

逐次検定のための SAS/IML 関数の紹介と

中間解析を伴う生存時間データの例数設計への適用

橋本 隆正、 宇野 一

武田薬品工業株式会社

医薬開発本部 統計解析部統計グループ

1. はじめに

宇野(1998)では SAS/DATA ステップにおいて独自にプログラミングすることにより群逐次検定における棄却限界値の算出等を行ったが、実は SAS/IML の中に、逐次検定のためのサブルーチンが3つあり、宇野(1998)で例示された計算結果は、これらのサブルーチンを用いることで再現できる。本報告では、逐次検定のためのサブルーチンを紹介し、これらを用い中間解析を伴う生存時間データの例数設計に適用することを試みる。

2. 問題の定式化と記号

本報告では宇野(1998)と同じ記号を用いる。即ち：

A 群、B 群の 2 群間の平均値の差の比較を考える。各群のデータは、互いに独立に平均 μ_A 、 μ_B 、両群共通の既知の分散 σ^2 の正規分布に従うものとする。また両群併せて $2n$ 例（各群 n 例）集積される毎に解析し、最大 K 回の解析を行う（すなわち各群で nK 例集積されるまで行う）。 $i-1$ 回目の解析時点から i 回目の解析時点までに観測される患者の応答を X_{Aij} 、

X_{Bij} ($j=1, \dots, n$) とする。

A 群: $X_{Aij} \sim N(\mu_A, \sigma^2)$ 、B 群: $X_{Bij} \sim N(\mu_B, \sigma^2)$ ($i=1, \dots, K; j=1, \dots, n$)

$i-1$ 回目の解析時点から i 回目の解析時点までに得られる $2n$ 個のデータから作られる統計量 Z_i を、

$$Z_i \equiv \frac{\sqrt{n}(\bar{X}_{Ai} - \bar{X}_{Bi})}{\sqrt{2}\sigma}$$

とする。ここで $\bar{X}_{Ai} = \frac{\sum_{j=1}^n X_{Aij}}{n}$ 、 $\bar{X}_{Bi} = \frac{\sum_{j=1}^n X_{Bij}}{n}$ である。

この時、 Z_1, \dots, Z_K は互いに独立に平均 $\frac{\sqrt{n}(\mu_A - \mu_B)}{\sqrt{2\sigma}}$ 、分散 1 の正規分布に従う。ここで、

$\Delta \equiv \frac{\sqrt{n}(\mu_A - \mu_B)}{\sqrt{2\sigma}}$ とおくと、 $Z_i \sim N(\Delta, 1)$ と表される。さらに、正規分布の再生性より、

$S_i \equiv \sum_{j=1}^i Z_j \sim N(\Delta i, i)$ が成り立つ。

帰無仮説 $H_0: \mu_A = \mu_B$ 、対立仮説 $H_A: \mu_A \neq \mu_B$ の両側検定を考える。この仮説は $H_0: \Delta = 0, H_A: \Delta \neq 0$ と同じである。各 i 回目の解析時の棄却限界値を $c_i (i = 1, \dots, K)$ とし、

$$|S_i| \geq c_i \Rightarrow \text{試験を中止し } H_0 \text{ を棄却する } (i = 1, \dots, K-1)$$

$$|S_i| < c_i \Rightarrow \text{試験を継続する } (i = 1, \dots, K-1)$$

$$|S_K| \geq c_K \Rightarrow H_0 \text{ を棄却する}$$

$$|S_K| < c_K \Rightarrow H_0 \text{ を受容する}$$

とすると、この一連の解析において、帰無仮説が棄却される確率 P_K は以下のように表せる。

$$\begin{aligned} P_K &= \Pr\{|S_1| \geq c_1 \text{ or } |S_2| \geq c_2 \text{ or } \dots \text{ or } |S_K| \geq c_K\} \\ &= 1 - \Pr\{|S_1| < c_1 \text{ and } |S_2| < c_2 \text{ and } \dots \text{ and } |S_K| < c_K\} \end{aligned}$$

