

生存時間解析入門

汪 金芳

千葉大学 大学院自然科学研究科

平成 17 年 5 月 13 日

目次

1	生存時間解析とは	2
2	パラメトリック・モデル	3
2.1	生存関数とハザード関数	3
2.2	生存時間のモデル	3
2.3	生存関数の最尤推定量	4
3	ノンパラメトリック推定と検定	5
3.1	カプラン・マイヤー推定量	5
3.1.1	打ち切りがない場合	5
3.1.2	打ち切りがある場合	6
3.2	信頼バンド	8
3.3	ログ・ランク検定	8
4	比例ハザード・モデル	10

1 生存時間解析とは

表 15-1 では、ある治療法（群 1）とプラシーボ（群 2）を、それぞれ 21 人の白血病患者に対して行い、治療開始から死亡するまでの時間（生存週数）を記録したものである。たとえば、新しい治療を受けてから、6 番目の患者は 23 週間後死亡し、またプラシーボを受けた 1 番目の患者も同じ 23 週間後死亡した。生存時間に “+” がついているものは、打切り censoring を意味する。すなわち、研究が終了した時点で、死亡が観察されないか、もしくは試験の途中で脱落したなどを意味する。たとえば、群 1 の 1 番目の患者の生存時間は 35+ なので、この患者は少なくとも 35 週間生存したことを意味する。

表 15-1 白血病患者の週生存時間（出典：Kleinbaum, 1996, p.75）

患者 番号	生存 時間	白血球数 の対数	性 別	患者 番号	生存 時間	白血球数 の対数	性 別		
群 1	1	35+	1.45	1	群 2	1	23	1.97	1
	2	34+	1.47	1		2	22	2.73	0
	3	32+	2.20	1		3	17	2.95	0
	4	32+	2.53	1		4	15	2.30	0
	5	25+	1.78	1		5	12	1.50	0
	6	23	2.57	1		6	12	3.06	0
	7	22	2.32	1		7	11	3.49	0
	8	20+	2.01	1		8	11	2.12	0
	9	19+	2.05	0		9	8	3.52	0
	10	17+	2.16	0		10	8	3.05	0
	11	16	3.60	1		11	8	2.32	0
	12	13	2.88	0		12	8	3.26	1
	13	11+	2.60	0		13	5	3.49	1
	14	10+	2.70	0		14	5	3.97	0
	15	10	2.96	0		15	4	4.36	1
	16	9+	2.80	0		16	4	2.42	1
	17	7	4.43	0		17	3	4.01	1
	18	6+	3.20	0		18	2	4.91	1
	19	6	2.31	0		19	2	4.48	1
	20	6	4.06	1		20	1	2.80	1
	21	6	3.28	0		21	1	5.00	1

このように、ある時点から、興味のあるイベント event が発生するまで、個体を観察することが、医学や工学などの分野でしばしば行われる。白血病の場合のイベントは患者の死亡であり、また、たとえば、製品の信頼性実験の場合、製品の故障・破壊をイベントとすることが多い。イベントは、故障 failure や死亡などとも呼ばれる。イベントが観察されるまでの時間 T を生存時間 survival time といい、 T は確率変数である。

生存時間解析の目的は、表 15-1 で示されたような生存時間と関連情報についてのデータを用いて、生存率の推定や群 2 の生存率の比較、さらに生存率と共変量の関係（たとえば白血球数と生存時間の関係）の解明などである。

2 パラメトリック・モデル

2.1 生存関数とハザード関数

生存時間解析の主要な目的の 1 つは、次の生存関数 survival function

$$S(t) = \Pr(T > t) = \int_x^\infty f(t) dt \quad (1)$$

の推定や比較である。ここで、 $f(t)$ は T の密度関数である。生存関数 $S(t)$ は故障・死亡するまでの時間が t を超える確率を表している。

ところで、 T の分布に対するモデルを構築するとき、次のハザード関数 hazard function

$$h(t) = \lim_{\Delta t \rightarrow 0^+} \frac{\Pr(t \leq T < t + \Delta t | T \geq t)}{\Delta t} \quad (2)$$

を使用するのが便利である。ハザード関数 $h(t)$ は、 t 時まで生存した条件の下で、次の時刻に死亡する、瞬間死亡率を表している。人口データなどの解析において、ハザード関数 $h(t)$ の定性的な性質についてある程度経験的に知られていることが多い。 $h(t)$ と $S(t)$ は

$$S(t) = \exp \left\{ - \int_0^t h(u) du \right\}, \quad h(t) = - \frac{S'(t)}{S(t)} \quad (3)$$

