Title	Some Thoughts in Sequential Two Sample Problems with Data Dependent Allocation Rule
Author(s)	Takahashi, Hajime
Citation	Hitotsubashi Journal of Economics, 28(2): 173-181
Issue Date	1987-12
Type	Departmental Bulletin Paper
Text Version	publisher
URL	http://hdl.handle.net/10086/7851
Right	

Hitotsubashi Journal of Economics 28 (1987) 173-181. © The Hitotsubashi Academy

SOME THOUGHTS IN SEQUENTIAL TWO SAMPLE PROBLEMS WITH DATA DEPENDENT ALLOCATION RULE*

HAJIME TAKAHASHI

I. *Introduction*

Let $x_1, x_2, \ldots, x_m, \ldots$ and $y_1, y_2, \ldots, y_n, \ldots$ be independent normally distributed random variables with means $Ex_m=\mu_1$, $Ey_n=\mu_2$ and variances Var $x_m=\sigma_1^2$, Var $y_n=\sigma_2^2$ $(m, n=1, 2, \ldots)$. A sequential decision problem of testing $H_1:\delta>0$ against $H_2:\delta<0$, $\delta=\mu_1-\mu_2$ with data dependent allocation rule has been considered by many authors including Robbins and Siegmund (1974), Louis (1975), Hayre (1979) and Hayre and Gittins (1981). The problem is invariant with respect to the common change of location for both x and y, it is reasonable to restrict to the invariant procedure. So after having observed $x_1, \ldots, x_m$, and $y_1, \ldots, y_n$, we shall consider

$$Z_{m,n}=\frac{mn}{\sigma_2^2 m+\sigma_1^2 n}(\bar{x}_m-\bar{y}_n), \quad \bar{x}_m=\frac{1}{m}\sum_1^m x_i, \ \bar{y}_n=\frac{1}{n}\sum_1^n y_i \tag{1}$$

[Robbins and Siegmund (1974), Hayre and Gittins (1981)]. By simple algebra it follows that

$$E_\delta\{Z_{m,n}\}=[mn/(\sigma_2^2 m+\sigma_1^2 n)]\delta$$

and

$$\mathrm{Var}_\delta\{Z_{m,n}\}=[mn/(\sigma_2^2 m+\sigma_1^2 n)]\,.$$

Now by applying Wald's SPRT to the statistic $Z_{m,n}$, Robbins and Siegmund (1974) proved that (to the extent that we disregard the excess over the boundaries) the error probabilities and $E\{MN/\sigma_2^2 M+\sigma_1^2 N\}$ are independent of the allocation rules used (here (M,N) denotes 2-dimensional stopping time), provided they depend on previous x's and y's only through their differences. Related to this result, they have proved the following technical lemma

Lemma 1. Let $W(t)$ be a Brownian motion with drift δ and variance 1 per unit time. Let $\tau_{m,n}=mn/(m\sigma_2^2+n\sigma_1^2)$. Then for any allocation rules which depend on x's and y's only through the $Z_{i,j}$'s, the joint distribution of $\{Z_{m,n}, m,n=1,2,\ldots\}$ is equal to that of $\{W(\tau_{m,n}), m,n=1,2,\ldots\}$.

Since Robbins and Siegmund (1974) considers the two-population analogue of Wald's SPRT, their result shares the same drawback as SPRT; the test is open ended. To remedy this, Hayre (1979) considered a sequential decision problem of choosing one of three hypo-

* Research supported in part by the Grant-in-Aid for Scientific Research of the Ministry of Education, Science and Culture of Japan. The author thanks to Mrs. Akiko Ishikawa for her compuattional assistance.

theses, $H_0:\delta=0$, $H_1:\delta>0$ and $H_2:\delta<0$. He proposed to use the boundary suggested by Armitage (1957), which is a version of truncated SPRT. Continuing in this direction, Siegmund (1985, Ch. 6) applied the boundary for the repeated significance test of Armitage (1975). To be more specific, let $b>0$ and $0\leq K_0<K$ are given constant. We shall let

$$(M,N)=\text{first}\{(m,n):\tau_{m,n}>K_0,\quad |Z_{m,n}|>b\sqrt{\tau_{m,n}}\}\ . \tag{2}$$

