Advances and Applications in Statistics
Volume …, Number …
This paper is available online at http://www.pphmj.com

ε-BAYES SEQUENTIAL PLAN

SUMIHIRO SUZUKI and MICHAEL I. BARON

Department of Biostatistics
University of North Texas Health Science Center
Fort Worth, Texas
U. S. A.
e-mail: ssuzuki@hsc.unt.edu

Department of Mathematical Sciences
University of Texas at Dallas
Richardson, Texas
U. S. A.

Abstract

An optimal (Bayes) sequential plan minimizes the (Bayes) risk function which takes into account the decision loss, observation (variable) cost, and group (fixed) cost. In general, determining the optimal sequential plan remains an open problem mainly because it requires risk-optimization over a rather unstructured set of all plans. An ε-Bayes sequential plan is a sequential plan comparable to the Bayes plan yet in general more computationally feasible. Methods of deriving upper and lower bounds for the Bayes sequential plan are shown which are then used to defined an ε-Bayes sequential plan.

1. Introduction

Classical (pure) sequential procedures are often impractical because of their requirement to take a single observation at a time. Such sampling schemes are often expensive and time consuming. Sequentially

2000 Mathematics Subject Classification: **Kindly Provide.**

Keywords and phrases: ε-Bayes, risk function, sequential plan.

Communicated by Karan P. Singh

Received September 4, 2008

planned procedures, or simply sequential plans, extend and generalize these schemes by allowing observations to be collected in groups of variable sizes. After every group, all the previously collected data are used to determine the next course of action. Formally, a sequential plan $\mathcal{N}$ is a family of integer-valued functions

$$\mathcal{N} = \{N^{(t)} : \mathbf{X}^{(t)} \to \{0, 1, 2, \ldots\}\},$$

where $N^{(t)}$ is a random variable which returns the sample size of the next group given a sample of size t, $\mathbf{X}^{(t)} = (X_1, X_2, ..., X_t)$. The ultimate goal is to find an optimal sequential plan out of all possible plans. We treat optimality in the decision theoretic sense of minimizing the overall risk function of the sequential plan. Let $\mathcal{N}$ be any non-random sequential plan. Let $\delta(\mathbf{X})$ be the terminal decision given $\mathbf{X} = (X_1, X_2, \cdots)$. Let $L(\theta, \delta)$ be the loss function under the parameter (true state of nature) $\theta \in \Theta$ and terminal decision $\delta \in \mathcal{A}$. Let C_j and T_j be the cost of sampling a group of size j and the total number of groups of size j sampled, respectively. The cost is fixed whereas the number of groups is (usually) random. The (frequentist) risk function is given by

$$R(\theta, \mathcal{N}, \delta) = \mathbf{E}_\theta\left\{L(\theta, \delta(\mathbf{X})) + \sum_j T_j C_j\right\}.$$

Given a prior distribution $\pi(\theta)$ on θ, we can define the Bayes risk,

$$r(\pi, \mathcal{N}, \delta) = \mathbf{E}_\pi R(\theta, \mathcal{N}, \delta).$$

A sequential plan is called *Bayes* iff it minimizes the Bayes risk over all possible sets of sequential plans. In general, determining the optimal sequential plan remains an open problem mainly because it requires risk-optimization over a rather unstructured set of all plans.

2. *M*-truncated Sequential Plans

One of the difficulties of determining the Bayes sequential plan is the fact that in general there is infinitely many possible futures to consider.

Even when the maximum number of observations needed is known in advance, there may still be an overwhelming number of ways to allocate different group sizes. An obvious method of bypassing this difficulty is to consider a class of sequential plans which terminates no later than after a predetermined number of groups have been collected. Hence we define an *M-truncated sequential plan* as a plan which terminates after at most M groups have been taken. That is to say, an M-truncated sequential plan $\mathcal{N}_{\mathcal{MT}}$ is such that

$$\mathcal{N}_{\mathcal{MT}} = \begin{Bmatrix} N_j^{(t)} : \mathbf{X}^{(t)} \to \{0, 1, 2, \ldots\} & \text{for all } j \le M \\ N_j^{(t)} : \mathbf{X}^{(t)} \to \{0\} & \text{for all } j > M \end{Bmatrix},$$

where $N_j^{(t)}$ is the size of the jth group given $\mathbf{X}^{(t)}$ and M is some predeter-mined constant. This sequential plan is commonly referred to as a multistage sampling plan (e.g., Ghosh and Sen [7, Chapter 6]), but we use a different name in order to emphasize the fact that this sequential plan takes no more than M groups. A natural goal when using an M-truncated sequential plan is to seek the optimal sequential plan among all M-truncated sequential plans in the hope that it performs comparably to the overall Bayes plan for a suitable choice of M. Hence a *Bayes M-truncated sequential plan* minimizes the Bayes risk among all M-truncated sequential plans.