という関係があり、1 つが決まればもう 1 つの方を求めることができる。

2.2 生存時間のモデル

もっとも良く使われるモデルは次のものである。

(i) 指数分布 これはハザードが一定のモデルで（図 15-1(a)）、観察期間中に健康状態が安定な人を観察するときなどに相当する。定数 $\lambda > 0$ に対して、 $h(t) = \lambda$ とすれば、(3) より生存関数と密度関数が次のように得られる。

$$S(t) = e^{-\lambda t}, \quad f(t) = \lambda e^{-\lambda t} \quad (4)$$

(ii) ワイブル分布 これは指数分布の一般化で、ハザード関数が $h(t) = \lambda p (\lambda t)^{p-1}$ と表され、 $p = 1$ のとき $h(t)$ は定数で、 $p > 1 (< 1)$ のとき $h(t)$ は t の単調増加（減少）関数となる。図 15-1(b)(c) 参照。たとえば、ある病気に対して、まったく治療を受けてない患者の死亡するハザードは時間と共に増加し、また手術などを受けた患者に対してはハザードが単調減少すると考えられる。(3) より、生存関数と密度関数は次のようになる。

$$S(t) = e^{-(\lambda t)^p}, \quad f(t) = \lambda p (\lambda t)^{p-1} e^{-(\lambda t)^p} \quad (5)$$

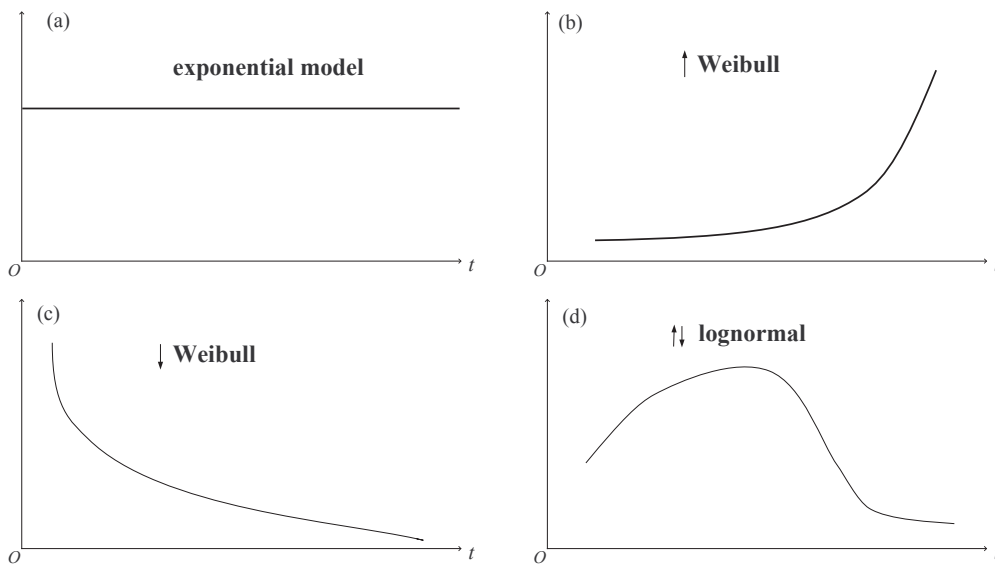


図 15-1: いろいろなハザード関数

(iii) 対数正規分布 $\log T$ が正規分布 $N(\mu, \sigma^2)$ に従い, T の密度関数が次のようになる.

$$f(t) = \frac{1}{\sqrt{2\pi\sigma t}} e^{-\frac{(\log t - \mu)^2}{2\sigma^2}}, \quad t > 0$$

$\Phi(\cdot)$ を標準正規分布の密度関数とすれば, 生存関数は

$$S(t) = 1 - \Phi\left(\frac{\log t - \mu}{\sigma}\right) \quad (6)$$

となる. ハザード関数は $h(t) = -S'(t)/S(t)$ により求められるが, 式が煩雑なため省略する. 図 15-1(d) を参照. この場合のハザードは, 増加から減少に転じる関数で, 肺結核などの慢性疾患をもつ患者に対して適切なモデルであろう.

他のモデルとして, ガンマ分布や対数ロジスティック分布などもしばしば利用される.