We stop sampling at $\min\{\tau_{M,N},K\}$ and

$$\begin{aligned}&\text{Accept } H_1 \text{ if } Z_{M,N}>b\sqrt{\tau_{M,N}} \quad \text{and} \quad \tau_{M,N}\leq K\\ &\text{Accept } H_2 \text{ if } Z_{M,N}<-b\sqrt{\tau_{M,N}} \quad \text{and} \quad \tau_{M,N}\leq K\\ &\text{Accept } H_0 \text{ if } Z_{M,N}>K.\end{aligned} \tag{3}$$

Knowing that the error probabilities and the expected sample size $E\{\tau_{M,N}\wedge K\}$ are approximately the same for all invariant allocation rules, Siegmund (1985) has calculated these characteristics under the pairwise allocation rule.

Now the extension to the regression model may be of interest. Let $x_m=\mu_1+\beta\xi_m+\varepsilon_{x_m}$, $m=1, 2, \ldots$ and $y_n=\mu_2+\beta\eta_n+\varepsilon_{y_n}$, $n=1, 2, \ldots$, where μ_1 and μ_2 are unknown intercepts and β denotes an unknown common slope, $\{\xi_m,\ m\geq 1\}$ and $\{\eta_n,\ n\geq 1\}$ are sequences of known constants and $\{\varepsilon_{x_m},\ m\geq 1\}$ and $\{\varepsilon_{y_n},\ n\geq 1\}$ are independent and normally distributed random variables with means 0 and known variances $\text{Var}\,\varepsilon_{x_m}=\sigma_1^2$ and $\text{Var}\,\varepsilon_{y_n}=\sigma_2^2$. By long and tedious algebra, after having observed $(x_1,\xi_1), \ldots, (x_m,\xi_m)$ and $(y_1,\eta_1), \ldots, (y_n,\eta_n)$, our invariant test statistic is

$$\tilde{Z}_{m,n}=\frac{mn}{m\sigma_2^2+n\sigma_1^2}[\bar{x}_m-\bar{y}_n-b_{m,n}(\bar{\xi}_m-\bar{\eta}_n)] \tag{4}$$

where

$$\bar{x}_m=\frac{1}{m}\sum_1^m x_i,\quad \bar{y}_n=\frac{1}{n}\sum_1^n y_i,\quad \bar{\xi}_m=\frac{1}{m}\sum_1^m \xi_i,\quad \bar{\eta}_n=\frac{1}{n}\sum_1^n \eta_i$$

and

$$\begin{aligned}b_{m,n}=&\{(\sigma_2^2\sum_1^m \xi_i x_i+\sigma_1^2\sum_1^n \eta_i y_i)(m\sigma_2^2+n\sigma_1^2)-(m\sigma_2^2\bar{\xi}_m+n\sigma_1^2\bar{\eta}_n)(m\sigma_2^2\bar{x}_m+n\sigma_1^2\bar{y}_n)\}\\ &\div\{(\sigma_2^2\sum_1^m \xi_i^2+\sigma_1^2\sum_1^n \eta_i^2)(m\sigma_2^2+n\sigma_1^2)-(m\sigma_2^2\bar{\xi}_m+n\sigma_1^2\bar{\eta}_n)^2\}\end{aligned}$$

(See Takahashi (1977) for the special case $\sigma_1^2=\sigma_2^2=1$).

Note that

$$E\{\tilde{Z}_{m,n}|\xi,\eta\}=\frac{mn}{m\sigma_2^2+n\sigma_1^2}C_{m,n}\delta,\quad \text{and}$$

$$\text{Var}\{\tilde{Z}_{m,n}|\xi,\eta\}=\frac{mn}{m\sigma_2^2+n\sigma_1^2}C_{m,n}$$

where $\delta=\mu_1-\mu_2$ and

$$\begin{aligned}C_{m,n}=&\{\sigma_2^2\sum_1^m \xi_i^2+\sigma_1^2\sum_1^n \eta_i^2-(m\sigma_2^2\bar{\xi}_m^2+n\sigma_1^2\bar{\eta}_n^2)\}\\ &\div\{\sigma_2^2\sum_1^m \xi_i^2+\sigma_1^2\sum_1^n \eta_i^2-(m\sigma_2^2\bar{\xi}_m+n\sigma_1^2\bar{\eta}_n^2)^2/(m\sigma_2^2+n\sigma_1^2)\}\ .\end{aligned} \tag{5}$$

Since $C_{m,n}$ is increasing in m and n [Lai and Robbins (1977)], we have the analogous result of Lemma 1.