帰無仮説のもとでは、上記の P_K が試験全体の第一種の過誤の確率 (α) となり、対立仮説のもとでは、検出力 ($1 - \beta$) となる。

帰無仮説のもとで、上記 P_K が試験全体の第一種の過誤の確率 (α) となるような、棄却限界値の列 (c_1, \dots, c_K) は無限に存在し、これらを求めるには何らかの工夫が必要である。本報告では、EaSt でも採用されている Wang and Tsiatis(1987)の one parameter boundaries アプローチを用いて逐次検定のためのサブルーチンの紹介を行う。このアプローチは棄却限界値の列 c_1, \dots, c_K に以下の制約を入れるものである。即ち：

$$c_i = C_w \times i^\lambda$$

ここで、 C_w 、 λ は定数である。 λ は boundary shape parameter と呼ばれ、0 から 0.5 の間で任意に設定できる。特に $\lambda = 0$ の時は O'Breien and Fleming(1979)、 $\lambda = 0.5$ の時は Pocock(1977)で提案されている棄却限界値の列に一致することが知られている。従って、これらより以下の式を満足する C_w を求め、棄却限界値 c_1, \dots, c_K を得ることができる。

$$P_K = 1 - \Pr\{|S_1| < C_w \text{ and } |S_2| < C_w 2^\lambda \text{ and } \dots \text{ and } |S_K| < C_w K^\lambda\} = \alpha \quad \text{under } H_0$$

宇野 (1998) では、Armitage et al.(1969)で提案された Recursive Integration 法、即ち検定統計量の独立増分構造を利用することにより、確率量を convolution の形で表現し積分の次数を減らし逐次的に棄却限界値を求める方法で、SAS/DATA ステップでプログラミングすることによって上記関係式を満足する C_w (あるいは c_1, \dots, c_K) を求めた。ところで、実は SAS/IML の中に同じ計算アルゴリズム、即ち Recursive Integration 法を用いて逐次検定を行うためのサブルーチンが用意されている。次章ではそれらのサブルーチンを紹介する。

3. SAS/IML で提供されている逐次検定のためのサブルーチンの紹介

SAS/IML には逐次検定のためのサブルーチンが3つ存在する。それらは SEQ、SEQSCALE、SEQSHIFT というサブルーチンであり、各サブルーチンの機能を大まかに述べると (well-defined でない表現もあるかもしれないが詳細の説明で明らかになると思われる) :

SEQ : 各解析における棄却限界値を具体的に数値として与えることで、各解析において検定統計量が棄却限界値を越えない確率を返す ($=0$ のもとで)

SEQSCALE : 全体の有意水準と domain という量を与えることで、各解析における棄却限界値と各解析において検定統計量が棄却限界値を越えない確率を返す ($=0$ のもとで)

SEQSHIFT : 全体の検出力、棄却限界値を具体的に数値として与えることで、その条件にて検出される基準化された差と、その場合に各解析において検定統計量が棄却限界値を越えない確率を返す ($=0$ のもとで)

というものである。

以下では、各サブルーチンの機能を詳細に紹介する。なお、臨床試験の中間解析での適用を念頭においているので、検定統計量や棄却限界値などの記号や parametrization は2章と同じものを用い限定された場面でのみの紹介になっていることに注意されたい。

● サブルーチン SEQ の説明

CALL SEQ(prob, domain) < tscale, eps, den >

第2引数指定必須

第1引数 (prob) : 各 stage での sequential test の確率 (ただし $=0$ のもとで、戻り値 : $(m+1) \times n$ の行列)

第2引数 (domain) : 棄却限界値を指定 (入力値 : $m \times n$ 行列)

tscale オプション : 隣り合う stage の時間間隔 (入力値 : デフォルトは1)

eps オプション : 収束条件 (入力値)

den オプション : 計算された各ステージでの密度の値 (戻り値)

入力パラメータ (第2引数) の指定方法

本報告で扱う parametrization においては、 m は 2、 n は最大解析回数 K であり、具体的

$$\text{には、 } domain = \begin{pmatrix} -c_1 & -c_2 & \cdots & -c_K \\ c_1 & c_2 & \cdots & c_K \end{pmatrix} = \begin{pmatrix} -C_w \times 1^\lambda & -C_w \times 2^\lambda & \cdots & -C_w \times K^\lambda \\ C_w \times 1^\lambda & C_w \times 2^\lambda & \cdots & C_w \times K^\lambda \end{pmatrix}$$

で指定する。つまり、各列は各解析時点を表し、1 行目は各解析時点の棄却限域の下限を 2 行目は各解析時点の棄却域の上限を指定する（下限と上限は原点对称である必要はない）。なお、 $-\infty$ を指定したい場合は **M** で表し、 $+\infty$ を指定したい場合は **P** で表す。一般的には各回における棄却域は互いに素な複数の区間が指定でき、この場合各区間を縦方向に加えて指定する。詳細はマニュアルを参照のこと。

