数理解析研究所講究録
第384巻 1980年 122-142

脱落のある寿命データの数理解析

九大 理　柳川 堯

1. 序

1.1. 脱落のある寿命データ　生命科学における寿命データは動物実験データと臨床医学データに大別できる。両者の特長はつぎの通りである。

動物実験データ：例として発がん性が疑われている化学物質の発がん性テストを考えよう。実験動物はランダムに k 群に分けられ、各群に対して相異なる量の化学物質が割りつけられ、ランダムに各個体に化学物質が投与される。厳密な管理の下で腫瘍発症までの日数が観察される。かかる状況の下では各群内の観測値は互いに独立、同一分布に従う確率変数の実現値とみなすことができ数学モデルを導入できる。しかしながら、実験途中で実験動物が他の原因で死亡したり、ある期間を経過したら観察が打切られることが多く得られるデータは不完全データである場合が多い。

臨床医学データ：ある事象（治癒、がんの転移、死亡など）が発生するまでの期間が観察され、得られるデータに基づいて治療効果が比較される。このとき、患者が突然来院しなくなったり、転院したりするため動物実験の場合と同様なデータの脱落、打切りの問題が頻繁に生じるほか、ヒトの生存には年令、性別、生活環境などさまざまな因子が関与してくるからこれらを考慮した数学モデルをたてデータ解析を行うことが必要である。

これら2種類のデータは、寿命の長さを表わす変量のほかに共変量 $\boldsymbol{z}=(z_1, z_2, \cdots, z_p)$ を導入することによって統一的に取り扱うことができる。例えば、動物実験データでは $p=1$ とし、$z_1=1$（群1），$z_1=2$（群2），…，$z_1=K$（群K）なるダミー変数 z_1 を導入すればよい。また、臨床医学データでは、このほか、z_2 を性別、z_3 を年令，… などを表わす変量に定めればよい。

2節で示されるが、打切りは脱落の特別の場合として定式化できる。従がって、本稿では脱落や打切りのあるデータを総称して脱落のあるデータとよぶことにする。脱落のあるデータは、少なくとも個体がある期間生存したという情報を持っている。これを捨てることは、せっかく得られた情報を利用しないばかりか、結果にカタヨリを与えることにも

なりかねない。

1.2. 本稿の目的　Cox (1972) は脱落のある寿命データを、共変量を考慮しながら解析する強力な方法を提案した。Cox法による生命科学データの解析は英米で大流行である。Cox法の数学的側面を明らかにする研究も盛んになりつつある。特に従来のノンパラメトリック法の流れの中でCox法を位置づける試みは面白そうであるがまだ部分的にしか成功していない。本稿では、寿命分布が連続である場合に話を制限し、Cox法の数学的側面に焦点を当て、その特長を考察することを目的とする。

2. 定式化

脱落のある寿命データを数学的につぎのように定式化する。

- $T_1^\circ, T_2^\circ, \cdots, T_N^\circ$; 互いに独立な確率変数

T_i° は個体 i の真の寿命に対応する確率変数である。脱落があるため観測可能とは限らない。

- $F_i(t) = P\{T_i^\circ \leq t\}$; 個体 i の生存関数、連続性を仮定する。
- $f_i(t) = -dF_i(t)/dt$
- $\lambda_i(t) = \dfrac{f_i(t)}{F_i(t)} = \lim_{\Delta \to 0} \dfrac{P\{t \leq T_i^\circ < t+\Delta \mid T_i^\circ \geq t\}}{\Delta}$

$\lambda_i(t)$ はハザード関数とよばれる。個体 i の時刻 t における瞬間死亡率*である。つぎの関係式が成立することは明らかである。

$$F_i(t) = \exp\left\{-\int_0^t \lambda_i(x)\,dx\right\} \qquad (2.1)$$

- $U_1, U_2, \cdots, U_N$; 互いに独立な確率変数、T_i^0 とも独立。
- $H_i(t) = P\{U_i \geq t\}$; 連続性を仮定する。
- $h_i(t) = -dH_i(t)/dt$

U_i は個体 i の脱落に関する確率変数である。

- $T_i = \min(T_i^0, U_i)$
- $\varepsilon_i = \begin{cases} 1 & ; \quad T_i = T_i^0 \quad (\text{死亡}) \\ 0 & ; \quad T_i > U_i \quad (\text{脱落}) \end{cases}$

