Fast Selection From Multiple Treatments: Online Appendix

August 3, 2026

A Lemma 1 and Theorem 1 Proofs

Before proving Theorem 1, which establishes our algorithm's statistical properties, we prove Lemma 1, which is instrumental in the proof of Theorem 1.

Proof of Lemma 1: From Equations (4) and (5), the sample weighted average corresponding to the i^{th} alternative is

$$\begin{aligned}
\bar{X}_{iN_i} &= \sum_{j=1}^{N_i} a_{ij} X_{ij}, \\
a_{i1} = a_{i2} = \ldots = a_{in_0} &= \frac{c_{1i}}{n_0}, \\
a_{i(n_0+1)} = \ldots = a_{iN_i} &= \frac{1-c_{1i}}{N_i - n_0}, \text{ and} \\
c_{1i} &= \frac{n_0}{N_i}\left(1 + \sqrt{1 - \frac{N_i}{n_0}\left(1 - \frac{(N_i - n_0)\psi_2^2}{s_{N_i}^2 d_1^2 h_1^2}\right)}\right). \qquad (1)
\end{aligned}$$

The expression is similar for $\bar{Y}_M$ involving weights $b_1, b_2, \ldots, b_M$. First of all, $c_{1i}, i = 1, 2, \ldots, K$ and c_{2i} are non-negative. Note that the term within the square root in the expression of c_{1i} is

$$1 - \frac{N_i}{n_0}\left(1 - \frac{(N_i - n_0)\psi_2^2}{s_i^2 d_1^2 h_1^2}\right) \geq 1 - \frac{N_i}{n_0}\left(1 - \frac{(N_i - n_0)}{N_i}\right) = 0 \ a.s. \qquad (2)$$

This is true since we have $N_i \geq \frac{s_i^2 d_1^2 h_1^2}{\psi_2^2}$ and $N_i > n_0$ almost surely from the stopping rule. Similarly, we can show that the square root in the c_2 expression is non-negative.

Next, we show that:

(i) $a_{i1}, a_{i2}, \ldots, a_{iN_i}$ adds up to 1 for each $i = 1, 2, \ldots, K$ and likewise $b_1, b_2, \ldots, b_M$ adds up to 1.

(ii) The denominators of Z_i and Z_Y are constants, that is $s_i^2 \sum_{j=1}^{N_i} a_{ij}^2 = \frac{\psi_2^2}{d_1^2 h_1^2}, i = 1, 2, \ldots, K$ and $s_M^2 \sum_{j=1}^{M} b_j^2 = \frac{\psi_2^2}{d_2^2 h_2^2}$.

Now, for $i = 1, 2, \ldots, K$,

$$\begin{aligned}\sum_{j=1}^{N_i} a_{ij} &= \sum_{j=1}^{n_0} a_{ij} + \sum_{j=n_0+1}^{N_i} a_{ij} \\ &= n_0 \frac{c_{1i}}{n_0} + (N_i - n_0)\frac{1 - c_{1i}}{N_i - n_0} = 1\end{aligned}$$

In the same way, it can be shown that $\sum_{j=1}^{M} b_j = 1$. Hence, part (i) is proved.

For ease of calculation, let $c_{1i} = \frac{n_0}{N_i}(1 + z)$, $z = \sqrt{1 - \frac{N_i}{n_0}\left(1 - \frac{(N_i - n_0)\psi_2^2}{s_{N_i}^2 d_1^2 h_1^2}\right)}$. Note that,

$$\sum_{j=1}^{N_i} a_{ij}^2 = \sum_{j=1}^{n_0} a_{ij}^2 + \sum_{j=n_0+1}^{N_i} a_{ij}^2 = \frac{c_{1i}^2}{n_0} + \frac{(1 - c_{1i})^2}{N_i - n_0} = \frac{N_i c_{1i}^2}{n_0(N_i - n_0)} + \frac{1}{N_i - n_0} - \frac{2c_{1i}}{N_i - n_0}$$

So,

$$\begin{aligned}(N_i - n_0)\sum_{j=1}^{N_i} a_{ij}^2 &= \frac{N_i c_{1i}^2}{n_0} + 1 - 2c_{1i} = \frac{n_0}{N_i}(1 + z^2 + 2z) + 1 - \frac{n_0}{N_i}(2 + 2z) = \frac{n_0}{N_i} z^2 - \frac{n_0}{N_i} + 1 \\ &= \frac{n_0}{N_i} - 1 + \frac{n_0(N_i - n_0)\psi_2^2}{N_i s_i^2 d_1^2 h_1^2} - \frac{n_0}{N_i} + 1 = (N_i - n_0)\frac{\psi_2^2}{s_i^2 d_1^2 h_1^2} \qquad (3)\end{aligned}$$

In the same way, it can be shown that holds $s_M^2 \sum_{j=1}^{M} b_j^2$ is constant. Thus, we prove parts (i) and (ii).

Let us consider the case when the population variances are known. The required optimal sample sizes are fixed and not random in such a situation. Then the optimal sample size from the i^{th} treatment and the selected are given by

$$C_i = \frac{d_1^2 h_1^2}{\psi_2^2}\sigma_i^2 \text{ for } i = 1, 2, \ldots, K \text{ and } C_S = \frac{d_2^2 h_2^2}{\psi_2^2}\sigma_S^2 \qquad (4)$$

The weights will also involve fixed values. Hence, applying Lyapunov CLT to the triangular array $a_{ij}(X_{ij} - \mu_i)$, $Z_{C_i} = \frac{(\bar{X}_{C_i} - \mu_i)}{\sqrt{\sigma_i^2 \sum_{j=1}^{C_i} a_{ij}^2}}$ asymptotically follows $N(0, 1), i = 1, 2, \ldots, K$. Similarly,

$Z_{C_S} = \frac{(\bar{Y}_{C_S} - \mu_{(K)})}{\sqrt{\sigma_S^2 \sum_{j=1}^{C_i} b_j^2}}$ asymptotically follows $N(0,1)$.

Now, the sample variances are U-statistics. So as $\psi_2 \to 0$, $N_i \to \infty$ a.s. and $s^2_{iN_i} \underset{a.s.}{\to} \sigma_i^2$. Using the stopping rule, we have

$$\frac{d_1^2 h_1^2}{\psi_2^2} s^2_{iN_i} \leq N_i \leq n_0 I(N_i = n_0) + \frac{d_1^2 h_1^2}{\psi_2^2} \left(s^2_{i(N_i-1)} + \frac{1}{N_i - 1} \right) \tag{5}$$

Dividing both sides of equation (5) by C_i, we have

$$\frac{s^2_{iN_i}}{\sigma_i^2} \leq \frac{N_i}{C_i} \leq \frac{n_0}{C_i} + \frac{s^2_{i(N_i-1)}}{\sigma_i^2} + \frac{1}{\sigma_i^2 (n_0 - 1)} \tag{6}$$

Taking limit as $\psi_2 \to 0$, both the LHS and RHS of the equation (6) converge to 1 almost surely, i.e., $N_i/C_i \underset{a.s.}{\to} 1$.

Let us define $n_1 = (1-\rho)C_i$ and $n_2 = (1+\rho)C_i$ for $0 < \rho < 1$. Note that $Z_{N_i} = Z_{C_i} + (Z_{N_i} - Z_{C_i})$. To prove that $Z_{N_i} \underset{d}{\to} N(0,1)$ follows the standard normal distribution, it is enough to show that $(Z_{N_i} - Z_{C_i}) \underset{P}{\to} 0$ since $Z_{C_i} \underset{d}{\to} N(0,1)$.