Note that as with any other sequential plan, the optimal terminal decision at any given time t is equivalent to the fixed sample Bayes decision given $\mathbf{X}^{(t)}$ (Schmitz [14, Theorem 4.4]). Also in general, there is an upper bound on the largest group size which can be taken at any given stage for the Bayes plan because we assume that the cost associated to sampling a group increases as group size increases. That is to say, if C_k the cost of group of size k is higher than the terminal risk of making an immediate decision, a group of that size should not be sampled. Hence the search for the Bayes M-truncated sequential plan only has to be done over a finite set. Then a Bayes M-truncated sequential plan is theoretically calculable even in situation where the overall Bayes plan is not.

Clearly, the class of all M-truncated sequential plans is a subclass of all possible sequential plans. Then because we are optimizing over a smaller class, the Bayes risk of any M-truncated sequential plan is no less than that of the overall Bayes plan. Hence for a suitably large M, the Bayes M-truncated sequential plan would provide a tight upper bound for the overall minimum Bayes risk. Now then if we can also come up with a lower bound to this minimum Bayes risk, we would have a method of estimating the Bayes sequential plan.

3. *M*-inner Truncated Sequential Plan

The Bayes sequential plan, by definition, has the lowest Bayes risk among all sequential plans. Then in order for us to come up with a lower bound, we must make some modifications to the premise of the original problem.

In particular, we modify the loss function.

Let $L(\theta, \delta, T, nT)$ be the loss plus cost of some sequential problem where θ is the true state of nature, δ is the terminal decision taken, T is the total number of groups taken, and n_T is the total number of observations collected through T groups. For any sequential problem given $L(\theta, \delta, T, n_T)$, the *M-inner truncated loss function*, $L^M(\theta, \delta, T, n_T)$, is defined as

$$L^M(\theta, \delta, T, n_T) = \begin{cases} L(\theta, \delta, T, n_T) & \text{if } T < M, \\ \inf_\delta L(\theta, \delta, M, n_M) & \text{if } T \geq M. \end{cases}$$

Any sequential plan based on L^M instead of L is called an *M-inner truncated sequential plan*. In other words, the expected loss plus cost (frequentist risk) as well as the Bayes risk for an M-inner truncated sequential plan is calculated based on L^M rather than the usual L. Hence a *Bayes M-inner truncated sequential plan* is a sequential plan which minimizes the Bayes risk based on L^M. It should be clear that as M tends to infinity, L^M converges to L. A similar idea was introduced for pure sequential procedures by Herman Rubin (Berger [3, Section 7.4.7]).

When the total number of groups reaches M, we assume that the optimal terminal decision is taken. That is to say, we assume that there is no more to gain by sampling more groups (observations). Because we assume that the optimal terminal decision is taken, we can further assume that when $T = M$, the terminal decision loss is 0. Hence the M-inner truncated loss function can be written as

$$L^M(\theta, \delta, T, n_T) = \begin{cases} L(\theta, \delta) + \sum_{i=1}^{T} C_{T_i} T_i & \text{if } T < M, \\ \sum_{i=1}^{M} C_{T_i} T_i & \text{if } T \geq M, \end{cases}$$

where $L(\theta, \delta)$ is the terminal decision loss, T_i is the number of observations in the ith group, $\sum_{i=1}^{T} T_i = n_T$, and C_{T_i} is the cost of one group of size T_i, $i = 1, 2, \ldots, M$.

Essentially this means that the decision loss is waived if the Mth group is taken. Therefore we see that

$$L^M(\theta, \delta, T, n_T) \leq L(\theta, \delta, T, n_T),$$

for any M, and hence the Bayes M-inner truncated sequential plan has a Bayes risk no higher than any other sequential plan. Then for a suitably large M, we can use the M-inner truncated sequential plan to come up with a tight lower bound for the overall minimum Bayes risk. Note that the computational complexity here is equivalent to that of M-truncated sequential plans. Because there is no advantage in taking any more than M groups, we only need to consider plans which are truncated at M.