脱落のある寿命データでは (T_i, ε_i), $i=1,2,\cdots,N$ が観測可能である。問題は (T_i, ε_i), $i=1,2,\cdots,N$ を観測して得られるデータに基づいて $\lambda_i(t)$ の推定, $\lambda_i(t)$ に関する仮説の検定、および リスク要因の分析を行なうこと、いいかえれば、リスク要因を分析し、それによって調整しながら生存関数 $F_i(t)$ を推定したり、検定したりすることである。ここで、$F_i(t)$ として、工業製品の寿命試験の類似として、未知母数を持つ指数分布、ワイブル分布、対数正規分布など、$H_i(t)$ として指

---

* 本稿では事象の発生を“死亡”とよぶことにする。

数分布や一様分布などが考えうれる。しかしながら、一般に生命科学データでは生物の個体差、実験や観察が理想状態で行なえるなどのため特定の分布型が指定ができず F_i, H_i は未知とされることが多い。ここでは、F_i, H_i に特定の分布型を指定しない、いいかえればノンパラメトリックな枠組で問題を考えようという立場である。

PROPOSITION

$$P\{T_j \geq t, \varepsilon_j = 1\} = \int_t^\infty f_j(x) H_j(x)\, dx$$

$$P\{T_j \geq t, \varepsilon_j = 0\} = F_j(t) H_j(t) - \int_t^\infty f_j(x) H_j(x)$$

$$\lim_{\Delta \to 0} \frac{P\{t \leq T_j < t+\Delta, \varepsilon_j = 1\}}{\Delta} = f_j(t) H_j(t)$$

$$\lim_{\Delta \to 0} \frac{P\{t \leq T_j < t+\Delta, \varepsilon_j = 0\}}{\Delta} = F_j(t) h_j(t)$$

(証明は簡単なので省略)

Proposition より (T_i, ε_i), $i = 1, 2, \cdots, N$ の同時密度関数は次式で与えられる。

$$L_0(t_1, \cdots, t_N, \varepsilon_1, \cdots, \varepsilon_N) = \prod_{j=1}^N [f_j(t_j) H_j(t_j)]^{\varepsilon_j} [F_j(t_j) h_j(t_j)]^{1-\varepsilon_j}$$

$$= \left\{ \prod_{j=1}^N [\lambda_j(t_j)]^{\varepsilon_j} \exp\left[-\int_0^{t_j} \lambda_j(t)\, dt \right] \right\} \left\{ \prod_{j=1}^N [H_j(t_j)]^{\varepsilon_j} [h_j(t_j)]^{1-\varepsilon_j} \right\}$$

これが $(t_i, \varepsilon_i)_{i=1 \sim N}$ が与えられたときの全尤度関数である。H_j に依存しない推論方式を得るには U_i, $i = 1 \sim N$ を与えたときの (T_i, ε_i), $i = 1 \sim N$ の条件付き尤度関数から出発する。これは簡単な計算より次式で与えられる。以後、本稿ではこの L_C に基づいて考える。

$$L_C(t_1,\cdots,t_N;\varepsilon_1,\cdots,\varepsilon_N)=\prod_{j=1}^{N}[\lambda_j(t_j)]^{\varepsilon_j}\exp\left[-\int_0^{t_j}\lambda_j(t)dt\right] \quad (2.2)$$

3. Cox の方法

Cox (1972) の方法は2つの斬新なアイディアよりなる。その一つは PH モデルの導入、他は部分尤度法の導入である。

3.1. PH モデル

$$\ln(\lambda_i(t)/\lambda_0(t))=\beta_1 z_{1i}+\cdots+\beta_p z_{pi}=\beta' z_i \quad (3.1)$$

とおき、対数変化された瞬間死亡率の対数を共変量 $z_1,\cdots,z_p$ の線形結合で説明しようというものである。Cox (1958) のロジスティックモデルの応用にほかならない。λ_0 は z の基準値に対する瞬間死亡率である。全国平均瞬間死亡率などが利用できる場合、それを λ_0 とすれば問題は簡単であるが、ここでは λ_0 は未知である。z_i から平均を差引いた $z_i-\bar{z}$ を改めて z_i とおき、$z_i=0$ のときの瞬間死亡率が λ_0 であると理解すれば便利がある。Cox は z が時間に関係する変量を含むことを許しているが、モデルの一意性という点で若干の問題があるようである。ここでは、z は時間に関係する項を含まないものとしておく。

(3.1) を (2.2) 式に代入し対数部分尤度関数を求めると

$$\ln L_c(t,\varepsilon) = \sum_{j=1}^{N}[\varepsilon_j \beta' z_j + \varepsilon_j \ln \lambda_0(t_j)] - \sum_{j=1}^{N} e^{\beta' z_j} \int_0^{t} \lambda_0(x)dx \quad (3.2)$$