Lemma 2. Given ξ's and η's, the joint distribution of $\{Z_{m,n},\ m,n=1, 2, \ldots\}$ is the same as that of $\{W(\tilde{\tau}_{m,n}),\ m,\ n=1, 2, \ldots\}$, where $\tilde{\tau}_{m,n}=\tau_{m,n}C_{m,n}$.

By the simple algebra $0\leq C_{m,n}\leq 1$ and $C_{m,n}=1$ if $\bar{\xi}_m=\bar{\eta}_n$ and they are not degenerated. It follows that we can get the full efficiency (with respect to the simple two sample problem) by taking the mean of ξ's and η's about the same but not degenerated to the same point.

Hence the sequential decision problem of choosing which of these three hypotheses $H_0:\delta=0$, $H_1:\delta>0$ and $H_2:\delta<0$ may be carried out exactly in the same manner as above up to the Brownian approximation.

II. *Allocation Rule*

Numerous allocation rules have been proposed since the pioneering work by Louis (1972). Among them the one proposed by Hayre (1979) is of interest. Given δ the cost of taking one observation on x is $g(\delta)>0$ and the cost of taking y is $h(\delta)>0$ for which $h(\delta)-g(\delta)$ has the same sign of δ. A specific example considered below is

$$g(\delta)=h(-\delta)=\begin{cases}1 & \delta>0\\ 1+d|\delta| & \delta<0\end{cases}. \tag{6}$$

The optimal allocation rule is defined to be one which minimizes the risk

$$R(\delta)=g(\delta)E_\delta(M)+h(\delta)E_\delta(N). \tag{7}$$

The answer to this question is in

Lemma 3. [Hayre (1979)]. Let $w(\delta)=\{h(\delta)/g(\delta)\}^{1/2}$. For the test (3), for any invariant allocation rule

$$R(\delta)\geq(\sigma_1 g^{1/2}+\sigma_2 h^{1/2})^2 E_\delta\{\tau_{m,n}\} \tag{8}$$

and equality holds if and only if

$$P_\delta\{M/N=(\sigma_1/\sigma_2)w(\delta)\}=1. \tag{9}$$

Since δ is unknown to us, the suggested allocation rule would be:

Allocation Rule [HA]. After having observed $x_1, \ldots, x_m$ and $y_1, \ldots, y_n$, if $|Z_{m,n}|<b\sqrt{\tau_{m,n}}$, then we choose x_{m+1} next if and only if

$$m/n\leq(\sigma_1/\sigma_2)w(\hat{\delta}_{m,n}),\quad \hat{\delta}_{m,n}=Z_{m,n}/\tau_{m,n}. \tag{10}$$

The law of large number implies that (9) will hold approximately in large samples, then it is not difficult to see

$$\begin{aligned}E_\delta\{M\}&\simeq\sigma_1(\sigma_1+\sigma_2 w(\delta))E_\delta\{\tau_{M,N}\}\\ E_\delta\{N\}&\simeq\sigma_2(\sigma_2+\sigma_1 w^{-1}(\delta))E_\delta\{\tau_{M,N}\}.\end{aligned} \tag{11}$$

The situation is quite the same for the regression problem. Analogous to (8) we obtain

$$R(\delta)\geq E_\delta\{[\sigma_1(g/C_{M,N})^{1/2}+\sigma_2(h/C_{M,N})^{1/2}]^2\tilde{\tau}_{M,N}\} \tag{12}$$

and equality holds if and only if (9) holds. Hence, we may use the same sampling

TABLE 1

Simple two sample problems with known variance $\sigma_1^2=\sigma_2^2=1$
$b=2.95$, $K=49$, $K_0=0$
Pairwise sampling=1st raw, Hayre's rule=2nd raw ($d=20$)