We should mention that when searching for a lower bound, we must use the Bayes (optimal) M-inner truncated sequential plan as oppose to any other M-inner truncated plan. In general there is no guarantee that the M-inner truncated sequential plan produces a lower Bayes risk than the overall Bayes plan unless it is the best sequential plan among all M-inner truncated sequential plans. This is different than searching for an upper bound. Any sequential plan based on L can be used to calculate an upper bound for the Bayes risk because under L, there is no sequential plan which can perform better than the overall Bayes plan.

4. ε-Bayes Sequential Plan: Method One

Now that we have a method of deriving both an upper and a lower bound for the minimum Bayes risk, we can come up with a method to find ε-*Bayes sequential plans*. We say that a sequential plan, $\mathcal{N}_\varepsilon$, is an ε-*Bayes Sequential Plan* iff its Bayes risk is within ε of the overall Bayes sequential plan $\mathcal{N}_B$, i.e.,

$$r(\pi, \mathcal{N}_\varepsilon, \delta) - r(\pi, \mathcal{N}_B, \delta) \leq \varepsilon,$$

for some $\varepsilon > 0$. To calculate an ε-Bayes sequential plan, we first have to choose a suitable M such that

$$r(\pi, \mathcal{N}_{MT}, \delta) - r(\pi, \mathcal{N}_{MIT}, \delta) \leq \varepsilon,$$

where $\mathcal{N}_{MT}$ and $\mathcal{N}_{MIT}$ are the Bayes M-truncated and M-inner truncated sequential plans, respectively. Because both of these sequential plans converge to the overall Bayes sequential plan as M tends to infinity, this difference tends to 0 as M tends to infinity. After finding an M which produces the desired ε value, the Bayes M-truncated sequential plan is an ε-Bayes sequential plan. We should note that in general the Bayes M-inner truncated sequential plan which was used to come up with a lower bound is not an ε-Bayes sequential plan because the Bayes risk for $\mathcal{N}_{MIT}$ is calculated using the M-inner truncated loss function.

The only problem which may arise with this methodology is that in order to achieve the desired ε value, the necessary M may have to be rather large. In such cases the computational complexity could still be too much to effectively implement an ε-Bayes sequential plan. To circumvent this problem we introduce another method of deriving an upper and a lower bound, and use them to come up with a more practical approach to ε-Bayes sequential plans.

5. Group Look-ahead Sequential Plan

A k-group sequential plan samples groups of size k at every stage until the termination of the experiment. We consider any sequential problem where the optimal k-group sequential plan is known for any positive integer k and the Bayes risk can be computed. Such sequential problems include problems where an optimal stopping boundary or rule can be calculated.

The idea of the group look-ahead sequential plan is that at every stage we assume we are following a k-group sequential plan and try to find the optimal k. However, after sampling one group, we again look for the optimal k, and repeat until we stop. Imagine that we are looking to travel by train from Station A to Station B with a stop at another station in between. From Station A, one particular train may lead us to Station B in the least amount of time. However, we can further cut down on travel time if we switch to another train after we stop at the station in between. Then provided everything being equal, it would be more efficient to switch trains.

Formally we can define a *group look-ahead sequential plan*, $\mathcal{N}_{GL}$, as the following

$$\mathcal{N}_{GL} = \{N_j^{(t)} : \mathbf{X}^{(t)} \to \{0, 1, \ldots\}\},$$

where $N_j^{(t)}$ is the jth group size given $\mathbf{X}^{(t)}$,

$$N_j^{(t)} = \{n \geq 0 : \arg\min_k r(\pi^{(t)}, \mathcal{N}_k, \delta(\mathbf{X}^{(t)})) = n\}.$$

Here $\pi^{(t)}$ is the posterior distribution of the parameter of interest given $\mathbf{X}^{(t)}$ (prior distribution whenever $t = 0$), $\mathcal{N}_k$ is a k-group sequential plan, and $r(\cdot)$ is the Bayes risk of the sequential plan. The algorithm to find a group look-ahead sequential plan is to first determine the best k-group sequential plan based on the prior distribution $\pi^{(0)}$. Let $N_1^{(0)} = k_0$ (say), and sample one group of k_0 observations. Next we calculate the best k-group sequential plan again, but this time uses the posterior distribution given $\mathbf{X}^{(k_0)}$. Let $N_2^{(k_0)} = k_1$ (say), and sample one group of k_1 observations. We repeat this process until $N_j^{(t)} = 0$ for some j and t.

