The Korean Communications
in Statistics Vol. 6, No. 1, 1999
pp. 143-149

# Adjusting Practical Aims in Optimal Extended Double Sampling Plans

Seoung-gon Ko[1)]

## Abstract

Ko(1998) proposed a procedure to enhance the efficiency of double sampling plans by allowing second-stage sample size and critical region to depend on first-stage evidence using constraint optimization approaches. In this study, further developments of such plans by incorporating several practically possible researcher's aims into the optimization are considered. Comparisons are made with the optimal ordinary double sampling plan and also among them. It is observed that it is to some extent possible to match the details of the optimization to certain qualitative methodological aims.

## 1. Introduction

Ko(1998) considers two-stage double sampling with variable second-stage. For such plans, involving a second-stage, it is then possible to talk about error rate both of the entire two-stage procedure, and also, in keeping with the "group sequential(Pocock(1982))" point of view, of each of its two stages. These plans are "extended" double sampling plans, which constitutes an extension of classical double sampling plans, in the sense that the second-stage sampling effort and second-stage critical value are allowed to depend on the point at which the first-stage continuation region is traversed. At the end of the first-stage an interim analysis is performed with the objective of deciding whether or not to continue the study based on results of the interim analysis. If the study is continued, the first-stage information is systematically put to work in conducting the second-stage, including its sample size and critical region, with the goal of achieving agreed-upon overall, as well as stage-specific operating characteristics.

Instead of the idea on previous work(Stein's(1945)) which focuses on a nuisance parameter ($\sigma$), this study adapts the utilization of the parameter($\mu$) of interest in first-stage information. Such focus on the parameter of interest is particularly reminiscent of the approach to sequential estimation indicated by Birnbaum and Healy(1960); and also of Miller and Freund(1977)'s more recent proposal of an ad hoc method for determining the second-stage sample size in binomial estimation problems. Where our approach differs from

1) Assistant Professor, Department of Applied Statistics, Kyungwon University, Sungnam, 461-701, Korea

previous work is in our casting of the design of extended double sampling plans in the form of constrained optimization problems. The objective functions for optimization problems are conditional averages of sampling effort measures. In this research we consider and try to demonstrate extended double sampling plans conducted in the two stages with several practically possible objective functions.

Section 2 defines four possible objective functions which are considered by other authors and brief reviews of the optimization. Section 3 gives and shows the $OC$ and $AST$(Average Sampling Time) functions for optimal plans defined in Section 2 and discusses the results.

## 2. Formulation and optimization

Wiener process and specifically, discriminating between the Wiener process with drift parameter $\mu_0$ and the Wiener process with drift parameter $\mu_1$, $\mu_1 > \mu_0$, with assuming a unit scale parameter, is considered for testing $H_0: \mu = \mu_0$ against $H_1: \mu = \mu_1 > \mu_0$, with error probabilities $P\{\text{Reject } H_0 | \mu = \mu_0\} = \alpha$ and $P\{\text{Accept } H_0 | \mu = \mu_1\} = \beta$.

In accordance with the group-sequential point of view, $\alpha$ and $\beta$ are equally allocated to the two stages. Such allocations which is to be distinguished from the repeated-significant-test usage(Armitage et al.(1969)) make possible independent assignment of error rates to both stages. Under the given $\alpha/2$ and $\beta/2$ in the first-stage, the continuation region, $(l, u)$, is determined by initial sampling time $T_0$. To satisfy the four stage-specific error rate restrictions, $T_0$ must be taken from a feasible interval $(0, \overline{T}]$. Given $(T_0, l, u)$, $\mu_1$ (respectively, $\mu_0$) is to be accepted at time $T_0$ if the process exceeds $u$ (respectively, is below $l$). Further, if it equals an intermediate value $s$, $l < s < u$, then a predetermined second-stage sampling plan is implemented, with sampling time $T_s$ and critical value $k_s$ depending on $s$. Thus the plan is determined by three numbers $(T_0, l, u)$, plus the two functions $\{T_s: l < s < u\}$ and $\{k_s: l < s < u\}$.

