検討者の一人が，この問題を処理する異なった方法を示唆した．[2]

The general statistical development needed is given below and, since it is extremely mathematical, the reader may wish to read it in a cursory manner and refer back to it while concentrating on the subsequent sections.

必要とされる一般的な統計学的展開は下記に示すが，これは高度に数学的なものなので，一通り目を通しておいて，それに続く三つの章を読むとき改めて検討するとよいであろう．

It is assumed throughout that all denominators are nonzero and the regularity conditions for maximum likelihood estimation are satisfied. The proposed model is developed using likelihood ratio tests, however, other testing procedures or estimation methods are equally appropriate for this development.

これらの方法では分母はすべてゼロではなく，最大尤度推定の正則性条件は十分なものであると仮定する．提案のモデルは尤度比検定を用いて開発したものであるが，この展開には他の検定法や推定法も同様に使用できる．

NOTATION AND THE BASIC FRAMEWORK

Consider the situation of only one failure type of interest. The generalization to multiple failure types is straightforward and will be briefly discussed later. Assume there are N individuals at risk at time 0 and for the ith individual the triple $(t_i, I_i, \underset{\sim}{z}_i)$ is observed, where t_i is the survival or censor time, I_i is an indicator of censoring, and $\underset{\sim}{z}_i$ is a J dimensional vector of possibly time-dependent covariates.

記数法及び基本構造

対象とする疾患がただ一つの failure 型である場合を考える．複数の failure 型の疾患の一般化は容易であり，後に略述する．0時間における観察人数をN人とし，i 番目の対象者については三つの事項 (t_i, I_i, z_i) が観察されると仮定する．ただし，t_i は生存時間若しくは打ち切り時間，I_i は打ち切り標識，z_i は時間依存の共変量と考えられるJ次元ベクトルである．

For survival time, t, and covariate vector, $\underset{\sim}{z}$, define the hazard function, $\lambda(t|\underset{\sim}{z})$, in the usual manner.[3] Assuming the censoring is uninformative and dropping the proportionality constant, the log likelihood becomes

生存時間 t 及び共変量ベクトル z についてハザード関数 $\lambda(t|z)$ を通常の方法で定義する．[3] 打ち切りが非情報的で比例関係定数がないと仮定すると対数尤度は次のようになる．

$$\sum_{i=1}^{N}\left[I_i \ln\left[\lambda(t_i|\underset{\sim}{z}_i)\right] - \int_0^{t_i}\lambda(s|z_i)\,ds\right] \qquad (1)$$

Suppose that $\lambda(t|\underset{\sim}{z})$ has three continuous derivatives in all (J+1) arguments. This assumption is slightly stronger than the usual regularity condition regarding three derivatives normally required for maximum likelihood methods.[4] Applying Taylor's theorem[5] to $\lambda(t|\underset{\sim}{z})$ with expansion point $\binom{t_o}{\underset{\sim}{z}_o}$ yields

$\lambda(t|z)$ はすべての（J＋1）引き数において三つの連続導関数をもつと考える．この仮定は，通常最尤法で必要とされる三つの導関数に関しては通常の正規性条件よりもやや強力である．[4] テイラーの定理[5] を展開点 $\binom{t_o}{z_o}$ をもって $\lambda(t|z)$ に適用すると，

$$\lambda(t|\underset{\sim}{z}) = \alpha + \beta(t-t_o) + \sum_{j=1}^{J}\gamma_j(z_j - z_{oj}) + \delta(t-t_o)^2 + \sum_{j=1}^{J}\epsilon_j(t-t_o)(z_j-z_{oj})$$
$$+\sum_{j=1}^{J-1}\sum_{l>j}\eta_{jl}(z_j-z_{oj})(z-z_{ol}) + \sum_{j=1}^{J}\theta_j(z_j-z_{oj})^2 + R(\underset{\sim}{c}, t, \underset{\sim}{z}) \qquad (2)$$

where

となる．ただし，

$$\alpha=\lambda(t_o|z_o)$$

$$\beta=\left.\frac{\partial\lambda(t|\underset{\sim}{z})}{\partial t}\right|_{\binom{t_o}{\underset{\sim}{z}_o}}$$

$$\gamma_j=\left.\frac{\partial\lambda(t|\underset{\sim}{z})}{\partial z_j}\right|_{\binom{t_o}{\underset{\sim}{z}_o}} \qquad j=1,\ldots,J$$

$$\delta=½\left.\frac{\partial^2\lambda(t|\underset{\sim}{z})}{\partial t^2}\right|_{\binom{t_o}{\underset{\sim}{z}_o}}$$

$$\epsilon_j=\left.\frac{\partial^2\lambda(t|\underset{\sim}{z})}{\partial t\partial z_j}\right|_{\binom{t_o}{\underset{\sim}{z}_o}} \qquad j=1,\ldots,J$$

$$\eta_{jl}=\left.\frac{\partial^2\lambda(t|\underset{\sim}{z})}{\partial z_j\partial z_l}\right|_{\binom{t_o}{z_o}} \qquad 1\leqq j<l\leqq J$$

$$\theta_j=½\left.\frac{\partial^2\lambda(t|\underset{\sim}{z})}{\partial z_j^2}\right|_{\binom{t_o}{z_o}} \qquad j=1,\ldots,J$$

$\underset{\sim}{c}$ is a point on the line segment connecting $\binom{t_o}{\underset{\sim}{z}_o}$ and $\binom{t}{\underset{\sim}{z}}$, and $R(\underset{\sim}{c}, t, \underset{\sim}{z})$ is the remainder term involving third derivatives from Taylor's theorem. The parameters of interest are the coefficients in (2), that is, α, β, etc. Since the expansion point does not correspond to a "true parameter value" in any sense, equation (2) is an unusual statistical usage of Taylor's theorem. However, as will be shown in the subsequent sections, asymptotic likelihood ratio tests of various statistical models can be constructed using standard methods.

Unfortunately, without some additional assumptions regarding $R(\underset{\sim}{c}, t, \underset{\sim}{z})$ formula (2) is not directly applicable in analyzing the hazard function. A problem develops since maximum likelihood estimates of the parameters must depend upon the $R(\underset{\sim}{c}, t, \underset{\sim}{z})$'s, even though the parameters themselves do not. In order to circumvent this difficulty, we assume the survival time axis can be broken into disjoint intervals, such that the remainder term is reasonably

$\underset{\sim}{c}$は$\binom{t_o}{z_o}$と$\binom{t}{\underset{\sim}{z}}$を結ぶ線分上の点，$R(\underset{\sim}{c}, t, \underset{\sim}{z})$はテイラーの定理による三次の剰余項である．対象とするパラメーターは(2)の係数，すなわちα，β等である．展開点はあらゆる意味で「真のパラメーター値」とは対応しないので，方程式(2)はテイラーの定理の統計学的使用法としては異例のものである．しかし後述の章で示すように，種々の統計モデルの漸近的尤度比検定は標準的な方法を用いて構成できる．

残念ながら，$R(\underset{\sim}{c}, t, \underset{\sim}{z})$に関する仮説を更に加えなければ公式(2)をハザード関数の解析に直接応用することはできない．パラメーター自体は$R(\underset{\sim}{c}, t, \underset{\sim}{z})$の最大尤度推定値に依存しなくても，パラメーターの最大尤度評価量は$R(\underset{\sim}{c}, t, \underset{\sim}{z})$のそれに依存せざるを得ないので，問題が起こってくる．こうした困難を避けるために，生存時間軸は互いに交わりのない期間に分断することが可能で，剰余項が生存時間と共変量の双方において三次となるように各

approximated on each interval as being cubic in both survival time and covariate. Implicit in this framework is the existence of a different t_o for each interval, say t_{ok} for the kth interval. For convenience, we take $\underset{\sim}{z}_o$ to be the same for all intervals; the extension to $\underset{\sim}{z}_{ok}$'s is trivial. This assumption is based upon two beliefs. First, Taylor's theorem says that $R(\underset{\sim}{c}, t, \underset{\sim}{z})$ is exactly a cubic polynomial in t and z, but where the coefficients are proportional to the third derivatives of $\lambda(t|\underset{\sim}{z})$ evaluated at $\underset{\sim}{c}$. Thus, it is felt that within intervals the highest order terms, the cubics, will be a good approximation to the remainder. Secondly, it is felt that within intervals the highest order derivatives will be approximately independent of $\underset{\sim}{c}$, especially if the intervals are short and/or the hazard itself is reasonably smooth.

期間に合理的に近似するものと仮定する．この構造では当然各期間に異なる t_o が存在する．例えば第 k 期間では t_{ok} である． $\underset{\sim}{z}_o$ は便宜上全期間で等しく，$\underset{\sim}{z}_{ok}$ への拡張は自明であるとした．この仮説は次の二つの考え方に基づいている．第一に，テイラーの定理では $R(\underset{\sim}{c}, t, \underset{\sim}{z})$ は t 及び z においては正しく三次多項式であるが，この場合係数は $\underset{\sim}{c}$ 点で評価した $\lambda(t|\underset{\sim}{z})$ の第三導関数と比例する．したがって期間内では最高次，すなわち三次の項は剰余項によく近似すると思われる．第二に期間内では，特にその期間が短い場合，ないしはハザード自体がかなりの程度まで滑らかな場合には，最高次の導関数がほぼ $\underset{\sim}{c}$ に対して独立していると思われる．

Formally, for t in the kth interval assume

形式的に，第 k 期間の t については

$$R(\underset{\sim}{c}, t, \underset{\sim}{z}) = R_{k1}(t-t_{ok})^3 + \sum_{j=1}^{J} R_{k2j}(z_j-z_{oj})^3 \tag{3}$$

where R_{k1} and R_{k2j} $(1 \leq j \leq J)$ are constants.

