

*Group Sequential Methods
with Applications to Clinical Trials*

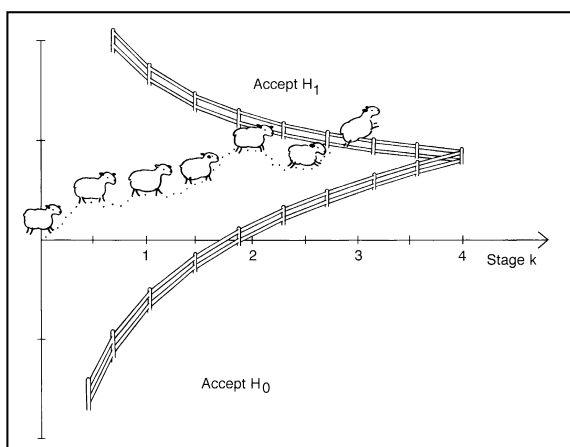
臨床試験における群逐次法：
理論と応用

Christopher Jennison & Bruce W. Turnbull

森川敏彦・山中竹春 訳

CRC

GROUP SEQUENTIAL METHODS with APPLICATIONS to CLINICAL TRIALS



Christopher Jennison

Professor of Statistics
Department of Mathematical Sciences
University of Bath, UK.

and

Bruce W. Turnbull

Professor of Statistics
School of Operations Research & Industrial Engineering
Department of Statistical Science
Cornell University, USA.

CHAPMAN & HALL/CRC

Boca Raton London New York Washington, D.C.

訳者まえがき

本書は Christopher Jennison (英バース大学) および Bruce W. Turnbull (米コーネル大学) による Group Sequential Methods with Applications to Clinical Trials (Chapman & Hall/CRC, 2000) の全訳である。1 群のまとまったデータが得られるたびに中間解析を実施して研究の中止・継続を判断する方法のことを「群」逐次法と呼ぶが、原著は臨床試験への応用を念頭に群逐次法の理論と実践を解説した世界的に定評のあるテキストである。

原著者の Jennison-Turnbull 両博士は群逐次法の理論的發展を大きく牽引した研究者である。本書の内容に関するロードマップは 1 章に述べられているので詳細は割愛するが、Pocock 検定や O'Brien-Fleming 検定、Wang-Tsiatis 検定族、 α 消費関数アプローチ、早期中止後の統計的推測、繰り返し信頼区間、確率打ち切りといった基本的事項から、同定性検定、2 値データ、生存時間データ、局外パラメータに基づく被験者数の再計算、多重エンドポイント、3 アーム以上の比較などの実用的なトピックスがそれぞれ 1 つの章を割いて丁寧に解説されている。原著に特徴的なのは統一したフレームワークに基づく定式化であろう。正規反応の平均の比較といった単純な設定からかなり一般的な問題に至るまで、群逐次デザインから得られる検定統計量の列がデータタイプや統計モデルの違いを超えて正確または漸近的にある同時分布に従い、したがって、原理的に同一の解析アプローチが可能であることが見通しよく説明されている。群逐次法は確率過程論に基づいて説明されることが多いが、その類いの予備知識は一切不要としながらも数学的厳密性は保ちつつ緻密に議論が展開されている。さらに具体的な応用例や数値計算例を随所に挿入することで実際の試験のための指南役として活用することができるような配慮もなされている。優秀な理論家でありながら、実務経験の豊富な両博士ならではの見識を伺うことができる。

原著は 400 ページに近い大冊であるが、分担訳ではなく、訳者 2 名の間で議論と相互確認を繰り返しながら充実した訳となるように努めた。上述したように原著は総じて明快なテキストであるが、理論書ゆえの緻密な記載や言語構造の違いなどから原著に逐一忠実に訳すとかえって意味がとりにくいと予想される箇所が少なくなく、悩んだ末に日本語としての読みやすさを優先させる方針をとることにした。論旨や記載が不明確な箇所についてはときに原著者共に議

論を交わしながら翻訳に反映させた。結果として意識の部分が多いが、原著における記載や論理の筋は一切損なわないように慎重な配慮をしたつもりである。この点については読者諸氏のご意見、ご批判を仰ぎたい。訳者の浅学非才からこれらの対応に時間を要し、脱稿に至るまで当初の予定を大幅に越える年月が経過してしまった。株式会社シーエーシーさんに多大なご迷惑をお掛けしたことを深くお詫びしたい。