Because the maximum group size is bounded at every stage, the computational complexity now significantly reduces. At each stage we only need to compare the Bayes risks of finitely many k-group sequential plans. By assuming that we continue according to a k-group sequential plan, we essentially bypass the difficulty which comes from the uncertainty of future group size allocations. Furthermore, because any one k-group sequential plan is a subclass of the group look-ahead sequential plan, it can only improve k-group sequential plans. Lastly, and most importantly for us, the group look-ahead sequential plan is still a subclass of all sequential plans, and hence its Bayes risk is an upper bound for the overall Bayes plan.

6. Zero Group Cost Sequential Plan

Now we derive another method of coming up with a lower bound for the minimum Bayes risk. As before, this can only be done if we make some modifications to the initial premise of the problem. Up to this point, we have always denoted cost of a group of size k as C_k. Recall that this C_k, accounts for two cost factors. One is the actual cost of k observations, and the other is the cost associated to the group itself.

Consider a sequential problem where the cost associated to the group itself is either zero or waived. The optimal sequential plan in such a scenario would be a pure sequential plan. Because there is no group cost, there is no penalty in sampling one observation at a time, or equivalently there is no advantage in sampling in batches. We have essentially modified the original problem so that a pure sequential plan is the Bayes sequential plan. We call such a sequential plan a *zero group cost sequential plan*.

The nice thing here is that the zero group cost sequential plan can be easily computed in every case where the group look-ahead plan can be computed (a pure sequential plan is equivalent to a k-group sequential plan with $k = 1$). The Bayes risk of the optimal zero group cost sequential plan is no greater than that of the overall Bayes plan, and hence serves as a lower bound. The sharpness of this lower bound depends on the original cost associated to the group itself. The lower the said cost, the tighter the bound. That is if a significant portion of the Bayes risk comes from the group cost, then waiving this cost significantly affects the reliability of the resulting lower bound.

7. ε-Bayes Sequential Plan: Method Two

We again use the derived upper and lower bounds for the minimum Bayes risk to come up with an ε-Bayes sequential plan. Let $\mathcal{N}_{GL}$ and $\mathcal{N}_{ZG}$ denote the Bayes group look-ahead and Bayes zero group cost sequential plans, respectively. Let ε be the difference in the Bayes risks of these two plans, i.e.,

$$\varepsilon = r(\pi, \mathcal{N}_{GL}, \delta) - r(\pi, \mathcal{N}_{ZG}, \delta).$$

Then $\mathcal{N}_{GL}$ is an ε-Bayes sequential plan. The big difference with this method as oppose to the previous one is that the value of ε cannot be controlled. Though the two sequential plans used here are upper and lower bounds, unlike the truncated procedures, they do not converge to the overall Bayes sequential plan. This is the price we pay in return for using a computationally simple methodology. Because the value of ε cannot be preset, the experimenter must decide whether or not the resulting ε is acceptable or not.

Also, we can suggest one slight modification in the calculation of ε-Bayes sequential plans which may lead to a better estimate of the overall Bayes sequential plan. Essentially all we need is an upper bound and a lower bound, hence there is nothing preventing us from combining both methods. In particular we can compute two upper bounds and two lower bounds, then use the smaller of the upper bounds and the higher of the lower bounds.