と仮定する．ただし，R_{k1} 及び R_{k2j} $(1 \leq j \leq J)$ は定数である．

It should be emphasized that estimates of the remainder coefficients contain valuable information regarding the adequacy or "goodness of fit" of an assumed model. For example, the most general dose-response model considered subsequently has a vanishing remainder term. Consequently, a nonvanishing remainder would indicate that the assumed class of models may be inappropriate. Notice that jointly equations (2) and (3) allow the integrals in (1) to be evaluated in closed form and the result is a partially parametric log likelihood. The parametrization is only partial because the parameters are only values of the unknown hazard function or its derivatives at the expansion points. Therefore, certain attributes of the hazard function can be examined with likelihood ratio tests without specifying the hazard because specific models for the hazard will require relationships between the parameters of equation (2). Consequently, there will be fewer "free" parameters in the specific model than in the general model and the likelihood ratio test will depend upon this difference in the number of parameters.

剰余係数の推定値は，仮定したモデルの「適合度」に関する重要な情報を含んでいることを強調しておかなければならない．例えば，次に考えられる最も一般的な線量反応モデルはゼロになる剰余項をもつ．したがって剰余項がゼロでなかった場合は，仮定したモデルのクラスが不適当であるかもしれないことを示す．方程式(2)と(3)を結合すると，(1)の積分は閉じた形で評価でき，その結果は部分的にパラメトリックな対数尤度になることに注目されたい．パラメーターは不明のハザード関数若しくはその展開点における導関数の値にすぎないので，パラメーターの一部しか定められない．したがってハザード関数の特定の属性は，そのハザードに対する特定モデルに方程式(2)のパラメーター間の関係が必要なことから，そのハザードを具体的に定めないで尤度比検定で調べることができる．その結果「自由」パラメーターの数は，一般モデルより特定モデルの方が少なく，尤度比検定はこのパラメーター数の差によって左右される．

Suppose there are K intervals and N_k individuals at risk at the beginning of the kth interval. Take

期間の数をK，第 k 期間の開始時点での観察人数を

N_{K+1} to be the number of survivors of the Kth interval. Substitution of equations (2) and (3) into (1) yields for log likelihood, ℓ,

N_k 人とする．N_{k+1} は第K期間の生存者数であるとして，方程式(2)及び(3)を(1)に代入すると対数尤度 ℓ は

$$\ell = \ell_1 + \ell_2 + \cdots + \ell_k \tag{4}$$

where ℓ_k is the log likelihood corresponding to interval k. Each ℓ_k is based solely upon the N_k individuals at risk for interval k and has the N_{k+1} survivors treated as censored at the end of the interval. Thus, unless there are parameter constraints across intervals, standard maximum likelihood theory applies to each ℓ_k individually and estimates from different intervals are asymptotically independent. Also, each ℓ_k can be maximized separately and problems associated with large numbers of parameters are minimized. For each interval, under the general model of equation (2), there are $4 + \frac{J(J+7)}{2}$ parameters. Hence, for likelihood ratio tests involving all K intervals, the most general alternative model has $\left(4 + \frac{J(J+7)}{2}\right) \cdot K$ parameters.

となる．ただし，ℓ_k は期間kに対応する対数尤度である．各 ℓ_k は期間kの観察人数 N_k 人のみに基づくものであり，N_{k+1} 人の被爆者は期間の終了時点で打ち切られたものとして処理する．したがって，各期間にまたがるパラメーターの制約がなければ，標準的な最尤法を各 ℓ_k に適用でき，異なる期間から得られた推定値は漸近的に独立したものとなる．また，各 ℓ_k は別個に最大にすることができ，パラメーター数の多いことからくる問題を最少限に抑えられる．各期間について，方程式(2)の一般モデルではパラメーターの数は $4 + \frac{J(J+7)}{2}$ 個である．ゆえにすべてのK個の期間の尤度比検定では最も一般的な代替モデルは $\left(4 + \frac{J(J+7)}{2}\right) \cdot K$ 個のパラメーターをもつことになる．

In closing this section, we remark that the only assumptions made thus far regarding the hazard function are in regards to smoothness of the function itself.

最後に，ハザード関数について現在までに立てられた仮説は，関数自体の滑らかさに関するもののみであることを述べておく．

A PARTIAL ANALYSIS OF LEUKEMIA MORTALITY

白血病死亡率の部分解析

Here we demonstrate a partial analysis of leukemia mortality in members of LSS sample. A complete analysis, using the T65 revised doses (T65DR), is given elsewhere.[6]

ここでは寿命調査対象者の白血病死亡率の部分解析について述べる．T65改訂線量(T65DR)を使用する完全な解析については別記する．[6]

All 79,856 members of the LSS sample who had estimated T65DR were considered. However, in this example each individual's T65DR was converted to an approximate Lawrence Livermore National Laboratory (LLNL) dose by using the most currently available dose information.[7,8]

T65改訂線量(T65DR)の判明している寿命調査対象者79,856人全員を対象とした．しかし，本解析では各個人のT65DR線量は最新の線量情報を用いてLawrence Livermore研究所(LLNL)の線量に近似したものに変換した．[7,8]

The principal aim of this analysis was to distinguish between the usual dose-response models[9]: L-Q-L (linear and quadratic in gamma rays and linear in neutrons), L-L (linear in both gamma rays and neutrons), and Q-L (quadratic in gamma

本解析の主目的は通常の線量反応モデル，[9] すなわちL-Q-L(ガンマ線に関して線型及び二次的で，中性子線に関して線型)，L-L(ガンマ線及び中性子線の双方に関して線型)及びQ-L(ガンマ線に関して二次的で，中性子線に関して線型)を識別することで

rays and linear in neutrons). In the complete analysis[6] four intervals were considered, though here we shall present the results of the first interval (1 October 1950 – 30 September 1957) only. Since the L-Q-L model includes the other models as special cases, the hazard function was assumed to be of the form

あった．完全な解析[6]では四つの期間について検討したが，ここでは第一期間（1950年10月1日～1957年9月30日）の結果のみを示す．L-Q-Lモデルは特例として他のモデルも含むので，ハザード関数は次の形になると仮定した．

$$\lambda(t|\underset{\sim}{z}) = \alpha + B_1 D + B_2 D^2 + \left\{ \begin{array}{ll} \text{other possible covariate} & \text{他の可能な共変量効果} \\ \text{effects and interactions} & \text{及び相互作用} \end{array} \right\} \quad (5)$$

where $D = (Cg_{1d} + (1-C)g_{od})G' + (Cn_{1d} + (1-C)n_{od})N'$
= LLNL total dose
$D^2 = (Cg_{1d} + (1-C)g_{od})^2 (G')^2$
= LLNL gamma dose squared
G' = T65DR gamma dose
N' = T65DR neutron dose
d = T65DR total dose
and C = 0, 1 for Hiroshima and Nagasaki, respectively.

ただし，$D = (Cg_{1d} + (1-C)g_{od})G' + (Cn_{1d} + (1-C)n_{od})N'$
=LLNLの総線量
$D^2 = (Cg_{1d} + (1-C)g_{od})^2 (G')^2$
=LLNLのガンマ線量の二乗
G' =T65DRガンマ線量
N' =T65DR中性子線量
d =T65DR総線量
C =広島0，長崎1

The g's and n's are T65DR to LLNL dose conversion constants. The assumptions regarding D and D^2 explicitly control for the dose conversions and require the L-Q-L model to be the most general possible. Mathematically, the model in (5) is equivalent to

g及びnはT65DRのLLNL線量への換算定数である．D及びD^2に関する仮説は明らかに線量換算を制御し，L-Q-Lモデルができるかぎり最も一般的なものであることを求める．数学的には(5)におけるモデルは

$$\lambda(t|\underset{\sim}{z}) = \alpha + b_1 G + b_2 G^2 + b_3 N + \left\{ \begin{array}{ll} \text{other possible covariate} & \text{他の可能な共変量効果} \\ \text{effects and interactions} & \text{及び相互作用} \end{array} \right\} \quad (6)$$

for G and N being LLNL gamma and neutron doses, respectively.