The plan $(T_0, l, u, \{T_s; s \in (l, u)\}, \{k_s; s \in (l, u)\})$ is to be chosen in such a way as to minimize an objective function measuring sampling effort, subject to stage-specific error rates restrictions. Among the objective functions so considered are the following :

i) $F_1 = \overline{AST} = (\mu_1 - \mu_0)^{-1} \int_{\mu_0}^{\mu_1} AST_\mu d\mu$

where $AST_\mu = T_0 + \int_l^u \frac{T_s}{\sqrt{2\pi}\sqrt{T_0}} \exp[-\frac{1}{2}(\frac{s-\mu T_0}{\sqrt{T_0}})^2] ds$,

$$\text{ii)}\ F_2 = AST_{\bar{\mu}},\ \text{where}\ \bar{\mu} = \frac{\mu_0+\mu_1}{2},$$

$$\text{iii)}\ F_3 = AST_{\mu_1},$$

$$\text{iv)}\ F_4 = \overline{AST}_N = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \exp[-\frac{1}{2}(\mu-\bar{\mu})^2]\, AST_\mu d\mu. \qquad (2.1)$$

The first objective function is the unweighted average of expected sampling times for the values of $\mu$ between $\mu_0$ and $\mu_1$. The second is the expected sampling time for $\mu=\bar{\mu}$. The third is the expected sampling time for $\mu=\mu_1$. The fourth objective function is the weighted average of expected sampling times for all values of $\mu$ in the interval $(-\infty, \infty)$, the weights being provided by the (conjugate) normal prior distribution of $\mu$ centered at $\bar{\mu}$ with unit variance.

These objective functions were considered previously by several authors in deriving their optimal plans. Especially, Colton and McPherson(1976) consider the third one for deriving their optimal plans which emphasize clinical efficiency when actual treatment differences exist. Also, minimizing the second one, known as the Keifer-Weiss problem, is considered by Jennison(1987). The first one is a generalization of a simple average of average sampling efforts at two hypotheses points as considered by Thall et al.(1988). The last one is, in fact, a common choice of many Bayesian approaches because of its conjugation.

Brief review of the conditional optimizations using objective functions in (2.1) is as follows: Given $T_0$, $l$ and $u$, we then minimize the function of $T_s$, $F's$ in (2.1), subject to the restriction that the second-stage error rates equal $\alpha/2$ and $\beta/2$. It is equally possible to consider second stages only, for example, $\overline{AST}^{(2)} \equiv \overline{AST} - T_0$ in case of $F_1$. This minimization is with respect to the second-stage sampling time function $T_s$, $l \le s \le u$, and second-stage critical value function $k_s$, $l \le s \le u$. We consider this restricted optimizations by subdividing the continuation region $(l, u)$, as determined by the first-stage sampling time $T_0$ into a set of equally spaced $m$ grid points where $m$ is sufficiently large enough to make sure the accuracy of numerical computations. Then, we simply abbreviate formulation with a objective function to

$$\text{minimize}_{\{T_i\},\{k_i\}}\ F(\{T_i\}_1^m)$$

$$\text{subject to}\ g_0(T_1, T_2, \ldots, T_m, k_1, \ldots k_m) - \alpha/2 = 0 \quad \text{and}$$

$$g_1(T_1, T_2, \ldots, T_m, k_1, \ldots k_m) - \beta/2 = 0$$

where $f$, $g_0$, $g_1$ are of additive form. (2.2)

The solutions of (2.2) is obtained by a standard Kuhn-Tucker argument of solving Lagrangian functions and separability of this Lagrangian functions. The details of this issues are due to section 2 of Ko(1998).

## 3. Comparisons and further comments

By following the outlined procedures of section 2 and we can find the optimal plans for $F_2$, $F_3$, $F_4$, given the same error restrictions as those of Section 2. At optimum, sampling effort is about equally divided between the first- and second-stages. Moreover, in the case of objective functions $F_1$, $F_2$ and $F_4$, $T_s$ is maximum for $s$ roughly half-way between $l$ and $u$, and drops to approximately half that maximum at $s = l$ and $s = u$; while, in the case of $F_3$, $T_s$ is largest at $s = l$, and as would be expected, smallest at $s = u$. These features are shown in Figure 1.

