

臨床比較試験における統計的逐次検定

神戸商船大学 応用数学教室 杉村 正彦

塩野義製薬(株) 解析センター 後藤 昌司

塩野義製薬(株) 解析センター 浅野長一郎

§ 1. 序

逐次検定的方式は *sampling inspection* の問題として、
Dodge-Romig らによる一回抜取、二回抜取、多回抜取検査
の延長として考按され、やがてこの方式が *Statistical Research Group (Columbia U.)*, Barnard や Wald らの論文・成書
に体系化されたことは周知である。

第2次大戦後、臨床医学の分野で異った治療法の科学的な
比較が普及し、ここに試験対象となる患者が期待的に少くて
すむなどの方法論的関心から統計的逐次検定の適用が考えら
れてきた。これらの種々の工夫は Sobel, Brass, Armitage,
Anderson, Schneiderman, Anscombe, Wetherill, Ghosh
らに代表されてみられる。そこでは、処理 A , B に関して、
 $H_1: A > B$, $H_2: B > A$, $H_3: A \approx B$ の何れかを判定する
truncated された *closed type* や横軸に平行な *open type* な
ど、 ASN への配慮をこめて、実際的な判定を得る方式が主
に論じられた。また、その後 *sampling* 方式については、
Hall や Read は *Partial SPRT* を論じ、また、Anderson,

Choi-Clark は *delayed data* の SPRT を検討した。さらに、Jackson-Bradley, Appleby-Freund, Grimm らは、多変量の SPRT を論じてきている。(今回はこれらの方角の論議を割愛する。以下に論ずる方向が標題にも、と重要な意味をもつと思われるからである。)

さて、これらの逐次検定では適用の目的を臨床試験にしているといっても、なお倫理性の立場を積極的にとり入れたとは言い難い。ASN は可成り少くて結論づけられるといっても、QC の場で考えられるものと、それ程異なるものではない。Colton は、期待的に少ない患者数で処理法に関する判定を得ることは当然であるが、「臨床試験に際し、劣性処理をうける患者はなるべく少なくなければならない」という試験対象への倫理を強調した。彼は、試験対象のうける *loss* とともに、その判定でうける残りの患者の *loss* をも考慮に入れて定式化し、*minimum expected loss* の方式を提唱した。Sugimura-Goto-Asano はより広い母集団分布族の場合にこれを拡張した。これらの逐次検定法はむしろ確率過程の問題として、*open type* にしろ *closed type* にしろ Markov process の定式化による最適方式の設定問題と考えられている。事実、これまでの SPRT では、 $(\alpha, \beta, \theta_0, \theta_1)$ の設定が困難であり、また Wald の解析的な lemma も本質的なことではなかつ

にわけである。

このような臨床試験での逐次検定を確率過程における *two-armed bandit problem* とし、Zelen は、Robbins, Bradt-Johnson-Karlin, Feldman, Smith-Pyke, Samuel を引用しながら、*Play-the-Winner Rule* を導入した。彼は劣性処理をうける患者をできる限り少なくするのに PWR 方式が有効であり、とくに従来の *Vector-at-a-time* 方式よりも優れた場合の多いこと、また、ある種の *delayed data* には *Modified PWR* を提唱した。これを契機に、Sobel-Weiss, Kiefer-Weiss, Hoel, Nebenzahl-Sobel, Asano-Sugimura-Goto (準備中) らは、*untruncated* や *truncated type*, PWR と VT plan, *Inverse Sampling plan* および *Fixed Sample-Size plan* など種種おりまぜた比較研究を展開している。

上の PWR をめぐる一連の研究は、何れも 2 処理間の効果の優劣や同等を判定するのであるが、Zelen と同じ頃、Beckhoffer-Kiefer-Sobel (1968) は処理法が 2 以上の場合に生ずる *ranking* と *identification* の問題を考察した。一般に、逐次検定の方法は、*Sampling rule*・*Stopping rule*・*Decision rule* で規定されるが、これらは 2-player, 3-player のもとで論じられる。まず、Sobel-Weiss (1969) は 2-player の場合の PWR に *analogous* な 3-player binary comparison

