

# 医学統計勉強会

東北大学病院循環器内科・東北大学病院臨床研究推進センター 共催

東北大学大学院医学系研究科 EBM 開発学寄附講座

宮田 敏

*Absence of evidence is not evidence of absence!*

- Carl Sagan -

## 第4回 生存時間解析

### —生存曲線, Cox 比例ハザードモデル—

#### 1. 生存時間解析

医学データの解析においては、何らかのイベントが発生するまでの時間 (=生存時間) を解析の対象とする場合があります。例えば、観察開始時から、ある疾患の発症や死亡、入院といったイベントの発生までの時間を記録し、生存時間に影響を与える要因を検討する場合などがこれに当たります。このような時間データを被説明変数とする解析手法を、**生存時間解析 (survival time analysis)** と呼びます。

生存時間解析においては、対象となる事象を**イベント (event)**、**エンドポイント (end point)**、**結果 (outcome)**、解析対象の時間を **生存時間 (survival time, failure time)** などと呼びます。生存時間に影響を与える因子を探索することは生存時間解析の重要な目的ですが、生存時間を説明する変数は、**説明変数 (explanatory variable)**、**独立変数 (independent variable)**、**共変量 (covariate)**、**危険因子 (risk factor)**、**予後因子 (prognostic factor)** などと呼ばれます。

生存時間解析では、観察期間中にイベントの発生が観察出来ず、正確な生存時間が不明となる場合があります。例えば、

- 観察期間終了時までイベントが起らなかった場合 (**censoring**)
- 研究に参加する同意が撤回された、患者さんが自主的に通院を中止した、などの理由により意図的に観察を打ち切る場合 (**withdraw**)
- 患者さんが行方不明になった、海外に移住して追跡不能となった、など意図せざる理由によって観察を打ち切る場合 (**lost to follow-up**)

などがあります。このような状況を「**打ち切り (censoring)**」が生じた、といいます。生存時間データは、**生存時間 (観察期間)** と**打ち切りの有無**の二つの情報のペアとして記録されることとなります。打ち切りが生じた場合、生存時間そのものは不明となりますが、「観察期間中にイベントが起らなかった」という情報は手に入るようになります。

なお、上に挙げた **censoring, withdraw, lost to follow-up** の三つは、現実には区別が難しい場合も有り本稿では同一視して扱います。また、ここで取り上げ

た打ち切りは、厳密には**右側打ち切り (right censoring)** と呼ばれます。このほかに、観察期間が始まる前にイベントが起きてしまった場合 (**左側打ち切り, left censoring**) や、観察期間中のある期間においてイベントが起こったことだけがわかっている場合 (**区間打ち切り, interval censoring**) もありますが、本稿では取り上げません。

## 2. 生存時間解析の基本概念

あるイベントが起こるまでの時間を記述するために、基本的な概念を定義します。まず、あるイベントが起こるまでの時間を表す確率変数を  $T (T > 0)$  とします。  $T$  の確率密度関数を  $f(t)$ 、累積分布関数を  $F(t)$  とします。

$$F(t) = P(T \leq t) = \int_{-\infty}^t f(x)dx = \int_0^t f(x)dx$$

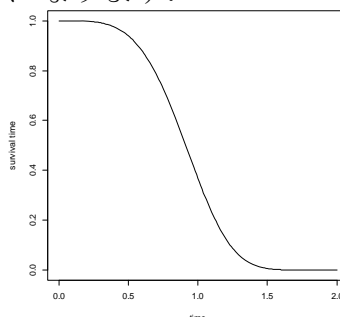
分布関数  $F(t)$  は「 $t$  時点までにイベントが起きる確率」 (= 「 $t$  時点以前に死亡する確率」) を表すのに対し、「 $t$  時点までイベントが起こらない確率」 (= 「 $t$  時点まで生きている (死なない) 確率」) を  $t$  の関数としてみたものを **生存関数  $S(t)$  (survival function)** と呼び、以下のように定義します。

$$S(t) = P(T > t) = 1 - F(t) = \int_t^{\infty} f(x)dx$$

確率密度関数  $f(t)$  と生存関数  $S(t)$  の間には、以下のような1対1対応があります。

$$\begin{aligned} \frac{d}{dt} S(t) &= \frac{d}{dt} [1 - F(t)] = -f(t) \\ \Leftrightarrow f(t) &= -\frac{d}{dt} S(t) \end{aligned}$$

生存関数のグラフを **生存曲線 (survival curve)** と呼びます。生存関数は  $t = 0$  のとき  $S(0) = 1.0$ 、 $t = \infty$  のとき  $\lim_{t \rightarrow \infty} S(t) = 0$  に収束する単調減少 (非増加) 関数で、生存曲線は必ず右下がりのグラフになります。