誤植についてはウェブ上(1.4 節の URL 参照)に若干公開されているが、実際にはこれを遥かに上回る数の誤植が発見されたため、原著者に確認をとった上で修正を施した。煩雑さを避ける意味で一部を除き、断りなく修正している。また、式の展開をはじめとして一読では理解しにくいと思われる箇所を中心に脚注を、あるいは説明が長くなる場合は訳者補遺を巻末に付したので、本書を読み進める上での参考として頂ければ幸いである。

中間解析の実施にあたっては試験の完全性を損ないという問題が色々と生じるため、統計家は中間解析の理論と性質をよく理解して様々な実践的狀況に対応することが求められる。その観点から本書が中間解析に携わる本邦の企業・規制当局・アカデミアの統計家に広く活用されることを願っている。

Jennison-Turnbull 両博士には訳者からの不定期な問い合わせに対して常に丁寧なご対応を頂いたばかりか、日本語版への序文もご寄稿頂いた。両博士に謹んで感謝の意を表したい。上坂浩之博士(大阪大学)と杉本知之博士(弘前大学)は大変ご多忙の中、拙い初稿を入念に校閲して下さい、数多くの貴重なご指摘を頂戴した。ここに記して感謝する。最後に今回の出版を快くお引き受け下さり、かつ翻訳に関わる数々のわがままをお許し頂いた株式会社シーエーシー執行役員の高橋久氏、原著版元との交渉から出版に至るまで何かとお世話になった同社西日本医薬 BTO 部の藤本大祐氏に心から御礼を申し上げる。

2011 年 9 月 大阪および福岡にて

森川敏彦 山中竹春

日本語版への序文

このたび、我々の著書の日本語版が上梓されたことを大変嬉しく思います。翻訳を担われた森川敏彦・山中竹春の両博士に心から感謝を申し上げます。両博士は翻訳という骨の折れる作業のみならず、原著の数式や論旨を丁寧に追って下さり、我々が見落としていた数多くの誤りを発見して頂きました。日本語版ではこれらの誤りがすべて訂正されています。

臨床試験の統計的方法論は過去 10 年間で大きな発展を遂げました。しかしながら、本書の内容が中間解析の核心部分となり、最先端(state of the art)のアプローチであることに今も変わりはありません。試験途中の意思決定における群逐次法の重要性は適応型デザインに関する最近の FDA ガイダンス(2010 年 2 月)の中でも明記されています。近年は統計ソフトも発達し、比較的容易に群逐次デザインが計画できるようになりました。

生物統計学の研究は日本の大学や製薬企業でも盛んですが、幸いにも著者の 1 人(CJ)は 2009 年 3 月に東京で開催された第 6 回多重比較手法国際会議(International Conference on Multiple Comparison Procedures)に参加し、現地にてその活発な研究環境を味わう機会に恵まれました。この来日の際には皇居周辺の美しい桜を鑑賞できたことも印象に残っています。

原著はすでに米国や欧州の読者から高い評価を頂いていますが、この翻訳書が日本の読者にとっても等しく有用なものになることを願ってやみません。

Christopher Jennison
Bruce Turnbull

原著序文

我々の好きな本の一つに米国人作家 Bill Bryson の『Notes from a small island*1』がある。この本の中に、スコットランド最北端の町 Thurso の「地元有力者」John Sinclair 卿が statistics という語を生み出したことについて述べたくだりがある。1834 年のことだったそうだが「しかし、田舎町発のこの一大事件の波紋は長くは続かなかった」と述べられている*2。語源辞典を調べてみれば、このゲール語よりもさらに古いルーツに辿り着けるかもしれないが、それはともかくとして、statistics という語にはこのように長い歴史がある。同様に本書執筆までの歴史も長く時間のかかるもので、世界のあちこちで執筆活動をするようになってしまった。筆者の 1 人(BWT)は本書の起源を 1971 年まで遡らなければならない。スタンフォード大学医学部に新米助手として勤務し、ある重要な臨床試験のデータ管理を任されていた頃である。その試験の主任研究者は有名な外科医であったが、彼は新たな再発や死亡を確認するとその事実を伝えるために BWT のオフィスへ立ち寄り、「これでもう結果を報告してよいか？」という質問を投げかけるのであった。このようなやり取りが数日おきに繰り返され、試験はたちまち逐次的な性質を帯びていくことになった。本書の大半はニューヨーク州北部の Ithaca(イサカ)と英国 Bath(バース)の Claverton Down において執筆されているが、しかし、それ以外の土地——ウィルトシャー州 Bradford on Avon からアリゾナ州 Tucson, ニューサウスウェールズ州 North Ryde からテネシー州 Oak Ridge——でまとめられた部分も随所にある。