λ_0 に対応する密度関数を f_0 とする。f_0 として、z に無関係な未知母数を含む指数分布，ワイブル分布，対数正規分布などを想定し、これらの未知母数と β に関して (3.2) 式を最大にすることによって β の推定量を求めることが出来る。実際松原・後藤 [1979] は f_0 として各正規分布を仮定し、適応的に分布型を推定しながら β の推定を行なう意欲的な方法を提案している。しかしながら，これらはつぎの2点でやっかいな問題を含んでいる。

(1) 計算が（Cox 法にくらべ）極めて面倒である。

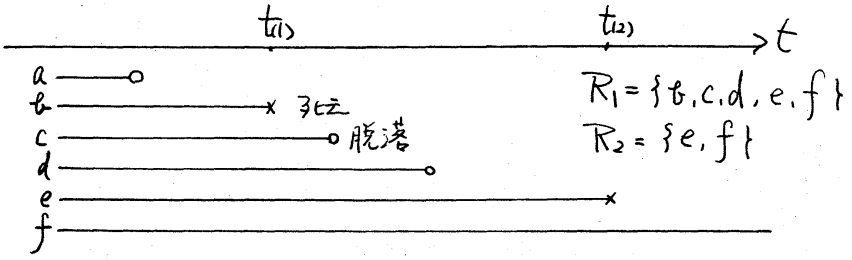
(2) PH モデルでは我々の関心は、主に β に関する推定、検定にある。このとき、λ_0 は局外母数である。局外母数が存在するときの最尤推定量は，特にサンプル数が少ないとき、良く知られているように理論的に問題があり信頼できない場合が多い。また、分布形の指定から生じる specification error の影響も無視できない。

3.2 部分尤度法

まず、つぎのようにデータをいわゆる生命表形式にまとめることより出発する。

(3.3) $\begin{cases} t_{(1)} < t_{(2)} < \cdots < t_{(\ell)} \ ; \text{死亡時刻} \\ i_j \ ; \ t_{(j)} \text{で死亡した個体名} \\ R_j = R(t_{(j)}) \ ; \text{時刻 } t_{(j)}-0 \text{で生存しており，かつ観察下にある個体名のリスト，リスク集合という。} \end{cases}$

例>



注1> F_Xの連続性が仮定されているが、実際には死亡時刻は日で記録されたりするため同時刻に複数個の個体が死亡することも多い。このときの取り扱いは付録にゆずり、ここでは簡単のため重複がないものとしておく。

注2> 上のようにデータをまとめることは、区間$(t_{(j)}, t_{(j+1)})$で生じたすべての脱落を時刻$t_{(j)}+0$で生じたと見做し、脱落に関する情報を失っていることに注意しよう。

(3.3)のデータを、R_jが所与で、集合R_jの中からランダムに1個個体を抽出するとき抽出された個体がi_jであったと見做す。R_jの中からi_jが抽出される確率は

$$\lambda_{i_j}(t_{(j)}) \Big/ \sum_{i \in R_j} \lambda_i(t_{(j)})$$

これを $j=1,2,\cdots,l$ について掛け合せたのが Cox の部分尤度関数である。即ち、部分尤度関数は次式で与えられる。

$$L_p(\beta)=\prod_{j=1}^{l}\frac{\lambda_{i_j}(t_{(j)})}{\sum_{i\in R_j}\lambda_i(t_{(j)})}$$

PH モデル (3.1) を代入し、対数尤度を求めると

$$\ln L_p(\beta)=\sum_{j=1}^{l}\beta' z_{i_j}-\sum_{j=1}^{l}\ln\sum_{i\in R_j}e^{\beta' z_i}$$

Cox の部分尤度法は、β に関する推定や検定を行なう為には $\ln L_p(\beta)$ を普通の対数尤度関数と見做し、従来の最尤法を適用せよと主張する。上式は λ_0 を含んでいない。よって、Cox の部分尤度法では、未知の非関数 λ_0 に依存することなく β に対する統計的推測を行なうことができるわけである。

4. 生存関数の推定

β の推定値が求められたとき、つぎに問われるのは共変量 z で調整したときの生存関数の推定である。それには λ_0 をデータから推定しておかねばならない。

(2.2) の条件つき尤度 L_c は、データが (3.3) の形にまとめられるとき、注2で述べたことがらに注意して計算すると、つぎのように表現できる。

$$L_c^* = \prod_{j=1}^{l} \lambda_{i_j}(t_{(j)}) \exp\left\{-\sum_{j=1}^{l} \int_{t_{(j-1)}}^{t_{(j)}} \sum_{i \in R_j} \lambda_i(t)\, dt\right\} \qquad (4.1)$$