ここで、「 $t$ 時点の直前までイベントが起こらなかったとき、続く $\Delta t$ の期間にイベントが起こる確率」(=「 $t$ 時点まで生きていたという条件の下で、続く $\Delta t$ の期間に死亡イベントが起こる条件付き確率」)を考えます。

$$P(t \leq T < t + \Delta t | T \geq t)$$

上の条件付き確率は $\Delta t$ に依存するので、 $\Delta t$ で標準化した値を考えます。

$$\frac{P(t \leq T < t + \Delta t | T \geq t)}{\Delta t}$$

$\Delta t \rightarrow 0$ としたときの極限を $h(t)$ とすると、

$$\begin{aligned} h(t) &= \lim_{\Delta t \rightarrow 0} \frac{P(t \leq T < t + \Delta t | T \geq t)}{\Delta t} = \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} \frac{P[(t \leq T < t + \Delta t) \cap (T \geq t)]}{P(T \geq t)} \\ &= \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} \frac{P(t \leq T < t + \Delta t)}{P(T \geq t)} \\ &= \frac{1}{P(T \geq t)} \lim_{\Delta t \rightarrow 0} \frac{P(t \leq T < t + \Delta t)}{\Delta t} = \frac{1}{S(t)} \lim_{\Delta t \rightarrow 0} \frac{S(t) - S(t + \Delta t)}{\Delta t} = \frac{1}{S(t)} \lim_{\Delta t \rightarrow 0} \frac{F(t + \Delta t) - F(t)}{\Delta t} \\ &= \frac{1}{S(t)} \left( \frac{d}{dt} F(t) \right) = \frac{f(t)}{S(t)} = -\frac{d}{dt} \log[S(t)]. \end{aligned}$$

このとき、 $h(t)$ を**ハザード関数 (hazard function)**と呼び、「 $t$ 時点直前までイベントが起こらなかったという条件の下で、 $t$ 時点でイベントが起こる瞬間的な確率」あるいは「瞬間死亡率」を表します。ハザード関数は $t$ 時点におけるイベントのリスクを表しており、必ず非負の関数になります。

$$h(t) = \frac{f(t)}{S(t)} = -\frac{d}{dt} \log[S(t)].$$

また、これと関連して、**累積ハザード関数 (cumulative hazard function)**を以下のように定義します。

$$\begin{aligned} H(t) &= \int_0^t h(x) dx = -\log[S(t)], \\ S(t) &= e^{-H(t)} = \exp\left[-\int_0^t h(x) dx\right] = \exp[-H(t)] \end{aligned}$$

ここで、 $H(t) = \int_0^t h(x) dx$ は累積ハザード (cumulative hazard)。

以上の定義から、確率変数 $T$ に対しては確率密度関数 $f(t)$ 、累積分布関数 $F(t)$ 、生存関数 $S(t)$ 、ハザード関数 $h(t)$ 、累積ハザード関数 $H(t)$ が一意的に定まります。

### 3. 生存率と Kaplan-Meier 推定量

生存曲線を考えるため、例として生存日数が

$$2, 4, 4, 5, 7, 8, 10$$

である生存時間データを考えます。

1. 観察開始時0日から2日目まで、イベントは起こらず全て生存していたので、

$$S(t) = 1.0, 0 \leq t \leq 2$$

2. 2日目で1件イベントが発生（死亡）したので、2日目が過ぎた瞬間を  $2+0$  とすると、

$$S(2+0) = \frac{6}{7}$$

3.  $2+0$  日目から4日目まではイベントが起こらないので

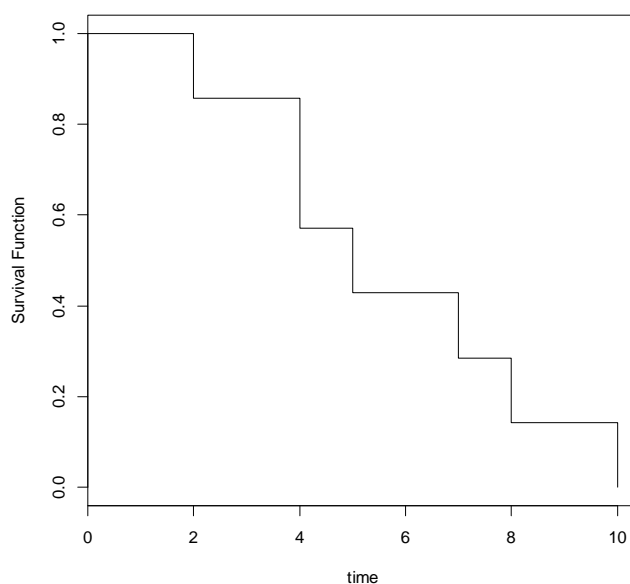
$$S(t) = \frac{6}{7}, 2 < t \leq 4$$

4. 4日目には2件イベントが起きています。このように同じ時点で複数のイベントが起こることをタイ(tie)といいます。4日目が過ぎた瞬間のでは、

$$S(4+0) = \frac{4}{7}$$

5. これ以降も同様に、生存曲線は以下ようになります。

図 1



ここで4日目の状況を考えると、4日目までは当初7件あった症例のうち6件が生存中であるのに対して、4日目にタイのイベント2件が起こり4/6の割合でイベントが発生したことになります。このことを踏まえ、

$$S(4+0) = \frac{4}{7} = \frac{4}{6} \times \frac{6}{7}$$

と書き直してみます。すなわち、時点4+0における生存率は、

(4日目に生存中の個体のうち、4日目を過ぎた瞬間生存している個体の割合)  
× (4日目までの生存率)