出力パラメータ (prob) の内容

本報告で扱う parametrization においては、 m は 2、 n は最大解析回数 K であり、出力される $(m+1) \times n$ 行列の各要素に格納される量は以下の通りである。この行列も上の domain の一般的な指定に対応して複雑な棄却域の場合においても確率を計算してくれる。なお、本報告では domain を原点对称に指定しているので、スペースの都合上求められる確率を表す際、例えば $|S_1| < c_1$ と記載しているが、domain が非対称で指定されていれば、例えば、

$-\infty < S_1 < c_1$ などと指定されていれば（ $-\infty$ は **M** で指定されている）片側検定に対応した確率 $\Pr\{S_1 < c_1, \dots, S_{K-1} < c_{K-1}, S_K < c_K\}$ も求めることができる。詳細はマニュアルを参照のこと。

$$prob = \begin{pmatrix} \Pr\{-\infty < S_1 < -c_1\} & \Pr\{|S_1| < c_1, -\infty < S_2 < -c_2\} & \cdots & \Pr\{|S_1| < c_1, \dots, |S_{K-1}| < c_{K-1}, -\infty < S_K < -c_K\} \\ \Pr\{-\infty < S_1 < c_1\} & \Pr\{|S_1| < c_1, -\infty < S_2 < c_2\} & \cdots & \Pr\{|S_1| < c_1, \dots, |S_{K-1}| < c_{K-1}, -\infty < S_K < c_K\} \\ \Pr\{-\infty < S_1 < \infty\} & \Pr\{|S_1| < c_1, -\infty < S_2 < \infty\} & \cdots & \Pr\{|S_1| < c_1, \dots, |S_{K-1}| < c_{K-1}, -\infty < S_K < \infty\} \end{pmatrix}$$

なお、

$$P_K = 1 - \Pr\{|S_1| < c_1 \text{ and } |S_2| < c_2 \text{ and } \cdots \text{ and } |S_K| < c_K\} \\ = 1 - [\Pr\{|S_1| < c_1 \text{ and } |S_2| < c_2 \text{ and } \cdots \text{ and } -\infty < S_K < c_K\} - \Pr\{|S_1| < c_1 \text{ and } |S_2| < c_2 \text{ and } \cdots \text{ and } -\infty < S_K < -c_K\}]$$

となることに注意。

例 3-1 (Pocock の棄却限界値が Type I error の確率を満たしているか確認)

今、全解析回数を 5 回とする ($K = 5$)、この時、有意水準 5% (両側) としたときの Pocock の棄却点の列は、 $c_i = C_w \times i^\lambda = 2.413 \times i^{0.5}$ ($i = 1, 2, 3, 4, 5$) であることが知られている。この棄却限界値の列が Type I error の確率を満たしているか確認する。なお、以下のプログラムでは SAS/IML にある行列演算の様々な便利な機能はなるべく利用せず、SAS/IML に

馴染みでない方が本質的なサブルーチンの機能の理解をする前に行列演算で苦しんで挫折ないように心掛けたため、詳しくご存知な方にはかなり冗長で御見苦しい点もあるかもしれないがご容赦願いたい。

(プログラム)

```
proc iml ;
  m={2.413 2.413 2.413 2.413 2.413}#sqrt({1 2 3 4 5}) ;
  domain=-m//m ;
  print domain ;
  call seq(prob, domain) ;
  print prob ;
  alpha={1}-(prob[2,5]-prob[1,5]) ;
  print alpha ;
quit;
```

(アウトプット)

```
DOMAIN
  -2.413 -3.412497 -4.179439 -4.826 -5.395632
   2.413 3.4124973 4.1794386 4.826 5.395632
PROB
0.0079109 0.0058585 0.0045113 0.0036566 0.0030738
0.9920891 0.9783197 0.9679499 0.959782 0.9530516
      1 0.9841782 0.9724612 0.9634386 0.9561254
ALPHA
0.0500222
```

(説明)

DOMAIN を $\begin{pmatrix} -2.413 \times \sqrt{1} & -2.413 \times \sqrt{2} & -2.413 \times \sqrt{3} & -2.413 \times \sqrt{4} & -2.413 \times \sqrt{5} \\ 2.413 \times \sqrt{1} & 2.413 \times \sqrt{2} & 2.413 \times \sqrt{3} & 2.413 \times \sqrt{4} & 2.413 \times \sqrt{5} \end{pmatrix}$ と指定することにより確率

行列 PROB を得て、その第 2 行第 5 列から第 1 行第 5 列の値を引いたものを 1 から引いたものが試験全体の有意水準 ALPHA になり、Pocock の棄却限界値が第一種の過誤の確率を満たしていることが分かった。