Fix some $\epsilon > 0$ we have

$$\begin{aligned} P\left(|Z_{N_i} - Z_{C_i}| > \epsilon\right) &\leq P\left(|(Z_{N_i} - Z_{C_i}| > \epsilon, |N_i - C_i| < \rho C_i\right) + P(|N_i - C_i| > \rho C_i) \\ &\leq P\left(\underset{n_1 < n < n_2}{max} |(Z_n - Z_{C_i}| > \epsilon \right) + P(|N_i - C_i| > \rho C_i) \end{aligned} \tag{7}$$

Since the uniform continuity in probability holds, so for given $\epsilon > 0$, there exist $\eta > 0$ and $\psi_0 > 0$ such that

$$P\{ \underset{n_1 < n < n_2}{max} |(Z_n - Z_{C_i}| > \epsilon\} < \eta \text{ for all } \psi_2 < \psi_0. \tag{8}$$

Thus, using the above uniform continuity in probability condition and the fact that $N_i/C_i \underset{P}{\to} 1$, we conclude that $Z_{N_i} - Z_{C_i} \underset{P}{\to} 0$ and hence $Z_{N_i} \underset{d}{\to} N(0,1)$. Thus, the Lemma 1 holds.

Proof of Theorem 1 First we prove that the probability of correct selection converges to $1 - \beta$. Let $P_{(i)}$ denote the population having population mean $(\mu_{(i)})$ with corresponding weighted sample mean $\bar{X}_{(i)}$. The subscripts (i) index populations ordered by their true means $\mu_{(1)} \leq \ldots \leq \mu_{(K)}$ and

do not denote order statistics of the observed data. Consequently, $\bar{X}_{(1)}, \ldots, \bar{X}_{(K)}$ are sample means from independent populations, and $\bar{Y}_M$ is obtained from an independent second-stage sample.

The procedure selects the population $P_{(K)}$ and rejects the null hypothesis if and only if

$$\bar{X}_{(K)} > \bar{X}_{(i)}, \quad i = 1, \ldots, K-1, \tag{9}$$

$$\bar{X}_{(K)} > \mu_0 + \frac{\psi_2}{d_1}, \tag{10}$$

$$\text{and } \frac{\bar{X}_{(K)} + \bar{Y}_M}{2} > \mu_0 + \frac{\psi_2}{d_2}. \tag{11}$$

Define the standardized statistics

$$W_i = \frac{\bar{X}_{(i)} - \mu_{(i)}}{\psi_2/(d_1 h_1)} \text{ for } i = 1, 2, ..., K, \text{ and } W_Y = \frac{\bar{Y}_M - \mu_{(K)}}{\psi_2/(d_2 h_2)}.$$

As noted earlier, since the indexing (i) is determined by the fixed parameters $\mu_{(1)} \leq \ldots \leq \mu_{(K)}$ and not by the data, the sample means $\bar{X}_{(1)}, \ldots, \bar{X}_{(K)}$ and $\bar{Y}_M$ are independent, and hence $W_1, \ldots, W_K, W_Y$ are mutually independent. Using conditions in (11), we have,

$$\begin{aligned} W_K &> W_i - \frac{d_1 h_1}{\psi_2}(\mu_{(K)} - \mu_{(i)}), \quad i = 1, \ldots, K-1, \\ W_K &> \frac{d_1 h_1}{\psi_2}\left(\frac{\psi_2}{d_1} + \mu_0 - \mu_{(K)}\right), \text{ and} \\ \frac{\psi_2}{2d_1 h_1} W_K + \frac{\psi_2}{2d_2 h_2} W_Y &> \frac{\psi_2}{d_2} + \mu_0 - \mu_{(K)}. \end{aligned} \tag{12}$$

Now, under LFC, $\mu_{(K)} - \mu_{(i)} = \psi_2 - \psi_1$ and $\mu_{(K)} - \mu_0 = \psi_2$. Thus, we have,

$$\begin{aligned} &P(\text{Selecting } P_{(K)} \mid LFC) \\ &= P\left(W_K > W_i - \frac{d_1 h_1}{\psi_2}(\mu_{(K)} - \mu_{(i)}), i = 1, 2, ..., K-1; W_K > \frac{d_1 h_1}{\psi_2}\left(\frac{\psi_2}{d_1} + \mu_0 - \mu_{(K)}\right); \right. \\ &\quad \left. \frac{\psi_2}{2d_1 h_1} W_K + \frac{\psi_2}{2d_2 h_2} W_Y > \mu_0 - \mu_{(K)} + \frac{\psi_2}{d_2} \mid LFC\right) \\ &= P\left(W_K > W_i - \frac{d_1 h_1}{\psi_2}(\psi_2 - \psi_1), i = 1, 2, ..., K-1; W_K > \frac{d_1 h_1}{\psi_2}(\psi_2/d_1 - \psi_2); \right. \\ &\quad \left. \frac{\psi_2}{2d_1 h_1} W_K + \frac{\psi_2}{2d_2 h_2} W_Y > \frac{\psi_2}{d_2} - \psi_2\right) \end{aligned} \tag{13}$$

Conditioning on W_K, Equation (13) becomes

$$\int_{h_1(1-d_1)}^{\infty} P\left(W_i < z + \frac{d_1 h_1}{\psi_2}(\psi_2 - \psi_1), i = 1, 2, ..., K-1; W_Y > \frac{2d_2 h_2}{\psi_2}\left(\frac{\psi_2}{d_2} - \psi_2 - \frac{\psi_2 t}{2d_1 h_1}\right)\right) \times$$

$$f_{W_K}(z)dz$$

$$= \int_{h_1(1-d_1)}^{\infty} \Pi_{i=1}^{K-1} P\left(W_i < z + \frac{d_1 h_1}{\psi_2}(\psi_2 - \psi_1)\right) P\left(W_Y < -2h_2 + 2h_2 d_2 + \frac{d_2 h_2 z}{d_1 h_1}\right) f_{W_K}(z)dt \tag{14}$$

By Lemma 1 and the Dominated Convergence Theorem,

$$P(\text{Selecting } P_{(K)} \mid \text{LFC}) \longrightarrow \int_{h_1(1-d_1)}^{\infty} \Phi^{K-1}\left(z + \frac{d_1 h_1}{\psi_2}(\psi_2 - \psi_1)\right) \times$$
$$\Phi\left(-2h_2 + 2h_2 d_2 + \frac{d_2 h_2 z}{d_1 h_1}\right) \phi(z)\, dz = 1 - \beta. \tag{15}$$

Next, we show that the probability of rejection under the null hypothesis converges to α. Under the null hypothesis, $H_0 : \mu_1 = \cdots = \mu_K = \mu_0$. Hence, $\mu_{(K)} - \mu_{(i)} = 0, \mu_{(K)} - \mu_0 = 0$.

The Type I error probability (size) of the procedure is $P(\text{the procedure rejects } H_0 \mid H_0)$ and the rejection of null hypothesis occurs only after an alternative is selected. So, if A_i denotes the event that P_i is selected and the rejection criteria are satisfied, then

$$P(\text{The procedure rejects } H_0 \mid H_0) = P\left(\bigcup_{i=1}^{K} A_i \,\middle|\, H_0\right) = \sum_{i=1}^{K} P(A_i \mid H_0) = K\, P(A_1 \mid H_0) \tag{16}$$

The above is true since exactly one treatment can be selected, the events $A_1, \ldots, A_K$ are disjoint. Moreover, under H_0, all treatments are statistically identical, implying $P(A_1 \mid H_0) = \cdots = P(A_K \mid H_0)$.

