

An Efficient Adaptive Sequential Procedure for Simple Hypotheses with Expression for Finite Number of Applications of Less Effective Treatment

Sampurna Kundu*, Jayant Jha[†] and Subir Kumar Bhandari[‡]

Interdisciplinary Statistical Research Unit, Indian Statistical Institute, Kolkata, India

Abstract

We propose an adaptive sequential framework for testing two simple hypotheses that analytically ensures finite exposure to the less effective treatment. Our proposed procedure employs a likelihood ratio-driven adaptive allocation rule, dynamically concentrating sampling effort on the superior population while preserving asymptotic efficiency (in terms of average sample number) comparable to the Sequential Probability Ratio Test (*SPRT*). The foremost contribution of this work is the derivation of an explicit closed-form expression for the expected number of applications to the inferior treatment. This approach achieves a balanced method between statistical precision and ethical responsibility, aligning inferential reliability with patient safety. Extensive simulation studies substantiate the theoretical results, confirming stability in allocation and consistently high probability of correct selection (*PCS*) across different settings. In addition, we demonstrate how the adaptive procedure markedly reduces inferior allocations compared with the classical *SPRT*, highlighting its practical advantage in ethically sensitive sequential testing scenarios. The proposed design thus offers an ethically efficient and computationally tractable framework for adaptive sequential decision-making.

Keywords. Adaptive sequential design; Adaptive allocation; Average sample number; Number of applications of inferior treatment; Probability of correct selection; Sequential probability ratio test.

MSC 2020 Classification. 62L05, 62L10.

*sampurna.kundu58@gmail.com, Corresponding Author.

[†]jayantjha@gmail.com

[‡]subirkumar.bhandari@gmail.com

1 Introduction

Sequential design, first formalized by Wald (1947) through the Sequential Probability Ratio Test (*SPRT* henceforth), is a cornerstone of modern statistical decision theory. It departs from fixed-sample schemes by allowing data to be analyzed as they are collected, and the decision to continue or stop sampling depends on accumulated information. The principal goal is to achieve efficient inference — minimizing the expected or average sample number (*ASN*) for given Type *I* and Type *II* error probabilities. This framework has since evolved into a variety of methodologies used in industrial quality control (Teoh et al., 2022; Li and Nenes, 2024), reliability testing (Rasay and Alinezhad, 2022; Jain and Jain, 1994), and clinical trials (Li et al., 2012; Martens and Logan, 2024), where resources are limited and early stopping can save both cost and ethical burden. In the classical sequential paradigm, the experimenter observes an equal number of samples from each of two or more populations at each stage and uses likelihood ratios computed from the collected data to update inference until a stopping criterion is satisfied.

However, adaptive sequential designs allocate experimental units (or observations) to competing treatments or populations sequentially, using information accrued so far to guide future allocation. At each stage, the probability of selecting a population depends on its estimated superiority. Adaptive sequential design embodies the dual objectives — achieving precise inference while safeguarding subjects from excessive exposure to inferior treatments, which has driven a large literature spanning bandit problems, response-adaptive randomization, and sequential testing (see Berry and Fristedt, 1985; Friedman et al., 2010; Ivanova and Rosenberger, 2000; Rosenberger et al., 2001). Typical approaches emphasize either optimality of inference or desirable allocation properties, but rarely provide closed-form control of the total use of the inferior option.

The central problem in adaptive allocation lies in balancing statistical optimality and ethical constraints. Ideally, an adaptive procedure should satisfy the following criteria:

- It should asymptotically allocate a higher proportion of samples to the better population and the number of samples allocated to the inferior population should remain small or, if possible, finite in expectation and higher moments.
- It should retain inferential efficiency comparable to the *SPRT*, ensuring that Type *I* and Type *II* error probabilities remain controlled.

Most existing methods satisfy only part of these goals. Urn and play-the-winner-type schemes tend to concentrate sampling on the superior treatment as trials progress, but typically the absolute number of allocations to the inferior treatment continues to grow with the total sample size. Moreover, very few of these studies provide analytical expressions for the number of inferior allocations or connect their procedures to formal inferential properties such as *ASN* or error control.

A number of studies have examined aspects of adaptive allocation and ethical sequential sampling. Biswas et al. (2007) provided a comprehensive treatment of response-adaptive designs in clinical trials, highlighting the ethical dimension of minimizing allocations to inferior treatments. Recent contributions by Das et al. (2023), Biswas et al. (2020), Bandyopadhyay et al. (2020), Das (2024a), and Das (2024b) have proposed new adaptive rules incorporating covariate adjustments, ordinal responses, crossover trials, multi-treatment response adaptive design, misclassifications, and adaptive interim decisions, all designed to enhance ethical allocation properties.

Earlier studies explored related versions of the two-treatment adaptive problem, but with important limitations. [Bhandari et al. \(2007\)](#) considered the case with known face value of the parameters, indicating that under a specific adaptive procedure, the expected number of allocations to the less effective treatment could be finite, although no closed-form expression or rigorous distributional analysis was provided, and inferential aspects such as the *ASN* or probability of correct selection (*PCS* henceforth) were not addressed. [Bhandari et al. \(2009\)](#) extended the problem to unknown parameters and obtained that the expected number of inferior allocations grows logarithmically with the total sample size, but the study did not link the allocation mechanism with inferential efficiency. More recently, [Kundu et al. \(2025\)](#) revisited the problem in an adaptive sequential context, proving that the number of inferior allocations is a finite random variable with finite moments. However, their proof relied on a subset of the sample space corresponding to correct selection events and did not yield a general closed-form expression, while the procedure’s inferential efficiency was only qualitatively observed but not comparable with the *SPRT* framework.