δ	H_2	H_0	H_1	$E(M)$	$E(N)$	$E(M+N)$	Risk
0.0	0.02	0.97	0.01	97.5	97.5	195.0	195.0
	0.03	0.95	0.02	100.5	105.0	205.5	205.5
0.2	0.00	0.87	0.13	93.4	93.4	186.8	560.5
	0.00	0.84	0.16	143.4	72.2	215.6	504.4
0.4	0.00	0.42	0.58	71.5	71.5	143.0	715.4
	0.00	0.45	0.55	138.8	50.0	188.8	588.8
0.6	0.00	0.05	0.95	42.3	42.3	84.6	592.2
	0.00	0.06	0.94	100.9	30.4	131.3	496.1
0.8	0.00	0.00	1.00	27.6	27.6	55.2	496.5
	0.00	0.01	0.99	59.3	16.2	75.5	334.7
1.0	0.00	0.00	1.00	20.0	20.0	40.0	437.8
	0.00	0.00	1.00	42.4	11.3	53.7	279.7

In Tables 1–6 all the values are obtained from the average of 400 repetition. We have used subroutine RANN2 in SSL II and the system FACOM M-360 at Hitotsubashi University.

TABLE 2

Regression case with $\beta=0.5$, known variance $\sigma_1^2=\sigma_2^2=1$
$b=2.95$, $K=49$, $K_0=0$
Pairwise sampling=1st raw, Hayre's rule=2nd raw ($d=20$)

δ	H_2	H_0	H_1	$E(M)$	$E(N)$	$E(M+N)$	Risk
0.0	0.02	0.97	0.01	97.5	97.5	195.0	195.0
	0.03	0.95	0.02	103.1	106.2	209.3	209.3
0.2	0.00	0.87	0.13	93.4	93.4	186.8	560.5
	0.00	0.86	0.14	143.9	73.2	217.1	509.8
0.4	0.00	0.42	0.58	71.5	71.5	143.1	715.4
	0.00	0.46	0.54	140.0	50.5	190.5	594.4
0.6	0.00	0.05	0.95	42.3	42.3	84.6	592.1
	0.00	0.05	0.95	92.3	28.1	120.4	458.1
0.8	0.00	0.00	1.00	27.6	27.6	55.2	496.5
	0.00	0.00	1.00	60.8	18.6	79.4	376.7
1.0	0.00	0.00	1.00	19.9	19.9	39.8	437.8
	0.00	0.00	1.00	38.2	13.4	51.6	319.6

rule [*HA*] here with $\hat{\delta}_{m,n}$ replaced by $\tilde{Z}_{m,n}/\tilde{\tau}_{m,n}$. If we may choose ξ's and η's for which $C_{M,N}\to\alpha\in(0,1)$, we can show

$$E_\delta\{M\}\simeq\alpha^{-1}\sigma_1(\sigma_1+\sigma_2 w(\delta))E_\delta\{\tilde{\tau}_{M,N}\} \tag{11'}$$
$$E_\delta\{N\}\simeq\alpha^{-1}\sigma_2(\sigma_2+\sigma_1 w^{-1}(\delta))E_\delta\{\tilde{\tau}_{M,N}\}$$

as $b\to\infty$. Some of the simulation results are presented in Table 1 and 2.

III. *Unknown Variance*

When variances are not known to us, we may replace all σ_i^2's by their unbiased esti-

TABLE 3
Simple two sample problems with unknown variance
$b=2.95$, $K=49$ $(M,N\geq 10)$
Pairwise sampling=1st raw, Heyre's rule=2nd raw $(d=20)$