PH モデル (3.1) を代入し，さらに変形すると

$$L_c^* = L_P(\beta) \cdot L_D(\beta, \lambda_0)$$

と表わせる。ここに、L_P は Cox の部分尤度関数、L_D は次式で与えられる。

$$L_D(\beta, \lambda_0) = \left[\prod_{j=1}^{l} \sum_{i \in R_j} e^{\beta' z_i} \lambda_0(t_{(j)})\right] \exp\left[-\sum_{j=1}^{l} \sum_{i \in R_j} e^{\beta' z_i} \int_{t_{(j-1)}}^{t_{(j)}} \lambda_0(t)\, dt\right]$$

L_P には λ_0 は含まれていない。よって、L_c^* を λ_0 について最大にするには，L_D を最大にする λ_0 を求めればよい。本来なら L_c^* を最大にする $(\beta, \hat{\lambda}_0)$ が (β, λ_0) の (条件つき) 最尤推定量であるが、ここでのミソは、β の推定値 β^* が部分尤度 $L_P(\beta)$ を最大にすることによって求められていると考え、$\beta = \beta^*$ を既知として $L_D(\beta^*, \lambda_0)$ を λ_0 について最大にすることによって λ_0 の推定量を求めようと考えることである。これは Breslow (1972) のアイディアである。Cox (1972) も λ_0 の推定法を提案しているが、理論的に根拠が薄弱であるように思われる。両者の比較はまだ行なわれていない。また、β が主役、λ_0 が脇役の母数であるとき Breslow 法は直観的に極めて妥当な推定法と思われるがその正当性は如何にして保証できるであろうか。λ_0 の推定に焦点を当てこ

の問題を考察することは面白そうである。これについて、Fraser (1968) が導入した marginal likelihood の考えが役に立ちそうである。実際、Kalbfleish and Prentice (1973) は marginal likelihood の立場から Cox の部分尤度関数を理論的に導びく試みを行なっている。

そで、

$$\lambda_0(t) = \lambda_i \qquad t_{(i-1)} < t \leq t_{(i)}, \quad i=1,2,\cdots,l$$

とおく。このとき

$$L_D(\beta^*, \lambda_0) = \left[\prod_{j=1}^{l} \sum_{i \in R_j} e^{\beta^{*\prime} z_i} \lambda_j\right] \exp\left[-\sum_{j=1}^{l} \sum_{i \in R_j} e^{\beta^{*\prime} z_i} \lambda_j (t_{(j)} - t_{(j-1)})\right]$$

$\partial L_D(\beta^*, \lambda_0)/\partial \lambda_\alpha = 0$ より

$$\hat{\lambda}_\alpha = \{[t_{(\alpha)} - t_{(\alpha-1)}] \sum_{i \in R_\alpha} e^{\beta^{*\prime} z_i}\}^{-1}$$

を得る。これを (2.1) 式に代入すると生存関数の推定量

$$\hat{F}_0(t_{(k)}) = \exp\left\{-\sum_{j=1}^{k} \frac{1}{\sum_{i \in R_j} e^{\beta^{*\prime} z_i}}\right\} \doteqdot \prod_{j=1}^{k} [1 - \hat{\pi}_j]$$

を得る。ここに、$\hat{\pi}_j = \left[\sum_{i \in R_j} e^{\beta^{*\prime} z_i}\right]^{-1}$

特に $\beta^* = 0$ のとき、$\hat{F}_0$ は良く知られた Kaplan-Meier 推定量 (1958) と一致する。即ち、Breslow 法は共変量で調整した場合の Kaplan-Meier 推定量の一般化を提供する。

$F_i(t)$ の推定量は (3.1) 式よりつぎのように与えられる。

$$\hat{F}_i(t_{(k)}) = \exp\{-e^{\beta^* Z_i}\left[\sum_{j=1}^{k} \frac{1}{\sum_{\ell \in R_j} e^{\beta^* Z_\ell}}\right]\}$$

5. 部分尤度法の正当化

実は Cox (1972) では Lp は "条件つき尤度" とよばれている。それがどのような意味で条件つき尤度であるのか，Lp をふつうの尤度関数とみなし通常の最尤法を適用すればよいという根拠は何かについて数多くの疑問が集中した。これに対して Cox (1975) は "条件つき尤度" という言葉を取り消し，部分尤度という概念を導入して理論的正当化を行なった。その根幹はつぎの通りである。