Building on that theoretical foundation, the present paper addresses the above gap by developing a new adaptive sequential procedure for two-sample/simple-hypotheses testing problem and employs a likelihood ratio–driven adaptive rule that determines at each stage which population to sample next, based on cumulative log-likelihood comparisons of the data collected so far. The method ensures that sampling effort is increasingly concentrated on the better-performing population, thereby simultaneously achieving statistical efficiency and ethical prudence. Secondly, and most importantly, we derive an explicit analytical expression for the expected number of allocations to the less effective treatment (valid for large sample regimes) and we prove that this count is a finite random variable with all moments finite. While the existence of a finite bound was hinted at in our earlier work ([Kundu et al., 2025](#)), no closed-form expression had been obtained. In the current framework, using the asymptotic behavior of cumulative likelihood ratios and the distributional properties of their standardized sums, we elucidate that this expected number converges to a finite value that depends on the mean and variance of the underlying log-likelihood ratio statistics. This formula provides a quantifiable measure of ethical efficiency, representing the expected finite number of applications of the less effective treatment. Thirdly, we demonstrate that the proposed adaptive rule retains the asymptotic efficiency same as that of *SPRT* in terms of *ASN*.

The remainder of this paper is organized as follows. Section 2 introduces the formal preliminaries and describes the proposed adaptive allocation rule (Method \mathcal{M}). Section 3 explicates the main theoretical results, including the proof of finiteness and the derivation of the closed-form expression for the number of inferior allocations, along with the efficiency comparison with the *SPRT*. Section 4 assesses the results of comprehensive simulation studies under various parameter choices and distributions. Finally, Section 5 concludes the paper.

2 Preliminaries

Let $\mathcal{X} = X_1, X_2, X_3, \dots$ and $\mathcal{Y} = Y_1, Y_2, Y_3, \dots$ two independent data streams be generated. X_i and Y_i have densities from $\{f_0, f_1\}$ with respect to some σ -finite measure. It is not known which f_i is assigned with X_i or Y_i .

2.1 Adaptive Sequential Procedure

We start with one sample each from \mathcal{X} and \mathcal{Y} . At the step n , using past data, we use the method \mathcal{M} to select the population from which to collect a sample next. At step n , let we have $N_{0,n}$ and $N_{1,n}$ samples from \mathcal{X} and \mathcal{Y} respectively, with $N_{0,n} + N_{1,n} = n$. Let us define the following:

$$n_{max} = \max\{N_{0,n}, N_{1,n}\} \quad \text{and} \quad n_{min} = \min\{N_{0,n}, N_{1,n}\}.$$

When $N_{0,n} = N_{1,n}$, we assign them to n_{max} and n_{min} with probability $\frac{1}{2}$ each.

Consider two simple null and alternative hypotheses: $H_0 : (f_0, f_1)$ vs $H_1 : (f_1, f_0)$, where the first coordinate represents density that corresponds to \mathcal{X} .

Lemma 2.1. *At step n , the samples collected from \mathcal{X} is $(X_1, X_2, \dots, X_{N_{0,n}})$ and the samples collected from \mathcal{Y} is $(Y_1, Y_2, \dots, Y_{N_{1,n}})$. n samples together conditioned by $(N_{0,n}, N_{1,n})$ are independent and conditional distribution of $X_1, X_2, \dots, X_{N_{0,n}}$ is i.i.d. and that of $Y_1, Y_2, \dots, Y_{N_{1,n}}$ is also i.i.d..*

Proof.

$$\mathbb{P}(N_{0,n}, N_{1,n}) = \sum_{path} \mathbb{P}(\text{path leading to } (N_{0,n}, N_{1,n})).$$

In the below, let f denote respective densities (or, probabilities) for the random variables (or, events) given after it.

$$\begin{aligned} & f(X_1, X_2, \dots, X_{N_{0,n}} | N_{0,n}, N_{1,n}) \\ &= \frac{f(X_1, X_2, \dots, X_{N_{0,n}}, N_{0,n}, N_{1,n})}{f(N_{0,n}, N_{1,n})} \\ &= \frac{\sum_{path} f(X_1, X_2, \dots, X_{N_{0,n}}, \text{path leading to } (N_{0,n}, N_{1,n}))}{f(N_{0,n}, N_{1,n})} \\ &= \frac{\sum_{path} f(X_1, X_2, \dots, X_{N_{0,n}} | \text{path leading to } (N_{0,n}, N_{1,n})) \cdot \mathbb{P}(\text{path leading to } (N_{0,n}, N_{1,n}))}{f(N_{0,n}, N_{1,n})} \\ &= \frac{f(X_1)f(X_2) \dots f(X_{N_{0,n}}) \cdot \sum_{path} \mathbb{P}(\text{path leading to } (N_{0,n}, N_{1,n}))}{f(N_{0,n}, N_{1,n})} \\ &= f(X_1)f(X_2) \dots f(X_{N_{0,n}}), \end{aligned}$$

where, f is f_0 (under H_0) and f_1 (under H_1).

Similarly, this holds for Y_i 's. □

Remark 1. *We consider without loss of generality, f_0 to be better distribution. Under H_0 , $X \sim f_0$ and under H_1 , $Y \sim f_0$. Without loss of generality, we consider that H_0 is true.*

2.2 Method \mathcal{M}

One observes that $n_{max} \geq \frac{n}{2}$. At step n , we consider $U_1, U_2, \dots, U_{n_{max}}$ i.i.d. sample where U_i (X_i or Y_i) corresponds to n_{max} .

- (i) If $\log \left[\frac{\prod_{i=1}^{n_{max}} f_0(U_i)}{\prod_{i=1}^{n_{max}} f_1(U_i)} \right] > 0$, we draw one more sample from f_i corresponding to n_{max} .

- (ii) If $\log \left[\frac{\prod_{i=1}^{n_{\max}} f_0(U_i)}{\prod_{i=1}^{n_{\max}} f_1(U_i)} \right] < 0$, we draw one more sample from f_i corresponding to n_{\min} .
- (iii) If $\log \left[\frac{\prod_{i=1}^{n_{\max}} f_0(U_i)}{\prod_{i=1}^{n_{\max}} f_1(U_i)} \right] = 0$, we draw a sample from f_0 or f_1 with probability $\frac{1}{2}$ each.