The event A_1 occurs if and only if

$$\bar{X}_1 > \bar{X}_j, \quad j = 2, \ldots, K, \bar{X}_1 > \mu_0 + \frac{\psi_2}{d_1}, \frac{\bar{X}_1 + \bar{Y}_M}{2} > \mu_0 + \frac{\psi_2}{d_2}. \tag{17}$$

Define the standardized statistics

$$W_i = \frac{\bar{X}_i - \mu_0}{\psi_2/(d_1 h_1)}, \quad i = 1, \ldots, K, \qquad W_Y = \frac{\bar{Y}_M - \mu_0}{\psi_2/(d_2 h_2)}.$$

Then using Equation (17), the event A_1 is equivalently expressed as

$$W_1 > W_j, \quad j = 2, \ldots, K, W_1 > h_1, \frac{\psi_2}{2d_1h_1}W_1 + \frac{\psi_2}{2d_2h_2}W_Y > \frac{\psi_2}{d_2}. \tag{18}$$

By Lemma 1, $(W_1, \ldots, W_K, W_Y) \xrightarrow{d} (Z_1, \ldots, Z_K, Z_Y)$, where $Z_1, \ldots, Z_K, Z_Y$ are independent standard normal random variables. Therefore,

$$\begin{aligned} P(A_1 \mid H_0) \to & P\bigg(Z_1 > Z_j,\ j = 2, \ldots, K;\ Z_1 > h_1;\ \frac{2}{d_1h_1}Z_1 + \frac{2}{d_2h_2}Z_Y > 2\bigg) \\ &= \int_{h_1}^{\infty} P(Z_j < z,\ j = 2, \ldots, K)\, P\bigg(Z_Y > \frac{d_2h_2}{h_1d_1}z - 2h_2\bigg)\, \phi(z)\, dz \\ &= \int_{h_1}^{\infty} \Phi^{K-1}(z)\, \Phi\bigg(\frac{h_2d_2}{h_1d_1}z - 2h_2\bigg)\, \phi(z)\, dz. \end{aligned} \tag{19}$$

Combining Equations (16) and (19), the asymptotic size control of the procedure is achieved. □

B Weights: Derivation and Discussion

B.1 Derivation of c_{1i} in Equation (5)

Here, we set the $Var(\bar{X}_{iN_i})$ equal to $\frac{\psi_2^2}{d_1^2h_1^2}$. Thus,

$$\mathrm{Var}(\bar{X}_{iN_i}) = \sigma_i^2 \sum_{j=1}^{N_i} a_{ij}^2 = \sigma_i^2 \left(\frac{c_{1i}^2}{n_0} + \frac{(1 - c_{1i})^2}{N_i - n_0}\right) = \frac{\psi_2^2}{d_1^2h_1^2}.$$

Notice, σ_i^2 is unknown, hence, we estimate by $s_{N_i}^2$. Thus, we have

$$\frac{c_{1i}^2}{n_0} + \frac{(1 - c_{1i})^2}{N_i - n_0} = \frac{\psi_2^2}{s_{N_i}^2 d_1^2 h_1^2}. \tag{20}$$

With some algebraic steps, we have

$$\begin{aligned} & c_{1i}^2(N_i - n_0) + (1 - c_{1i})^2 n_0 = \frac{\psi_2^2}{s_{N_i}^2 d_1^2 h_1^2} n_0(N_i - n_0) \\ & \implies N_i c_{1i}^2 - 2n_0 c_{1i} + \left[n_0 - \frac{\psi_2^2}{s_{N_i}^2 d_1^2 h_1^2} n_0(N_i - n_0)\right] = 0. \end{aligned} \tag{21}$$

Solving Equation (21) and taking into account the fact that $0 < c_{1i} < 1$, we obtain

$$c_{1i} = \frac{2n_0 + \sqrt{(2n_0)^2 - 4N_i\left(n_0 - \frac{\psi_2^2}{s_{N_i}^2 d_1^2 h_1^2} n_0(N_i - n_0)\right)}}{2N_i}$$
$$= \frac{n_0}{N_i}\left(1 + \sqrt{1 - \frac{N_i}{n_0}\left(1 - \frac{(N_i - n_0)\psi_2^2}{s_{N_i}^2 d_1^2 h_1^2}\right)}\right),$$

B.2 Discussion on Sample Mean Weights a_{ij} and b_j

The weights a_{ij} are structured such that the first n_0 observations and the remaining $N_i - n_0$ observations receive different allocations as in Equation 5. We note that $c_{1i} \in (0, 1)$ determines the proportion of total weight assigned to the initial subset of the data. The value of c_{1i} is derived analytically to ensure that the variance of the weighted means $\bar{X}_{iN_i}$ are constant and equal to the pre-specified value ψ_2^2 across all alternatives, that is, $Var(\bar{X}_{1i}) = \psi_2^2$ for all $i = 1, 2, \ldots, K$. We now detail the importance of this weighting mechanism, and hence of the constants c_{1i}s from various dimensions.

Setting the target variance of the estimator to ψ_2^2 serves a critical design and decision-theoretic purpose. It is crucial to estimate the performance of each treatment with comparable precision to enable fair and accurate comparisons. By choosing c_{1i} to equalize the variance of each $\bar{X}_{iN_i}$ to ψ_2^2, the method uses c_{1i} as an adaptive weighting factor that balances the contributions of pilot stage data and the remaining data based on their sample variance and target variance and the pilot and the final sample size. This adjustment ensures that comparison of weighted means reflects true comparison of treatment effects rather than mere variations in sample size or variability.

Moreover, in treatment selection, $\mu_0 + \psi_2$ is the threshold above which a treatment is deemed acceptable. If the standard deviation of the estimator exceeds ψ_2, even treatments whose true means surpass this threshold may fail to appear effective due to sampling variability. Conversely, treatments with true means below the threshold could erroneously appear superior purely by chance. By setting the estimator's variance to ψ_2^2, the procedure calibrates estimation precision to the decision context, ensuring that one standard error corresponds exactly to the practical significance threshold. This alignment guarantees that the sampling distribution is narrow enough to detect effects of size ψ_2 with high probability, while also avoiding unnecessary oversampling beyond what

is needed to distinguish practically meaningful differences.

Additionally, selecting c_{1i} to equalize variance achieves homoscedasticity, which was essential for deriving the Random Central Limit Theorem (Lemma 1) needed to control error probabilities for the proposed procedure. Similar variance-standardizing weight adjustments have been employed in previous selection and ranking frameworks, such as those described in Dudewicz and Nelson (2003) and Nelson and Dudewicz (2002). A similar justification applies to the weighting scheme used in Phase 2 as well.

B.3 Dependence Between Weights and Data in Sequential Designs

We note that in the realm of sequential analysis, the analysis of data is done on the fly and a decision is taken as and when an optimal number of observations arrive. Consequently, the final sample size becomes a random variable that depends on the observed data. Even for the simple sample mean, when the sample size is random, the weight assigned to each observation (that is, the reciprocal of the sample size, $1/N_i$) is data-dependent. This introduces a correlation between the weights and the observations themselves. Such dependence makes the derivation of moments, consistency, and asymptotic normality properties more mathematically involved compared to fixed sample size settings. For a detailed exposition on these technical complexities, one can refer to Chapters 6–14 of Mukhopadhyay and De Silva (2009), where the theory of sequential estimation is rigorously treated. Similarly, in the current weighted mean setup, $s^2_{N_i}$ is an estimate from the data, and therefore c_{1i}, and subsequently the weights a_{ij}, are also random variables as they depend on the sample. Despite this dependence, the design remains valid because the construction aims to achieve constant variance across populations, which is utilized to derive the Random Central Limit Theorem in Lemma 1 in the Appendix section under the proposed framework. This asymptotic result underpins the validity of subsequent hypothesis testing procedures, ensuring that Type I Error and power are controlled even though the weights and data are dependent.

C Simulation Details and Additional Simulation Results

This section contains additional technical notes on the implementation of our method and the comparison methods in the simulation studies (C.1), followed by a table extending the parameter

space explored in Section 5.1 (C.2) and simulation results for comparisons under non-normally distributed conditions (C.3). Finally, there is a simulation study comparing the performance of our method with that of Rollin and Chen (2005) (C.4).

C.1 Implementation Details

C.1.1 Evaluation of Our Method in Table 2

Sample means $\bar{n_1}$, $\bar{n_2}$, and $\bar{n_3}$ were always found to be nearly equal regardless of parameter condition ($\boldsymbol{\theta}_0$ or $\boldsymbol{\theta}_{LFC}$), so the presented values were calculated using the 10,000 samples from the null condition, $\boldsymbol{\theta}_0$. Phase II sample sizes are typically equal to 0 in the condition $\boldsymbol{\theta}_0$, as most of the $\boldsymbol{\theta}_0$ simulations do not proceed to Phase II. Thus, the simulation estimate of m comes strictly from the $\boldsymbol{\theta}_{LFC}$ condition.

C.1.2 Implementation of Comparison Methods

The KN++ algorithm is designed to handle dependence across samples over time, which is not present in our setting. However, it still performs competitively. For settings with dependence over time, this approach may be superior to the other approaches compared.