$x_i = (R_i, t_i)$ とおく。事象は経時的に $(x_1, i_1), (x_2, i_2), \cdots, (x_\ell, i_\ell)$ と観察されることより，(3.3) のデータを得る同時確率密度関数は

$$f_{\beta,\lambda_0}(i_1, \cdots, i_\ell ; R_1, \cdots, R_\ell ; t_1, \cdots, t_\ell) = f_{\beta,\lambda_0}(x_1) f_{\beta,\lambda_0}(i_1 | x_1) f_{\beta,\lambda_0}(x_2 | i_1, x_1) f_{\beta,\lambda_0}(i_2 | i_1, x_1, x_2) \cdots = \prod_{\alpha=1}^{\ell} f_{\beta,\lambda_0}(i_\alpha | i_1, \cdots, i_{\alpha-1} ; x_1, \cdots, x_\alpha) \prod_{\alpha=1}^{\ell} f_{\beta,\lambda_0}(x_\alpha | i_1, \cdots, i_{\alpha-1} ; x_1, \cdots, x_{\alpha-1})$$

ここで

$$f_{\beta,\lambda_0}(i_\alpha | i_1, \cdots, i_{\alpha-1} ; x_1, \cdots, x_\alpha) = f_{\beta}(i_\alpha | x_\alpha)$$

$$f_{\beta,\lambda_0}(x_\alpha | i_1, \cdots, i_{\alpha-1} ; x_1, \cdots, x_{\alpha-1}) = f_{\beta,\lambda_0}(x_\alpha | x_{\alpha-1})$$

であることに注意する。明らかにつぎの対応が成立する。

$$L_C^* = f_{\beta,\lambda_0}(i_1,\cdots,i_\ell;R_1,\cdots,R_\ell;t_1,\cdots,t_\ell)$$

$$L_P(\beta) = \prod_{\alpha=1}^{\ell} f_\beta \quad (i_\alpha \mid x_\alpha)$$

$$L_D(\beta,\lambda_0) = \prod_{\alpha=1}^{\ell} f_{\beta\lambda_0}(x_\alpha \mid x_{\alpha-1})$$

簡単のため $\beta \in R^1$ の場合について考える。

$$U_\alpha = \partial \ln f_\beta(i_\alpha \mid x_\alpha)/\partial\beta, \quad U = \sum_{\alpha=1}^{\ell} U_\alpha = \partial \ln L_P(\beta)/\partial\beta$$

$$J(\beta) = -\partial^2 \ln L_P(\beta)/\partial\beta^2$$

β^* : $L_P(\beta)$ より求められる最尤推定量

とおく。L_P を普通の尤度関数と考え、通常の最尤法が適用できるためには、つぎの定理が、条件つきでないもとの分布 $f_{\beta,\lambda_0}(i_1,\cdots,i_\ell;R_1,\cdots,R_\ell;t_1,\cdots,t_\ell)$ に対して成立することを示せばよい。

定理　$n \to \infty$ のとき

i) $U/\sqrt{J(\beta^*)} \longrightarrow N(0,1)$　in law

ii) $(\beta^*-\beta)\sqrt{J(\beta^*)} \longrightarrow N(0,1)$　in law

これを示すには、$f_{\beta\lambda}(i_1,\cdots,i_\ell:R_1,\cdots,R_\ell:t_1,\cdots,t_\ell)$ に関して期待値をとるとき $E(U)=0$, $V(U)=\sum_{\alpha=1}^{\ell} V(U_\alpha)$ が成立すること、すなわち

① $E(U_\alpha)=0$

② $E(U_\alpha U_{\alpha'})=0 \quad (\alpha<\alpha')$

が成立することを示せばよい。① について、U_α は x_α を所与としたときの条件つき密度によって定義されているから

$E(U_\alpha|\chi_\alpha)=0$ は明らかである。よって、$EU_\alpha=EE(U_\alpha|\chi_\alpha)=0$。

② について、$\alpha<\alpha'$ のとき

$$E(U_\alpha U_{\alpha'})=EE(U_\alpha U_{\alpha'}|U_\alpha=u_\alpha)=E[U_\alpha E(U_{\alpha'}|U_\alpha=u_\alpha)]$$

事象は経時的に生起するから $\chi_{\alpha'}$ を所与とするとき、それは $U_\alpha=u_\alpha$ を所与とすることをも含んでいる。よって

$$E(U_{\alpha'}|U_\alpha=u_\alpha)=E(U_{\alpha'}|\chi_{\alpha'})=0$$

これより ③ は明らかである。

6、部分尤度法の効率

β に関する情報は明らかに $L_D(\beta,\lambda_0)$ の方にも含まれている。にもかかわらずこれを考慮すると β の推定量が局外母数 λ_0 に依存し、やっかいな問題が生じるため Cox はこれを切り捨てた。しかしながら切り捨てられる情報が無視できないほど大きければ部分尤度法の適用は躊躇せざるを得ない。