と等しい．ただしGはLLNLガンマ線量，Nは中性子線量である．

The expansion point, $\begin{pmatrix} t_o \\ \underset{\sim}{z}_o \end{pmatrix}$, was taken to be the origin and four covariates were used: LLNL total dose (covariate #1), age at the time of the bomb (ATB) (covariate #2), sex (covariate #3), and city (covariate #4). Table 1 contains log likelihoods and fitted parameters for a series of hierarchical models. By the likelihood ratio criterion, the best fitting model has only three nonzero parameters: α, γ_1, and θ_1. Note that since the b's in (6) must be non-negative some of the likelihood ratio tests were performed in a one-sided manner. For constants $g_{o\cdot}$, $g_{1\cdot}$, $n_{o\cdot}$,

展開点$\begin{pmatrix} t_o \\ z_o \end{pmatrix}$を原点とし，LLNL総線量（共変量1），原爆時年齢（共変量2），性（共変量3）及び都市（共変量4）の四つの共変量を用いた．表1に一連の階段的モデルの対数尤度と適合パラメーターを示した．尤度比基準では，最もよく適合するモデルのもつ非ゼロパラメーターはα，γ_1及びθ_1のわずか3個である．(6)のbは負であってはならないので，尤度比検定の幾つかは片側検定で行った．定数$g_{o\cdot}$，

TABLE 1 PROPERTIES OF FITTED HIERARCHICAL MODEL
表1　固定階段的モデルの特性

Number of Parameters	Parameters	Minus Log Likelihood
12	$\gamma, \beta, \gamma_1, \gamma_2, \gamma_3, \delta, \epsilon_1, \eta_{13}$ $\eta_{14}, \theta_1, R_1, R_{21}$	661.0681
11	$\alpha, \beta, \gamma_1, \gamma_2, \gamma_3, \delta, \epsilon_1, \eta_{13},$ θ_1, R_1, R_{21}	661.3441
7	$\alpha, \gamma_1, \gamma_2, \eta_{14}, \theta_1, R_1, R_{21}$	665.6681
6	$\alpha, \gamma_1, \eta_{14}, \theta_1, R_1, R_{21}$	665.7919
5	$\alpha, \gamma_1, \eta_{13}, \theta_1, R_1$	666.1425
5	$\alpha, \gamma_1, \eta_{13}, \eta_{14}, \theta_1$	666.6200
4	$\alpha, \gamma_1, \eta_{13}, \theta_1$	666.6219
4	$\alpha, \gamma_1, \eta_{14}, \theta_1$	667.3301
3	$\alpha, \gamma_1, \theta_1$	667.3334
3	$\alpha, \eta_{13}, \theta_1$	669.2085
3	$\alpha, \gamma_1, \eta_{13}$	669.4125
2	α, θ_1	669.2195
2	α, γ_1	670.4748
1	α	732.3226

and $n_{1\cdot}$, equations (2), (5), and (6) imply via the Chain Rule that

$g_{1\cdot}$, $n_{o\cdot}$ 及び $n_{1\cdot}$ は，方程式(2)，(5)及び(6)では連鎖律を用いると次のようになる．

$$\alpha = \lambda(0\,|\,\underset{\sim}{0}) = \text{background leukemia rate at 1 October 1950}$$

1950年10月1日における自然白血病率

$$\gamma_1 = \frac{\partial\lambda(t\,|\,\underset{\sim}{z})}{\partial D}\Bigg|_{\underset{\sim}{\binom{0}{0}}} = [b_1(Cg_{1\cdot} + (1-C)g_{o\cdot}) + b_3(Cn_{1\cdot} + (1-C)n_{o\cdot})$$

$$+ 2b_2(Cg_{1\cdot} + (1-C)g_{o\cdot})^2 G'] \Bigg|_{\underset{\sim}{\binom{0}{0}}} = b_1 g_{o\cdot} + b_3 n_{o\cdot}$$

$$\eta_{14} = \frac{\partial^2\lambda(t\,|\,\underset{\sim}{z})}{\partial D \partial C}\Bigg|_{\underset{\sim}{\binom{0}{0}}} = b_1(g_{1\cdot} - g_{o\cdot}) + b_3(n_{1\cdot} - n_{o\cdot})$$

and

$$\theta_1 = \tfrac{1}{2}\frac{\partial^2(t\,|\,\underset{\sim}{z})}{\partial D^2}\Bigg|_{\underset{\sim}{\binom{0}{0}}} = b_2 g_{o\cdot}^2.$$

Because of the linear relationships between $\binom{\gamma_1}{\eta_{14}}$ and $\binom{b_1}{b_3}$, and θ_1 and b_2, once the parameters are estimated it is straightforward to obtain the corresponding estimates of the b's.

$\binom{\gamma_1}{\eta_{14}}$ 及び $\binom{b_1}{b_3}$，並びに θ_1 及び b_2 の間には線型関係があるので，パラメーターが推定されると b の対応推定値を求めるのは簡単である．

TABLE 2 COEFFICIENTS OF THE BEST FITTING MODEL
表2　最良に適合するモデルの係数

Quantity	Estimate
$\alpha \times 10^5$	.23895
$SD(\alpha) \times 10^5$	.07327
$b_1 \times 10^6$	.13222
$SD(b_1) \times 10^6$	.07312
$b_2 \times 10^8$	.08972
$SD(b_2) \times 10^8$	.03960
$b_3 \times 10^6$	.04740
$SD(b_3) \times 10^6$	.02573
$\frac{b_1}{b_2}$	118.68
$SD(\frac{b_1}{b_2})$	122.03
log likelihood	−667.3334

Table 2 contains the best fitting model's estimates of α, b_1, b_2, b_3, and the quantity $\frac{b_1}{b_2}$. This latter ratio corresponds to the gamma dose at which the quadratic term begins to dominate the linear term in the hazard function. This dose is species dependent but for mammals it is believed to be approximately 100 rad.[10] It should be emphasized that simply dividing the estimate of b_1 by the estimate of b_2, $\hat{b}_2$, will result in a ratio estimate which is biased by the amount $\frac{b_1}{b_2}\mathrm{Var}(\hat{b}_2)$. The estimate reported in Table 2 (118.68) is bias corrected and in agreement with this preconceived value.

表2に最もよく適合するモデルのα, b_1, b_2及びb_3の推定値並びに$\frac{b_1}{b_2}$量を示した．後者の比は二次項がハザード関数の線型項を支配し始めるときのガンマ線量と対応する．この線量は動物の種族によって異なるが，哺乳類では約100radとされている.[10] b_1の推定値をb_2の推定値である$\hat{b}_2$で単純に割ると，$\frac{b_1}{b_2}\mathrm{Var}(\hat{b}_2)$の量によって偏りの生ずる推定比が得られることを強調しなければならない．表2に示した推定値(118.68)は偏りを訂正したもので，この予測値と一致する．

THE GENERAL REGRESSION MODEL

As implied in the introduction, one purpose of this paper is to investigate modeling assumptions regarding the hazard function. In particular, we are interested in examining how the proportional hazards methodology behaves under model violations.

In some situations it is believed that the covariates affect the hazard mainly through a multiple regression and the vector of regression coefficients, $\underset{\sim}{b}$, is of interest. Call this situation the general regression model and denote the hazard as

一般回帰モデル

緒言で述べたように本稿の目的は，ハザード関数に関する仮説のモデル化について調査することである．特に，比例ハザード法がモデルが崩れた場合にどうなるかを調べることに関心がある．

幾つかの調査では，共変量は主として多重回帰によってのみハザードに影響を与えると考えられており，回帰係数のベクトル$\underset{\sim}{b}$が関心の対象となっている．こうした状態を一般回帰モデルと呼び，その

$\lambda(t|\underset{\sim}{b}\ \underset{\sim}{z})$. Notice the general regression model includes proportional hazards and log linear models as special cases.

ハザードを $\lambda(t|\underline{b}\ \underline{z})$ とする．一般回帰モデルには特例として比例ハザード及び対数線型モデルが含まれることに注目されたい．

The general regression model implies certain relationships among the parameters of equation (2). These relationships follow since some of the derivatives in equation (2) can be further detailed. For example by the Chain Rule γ_{kj} becomes

一般回帰モデルは方程式(2)のパラメーター間に特定の関係の存在することを示す．これらの関係は，方程式(2)の導関数の一部は更に詳述できることから当然存在する．例えば，連鎖律を用いると γ_{kj}は

$$\gamma_{kj} = \frac{\partial \lambda_k(t|x)}{\partial x} \cdot \frac{\partial x}{\partial z_j} \Bigg|_{(t_{ok},\ \underset{\sim}{b}\ \underset{\sim}{z}_o)} = b_{kj} \frac{\partial \lambda_k(t|x)}{\partial x} \Bigg|_{(t_{ok},\ \underset{\sim}{b}\ \underset{\sim}{z}_o)}$$

where the possibility of time dependent regression coefficients (i.e., one type of time dependent covariates) is expressed as interval dependent coefficients. Methods for treating the b_{kj}'s as explicit functions of time are given under "Proportional Hazards Models." Assuming all denominators are nonzero, within each interval the totality of these relationships is for $j \neq l \neq m$

となる．ただし，時間に依存する回帰係数の可能性(すなわち，時間依存性共変量の一つの型)は期間依存係数として表される．b_{kj}を時間の明らかな関数として扱う方法は「比例ハザードモデル」の章で述べる．すべての分母がゼロでないとすると，各期間内ではこれらの関係の全体は $j \neq l \neq m$ で

$$\frac{\gamma_l}{\gamma_j} = \frac{\epsilon_l}{\epsilon_j} = \frac{\eta_{jl}}{2\theta_j} = \frac{\eta_{lm}}{\eta_{jm}} = \sqrt{\frac{\theta_l}{\theta_j}} \qquad (7)$$

The subscript k has been dropped in (7) solely for notational convenience and (7) is not intended to imply a lack of time dependence of the covariates. In fact, the null hypothesis of time independent covariates can be tested and the development within the proportional hazards framework of such a test, which is directly applicable to the general regression model, will be detailed later.

となる．下の添字 k は(7)では単に表記の便宜上省略したもので，(7)は共変量の時間依存がないことを示すものではない．事実，時間独立共変量の帰無仮説は検定が可能で，一般回帰モデルに直接応用できるような検定の比例ハザード構造内での展開は後に詳述する．

It needs to be emphasized that the general regression model implies equations (7), but (7) in general does not imply the general regression model. This follows because equations (7) only constrain certain first and second partial derivatives but it is only for those hazards with all third partials vanishing that (7) implies the general regression model.