2.3 生存関数の最尤推定量

無作為標本 t_1, \dots, t_n に打ち切りがなく, またパラメトリック・モデルの仮定が妥当な場合を考える. 生存関数を $S(t) = S(t|\theta)$ と書くと, 母数 θ の最尤推定量を $\hat{\theta}$ を用いて, 生存関数を $\hat{S}(t) = S(t|\hat{\theta})$ で推定できる.

たとえば, 指数分布モデルに対して, $\bar{T}_n = n^{-1} \sum_{i=1}^n t_i$ を標本平均とすると, λ の最尤推定量は $\hat{\lambda} = 1/\bar{T}_n$ となる. したがって生存関数の最尤推定量は $\hat{S}(t) = e^{-\hat{\lambda}t}$ となる. 表 15-1 の群 2 のデータに対して指数分布モデルを適用してみよう. このとき $\bar{T}_n = 8.667$ で, $\hat{\lambda} = 0.115$ となる. 生存関数のグラフは図 15-2 で示されている.

3 ノンパラメトリック推定と検定

前節で紹介したパラメトリック・モデルの適用が難しい場合、生存関数の推定量を次のように構成することができる。

3.1 カプラン・マイヤー推定量

3.1.1 打ち切りがない場合

まず n 個の無作為標本に対して、打ち切りがない場合を考える。生存関数と分布関数の関係 $S(t) = 1 - F(t)$ により、経験分布 $F_n(t)$ を用いて、 $S(t)$ を

$$\hat{S}(t) = 1 - F_n(t) = \frac{1}{n} \sum_{i=1}^n \delta(t_i > t) \quad (7)$$

で推定することが考えられる。ここで $\delta(t_i \geq t)$ は指標関数で、 $t_i \geq t$ のときに 1、 $t_i < t$ のときに 0 である。(7) 式より、 $t \leq t_1$ に対して $\hat{S}(t) \equiv 1$ で、また $t > t_k$ であれば $\hat{S}(t) \equiv 0$ である。推定量 $\hat{S}(t)$ は後述で述べるカプラン・マイヤー推定量の特殊な場合である。

(7) 式を表 15-1 の群 2 のデータに対して適用し得られた生存関数の推定量が図 15-2 である。このように、タイ tie がなければ、 $\hat{S}(t)$ は死亡時刻毎に $1/n$ ずつ減少する階段関数である。

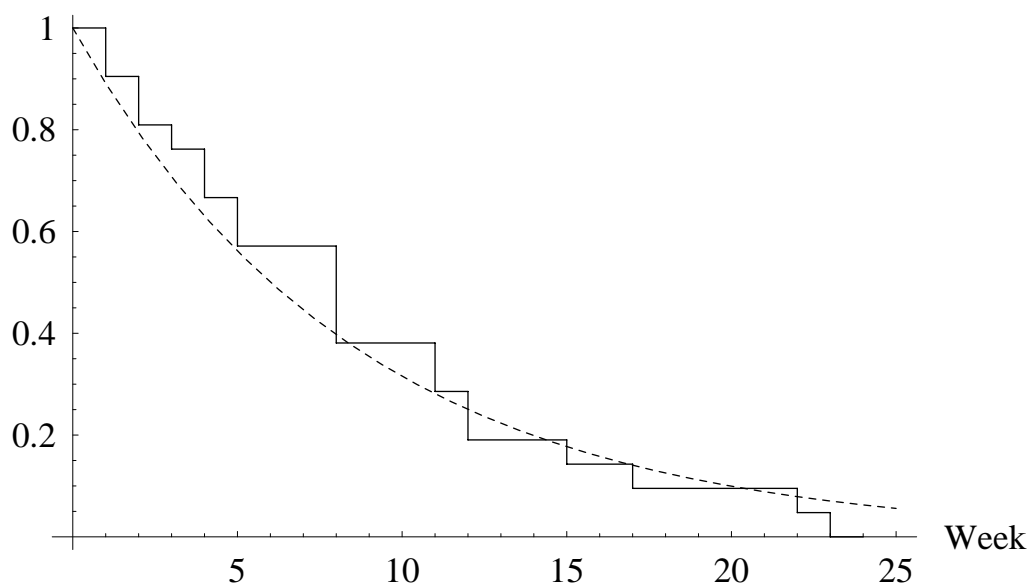


図 15-2: 表 15-1 の群 2 のデータに対する生存関数の推定量. 実線 : カプラン・マイヤー推定量, 点線 : 指数分布を仮定したときの最尤推定量.