● サブルーチン SEQSCALE の説明

CALL SEQSCALE(prob, gscale, domain, level) < iguess, tscale, eps, den >

第 2 引数指定必須

第 1 引数 (prob): SEQ で説明したものと同一。即ち各 stage での sequential test の確率 (m+1)

$\times n$ の行列)が計算される ($=0$ のもとで : 戻り値)

第 2 引数(*gscale*) : 本報告の *parameterization* においては C_w が計算される (戻り値)

第 3 引数(*domain*) : 本報告の *parameterization* においては $\begin{pmatrix} -1^\lambda & -2^\lambda & \dots & -K^\lambda \\ 1^\lambda & 2^\lambda & \dots & K^\lambda \end{pmatrix}$ を指定

(入力値)

第 4 引数(*level*) : $1 - \alpha$ を指定 (入力値)

iguess オプション : 計算のための初期値 (入力値)

tscale オプション : 隣り合う *stage* の時間間隔 (入力値 : デフォルトは 1)

eps オプション : 収束条件 (入力値)

den オプション : 計算された各ステージでの密度の値 (出力値)

例 3-2 (Pocock の棄却限界値を求め Type I error の確率を満たしているか確認)

今、全解析回数を 5 回 ($K = 5$)、全体の有意水準 $\alpha = 0.05$ とする。この時、Pocock の棄却点の列は、 $c_i = C_w \times i^{0.5}$ ($i = 1, 2, 3, 4, 5$) であることが知られている。SEQSCALE では、 $1 - \alpha$ を与えると one parameter boundaries アプローチのフレームワークでの C_w が計算される。なお、*domain* の指定は SEQ と異なることに注意。

(プログラム)

```
proc iml;
  m={1 2 3 4 5}##{0.5} ;
  domain=-m//m ;
  level=1-0.05 ;
  call seqscale(prob,cw,domain,level) ;
  print cw ;
  print prob ;
  alpha={1}-(prob[2,5]-prob[1,5]) ;
  print alpha ;
quit;
```

(アウトプット)

```
CW
2.4131761
PROB
0.0079071 0.0058559 0.0045093 0.0036551 0.0030726
0.9920929 0.9783299 0.9679647 0.9598002 0.9530726
      1 0.9841858 0.972474 0.9634553 0.9561451
```

ALPHA

0.05

(説明)

DOMAIN を $\begin{pmatrix} -\sqrt{1} & -\sqrt{2} & -\sqrt{3} & -\sqrt{4} & -\sqrt{5} \\ \sqrt{1} & \sqrt{2} & \sqrt{3} & \sqrt{4} & \sqrt{5} \end{pmatrix}$ と指定することにより Wang and Tsiatis(1987)の one

parameter boundaries アプローチに基づく $CW=2.4131761$ を得た。これは、Pocock の棄却限界値に一致しており、対応する確率行列 PROB も計算される。PROB の第 2 行第 5 列から第 1 行第 5 列の値を引いたものを 1 から引いたものが試験全体の有意水準 ALPHA になり、第一種の過誤の確率を満たしていることが分かる。

● 基準化された差を求めるサブルーチン SEQSHIFT の説明

CALL SEQSHIFT(prob, shift, domain, level) < iguess, tscale, eps, den >

第 2 パラメータ指定必須

第 1 引数 (prob) : SEQ で説明したものと同一。即ち各 stage での sequential test の確率 $(m+1) \times n$ の行列が計算される (ただし、 のもとで : 戻り値)

第 2 引数(shift) : 本報告の parameterization においては Δ が計算される (戻り値)

第 3 引数(domain) : 本報告の parameterization においては

$$\begin{pmatrix} -C_w \times 1^2 & -C_w \times 2^2 & \dots & -C_w \times K^2 \\ C_w \times 1^2 & C_w \times 2^2 & \dots & C_w \times K^2 \end{pmatrix} \text{を指定 (入力値)}$$

第 4 引数(level) : β を指定 (入力値)

iguess オプション : 計算のための初期値 (入力値)

tscale オプション : 隣り合う stage の時間間隔 (入力値 : デフォルトは 1)

eps オプション : 収束条件 (入力値)

den オプション : 計算された各ステージでの密度の値 (戻り値)

例 3-3 (Pocock の検定における基準化された 2 群の平均値の差の算出)

今、全解析回数を 5 回とする ($K = 5$)。この時、有意水準 5% (両側) としたときの Pocock

の棄却限界値の列、 $c_i = C_w \times i^{0.5}$ ($i = 1, 2, 3, 4, 5$) は、例 3-2 で示したように SEQSCALE