と書き直すことが出来たこととなります。この考えは、「打ち切り」がある場合の生存率の定義に一般化することが出来ます。

ここで「打ち切り」がある場合の生存率を考えるため、7日目と10日目のデータが打ち切りデータであるとします。これを以下のように表記します。

2, 4, 4, 5, 7+, 8, 10+

5日目までに4件のイベントが起こっていますので、5日の時点の生存率は3/7です。7日目に最初の打ち切りが起こりますので、打ち切り例の生存時間は「7以上」ということ以外は不明で、5日から次のイベントが起こる8日までの生存率は3/7のままです。

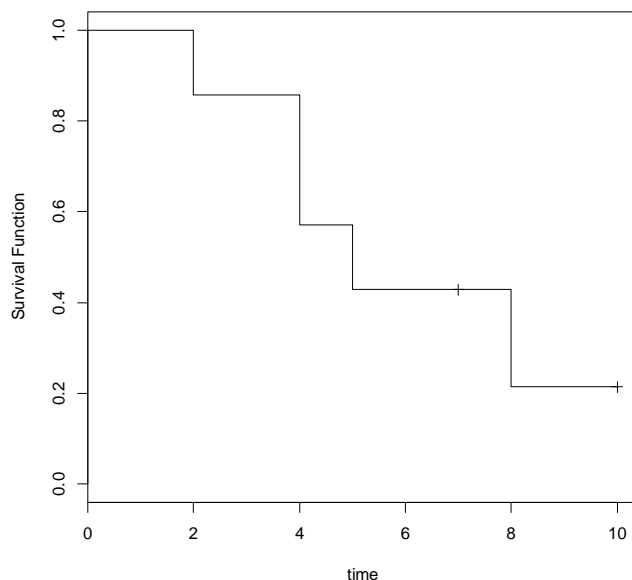
$$S(t) = \frac{3}{7}, 5 < t \leq 8$$

一方、8日目には1件イベントが起こりますが、それ以前7日目に打ち切りが1件ありますので8日目に生存中の個体は3件から2件に減り、8日目を過ぎた瞬間生存しているのは1件になります。結局、8+0時点における生存率は以下のようになります。

(8日目に生存中の個体のうち、8日目を過ぎた瞬間生存している個体の割合)  
× (8日目までの生存率)  
= (1/2) × (3/7) = 3/14 = 1.5/7 = 0.214.

10日目に2回目の打ち切りが有り、打ち切り例の生存時間は不明ですから、8日目以降の生存曲線の値は3/14のままになります。なお、イベント発生時の個体数を“**number at risk**”と言います(=リスクに直面する個体数 = イベントが起こる直前の個体数)。7日目と10日目が打ち切りであった場合の生存曲線は、以下の図2のようになります。

図 2



いま，一般に生存時間  $t_1 < t_2 < \dots < t_m$  においてイベントが観察されたとします。観察開始時点  $t_0 = 0$  とし，  $t_j$  時点におけるイベントの数を  $d_j \geq 1$ ， number at risk を  $n_j \geq 1$ ，ただし  $d_0 = 0, n_0 = n$  とします。期間  $[t_j, t_{j+1})$  に観察された打ち切りを  $w_j \geq 0$  とすると，  $n_{j+1} = n_j - (d_j + w_j)$ ，  $t_j$  時点を過ぎた瞬間の個体数は  $(n_j - d_j)$  となります。前ページの議論を踏襲して，  $t_j + 0$  時点の生存率は以下で与えられます。

$$\begin{aligned}
 S(t_j + 0) &= \frac{n_j - d_j}{n_j} S(t_{j-1} + 0) = \frac{n_j - d_j}{n_j} \frac{n_{j-1} - d_{j-1}}{n_{j-1}} S(t_{j-2} + 0) = \dots \\
 &= \prod_{1 \leq i \leq j} \left( \frac{n_i - d_i}{n_i} \right)
 \end{aligned}$$

一般に，以下に定義されるものを生存関数に対する **Kaplan-Meier (Product limit) 推定量** と呼びます。

**Kaplan-Meier 推定量** : 
$$S(t) = \prod_{j: t_j < t} \left( \frac{n_j - d_j}{n_j} \right)$$

#### 4. 2群の生存関数の差の検定

いま、治療群と対照群のように、二つの群（例えば、薬剤の投与群と非投与群など）の間で生存関数に差があるかどうかを検定したいとします。このとき一般的に用いられる検定は、以下に述べる **log-rank 検定** になります。まず検定の帰無仮説と対立仮説は、以下のようになります。

$$H_0 : S_1(t) = S_2(t), \forall t$$

$$H_1 : S_1(t) \neq S_2(t)$$

帰無仮説が正しいと仮定すると、治療群と対照群に差はなくなりますから、全てのデータを統合して一つのデータセットを作ります。そして、イベントが起こるごとに、以下の  $2 \times 2$  分割表を考えます。