3.1.2 打ち切りがある場合

いま n 個のデータに打ち切りの可能性がある場合を考える。死亡があった時刻を

$$t_1 < t_1 < \cdots < t_k, \quad k \leq n$$

とする。タイ或いは打ち切りがある場合, $k < n$. ここで時刻 t_j における死亡数を d_j とすると, $D = \sum_{j=1}^n d_j$ は総死亡数を表し, 打ち切りがなければ, $n = D$ で, そうでなければ $n < D$ となる。

次に打ち切り標本数を考える。区間 $[t_j, t_{j+1})$ における打ち切り標本数を $m_j, j = 1, \dots, k$ とし, 時間 t_j まで (t_j を含まない) の生存者数を n_j とすると, n_j は時刻 t_j におけるリスク集合の大きさと呼ばれる。次が成り立つことに注意する。

$$n_j = \sum_{i=j}^k (d_i + m_i), \quad i = 1, \dots, k$$

生存関数 $S(t)$ のカプラン・マイヤー推定量 (積極限推定量) は次のように定義される。

$$\hat{S}(t) = \begin{cases} 1 & t < t_1 \text{ のとき} \\ \prod_{t_i \leq t} \frac{n_i - d_i}{n_i} & t \geq t_1 \text{ のとき} \end{cases} \quad (8)$$

すなわち, $t_\ell \leq t < t_{\ell+1}$ のとき, 生存関数は次のように計算される。

$$\hat{S}(t) = \frac{n_1 - d_1}{n_1} \times \cdots \times \frac{n_\ell - d_\ell}{n_\ell}$$

特に $n_k = d_k + m_k$ なので, $m_k = 0$ であれば, $t > t_k$ に対して $\hat{S}(t) = 0$ となる。逆に, $m_k > 0$ であれば, $t > t_k$ に対して, $\hat{S}(t) > 0$ となる。

打ち切りがまったくない標本に対しては, $n_j = \sum_{i=j}^k d_i$ より, $n_j - d_j = n_{j+1}$ となる。したがって, $t_\ell \leq t < t_{\ell+1}$ に対して, (8) より次が成り立つ。

$$\hat{S}(t) = \frac{n_2}{n_1} \times \frac{n_3}{n_2} \times \cdots \times \frac{n_{\ell+1}}{n_\ell} = \frac{n_{\ell+1}}{n_1}$$

$n_1 = n$ に注意すると, 上の式は (7) とまったく同じものとなることが分かる。故に, (7) 式はカプラン・マイヤー推定量の特殊な場合に過ぎない。

時刻 t_j におけるカプラン・マイヤー推定量は, 直前の時刻 t_{j-1} における推定量と, t_j まで生きていた条件のもとでの t_j を乗り越える確率との積で表すことができる。すなわち

$$\hat{S}(t_j) = \hat{S}(t_{j-1}) \times \widehat{\Pr} [T > t_j | T \geq t_j]$$

という関係が成立する。 $S(t_{j-1}), S(t_{j-2}), \dots, S(t_1)$ についても同様な式を当てはめると, 次が得られる

$$\hat{S}(t_j) = \prod_{i=1}^j \widehat{\Pr} [T > t_i | T \geq t_i] \quad (9)$$

このようにカプラン・マイヤー推定量は、条件付生存確率の積として表現できることがわかる。

プラン・マイヤー推定量 (8) を表 15-1 の白血病患者データに対して適用し、各死亡時刻における生存関数の推定値を示したのが表 15-2 である。

表 15-2 白血病患者データに対するカプラン・マイヤー推定量

群 1					群 2				
t_j	d_j	m_j	n_j	$\hat{S}(t_j)$	t_j	d_j	m_j	n_j	$\hat{S}(t_j)$
6	3	1	21	0.857	1	2	0	21	0.905
7	1	1	17	0.807	2	2	0	19	0.810
10	1	2	15	0.753	3	1	0	17	0.762
13	1	0	12	0.690	4	2	0	16	0.667
16	1	3	11	0.627	5	2	0	14	0.571
22	1	0	7	0.538	8	4	0	12	0.381
23	1	5	6	0.448	11	2	0	8	0.286
					12	2	0	6	0.190
					15	1	0	4	0.143
					17	1	0	3	0.095
					22	1	0	2	0.048
					23	1	0	1	0.000

表 15-2 の計算結果に基づいて、2 群の生存関数のカプラン・マイヤー推定量を示したのが図 15-3 である。すべての時間において、群 1 の生存関数の推定量が群 2 のそれに比べると明らかに高くなっていることから、治療効果があることが伺える。