で求めることができる。SEQSHIFT では、棄却限界値の列 c_i を定め、TypeII error の確率 β (1-検出力) を与えた下で、基準化された 2 群の平均値の差を求めることができる。こ

れは、2 章で定義した $\Delta \equiv \frac{\sqrt{n}(\mu_A - \mu_B)}{\sqrt{2}\sigma}$ に相当する。以下では、 $\beta = 0.1$ とする。

(プログラム)

```
proc iml ;
```

```

m={1 2 3 4 5}##{0.5} ;
domain1=-m//m ;
level=1-0.05 ;
call seqscale(prob,cw,domain1,level) ;
domain2=domain1*cw ;
plevel=0.1;
call seqshift(prob,delta,domain2,plevel) ;
print delta ;
print prob ;
power={1}-(prob[2,5]-prob[1,5]) ;
print power ;
quit ;

```

(アウトプット)

```

DELTA
1.5922877
PROB
0.0000309 1.2109E-6 7.4933E-8 5.9919E-9 5.648E-10
0.7941451 0.533881 0.3252798 0.1850807 0.1000009
      1 0.7941141 0.5338798 0.3252798 0.1850807
POWER
0.8999991

```

(説明)

DOMAIN2 を $\begin{pmatrix} -2.413 \times \sqrt{1} & -2.413 \times \sqrt{2} & -2.413 \times \sqrt{3} & -2.413 \times \sqrt{4} & -2.413 \times \sqrt{5} \\ 2.413 \times \sqrt{1} & 2.413 \times \sqrt{2} & 2.413 \times \sqrt{3} & 2.413 \times \sqrt{4} & 2.413 \times \sqrt{5} \end{pmatrix}$ 、第二種の過誤の確率を

0.1(plevel)と指定することにより基準化された群間差 DELTA=1.5922877 を得た。これは、Pocock(1977)あるいは宇野(1998)に記載されている に一致しており、対応する確率行列 PROB も計算される。PROB の第2行第5列から第1行第5列の値を引いたものを1から引いたものが検出力 POWER になり、それが90%になっていることが分かる。

4. 中間解析を伴わない生存時間データの例数設計

以下の例題を考えながら、まず中間解析を行わない場合の生存時間データの例数設計を考える (Shoenfeld and Richter(1982))。

例 4-1

ある疾患のあるステージの患者が無治療で放置された場合の3年後の死亡率は 15.6%と

報告されている。ある治験のプラセボ群（P 群）における 3 年間後の累積死亡率がこれに等しいと仮定し、実薬群（A 群）では死亡率がその半分（7.8%）となることを期待する。このとき、主要評価項目である「死亡」までの期間が指数分布に従うと仮定すると、時刻 t （今回は月とする）までに死亡する確率は、

$$1 - e^{-\lambda t}$$

となる。指数分布の場合、分布はハザードのみで特定されるので、群間比較はハザードに基づいて行えばよい。今、ハザードは P 群で 0.0047 / 月、A 群で 0.0023 / 月となる。これは例えば、P 群のハザード 0.0047 / 月は 3 年後の死亡率は 15.6% と想定しているので、 $\Pr(T_1 < 36) = 1 - e^{-\lambda_1 \times 36} = 0.156$ を λ_1 について解けば求まる。

今、P 群および A 群のハザードをそれぞれ λ_1 、 λ_2 で表すことにする。 λ_1 、 λ_2 の最尤推定量は、それぞれ

$$\hat{\lambda}_1 = \frac{n_1^*}{\sum_{i=1}^{n_1} T_{i1}}, \quad \hat{\lambda}_2 = \frac{n_2^*}{\sum_{i=1}^{n_2} T_{i2}}$$

で表せる。ここで n_1 、 n_2 は両群の登録症例数、 n_1^* 、 n_2^* は両群の死亡発生例数、 T_{i1} 、 T_{i2} は各群の各症例の死亡までの時間または打ち切りまでの時間を表す。

この時、最尤推定量を対数変換したものは、P 群を例にとると以下のような正規分布で近似できる。

$$\log \hat{\lambda}_1 \sim N\left(\log \lambda_1, \frac{1}{n_1^*}\right)$$

今、対比係数 $C' = (1, -1)$ を用いて、検定統計量として以下のものを考える。

$$\frac{C'\Lambda}{\sqrt{\frac{1}{n_1^*} + \frac{1}{n_2^*}}} = \frac{\log \hat{\lambda}_1 - \log \hat{\lambda}_2}{\sqrt{\frac{1}{n_1^*} + \frac{1}{n_2^*}}}$$