3 Main Result

We aim to allocate more samples to a better density (f_0). $\mathbb{P}\mathbb{I}_n$ denotes the probability of incorrect allocation in this context at step n . We consider f_0 and f_1 to be continuous. If they are not continuous, we need to adjust a little bit (not shown in the paper). Here, in this context, let

$$\begin{aligned}
\mathbb{P}\mathbb{I}_n &= \begin{cases} \mathbb{P} \left[\log \left(\prod_{i=1}^{n_{\max}} \frac{f_0(U_i)}{f_1(U_i)} \right) < 0 \right], & \text{if } U \sim X, \\ \mathbb{P} \left[\log \left(\prod_{i=1}^{n_{\max}} \frac{f_0(U_i)}{f_1(U_i)} \right) > 0 \right], & \text{if } U \sim Y \end{cases} \\
&= \begin{cases} \mathbb{P} \left[\sum_{i=1}^{n_{\max}} Z_i^{(X)} < 0 \right], & \text{if } U \sim X \text{ where } Z_i^{(X)} = \log \left(\frac{f_0(X_i)}{f_1(X_i)} \right), \\ \mathbb{P} \left[\sum_{i=1}^{n_{\max}} Z_i^{(Y)} > 0 \right], & \text{if } U \sim Y \text{ where } Z_i^{(Y)} = \log \left(\frac{f_0(Y_i)}{f_1(Y_i)} \right) \end{cases} \\
&= \begin{cases} \mathbb{P} \left[\sum_{i=1}^{n_{\max}} \frac{Z_i^{(X)} - \eta_x}{\sigma_x} < -\frac{\eta_x}{\sigma_x} \cdot n_x \right], & \text{if } U \sim X, \\ \mathbb{P} \left[\sum_{i=1}^{n_{\max}} \frac{Z_i^{(Y)} - \eta_y}{\sigma_y} > -\frac{\eta_y}{\sigma_y} \cdot n_y \right], & \text{if } U \sim Y \end{cases} \\
&\rightarrow \begin{cases} 1 - \Phi \left(\frac{\eta_x}{\sigma_x} \cdot \sqrt{n_x} \right), & \text{if } U \sim X, \\ \Phi \left(\frac{\eta_y}{\sigma_y} \cdot \sqrt{n_y} \right), & \text{if } U \sim Y. \end{cases}
\end{aligned}$$

Here, η_x , σ_x and η_y , σ_y are respective means and standard deviations of $Z_i^{(X)} = \log \left(\frac{f_0(X_i)}{f_1(X_i)} \right)$ and $Z_i^{(Y)} = \log \left(\frac{f_0(Y_i)}{f_1(Y_i)} \right)$, with $n_x = N_{0,n}$ and $n_y = N_{1,n}$. The expected number of allocations to the less effective treatment is

$$\begin{aligned}
\mathbb{E}(N_{1,n}) &= \sum_{m=2}^n \mathbb{P}\mathbb{I}_m \\
&\approx \sum_{i=1}^{n_x} \left(1 - \Phi \left(\frac{\eta_x}{\sigma_x} \sqrt{i} \right) \right) + \sum_{j=1}^{n_y} \Phi \left(\frac{\eta_y}{\sigma_y} \sqrt{j} \right) \\
&\leq \sum_{i=1}^{\infty} \left(1 - \Phi \left(\frac{\eta_x}{\sigma_x} \sqrt{i} \right) \right) + \sum_{j=1}^{\infty} \Phi \left(\frac{\eta_y}{\sigma_y} \sqrt{j} \right) < \infty. \tag{1}
\end{aligned}$$

As $\eta_x > 0$ and $\eta_y < 0$, $N_{1,n} < \infty \forall n$, and $\lim_{n \rightarrow \infty} \mathbb{E}(N_{1,n}) < \infty$.

With $n_x = N_{0,n}$ and $n_y = N_{1,n}$ moderately large, the expected number of allocations to the less effective treatment can be given by the following result.

Theorem 3.1 (Expression for the expected number of inferior allocations). *For large n , the expected number of allocations to the less effective treatment under the proposed adaptive rule (and with the assumptions given in Section 2) satisfies*

$$E(N_{1,n}) \approx \frac{1}{2} \left(\frac{\sigma_x^2}{\eta_x^2} + \frac{\sigma_y^2}{\eta_y^2} \right),$$

which represents a finite constant depending only on the first two moments of the log-likelihood ratio statistics. Also, as $N_{1,n}$ is the sum of independent Bernoulli variables, it follows that all the moments of $N_{1,n}$ are bounded.

Proof. Starting from the preceding summation (Equation 1) and applying the normal approximation to the tail probabilities,

$$\begin{aligned} \mathbb{E}(N_{1,n}) &\approx \sum_{i=1}^{\infty} \Phi\left(-\frac{\eta_x}{\sigma_x} \cdot \sqrt{i}\right) + \sum_{j=1}^{\infty} \Phi\left(\frac{\eta_y}{\sigma_y} \cdot \sqrt{j}\right) \\ &\approx \int_{-\infty}^0 \Phi\left(-\frac{\eta_x}{\sigma_x} \cdot \sqrt{-t}\right) dt + \int_{-\infty}^0 \Phi\left(\frac{\eta_y}{\sigma_y} \cdot \sqrt{-t}\right) dt \\ &= \frac{1}{2} \left(\frac{\sigma_x^2}{\eta_x^2} + \frac{\sigma_y^2}{\eta_y^2} \right) \quad [\text{By Integration by parts}] \end{aligned}$$

where $N_{0,n}$ and $N_{1,n}$ are moderately large.

This is finite. This is the approximate value of the number of applications of the less effective treatment for large n .

As $\sum_{m=2}^{\infty} \mathbb{P}\mathbb{I}_m < \infty$, moment generating function of $N_{1,n}$ is finite for finite domain by a constant function not depending on n . Thus, all the moments of $N_{1,n}$ are similarly bounded. \square

Remark 2. $\mathbb{E}(N_{1,n})$ increases to a finite quantity, as $n \rightarrow \infty$. Also, we proved that all the moments of $N_{1,n}$ are bounded. Hence, $\frac{N_{1,n}}{N_{0,n}} \rightarrow 0$ in probability, as $n \rightarrow \infty$, by Markov inequality.

3.1 Stopping Rule

We will perform the *SPRT* with U_1, U_2, \dots to test hypotheses $K_0 : U_i \sim f_0$ Vs $K_1 : U_i \sim f_1$, where U_i corresponds to data stream for n_{max} . Under H_0 , if K_0 is accepted, we have correct selection. Here, we derive the expression for probability of incorrect selection.