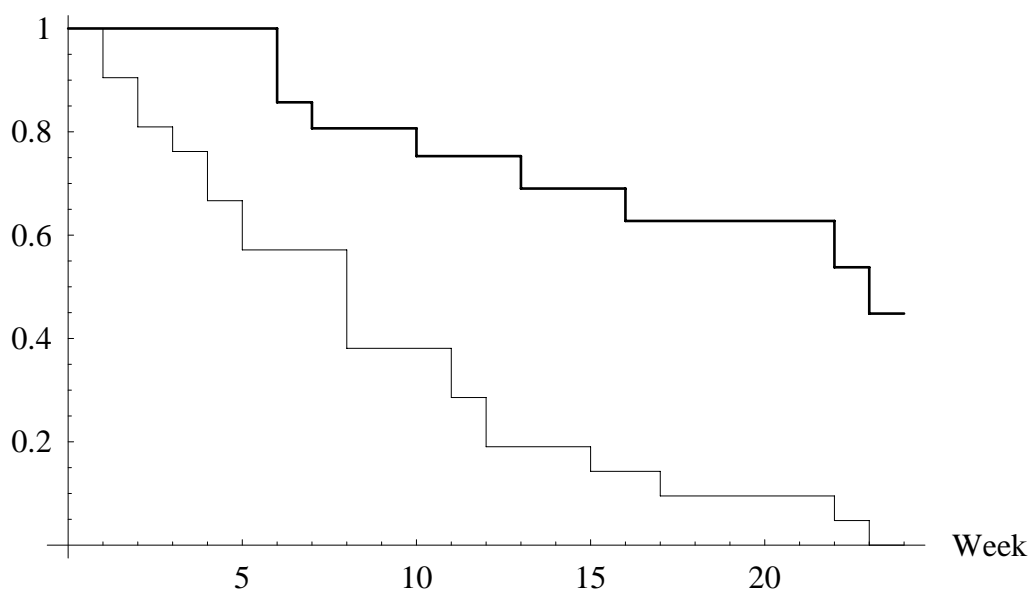


図 15-3: 白血病患者データに対するカプラン・マイヤー推定量 (太線: 群 1; 細線: 群 2)

3.2 信頼バンド

カプラン・マイヤー推定量 $\hat{S}(t)$ は漸近的に正規分布 $N(S(t), V(t))$ に従うことが知られている。ここで $V(t)$ を次のように推定することができる。

$$\hat{V}(t) = \hat{S}(t) \sum_{t_i \leq t} \frac{d_i}{n_i(n_i - d_i)} \quad (10)$$

この式を通常グリーンウッド Greenwood の公式と呼ばれている。(10)により、 $S(t)$ の信頼係数 $1 - 2\alpha$ の近似信頼バンドは次に与えられる

$$\left[\hat{S}(t) + z_\alpha \sqrt{\hat{V}(t)}, \hat{S}(t) - z_\alpha \sqrt{\hat{V}(t)} \right] \quad (11)$$

信頼バンド(11)は推定量 $\hat{S}(t)$ に関して対称である。公式(11)を表15-1の群1のデータに当てはめ、得られた生存関数の信頼バンドを示したのが図15-4である。