検定統計量は、帰無仮説の下で、以下の正規分布に従う。

$$\frac{C'\Lambda}{\sqrt{\frac{1}{n_1^*} + \frac{1}{n_2^*}}} \sim N(0, 1)$$

検定統計量は、対立仮説の下では、以下の正規分布に従う。

$$\frac{C'\Lambda}{\sqrt{\frac{1}{n_1^*} + \frac{1}{n_2^*}}} \sim N\left(\log \frac{\lambda_1}{\lambda_2} / \sqrt{\frac{1}{n_1^*} + \frac{1}{n_2^*}}, 1\right)$$

これから、正規分布の議論が適用でき、検出力と全群の必要イベント数について以下の関係式が導かれる。この時、 $n_1^* \approx n_2^*$ を仮定する。

$$2n_1^* = D = 4 \frac{(Z_\alpha + Z_\beta)^2}{\left(\log \frac{\lambda_1}{\lambda_2}\right)^2} \quad (1)$$

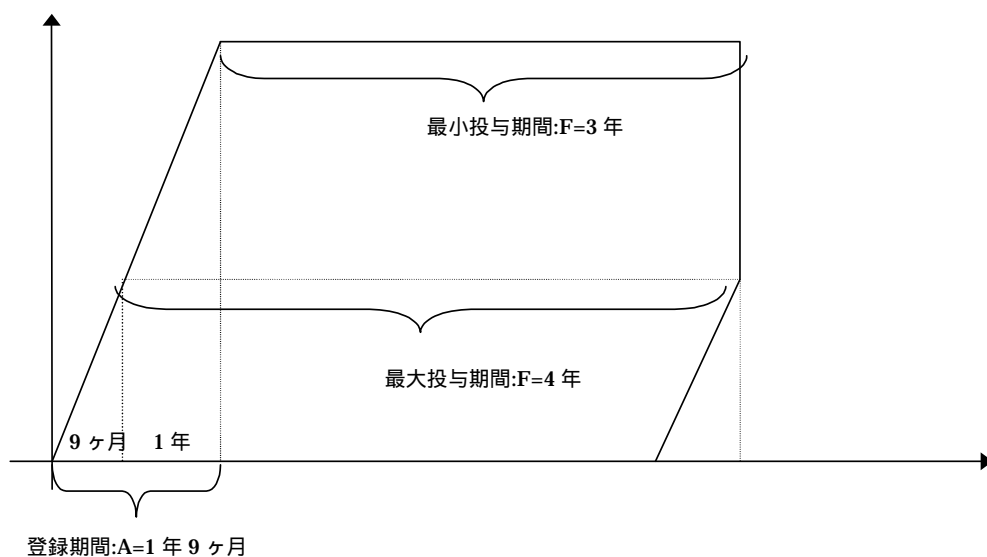
ここで Z_α 、 Z_β は標準正規分布の α 点、 β 点であり、 α は第一種の過誤の確率を表し、 β は第二種の過誤の確率を表す。

なお $n_1^* \approx n_2^*$ の下、中間解析を考慮した場合の例数設計を5章で考えるので、その時必要な

基準化した2群の平均値の差を2章で示した記号で書けば $\Delta = \frac{\sqrt{n_1^*} \times \log \lambda_1 / \lambda_2}{\sqrt{2}}$ であるこ

とに注意。

登録期間を $A=1$ 年9ヶ月(=21週)、投与期間(追跡期間)を F =最長4年(最終症例エントリー終了後3年が経過した時点で治験全体を終了する)。すなわち、以下の図のように治験期間は4年9ヶ月=57ヶ月となり、前半9ヶ月(下図の下半分の平行四辺形の に相当)までにエントリーされた症例は4年間追跡され、後半12ヶ月(下図の上半分の台形の に相当)にエントリーされた症例は3~4年間追跡される。なお、リクルート期間中症例は一律に登録されると仮定する。この時、両群の対数ハザードの差を $=0.05$ (両側)、 $=0.1$ で検出したいとする。



今、 の区間と の区間に分けて考える。

両群で N 例の登録があるとすると、前半 9 ヶ月で登録された $\frac{9}{21}N$ については全例 4 年間追跡されるので、各群の死亡率は $t = 48$ (週) における分布関数の値である。つまり、各群の分布関数 $F_i(t) = 1 - \exp(-t \times \lambda_i)$ ($i = 1, 2$) とおくと、

$$\text{P 群} : F_1(48) = 1 - \exp(-48 \times \lambda_1)$$

$$\text{A 群} : F_2(48) = 1 - \exp(-48 \times \lambda_2)$$

となる。従って、前半に登録された症例の両群併せた期待死亡数は

$$FEN = \frac{9}{21}N \times \frac{1}{2} \times F_1(48) + \frac{9}{21}N \times \frac{1}{2} \times F_2(48)$$