Remark 3. In summary, at each step n , we get $U_1, U_2, \dots, U_{n_{max}}$ and apply method \mathcal{M} and get the population to which to allocate next treatment and we update $U_1, U_2, \dots, U_{n_{max}+1}$. So, we continue with the adaptive rule to get U -data stream and ultimately stop when the *SPRT* between K_0 and K_1 stop. Hence, if we get *PICS* of that *SPRT*, we get *PICS* of the adaptive rule.

3.2 Details of Adaptive *SPRT* Rule

Suppose $\alpha = \mathbb{P}(\text{Type I error}) = PICS_I$ and $\beta = \mathbb{P}(\text{Type II error}) = PICS_{II}$. Let $a \approx \log\left(\frac{1-\beta}{\alpha}\right)$ and $b \approx \log\left(\frac{\beta}{1-\alpha}\right)$. We consider the following adaptive *SPRT* rule:

- We continue adaptive sampling if $b < \sum_{i=1}^{n_{\max}} \log\left(\frac{f_1(U_i)}{f_0(U_i)}\right) < a$.
- We stop adaptive sampling in favour of K_1 if $\sum_{i=1}^{n_{\max}} \log\left(\frac{f_1(U_i)}{f_0(U_i)}\right) \geq a$.
- We stop adaptive sampling in favour of K_0 if $\sum_{i=1}^{n_{\max}} \log\left(\frac{f_1(U_i)}{f_0(U_i)}\right) \leq b$.

Let ASN denote the average sample number of the proposed adaptive rule when coupled with the $SPRT$ stopping rule above, and let ASN_{K_0} and ASN_{K_1} denote the corresponding ASN s under K_0 and K_1 , respectively.

From Rao (1973) (pp. 479), we get approximate expressions for ASN of the $SPRT$ as

$$ASN_{K_0} \approx \frac{b(1 - \alpha) + a\alpha}{-\eta_x} \quad \text{and,} \quad ASN_{K_1} \approx \frac{b\beta + a(1 - \beta)}{-\eta_y}, \quad (2)$$

α and β are small (tend to 0) and accordingly $a \rightarrow \infty$ and $b \rightarrow -\infty$.

Hence, from Equation 2,

$$\begin{aligned} \log(PICS_{II}) &\approx -\eta_x \cdot ASN_{K_0} + o(ASN_{K_0}) \\ \log(PICS_I) &\approx \eta_y \cdot ASN_{K_1} + o(ASN_{K_1}). \end{aligned}$$

Thus, we get the expression for $\log(PICS)$ using our stopping rule and our selection procedure.

Remark 4. If we make $SPRT$ with X -data stream (or, Y -data stream) instead of U -data stream for hypotheses $K_0 : X \sim f_0$ Vs $K_1 : X \sim f_1$ (or, $K_0 : Y \sim f_0$ Vs $K_1 : Y \sim f_1$), we get similar expression for $\log(PICS_{II})$ and $\log(PICS_I)$. Only difference will be in ASN . In that case,

$$ASN \approx ASN_{K_0} + N_1^* \quad (\text{or, } \approx ASN_{K_1} + N_1^*) \quad (3)$$

where, $N_1^* = \lim_{n \rightarrow \infty} \mathbb{E}(N_{1,n})$.

Theorem 3.2 (Efficiency of the proposed selection procedure). *Under the assumptions in Section 2, the proposed adaptive $SPRT$ rule using $U_1, U_2, \dots, U_{n_{\max}}$ is efficient, and for the same value of $PICS$,*

$$\frac{ASN}{ASN_{K_0}} \rightarrow 1 \quad \text{or,} \quad \frac{ASN}{ASN_{K_1}} \rightarrow 1.$$

Proof. The proof can be directly followed from the Equation 3 for the adaptive $SPRT$ rule. \square

3.3 Example

We illustrate the results with an example of $f_0 \sim N(\theta_0, 1)$ (first data stream) and $f_1 \sim N(\theta_1, 1)$ (second data stream). Then with the notations of Section 3 we have, $\eta_x = \frac{1}{2} \cdot (\theta_0 - \theta_1)^2 > 0$ and $\eta_y = -\frac{1}{2} \cdot (\theta_0 - \theta_1)^2 < 0$ are the means of $Z_i^{(X)} = \log\left(\frac{f_0(X_i)}{f_1(X_i)}\right)$ and $Z_i^{(Y)} = \log\left(\frac{f_0(Y_i)}{f_1(Y_i)}\right)$ respectively. Also, the calculated variances for both are same, i.e., $\sigma_x^2 = \sigma_y^2 = (\theta_0 - \theta_1)^2$.

In this context, from [Theorem 3.1](#), the expression of N_1^* is,

$$\begin{aligned} N_1^* &\approx \sum_{m=1}^{\infty} \Phi\left(-\frac{\eta_x}{\sigma_x} \cdot \sqrt{m}\right) + \sum_{m=1}^{\infty} \Phi\left(\frac{\eta_y}{\sigma_y} \cdot \sqrt{m}\right) \\ &= \frac{1}{2} \left(\frac{\sigma_x^2}{\eta_x^2} + \frac{\sigma_y^2}{\eta_y^2} \right) \\ &= \frac{4}{(\theta_0 - \theta_1)^2} < \infty. \end{aligned}$$

Our stopping rule in case of samples from normal populations is to be adapted from the *SPRT* stopping rule given in the earlier section. After adapting the selection procedure in that context, we will find the example exhibited is efficient also.

Remark 5. With $X_i \sim N(-\theta_0, 1)$ and $Y_i \sim N(\theta_0, 1)$, we have, $\eta_x = 2\theta_0^2$ and $\eta_y = -2\theta_0^2$, whereas, $\sigma_x^2 = \sigma_y^2 = 4\theta_0^2$ for doing adaptive sequential testing as discussed earlier. Here, we can note that, the allocation rule in this context does not depend on parameters and it depends only on $\bar{U}_n = \frac{1}{n} \sum_{i=1}^n U_i$ and n . In that case, $N_1^* \approx \frac{1}{\theta^2} < \infty$.