として Drop-the-Loser rule を導入し、3 処理間の best なものの選択法を論じた。彼は DLR によって、 R_F (single-stage procedure of a fixed sample size), R_I (inverse sampling procedure), R_S (sequential sampling procedure), R_E (elimination procedure) をとりあげた。ついで、Beckhofer (1969) は、前の 3-player-tournament での判定が 3 者のうちの best な唯一つを選ぶのに対し、3 者のうちどれが best とも言えない場合があるとして、best なものまたは何れも同等という 4 つの判定を下しうる手順を、ranking と identification の立場から論じている。一方、Hoel (1971) は、Paulson-type の 2 処理効果間の逐次検定を、truncated と open について modified SPRT により劣性処理を除去していく方式で構成している。

次節以降では、上のような臨床試験における逐次選別の問題を、主として、Zelen 以降の PWR に関連した諸研究の総合的な報告と、あわせて著者らの若干の結果や今後の諸問題にふれる。

§ 2. Colton, 著者らの Procedures

標本サイズ固定計画では、 N 人の患者集団からの $2n$ 人について、その各 n 人に処理法 A と B を施し、得られた観測値をもとに、一定の優劣判定の規準によって選ばれた処理法を残

り $(N-2n)$ 人に施すというプランのもとで、Expected Loss

$$\begin{aligned} [E \text{ Loss}] &= C \delta \{n + (N-2n) - \Pr(\text{select inferior})\} \\ &= K_1 \delta N - K_2 [E \text{ Net Gain}] \end{aligned}$$

(但し, $\delta = d(\theta_A, \theta_B)$ は A, B の効果 θ_A, θ_B 間の差違を示す既知関数, C は比例定数)

の minimax 解および δ の先験分布のもとでベイズ解を与えた。また逐次計画では、2人の患者に A, B を施した vector-at-a-time での観測対を逐次的に取り、一定の停止・判定規則で選ばれた処理を残り $N-2E_\delta(n)$ に施すという方式の $E \text{ Loss}$ を固定計画と同様に考察した。更に同一条件下で最小 $E \text{ Loss}$ を比較して、逐次計画の固定計画に対する優位性を主張した。

Colton は反応変数が正規分布の場合、Asano-Sugimura-Goto は Koopman-Darmois family の場合および正規、2項分布の場合について詳細に考察した。Cornfield-Greenhouse [32] は固定計画での患者割当を δ の先験分布によることを唱え、Canner [32] は反応が dichotomous の場合を考察した。

§ 3. Zelen の Procedures

Zelen は臨床比較試験の倫理性を強調して、sampling 方式に Play-the-Winner Rule を提唱した。固定計画では Colton の Expected Net Gain の最大化を考察し、全実験中での success の期待回数が、 N, n が大きいときには、各処理法に等しい

固定標本数を割当てるという Colton 方式より多くなならないことを示し、最適な標本数の近似値が Colton での Maximin 解_{N/3}にほぼ等しいことを示した。逐次計画に関しては、Stopping Rule として「予め定めた自然数 f に対し、両処理法の failure 回数が共に f に達すれば実験を停止」し、 $E_f(N_A)/E_f(N_B) = g_B/g_A$ (但し、 g_i , N_i は処理法 i の failure の確率、実験停止までの患者数を表わす ($i=A, B$)) なることを示した。更に、delayed observation の場合の Modified PWR 方式を提案し、また両処理の success の確率 p_A , p_B が患者の症状別、年令別、性別や処理実施上の技術の上達などにより変化する場合の PWR 方式をも提唱した。

§ 4. 種々の PWR-および VTR-Procedures の比較

(イ) 一方の処理が他方より十分に優れて ($p_A - p_B = \Delta \geq \Delta^*$) いる時、その処理を選択する確率をある値以上にしたい ($P\{A\text{-selection}\} = P\{\text{Correct Selection}\} \geq P^*$)、同時に、(ロ) 劣性処理を受ける平均患者数 (すなわち、Expected Loss) を可能な限り少なくしたいという「2つの2項母集団間の優劣判定の逐次手順」に関し、Sobel, Weiss, Hoel, Kiefer, Nebenzahl および著者らの研究経過について述べる。

N を最終判定規則を下すまでに必要な患者数; N_i , S_i , F_i , p_i をそれぞれ、判定までに処理 i を受けた患者数、その中の

success 回数, failure 回数, 母 success 確率を表わすものとする. ($i = A, B$, $p_A > p_B$ と仮定する.) この時, Expected Loss を

$$[E \text{ Loss}] = (p_A - p_B) \cdot E(N_B | \text{Procedure } R)$$

と定義し, 表 1 で示した種々の Procedures を考える.