一般回帰モデルは方程式(7)を示すが，(7)は普通一般回帰モデルを示さないということを強調しておかねばならない．このことは，方程式(7)が第一及び第二部分導関数だけを束縛するのでそう言われる．三つの部分全部がゼロになるハザードの場合に限り，(7)は一般回帰モデルを示すことになる．

A second point worth emphasizing is that equations (7) and all subsequent modeling relationships do not depend explicitly on the chosen expansion points. In fact, equations (7) must hold for all possible choices of the expansion points. Thus, if the assumptions regarding

強調すべき第2点は，方程式(7)とそれらから生ずるすべてのモデル関係は選択した展開点にはっきりとは依存しないということである．事実，方程式(7)は選択可能のすべての展開点について適用できるものでなければならない．$R(\underline{c}, t, \underline{z})$ に関する仮定が

$R(\underset{\sim}{c},\ t,\ \underset{\sim}{z})$ are reasonable, tests under most situations should be insensitive to both choice of interval and choice of t_{ok}.

合理的なものであれば，ほとんどの状況下において検定は期間の選択と t_{ok} の選択の双方に対して影響を受けないはずである．

Equations (7) imply within intervals there are only (J+2) independent parameters among the γ's, ϵ's, η's, and θ's. For convenience, we chose $\gamma_1, \gamma_2, \ldots, \gamma_J, \epsilon_1$, and θ_1 to be these parameters. Obviously, other choices are possible. Substitution of equations (7) into (2) yields for t in the kth interval

方程式(7) においては，期間内では γ, ε, η 及び θ の間にはわずか (J+2) 個の独立パラメーターしかない．便宜上 γ_1, $\gamma_2,\ldots,$ γ_J, ε_1 及び θ_1 をパラメーターとして選ぶ．無論，ほかの選択も可能である．方程式(7) を(2) に代入すると第 k 期間の t は

$$\begin{aligned}\lambda_k(t|\underset{\sim}{b}\ \underset{\sim}{z}) &= \alpha_k + \beta_k(t-t_{ok}) + \sum_{j=1}^{J}\gamma_{kj}(z_j - z_{oj}) + \delta_k(t-t_{ok})^2 \\ &+ \frac{\epsilon_{k1}}{\theta_{k1}}\sum_{j=1}^{J}\gamma_{kj}(t-t_{ok})(z_j - z_{oj}) + \frac{2\theta_{k1}}{\gamma^2_{k1}}\sum_{j=1}^{J-1}\sum_{l>j}\gamma_{kj}\gamma_{kl}(z_j - z_{oj})(z_l - z_{ol}) \\ &+ \frac{\theta_{k1}}{\gamma^2_{k1}}\sum_{j=1}^{J}\gamma^2_{kj}(z_j - z_{oj})^2 + R_{k1}(t-t_{ok})^3 + \sum_{j=1}^{J}R_{k2j}(z_j - z_{oj})^3 \ . \qquad (8)\end{aligned}$$

Also, for interval k and
となる．また，期間 k 及び

S_k = beginning of the kth interval　第 k 期間の始まり

$$I_{ki} = \begin{cases} 1 & \text{if individual i failed from the cause of interest during the kth interval} \\ & \text{第 k 期間中に対象者 i が対象とする死因によって死亡した場合} \\ 0 & \text{otherwise} \\ & \text{それ以外} \end{cases}$$

$$T_{ki} = \begin{cases} S_k & \text{if } t_i < S_k \\ \min(t_i, S_{k+1}) & \text{if } t_i \geqslant S_k \end{cases}$$

$$\begin{aligned}\ell_k = \sum_{i=1}^{N}\Bigg[I_{ki}\ln\left(\lambda_k(T_{ki}|\underset{\sim}{b}\ \underset{\sim}{z}_i)\right) - [A_{ki}(T_{ki}-S_k) \\ + \left(\beta_k + \frac{\epsilon_{k1}}{\gamma_{k1}}\sum_{j=1}^{J}\gamma_{kj}(z_{ij}-z_{oj})\right)\frac{(T_{ki}-t_{ok})^2-(S_k-t_{ok})^2}{2} \\ + \delta_k\cdot\frac{(T_{ki}-t_{ok})^3-(S_k-t_{ok})^3}{3} + R_{k1}\cdot\frac{(T_{ki}-t_{ok})^4-(S_k-t_{ok})^4}{4}\ ]\Bigg] \qquad (9)\end{aligned}$$

where
ただし，

$$\begin{aligned}A_{ki} &= \alpha_k + \sum_{j=1}^{J}\gamma_{kj}(z_{ij}-z_{oj}) + \frac{2\theta_{k1}}{\gamma^2_{k1}}\sum_{j=1}^{J}\sum_{l>j}\gamma_{kj}\gamma_{kl}(z_{ij}-z_{oj})(z_{il}-z_{ol}) \\ &+ \frac{\theta_{k1}}{\gamma^2_{k1}}\sum_{j=1}^{J}\gamma^2_{kj}(z_{ij}-z_{oj})^2 + \sum_{j=1}^{J}R_{k2j}(z_{ij}-z_{oj})^3\end{aligned}$$

By equations (1)-(4) and (9), the likelihood ratio statistic for testing whether the data are consistent with the general regression model on interval k is approximately χ^2 with $\frac{J(J+3)}{2} - 2$ degrees of freedom. If there is only one covariate the test is undefined because the general regression model is undefined. Another model which can be tested is whether a subset of the covariates affect the hazard through a multiple regression and the remainder of the covariates behave as under the general model. The procedure for testing such a situation is straightforward to obtain and we do not pursue this issue further (Example 1).

方程式(1)～(4)及び(9)においては，資料が期間kにおける一般回帰モデルと一致するか否かを検定する尤度比統計量は，自由度$\frac{J(J+3)}{2} - 2$で近似的にχ^2である．共変量が一つだけであれば，一般回帰モデルは定義できないので検定も定義できない．検定できるモデルは，共変量の小部分が多重回帰によってハザードに影響を与え，その他の共変量が一般モデルの場合と同様の反応を示すか否かを検定するものである．そのような状態を検定する方法は簡単に得られるのでここではこれ以上言及しない(例1)．

Let us turn to the question of estimation in the general regression model. We assume that the likelihood equations (A.1 of the Appendix) have an allowable solution in the parameter space. Then, under suitable regularity conditions standard maximum likelihood theory implies that for large samples the maximum likelihood estimates are consistent. However, it is not true that the estimates of the regression coefficients under the general regression model are consistent. In fact, these coefficients cannot be estimated. Nonidentifiability follows from the differential equations themselves:

一般回帰モデルにおける推定の問題に移りたい．本解析では尤度方程式(附録A.1)がパラメーター空間に許容できる解をもつものと仮定する．したがって，適切な正規性条件の下では標準最尤度理論による大規模集団の最大尤度評価量は一定している．しかし，一般回帰モデルにおいて回帰係数の推定値が一定しているわけではない．実際，これらの係数は推定できない．微分方程式自体から識別は不可能である．

$$\alpha_k = \lambda_k(t_{ok} | \underset{\sim}{b}\, \underset{\sim}{z}_o)$$

$$\beta_k = \left(\frac{\partial \lambda_k(t|x)}{\partial t} \right|_{(\underset{\sim}{b}\, \underset{\sim}{z}_o^{t_{ok}})}$$

$$\gamma_{kj} = b_{kj} \left(\frac{\partial \lambda_k(t|x)}{\partial x} \right|_{(\underset{\sim}{b}\, \underset{\sim}{z}_o^{t_{ok}})} \qquad j=1, \ldots, J$$

$$\delta_k = \tfrac{1}{2} \left(\frac{\partial^2 \gamma_k(t|x)}{\partial t^2} \right|_{(\underset{\sim}{b}\, \underset{\sim}{z}_o^{t_{ok}})}$$

$$\epsilon_{k1} = b_{k1} \left(\frac{\partial^2 \lambda_k(t|x)}{\partial t \partial x} \right|_{(\underset{\sim}{b}\, \underset{\sim}{z}_o^{t_{ok}})}$$

$$\theta_{k1} = \frac{b^2{}_{k1}}{2} \left(\frac{\partial^2 \lambda_k(t|x)}{\partial x^2} \right|_{(\underset{\sim}{b}\, \underset{\sim}{z}_o^{t_{ok}})} \qquad (10)$$

From (10), it is seen that the ratios $\frac{b_{kj}}{b_{k1}}$ are consistently estimated as the ratios of the estimates of the respective γ_{kj}'s. However, all the information about b_{k1} contained in $\hat{\gamma}_{k1}$, $\hat{\epsilon}_{k1}$, and $\hat{\theta}_{k1}$ is confounded with information about partial derivatives of the hazard function. Thus, unless there is some additional assumption regarding these derivatives the ratios of the regression coefficients can be estimated but the coefficients themselves cannot. In general, there are three estimable quantities among γ_{k1}, ϵ_{k1}, θ_{k1}, and b_{k1} but the J-1 estimable γ_{kj}'s ($j \geqslant 2$) do not correspond to J-1 estimable b_{kj}'s ($j \geqslant 2$). Therefore, a reparametrization can make b_{k1} estimable but not $b_{k2}, \ldots, b_{kJ}$.