δ	H_2	H_0	H_1	$E(M)$	$E(N)$	$E(M+N)$	Risk	σ_x^2 / σ_y^2
	0.01	0.98	0.01	109.5	109.5	219.0	219.0	1.12
0.0								1.13
	0.01	0.98	0.01	116.8	116.9	233.7	233.7	1.11
								1.11
	0.01	0.88	0.11	105.5	105.5	211.0	633.0	1.11
0.2								1.13
	0.00	0.86	0.14	162.0	80.6	242.6	564.8	1.12
								1.14
	0.00	0.44	0.56	83.9	83.9	167.8	839.0	1.16
0.4								1.15
	0.00	0.41	0.59	131.8	43.4	175.2	655.8	1.12
								1.12
	0.00	0.04	0.96	51.2	51.2	102.4	716.8	1.12
0.6								1.22
	0.00	0.05	0.95	109.9	33.8	143.7	549.4	1.15
								1.19
	0.00	0.00	1.00	34.2	34.2	68.4	615.6	1.27
0.8								1.30
	0.00	0.00	1.00	70.1	19.4	89.5	399.5	1.18
								1.25
	0.00	0.00	1.00	22.8	22.8	45.6	501.6	1.29
1.0								1.32
	0.00	0.00	1.00	49.6	13.0	62.6	323.5	1.21
								1.26

mator $\hat{\sigma}_i^2$ $(i=1, 2)$. By the strong law of large numbers, $\hat{\sigma}_i^2$ converges strongly to σ_i^2 as $b\to\infty$. Hence at least asymptotically, the modified procedure posseses the same properties discussed in the previous section. Probabilistically the problem is too complicated and we do not discuss in detail here. Some of their small sample properties are obtained by Monte Carlo studies. We first use unbiased estimator of σ_i^2 $(i=1, 2)$ in both simple and regression cases. In the simple two sample problem, it turns out that usual unbiased estimator seriously underestimate the true σ_i^2 for all values of δ. We then used the estimator proposed by Lai et al. (1975). The results are satisfactory as shown in Table 3. As to the regression case, the ususal unbiased estimator behaves fairly well (Table 4). We do not know why these phenomena occurs.

IV. *Estimation after Testing*

In this section we shall suppose that $\sigma_1^2=\sigma_2^2=1$ and consider testing sequentially $H_0:\delta=0$ against $\delta\neq 0$ with data dependent allocation rule $[HA]$. When H_0 is rejected, we are interested in estimating the value of δ based on the data collected so far. The problem of estimating δ after sequential testing has been considered by Siegmund (1978), Takahashi (1987), Woodroofe and Keener (1987), where they have utilized the non-linear renewal theory

TABLE 4
Regression case with $\beta=0.5$, unknown variance
$b=2.95$, $K=49$ ($M,N\geq 10$)
Pairwise sampling=1st raw, Hayre's rule=2nd raw ($d=20$)

δ	H_1	H_0	H_2	$E(M)$	$E(N)$	$E(M+N)$	Risk	σ_x^2/σ_y^2
	0.02	0.97	0.01	114.3	114.3	228.6	228.6	1.17
0.0								1.18
	0.03	0.94	0.03	120.4	121.6	242.0	242.0	1.16
								1.17
	0.01	0.87	0.12	109.6	109.6	219.2	657.4	1.17
0.2								1.17
	0.00	0.85	0.15	166.6	83.8	250.4	585.5	1.20
								1.13
	0.00	0.44	0.56	84.6	84.6	169.2	845.9	1.13
0.4								1.14
	0.00	0.44	0.56	164.1	58.4	222.5	689.9	1.18
								1.10
	0.00	0.05	0.95	52.4	52.4	104.8	734.1	1.09
0.6								1.09
	0.00	0.04	0.96	97.0	29.5	126.5	480.0	1.11
								1.00
	0.00	0.00	1.00	29.4	29.4	58.8	528.8	1.04
0.8								1.04
	0.00	0.00	1.00	62.6	18.4	81.0	375.6	1.10
								0.99
	0.00	0.00	1.00	20.9	20.9	41.8	45818	1.03
1.0								1.02
	0.00	0.00	1.00	40.6	13.6	54.2	326.4	1.03
								0.97

to approximate the excess over the boundary. In this section, however, we shall use the Brownian motion $W(t)$ to approximate the original process $Z_{m,n}$, which is tantamount to disregarding the excess over the boundary. As has been pointed out by several authors, Brownian approximation usually does not give us a good numerical accuracy [Siegmund (1985), Ch. 3], which is the consequence of disregarding the excess over the boundary. But in this case, the situation is much better. After having observed $x_1, \ldots, x_m$ and $y_1, \ldots, y_n$, if we take one more x sample (y sample), then the average increment would be $[n^2/(m+n)(m+n+1)]\delta$ ($[m^2/(m+n)(m+n+1)]\delta$). This is substantially smaller than δ when $m \gg n$ ($n \gg m$), which may explain the good approximation we have below (see Table 5).