Remark 6. Let f_0, f_1 be in the same MLR-family (with parameter θ) with \bar{X}_n and \bar{Y}_n as sufficient statistics respectively. Note that, for testing composite hypotheses $H'_0 : \theta \leq \theta_0$ Vs $H'_1 : \theta > \theta_1$, if we try to apply adaptive sequential rule as discussed earlier, allocation rule will depend on \bar{U}_n and parameters, and a, b, α, β depend on parameters. Then, $N_1^* \approx \frac{1}{2} \cdot \left(\frac{\sigma_x^2(\theta'_0, \theta'_1)}{\eta_x^2(\theta'_0, \theta'_1)} + \frac{\sigma_y^2(\theta'_0, \theta'_1)}{\eta_y^2(\theta'_0, \theta'_1)} \right)$, where $\theta'_0 \in H'_0$ and $\theta'_1 \in H'_1$. In that case, for $\theta'_0 \in H'_0$ and $\theta'_1 \in H'_1$, the worst (i.e., highest) $N_1^* \approx \frac{1}{2} \cdot \left(\frac{\sigma_x^2(\theta_0, \theta_1)}{\eta_x^2(\theta_0, \theta_1)} + \frac{\sigma_y^2(\theta_0, \theta_1)}{\eta_y^2(\theta_0, \theta_1)} \right)$.

4 Simulation

In this section, we present a comprehensive simulation analysis to assess the performance of the proposed adaptive *SPRT* procedure under a range of underlying distributional settings. For each configuration, we estimate the *PCS*, the expected number of allocations to the inferior population, and the *ASN*. Simulations are conducted under Normal, Poisson, and Asymmetric Laplace distributions, with 1000 replications in each scenario to ensure numerical stability. These results collectively convey the operational behaviour of the procedure across distinct distributional regimes.

Across all experiments, the decision thresholds are computed as $a = \log\left(\frac{1-\beta}{\alpha}\right)$, $b = \log\left(\frac{\beta}{1-\alpha}\right)$, for each specified pair (α, β) . Sampling commences when one observation is drawn independently from each population, forming the initial likelihood contributions. Subsequent sampling proceeds according to the adaptive allocation rule described in [Section 3.2](#). At each stage, the accumulated sample sizes from the two populations are examined, and the next observation is drawn from the population whose cumulative log-likelihood ratio $L_n = \sum_{i=1}^{n_{\max}} \log\left(\frac{f_0(U_i)}{f_1(U_i)}\right)$, provides weaker support for K_1 (hypothesis as defined in [Section 3.1](#)). This mechanism ensures that sampling is dynamically steered toward the superior population, thereby allowing the likelihood ratio to evolve in an efficient manner towards getting more and more samples from the superior population and highlighting the implication of the adaptive nature of the design.

After each new observation, the log-likelihood ratio is updated based on the stream currently yielding the larger sample size. The procedure terminates once the statistic crosses one of the two boundaries: the alternative K_1 is accepted when the statistic exceeds a , and the null K_0 is accepted when it falls below b . For every replication, we record whether the final decision corresponds to the truly superior population, the total number of observations drawn, and the frequency of allocations to the inferior population. Averages across replications yield the performance metrics PCS , $\mathbb{E}(N_{1,n})$, and ASN .

The subsequent subsections outline the distributional scenarios considered and summarize the corresponding numerical results.

4.1 Normal Distributions (Adaptive $SPRT$)

We first consider, $f_0 \sim N(\theta_0, 1)$, $f_1 \sim N(\theta_1, 1)$, with following mean pairs

$$(\theta_0, \theta_1) \in \{(0.1, 0), (0.2, 0), (0.3, 0), (0.4, 0), (0.5, 0)\}.$$

The adaptive $SPRT$ procedure continues to perform reliably, producing high PCS and maintaining small inferior allocations across these configurations. The expected number of allocations to the inferior population remains small and aligns closely with the theoretical benchmark N_1^* derived in Section 3.3. As the separation $|\theta_0 - \theta_1|$ increases, ASN declines substantially and as (α, β) decreases $\mathbb{E}(N_{1,n})$ increases to the finite limit N_1^* , reflecting good discrimination between the two populations. Detailed results are reported in Table 1.

4.2 Poisson Distributions (Adaptive $SPRT$)

Next, we consider $f_0 \sim P(\lambda_0)$, $f_1 \sim P(\lambda_1)$, with several contrasting parameter pairs

$$(\lambda_0, \lambda_1) \in \{(2.5, 2), (3, 2.5), (3.5, 2.5), (2, 1), (1.5, 0.5), (2.5, 1)\}.$$

The adaptive procedure continues to exhibit good behaviour under these discrete distributions, yielding high PCS values and maintaining modest inferior allocations. As in the Normal case, ASN decreases as the divergence between λ_0 and λ_1 widens. The corresponding numerical outcomes are displayed in Table 2.

4.3 Asymmetric Laplace Distributions (Adaptive $SPRT$)

To evaluate performance under skewed and asymmetric settings, we consider the Asymmetric Laplace family (Kotz et al., 2001)

$$f(x; m, \lambda, \kappa) = \frac{\lambda}{\kappa + \kappa^{-1}} \begin{cases} \exp\left(\frac{\lambda}{\kappa}(x - m)\right), & x < m, \\ \exp(-\lambda\kappa(x - m)), & x \geq m, \end{cases}$$

where m is a location parameter, $\lambda > 0$ is a scale parameter, and $\kappa > 0$ governs the degree of asymmetry.

A variety of contrasting parameter pairs $(m_0, \lambda_0, \kappa_0)$ and $(m_1, \lambda_1, \kappa_1)$ are examined. Across all configurations, the adaptive $SPRT$ continues to achieve strong PCS performance and small inferior allocations. As expected, ASN decreases monotonically with increasing separation between the distributions. The resulting performance measures are summarized in Table 3.

Table 1: Simulation results for two Normal populations with distinct mean pairs (θ_0, θ_1) and common variance $(\sigma_0^2, \sigma_1^2) = (1, 1)$ conducted within the adaptive *SPRT* framework.