表 1. Procedures R

Plan	Terminal Decision Rule	Sampling 方式	PWR	VTR
Un-truncated Plan	$S_i - S_j = r$ となれば 実験を停止して、 処理 i をベターと判定。 (S-方式と名づける。)		$R_s [55]$	$R'_s [55]$
	一方の処理 i に対して、 $S_i = r$ となれば 実験を停止して、処理 i をベターと判定。 (I-方式と名づける。)		$R_I [56]$	$R'_I [56]$
	一方の処理 i に対して、 $S_i + F_j = r$ となれば 実験を停止して、処理 i をベターと判定。 (H-方式と名づける。)		$R_H [41]$	—
Truncated Plan	D-方式であるが、最大標本数 N_0 までに 判定がつかなければ $A = B$ と判定する。		—	$R'_{Ts} [46]$ (R''_{Ts})
	I-方式		$R_{TI} [57]$	$R'_{TI} [57]$
Fixed Sample Size Plan	$S_i > S_j$ なら処理 i をベターと判定し、 $S_i = S_j$ なら確率 $1/2$ でいずれか一方を ベターと判定する。		$R_F [47]$	$R'_F [47]$
(i, j = A, B, i ≠ j ; r は上記(1)の条件, すなわち (Δ*, P*)-条件を満たす与えられた自然数)				

(1) R_s と R'_s との比較 : (イ) p_A, p_B の先験情報が知れ、 $p_A > 0.75, p_B > 0.75$ であれば $Loss, E(N_B)$ と平均患者数 $E(N)$ の観点から R_s の方が R'_s よりよいが、 p_A, p_B が共に小さいなら逆に R'_s の方が R_s よりよい。(ロ) p_A, p_B の先験情報が何もない時は予備的推定を必要とするが、この場合の R_s, R'_s 間の優劣はまだ不明で、どのような性質の先験情報が各 Procedure 間の優劣を変えるかについても今後の問題がある。

(2) R_I と R'_I との比較 : (イ) $Pr\{\text{Correct Selection}\}$ は両 Procedures で等しい。また、 $E(N_B)$ は p_A, p_B の値に関して一様に R_I の方が
 (ロ) r が大であれば R'_I よりよい。

(3) (R_I, R'_I) と (R_s, R'_s) との比較 : $E(N)$ は、 $\Delta^* \neq 0$ なら後者が前者よりよいが、 $p^* \neq 1$ の時の特定の Δ^* に対してはその優劣が逆になる。すなわち、両者の優劣には (Δ^*, p^*) に関して一様に結果づけできない。

(4) R_H と R_I との比較 : $P^* = 0.95, \Delta = \Delta^* = 0.2$ の場合の数値結果として、(イ) $(p_A, p_B) = (0.2, 0) \sim (0.6, 0.4)$ なら $E(N_B), E(N)$ から R_H の方が R_I よりもよく、しかも (p_A, p_B) の値が小さい程 R_H が非常に良くなる。(ロ) $(p_A, p_B) = (0.7, 0.5) \sim (1, 0.8)$ なら、 $E(N_B)$ に関し逆に R_I の方が R_H よりもよく、 $E(N)$ では R_I の方が R_H よりもややよい。(ハ) (p_A, p_B) の値が小から大に変化するに伴って、 $E(N)$ は、 R_H の場合平坦に減少するが、 R_I は R_H より大きい値か

ら急激に小さくなり、遂に R_H よりやや小さくなる。従って、
 (p_A, p_B) に関する先験情報が何もない場合には、 $E(N)$ の観点
 から R_H を採用した方が安全であろう。

(5) R'_{Ts} と R'_S との比較 : N_0 が大であれば、(1) $\Pr\{\text{Correct Selection}\}$ は R'_S がよいが、(2) 平均観測対数は殆ど変わらず、
 結局(1)、(2)を総合すれば N_0 が大のときには R'_S が R'_{Ts} よりも望
 ましい。しかし、 R'_{Ts} は $A=B$ という判定を下すことができる
 という特長を持っている。

(6) R_F と R'_F との比較 : $\Pr\{\text{Correct Selection}\}$ から、 $E(N_B)$ か
 らも R_F が R'_F よりすぐれている。