$p=1$ のとき、L_C^*, L_P より求められる Fisher の情報量は

$$I_C^*(\beta)=E\left[-\frac{\partial^2 \ln L_C^*}{\partial\beta^2}\right]=\sum_{j=1}^{l}E\left\{\sum_{i\in R_j}z_i^2e^{\beta z_i}\int_{t_{(j-1)}}^{t_{(j)}}\lambda_0(t)\,dt\right\}$$

$$I_P(\beta)=E\left[-\frac{\partial^2 \ln L_P}{\partial\beta^2}\right]=\sum_{j=1}^{l}E\left\{\frac{\sum_{i\in R_j}z_i^2e^{\beta z_i}}{\sum_{i\in R_j}e^{\beta z_i}}-\left[\frac{\sum_{i\in R_j}z_ie^{\beta z_i}}{\sum_{i\in R_j}e^{\beta z_i}}\right]^2\right\}$$

L_C^* の替りに L_P を用いることより生じる情報の損失は

$$R(\beta) = \frac{I_p(\beta)}{I_c^*(\beta)}$$

で評価できる。$R(\beta)$ を部分尤度法の効率とみなす。

特に2標本のとき

$z_i = \begin{cases} 0 : \text{標本1} \\ 1 : \text{標本2} \end{cases}$　　$n_0(t_{(j)})$: R_j に属する標本1の個数

$n_1(t_{(j)})$: R_j に属する標本2の個数

$\lambda_0(t) = 1$，即ち指数分布のとき

とおくと

$$R(\beta) = \frac{\sum_{j=1}^{l} E\{ n_1(t_{(j)})[t_{(j)} - t_{(j-1)}] \}}{\sum_{j=1}^{l} E\left\{ \frac{n_0(t_{(j)})\, n_1(t_{(j)})}{[n_0(t_{(j)}) + n_1(t_{(j)}) e^{\beta}]^2} \right\}}$$

Kalbfleish (1974) は脱落がない場合、Efron [1977] は数学的な近似、Kay (1979) はモンテカルロ法によって $R(\beta)$ の評価を行なっているが、彼らの方法は極めて複雑、難解で筆者の力不足のため容易く納得できない。理論的にもまだ多くの問題点が残されており、効率の評価は今後の課題である。彼らによって得られた結果はほぼつぎのようにまとめることができる。

(i) 脱落がないとき，$\beta = 0$ のときの評価

標本数 (N)	10	20	40	60	∞
$R(0)$	0.89	0.94	0.95	0.97	1

(ii) $R(\beta)$ は $\beta=0$ のとき最大値をとり β が増加するにつれて減少する。

(iii) $N\to\infty$ のとき、脱落のパターンが極端でなければ $\beta=0.5$ ぐらいのとき $R(\beta)\geq 0.80$ と考えてよい。多くの場合、この値は 95% 以上になる。N が有限のときは、上の表に $N\to\infty$ のときの値をかけた値がほぼ $R(\beta)$ の値に近いと見なしてよいのではないかと考えられている。

7. 今後の問題点

前節の2標本問題は

第1標本の分布： $F(t)=\exp\{-\int_0^t \lambda_0(x)dx\}$

第2標本の分布： $G(t)=\exp\{-e^{\beta}\int_0^t \lambda_0(x)dx\}$

であるとき、仮説 $H_0:\beta=0$ を検定する問題にほかならない。λ_0 は任意だから上のように $F(t)$ に特別の型を指定しても意味はない。したがって、問題は第1，第2標本の母集団分布を F，G とし，F，G を未知とするとき、$G=F^{\psi}$ ($\psi=e^{\beta}$) と表わされることから 仮説：$\psi=1$ の検定問題に帰着する。この問題は、Lehmann (1953) の対立仮説に対する検定問題としてノンパラメトリック法では古くから良く研究されている。2標本問題に於ても同様な対応がつく。従って Cox 法は脱落のあるデータに基づいて Lehmann 対立仮説を検定す

る方法をその特別の場合として提供する。他方，ノンパラメトリック法の枠内でも局所的最強な順位検定という基準に基づき脱落のあるデータに対する順位検定が研究されている（例えば Johnson and Mehrota (1972), Mehrota and Johnson (1976) など）。Cox の部分尤度法は極めて広範な問題に適用可能であるが，上のように問題を制限し，ノンパラメトリック法の立場から脚光を当て，その最適性を何らかの意味で数学的に証明することは可能であるように思われる。Cox 法を数学的に明確に位置づけることは今後に残された課題である。