(a) $(\theta_0, \theta_1) = (0.1, 0)$, $N_1^* = 400$				(b) $(\theta_0, \theta_1) = (0.2, 0)$, $N_1^* = 100$			
$\alpha(=\beta)$	<i>PCS</i>	$\mathbb{E}(N_{1,n})$	<i>ASN</i>	$\alpha(=\beta)$	<i>PCS</i>	$\mathbb{E}(N_{1,n})$	<i>ASN</i>
10^{-3}	0.908	349.550	1575.253	10^{-3}	0.918	91.237	418.085
5×10^{-5}	0.955	394.827	2289.517	5×10^{-5}	0.955	98.748	574.303
10^{-5}	0.969	391.828	2657.338	10^{-5}	0.979	97.999	663.545
5×10^{-6}	0.975	398.996	2756.844	5×10^{-6}	0.976	99.725	694.320
10^{-6}	0.986	394.102	3093.543	10^{-6}	0.984	100.309	781.425

(c) $(\theta_0, \theta_1) = (0.3, 0)$, $N_1^* = 44.444$				(d) $(\theta_0, \theta_1) = (0.4, 0)$, $N_1^* = 25$			
$\alpha(=\beta)$	<i>PCS</i>	$\mathbb{E}(N_{1,n})$	<i>ASN</i>	$\alpha(=\beta)$	<i>PCS</i>	$\mathbb{E}(N_{1,n})$	<i>ASN</i>
10^{-3}	0.909	38.350	180.370	10^{-3}	0.919	20.496	100.263
5×10^{-5}	0.969	38.488	252.804	5×10^{-5}	0.952	23.338	143.625
10^{-5}	0.974	46.189	296.015	10^{-5}	0.971	25.552	168.346
5×10^{-6}	0.975	43.336	309.886	5×10^{-6}	0.987	23.845	177.570
10^{-6}	0.989	40.482	351.835	10^{-6}	0.984	22.652	194.418

(e) $(\theta_0, \theta_1) = (0.5, 0)$, $N_1^* = 16$			
$\alpha(=\beta)$	<i>PCS</i>	$\mathbb{E}(N_{1,n})$	<i>ASN</i>
10^{-3}	0.930	13.422	66.488
5×10^{-5}	0.961	15.735	94.617
10^{-5}	0.980	14.324	105.348
5×10^{-6}	0.979	14.463	110.278
10^{-6}	0.985	15.161	124.351

4.4 Classical *SPRT*

For comparative context, we also examine the classical *SPRT* under the Normal distribution, where it is naturally applicable. It utilizes the same threshold pair (a, b) , i.e., same pair of (α, β) but employs deterministic alternating sampling from the two populations, without any adaptive allocation. The test statistic is $Z_n = \sum_{i=1}^n \log\left(\frac{f_1(X_i)}{f_0(X_i)}\right)$, and the sampling continues until Z_n crosses one of the stopping boundaries.

Although the classical *SPRT* often achieves reasonably small *ASN*, it necessarily allocates a substantial number of samples to the inferior population. This stands in sharp contrast with the adaptive *SPRT*, which significantly curtails inferior allocations by design. The numerical results (Table 4) show that the *PCS* of classical *SPRT* is higher than that of our adaptive *SPRT* method, although both methods are efficient.

Remark 7. $N_1^* \approx \frac{1}{2} \left(\frac{\sigma_x^2}{\eta_x^2} + \frac{\sigma_y^2}{\eta_y^2} \right)$ is the limit of $\mathbb{E}(N_{1,n})$, i.e., the expression for finite increasing limit of applications to the inferior treatment. For adaptive *SPRT*, in all the three sets of the above tables, we have found that $\mathbb{E}(N_{1,n})$ conforms to N_1^* as (α, β) decreases. There may be small fluctuations due to sampling error but overall $\mathbb{E}(N_{1,n})$ goes close to N_1^* .

Table 2: Simulation results for two Poisson populations with distinct mean pairs (λ_0, λ_1) conducted within the adaptive *SPRT* framework.

(a) $(\lambda_0, \lambda_1) = (2.5, 2)$, $N_1^* = 35.851$				(b) $(\lambda_0, \lambda_1) = (3, 2.5)$, $N_1^* = 43.879$			
$\alpha(=\beta)$	<i>PCS</i>	$\mathbb{E}(N_{1,n})$	<i>ASN</i>	$\alpha(=\beta)$	<i>PCS</i>	$\mathbb{E}(N_{1,n})$	<i>ASN</i>
10^{-3}	0.924	32.295	146.120	10^{-3}	0.909	39.151	176.486
5×10^{-5}	0.959	35.314	203.890	5×10^{-5}	0.971	41.286	246.989
10^{-5}	0.984	32.991	229.243	10^{-5}	0.980	42.927	286.710
5×10^{-6}	0.982	35.608	245.363	5×10^{-6}	0.978	43.234	303.215
10^{-6}	0.987	34.219	271.723	10^{-6}	0.986	42.506	333.556

(c) $(\lambda_0, \lambda_1) = (3.5, 2.5)$, $N_1^* = 11.888$				(d) $(\lambda_0, \lambda_1) = (2, 1)$, $N_1^* = 5.771$			
$\alpha(=\beta)$	<i>PCS</i>	$\mathbb{E}(N_{1,n})$	<i>ASN</i>	$\alpha(=\beta)$	<i>PCS</i>	$\mathbb{E}(N_{1,n})$	<i>ASN</i>
10^{-3}	0.931	10.940	49.169	10^{-3}	0.935	5.452	24.310
5×10^{-5}	0.967	10.668	66.835	5×10^{-5}	0.969	5.548	32.654
10^{-5}	0.980	11.002	76.338	10^{-5}	0.981	5.793	37.123
5×10^{-6}	0.987	10.611	80.323	5×10^{-6}	0.986	5.962	39.403
10^{-6}	0.996	10.525	89.907	10^{-6}	0.992	5.518	42.659