(10) から比$\frac{b_{kj}}{b_{k1}}$はそれぞれの γ_{kj} の推定値の比と同様に常時推定されることが分かる．しかし $\hat{\gamma}_{k1}$, $\hat{\epsilon}_{k1}$ 及び $\hat{\theta}_{k1}$ に含まれる b_{k1} に関するすべての情報は，ハザード関数の部分導関数に関する情報と混同される．したがって，これらの導関数に関して更に幾つかの仮説を加えなければ，回帰係数の比は推定できても回帰係数自体は推定できない．一般的に，γ_{k1}, ϵ_{k1}, θ_{k1} 及び b_{k1} の間には推定できる三つの量があるが，(J−1)個の推定できる γ_{kj}($j \geqslant 2$)は(J−1)個の推定できる b_{kj}($j \geqslant 2$)と対応しない．したがって，パラメーターの再表示によって b_{k1} を推定可能にできるが，$b_{k2}, \ldots, b_{kJ}$ は推定可能にできない．

There are situations where some of the ratios, $\frac{b_{kj}}{b_{k1}}$, are entities of interest. An example is estimation of the relative biological effectiveness (RBE) of neutron vs gamma radiation. Call gamma dose covariate number 1 and neutron dose covariate number 2. If the L-L model is correct and z_{o1} and z_{o2} are both zero, then the RBE for interval k is precisely $\frac{b_{k2}}{b_{k1}}$ (Note that these b's are different from the b's used previously). Furthermore, it is of interest whether the RBE varies with time. This question can be examined using the methods for time dependent covariates to be described.

比$\frac{b_{kj}}{b_{k1}}$のうちの幾つかが対象となる要素である場合もある．中性子線対ガンマ線の相対的生物学的効果比(RBE)の推定がその例である．ガンマ線量の共変量を1番，中性子線量の共変量を2番とする．L-Lモデルが正しく，z_{o1} 及び z_{o2} が両方ともゼロであれば期間kのRBEは正確に$\frac{b_{k2}}{b_{k1}}$となる(これらのbは以前に用いたbとは異なる)．更に，RBEが経時的に変化するか否かに興味がもたれる．この問題は後述する時間依存共変量のための方法を用いて調べることができる．

Example: Reconsidering the Previous Dose-response Models

Under "A Partial Analysis of Leukemia Mortality" the L-L, Q-L, and L-Q-L models were introduced. However, because of the T65DR to LLNL dose conversion, the actual model analyzed explicitly included the dose conversion constants. Now we will demonstrate how to test the adequacy of the different models when the conversion constants are not included. Discriminating between the models is important because the different models can give rise to substantially different low dose results.[9] The models being considered are:

例：前述の線量反応モデルの再検討

「白血病死亡率の部分解析」ではL-Lモデル，Q-Lモデル及びL-Q-Lモデルを紹介した．しかし，T65DRをLLNL線量に換算するため，実際に解析したモデルは明らかに線量変換定数を含んでいた．ここでは，変換定数が含まれない場合に異なるモデルの適合性をいかに検定するかを示す．異なるモデルは実質的に異なる低線量結果を生じる可能性があるので，[9] モデルの識別は重要である．考慮するモデルは

$$\text{Model L-L} \quad \lambda(t|\underset{\sim}{z}) = a + bD_\gamma + cD_n$$

$$\text{Q-L} \quad \lambda(t|\underset{\sim}{z}) = a + b(D_\gamma)^2 + cD_n$$

$$\text{L-Q-L} \quad \lambda(t|\underset{\sim}{z}) = a + b_1 D_\gamma + b_2(D_\gamma)^2 + cD_n$$

where D_γ is administered gamma radiation dose; D_n is administered neutron radiation dose; and

である．ただし D_γ は照射されたガンマ線量，D_n は

TABLE 3 PARAMETER RESTRICTIONS FOR ALL T*
表3　すべてのT*に関するパラメーターの制限

Model	η_{12}	θ_1	θ_2	R_{21}	R_{22}	Restricted Parameters for K Intervals
L-L	0	0	0	0	0	5K
Q-L	0	$\frac{\gamma_1}{2z_{o1}}$	0	0	0	5K
L-Q-L	0		0	0	0	4K

*Tabled values are the values each indicated parameter must equal for each value of t.
表中の値は，それぞれの示すパラメーターが各 t の値に関して等しくなければならない．
The blank cell represents an unrestricted parameter.
空欄は制限を受けないパラメーター．

the a's, b's, and c's are regression coefficients. Also, the a's may be dependent upon other covariates. The restrictions under each model for all t are shown in Table 3.

照射された中性子線量，a，b及びcは回帰係数である．また，aはほかの共変量に依存するかもしれない．各モデルにおけるすべてのtの制限を表3に示した．

The likelihood ratio statistics for these models against the general model have degrees of freedom 5K, 5K, and 4K, respectively.

これらのモデルの一般モデルに対する尤度比統計量はそれぞれ自由度5K，5K，4Kとなる．

Two slightly different testing approaches to this example could be taken. One could test the L-Q-L model against the general model and if it is not rejected, the other models could be tested against it. Alternatively and/or complementary to this approach, one could test within intervals. This procedure would be to test for each interval the L-Q-L model against the general model with a χ^2_4 variate. If L-Q-L is not rejected, then other models could be tested against it with χ^2_1 variates. Other modifications are also possible.

この例に関しては二つの少し異なる検定法を用いることもできる．一つは一般モデルに対するL-Q-Lモデルを検定するもので，棄却されなければそれと他のモデルとの検定を行うことができる．この方法と択一的，ないしはこの方法に対して補完的な方法として期間内検定がある．この方法は各期間についてχ^2_4変量を用いて一般モデルに対してL-Q-Lモデルを検定するものである．L-Q-Lモデルが棄却されなければ，χ^2_1変量を用いてほかのモデルを検定できる．このほかの修正法も可能である．

PROPORTIONAL HAZARDS MODELS

We now turn attention to models based upon the proportional hazards assumption.[1] There has been an extensive amount of work based upon the partial likelihood.[3] We shall demonstrate an alternative approach using the full likelihood and the Taylor series approximation of equations (2) and (3).

比例ハザードモデル

次に比例ハザード仮説に基づくモデルについて考察する．[1] 部分尤度に基づく調査は数多くなされてきた．[3] 本調査では完全尤度並びに方程式(2)及び(3)のテイラー級数近似法を用いた代わりの方法を示す．

The most general form of the proportional hazards assumption states

比例ハザード仮説の最も一般的な形は，幾つかの関数$\lambda_0(\cdot)$及び$g(\cdot,\ldots,\cdot)$について

$$\lambda(t|\underset{\sim}{z}) = \lambda_o(t)g(z_1, \ldots, z_J) \tag{11}$$

for some functions $\lambda_o(\cdot)$ and $g(\cdot, \ldots, \cdot)$. However, proportional hazards has commonly referred

と表される．しかし，比例ハザードは通例共変量

to the situation where the covariate function, $g(\cdot, \ldots, \cdot)$, depends solely upon a linear combination of the covariates. Thus, proportional hazards has routinely meant

関数 $g(\cdot,\ldots,\cdot)$ が共変量の線型結合のみに依存する場合に当てはまる．したがって，比例ハザードは関数 $\lambda_1(\cdot)$ 及び $g(\cdot)$ について通常

$$\lambda(t|\underset{\sim}{b}\,\underset{\sim}{z}) = \lambda_o(t)g(b\,\underset{\sim}{z}) \tag{12}$$

for some functions $\lambda_1(\cdot)$ and $g(\cdot)$. We shall follow this convention and refer to (12) as the proportional hazards assumption. Notice that equation (11) is testable because it implies for all t

となる．本解析ではこの条件に従って比例ハザード仮説として(12)を引用する．方程式(11)はすべての t に関して

$$\epsilon_j = \frac{\beta}{\alpha}\gamma_j \qquad j=1, \ldots, J \tag{13}$$

We do not pursue this point further.

であるので，検定可能であることに注目されたい．この点に関してはこれ以上詳述しない．

Tests of the Proportional Hazards Assumption

Since equation (12) is the special case of the simultaneous occurrence of the general regression model and equation (11), proportional hazards implies the simultaneous restrictions of both equations (7) and (13). Notice there are no restrictions on the covariate function except the regularity conditions.

比例ハザード仮説の検定

方程式(12)は一般回帰モデルと方程式(11)が同時に成立するという特殊な場合なので，比例ハザードは方程式(7)及び(13)の双方を同時に制限する．正規性条件以外に共変量関数に対する制限はない．

Substitution of (13) into (8) yields for t in interval k

(13)を(8)に代入すると期間 k の t は

$$\lambda_k(t|\underset{\sim k}{b}\,\underset{\sim}{z}) = \alpha_k + \beta_k(t-t_{ok}) + \sum_{j=1}^{J}\gamma_{kj}(z_j - z_{oj}) + \delta(t-t_{ok})^2$$

$$+ \frac{\beta_k}{\alpha_k}(t-t_{ok})\sum_{j=1}^{J}\gamma_{kj}(z_j-z_{oj}) + \frac{2\theta_{k1}}{\gamma^2{}_{k1}}\sum_{j=1}^{J-1}\sum_{l>j}\gamma_{kj}\gamma_{kl}(z_j-z_{oj})(z_l-z_{ol})$$

$$+ \frac{\theta_{k1}}{\gamma^2{}_{k1}}\sum_{j=1}^{J}\gamma_{jk}^2(z_j-z_{oj})^2 + R_{k1}(t-t_{ok})^3 + \sum_{j=1}^{J}R_{k2j}(z_j-z_{oj})^3$$

for
ただし

$$\alpha_k = \lambda_o(t_{ok})g(\underset{\sim k}{b}\,\underset{\sim}{z})$$

$$\beta_k = \lambda_o'(t_{ok})g(\underset{\sim k}{b}\,\underset{\sim o}{z})$$

$$\gamma_{kj} = b_{kj}\lambda_o(t_{ok})g'(\underset{\sim k}{b}\,\underset{\sim o}{z}) \qquad j=1, \ldots, J \tag{14}$$

$$\delta_k = ½\lambda_o''(t_{ok})g(\underset{\sim k}{b}\,\underset{\sim o}{z})$$

$$\theta_{kl} = ½ b_{kl}^2\lambda_o(t_{ok})g''(\underset{\sim k}{b}\,\underset{\sim o}{z})$$

$$R_{kl} \simeq \frac{1}{6}\lambda_o'''(t_{ok})g(\underset{\sim k}{b}\,\underset{\sim o}{z})$$

$$R_{k2j} \simeq \frac{1}{6}b_{kj}^3\lambda_o(t_{ok})g'''(\underset{\sim k}{b}\,\underset{\sim o}{z}) \qquad j=1, \ldots, J\,.$$

Substituting the explicit dependence of the parameters on $\lambda_0(\cdot)$ and $g(\cdot)$ into (14) will show with the exception of the remainder term that (14) depends upon the covariates solely through $\underset{\sim}{b}_k\underset{\sim}{z}$.