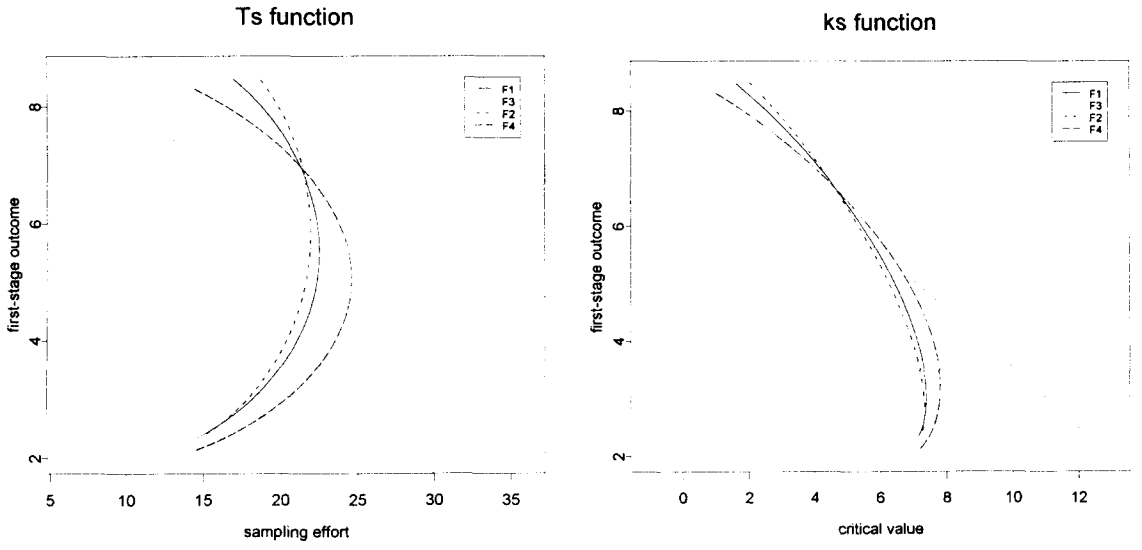


**Figure 1** : Optimal second-stage sampling efforts and critical values

We also note this in contemplating our optimum sampling plans minimizing $F_2$ and $F_1$: even optimizing expected sampling effort right between $\mu_0$ and $\mu_1$ cannot counteract the natural tendency of optimal sequential plans to call for maximal sampling effort in this intermediate zone.

For the case of $(\alpha, \beta, \mu_0, \mu_1) = (0.05, 0.10, 0, 0.50)$, the $OC$ and $AST$ functions for the optimal extended double sampling plans corresponding to our four objective functions are given in Table 1 and illustrated in Figure 2. The $OC$ functions for all our plans are essentially same. Actually, they are coincided within the third decimal points. $AST$ functions are alike, except for the case of $F_3$. The plan corresponding to $F_3$ is clearly best near $\mu_1$, and worst near $\mu_0$. Note that the $AST$ of $F_3$ is not uniformly smaller than that of the optimal Ordinary Double Sampling plan( $ODS^*$ ) in the neighborhood of $\mu_0$. This is due to different criteria of both optimizations. The $AST$ function for $F_3$ also dominates the $AST$ function of $ODS^*$ derived by using the analogue of $F_3$.

Parallel results for the case $(\alpha, \beta, \mu_0, \mu_1) = (0.05, 0.10, 0, 1.0)$ also illustrated in Figure 3.

It is equally possible to develop such designs for the other weights over parameter space or hypotheses points and different allocations of error rates according to researcher's aim.

**Table 1** : $OC_\mu$ and $AST_\mu$ of optimal extended double sampling plans( $EDS^*$ ) and optimal ordinary double sampling plan( $ODS^*$ ) for the Wiener process $w(0,1)$and $w(0.5,1)$ with $(\alpha, \beta) = (0.05, 0.10)$ and $m=64$.