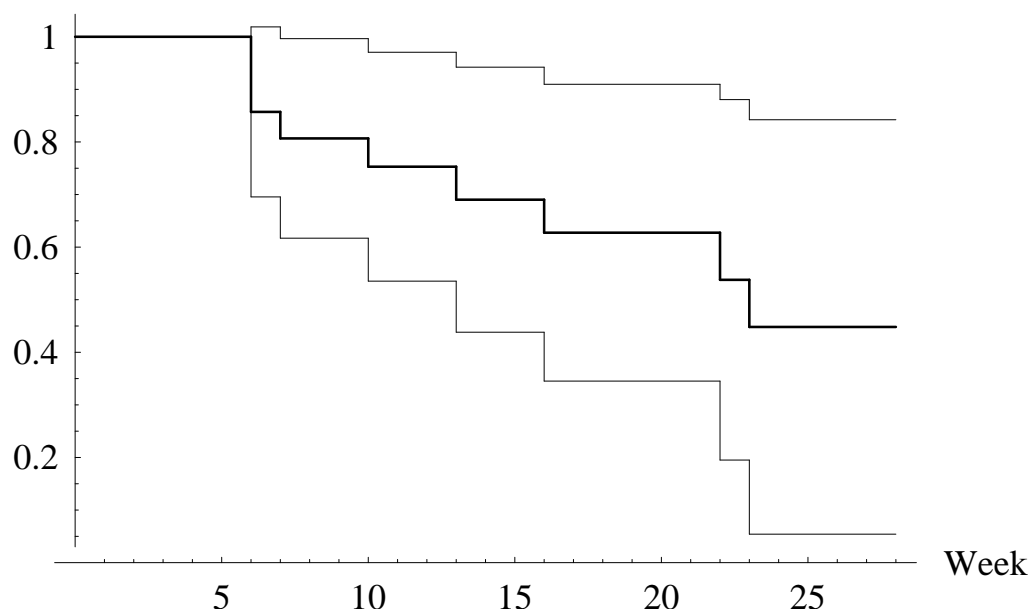


図 15-4: 白血病データにおけるカプラン・マイヤー推定量(太線)とグリーンウッドの公式による信頼バンド(細線)。

3.3 ログ・ランク検定

図15-3から治療群における生存確率がより高い可能性を示唆しているが、ここで2つの生存関数の差の有無の検定について考えてみよう。そのために、2群における死亡のあったすべての時刻を、 $t_1 < t_2 < \dots < t_K$ とする。第1群の時刻 t_j における死亡数、打ち切り数、およびリスク集合の大きさを、それぞれ d_{1j}, m_{1j}, n_{1j} とし、第2群対応する量を d_{2j}, m_{2j}, n_{2j} とする。白血病患者データに対してこのように整理したのが表15-3である。

このとき, $d_{1j} + d_{2j}$ は時刻 t_j における 2 群の総死亡数を表し, $n_{1j} + n_{2j}$ は t_j における総リスク集合の大きさを意味する. 2 つの生存曲線に差がなければ, t_j 時におけるリスク集合の相対的大きさ $n_{1j}/(n_{1j} + n_{2j})$ を用いて, 時刻 t_j における群 1 と第 2 群の期待される死亡数を, それぞれ次のように表すことができよう.

$$D_{1j} = \frac{n_{1j}}{n_{1j} + n_{2j}} (d_{1j} + d_{2j}), \quad D_{2j} = \frac{n_{2j}}{n_{1j} + n_{2j}} (d_{1j} + d_{2j}), \quad j = 1, \dots, K$$

表 15-3 すべての白血病患者における死亡時刻, 打ち切り数とリスク集合

t_j	群 1			群 2		
	d_{1j}	m_{1j}	n_{1j}	d_{2j}	m_{2j}	n_{2j}
1	0	0	21	2	0	21
2	0	0	21	2	0	19
3	0	0	21	1	0	17
4	0	0	21	2	0	16
5	0	0	21	2	0	14
6	3	1	21	0	0	12
7	1	0	17	0	0	12
8	0	1	16	4	0	12
10	1	1	15	0	0	8
11	0	1	13	2	0	8
12	0	0	12	2	0	6
13	1	0	12	0	0	4
15	0	0	11	1	0	4
16	1	0	11	0	0	3
17	0	3	10	1	0	3
22	1	0	7	1	0	2
23	1	5	6	1	0	1

ここでそれぞれの群における観測死亡数と期待死亡数を, すべての死亡時刻に対して和をとると次のようになる.

$$O_1 = \sum_{j=1}^K (d_{1j} - D_{1j}), \quad O_2 = \sum_{j=1}^K (d_{2j} - D_{2j})$$

もし両群の生存関数にあまり差がなければ, O_1 も O_2 も大きくなると想像できよう. 一方, 簡単な計算より, $O_1 = -O_2$ となることが分かる.

さて, $O_1 = -O_2$ は漸近的に平均ゼロの正規分布に従うことが知られ, また O_1 と O_2 の分散を次の式で推定することができる.

$$\hat{V} = \sum_{j=1}^K (d_{1j} + d_{2j}) \frac{n_{1j}}{n_{1j} + n_{2j}} \left(1 - \frac{n_{1j}}{n_{1j} + n_{2j}} \right) \frac{n_{1j} + n_{2j} - d_{1j} - d_{2j}}{n_{1j} + n_{2j} - 1}$$

このように、2群の生存関数に差がないという帰無仮説に対して、次の統計量

$$\chi^2 = \frac{O_1^2}{\hat{V}} = \frac{O_2^2}{\hat{V}} \quad (12)$$

を利用することができる。 χ^2 をログ・ランク検定統計量といい、帰無仮説のもとで、 χ^2 は漸近的に自由度1のカイ二乗分布に従う。

表15-3の白血病データに基づいて、計算してみると、 $O_1 = -10.2505$, $\hat{V} = 6.2570$, $\chi^2 = 16.7929$ となる。このときの p -値は 4.1688×10^{-5} となり、したがって治療効果が極めて有意であることが分かる。