参考文献

Cox, D.R. (1972) : Regression models and life tables (with discussion) J.R.S.S. B34, 187-220

Breslow, N. (1972) : 上の論文の Discussion

Fraser, D.A.S. (1968) : The structure of inference, New York, Wiley

Kalbfleish, J.D. and Prentice, R.L. (1973) : Marginal likelihood based on Cox's regression and life model. Biometrika 60, 267-278

Kaplan, E.L. and Meier, P. (1958) : Nonparametric estimation from incomplete observation J.A.S.A. 53, 457-481

Cox, D.R. (1975) : Partial likelihood, Biometrika, 62, 269-279

Halbfleish, J.D. (1974) : Some efficiency calculations for survival distributions. Biometrika 61, 31-38

Efron, B. (1977) : The efficiency of Cox's likelihood function for censored data. J.A.S.A. 359, 557-565

Kay, R. (1979) : Some further asymptotic efficiency calculations for survival data regression models. Biometrika 66, 91-96

Lehmann, E.L. (1953) : The power of rank tests. Ann. Math. Statist. 24, 23-43

Johnson, R.A. and Mehrota, K.G. (1972) : Locally most powerful rank tests for the two-sample problem with censored data. Ann. Math. Statist. 43, 823-831

Mehrota, K.G. and Johnson, R.A. (1976) : Asymptotic sufficiency and asymptotically most powerful tests for the two sample censored situation. The Ann. of Statist. 4, 589-596

松原義弘・後藤昌司 (1979) : センサード回帰による生存時間データの解析, 応用統計シンポジウム予稿集「医学・生物学における統計的諸問題とデータ解析」

付録 [コンピュータプログラム化へのスケッチ
— 重複度のある一般の場合 —]

(1) データの整理

死亡時刻	重複度	共変量			
$t_{(j)}$	$m_{(j)}$	$Z_{1(j)}$	$Z_{2(j)}$	...	$Z_{p(j)}$

の output。$Z_{k(j)}$ よりその平均 $\bar{Z}_k$ を差し引き, それを改

ために Z_{kij} と定めておく，$k=1,2,\cdots,p$ ．

(2) t_{ij} ―― Z_{ij_1}
重複 t_{ij} } m_{ij} ― Z_{ij_2}
$\vdots$ $\vdots$
t_{ij} ―― $Z_{ij_{m_{ij}}}$

$$\Delta_{ij}=\sum_{l=1}^{m_{ij}} Z_{ij_l} \quad \text{を求める。} \tag{A1}$$

リスク集合 R_j の中から相異なる m_{ij} 個の要素の組を抽出しその各々について，上の Δ_l を計算する。

R_j → ○m_{ij} → Δ_{j_1}
→ ○m_{ij} → Δ_{j_2}
$\vdots$
→ ○m_{ij} → $\Delta_{j_\cdot}$

R_j の要素の個数を r_j とするとこの個数は $\binom{r_j}{m_{ij}}$ 個 (A2)

対数尤度関数は

$$\ln L_p(\beta)=\sum_{j=1}^{l}\beta'\Delta_j-\sum_{j=1}^{l}\ln\Big[\sum_{k\in R_j}^{*}\exp\beta'\Delta_{jk}\Big]$$

ここに $\sum_{k\in R_j}^{*}$ は (A2) の $\binom{r_j}{m_{ij}}$ 個に関する和を意味する。

(3) β の推定，検定

$\beta=0$ を初期値として，くり返し計算で $\ln L_p(\beta)$ を最大にする $\hat{\beta}$ を求める。$\partial \ln L_p^*(\beta_d)/\partial\beta_d|_{\beta=\hat{\beta}}$，Fisher の情報行列，(分散共分散行列の推定値) {スチューデント化された β などの output。β に関する仮説の検定は通常の尤度比検定による。或いは AIC 基準を適用する。

(4) 生存関数の推定

(3) により共変量の個数 p と $\beta^*=\hat{\beta}$ を定め，それを input

データに用いる。

$$\hat{\pi}_j = \frac{m_{(j)}}{\sum_{k \in R_j} \exp(\beta^{*\prime} Z_k)}$$

$$\hat{F}_0(t_{(k)}) = \prod_{j=1}^{k} (1 - \hat{\pi}_j)$$

$\hat{F}_0$ の分散の1つの近似的な推定量は次式が与えられる。

$$\hat{V}[\hat{F}_0(t_{(k)})] \doteqdot \left[1 - \frac{m_k}{\sum_{\ell \in R_k} e^{\beta^{*\prime} Z_\ell}}\right]^2 \hat{V}[\hat{F}_0(t_{(k-1)})] + [\hat{F}(t_{(k)})]^2 \frac{m_k}{\left(\sum_{j \in R_k} e^{\beta^{*\prime} Z_j}\right)^2} \quad (A.1)$$