8. Conclusion

The idea behind an ε-Bayes sequential plan is to bound the minimum Bayes risk from above and below in hopes of finding a sequential plan whose Bayes risk is no more than some $\varepsilon > 0$ away from the overall

Bayes sequential plan. One approach to coming up with an upper and a lower bound is to use an M-truncated and M-inner truncated sequential plans, respectively. The M-truncated sequential plan terminates sampling no later than after the Mth group is taken. The M-inner truncated sequential plan waives the terminal decision loss if the Mth group is taken. To compute an ε-Bayes sequential plan, a suitable M which makes the difference in the Bayes risk of the two truncated plans no greater than ε is chosen. Then the M-truncated sequential plan is an ε-Bayes sequential plan. As M tends to the Bayes risk of this plan tends to the minimum Bayes risk. A second approach is to use the group look-ahead and zero group cost sequential plans to bound the minimum Bayes risk. In situations where a group sequential solution exists and is easy to calculate, this second approach is much more computationally feasible. The group look-ahead sequential plan assumes that at every stage, sampling will continue according to a k-group sequential plan. However, instead of a fixed k, at every stage k which produces the minimum Bayes risk is computed. The zero group cost sequential plan assumes that the cost associated to the group itself is waived, and hence only the cost of observations need to be considered. Under such an assumption, the Bayes sequential plan is a pure sequential scheme because there is no advantage in sampling in batches. Furthermore, anytime a component of the risk, in this case group cost, is waived, the minimum Bayes risk is lower than that of the original setting. This sharpness of this lower bound depends the original group cost, the lower the group cost the tighter the bound. By letting the difference in Bayes risks of these two plans be ε, the group look-ahead sequential plan is an ε-Bayes sequential plan. Though the second approach is much more computationally feasible, the value of ε cannot be set in advance.

References

[1] D. Assaf, A dynamic sampling approach for detecting a change in distribution, Ann. Statist. 16 (1988), 236-253.

[2] W. Bartky, Multiple sampling with constant probability, Ann. Math. Statist. 14 (1943), 363-377.

[3] J. O. Berger, Statistical Decision Theory, Springer-Verlag, New York, 1985.

[4] D. A. Berry and C. H. Ho, One-sided sequential stopping boundaries for clinical trials, Biometrics 44 (1988), 219-227.

[5] P. J. Bickel and K. A. Doksum, Mathematical Statistics. Basic Ideas and Selected Topics, Vol. 1, 2nd Edition, Prentice-Hall, Upper Saddle River, New Jersey, 2001.

[6] P. Billingsley, Probability and Measure, 3rd Edition, John Wiley & Sons, 1995.

[7] B. K. Ghosh and P. K. Sen, Handbook of Sequential Analysis, Marcel-Dekker, Inc., 1991.

[8] Z. Govindarajulu, The Sequential statistical analysis of hypothesis testing, Point and Interval Estimation, and Decision Theory, Amer. Sci. Press, Columbus, Ohio, 1981.

[9] L. S. Hayre, Group sequential sampling with variable group sizes, J. Roy. Statist., Soc. B 47(1) (1985), 90-97.

[10] C. Jennison and B. W. Turnbull, Interim analyses: the repeated confidence interval approach, J. Roy. Statist., Soc. B 51(3) (1989), 305-361.

[11] C. Jennison and B. W. Turnbill, Group Sequential Methods with Applications to Clinical Trials, Chapman & Hall, Boca Raton, FL, 2000.

[12] T. L. Lai, Sequential analysis: some classical problems and new challenges, Statist. Sin. 11 (2001), 303-408.

[13] C. Schmegner, Decision Theoretic Results for Sequentially Planned Statistical Procedures, Ph. D. Dissertation, The University of Texas at Dallas, 2003.

[14] N. Schmitz, Optimal Sequentially Planned Decision Procedures, Springer-Verlag, New York, 1993.

[15] T. M. Therneau and H. S. Wieand, Optimal design for a grouped sequential binomial trial, Biometrics 46 (1990), 771-781.

[16] A. Wald, Sequential tests of hypotheses, Ann. Math. Statist. June (1945), 117-186.

[17] G. B. Wetherill, The most economical sequential sampling scheme for inspection by variables, J. Roy. Statist. Soc., B 21(2) (1959), 400-408.

■

Kindly return the proof after correction to: *The Publication Manager* *Adv. & Appl. in Stat.* *Pushpa Publishing House* *Vijaya Niwas* 198, *Mumfordganj* *Allahabad*-211002 (*India*) along with the print charges* by the fastest mail ***Invoice attached**	Proof read by: …………………………………… Signature: ………………………………………. Date: …………………………………………….. Tel: ……………………………………………… Fax: ……………………………………………… E-mail: ……………………………………… Number of additional reprints required …………………………… Cost of a set of 25 copies of additional reprints @ U. S. Dollars 15.00 per page. (25 copies of reprints are provided to the corresponding author ex-gratis)