Because lil'UCB makes the strong assumption that variances are known, we give it an advantage by providing the true variance of each alternative treatment. This is a key component of the challenge of this setting, making lil'UCB less useful in this context. lil'UCB additionally requires the input of two hyperparameters — we use the values recommended by the authors.

Neither the KN++ nor the lil'UCB algorithms define a standard or baseline performance in their settings. However, by generating a zero-variance treatment arm with $\mu = \mu_0$, both algorithms are able to compare the alternative treatments with this baseline. The number of samples drawn from this zero-variance arm are naturally not included in the sample size evaluations.

C.1.3 Additional Notes on Normal Comparisons Results

In Figure 3, lil'UCB selects the best treatment in every simulated case. However, this comes at the cost of substantial over-sampling relative to the other methods. We do not simulate it in the smaller best-treatment cases due to its larger sample sizes which become extremely time-consuming.

In Figure 6, Lil'UCB exhibits very little bias for the best treatment, but this is a result of its relatively large sample size. Bias often comes from not sampling a treatment because of chance negative errors in performance (Nie et al. 2018). Thus, if the best treatment is always selected and heavily sampled, then no bias is expected in the estimation of that arm. In contrast, lil'UCB's estimates of the inferior treatments are heavily biased downward (Figure 7).

C.2 Sample Sizes and Error Rates Across Additional Parameter Settings for Our Method

Table C.1 shows additional parameter conditions, extending the parameter space explored in Section 5.1. In each condition simulated, the equations in Theorem 1 were solved for a set of experimenter-chosen parameters (α, β, K, ψ_1, and ψ_2) and data was generated either under the null hypothesis ($\boldsymbol{\theta}_0 : \mu_1 = \mu_2 = \mu_3 = \ldots = \mu_0$) or the alternative hypothesis with the least favorable configuration ($\boldsymbol{\theta}_{LFC} : \mu_0 \leq \mu_{(1)} = \mu_{(2)} = \ldots = \mu_0 + \psi_1 < \mu_0 + \psi_2 = \mu_{(K)}$). In both cases, the data-generating process is assumed to be normal.

C.3 Non-Normally Distributed Comparisons

Here we show additional performance comparisons under non-normally distributed conditions.

C.3.1 Bernoulli Distribution

Digital experimentation often focuses on Bernoulli outcomes, where treatments seek to affect the chances of individuals making a binary decision. In these settings, the probability of a positive outcome p (e.g., clickthrough, conversion, etc.) tends to be small even under effective treatments. This setting can be challenging for methods relying on the Central Limit Theorem, as it takes longer for the distribution of the sample mean in this setting to converge to the normal distribution. Because of this challenge, we compare performance in this setting. For each condition, the p for $K - 1$ alternatives and the baseline are set equal to 0.005, while a single alternative has some shift above the baseline.

Several of the methods we compare with either assume normality of outcomes or rely on a variant of the CLT (as ours does). For each, we use a larger starting sample size ($n_0 = 100$) to help the CLT engage before sequential sampling begins. We also compared with the median elimination

ψ_1	ψ_2	K	α	$1-\beta$	σ_1	σ_2	σ_3	$\bar{n_1}$ $s(n_1)$	$\bar{n_2}$ $s(n_2)$	$\bar{n_3}$ $s(n_3)$	$\bar{m}$ $s(m)$	**Type I Error**	P(CS)
0.1	0.5	2	0.05	0.8	3	3		292.6 (24.4)	293.7 (17.7)		50.6 (16.8)	0.0539	0.790
0.1	0.5	2	0.05	0.8	10	3		3260.8 (81.5)	294.0 (7)		553 (182)	0.0517	0.799
0.1	1.0	2	0.05	0.8	3	3		73.2 (12.9)	74.5 (10)		28 (7.5)	0.0531	0.799
0.1	1.0	2	0.05	0.8	10	3		828.6 (40.5)	75.2 (4.2)		124 (39.4)	0.0491	0.804
0.0	0.5	3	0.05	0.8	3	3	3	389.1 (28.2)	390.5 (19.9)	391.1 (16.6)	36.6 (10.8)	0.0490	0.799
0.0	0.5	3	0.05	0.8	3	5	10	389.4 (28.4)	1085.0 (40.3)	4336.8 (81.5)	39 (30)	0.0475	0.792
0.0	0.5	3	0.05	0.8	10	3	5	4338.7 (94.7)	391.0 (8.15)	1085.4 (20.2)	400 (103)	0.0539	0.797
0.0	1.0	3	0.05	0.8	3	3	3	96.3 (14.5)	97.8 (10.7)	98.1 (9.15)	28.6 (6.36)	0.0494	0.794
0.0	1.0	3	0.05	0.8	3	5	10	96.2 (14.6)	271.0 (21.0)	1084.6 (41.1)	29.1 (9.05)	0.0525	0.797
0.0	1.0	3	0.05	0.8	10	3	5	1084.1 (47.2)	98.3 (4.19)	271.9 (10.3)	100 (27.5)	0.0483	0.794
0.1	0.5	3	0.05	0.8	3	3	3	379.1 (27.7)	380.6 (19.7)	380.6 (16.2)	44.9 (12.7)	0.0494	0.795
0.1	0.5	3	0.05	0.8	3	5	10	379.7 (27.7)	1056.4 (39.9)	4226.8 (80.2)	57.6 (73.3)	0.0488	0.800
0.1	0.5	3	0.05	0.8	10	3	5	4225.2 (91.5)	381.0 (7.97)	1057.1 (19.7)	486 (140)	0.0497	0.802
0.1	1.0	3	0.05	0.8	3	3	3	95.6 (14.5)	96.9 (10.8)	97.2 (9.06)	28.7 (6.05)	0.0501	0.797
0.1	1.0	3	0.05	0.8	3	5	10	96 (14.7)	268 (20.9)	1076 (40.7)	29.8 (11.7)	0.0534	0.793
0.1	1.0	3	0.05	0.8	10	3	5	1074.4 (46.3)	97.5 (4.07)	269.5 (10.2)	106 (29.9)	0.0453	0.798

Table C.1: Error and Sample Size Estimates Using Sequential Best Treatment Selection
(i) $\mu_0 = 0$ for all simulations
(ii) P(CS): Probability of Correct Selection, i.e., empirical rate of rejection of null hypothesis in Phase II for the true best alternative. For estimation, $\boldsymbol{\theta}_{LFC} = \{\mu_1 = \mu_0 + \psi_2,\ \mu_2 = \mu_3 = \mu_0 + \psi_1\}$
(iii) The column m contains average sample size for the second phase under $\boldsymbol{\theta}_{LFC}$. The second phase usually has zero or far fewer samples under $\boldsymbol{\theta}_0$.

algorithm Even-Dar et al. (2006) and the exponential gap algorithm Karnin et al. (2013). However, these are not included in the graphs because they sample many orders of magnitude more, as they sacrifice practical sampling for far more robust theoretical guarantees on error rates. In all of our simulations, they sampled sufficiently such that they never once failed to select the true best arm. For digital experimentation, this certainty of correct selection is not worth the enormous number of samples.

In Figures C.1 and C.2, our approach is shown to oversample somewhat in this condition, but the desired success rates are met or exceeded in each case. As the desired P(CS) approaches 1, our competitors have some difficulty maintaining the desired rates, and fall short to varying degrees,

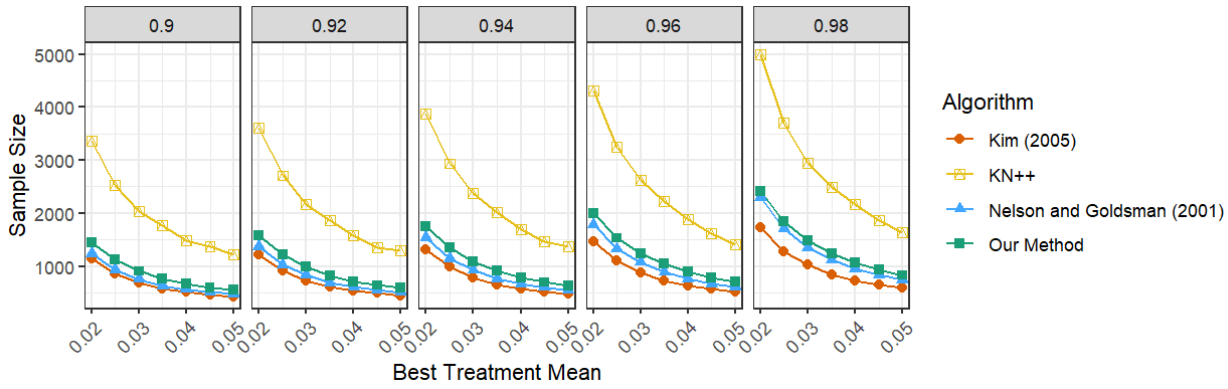

Figure C.1: Sample Size With Bernoulli-Distributed Outcomes

whereas our method's error rates approach the desired P(CS) from above. KN++ succeeds in almost all conditions to maintain the desired selection rate, although it samples substantially more.