(7) R_{Ts}'' : Asano-Sugimura-Goto での Sampling 方式は
 VTR で、最終判定規則は D-方式で、Truncated Plan のもと
 での、Colton 方式の Minimum Expected Loss に対する最適計
 画を求める問題を解決し、種々の数値例を示している。

§ 5. Delayed Data の場合の 2 段逐次判定手順

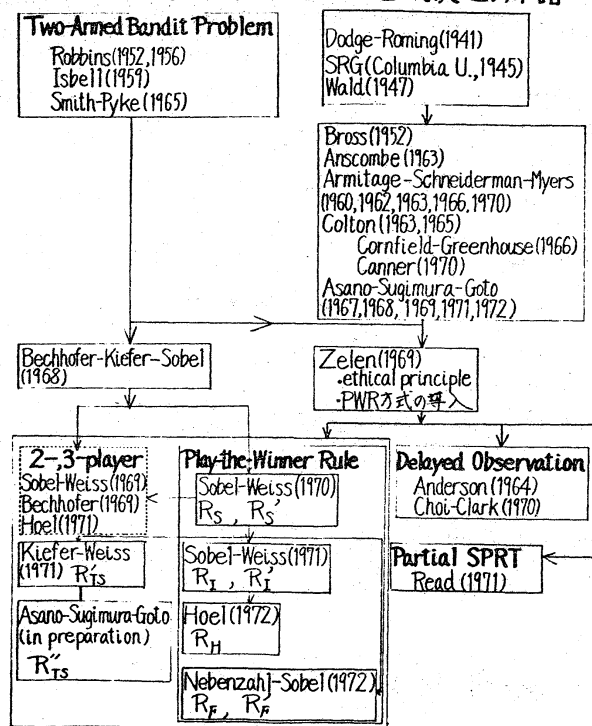
反応変数が 2 項型の場合
 Choi-Clark [33] は、delayed data をテストに組みこんだ
 2 段構成の SPRT Procedure を構成し、 $\alpha_k = \Pr\{\text{仮の判定で } H_k \text{ を棄却} \mid H_k\}$, $\alpha_{d,k} = \Pr\{\text{最終判定で } H_k \text{ を棄却} \mid H_k\}$, $\gamma_k = \Pr\{H_k \text{ を棄却から採択へと判定を変える} \mid H_k\}$, $\gamma'_k = \Pr\{H_k \text{ を採択から棄却へと判定を変える} \mid H_k\}$, 仮の判定による $OC: L(p)$, 最
 終判定による $OC: L_d(p)$ の概念を導入して、 $\alpha_{d,k}$, γ_k , OC

の観点から Procedure の有用性を主張した。また、Anderson [1] は反応変数が正規型の場合について論じた。

§ 6. Partial SPRT, [50]

パラメータ θ に関する、 $H_1: \theta = \theta_1$ に対する $H_0: \theta = \theta_0$ の SPRT では、例えば $\theta = (\theta_1 + \theta_0)/2$ における $E(N|\theta)$ は同じ α, β をもった固定計画での標本数より大きくなってしまう。Partial SPRT はこのような欠点を除く。また、サンプリングと試験の相対的な経済性によって Partial SPRT の適用が必要な場合がある。