後半 12 ヶ月で登録された $\frac{12}{21}N$ については、1 年間の受け入れ期間で一様に被験者を受け入れた後、最低 3 年間の追跡期間を設けた場合に相当するので、各群の死亡率は

$$\text{P 群} : H_1 = 1 - \frac{1 - \exp(-\lambda_1 \times 12)}{\lambda_1 \times 12} \times \exp(-\lambda_1 \times 36)$$

$$\text{A 群} : H_2 = 1 - \frac{1 - \exp(-\lambda_2 \times 12)}{\lambda_2 \times 12} \times \exp(-\lambda_2 \times 36)$$

となる。

このことを生存時間がパラメータ λ の指数分布に従うという状況下での一般論で示す。受け入れ期間 (A)、追跡期間 (F) という状況で考える。即ち治験期間は $A+F$ となる。受け入れ期間中のある時点 τ で登録され、受け入れ期間終了時点で死亡していないという条件付きで、追跡期間終了時点でも死亡していない確率は、

$$\Pr(T \geq F + A - \tau | T \geq A - \tau) = \frac{1 - \Pr(T < F + A - \tau)}{1 - \Pr(T < A - \tau)} = \frac{\exp(-\lambda(F + A - \tau))}{\exp(-\lambda(A - \tau))} = \exp(-\lambda F)$$

となる。

また、受け入れ期間のどこかの時点で登録されたのち受け入れ期間終了時点で死亡していない確率は、登録の確率に一様分布を仮定すると、

$$\begin{aligned} \Pr(T \geq A - \tau) &= 1 - \Pr(T < A - \tau) = 1 - \int_{\tau \in A} \Pr(T < A - \tau | \tau \text{ で登録}) \Pr(\tau \text{ で登録}) d\tau \\ &= 1 - \int_0^A \{1 - \exp(-\lambda(A - \tau))\} \times \frac{1}{A} d\tau = 1 - \left\{1 - \frac{1 - \exp(-\lambda A)}{\lambda A}\right\} = \frac{1 - \exp(-\lambda A)}{\lambda A} \end{aligned}$$

となる。

従って、受け入れ期間のどこかの時点で登録されたのち追跡終了時点で、死亡していない確率は、上の二つの量の積をとって、

$$\Pr(T \geq F + A - \tau) = \frac{1 - \exp(-\lambda A)}{\lambda A} \times \exp(-\lambda F)$$

となる。よって、死亡確率は、

$$\Pr(T < F + A - \tau) = 1 - \frac{1 - \exp(-\lambda A)}{\lambda A} \times \exp(-\lambda F)$$

となる。以上より、受け入れ期間があり追跡期間が存在する場合の死亡の確率が計算することができ、これを P 群ならびに A 群の死亡確率の計算に適用した。

従って、後半に登録された症例の両群併せた期待死亡数は

$$LEN = \frac{12}{21} N \times \frac{1}{2} \times H_1 + \frac{12}{21} N \times \frac{1}{2} \times H_2$$

従って、より前半、後半併せた期待死亡数は FEN + LEN となる。

これと(1)式の D が等しいので、これを解いて、必要登録例数は全群で $N=544$ (1 群あたり $n=272$) となる。

5. 中間解析を伴う生存時間データの例数設計

この章では、中間解析を伴う生存時間データの例数設計を考える。

例 5-1

例 4-1 と同じ状況下で治験の途中で O'Brien-Fleming 型の中間解析を 5 回行う場合を考える。

4 章で算出した、 $\Delta = \frac{\sqrt{n_1^*} \times \log \lambda_1 / \lambda_2}{\sqrt{2}}$ を用いれば、2 群のハザード比 ($\log \lambda_1 / \lambda_2$) なら

びに Δ が求めれば、解析 1 回あたりの必要イベント数 $2 \times n_1^*$ が算出できる。解析回数は 5 回であるから、各解析でのイベント数が等しいという仮定をおき、後は 4 章で述べたように、 $5 \times 2 \times n_1^* = FEN + LEN$ を解けば必要症例数 N が求まる。このための SAS プログラムを以下に示す。なお、3 章の時と同様プログラムの効率や汎用性は全く考慮に入れず数式との対比に重点をおいてプログラムを書いた。お見苦しいところをご容赦願いたい。