かわいらしい羊の表紙イラストを作成してくれた Alan Peacock, 表 9.3 の計算を手伝ってくれた Stuart Barber, 10.5 節の例の作成を手伝ってくれた Catalin

*1 Bill Bryson, *Notes from a Small Island*, Avon Books, New York, 1995

*2 (訳注) 原著序文の冒頭では、Bill Bryson 著『*Notes from a Small Island*』の 292 ページにある「*The big event in Thurso, according to civic records, was in 1834 when Sir John Sinclair, a local worthy, coined the term statistics in the town, though things have calmed down pretty considerably since.*」という一文が材料とされている。John Sinclair (1754-1835) はスコットランド農業省の長官を務めるなど多方面で足跡を残した人物であり、なかでも 20 巻に及ぶ『*Statistical Account of Scotland*』の出版を主導した功績は広く知られている。1834 年の Statistical Society of London (英国 Royal Statistical Society の前身) 設立当時の最古参の会員でもあった Sinclair が統計学史上の重要人物であることは明らかで、一介の地元有力者(local worthy)に留まるものではない。原著序文の中で「地元有力者」の部分をあえて強調表示しているのはこのような事情を汲んでのことと思われる。

Starica, そして草稿段階の原稿に貴重な助言を寄せてくれた Rebecca Betensky, John Bryant, Jonathan Denne, Patricia Grambsch の諸氏に心からの御礼を申し上げる. 米国 National Cancer Institute (R01 CA-66218), 英国 Science and Engineering Research Council および Engineering and Physical Sciences Research Council からは研究費の助成を受けた.

これまで強い影響を与えてくれた方々にもあらためて感謝の意を捧げたい. コーネル大学時代に Bob Bechhofer から学問としての統計学の魅力を学び, Jack Kiefer から大学院セミナーを通じて逐次解析の手ほどきを受けることができたのは幸いであった. また, Rupert Miller と Toby Mitchell は現実の生物医学問題の解決のために統計学の理論と実践をどう融合させるべきかを具体例でもって BWT に伝えてくれた. ここにお名前を挙げさせて頂いた 4 名は全員がすでに逝去されている. 皆, いよいよこれからというときであったことを思うと残念でならない. アリゾナ大学がんセンターの Larry Clark 博士にもこの場を借りて御礼を申し上げる. 博士は臨床試験や疫学研究の日々の実践を通じ, 長年にわたって統計学上の諸問題を見出され, 解決を BWT に促してこられた. いずれも重要かつ困難で意欲的に取り組める問題であった. 最後になるが, 我々を粘り強く支えてくれた家族の Debbie, Marty, Kenneth, Douglas に最大限の感謝の言葉を述べたいと思う. 本書がついに完成に至り, そのことを報告したときには家族中で本当に驚いてくれた. ただし, この一大事件の波紋はここでも長くは続かなかったのだが.

1999 年 6 月 バースおよびイサカにて

Christopher Jennison
Bruce Turnbull

目 次

訳者まえがき	v
日本語版への序文	vii
原著序文	ix
略語・記号集	xv
第1章 はじめに	1
1.1 本書について	1
1.2 なぜ逐次的な方法が必要なのか？	4
1.3 逐次法および群逐次法の簡潔な歴史	7
1.4 いくつかの例	15
1.5 各章の構成：本書のロードマップ	20
1.6 文献と注釈	24
第2章 両側検定：基本的な群逐次法	27
2.1 正規反応の平均の差に関する両側検定	27
2.2 固定標本検定	28
2.3 群逐次検定	30
2.4 Pocock 検定	32
2.5 O'Brien-Fleming 検定	37
2.6 Pocock 検定と O'Brien-Fleming 検定の性質	40
2.7 その他の検定	49
2.8 ま と め	57
第3章 両側検定：統一的定式化と様々な応用	61
3.1 統一的定式化	61
3.2 情報列が等間隔な場合	66
3.3 情報列が等間隔でない場合	71
3.4 正規線形モデル	76
3.5 一般的なパラメトリックモデル	86
3.6 2値データ：割合に関する群逐次検定	91