臨床比較試験における統計的逐次検定の系譜



References

- 1) Anderson, T.W., "Sequential analysis with delayed observation," JASA, 59, 1006-1015(1964).
- 2) Anscombe, F.J., "Sequential medical trials," JASA, 58, 365-383(1963).
- 3) Appleby, R.H. and Freund, R.J., "An empirical evaluation of multivariate sequential procedures for testing means," AMS, 33, 1413-1420(1962).
- 4) Armitage, P., "Some sequential tests of student's hypothesis", J.R.S.S. Suppl. 9, 250-63(1947).
- 5) Armitage, P., "Restricted sequential procedures," Biometrika, 44, 9-26 (1957).
- 6) Armitage, P., "Numerical studies in the sequential estimation of binomial parameter," Biometrika, 45, 1-15(1958).
- 7) Armitage, P., "The comparison of survival curves" JRSS, A, 122, 279-293 (1959).
- 8) Armitage, P., "Sequential Medical Trials" Springfield(1960).
- 9) Armitage, P. and Schneiderman, M.A., "A family of closed sequential procedures," Biometrika, 49, 41-56(1962).
- 10) Armitage, P. and Schneiderman, M.A., "Closed sequential t-tests". Biometrika, 49, 359-366(1962).
- 11) Armitage, P., "Sequential medical trials: Some comments on F. J. Anscombe's paper," JASA, 58, 384-387(1963).
- 12) Armitage, P., "Some developments in the theory and practice of sequential medical trials," 791-804, 5th B. Symp.
- 13) Armitage, P., Myers, M.H. and Schneiderman, M.A., "Boundaries for closed sequential t-test plans," Biometrika, 53, 431-437(1966).
- 14) Armitage, P. and Schneiderman, M.A., "Corrections to "A family of closed sequential procedures", Biometrika, 56, 457(1969).
- 15) Armitage, P., "The combination of assay results," Biometrika, 57, 665-666 (1970).
- 16) Sugimura, M., "Optimum Designs for Selecting One of Two Medical Treatments, Fixed Sample Size Plan 1", Mem. Fac. Gen. Educ. Kumamoto Univ., Ser. Nat. Sci., 2, 1-7(1967).
- 17) Sugimura, M., "Optimum Designs for Selecting One of Two Medical Treatments, Fixed Sample Size Plan 2", Kumamoto J. Sci., Ser A, 7, No. 4, 87-94(1967).
- 18) Sugimura, M. and Asano, c., "Optimum Designs for Selecting One of Two Medical Treatments, Fixed Sample Size Plan 3", Kumamoto J. Sci., Ser. A, 7, No. 4, 95-102(1967).

- 19) Sugimura, M., "Optimum Designs for Selecting One of Two Medical Treatments, Fixed Sample Size Plan 4.1", J. Kumamoto Women's Univ., 19, 67-71 (1967).
- 20) Sugimura, M., "Optimum Designs for Selecting One of Two Medical Treatments, Fixed Sample Size Plan 5", J. Kumamoto Women's Univ., 20, 58-63 (1968).
- 21) Sugimura, M. and Asano, C., "Optimum Designs for Selecting One of Two Medical Treatments, Sequential Plan 1", Bull. Math. Stat., 12, No. 3-4, 1-9 (1967).
- 22) Sugimura, M., "Optimum Designs for Selecting One of Two Treatments, Sequential Plan 2", Kumamoto J. Sci., Ser. A, 8, No. 1, 11-19 (1967).
- 23) Sugimura, M. and Asano, C., "Optimum Designs for Selecting One of Two Treatments Fixed Sample Size Plan 6 and Sequential Plan 3", Kumamoto J. Sci., Ser. A, 8, No. 1, 21-57 (1967).
- 24) Sugimura, M., "Optimum Designs for Selecting One of Two Treatments, General Considerations", Kumamoto J. Sci., Ser. A, 8, No. 2, 63-83 (1967).
- 25) Sugimura, M., "Optimum Designs for Selecting One of Two Treatments, Fixed Sample Size Plan and Sequential Plan in a Population with One Parameter Exponential Distribution", Bull. Math. Stat., 13, No. 1-2, 1-23 (1968).
- 26) Sugimura, M., Goto, M. and Asano, C., "Optimum Sequential Designs Based on Markov Chains for Selecting One of Two Clinical Treatments", Mem. Fac. Engrg. Kumamoto Univ., XV, No. 3, 1-16 (1969).
- 27) Sugimura, M., Goto, M. and Asano, C., "Some Sequential Designs Based on Markov Chains for Selecting One Two Medical Treatments", J. Kobe Univ. of Mercantile Marine, 19, 159-172 (1971).
- 28) Goto, M., Sugimura, M. and Asano, C., "Numerical tables of optimum sequential designs based on Markov chains for selecting one of two medical treatments", Bull Math. Statist., 14, No. 3-4, 27-56 (1971).
- 29) Bechhofer, R.E., Kiefer, J. and Sobel, M., "Sequential Identification and Ranking Procedures", U. of Chicago Press (1968).
- 30) Bechhofer, R.E. "On ranking the players in a 3-player tournament", In 'Non-parametric Techniques in Statistical Inference' (M.L.Puri, ed.) 545-59, Cambridge U. Press (1970).
- 31) Bross, I., "Sequential medical plans", Biometrics, 8, 188-205 (1952).
- 32) Canner, F.L., "Selecting one of two treatments when the responses are dichotomous", JASA, 65, 293-306 (1970).
- 33) Choi, S.C. and Clark, V.A., "Sequential decision for a binomial parameter with delayed observations", Biometrics, 26, 411-420 (1970).
- 34) Colton, T., "A model for selecting one of two medical treatments", JASA, 58, 388-400 (1963).