4 比例ハザード・モデル

ところで、白血病患者の生存時間を示した表15-1には、患者の白血球数（の対数）などの情報も示されている。特に白血球数は、白血病患者の死亡をイベントとした場合、よく知られる予後因子 prognostic indicator である。すなわち白血病患者の生存時間は、治療効果のほか、白血球数という予後因子によって影響される可能性が考えられる。このようにいくつかの重要な共変量が存在するとき、それらを解析に入れ、交絡要因による影響を排除して、生存関数の比較などを行うことが重要である。

一般に患者に付随する共変量を $x = (x_1, \dots, x_p)$ とする。ここで x は時間に依存しないことを仮定する。白血病の例の場合、共変量を $x = (x_1, x_2)$ とする。ここで x_1 はダミー変数で、治療を受けた場合、 $x_1 = 0$ 、対象群の場合、 $x_1 = 1$ とし、また x_2 を白血球数の対数とする。この場合、 x_2 の影響を除いて、 x_1 の効果を調べるのが目的である。

さて従来の回帰分析の考え方を借りて、生存時間解析における回帰分析の考え方を述べよう。従来の回帰分析においては、 $E(Y) = g(\beta'x)$ などの仮定をおき、最小二乗法や最尤法などを用いて回帰母数 β の推定を行う。いまの場合、 $E(Y)$ の代わりに、ハザード関数を用いて考えるのが自然であろう。

たとえば、ハザードが時間に依存しなければ、次の単純なモデルが考えられる。

$$h(t; x) = \exp(\alpha + \beta'x) = h_0 e^{\beta'x}, \quad h_0 > 0 \quad (13)$$

モデル(13)におけるハザードは時間と無関係なので、生存時間 T の分布は指数分布に限られる。モデル(13)は指数回帰モデルと呼ばれる。密度関数 $f(t) = \lambda \exp(-\lambda t)$ にもつ指数分布のハザードは λ なので、 $\lambda = h_0 e^{\beta'x}$ として、最尤法で β を推定することができる。

指数回帰モデル(13)における h_0 を $h_0(t)$ で置き換えて得られたのが、次のコックス D.R. Cox による比例ハザードモデル proportional hazard model である。

$$h(t; x) = h_0(t) e^{\beta'x}, \quad h_0(t) > 0 \quad (14)$$

ここで $h(t; x = 0) = h_0(t)$ となることから、 $h_0(t)$ を基準ハザード baseline hazard と呼ばれる。 $h_0(t)$ は t の関数であるが、その形を全く指定しない。一方、(14)における共変量の効果については、明示的な関数を用いて規定している。このことから、比例ハザードモデルはセミ・パラメトリックなモデルと呼ばれる。

比例ハザードモデルは理論と応用の両面において極めて重要なモデルである。このモデルの一番の魅力は、基準ハザード関数 $h_0(t)$ の形については、如何なるモデルも仮定しない、また仮定する必要がないことである。このことは、生存時間 T の分布がどんな分布であっても、比例ハザードモデル(14)に基づいて解析を行えば、得られる結果は頑健的 robust であることを意味する。基準ハザード関数 $h_0(t)$ は、共変量を考慮しないときのハザード関数であり、実際のデータ解析においてこれを指定する必要がないことは非常に有難いことである。

比例ハザードモデルを適用するときの最大の注意点は比例ハザード性の仮定といえよう。いま共変量 x, x^* を持つ二人のハザードの比を考えると、(14) より

$$\frac{h(t; x)}{h(t; x^*)} = \exp [\beta'(x - x^*)] \quad (15)$$

となる。すなわち個人間のハザード比 hazard ratio は、時間によらず、共変量のみ関数となる。ハザード関数における(15)式の制約は通常比例ハザード性 proportional hazard assumption と呼ばれる。共変量の効果が時間と共に変化するなどの場合に、比例ハザード性は成り立たず、吟味せずに比例ハザードモデルを当てはめることは不適切である。

ところで、比例ハザードモデル(14)のもとで、生存関数は(3)より次のように表現できる。

$$\begin{aligned} S(t; x) &= \exp \left\{ - \int_0^t h(u; x) du \right\} \\ &= \exp \left\{ - \int_0^t h_0(u) \exp(\beta'x) du \right\} \\ &= \exp \left\{ - \exp\{\beta'x\} \int_0^t h_0(u) du \right\}. \end{aligned}$$