(プログラム)

```
proc iml ;
  m={1 2 3 4 5}##{0} ;
  domain1=-m//m ;
  level=1-0.05 ;
  call seqscale(prob,cw,domain1,level) ;
  domain2=domain1*cw ;
  plevel=0.1 ;
```

```

call seqshift(prob,delta,domain2,plevel) ;
names={'delta'} ;
create calc from delta[colname=names] ;
append from delta ;
close calc ;
quit ;

data calc ;
set calc ;
lambda1=-log(1-0.156)/36 ;
lambda2=-log(1-0.078)/36 ;
ne=2*2*delta**2/(log(lambda1/lambda2))**2 ;
ne=5*ne ; /* neに中間解析の回数をかけると必要イベント数が算出される */
fe1=1-exp(-48*lambda1) ;
fe2=1-exp(-48*lambda2) ;
h1=1-(1-exp(-12*lambda1))/lambda1/12*exp(-lambda1*36) ;
h2=1-(1-exp(-12*lambda2))/lambda2/12*exp(-lambda2*36) ;
N=ne/(9/21/2*(fe1+fe2)+12/21/2*(h1+h2)) ;

run ;

proc print data=calc ;
run ;

```

(アウトプット)

delta	lambda1	lambda2	ne	fe1	fe2	h1	h2	N
1.46873	.004711188	.002255835	79.5537	0.20239	0.10262	0.17941	0.090367	558.504

(説明)

両群併せての必要イベント数はne=79.5537例である。両群併せての必要症例数はN=558.504例より、登録例数として一群あたり280例必要となることが分かる。

6. まとめ

本報告では、SAS/IML から提供されている逐次検定を行うサブルーチンを紹介し、臨床試験で用いられる群逐次検定との紐付けを行った。宇野(1998)で提供されたマクロは片側検定の対応がされてないこと、中間解析の間隔は等間隔であるが、今回、3章で紹介したサブルーチンでは、片側検定、不等間隔の中間解析にも対応可能である。ただし、今回紹介し

たサブルーチンは Wang and Tsiatis(1987)のアプローチに当てはめて使うことは容易であるが、Lan and DeMets(1983)のアルファ消費関数を用いたアプローチを適用することは困難であり、即ち、データ解析時のツールとして用いるにはかなり限定された状況でしか使用できない。従って、実際上の適用としては試験計画時の例数設計での場面が適しているのではないかと考えている。そこで、中間解析を伴う生存時間データの例数設計を例としてサブルーチンの適用を試みた。なお、4章で紹介した例数設計では、staggerd entry で、1部の症例はある一定の期間で観察を終了し、その他の症例は治験の終了時まで観察されるという例を扱った。これは企業が行う臨床試験ではよく遭遇するケースとは思われるが市販の例数設計用のソフトではこのような少々複雑な観察期間を扱えるものは存在しないと思われ、計算方法のロジックを理解できるように説明したつもりである。今回紹介した例では登録が一樣に進むこと、生存時間が指数分布に従うことなどの仮定がいくつか入っているが、生存時間解析の例数設計の基本的な考え方は、より複雑なケースにも応用可能である。

[参考文献]

- [1] Armitage, P., McPherson, C.K. and Rowe, B.C. (1969). Repeated significance tests on accumulating data. *Journal of the Royal Statistical Society* **A132**, 235-244
- [2] Lan, K.K. and DeMets, D.L. (1983). Discrete sequential boundaries for clinical trials. *Biometrika* **70**, 659-663
- [3] O'Brien, P.C. and Fleming, T. R. (1979). A multiple testing procedure for clinical trials. *Biometrics* **35**, 549-556
- [4] Pocock, S. J. (1977). Group sequential methods in the design and analysis of clinical trials. *Biometrika* **64**, 191-199
- [5] SAS Institute Inc (1995). *SAS/IML Software: Changes and Enhancements through Release 6.11*, Cary NC:SAS Institute Inc.
- [6] Shoenfeld, D.A. and Richter, J. R. (1982). Nomograms for calculating the number of patients needed for a clinical trial with survival as an endpoint. *Biometrics* **38**, 163-170
- [7] 宇野 一(1998). SAS/DATA ステップによる群逐次計画における棄却限界値及び必要症例数の算出, 第 17 回日本 SAS ユーザー会総会および研究発表論文集,77-102
- [8] Wang, S. K. and Tsiatis, A. A. (1987). Approximately optimal one-parameter boundaries for group sequential trials. *Biometrics* **43**, 193-199