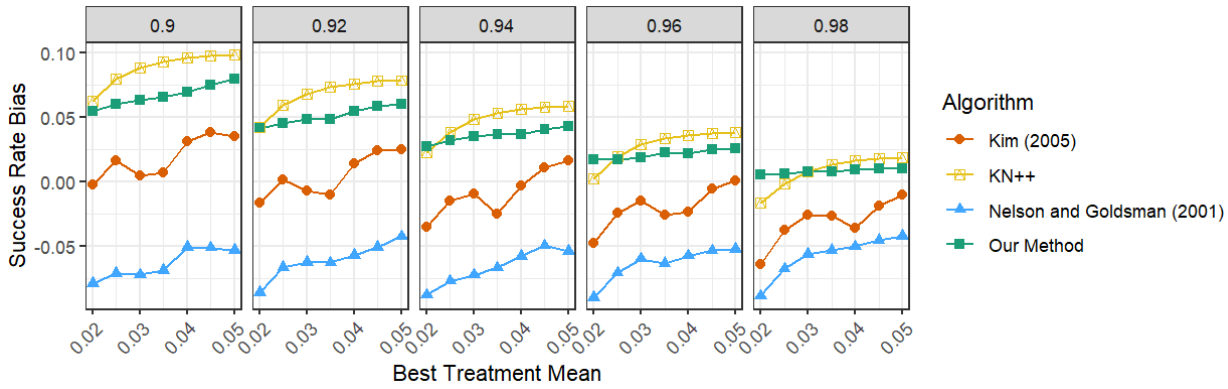

Figure C.2: P(CS) With Bernoulli-Distributed Outcomes

Figure C.3 shows the standard deviation of the final sample sizes for different methods. Figure C.4 shows the upper 95th percentile of sample size.

C.3.2 Gamma Distribution

We use the gamma distribution to simulate settings with skewed outcomes, showing that our method is still effective in this context. We include the algorithms of Kim (2005), Kim and Nelson (2006) and Nelson and Goldsman (2001) in this study.

The standard treatment mean, μ_0, is set to 1. The distributions used to sample the alternative treatments are shown in Figure C.5 with labels corresponding to the x-axis numbers in the gamma simulation figures (and larger numbers correspond to more rightward-shifted distributions and thus

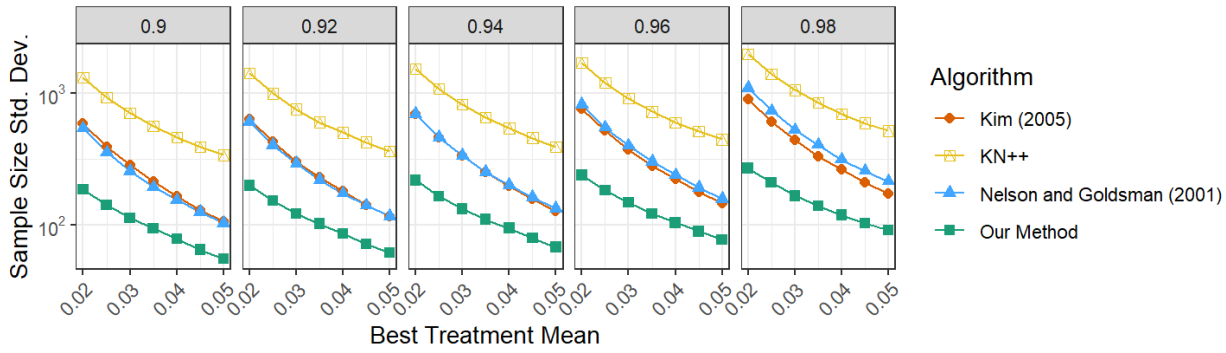

Figure C.3: Sample Size Standard Deviation for Binary Outcomes

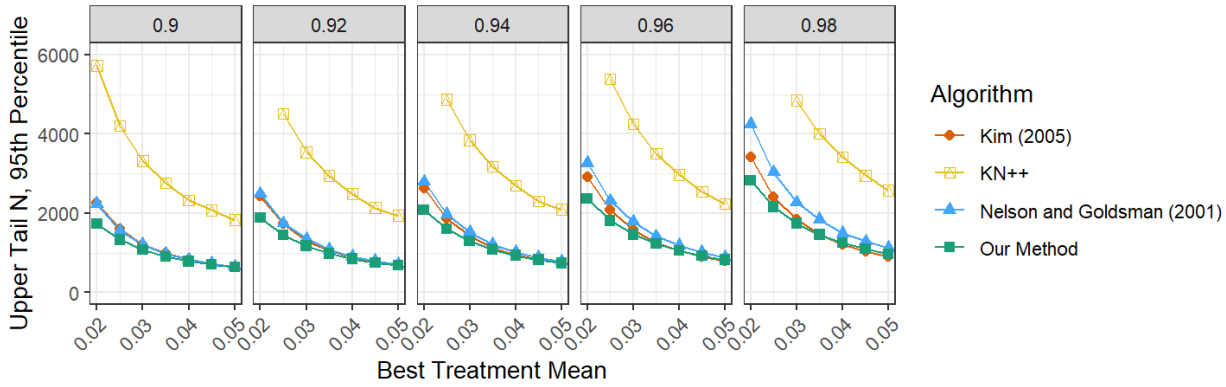

Figure C.4: Sample Size Upper 95th Percentile for Binary Outcomes

larger effect sizes). For all conditions, the inferior treatments are sampled from the $\alpha = 1, \beta = 1$ gamma, while the single superior treatment is varied across the other distributions depicted in Figure C.5. Overall, these simulations found that the algorithms we compared are not very affected by skewness, and in general, we observe that in the gamma-distributed simulations, the algorithms exhibit similar behavior as in the normally distributed conditions. Figures C.6, C.7 and C.8 show the results of the study.

C.4 Comparison With Rollin and Chen Best Population Selection Method

As noted earlier, Rollin and Chen (2005) propose a two-stage procedure that considers two stages in each phase. In the first stage of each phase, the population variance of the treatment is estimated using a pilot sample similar to ours, then the necessary sample size is estimated for each treatment, and all remaining samples are taken at once. A primary advantage of the sequential method over the two-stage procedure is that the final sample size is estimated using the sequential procedure to