The likelihood ratio test of the proportional hazards assumption is dependent upon the alternative hypothesis considered. If the alternative is the general regression model, equations (13) yield only one additional constraint per interval. Thus, for K intervals the test statistic is approximately χ^2_K. If the alternative is the general model, equations (7) and (13) total $\frac{J(J+3)}{2} - 1$ restrictions per interval. Interestingly, with only one covariate proportional hazards is testable because (13) must still hold even though equations (7) are undefined. Also, as before these tests can be performed on subsets of covariates and on an interval by interval basis.

$\lambda_0(\cdot)$ 及び $g(\cdot)$ に明白に依存するパラメーターを(14)に代入すると，剰余項を除いて(14)が $\underset{\sim}{b}_k\underset{\sim}{z}$ によって共変量に単独で依存する.

比例ハザード仮説の尤度比検定は考慮する択一的仮説によって異なる. 代わりの方法が一般回帰モデルであれば，方程式(13)では一つの期間当たりの追加定数はわずか一つである. したがって，K個の期間の検定統計量は約 χ^2_K となる. 一般モデルを用いると，方程式(7)及び(13)では1期間当たり合計 $\frac{J(J+3)}{2} -$ 1の制限となる. 面白いことに，たとえ方程式(7)が定義されていなくても(13)は成立するはずなので，一つの共変量を用いただけで比例ハザードは検定できる. また，先程と同様にこれらの検定は共変量の小部分について，また，各期間ごとにも行える.

Tests Related to Proportional Hazards

Once it has been decided that the data are consistent with a proportional hazards model, it is interesting to examine if the regression coefficients vary with time. The corresponding null hypotheses can be tested because (14) implies

比例ハザードに関連する検定

資料が比例ハザードモデルに一致することが決定すれば，回帰係数が経時的に変化するか否かを調べると面白い. (14)から

$$\frac{\gamma_{kj}}{\alpha} = b_{kj}\left[\frac{d}{dx}\ln g(x)\Big|_{x=\underset{\sim}{b}_k\underset{\sim}{z}_o}\right] \qquad j=1,\ldots,J;\ k=1,\ldots,K \qquad (15)$$

Thus, for interval k

であるので，対応する帰無仮説が検定できる. したがって，期間kにおいて

$$\gamma_{kj} = c_{kj}\alpha_k \qquad (16)$$

and to test for a time independent coefficient for covariate j is equivalent to testing c_{kj} being constant over time. Numerically, this test may pose some difficulties because it will require the evaluation of the entire log likelihood. Fortunately, there appears to be at least two possible solutions to this problem. First, if a Newton-Raphson type algorithm is to be applied, the method discussed in the Appendix can be used with some modifications. Since good initial values will be available as the maximum likelihood estimates needed for the tests of the proportional hazards assumption (above), this approach appears feasible if the Hessian matrix can be inverted. A second approach is to utilize a Monte Carlo optimizer with these initial values. We have had reasonable good experience[11] with

であり，共変量jについて時間依存係数を検定することは，経時的に一定である c_{kj} を検定することと等しい. この検定は完全な対数尤度の評価を必要とするので，数値的に若干困難な点があるかもしれない. しかし，幸いなことにこの問題に関しては少なくとも二つの解決法があるようである. 第一は，Newton-Raphson 型アルゴリズムを応用すると，付録で検討した方法を若干修正して用いることができる. 比例ハザード仮説(前述)の検定に必要な最大尤度評価量としてよい初期値が使用できるので，Hessian 行列を逆転できれば，この方法は使用可能と思われる. 第二の方法はこれらの初期値とともに Monte Carlo 最適化法を利用することである. 著者らは

one proposed by Bremmermann.[12] However, a major disadvantage of this latter approach is the lack of an estimated variance covariance matrix for the estimates.

Bremmermann[12] の提起した方法を用いてかなりよい結果を得たが,[11] この第二の方法の主な欠点は推定値に関する推定変量・共変量行列がないことである.

It must be emphasized that the three testing procedures considered thus far for proportional hazards are independent of the covariate function. We now consider testing the adequacy of specific choices of this function. From equations (14), it follows that without specifying some properties of the $\lambda_o(\cdot)$ function all information regarding the covariate function in the β_k's and δ_k's is confounded with information about derivatives of $\lambda_o(\cdot)$. However, there is useful information contained in the α_k's, γ_{kj}'s and θ_{k1}'s which can be extracted. Usually, a specific choice of g(x) will constrain only θ_1 but this may not be true in general. Three examples of g(x) are e^x, 1+x, and cos x + sin x. For these choices the restricted values for θ_{k1} are $\frac{\gamma^2_{k1}}{2\alpha_k}$, 0, $-\frac{\gamma^2_{k1}}{2\alpha_k}$, respectively. Thus, the test statistics for these choices of g(x) against a proportional hazards model each have χ^2_K distributions. Notice there is no formal method for deciding between two "acceptable" choices of the covariate function if both functions have the same number of restrictions. This problem is common to the partial likelihood methodology as well.

これまで考慮してきた三つの比例ハザード検定法は,共変量関数に依存していないことを強調しておかねばならない.ここではこの関数の特定選択の適合性の検定について考える.方程式(14)から,$\lambda_o(\cdot)$関数の幾つかの属性を指適しなければ,β_k及びδ_kにおける共変量関数に関するすべての情報が$\lambda_o(\cdot)$の導関数についての情報と混同されるようになる.しかし,α_k, γ_{kj}及びθ_{k1}には抽出できる有効な情報が含まれている.g(x)の特定選択はθ_1のみを束縛することが多いが,総体的にそうであるとは言い切れない.g(x)の三つの例はe^x, 1+x 及び cos x + sin x である.これらの選択に対しては,θ_{k1}の制限値はそれぞれ$\frac{\gamma^2_{k1}}{2\alpha_k}$, 0, $-\frac{\gamma^2_{k1}}{2\alpha_k}$である.したがって,これらg(x)の選択の各比例ハザードモデルに対する検定統計量はχ^2_K分布を有する.共変量関数の二つの「許容できる」選択については,双方が同数の制限をもつ場合これらを決定する形式的な方法はない.この問題は部分尤度法にも共通する.

Estimation Under Proportional Hazards

For the general regression model, it was shown that the regression coefficients could not be estimated without an additional assumption regarding the hazard function. The proportional hazards assumption implies

比例ハザードにおける推定

一般回帰モデルでは,ハザード関数に関してもう一つの仮説がないと回帰係数が推定できないことが示された.比例ハザード仮説では

$$\frac{\partial^2}{\partial t \partial x}\lambda(t|x) = \frac{[\frac{\partial}{\partial t}\lambda(t|x)]\,[\frac{\partial}{\partial x}(t|x)]}{\lambda(t|x)} \qquad \forall t, x \tag{17}$$

which yields equations (13) and is sufficient to guarantee identifiability of the regression coefficients. The identifiability follows because the general regression situation is analogous to having one variable more than independent equations and (17) is the additional equation. Unfortunately, any nonzero solution to (17) is a proportional hazards model and thus (17) does not yield a richer class of models than propor-

となり,これから方程式(13)が導かれて,回帰係数の識別が可能であることが十分証明される.識別が可能となるのは一般回帰状態が独立方程式よりももう一つ多くの変数をもっていることと同様であり,その方程式は(17)である.残念なことに(17)のゼロでない解は比例ハザードモデルであり,(17)は比例ハザードよりも豊かなクラスのモデルを導くことはない.した

tional hazards. Hence, if the proportional hazards assumption is violated within the class of regression models, one should seriously question the estimated values of the proportional hazards regression coefficients.