- 35) Colton, T., "A two-stage model for selecting one of two treatments", *Biometrics*, 21, 169-180 (1965).
- 36) Cornfield, J. and Greenhouse, S.W., "On certain aspects of sequential clinical trials", 5th Berkeley Symp., 813-29 (1966).
- 37) Dodge, H.F. and Romig, H.G., "Single Sampling and Double Sampling Inspection Table", *J. Bell Sys. Tech.*, 20-1-61 (1941).
- 38) Ghosh, B.K., "Sequential Tests of statistical Hypotheses", Addison-Wesley (1970).
- 39) Grimm, H., "Die Verwendung der mehrdimensionalen Sequenzanalyse beim Siebtestverfahren (Screening)", *Biomet. Zeit.*, 10, 1-18 (1968).
- 40) Hoel, D.G., "A method for the construction of sequential selection procedures", *AMS*, 42, 630-642 (1971).
- 41) Hoel, D.G., "An inverse stopping rule for play-the-winner sampling", *JASA*, 67, 148-151 (1972).
- 42) Isbell, J., *AMS*, 30, 606-10 (1959).
- 43) Jackson, J.E. and Bradley, R.A. *Technometrics*, 3, 519-34 (1961).
- 44) Jackson, J.E. and Bradley, R.A., "Sequential X - and T - tests", *AMS*, 32, 1063-1077 (1961).
- 45) Jackson, J.E. and Bradley, R.A., "Sequential multivariate procedures for means with quality control applications", *Multivariate analysis* (Edited by Krishnaiah), Academic Press (1966).
- 46) Kiefer, J.E. and Weiss, G.H., "A truncated test for choosing the better of two binomial populations", *JASA*, 66, 867-871 (1971).
- 47) Nebenzahl, E. and Sobel, M. "Play-the-winner sampling for a fixed sample size binomial selection problem", *Biometrika*, 1, 1-8 (1972).
- 48) Paulson, E., "A sequential procedure for selecting the population with the largest mean from k normal populations", *A.M.S.* 35, 174-80 (1964).
- 49) Paulson, E., "Sequential estimation and closed sequential decision procedures", *A.M.S.* 35, 1048-58 (1964).
- 50) Read, C.B., "The partial sequential probability ratio test", *JASA*, 66, 646-650 (1971).
- 51) Robbins, H. "Some aspects of the sequential design of experiments", *Bull. Am. Math. Soc.* 58, 527-35 (1952).
- 52) Robbins, H., "A sequential decision problem with a finite memory", *Proc. Nat. Acad. Soc. USA* 42, 920-23 (1956).
- 53) Smith, C.V. and Pyke, R., "The Robbins-Isbbins two-armed bandit problem with finite memory", *A.M.S.* 36, 1375-86 (1965).

- 54) Sobel, M. and Weiss, G., "Inverse sampling and other selection procedures for tournaments with two or three players", In 'Non-parametric Techniques in Statistical Inference' (M.L.Puri, ed.) 515-43, Cambridge U.Press(1970).
- 55) Sobel, M. and Weiss, G.H., "Play-the-winner sampling for selecting the better of two binomial populations", Biometrika, 57, 357-365(1970).
- 56) Sobel, M. and Weiss, G.H., "Play-the-winner rule and inverse sampling in selecting better of two binomial populations", JASA, 66, 545-551(1971).
- 57) Sobel, M. and Berry., Minnesota University. Tech. Rep., 154(1972).
- 58) Statistical Research Group, Columbia University, "Sequential analysis of statistical data : applications", Columbia U. Press(1945).
- 59) Statistical Research Group, Columbia University, "Sampling Inspection principles, Procedures, and Tables for Single, Double, and Sequential Sampling in Acceptance Inspection and Quality Control Based on Percent Defective", McGraw-Hill(1948).
- 60) Wald, A., "Sequential Analysis" John Wiley(1947).
- 61) Wetherill, G.B., "Sequential Methods in Statistics", Methuen(1966).
- 62) Zelen, M., "Play the winner rule and the controlled clinical trial", JASA, 64, 131-146(1969).