したがって、 $H_0(t) = \int_0^t h_0(u) du$ を基準累積ハザードとすると、

$$\log(-\log S(t; x)) = \beta'x + \log H_0(t) \quad (16)$$

が成り立つのである。(16)を利用して比例ハザード性を検証することができる。

いま表 15-1 の白血病の例において、 $x = x_1$ のみを考える。 x_1 はダミー変数で 0 か 1 の値をとる。(16)より次の関係式を得る。

$$\log(-\log S(t; 1)) = \beta_1 + \log(-\log S(t; 0)) \quad (17)$$

すなわち、比例ハザード性のもとでは、共変量の値で層別した時の生存関数の 2 重対数 $\log(-\log S)$ は層間で平行になる必要がある。さらに、生存関数の 2 重対数間の距離はその共変量の効果を表すパラメータの大きさそのものであることも分かる。

時間または時間の対数を横軸にとり、 $\log(-\log S)$ をプロットした図形を 2 重対数プロットという。層間で 2 重対数プロットが平行でないときに、比例ハザードモデルを適用しては誤った結論を招く恐れがある。

図 15-5 では、時間の対数を横軸にとり、表 15-1 の治療群と対象群における生存関数の Kaplan-Meier 推定量に対する 2 重対数プロットを示している。この図から治療効果は

時間によってあまり変化しないことが読み取れる。したがって、このデータに対して比例ハザードモデルが適用できそうである。さて、 $x = (x_1, x_2)$ とし、部分尤度 partial likelihood の最大化によってパラメータの値を推定してみると、 $\hat{\beta}_1 = 1.294$, $\hat{\beta}_2 = 1.604$ という結果をえる。この $\hat{\beta}_1 = 1.294$ の値が大体図 15-5 における 2 本の曲線間の距離になっていることが確認できる。

部分尤度法の解説や、市販のソフトによるパラメータの推定量の求め方などについては専門書（たとえば、大橋・浜田（1995）を参照）に譲ることにする。

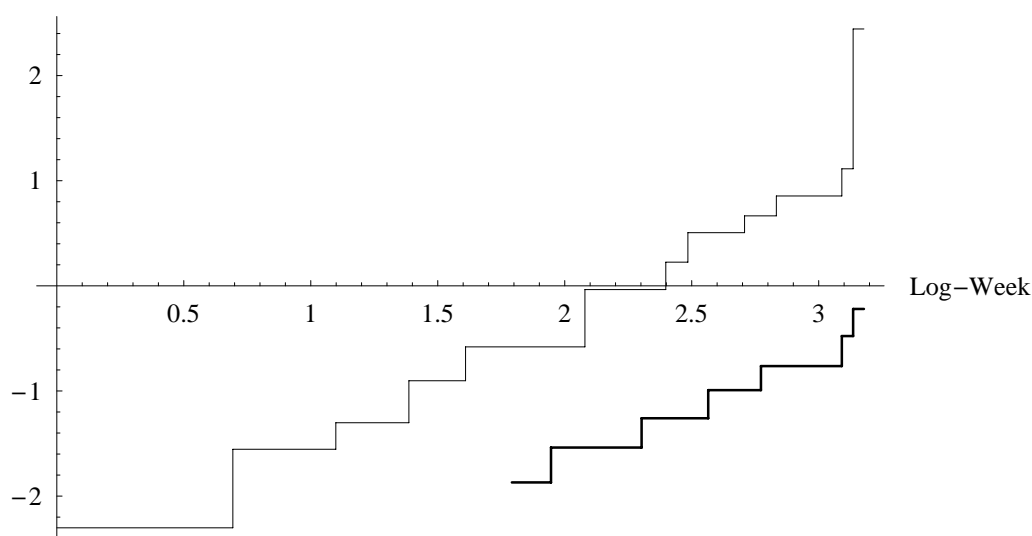


図 15-5: 表 15-1 の治療群（太線）と対象群（細線）における Kaplan・マイヤー推定量の 2 重対数プロット。

参考文献

D. G. Kleinbaum (1996). *Survival Analysis: A Self-Learning Text*, Springer: New York.

栗原考次 (2001). データの科学, 放送大学教育振興会: 東京.

松原望 (2000). 統計の考え方, 放送大学教育振興会: 東京.

大橋靖雄・浜田知久馬 (1995). 生存時間解析 SAS による生物統計, 東京大学出版会: 東京.