Surprisingly, in addition to the various testing procedures already discussed the Taylor series methodology yields maximum likelihood estimates of the proportional hazards regression coefficients which are fairly insensitive to particular choices of the covariate function. This insensitivity is in direct contrast with methods based upon the partial likelihood, which are known to be sensitive to particular choices of g(x). This sensitivity follows because the partial likelihood is explicitly dependent upon the chosen g(x) function[13] and if this choice is incorrect, standard maximum likelihood theory[5] implies possibly inconsistent estimates. In contrast, the sole dependence on the covariate function of the Taylor series estimates, $\hat{b}_{kj}$, is through the derivative on the right hand side of (15). Without loss of generality the $\underset{\sim}{z}_o$ vector can be taken to be the zero vector and then the dependence is via

がって，もし比例ハザード仮説が回帰モデルのクラス内で崩れれば，比例ハザード回帰係数の推定値について大いに疑いをもたねばならない．

驚くべきことに，既に取り上げた多くの検定法のほかに，テイラー級数法は共変量関数の特定選択に対して，かなり感度の低い比例ハザード回帰係数の最大尤度推定量を導きだす．この感度の低さは，g(x)の特定選択に対して感度が高いことで知られている部分尤度に基づく方法と直接的な対比をなしている．後者の方法の感度が高いのは，部分尤度が選択されたg(x)関数に明らかに依存しており，[13] この選択が正しくなければ，標準最尤度理論[5] による推定値は恐らく一致しないものになるからである．一方，テイラー級数推定値 $\hat{b}_{kj}$ の共変量関数に単独に依存するのは(15)の右側の導関数を通じてである．一般性を失わずに z_o ベクトルはゼロベクトルと考えられるので，依存は

$$\left.\frac{d}{dx}\ln g(x)\right|_{x=0} \tag{18}$$

Thus, within a set of covariate functions with equal values of (18) the $\hat{b}_{kj}$'s are entirely robust. Five disperate choices of g(x) with equal values of (18) are e^x, $1+x$, $\cos x + \sin x$, $(1-x)^{-1}$, and $\exp(e^x)$.

A second estimation area where the Taylor series methodology introduces an approach different from the standard one is in choosing time dependent covariates. For any particular covariate, the $\hat{c}_{kj}$'s based on (16) and (18) can be plotted vs time. From the resultant graph, a functional form in time, such as log t, for a time dependent covariate can be ascertained. As before, this ascertainment can be done without specifying the covariate function. Also, likelihood ratio tests can be constructed to test the adequacy of the chosen functional form. In contrast, it is common practice with partial likelihood methods to choose functional forms which augment the chosen covariate function. Clearly, the adequacy of these choices depends upon the chosen g(x).

によるものとなる．したがって，(18)の等しい値をもつ共変量関数の一組の中では，$\hat{b}_{kj}$ は完全にロバストである．(18)と等しい値をもつg(x)の五つのdisperate 選択は e^x，$1+x$，$\cos x + \sin x$，$(1-x)^{-1}$ 及び $\exp(e^x)$ である．

テイラー級数法で標準のものとは異なる方法が導入された第二の推定の行われる分野は，時間依存性共変量の選択である．どのような特定共変量についても，(16)及び(18)に基づく $\hat{c}_{kj}$ は時間に対してプロットできる．できたグラフから，時間依存性共変量について log t のような時間における関数形式を確認できる．先の場合と同様に，この確認は共変量関数を指定しなくてもできる．また，尤度比検定は選択した関数形式の適合性を検定するために構成できる．一方，選択した共変量関数を増す関数形式を選択することは，部分尤度法に共通した手順である．これらの選択の適合性は明らかに選択したg(x)に依存している．

CLOSING REMARKS

We have attempted to outline a general testing approach to analyzing survival data. Obviously, the developed techniques are dependent upon the assumption regarding the remainder term in equation (3). In our opinion this assumption combined with the interval approach is reasonable for many cases. Furthermore, the method could be modified to include the entire third term of the Taylor series and then assume the remainder is negligible. Given a large data set and good computing facilities if J is not too large this approach may be feasible. However, there is a point where accuracy within a statistical model exceeds the accuracy of the model in approximating the real world situation. Clearly, these types of decisions should be made on an individual basis.

With competing risks, it is well known that the log likelihood becomes a sum of cause-specific log likelihoods. Thus, extending the Taylor series methods to competing risks is very easy. Unless there are model constraints on the parameters across failure types, the analysis proceeds failure type by failure type and there are no additional difficulties.

Interestingly, if multiple risks can simultaneously cause failure, then constraints across these failure types can be utilized to develop tests of independent competing risks.[11]

In closing we point out that the Taylor series approach is sufficiently general to include many situations not considered here. Three such examples are multistage failure models, Markov Illness-Death Models, and stratification within proportional hazards models.

結　語

生存資料を解析するための一般的検定法を概説した．明らかに，開発された技法は方程式(3)の剰余項に関する仮説によって左右される．著者らの意見では，期間法とこの仮説を結合すると多くの場合に合理的な方法となる．更に，この方法を修正してテイラー級数の完全な第三項を含めるようにし，剰余項を無視できるものと仮定することができる．資料が数多く，よい計算設備があり，Jが大き過ぎなければこの方法は実行できるであろう．しかし統計モデル内の正確性が，実際の情況を近似するときの正確性よりも高くなる点がある．このような場合の決定は，個々の人を基にして行うべきであることは明白である．

競合リスクに関しては，対数尤度が死因別対数尤度の合計となることはよく知られている．したがって，テイラー級数法を競合リスクに適用することは極めて容易である．種々のfailure型にわたるパラメーターに関するモデルの制限がなければ，解析はfailure型ごとに進めることができ，それ以上の問題はない．

面白いことに，複数のリスクが同時にfailureを導くことができるならば，これらのfailureの型に関する制限を独立競合リスクの検定法を開発するのに利用できる．[11]

最後に，テイラー級数法は十分に普遍的なものであって，本稿で検討しなかった多くの情況を含んでいる．その例としては，多段階failureモデル，Markovの疾患一死亡モデル及び比例ハザードモデル内の層化の三つがある．

APPENDIX
付　録

The purpose of this appendix is to demonstrate a computer adaptable method for obtaining first and second order partial derivatives of the log likelihood. The method is to first obtain the derivatives under the general model and then use the Chain Rule from Calculus to "correct" for other models. We consider an arbitrary interval k and use the notation of the text.

この付録の目的は，対数尤度の第1位及び第2位の部分導関数を得るためコンピューターに適応する方法を提示することである．この方法としては，まず一般的モデルのもとで導関数を得，次にその他のモデルに対しては微積分による連鎖法則(Chain Rule from Calculus)を用いて"修正"を行う．任意の区間kを考え，テキストの記法を使用する．

Consider the general model. For log likelihood, ℓ_k, the first order partial derivatives are:

一般的なモデルを考えてみたい．対数尤度ℓ_kについては，第1位部分導関数は次のとおりである：

$$\frac{\partial \ell_k}{\partial \alpha_k} = \sum_{i=1}^{N} \frac{I_{ki}}{\lambda_k(T_{ki}|\underset{\sim}{z}_i)} - \sum_{i=1}^{N} (T_{ki} - S_k)$$

$$\frac{\partial \ell_k}{\partial \beta_k} = \sum_{i=1}^{N} \frac{I_{ki}}{\lambda_k(T_{ki}|\underset{\sim}{z}_i)} (T_{ki} - t_{ok}) - \tfrac{1}{2} \sum_{i=1}^{N} [(T_{ki} - t_{ok})^2 - (S_k - t_{ok})^2]$$

$$\frac{\partial \ell_k}{\partial \gamma_{kj}} = \sum_{i=1}^{N} \frac{I_{ki}}{\lambda_k(T_{ki}|\underset{\sim}{z}_i)} (z_{ij} - z_{oj}) - \sum_{i=1}^{N} (z_{ij} - z_{oj})(T_{ki} - S_k), \qquad j=1, \ldots, J$$

$$\frac{\partial \ell_k}{\partial \delta_k} = \sum_{i=1}^{N} \frac{I_{ki}}{\lambda_k(T_{ki}|\underset{\sim}{z}_i)} (T_{ki} - t_{ok})^2 - \frac{1}{3} \sum_{i=1}^{N} [(T_{ki} - t_{ok})^3 - (S_k - t_{ok})^3]$$

$$\frac{\partial \ell_k}{\partial \epsilon_{kj}} = \sum_{i=1}^{N} \frac{I_{ki}}{\lambda_k(T_{ki}|\underset{\sim}{z}_i)} (T_{ki} - t_{ok})(z_{ij} - z_{oj}) - \tfrac{1}{2} \sum_{i=1}^{N} (z_{ij} - z_{oj})[(T_{ki} - t_{ok})^2 - (S_k - t_{ok})^2], \qquad j=1, \ldots, J$$

$$\frac{\partial \ell_k}{\partial \eta_{kjl}} = \sum_{i=1}^{N} \frac{I_{ki}}{\lambda_k(T_{ki}|\underset{\sim}{z}_i)} (z_{ij} - z_{oj})(z_{il} - z_{ol}) - \sum_{i=1}^{N} (z_{ij} - z_{oj})(z_{il} - z_{ol})(T_{ki} - S_k), \quad 1 \leq j < l \leq J \qquad \text{(A.1)}$$

$$\frac{\partial \ell_k}{\partial \theta_{kj}} = \sum_{i=1}^{N} \frac{I_{ki}}{\lambda_k(T_{ki}|\underset{\sim}{z}_i)} (z_{ij} - z_{oj})^2 - \sum_{i=1}^{N} (z_{ij} - z_{oj})^2 (T_{ki} - S_k), \qquad j=1, \ldots, J$$

$$\frac{\partial \ell_k}{\partial R_{k1}} = \sum_{i=1}^{N} \frac{I_{ki}}{\lambda_k(T_{ki}|\underset{\sim}{z}_i)} (T_{ki} - t_{ok})^3 - \tfrac{1}{4} \sum_{i=1}^{N} [(T_{ki} - t_{ok})^4 - (S_k - t_{ok})^4]$$

$$\frac{\partial \ell_k}{\partial R_{k2j}} = \sum_{i=1}^{N} \frac{I_{ki}}{\lambda_k(T_{ki}|\underset{\sim}{z}_i)} (z_{ij} - z_{oj})^3 - \sum_{i=1}^{N} (z_{ij} - z_{oj})^3 (T_{ki} - S_k), \qquad j=1, \ldots, J$$

Notice that all derivatives are composed of two sums. The first sum involves the reciprocal of the hazard of only those individuals who fail from the cause of interest during interval k and the second sum is a constant. The second sum, which is based upon all individuals at risk at the beginning of the kth interval, needs only to be evaluated once. Thus, for a derivative-based iterative algorithm the hazards of those individuals who fail from the cause of interest in interval k need to be evaluated only once per iteration and then the reciprocals multiplied by the proper factor for each derivative. For those individuals who do not fail from the cause of interest in the interval there is no computation required once the constant sums have been computed.