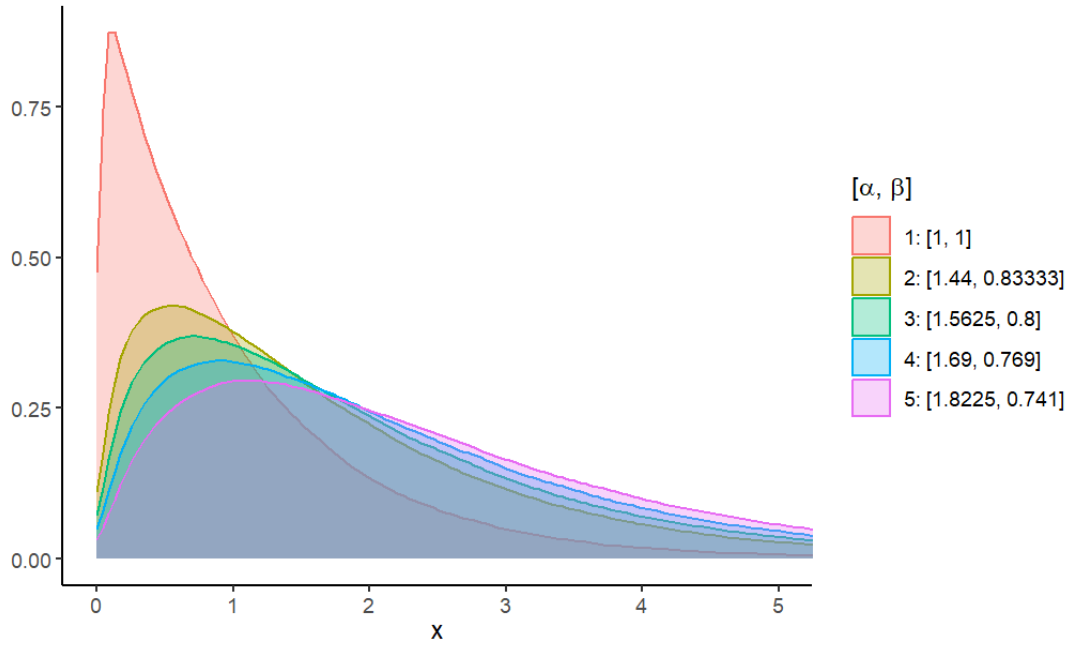

Figure C.5: Gamma Distributions Used in Simulations

be more precise. The benefit of this feature is demonstrated in the simulations described below.

Importantly, some of our contributions are also going to optimize Rollin and Chen's procedure. They give no recommendations for the selection of the design parameters d_1, d_2, h_1, and h_2. Our *dp-select* method, outlined in Section 4, can be applied to their problem. Our implementation of Rollin and Chen's method receives the benefits of this approach, as we solve their system of equations in the same way as ours. What remains to be compared is the mean and variance of the final sample sizes, where the sequential approach shows advantages.

For this comparison, we considered a small number of configurations, comparing the sequential method's sample size tendencies with the two-phase fixed-sample method of Rollin and Chen. For each simulation of the two-stage method, n_0 and m_0 were each set to 30, and the true effect size standard deviations were set equal to 10. The results are shown in Table C.2. For each combination of experiment parameters, method (Rollin and Chen two-step, Best treatment sequential), and population mean state ($\boldsymbol{\theta}_0$, $\boldsymbol{\theta}_{LFC}$), 10,000 simulations were run. For each configuration, the first treatment ($P = 1$) is the greatest under $\boldsymbol{\theta}_{LFC}$. This is why P_1 is sampled more in the $\boldsymbol{\theta}_{LFC}$ simulations with homogeneous variances — P_1 is nearly always selected as the best in Phase I and further sampled in Phase II.

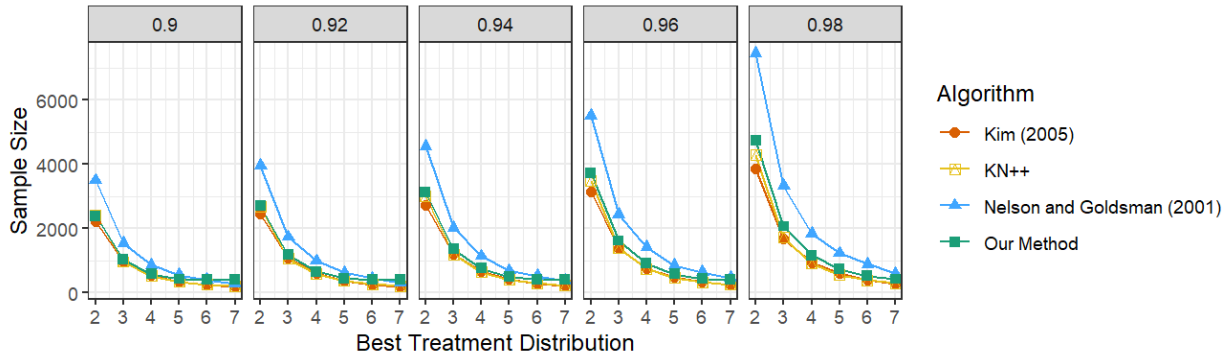

Figure C.6: Sample Size With Gamma-Distributed Outcomes

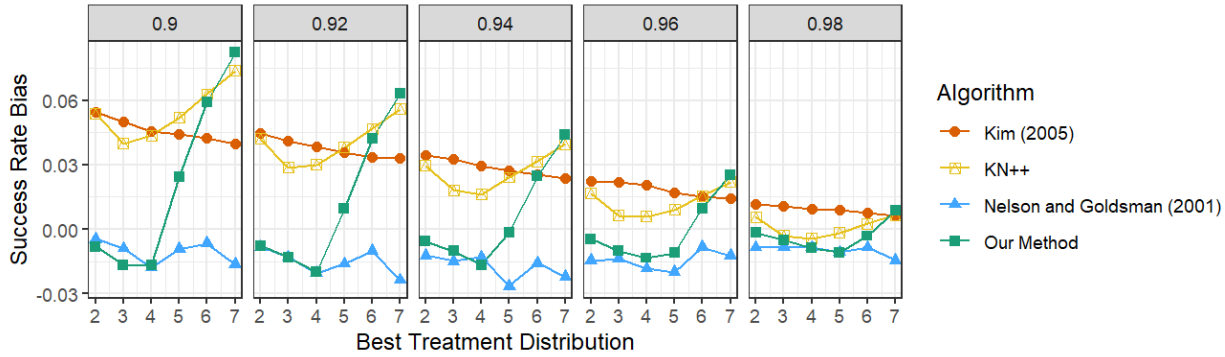

Figure C.7: P(CS) With Gamma-Distributed Outcomes

The methods performed comparably with respect to Type I error rate, while the sequential method occasionally overshot the probability of the correct selection requested. Our sequential method consistently outperforms Rollin and Chen's two-stage method with respect to sample size, and for some parameter configurations, Rollin and Chen conclude with a far larger sample size. Type I error rates for both methods fall well within Bradley's (1978) liberal criterion range of 0.025 to 0.075. It can be observed that the estimated final sample size depends on the estimate of the population variance. The final sample size of the sequential method has lower variance than the two-stage method, since the population variance(s) are estimated using a large sample for sequential procedure as opposed to the Rollin and Chen's two-stage procedure which estimates the necessary sample size using a sample of size equal to the pilot sample size only. Figure C.9 shows the first and second phase sample sizes joint empirical distribution for the first configuration in Table C.2, with data generated under θ_{LFC}. This illustrates the far greater spread in final sample sizes when the two-stage approach is used.

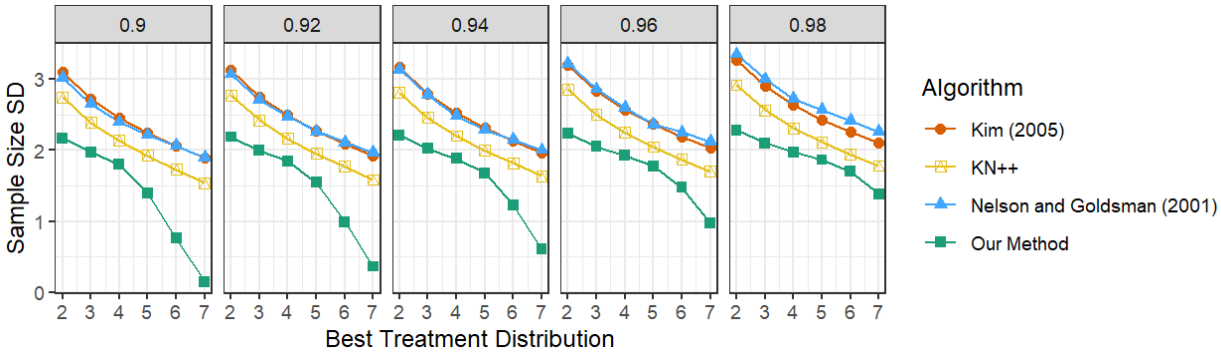

Figure C.8: Sample Size Std. Dev. With Gamma-Distributed Outcomes

Figure C.9: Phase I and II Sample Sizes From Repeated Simulations of a Single Experiment

D Practical Guidance for Implementing the Sequential Algorithm

In this section, we provide more guidance for how to think about selecting the parameter ψ_1 and then present a step-by-step stylized example of the method being implemented in practice.