As in the previous sections, we let $\{W(t), t\geq 0\}$ denote the Brownian motion with drift δ and variance 1 per unit time. For $b>0$ we shall define a stopping time

$$T=\inf\{t\geq 0;\ |W(t)|\geq b\sqrt{t}\}\ . \tag{13}$$

Then the continuous time analogue of the repeated signiffcance test for $\delta=0$ would be

$$\text{Reject } H_0\colon \delta=0 \quad \text{iff} \quad T\leq K \tag{14}$$

for some given constant $K>0$.

If $T\leq K$, then the naive estimator for δ would be

$$\bar{X}=W(T)/T\,, \tag{15}$$

which, however, has a serious upward (downward) bias when $\delta>0$ ($\delta<0$). In order to obtain the correction term, Siegmund (1978) utilizes our intuitive feelings that small value of the stopping time T are evidence in favor of large value of $|\delta|$. On the other hand Takahashi (1987) considers the asymptotic expansion of the distribution of $(W(T)-\delta T)/\sqrt{T}$ to get the correction terms. Although in the discrete time case these two approaches are based on completely different methods, they depend on the same technical result in this continuous time problem.

Lemma 4 [Siegmund (1985), Ch. 4]. Let $\delta>0$. Suppose $b\to\infty$ $k\to\infty$ such that $b/\sqrt{k}=\delta_0$ remain bounded away from 0 and ∞. Then uniformly in δ in the closed subintervals of $(0, \infty)$

$$P_\delta\{T<k\}=1-\Phi[\sqrt{k}(\delta_0-\delta)]+\{\phi[\sqrt{k}(\delta_0-\delta)]\delta/\sqrt{k}\}\{1+o(1)\} \tag{16}$$

as $b\to\infty$.

TABLE 5
90% confidence intervals of δ
Simple two sample problem with Hayre's rule ($d=20$)
$b=2.95$, $K=49$, $K_0=0$

δ	Native estimator	Siegmund's	Takahashi's
0.2		0.077, 0.707	−0.094, 0.398
	0.247	0.392	0.152
0.4		0.097, 0.856	0.113, 0.634
	0.503	0.477	0.388
0.6		0.125, 1.124	0.222, 0.934
	0.728	0.625	0.578
0.8		0.155, 1.404	0.286, 1.195
	0.929	0.779	0.740
1.0		0.189, 1.670	0.351, 1.473
	1.136	0.922	0.912
1.2		0.223, 1.954	0.455, 1.777
	1.378	1.089	1.116

TABLE 6
90% confidence intervals of δ
Simple two sample problem with Hayre's rule ($d=20$)
$b=3.45$, $K=148$, $K_0=0$

δ	Native estimator	Siegmund's	Takahashi's
0.2		0.122, 0.450	0.039, 0.321
	0.200	0.286	0.180
0.4		0.192, 0.705	0.260, 0.658
	0.466	0.449	0.459
0.6		0.270, 0.996	0.376, 0.960
	0.681	0.633	0.668
0.8		0.356, 1.311	0.506, 1.275
	0.918	0.833	0.890
1.0		0.441, 1.625	0.608, 1.586
	1.144	1.033	1.097
1.2		0.504, 1.864	0.709, 1.845
	1.346	1.184	1.277

Theorem 1. Under the same asymptotic relations of Lemma 4

$$P_\delta\left\{\frac{W(T)-\delta T}{\sqrt{T}}\leq x\right\}=\Phi(x)-\frac{1}{b}\phi(x)+o(b^{-1}) \tag{17}$$

Proof. Let T_+ denote the one sided version of T

$$T_+=\inf\{t\geq 0;\ W(t)\geq b\sqrt{t}\}.$$

It is easily seen that for any $\delta>0$

$$P_\delta\{T_+>T\}=o(b^{-2})\quad b\to\infty\ .$$

Hence,

$$(W(T_+)-\delta T_+)/\sqrt{T_+}=b-\delta\sqrt{T_+}\ .$$

Thus

$$\begin{aligned}P_\delta\left\{\frac{W(T)-\delta T}{\sqrt{T}}\leq x\right\}&=P_\delta\{T>(b-x/\delta)^2\}+o(b^{-2})\\&=\Phi(x)-\frac{1}{b}\phi(x)+o(b^{-1})\quad b\to\infty\ .\end{aligned}$$

Corollary 1.

$$P_\delta\left\{\frac{W(T)-\delta T}{\sqrt{T}}\leq x+\frac{1}{b}\right\}=\Phi(x)+o(b^{-1}) \tag{18}$$

as $b\to\infty$.