導関数はすべて二つの和から成ることに注目すべきである．第1の和は区間kのcause of interestに適合しない者のみのハザードの逆数を含み，第2の和は定数である．第k番目の区間の初めの観察対象全例に基づく第2の和は，1回のみ評価される必要がある．したがって導関数に基づく反復アルゴリズムについては，区間kのcause of interestに適合しない者のハザードは，反復当たり1回だけは評価され る必要があり，次いで逆数は各導関数ごとに適当な係数で乗じる必要がある．この区間においてcause of interestに適合しない者について，定数合計の算定が完了すれば計算の必要はない．

After the first partials are obtained, the second partials follow immediately. Let p_j represent parameter number j in the general model. Then

第１の部分が得られると，第２の部分はそれに続く．p_j が一般モデルのパラメーター番号 j を表すものとすれば，

$$\frac{\partial \ell_k}{\partial p_j} = \sum_{i=1}^{N} \frac{I_{ki}}{\lambda_k(T_{ki}|\underset{\sim}{z}_i)} \frac{\partial \lambda_k(T_{ki}|\underset{\sim}{z}_i)}{\partial p_j} - \text{const } (p_j) \tag{A.2}$$

where const (p_j) is the constant sum in the derivative with respect to p_j in (A.1). It is easily verified that for parameters, p_r and p_s,

ただし const(p_j)は，（A.1）における p_j に関する導関数の定数の和である．パラメーター p_r 及び p_s については，次のことが容易に証明される：

$$\frac{\partial^2 \ell_k}{\partial p_r \partial p_s} = -\sum_{i=1}^{N} \frac{I_{ki}}{\lambda^2 k(T_{ki}|\underset{\sim}{z}_i)} \cdot \frac{\partial \lambda_k(T_{ki}|\underset{\sim}{z}_i)}{\partial p_r} \cdot \frac{\partial \lambda_k(T_{ki}|\underset{\sim}{z}_i)}{\partial p_s} \tag{A.3}$$

Equations (A.1), (A.2), and (A.3) in concert are easily implemented on a computer. The second order partial with respect to p_r is obtained by setting s equal to r in (A.3).

方程式（A.1），（A.2）及び（A.3）はいずれもコンピューターで容易に演算できる．p_r に関する第２位の部分は s を方程式（A.3）の r に適用すれば得られる．

Now consider how to "correct" the derivatives for the other models. We detail an example for proportional hazards which is indicative of the general procedure. Equations (5) and (11) imply

さて，その他のモデルについてはいかに導関数を"修正"するかを考えたい．一般手続きを示す比例ハザードの１例について詳述する．方程式(5)及び(11)は，次のことを示唆する．

$$\epsilon_{kj} = \frac{\beta_k}{\alpha_k} \gamma_{kj} \qquad , j=1, \ldots, J$$

$$\eta_{kjl} = 2\,\frac{\gamma_{kj}\gamma_{kl}}{\gamma^2{}_{kl}}\,\theta_{kl} \qquad , 1 \leqq j < l \leqq J$$

$$\theta_{kj} = \left(\frac{\gamma_{kj}}{\gamma_{kl}}\right)^2 \theta_{kl} \qquad , j=1, \ldots, J \; .$$

For j>1, the Chain Rule yields

j > 1 の場合は，連鎖法則によって下記が得られる．

$$\begin{aligned} \frac{\Delta \lambda_k(T_{ki}|\underset{\sim}{z}_i)}{\Delta \gamma_{kj}} &= \frac{\partial \lambda_k(T_{ki}|\underset{\sim}{z}_i)}{\partial \gamma_{kj}} + \frac{\beta_k}{\alpha_k} \frac{\partial \lambda_k(T_{ki}|\underset{\sim}{z}_i)}{\partial \epsilon_{kj}} \\ &+ 2\,\frac{\theta_{kl}}{\gamma^2{}_{kl}} \sum_{l>j} \gamma_{kl} \frac{\partial \lambda_k(T_{ki}|\underset{\sim}{z}_i)}{\partial \eta_{kjl}} + 2\,\frac{\gamma_{kj}\theta_{kj}}{\gamma^2{}_{kl}} \frac{\partial \lambda_k(T_{ki}|\underset{\sim}{z}_i)}{\partial \theta_{kj}} \end{aligned} \tag{A.4}$$

where the left hand side is the "model corrected" partial derivative of $\lambda_k(T_{ki}|z_i)$ and the partial derivatives on the right hand side are obtained under the general model in equation (A.2). The proposed procedure is to evaluate the derivatives

ただし，左側は $\lambda_k(T_{ki}|z_i)$ の"モデル修正を行った"部分導関数であり，右側の部分導関数は方程式（A.2）の一般的モデルによって得られる．ここで述べる手続きは，各例に関する一般的モデルによってハザード

of the hazard under the general model for each individual and then "correct" for a specific model as indicated in equation (A.4). Letting n be the number of parameters in the general model, the Chain Rule yields

の導関数を評価し，また方程式(A.4)に示した特定のモデルについて"修正"を行うことである．一般モデルにおけるパラメーターの数をnとすれば，連鎖法則によって次のようになる．

$$\frac{\Delta\lambda_k(T_{ki}|\underset{\sim}{z}_i)}{\Delta p_r} = \sum_{j=1}^{N} \frac{\partial p_j}{\partial p_r} \frac{\partial\lambda_k(T_{ki}|\underset{\sim}{z}_i)}{\partial p_j} \tag{A.5}$$

$$\frac{\Delta\ell_k}{\Delta p_r} = \sum_{i=1}^{N} \frac{I_{ki}}{\lambda_k(T_{ki}|\underset{\sim}{z}_i)} \left(\frac{\Delta\lambda_k(T_{ki}|\underset{\sim}{z}_i)}{\Delta p_r}\right) - \sum_{j=1}^{N}\left(\frac{\partial p_j}{\partial p_r}\right) \text{const } (p_j) \tag{A.6}$$

and また，

$$\frac{\Delta^2\ell_k}{\Delta p_r \Delta p_s} = -\sum_{i=1}^{N} \frac{I_{ki}}{\lambda^2{}_k(T_{ki}|\underset{\sim}{z}_i)} \left(\frac{\Delta\lambda_k(T_{ki}|\underset{\sim}{z}_i)}{\Delta p_r}\right)\left(\frac{\Delta\lambda_k(T_{ki}|\underset{\sim}{z}_i)}{\Delta p_s}\right)$$

$$-\sum_{j=1}^{n} \frac{\partial^2 p_j}{\partial p_r \partial p_s}\left(\text{const } (p_j) - \sum_{i=1}^{N} \frac{I_{ki}}{\lambda_k(T_{ki}|\underset{\sim}{z}_i)}\right) \tag{A.7}$$

Therefore, once equations (A.1), (A.2), and (A.3) are implemented on a computer, the "correction" equations, (A.5)-(A.7) can be implemented with only minor difficulty.

したがって方程式(A.1)，(A.2)及び(A.3)がコンピューターで1回算出されると，"修正"方程式(A.5)～(A.7)はほとんど困難なく算出できる．

REFERENCES
参考文献

1. COX DR: Regression models and life-tables (with discussion). J R Statistical Soc B 34:187-220, 1972

2. TAULBEE J: A general model for the hazard rates with covariates. Biometrics 35:439-50, 1979

3. KALBFLEISH JD, PRENTICE RL: The statistical analysis of failure time data. New York, John Wiley and Sons. 1980

4. CRĀMER H: Mathematical methods of statistics. Princeton, Princeton University Press. 1946

5. ROSENLICHT M: Introduction to analysis. Glenview, Scott, Foresman, and Co. 1968

6. BRODSKY JB, LIDDELL R, ISHIMARU T, GROER PG, ICHIMARU M: Temporal analysis of a dose-response relationship: Leukemia mortality in atomic bomb survivors. RERF TR 5-82

7. LOEWE WE, MENDELSOHN E: Revised dose estimates at Hiroshima and Nagasaki. Health Physics, to appear.

8. MARSHALL E: Japanese A-bomb data will be revised. Science 214:31-2, 1981

9. NATIONAL ACADEMY OF SCIENCES – NATIONAL RESEARCH COUNCIL: The effects on population of exposure to low levels of ionizing radiation 1980. Report of the Advisory Committee on the Biological Effects of Ionizing Radiation. (BEIR Report 3, 1980)

10. BROWN JM: The shape of the dose-response curve for radiation carcinogenesis. Extrapolation to low doses. Radiat Res 71:34-50, 1977

11. BERLIN B, BRODSKY JB, CLIFFORD P: Testing disease dependence in experiments with serial sacrifice. JASA 74:5-14, 1979

12. BREMMERMANN H: A method of unconstrained global optimization. Math Biosciences: 9-15, 1970

13. COX DR: Partial likelihood. Biometrika 62:269-76, 1975