	死亡数	生存数	合計
治療群	$d_j$		$n_j$
対照群			$N_j - n_j$
合計	$D_j$	$N_j - D_j$	$N_j$

ただし、 $N_j$ :治療群，対照群を合わせたときの number at risk,  $n_j$ :治療群の number at risk,  $d_j$  : 治療群のイベント数,  $D_j$ :治療群，対照群を合わせたときのイベント数. このとき,  $d_j$ の期待値 $E_j$ と分散 $V_j$ を求めると, log-rank 検定の検定統計量は以下ようになります。(詳細は, 生物統計学の教科書をご覧ください.)

$$Z^2 = \frac{(D - E)^2}{V} \sim \chi(1) \text{ under } H_0, \text{ where } D = \sum d_j, E = \sum E_j, V = \sum V_j.$$



例: 白血病患者に対する寛解期間の臨床比較試験

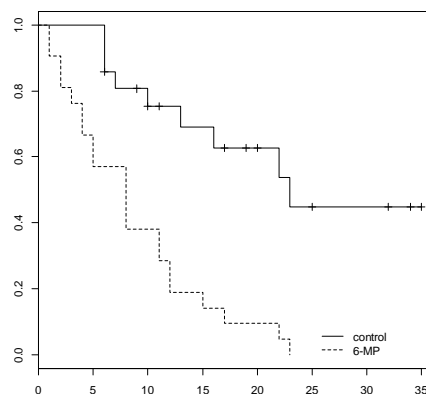
time: resimen time in weeks

cens: censoring, 0/1

treat: treatment, control or 6-MP (6-mercaptopurine)

pair	time	cens	treat
1	1	1	control
1	10	1	6-MP
2	22	1	control
2	7	1	6-MP
3	3	1	control
3	32	0	6-MP

Log-ran test p-value =  $4.17 \times 10^{-5}$



Gehan, E.A. (1965) A generalized Wilcoxon test for comparing arbitrarily single-censored samples. *Biometrika* **52**, 203–233.

#### 4-1. log-rank 検定の特徴と問題点

Log-rank 検定は、2群の生存関数を比較する代表的な手法ですが、いくつかの特徴とそれに伴う問題点が存在します。

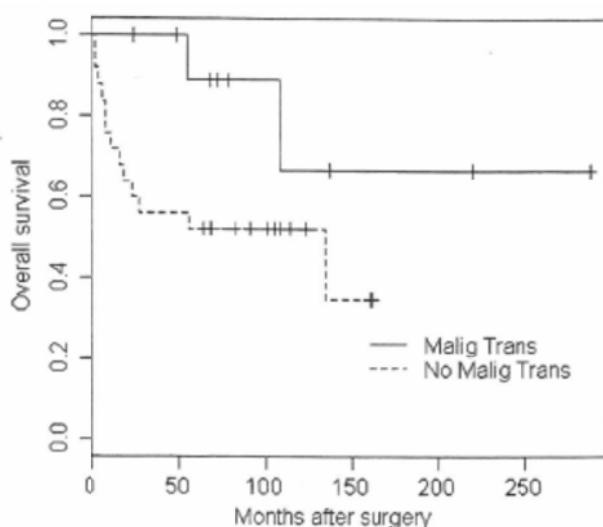
**Log-rank 検定は時間に依存しない:** 前に述べたとおり、log-rank 検定はイベントごとに作った分割表の検定を統合する形で構成されています。従ってイベントが起こる順序だけが重要であって、イベントが起こった特定の時点には依存しません。そのため、Kaplan-Meier 推定量で推定した生存曲線の形から受ける印象と、log-rank 検定の結果が一致しないことがあります。

**Log-rank 検定は単変量解析である:** log-rank 検定では、単一の因子によって群を場合分けし、群間で生存関数に有意差がないかを検定します。従って、生存時間に影響を与えるような第2, 第3の因子は存在しないことが大前提になります。ランダム化比較試験のように、背景因子に系統的な差がない場合には log-rank 検定は適していますが、多数の要因が関与するような場合は log-rank 検定は不適切です。そのような場合は、以下に述べる**層別 log-rank 検定** (stratified log-rank test) や、**Cox 比例ハザードモデル**などを利用する必要があります。

## 2群の生存関数の比較は log-rank 検定だけではない: (一般化 Wilcoxon 検定 (Generalized Wilcoxon test, Peto-Peto test))

2群の生存関数の比較には, log-rank 検定が代表的ですが, その他に**一般化 Wilcoxon 検定 (Generalized Wilcoxon test, Peto-Peto-test)**などが用いられることがある. 一般化 Wilcoxon 検定は, 特に, より早い時点での生存曲線の差に対して検出力が高いとされている.