(e) $(\lambda_0, \lambda_1) = (1.5, 0.5)$, $N_1^* = 3.642$				(f) $(\lambda_0, \lambda_1) = (2.5, 1)$, $N_1^* = 2.911$			
$\alpha(=\beta)$	<i>PCS</i>	$\mathbb{E}(N_{1,n})$	<i>ASN</i>	$\alpha(=\beta)$	<i>PCS</i>	$\mathbb{E}(N_{1,n})$	<i>ASN</i>
10^{-3}	0.962	3.662	15.269	10^{-3}	0.952	2.996	12.614
5×10^{-5}	0.981	3.759	20.489	5×10^{-5}	0.977	3.093	16.832
10^{-5}	0.989	3.689	22.976	10^{-5}	0.986	3.204	19.170
5×10^{-6}	0.986	3.650	24.098	5×10^{-6}	0.990	2.991	19.950
10^{-6}	0.994	3.701	26.718	10^{-6}	0.989	3.104	21.790

Remark 8. Regarding *PCS* of adaptive *SPRT* in Tables 1, 2 and 3, we have seen that obtained *PCS* is little higher than $1 - \alpha(= 1 - \beta)$, though it is found that *PCS* tends to 1 as *ASN* increases (conforming that adaptive *SPRT* is efficient). In classical *SPRT* (Table 4), the value of *PCS* well coincide with $1 - \alpha(= 1 - \beta)$. Effective *ASN* of adaptive *SPRT* is less than that of classical *SPRT* by the finite amount N_1^* . Hence, though being efficient, adaptive *SPRT* gives little higher value of probability of incorrect selection.

4.5 Summary

Across all distributional regimes, the adaptive *SPRT* maintains high *PCS* while drastically reducing sampling from the inferior population. *ASN* decreases systematically with increasing signal strength, and the procedure remains stable across symmetric, discrete, and skewed scenarios. In the Normal case, the classical *SPRT* underscores the ethical and operational advantages of the adaptive allocation mechanism as while retaining comparable inferential accuracy, the adaptive *SPRT* dramatically reduces the expected inferior sample size. These findings collectively highlight the strong practical merits of the adaptive *SPRT* framework in sequential decision-making problems.

Table 3: Simulation results for two Asymmetric Laplace populations with various parameter configurations $(m_0, \lambda_0, \kappa_0)$ and $(m_1, \lambda_1, \kappa_1)$ conducted within the adaptive *SPRT* framework.

(a) $(m_0, \lambda_0, \kappa_0) = (0.2, 2, 0.7)$, $(m_1, \lambda_1, \kappa_1) = (0, 1, 0.3)$, $N_1^* = 2.288$ (b) $(m_0, \lambda_0, \kappa_0) = (0.2, 1, 0.8)$, $(m_1, \lambda_1, \kappa_1) = (0, 2, 0.2)$, $N_1^* = 4.802$

$\alpha(=\beta)$	<i>PCS</i>	$\mathbb{E}(N_{1,n})$	<i>ASN</i>
10^{-3}	0.844	1.936	11.701
10^{-5}	0.920	1.996	19.366
5×10^{-6}	0.949	1.978	20.703
10^{-6}	0.955	2.020	23.325
10^{-7}	0.964	2.060	27.126

$\alpha(=\beta)$	<i>PCS</i>	$\mathbb{E}(N_{1,n})$	<i>ASN</i>
10^{-3}	0.943	3.259	9.819
10^{-5}	0.985	3.493	12.693
5×10^{-6}	0.989	3.416	12.720
10^{-6}	0.994	3.364	13.489
10^{-7}	0.995	3.511	14.804

(c) $(m_0, \lambda_0, \kappa_0) = (0.4, 1, 0.6)$, $(m_1, \lambda_1, \kappa_1) = (0, 1, 0.2)$, $N_1^* = 4.576$ (d) $(m_0, \lambda_0, \kappa_0) = (0, 2, 0.7)$, $(m_1, \lambda_1, \kappa_1) = (0.2, 2, 0.3)$, $N_1^* = 2.774$

$\alpha(=\beta)$	<i>PCS</i>	$\mathbb{E}(N_{1,n})$	<i>ASN</i>
10^{-3}	0.893	2.793	15.669
10^{-5}	0.959	2.890	24.958
5×10^{-6}	0.969	2.812	26.276
10^{-6}	0.971	3.005	29.531
10^{-7}	0.975	3.046	33.611

$\alpha(=\beta)$	<i>PCS</i>	$\mathbb{E}(N_{1,n})$	<i>ASN</i>
10^{-3}	0.940	2.323	9.663
10^{-5}	0.975	2.492	14.553
5×10^{-6}	0.986	2.539	15.371
10^{-6}	0.989	2.571	17.121
10^{-7}	0.992	2.514	18.929

Table 4: Simulation results for two Normal populations with distinct mean pairs (μ_0, μ_1) and common variance $(\sigma_0^2, \sigma_1^2) = (1, 1)$ conducted within the *SPRT* framework.

$\alpha(=\beta)$	$(\mu_0, \mu_1) = (0.1, 0)$		$(\mu_0, \mu_1) = (0.2, 0)$		$(\mu_0, \mu_1) = (0.3, 0)$		$(\mu_0, \mu_1) = (0.4, 0)$		$(\mu_0, \mu_1) = (0.5, 0)$	
	<i>PCS</i>	<i>ASN</i>	<i>PCS</i>	<i>ASN</i>	<i>PCS</i>	<i>ASN</i>	<i>PCS</i>	<i>ASN</i>	<i>PCS</i>	<i>ASN</i>
10^{-2}	0.989	928.385	0.991	231.763	0.990	104.929	0.992	57.987	0.989	38.152
10^{-3}	0.999	1370.521	1.000	346.063	1.000	155.557	0.999	89.838	0.998	56.901
10^{-4}	0.999	1867.227	1.000	462.753	1.000	206.995	0.999	117.981	1.000	76.770
10^{-5}	1.000	2335.468	1.000	571.699	1.000	265.242	1.000	146.960	1.000	95.605

5 Concluding Remarks

This work, from a broader perspective, addresses a longstanding gap in adaptive sequential analysis — the lack of an explicit, finite-form quantification of ethical performance in allocation-driven testing procedures. The proposed likelihood ratio-based adaptive sequential rule for testing $H_0 : (f_0, f_1)$ vs $H_1 : (f_1, f_0)$, provides a direct analytical expression for the expected number of allocations to the less effective treatment and establishes its finiteness. The procedure also has significant implications for experimental design in practice. In clinical trials or adaptive testing problems where sample collection incurs real-world ethical or economic costs, having an explicit upper bound on the expected number of allocations to the inferior option offers clear interpretability for regulators and practitioners. The procedure also retains the asymptotic efficiency of the *SPRT*.