| $\mu$ | $AST_\mu$ | | | | |
|---|---|---|---|---|---|
| | $ODS^*$ | $F_1$ | $F_2$ | $F_3$ | $F_4$ |
| -0.500 | 19.95320 | 18.964783 | 19.061801 | 18.33289 | 18.77126 |
| -0.375 | 20.16018 | 19.193370 | 19.282164 | 18.74168 | 19.01740 |
| -0.25 | 20.85626 | 19.914601 | 19.982812 | 19.89402 | 19.78220 |
| -0.125 | 22.55449 | 21.591626 | 21.623497 | 22.28006 | 21.53608 |
| 0.000 | 25.49407 | 24.401593 | 24.391660 | 25.80068 | 24.43582 |
| 0.125 | 28.89291 | 27.598187 | 27.565812 | 29.20155 | 27.68527 |
| 0.250 | 30.96834 | 29.568191 | 29.550958 | 30.62768 | 29.63280 |
| 0.375 | 30.34827 | 29.063548 | 29.088476 | 29.19697 | 29.04696 |
| 0.500 | 27.46079 | 26.417334 | 26.482561 | 25.84087 | 26.31535 |
| 0.625 | 24.08306 | 23.206890 | 23.293986 | 22.37129 | 23.05101 |
| 0.750 | 21.66190 | 20.804565 | 20.899243 | 19.98693 | 20.62395 |
| 0.875 | 20.46213 | 19.547288 | 19.644528 | 18.79896 | 19.35591 |
| 1.000 | 20.03515 | 19.068318 | 19.167046 | 18.35647 | 18.87170 |

| $\mu$ | $OC_\mu$ | | | | |
|---|---|---|---|---|---|
| | $ODS^*$ | $F_1$ | $F_2$ | $F_3$ | $F_4$ |
| -0.500 | 0.99999 | 0.999981 | 0.999982 | 0.99998 | 0.999980 |
| -0.375 | 0.99985 | 0.999819 | 0.999822 | 0.99978 | 0.999812 |
| -0.25 | 0.99865 | 0.998542 | 0.998555 | 0.99835 | 0.998512 |
| -0.125 | 0.99044 | 0.990237 | 0.990261 | 0.98976 | 0.990171 |
| 0.000 | 0.95000 | 0.949998 | 0.949999 | 0.95000 | 0.949998 |
| 0.125 | 0.82131 | 0.821766 | 0.821711 | 0.82324 | 0.821942 |
| 0.250 | 0.57307 | 0.573178 | 0.573168 | 0.57367 | 0.573246 |
| 0.375 | 0.29011 | 0.289651 | 0.289710 | 0.28830 | 0.289508 |
| 0.500 | 0.10000 | 0.100000 | 0.100000 | 0.10000 | 0.100000 |
| 0.625 | 0.02337 | 0.023772 | 0.023723 | 0.02479 | 0.023897 |
| 0.750 | 0.00392 | 0.004179 | 0.004149 | 0.00472 | 0.004254 |
| 0.875 | 0.00052 | 0.000593 | 0.000584 | 0.00073 | 0.000613 |
| 1.000 | 0.00006 | 0.000069 | 0.000068 | 0.00009 | 0.000073 |

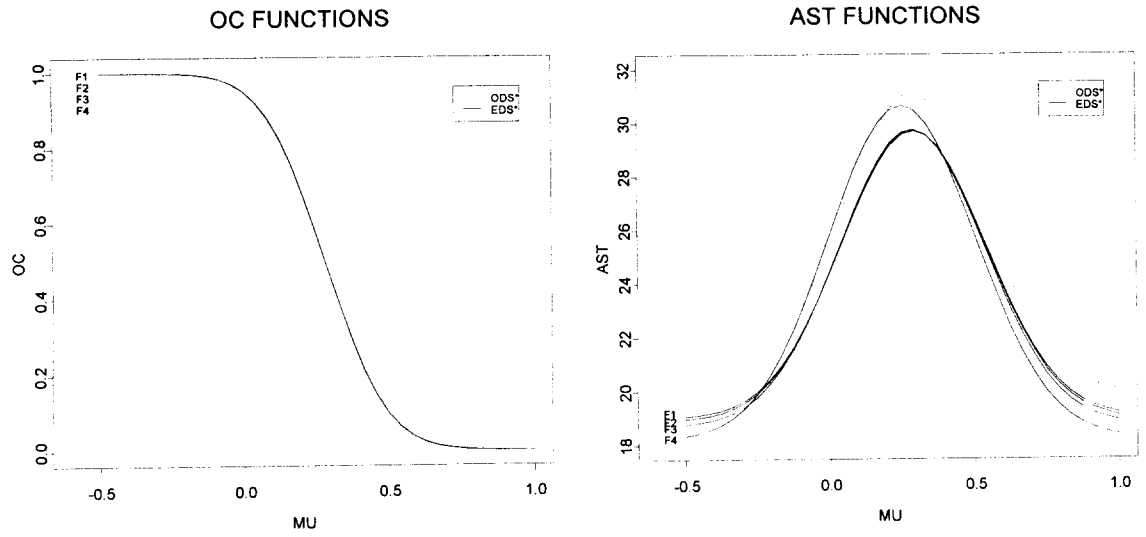


**Figure 2** : $OC_\mu$ and $AST_\mu$ of $EDS^*$ and $ODS^*$ for the Wiener process $w(0,1)$and $w(0.5,1)$ with $(\alpha, \beta) = (0.05, 0.10)$.