例: Abe, M. et al (2011) Malignant transformation of breast fibroadenoma to malignant phyllodes tumor: long-term outcome of 36 malignant phyllodes tumors, Breast Cancer (2011) 18:268-272



log-rank test  $p = 0.0551$ , Peto and Peto's test  $p = 0.0409$

## 層別 log-rank 検定 (stratified log-rank test):

Log-rank 検定で 2 群を比較する際, 群を分ける要因以外に, アウトカムに影響を与える因子が存在する場合があります. 上に述べたとおり, log-rank 検定は単変量解析ですから, 複数の共変量の影響を同時に解析することは出来ません. そのため, 第 2, 第 3 の危険因子があるときに log-rank 検定を用いることは, 不適切な解析となります.

主となる危険因子以外の危険因子によってデータがいくつかの層 (strata = subgroup) に分けられるとき, 層ごとに分割表を集計し, 後で全体を合成する**層別 log-rank 検定 (stratified log-rank test)**と呼ばれる方法が用いられます. この

層別を行うことによって、主となる危険因子以外の要因の影響を回避するわけ  
です。例えば、**多施設共同研究**などでは、実施施設の種類ごとに層別すること  
が行われます。

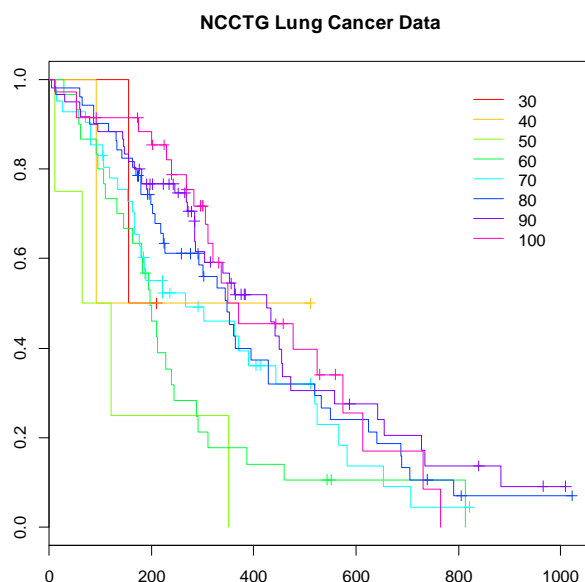
例: NCCTG Lung Cancer Data

inst: Institution code

time: Survival time in days

status: censoring status 1=censored, 2=dead

pat.karno: Karnofsky performance score (bad=0-good=100) as rated by  
patient



Log-rank 検定:  $p=0.00202$

層別 log-rank 検定 (実施施設で層別):  $p=0.00326$

ただし、層別 log-rank 検定には、以下の欠点が指摘されています。

- 層の分割が、恣意的になりがち。
- 層別のための危険因子として、あまり多くの数の因子を考えられない。
- 多数の層に分割すると、層ごとのサンプル数が少なくなる。

これらの欠点を克服するため、次項では Cox 比例ハザードモデルを検討します。

## 5. Cox 比例ハザードモデル (Cox proportional hazard model)

Cox 比例ハザードモデルは、多変量の生存時間解析モデルであり、複数の共変量に影響され症例ごとに生存時間関数が異なる場合を扱います。本項のはじめで、**ハザード関数 (hazard function)  $h(t)$** とは「t 時点直前までイベントが起こらなかったという条件の下で、t 時点でイベントが起こる瞬間的な確率」であると定義しました。そして、ハザード関数は確率密度関数、分布関数、生存関数などと共に、生存時間を表す確率変数  $T$  の確率分布を一意的に定めています。

Cox 比例ハザードモデルでは、ハザード関数  $h(t)$  と共変量  $x_1, \dots, x_k$  に対して、以下の関係を仮定します。

**Cox 比例ハザードモデル:** 
$$h(t | x) = h_0(t) \exp\{\beta_0 + \beta_1 x_1 + \dots + \beta_k x_k\}$$

上で定義した Cox 比例ハザードモデルでは、以下の二つの性質が仮定されています。

**比例ハザード性:** Cox 比例ハザードモデルでは、時間に依存する部分  $h_0(t)$  と、共変量に依存する部分  $\exp\{\beta_0 + \beta_1 x_1 + \dots + \beta_k x_k\}$  が分割されていて、それが乗法によって結合されています。  $h_0(t)$  を **ベースラインハザード (baseline hazard)**  $\exp\{\beta_0 + \beta_1 x_1 + \dots + \beta_k x_k\}$  を **相対危険度関数 (relative risk function)** と呼んでいきます。この性質のため、任意の時点  $t$  において、異なる共変量の値を持つ二つのハザード関数の間には、以下の通り比例関係が成立します。