$k = 1, 2, \cdots$

これより、初期値 $\hat{V}[\hat{F}_0(t_{(0)})] = 0$ より逐次に計算することにより $\hat{F}_0$ の標準偏差を求め、生命表形式に $t_{(k)}$, $\hat{\pi}_j$, $\hat{F}_0(t_{(k)})$ と $\hat{E}.(\hat{F}_0(t_{(k)}))$, $k = 1, 2, \cdots$ を output する。

いくつかの層に層別されているときは各層ごとの操作を行なって $\hat{F}_0(t_{(k)})$ を求めればよい。

[補遺] ((A.1) 式の導び き方)

筆者の見る限りにおいて $\hat{F}_0$ の分散の推定量を与えた文献は存在しない。(A.1) の導び き方を簡単に説明しておく。本文の枠組で、重複度があるとき $\ln L_D$ は次式で与えられる。

$$\ln L_D(\lambda_1, \cdots, \lambda_\ell) = \sum_{j=1}^{\ell} \ln \sum_{i \in K_j}^{*} e^{\beta' Z_i} + \sum_{j=1}^{\ell} m_j \ln \lambda_j - \sum_{j=1}^{\ell} \sum_{i \in K_j} \lambda_j (t_j - t_{j-1}) e^{\beta' Z_i}$$

$\partial \ln L_D / \partial \lambda_\alpha = 0$ より

$$\hat{\lambda}_\alpha = m_\alpha \Big/ \Big[\sum_{i \in R_\alpha} (t_{(i)} - t_{(i-1)}) e^{\beta' z_i} \Big]$$

$I(\lambda_\alpha) = - E\left[\partial^2 \ln L_D / \partial^2 \lambda_\alpha \right] = E\left[m_\alpha / \lambda_\alpha^2 \right]$, $E\left[\partial^2 \ln L_D / \partial \lambda_\alpha \cdot \partial \lambda_\gamma \right] = 0$

よって, $\hat{\lambda}_1, \cdots, \hat{\lambda}_k$ は漸近的に互いに独立、$\hat{\lambda}_\alpha$ は平均 λ_α 分散 $1/I(\lambda_\alpha)$ の正規分布に従がう。

$$\ln \hat{F}(t_{(K)}) = \ln \bar{F}(t_{(K)}) + \sum_{i=1}^{K} \ln \{ 1 + [t_{(i)} - t_{(i-1)}](\hat{\lambda}_i - \lambda_i) \}$$

$$\sim \ln \bar{F}(t_{(K)}) + \sum_{i=1}^{K} (t_{(i)} - t_{(i-1)})(\hat{\lambda}_i - \lambda_i)$$

$$\therefore \hat{F}(t_{(K)}) - \bar{F}(t_{(K)}) \sim \bar{F}(t_{(K)}) \Big\{ \sum_{i=1}^{K} [t_{(i)} - t_{(i-1)}][\hat{\lambda}_i - \lambda_i] \Big\}$$

よって、$\hat{F}(t_{(K)})$ の分散は次式で近似できる。

$$\hat{V}[\hat{F}(t_{(K)})] = \bar{F}^2(t_{(K)}) \sum_{i=1}^{K} \frac{1}{E\Big\{ \Big(\sum_{j \in R_i} e^{\beta' z_j} \Big)^2 \frac{1}{m_i} \Big\}}$$

$$\therefore \hat{V}[\hat{F}(t_{(K)})] \fallingdotseq \hat{\bar{F}}^2(t_{(K)}) \sum_{i=1}^{K} \frac{m_i}{\Big(\sum_{j \in R_i} e^{\beta' z_j} \Big)^2}$$

これを重み入の形に書き表わしたのが (A.I.) である。 (A.I.) で $\beta^* = 0$ とおくと

$$\hat{V}[\hat{F}(t_{(K)})] = \Big(\frac{r_K - m_K}{r_K} \Big)^2 \hat{V}[\hat{F}(t_{(K-1)})] + [\hat{F}(t_{(K)})]^2 \frac{m_K}{r_K^2}$$

右辺第一項は Kaplan-Meier 推定量の分散の近似推定式の第1項と一致する。第2項は Kaplan-Meier では $r_K(r_K - m_K)$ とされるところが r_K^2 に変っているが、ふつう m_K/r_K は十分小さいので これでも十分役に立つであろう。