3.7	生存時間データ：群逐次ログランク検定	95
3.8	両側群逐次 t 検定	98
第 4 章	片側検定	107
4.1	はじめに	107
4.2	べき型片側検定族	109
4.3	情報列が等間隔でない場合	118
4.4	片側群逐次 t 検定	122
4.5	Whitehead の 3 角検定	128
第 5 章	両側検定：早期の帰無仮説の受容を許す場合	137
5.1	はじめに	137
5.2	べき型両側インナーウェッジ検定族	139
5.3	Whitehead の 2 重 3 角検定	150
第 6 章	同等性検定	155
6.1	はじめに	155
6.2	片側同等性検定	157
6.3	両側同等性検定：バイオアベイラビリティの比較	157
6.4	個体内生物学的同等性：割合に関する片側検定	170
6.5	文献と注釈	173
第 7 章	柔軟なモニタリング：過誤消費アプローチ	175
7.1	予測不可能な情報列	175
7.2	両側検定	177
7.3	片側検定	194
7.4	データに依存した解析時期の決定	201
7.5	過誤消費検定に関する数値計算	204
第 8 章	群逐次検定後の解析	207
8.1	はじめに	207
8.2	分布論	208
8.3	点推定	212
8.4	P 値	217
8.5	信頼区間	220
第 9 章	繰り返し信頼区間	229
9.1	はじめに	229
9.2	例：正規反応の平均の差	234
9.3	派生検定：RCI を用いた早期中止の決定	236
9.4	繰り返し P 値	245

9.5 考察	246
第10章 確率打ち切り	249
10.1 はじめに	249
10.2 条件付き検出力アプローチ	250
10.3 予測検出力アプローチ	256
10.4 パラメータによらないアプローチ	259
10.5 事例研究：生存時間データ	262
10.6 文献と注釈	266
第11章 群逐次法のための一般分布論	269
11.1 はじめに	269
11.2 パラメータベクトルの列に関する標準同時分布	270
11.3 正規線形モデル	271
11.4 分散未知の正規線形モデル：群逐次 t 検定	274
11.5 例：正確な1標本群逐次 t 検定	276
11.6 一般的なパラメトリックモデル：一般化線形モデル	279
11.7 文献と注釈	284
第12章 2値データ	287
12.1 1標本問題	287
12.2 2標本問題	298
12.3 オッズ比および多層 2×2 表	307
12.4 マッチドペアによるケース・コントロール研究	311
12.5 ロジスティック回帰：共変量の調整	311
12.6 生存時間解析との関連	313
第13章 生存時間データ	315
13.1 はじめに	315
13.2 ログランク検定	316
13.3 層別ログランク検定	317
13.4 コックス回帰：共変量の調整	319
13.5 ハザード比に対する繰り返し信頼区間	322
13.6 例：中咽頭がんの臨床試験	325
13.7 生存率と分位点	333
13.8 文献と注釈	336
第14章 内部パイロット試験：被験者数の再計算	339
14.1 内部パイロットの果たす役割	339
14.2 固定標本検定における被験者数の再計算	342
14.3 群逐次検定における被験者数の再計算	357

第 15 章 多重エンドポイント	363
15.1 はじめに	363
15.2 ポンフェロー二手順	365
15.3 群逐次 Hotelling 検定	368
15.4 群逐次 O'Brien 検定	372
15.5 その他の包括的統計量に基づく検定	378
15.6 周辺規準に基づく検定	379
15.7 文献と注釈	383
第 16 章 3 種類以上の治療を比較する試験	385
16.1 はじめに	385
16.2 包括的検定	385
16.3 対比較のモニタリング	390
16.4 文献と注釈	394
第 17 章 適応型割り付け	397
17.1 多段階適応型デザイン	397
17.2 時間トレンドを含んだ多段階適応型デザイン	403
17.3 多段階適応型デザインの妥当性	405
17.4 文献と注釈	408
第 18 章 ベイズ流アプローチ	411
18.1 ベイズ流パラダイム	411
18.2 中止規則	413
18.3 事前分布の選択	417
18.4 考 察	420
第 19 章 群逐次法のための数値計算	423
19.1 はじめに	423
19.2 基本的な計算	424
19.3 過誤率と被験者数分布	430
19.4 過誤消費関数を用いる検定	433
19.5 群逐次検定後の解析	435
19.6 数値計算の更なる応用	438
19.7 統計ソフト	442
訳 者 補 遺	445
参 考 文 献	461
索 引	488
訳 者 紹 介	494