$$\frac{h(t | x')}{h(t | x)} = \frac{h_0(t) \exp\{\beta_0 + \beta_1 x'_1 + \dots + \beta_k x'_k\}}{h_0(t) \exp\{\beta_0 + \beta_1 x_1 + \dots + \beta_k x_k\}} = \exp\{\beta'(x'-x)\}$$

$$\Leftrightarrow$$

$$h(t | x') = \exp\{\beta'(x'-x)\} h(t | x)$$

下の式は、任意の時点  $t$  において、一方のハザード関数  $h(t | x')$  は他方のハザード

関数  $h(t|x)$  の定数倍になっており、比例定数である  $\exp\{\beta'(x'-x)\}$  のみが共変量の値に依存することを示しています。  $\exp\{\beta'(x'-x)\}$  を  $x$  の  $x'$  に対する **相対危険度 (relative risk of  $x'$  to  $x$ )** と読んでいます。

**対数線形性:** Cox 比例ハザードモデルでは、ハザード関数が共変量に依存する部分は相対危険度関数  $\exp\{\beta_0 + \beta_1 x_1 + \dots + \beta_k x_k\}$  であり、対数をとれば共変量の一次式  $\beta_0 + \beta_1 x_1 + \dots + \beta_k x_k$  になります。つまり、Cox 比例ハザードモデルは、共変量の一次式を通じてのみ共変量に依存していることになります。またこの性質は、Cox モデルの係数  $\beta$  の解釈も与えてくれます。すなわち、 $x' = (x_1 + 1, x_2, \dots, x_k)$  と  $x = (x_1, x_2, \dots, x_k)$  のように、 $x_1$  の値のみが 1 単位変化し、他の共変量の値が一定であった場合、ハザード関数の比は

$$\frac{h(t|x')}{h(t|x)} = \frac{h_0(t)\exp\{\beta_0 + \beta_1(x_1 + 1) + \dots + \beta_k x'_k\}}{h_0(t)\exp\{\beta_0 + \beta_1 x_1 + \dots + \beta_k x_k\}} = e^{\beta_1}$$

となります。つまり、Cox モデルにおいて  $x_1$  の値のみが 1 単位変化したときのハザード比は、 $\beta$  の指数変換に相当することになります。

## 6. Cox 比例ハザードモデルの推定と検定

Cox 比例ハザードモデルは、部分尤度法 (partial likelihood method) によって推定されます。最尤推定量の性質に従い、信頼区間の構成、仮説検定が行われます。統計解析ソフトからの出力は、回帰分析などの場合と似ており、同様の解釈が可能です。

例: 白血病患者に対する寛解期間の臨床比較試験 (続き)

	coef	exp(coef)	se(coef)	z	Pr(> z )	
treatcontrol	1.5721	4.8169	0.4124	3.812	0.000138	***
	exp(coef)	exp(-coef)	lower 0.95	upper 0.95		
treatcontrol	4.817	0.2076	2.147	10.81		
Likelihood ratio test= 16.35 on 1 df, p=5.261e-05						
Wald test = 14.53 on 1 df, p=0.0001378						
Score (logrank) test = 17.25 on 1 df, p=3.283e-05						

Cox 比例ハザードモデル:推定と検定のチェックポイント :

1. パラメターの推定値  $\beta_1 = 1.5721$
2. パラメターの有意性検定の p 値  $p = 0.000138$
3. ハザード比  $e^{\beta_1} = 4.8169$
4. ハザード比の信頼区間 (2.147, 10.81)
5. モデルの適合度検定 (以下の三通りあるが, 気にしない)

Likelihood ratio test = 16.35 on 1 df, p = 5.261e-05

Wald test = 14.53 on 1 df, p = 0.0001378

Score (logrank) test = 17.25 on 1 df, p = 3.283e-05

## 7. Cox モデルの比例ハザード性の検証

生存時間関数と累積ハザード関数を結びつける公式,

$$H(t) = \int_0^t h(x) dx = -\log[S(t)],$$

$$S(t) = e^{-H(t)} = \exp\left[-\int_0^t h(x) dx\right] = \exp[-\Lambda(t)]$$

を思い出すと, Cox 比例ハザードモデルの生存時間関数は

$$H(t) = \int_0^t h(x) dx = -\log[S(t)],$$

$$S(t) = \exp\left[-\int_0^t h(u) du\right] = \exp\left[-\int_0^t h_0(u) \exp\{\beta_0 + \beta_1 x_1 + \dots + \beta_k x_k\} dx\right]$$

$$= \exp\left[-\exp\{\beta_0 + \beta_1 x_1 + \dots + \beta_k x_k\} \int_0^t h_0(u) dx\right]$$

$$= \exp\left[-\int_0^t h_0(u) dx\right]^{\exp\{\beta_0 + \beta_1 x_1 + \dots + \beta_k x_k\}} = S_0(t)^{\exp\{\beta_0 + \beta_1 x_1 + \dots + \beta_k x_k\}}$$