D.1 Picking ψ_1

To readers unfamiliar with the IZ approach to R&S problems, ψ_1, which identifies the the lower end of the IZ, is likely the least immediately intuitive information that the experimenter must

ψ_1, ψ_2, K	α, $1-\beta$	**Type I Error**		**C.S. Rate**		P	σ	**Average Total Sample Size s(Total Sample Size)**			
								$\boldsymbol{\theta}_0$		$\boldsymbol{\theta}_{LFC}$	
		RC	**Seq**	**RC**	**Seq**			**RC**	**Seq**	**RC**	**Seq**
0.0, 0.5, 3	0.05, 0.8	0.0466	0.0517	0.81	0.80	1	10	2591	2425	4120	3860
								(1255)	(1052)	(1154)	(785)
						2	10	2587	2431	2041	1920
								(1247)	(1056)	(748)	(546)
						3	10	2582	2414	2036	1926
								(1243)	(1046)	(753)	(555)
0.0, 0.5, 3	0.05, 0.8	0.0528	0.0531	0.79	0.80	1	10	2582	2424	4131	3856
								(1247)	(1052)	(1155)	(790)
						2	7	1266	1182	1007	944
								(613)	(512)	(379)	(273)
						3	12	3746	3488	2940	2769
								(1818)	(1513)	(1066)	(795)
0.0, 1.0, 3	0.05, 0.8	0.0501	0.0479	0.81	0.80	1	10	642	603	1038	964
								(310)	(264)	(287)	(198)
						2	10	646	603	511	481
								(312)	(262)	(185)	(139)
						3	10	648	607	510	480
								(315)	(264)	(184)	(136)
0.0, 1.0, 3	0.05, 0.8	0.0521	0.0516	0.79	0.80	1	10	648	599	1032	964
								(314)	(261)	(294)	(199)
						2	7	319	298	252 (93)	235 (67)
								(154)	(129)		
						3	12	934	874	739	693
								(453)	(380)	(280)	(200)
0.1, 0.5, 3	0.05, 0.8	0.0496	0.0528	0.85	0.85	1	10	3049	2861	4751	4448
								(1337)	(1088)	(1204)	(725)
						2	10	3032	2862	2454	2324
								(1315)	(1086)	(793)	(511)
						3	10	3035	2854	2476	2320
								(1324)	(1083)	(801)	(502)
0.1, 0.5, 3	0.05, 0.8	0.0509	0.0491	0.85	0.85	1	10	3054	2866	4737	4463
								(1339)	(1091)	(1212)	(703)
						2	7	1496	1408	1207	1133
								(653)	(536)	(399)	(237)
						3	12	4379	4100	3541	3334
								(1932)	(1552)	(1162)	(709)
0.1, 1.0, 3	0.05, 0.8	0.0529	0.0508	0.83	0.82	1	10	694	653	1097	1027
								(323)	(270)	(292)	(191)
						2	10	690	648	552	518
								(321)	(267)	(192)	(131)
						3	10	694	650	552	519
								(324)	(268)	(190)	(132)
0.1, 1.0, 3	0.05, 0.8	0.0474	0.0504	0.82	0.82	1	10	691	647	1101	1023
								(324)	(268)	(293)	(195)
						2	7	337	317	271 (95)	255 (67)
								(157)	(131)		
						3	12	1004	937	794	745
								(467)	(386)	(277)	(186)

Table C.2: Comparison With Two-Stage Fixed-Sample Method
(i) RC: Rollin and Chen, Seq: Sequential Best Treatment Selection
(ii) Sample sizes and sample size standard deviations of the best treatment are emboldened.

provide. In short, for a given value of ψ_2, ψ_1 determines $\psi_2 - \psi_1$, how much worse the inferior treatments must be in order to distinguish them from the best treatment at the required rate. Note that a higher ψ_1 results in a smaller IZ. All else equal, bringing ψ_1 nearer to ψ_2 allows the test to distinguish the best treatment from increasingly effective/competitive inferior treatments at a given rate.

This margin serves two roles: first and foremost, it importantly controls the conditions where acceptable treatments are distinguished from unacceptable treatments. If an unacceptable treatment is selected over an acceptable treatment in the first Phase, then there is a risk that it is selected over the baseline, if the null hypothesis is rejected at the end of Phase II. This goes against the experimenter's preferences regarding acceptable treatments, as specified by their choice of ψ_2.

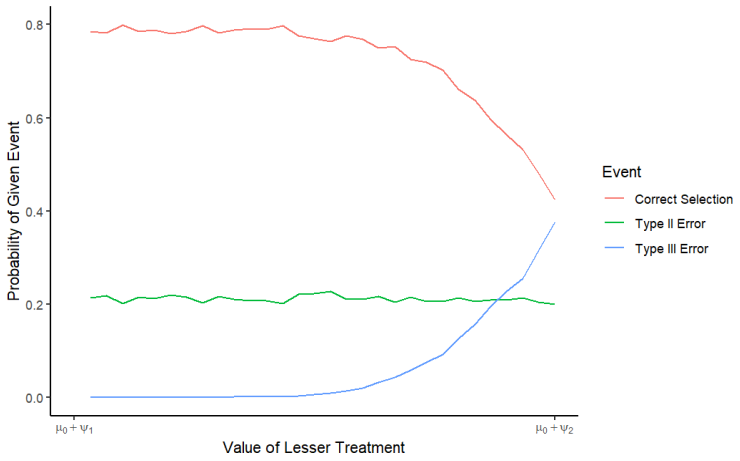

Figure D.10: Type III Error as Function of Lower Treatment's Mean in Indifference-Zone

This kind of mistake is sometimes called a Type III error (correctly rejecting H_0 but for the wrong reason).

Secondly, $\psi_2 - \psi_1$ is the margin at which treatments will be distinguished in general. Thus, if there are several acceptable treatments and they are nearer to each other than $\psi_2 - \psi_1$, then the experimenter has effectively expressed through the values of these parameters that distinguishing between these treatments is not worth the samples it costs.

In summary, given that $\mu_{(K)} \geq \mu_0 + \psi_2$ such that we should reject H_0, higher ψ_1 values will reduce the size of the indifference-zone, maintaining the desired power for a harder to detect case (higher-valued sub-optimal alternatives) but result in a higher sample size, all else equal. The experimenter should pick a ψ_1 for which Type III errors are less problematic (such that selecting a value between $\mu_0 + \psi_1$ and $\mu_0 + \psi_2$ is not so costly a mistake) and yet doesn't result in sample sizes which are so high as to be prohibitively expensive. Simultaneously, $\psi_2 - \psi_1$ can be specified such that it is the smallest difference among acceptable alternatives worth detecting.

Figure D.10 displays the consequences of an inferior alternative encroaching into the indifference-zone. For each simulation, the true best treatment was set equal to $\mu_0 + \psi_2$, and the value of the lesser treatment was varied across the IZ. When the lesser treatment is in the lower part of the indifference-zone, there is very little effect on the probability of correct selection. As the lesser treatment increases, the probability of correct selection decreases, as the lesser treatment is more often selected in Phase I, and is more often selected as the best treatment in Phase II. Note that as the value of the lesser alternative increases, likelihood of Type III error increases, while the cost

of the mistake decreases.

Importantly, the Type II error rate suffers little (if at all) across the graph. This means that, in the cases that an inferior alternative is selected, it is likely to still be selected over the standard by the test. For $\psi_1 \geq 0$, this is generally a desirable outcome, as any inferior treatment within the range $(\mu_0 + \psi_1, \mu_0 + \psi_2)$ is in fact superior to the standard. Conservative experimenters can then select ψ_1 as a rough "minimally acceptable/good enough" performance level, and ψ_2 as the substantial practical significance level for which the test is powered. In this way, if an inferior treatment falls in $(\mu_0 + \psi_1, \mu_0 + \psi_2)$, the experimenter will tend to select the best and occasionally select a "good enough" treatment at the rate of correct selection implicit in Phase I.