The study also opens up several directions for further research. In particular, when the parameters are completely unknown (composite hypotheses), it is of interest to design or modify the procedure to achieve a parameter-free decision rule while preserving the

essential goals of ethical and inferential efficiency. In this paper, we have achieved this property to some extent (as mentioned in Remark 5) for the symmetric case $H_0 : \theta = \theta_0$ Vs $H_1 : \theta = -\theta_0$. It is to be seen if the results can also be extended to settings involving more than two treatments, where similar finiteness and efficiency properties are expected to hold.

References

- Bandyopadhyay, U., Mukherjee, S., and Biswas, A. (2020). Adaptive two-treatment three-period crossover design for normal responses. *Brazilian Journal of Probability and Statistics*, 34(2):291 – 303. <https://doi.org/10.1214/18-BJPS427>.
- Berry, D. A. and Fristedt, B. (1985). *Bandit Problems: Sequential Allocation of Experiments*. Chapman and Hall, London.
- Bhandari, S. K., De, S. K., Mandal, S., Pradhan, S., and Ghosh, B. (2009). Study of optimal adaptive rule in testing composite hypothesis. *Sequential Analysis*, 28(3):394–405. <https://doi.org/10.1080/07474940903041852>.
- Bhandari, S. K., Dutta, R., and Niyogi, R. G. (2007). Study of optimal adaptive rule in testing problem. In *Advances in Multivariate Statistical Methods*. World Scientific.
- Biswas, A., Bandyopadhyay, U., and Bhattacharya, R. (2007). *Response-Adaptive Designs in Phase III Clinical Trials*. In *Statistical Advances in the Biomedical Sciences*, pages 33–53. John Wiley & Sons, Ltd.
- Biswas, A., Bhattacharya, R., and Das, S. (2020). A multi-treatment response adaptive design for ordinal categorical responses. *Statistical Methods in Medical Research*, 29(3):827–836.
- Das, R. (2024a). An optimal sequential design in ethical allocation with an adaptive interim analysis. *Journal of Nonparametric Statistics*, 36(3):547–572. <https://doi.org/10.1080/10485252.2023.2223322>.
- Das, S. (2024b). A two armed optimal response adaptive randomization for ordinal categorical responses with possible misclassification. *Sequential Analysis*, 43(2):131–146. <https://doi.org/10.1080/07474946.2023.2295264>.
- Das, S., Bhattacharya, R., and Biswas, A. (2023). Multi-arm covariate adjusted response adaptive designs for ordinal outcome clinical trials. *Statistical Methods in Medical Research*, 32(1):88–99.
- Friedman, L. M., Furgberg, C. D., and DeMets, D. L. (2010). *Fundamentals of Clinical Trials*. Springer.
- Ivanova, A. and Rosenberger, W. F. (2000). A comparison of urn designs for randomized clinical trials of $k > 2$ treatments. *Journal of Biopharmaceutical Statistics*, 10(1):93–107.
- Jain, S. and Jain, R. (1994). An application of spst for detecting change point in a reliability growth model. *Microelectronics Reliability*, 34(5):811–814. <https://www.sciencedirect.com/science/article/pii/002627149490006X>.

- Kotz, S., Kozubowski, T. J., and Podgórski, K. (2001). *Asymmetric Laplace Distributions*. In *The Laplace Distribution and Generalizations: A Revisit with Applications to Communications, Economics, Engineering, and Finance*, pages 133–178. Birkhäuser Boston. https://doi.org/10.1007/978-1-4612-0173-1_3.
- Kundu, S., Jha, J., and Bhandari, S. K. (2025). To study properties of a known procedure in adaptive sequential sampling design. *Sequential Analysis*, pages 1–13. <https://doi.org/10.1080/07474946.2025.2533328>.
- Li, C. and Nenes, G. (2024). Economic modelling and optimisation of sprt-based quality control schemes for individual observations. *International Journal of Production Research*, 62(24):8766–8789. <https://doi.org/10.1080/00207543.2024.2349252>.
- Li, J. X., Jeske, D. R., and Klein, J. A. (2012). Sequential analysis methodology for a poisson glmm with applications to multicenter randomized clinical trials. *Journal of Statistical Planning and Inference*, 142(12):3225–3234. <https://www.sciencedirect.com/science/article/pii/S0378375812001760>.
- Martens, M. J. and Logan, B. R. (2024). Statistical rules for safety monitoring in clinical trials. *Clinical Trials*, 21(2):152–161.
- Rao, C. R. (1973). *Linear Statistical Inference and its Applications: Second Edition*, pages 444–515. Wiley Series in Probability and Statistics. Wiley. <https://doi.org/10.1002/9780470316436.ch7>.
- Rasay, H. and Alinezhad, E. (2022). Developing an adaptable sequential probability ratio test applicable for lifetime analysis of different continuous distributions. *Quality Technology & Quantitative Management*, 19(4):511–530. <https://doi.org/10.1080/16843703.2021.2020954>.
- Rosenberger, W. F., Stallard, N., Ivanova, A., Harper, C. N., and Ricks, M. L. (2001). Optimal adaptive designs for binary response trials. *Biometrics*, 57(3):909–913. <https://doi.org/10.1111/j.0006-341X.2001.00909.x>.
- Teoh, J., Teoh, W., Khoo, M. B., Castagliola, P., and Moy, W. (2022). On designing an optimal sprt control chart with estimated process parameters under guaranteed in-control performance. *Computers & Industrial Engineering*, 174:108806. <https://www.sciencedirect.com/science/article/pii/S036083522200794X>.
- Wald, A. (1947). *Sequential Analysis*. Wiley Mathematical Statistics Series. John Wiley & Sons.