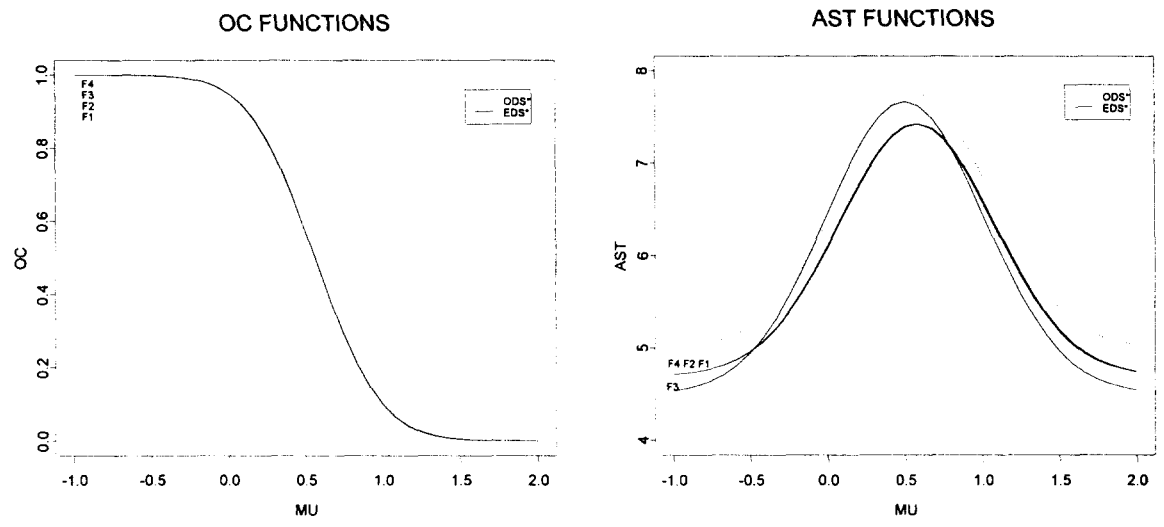


**Figure 3** : $OC_\mu$ and $AST_\mu$ of $EDS^*$ and $ODS^*$ for the Wiener process $w(0,1)$and $w(1,1)$ with $(\alpha,\ \beta) = (0.05,\ 0.10)$.

## References


[1] Armitage, P., McPherson, K. and Rowe, B.C. (1969). Repeated Significance tests on accumulating data. *Journal of the Royal Statistical Society*, Series A, 132, 235-244.

[2] Birnbarum,A. and Healy,W.C.Jr.(1960). Estimates with prescribed variance based on two-stage sampling. *The Annals of Mathematical Statistics* 31, 662-676.

[3] Colton, T. and McPherson, K. (1976). Two-stage plans compared with fixed sample size and Wald SPRT plans. *Journal of the American Statistical Association* 71, 80-86.

[4] Jennison, C. (1987). Efficient group seqeuntial tests with unpredictable group sizes. *Biometrika,* 74, 155-165.

[5] Ko, S. G.(1998). Hybrid Group-Sequential Conditional-Bayes Approaches to the Double Sampling Plans, *The Korean Communications in Statistics*, Vol. 5, No. 1, 107-120.

[6] Miller, I. and Freund, J. E.(1977). *Probability and Statistics for engineers*, 2nd edition, Englewood Cliff, NJ : Prentice Hall.

[7] Pocock, S. J.(1982). Interim analyses for randomized clinical trials : the group-sequential approach. *Biometrics* 38, 153-162.

[8] Stein, C.(1945). A two-sample test of linear hypothesis whose power is independent of the variance. *The Annals of Mathematical Statistics* 16, 243-258.

[9] Thall, P.F., Simon, R., Ellenberg, S.S., Shrager, R.(1988), Optimal two-stage designs for clinical trials with binary response, *Statistics in Medicine*, 7, 571-579.