略語・記号集

ASN	average sample number	平均被験者数
DSMB	data and safety monitoring board	データ・安全性モニタリング委員会
ITN	inferior treatment number	劣った方の治療に割り付けられる期待症例数
FWER	family-wise error rate	仮説族あたりの過誤率
LD	Lan & DeMets	Lan-DeMets
LR	likelihood ratio	尤度比
LSD	least significant difference	最小有意差
LSE	least squares estimator	最小 2 乗推定量
MLE	maximum likelihood estimator	最尤推定量
OBF	O'Brien & Fleming	O'Brien-Fleming
OS	one-sided	片側 (7.3 節参照)
RBO	relative betting odds	相対賭けオッズ
RCI	repeated confidence interval	繰り返し信頼区間
SCP	stochastically curtailed procedure	確率打ち切り手順
SPRT	sequential probability ratio test	逐次確率比検定
\mathcal{I}_k	第 k 解析における情報量	
\mathcal{I}_{max}	群逐次検定の最大情報量	
$\mathcal{I}_{f,2}$	両側固定標本検定の必要情報量	
$\mathcal{I}_{f,1}$	片側固定標本検定の必要情報量	
I_k	第 k 解析における繰り返し信頼区間(書体の違いにより情報量と区別する)	
J	アーム数 (3 アーム以上の試験)	
K	群逐次検定の最大解析数	
R	群逐次検定の最大情報量と固定標本検定の情報量の比 $\mathcal{I}_{max}/\mathcal{I}_f$ (α, β は同じ条件とする)	
$\mathbf{x}^T, \mathbf{X}^T$	ベクトル \mathbf{x} , 行列 \mathbf{X} の転置	
$t_{\nu, \alpha}$	自由度 ν のスチューデントの t 検定の上側 α 点	
z_α	標準正規分布の上側 α 点. すなわち, $z_\alpha = \Phi^{-1}(1 - \alpha)$	
α	第 I 種過誤率	
β	第 II 種過誤率	
β (太字)	回帰係数ベクトル (ベクトルおよび行列は太字で表す)	
Δ	べき型検定族の形状パラメータ	
Φ	標準正規累積分布関数	
ϕ	標準正規密度関数	

第 1 章

はじめに

1.1 本書について

本書は中間解析に用いられる**群逐次デザイン**の統計的方法論に関するテキストである。主に**臨床試験**の中間解析を念頭においているが、実験・調査を段階的に実施する研究領域や定期的に中間解析を実施する長期の研究であれば、ほとんど必ずと言っていいほど群逐次デザインの適用を考えられる場面がある。たとえば、医学分野では動物実験や疫学研究で群逐次デザインが活用されている。品質管理の抜き取り検査や信頼性研究の寿命試験ではすでに確立されたその分野の逐次デザインないし群逐次デザインが存在する。工業製品の信頼性を確認するための実地試験、更生プログラム施行後の仮釈放者の再犯率を評価する研究、職業コホートに対する慢性疾患の長期追跡調査などでも中間解析の実施は不可欠である。1.4 節で群逐次デザインの幅広い用途について説明する。

本書では臨床試験に関連した例を多く用いる。この分野に不案内な読者もいると思われるので、最初に基本的な用語をいくつか説明しておこう。一般に新しい治療が医療界に受け入れられるようになるためには動物実験を経て、その後、ヒトを対象とした数段階の臨床試験を実施する必要がある。まず、**第 I 相試験**として臨床薬理や毒性に関する検討が行われる。第 I 相試験の目的は、許容されないほどの有害反応は回避できる適切な用量レベルを探索することであり、通常は 20 例から 80 例の健常人ボランティアが対象とされる。健常人ボランティアは製薬企業の従業員や医学生から構成されることがある*1。第 I 相試験の終了後は有効性と安全性の両面のパイロット的な評価を目的に、

*1 (訳注) 以前は製薬企業の従業員や医学生がボランティアとして参加することもあったが、リスクの強制的につながりかねず、現行ではこのような制度や慣習はないはずである。GCP 運用通知(「医薬品の臨床試験の実施の基準に関する省令」の運用について)の第 44 条の解説も参照せよ。