と書き直せます. 両辺の対数をとると,

$$\log S(t) = \exp\{\beta_0 + \beta_1 x_1 + \dots + \beta_k x_k\} \log S_0(t)$$

最後に, 両辺にマイナスを掛けて再び log をとると, 以下を得ることが出来ます.

$$\log\{-\log S(t)\} = \log\{-\log S_0(t)\} + \{\beta_0 + \beta_1 x_1 + \dots + \beta_k x_k\}$$

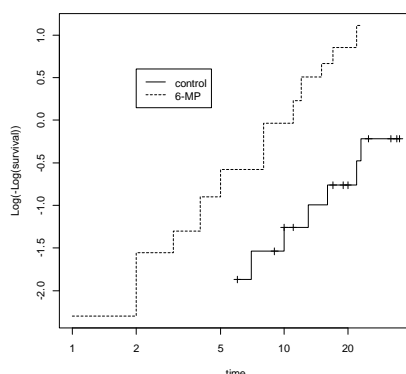
最後に式は、複数の群に対して生存曲線  $S(t)$  を当てはめた時、もし Cox モデルにおいて比例ハザード性が満たされていれば、任意の時点  $t$  に対して  $\log\{-\log S(t)\}$  の値は  $\log\{-\log S_0(t)\}$  (これは各群に共通) を定数分だけ平行に移動することを意味しています。

### 7-1. 補対数-対数プロット (complementary log-log plot)

$$\log\{-\log S(t)\} = \log\{-\log S_0(t)\} + \{\beta_0 + \beta_1 x_1 + \dots + \beta_k x_k\}$$

を検証するため、縦軸に  $\log\{-\log S(t)\}$ 、横軸に  $\log t$  を層ごとにプロットします。もし、比例ハザード性が成り立っているならば、各層のプロットは平行になるはずです。

例: Gehan の白血病データの complementary log-log plot



6-MP による治療群と、対照群で log-log plot を描きました。プロットは十分に平行であり、比例ハザード性は満たされていると考えられます。

### 7-2. シェーンフィールド残差 (Schoenfeld residual) に対する $\chi^2$ 検定

前項で検討した補対数-対数プロットは、サンプルが複数の群 (例えばプラセボ群と介入群) に分けられるとき、群間で比例ハザード性が成り立つかを検討する手法でした。逆に言えば、他の共変量の影響は全く考慮せず群の違いのみに着目した方法ということが出来ます。しかし、実際にはすべての共変量について Cox の比例ハザードモデルの比例ハザード性、

$$h(t | x) = h_0(t) \exp\{\beta_0 + \beta_1 x_1 + \dots + \beta_k x_k\}$$

が成り立っていることを確認する必要があります。このとき個々の共変量ごとに比例ハザード性を検定するための概念として、**シェーンフィールド残差 (Schoenfeld residual)** があります。シェーンフィールド残差の定義の詳細は省略するが、これを用いて各共変量ごとに比例ハザード性を検定できます。検定の帰無仮説は「比例ハザード性は成り立っている」になります。前項の Gehan の白血病データについて、統計解析ソフト R の `cox.zph()` コマンドでシェーンフィールド残差の検定を行ってみます。(survival パッケージが必要)

```
> fit.ph <- coxph(Surv(time, cens) ~ treat, data=gehan)
> cox.zph(fit.ph)

           rho  chisq    p
treatcontrol -0.0279 0.0229 0.88
```

ここでは群分けのフラグとなる共変量 `treat` に対して、検定の `p` 値が `0.88` ですから帰無仮説は棄却されません。すなわち「比例ハザード性は成り立っている」という仮説が受け入れられたことになります。

## 8. 非比例ハザード性への対処

前項最後の例では、補対数-対数プロットが平行になり、比例ハザード性が満たされていることがわかりました。しかし、実際のデータにおいて、比例ハザード性が破綻していた場合、どのように対処したら良いでしょうか。

比例ハザード性が破綻しているとは、ハザード比が時間に依存するということです。その原因としては、

1. データ集団の中に、時間への依存の仕方が異なる、複数のハザード関数が存在する。
2. 共変量が時間によって一定ではない。
3. 対数線形性が破綻しており、非線形な構造が存在する。

などの可能性が考えられます。これらへの対応には、

1. 層別に異なるベースラインハザードを適用する、層別 Cox 比例ハザードモデルを適用する。



2. 時間依存型の共変量を導入する.
3. Cox モデルの線形項に、非線形変換を導入した加法型モデルを検討する.