D.2 Experiment Example

Here, we give a walkthrough of an experiment to help with intuition on the algorithm and how one might provide the prerequisite design parameters.

Suppose a website design manager seeks to compare two proposed website designs against the currently running design. The manager wants to implement the website design that results in larger customer purchases. Since the current website design has been implemented for some time, and many customers have made purchases on the website and even more have visited, we assume that the average amount each visitor spends when visiting under the current website design is well-known to the manager. The manager could set μ_0 to this value, but instead, without any loss of generality, the manager sets $\mu_0 = 0$, and the outcomes of the alternative treatments are recorded as their difference from the mean of the standard website outcomes. Outcomes are then defined and measured as the "lift" of each alternative relative to the standard website outcome. The data for this example was generated using normal distributions.

First, the manager sets error rates at $\alpha = 0.05$ and $\beta = 0.2$, reasoning that the risk of mistakenly replacing the business-as-usual treatment is of greater concern than missing out on a marginally superior alternative. The website design manager then determines that, given the large volume of visitors to the website and the relatively low cost of implementing an alternative website set-up that has already been designed, it will be worthwhile to change the website design to an alternative if the alternative generates an average \$0.50 lift in sales per visitor. In addition to considering switching costs, the decision incorporates the risk of a false positive. The manager sets $\psi_2 = 0.50$.

The manager then sets $\psi_1 = 0.25$, reasoning that if one of the alternatives generates a lift of more than \$0.50, it is worth distinguishing alternatives which are inferior by \$0.25 at a rate of 0.8.

Lastly, since the manager is testing two alternatives against the baseline, $K = 2$. With this information, the system of equations in Theorem 1 can be solved, and then the sequential procedure can be executed.

D.2.1 Beginning the Experiment

Before beginning the sequential algorithm, each alternative treatment must be sampled n_0 times, so the website design manager redirects 30 site visitors each to Treatment A and Treatment B and stores each visitor's final purchase amount. Since this experiment concerns lift, the samples are stored relative to μ_0. In R, the samples are stored in a dataframe with a column specifying the arm name ("A" and "B") and a column for the outcomes.

```
# Initial sample outcomes
initial_data <- data.frame(
  arm = c(rep("A", 30), rep("B", 30)),
  outcome = c(1.51, -3.36, -0.49, 2.77, ..., 0.20,
              3.85, 3.44, -0.13, 5.41, ..., 1.23))
```

This information can then be submitted along with the experiment parameters to initialize the experiment. This initialization step automatically solves the system of equations to provide the h and d parameters, which are stored in the `state` object along with the data and the user-provided experiment parameters.

```
state <- initialize_experiment(
  initial_data = initial_data, alpha = 0.05, beta = 0.2, psi1 = 0.25, psi2 = 0.5, mu0 = 0
)
```

The h and d parameters can be accessed from `state`. In this example, they are shown below.

h_1	h_2	d_1	d_2
0.679814	1.21582	3.937859	1.882612

D.2.2 Phase I

Next, the `update_and_check()` function can be used to determine whether the stopping rules have been met. As shown in Section 3.4, each sampled treatment has a separate stopping rule in Phase I, so that the sampling for each treatment can be stopped independently when sufficient information on that treatment is gathered.

The function `update_and_check()` takes a dataframe containing the next batch of samples and the up-to-date `state` object to check whether the stopping rule is met and advise the user on which arms to continue sampling, if any. The `state` object is then updated by the `update_and_check()` output.

```
out <- update_and_check(state = state, new_batch = new_batch)
# Update "state" object:
state <- out$state
```

The updated `state` object stores the samples and current sufficient statistics for each arm under `state$arms`. The message from the output will specify which arms to continue sampling, for example: `"Continue phase 1: sample active arms {A, B}."`.

As neither stopping rule has been met, $n' = 1$ additional visitors are assigned to each treatment:

```
new_batch <- data.frame(arm = c("A", "B"), outcome = c(-1.97, 2.90))
out <- update_and_check(state = state, new_batch = new_batch)
state <- out$state
```

Note that the number of additional samples taken in each step (denoted n' in Phase I, m' in Phase II) can be any value chosen by the experimenter. Smaller values of n' and m' result in less over-sampling.

This loop of sampling and checking continues until both stopping rules are met. If they do not stop simultaneously, samples are added only to the treatment that has not yet met its stopping rule in the manner shown below. The samples are then passed again into the `update_and_check()` function.

```
new_batch <- data.frame(arm = c("A"), outcome = c(-1.21))
```

```
out <- update_and_check(state = state, new_batch = new_batch)
state <- out$state
```

Once the stopping rules are met, Phase I is complete. `update_and_check()` informs the user whether to continue to Phase II or to conclude. From this example experiment, the output from the end of Phase I shows that Treatment B is selected for continued sampling, as it has the larger of the two sample means and its mean is around 0.3387, which is greater than $\mu_0 + \psi 2/d_1 = 0 + 0.1270$.

```
"Phase I is complete. Treatment B is selected for sampling in Phase II. "
```

D.2.3 Phase II

Phase II is begun by assigning an additional $m_0 = 30$ website visitors to Treatment B, storing their data in a new double vector, and passing the data into the `update_and_check()` function, just as before.

```
new_batch <- data.frame(arm = c("B"), outcome = c(0.89))
out <- update_and_check(state = state, new_batch = new_batch)
state <- out$state


"Continue phase 1: sample active arms {B}."
```

This sampling is continued until the stopping rule is met. Once the stopping rule is met, the mean of the Phase I and Phase II sample means for Treatment B is compared against the critical value $\mu_0 + \psi_2/d_2 = 0.2658$.

The `state` object in our run example reports a final mean value of Treatment B's data as `0.38625`, which is greater than the Phase II critical value. Thus, we reject the null hypothesis, and Treatment B is deemed an acceptable, superior alternative to the current standard website condition.

In a real online digital experiment, a wrapper function connected to the website system should be used with this method, automatically directing the users to treatments until the stopping rules are met. The example above is presented in a by-hand format to illustrate the proposed method.

References

James Bradley. Robustness? *The British Journal of Mathematical and Statistical Psychology*, 31(2):144–152, 1978.

Edward J Dudewicz and Peter R Nelson. Heteroscedastic analysis of means (hanom). *American Journal of Mathematical and Management Sciences*, 23(1-2):143–181, 2003.

Eyal Even-Dar, Shie Mannor, and Yishay Mansour. Action elimination and stopping conditions for the multi-armed bandit and reinforcement learning problems. *Journal of Machine Learning Research*, 2006.

Zohar Karnin, Tomer Koren, and Oren Somekh. Almost optimal exploration in multi-armed bandits. In *Proceedings of the 30th International Conference on Machine Learning*, 2013.

Seong-Hee Kim. Comparison with a standard via fully sequential procedures. *ACM Transactions on Modeling and Computer Simulation*, 15(2), 2005.

Seong-Hee Kim and Barry L. Nelson. On the asymptotic validity of fully sequential selection procedures for steady-state simulation. *Operations Research*, 54(3), 2006.

Nitis Mukhopadhyay and Basil M De Silva. *Sequential Methods and Their Applications*. CRC Press, 2009.

Barry L. Nelson and David Goldsman. Comparisons with a standard in simulation experiments. *Management Science*, 2001.

Peter R Nelson and Edward J Dudewicz. Exact analysis of means with unequal variances. *Technometrics*, 44(2):152–160, 2002.

Xinkun Nie, Xiaoying Tian, Jonathan Taylor, and James Zou. Why adaptively collected data have negative bias and how to correct for it. In *Proceedings of the 21st International Conference on Artificial Intelligence and Statistics (AISTATS)*, 2018.

Linda Rollin and Pinyuen Chen. A two-stage design for choosing among experimental treatments in clinical trials. In *Advances in Ranking and Selection, Multiple Comparisons, and Reliability*, pages 385–409. Springer, 2005.