From (18) we may define $(1-\alpha)\times 100\%$ confidence interval of δ based on the data $\{T,W(T)\}$,

$$\begin{aligned}&[W(T)/T-z_{\alpha/2}/\sqrt{T},\ W(T)/T+z_{\alpha/2}/\sqrt{T}]-1/(b\sqrt{T})\quad\text{on}\quad[T\leq K,\ W(T)\geq 0]\\&[W(K)/K-z_{\alpha/2}/\sqrt{K},\ W(K)/K+z_{\alpha/2}/\sqrt{K}]\qquad\text{on}\quad[T>K]\\&(W(T)/T-z_{\alpha/2}/\sqrt{T},\ W(T)/T+z_{\alpha/2}/\sqrt{T}]+1/(b\sqrt{T})\quad\text{on}\quad[T\leq K,\ W(T)\leq 0]\end{aligned} \tag{19}$$

where $\Phi(z_\beta)=1-\beta$.

The numerical accuracy of (19) is compared with simulation results, Siegmund's estimator, and the naive estimator $W(T)/T$. On a whole Siegmund's estimator (center of 90% confidence interval) performs better than (19) $(W(T)/T-1/(b\sqrt{T}))$. But the width of the intervals of Siegmund's estimators are usually wider than (19). It suggests us to obtain higher order asymptotic expansion in (17), which will be considered in the next project. [cf. Takahashi (1987), Woodroofe and Keener (1987)].

HITOTSUBASHI UNIVERSITY

REFERENCES

Armitage, P. (1957), "Restricted sequential procedures," *Biometrika* 44, pp. 9–26.
Armitage, P. (1975), *Sequential Medical Trials, 2nd Edition*, Oxford, Blackwell.
Hayre, L.S. (1979), "Two-population sequential tests with three hypotheses," *Biometrika* 66, 465–474.

Hayre, L.S. and Gittins, J.C. (1981), "Sequential selection of the larger of two normal means," *J. Am. Statists. Assoc.* 76, pp. 696–700.

Lai, T.L. and Robbins, H. (1977), "Strong consistency of least-squares estimates in regression models," *Proc. Natl. Acad. Sci.* USA 74, pp. 2667–2669.

Lai, T.L., Robbins, H. and Siegmund, D. (1977), "Sequential decision about a normal mean," *In Statistical Decision Theory and Related Topics II* in: S.S. Gupta, ed., pp. 213–221. Academic Press.

Louis, T.A. (1972), *Two Population Hypothesis Testing with Data Dependent Allocation.* Ph.D. Dissertation, Columbia University, N.Y.

Louis, T.A. (1975), "Optimal allocation in sequential tests comparing the means of two Gaussian populations," *Biometrika* 62, pp. 359–369.

Robbins, H. and Siegmund, D. (1974), "Sequential tests involving two populations," *J. Am. Statists. Assoc.* 69, pp. 132–139.

Siegmund, D. (1978), "Estimation following sequential tests," *Biometrika* 65, pp. 341–349.

Siegmund, D. (1985), *Sequential Analysis*, Springer-Verlag.

Takahashi, H. (1977), "Sequential two sample problem in regression," *Hitotsubashi J. of Soc. Sci.* 2, pp. 28–39.

Takahashi, H. (1987), "Asymptotic expansions in Anscombe's theorem for repeated significance tests and estimation after sequential testing" *Ann. Statists.* 15, pp. 278–295.

Woodroofe, M. and Keener, R. (1987), "Asymptotic expansions in boundary crossing problems," *Ann. Probability* 15, pp. 102–114.