などの対処が考えられます。ただし、これらの方法は、理論的にも未だ発展途上の部分が有り、専門家でも意見の分かれるところです。一概に正解はないと思いますので、慎重なモデルの修正が必要です。

## 9. 一般化加法モデル (Generalized Additive Model: GAM)による非線形モデル

前項で述べたとおり、比例ハザード性の仮定が破たんした場合、共変量の非線形変換を導入することで問題を回避できる場合がある。

Cox 比例ハザードモデルの本質は、以下の二点にまとめられます。

1. ハザード関数が、時間に依存する部分と共変量に依存する部分に分かれている。
2. ハザード関数が共変量に依存する部分は、共変量の一次式に依存している。

$$h(t | x_1, \dots, x_k) = h_0(t) \exp\{\beta_0 + \beta_1 x_1 + \dots + \beta_k x_k\}$$

共変量に関する一次性が失われたとき、Cox 比例ハザードモデルに必要な比例ハザード性の仮定が破綻してしまうこととなります。そのような場合に、Cox モデルの線形項  $(\beta_0 + \beta_1 x_1 + \dots + \beta_k x_k)$  に非線形関数を導入したものが**一般化加法モデル (Generalized Additive Model, GAM)** になります。

$$h(t | x_1, \dots, x_k) = h_0(t) \exp\{\beta_0 + f_1(x_1) + \dots + f_k(x_k)\}$$

例: 在郷軍人の肺がんデータ

米国在郷軍人局による肺がん試験に関するデータ。

trt: 1=standard 2=test  
 time: survival time  
 karno: Karnofsky performance score (100=good)  
 age: in years

在郷軍人の肺がんデータに対して、まず Cox 比例ハザードモデルを当てはめてみると、Schoenfeld 残差の検定から karno と age で強い非比例ハザード性が認められます。

	rho	chisq	p
trt	-0.104	1.44	0.23093
karno		0.333	13.49 0.00024
age	0.183	4.90	0.02681
GLOBAL	NA	15.80	0.00124

そこで、karno と age に非線形関数を導入して一般化加法モデルを推定しました。その結果、karno (の非線形変換) のみが有意に予後に相関することが分かりました。

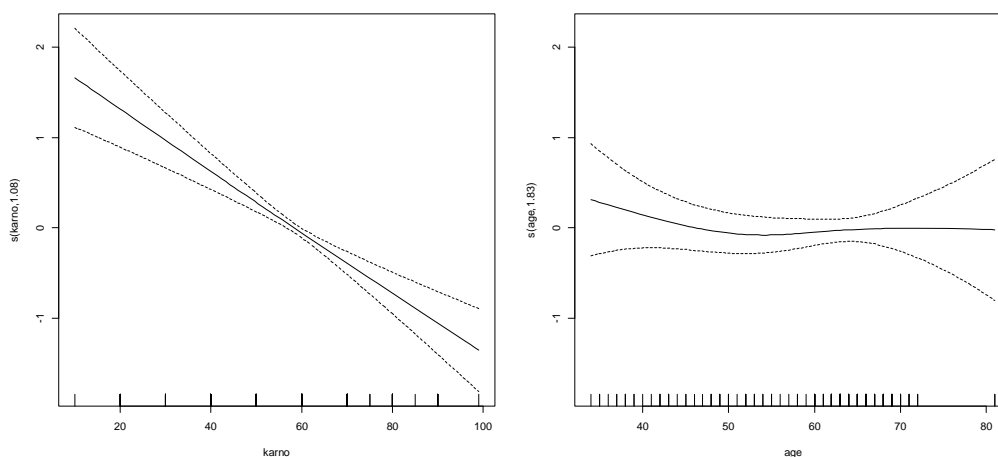
Parametric coefficients:

	Estimate	Std. Error	z value	Pr(> z )
trt	0.1970	0.1867	1.055	0.291

Approximate significance of smooth terms:

	edf	Ref.df	Chi.sq	p-value
s(karno)	1.080	1.153	39.604	1.16e-09 ***
s(age)	1.826	2.325	0.526	0.829

karno と age に非線形を図示すると、以下の通りとなります。縦軸は対数ハザードを表しています。



## Take Home Message

1. 生存時間解析
2. 生存時間解析の基本概念
3. 生存率と Kaplan-Meier 推定量
4. 2 群間の生存関数の差の検定
5. Cox 比例ハザードモデル
6. Cox 比例ハザードモデルの推定と検定
7. Cox モデルの比例ハザード性の検証
8. 非比例ハザード性への対処
9. 一般化加法モデル

### 参考文献：

- 中村 剛 (著), 丹後 俊郎 (編集)「Cox 比例ハザードモデル (医学統計学シリーズ)」朝倉書店 (2001/04) ISBN-13: 978-4254127539
- 大橋 靖雄 (著), 浜田 知久馬 (著)「生存時間解析—SAS による生物統計」東京大学出版会 (1995/04) ISBN-13: 978-4130602006
- T.J. Hastie, R.J. Tibshirani “Generalized Additive Models” Chapman and Hall/CRC; 00002 版 (1990/6/1) ISBN-13: 978-0412343902
- Simon Wood “Generalized Additive Models: An Introduction with R” Chapman and Hall/CRC (2006/2/27) ISBN-13: 978-1584884743

以上