Asymptotically Valid Single-Stage Multiple-Comparison Procedures

Marvin K. Nakayama

Computer Science Department
New Jersey Institute of Technology
Newark, NJ 07102

July 25, 2008

Abstract

We establish general conditions for the asymptotic validity of single-stage multiple-comparison procedures (MCPs) under the following general framework. There is a finite number of independent alternatives to compare, where each alternative can represent, e.g., a population, treatment, system or stochastic process. Associated with each alternative is an unknown parameter to be estimated, and the goal is to compare the alternatives in terms of the parameters. We establish the MCPs' asymptotic validity, which occurs as the sample size of each alternative grows large, under two assumptions. First, for each alternative, the estimator of its parameter satisfies a central limit theorem (CLT). Second, we have a consistent estimator of the variance parameter appearing in the CLT. Our framework encompasses comparing means (or other moments) of independent (not necessarily normal) populations, functions of means, quantiles, steady-state means of stochastic processes, and optimal solutions of stochastic approximation by the Kiefer-Wolfowitz algorithm. The MCPs we consider are multiple comparisons with the best, all pairwise comparisons, all contrasts, and all linear combinations, and they allow for unknown and unequal variance parameters and unequal sample sizes across alternatives.

Keywords: Central limit theorem, consistent estimation, simulation, quantile, pairwise comparison, contrasts, multiple comparisons with the best.

1 Introduction

Suppose that there are k alternatives to compare, where the ith alternative has an unknown parameter θ_i that needs to be estimated. Each alternative might represent, e.g., a population, treatment, system or stochastic process. The alternatives are to be compared relative to the θ_i, $i = 1, \ldots, k$. For example, we may be faced with 5 possible designs for a fault-tolerant computer system, and we want to compare the alternatives in terms of the 0.9-quantile of the time to failure.

To do this, we consider single-stage multiple-comparison procedures (MCPs). In this paper we focus on multiple comparisons with the best (MCB; Hsu (1984)), all pairwise comparisons, all contrasts, and all linear combinations of $\theta_1, \ldots, \theta_k$, but the same ideas also apply to other multiple-comparison procedures (Hochberg and Tamhane, 1987). MCB produces simultaneous confidence intervals for $\theta_i - \max_{j \neq i} \theta_j$, $i = 1, \ldots, k$, which provide useful information when one is interested in identifying the best alternative and larger parameter values are better.

We establish the asymptotic validity of our single-stage MCPs under the following two assumptions. First, each parameter θ_i has an estimator $\widehat{\theta}_i(n)$ based on a sample size of n from alternative i, where $\widehat{\theta}_1, \ldots, \widehat{\theta}_k$ are independent, and $\widehat{\theta}_i(n)$ satisfies a central limit theorem (CLT); i.e., there exists a constant η such that $n^\eta[\widehat{\theta}_i(n) - \theta_i] \Rightarrow N(0, \sigma_i^2)$ for each i as $n \to \infty$, where $0 < \sigma_i < \infty$, $\Rightarrow$ denotes weak convergence (Billingsley, 1999) and $N(a, b)$ denotes a normal distribution with mean a and variance b. In many applications, $\eta = 1/2$, but we also allow other values. Second, we assume there is a consistent estimator for the variance parameter σ_i^2 appearing in the CLT. Our framework encompasses comparing means (or other moments) of independent populations or treatments, functions of means, quantiles, steady-state means of stochastic processes, and optimal solutions to stochastic approximation via the Kiefer-Wolfowitz (1952) algorithm. We allow for unequal sample sizes and unequal variance parameters across alternatives.

While much of the previous work on MCPs focuses on finite-sample methods for comparing means of normally distributed populations by collecting independent and identically distributed (i.i.d.) samples within each population (e.g., see Hochberg and Tamhane (1987) and Hsu (1996)), there are also some large-sample procedures. Hjort (1988) considers the setting where the estimators $\widehat{\theta}_1, \ldots, \widehat{\theta}_k$ may be dependent (the present paper assumes independence), and they satisfy a joint CLT with $\eta = 1/2$ (we allow $\eta \neq 1/2$), and there is a consistent estimator of the covariance matrix in the CLT. He constructs large-sample Scheffé-type (1953) simultaneous confidence intervals for $f(\theta_1, \ldots, \theta_k)$ for all suitably smooth functions f. He provides analysis demonstrating that the resulting simultaneous confidence intervals are usually worse than those based on Bonferroni's inequality; in contrast, most of our proofs apply sharper inequalities than Bonferroni's, leading to intervals shorter than those using Bonferroni's inequality. Hochberg and Tamhane (1987, Section 10.1.1) and Piegorsch (1991) present large-sample MCPs for comparing treatments with Bernoulli responses. Nakayama (1997) establishes the asymptotic validity of single-stage MCPs based on standardized time series methods (Schruben, 1983) for steady-state simulations. Standardized time series methods yield variance estimators that are not consistent, so the results in the present paper do not cover those in Nakayama (1997).

The rest of the paper has the following organization. Section 2 describes the mathematical framework we adopt, and we present the single-stage MCPs in Sections 3. Section 4 shows that the assumptions (CLT and consistent estimator for the variance parameter) hold in a wide variety of contexts. We describe in Section 4.1 some settings where our assumptions do not hold. Section 5 contains numerical results from a Monte Carlo experiment. All proofs are collected in Section 6. Nakayama (2007) presents without proof the MCB procedure we develop.

2 Mathematical Framework

Suppose there are $k < \infty$ alternatives, where each alternative i has an unknown parameter θ_i to be estimated. We have an *estimation process* $\widehat{\theta}_i = [\widehat{\theta}_i(n) : n > 0]$, where $\widehat{\theta}_i(n)$ is the estimator of θ_i based on a sample size of n from alternative i. In the case of comparing populations, treatments, or stochastic systems relative to a terminating performance measure, n denotes the number of i.i.d. samples taken from an alternative. When comparing stochastic processes using simulations, n represents the run length of the simulation of a stochastic process. For example, when θ_i is the steady-state mean of a continuous-time stochastic process $X_i = [X_i(t) : t \geq 0]$, we can define an estimator $\widehat{\theta}_i(n) = (1/n) \int_0^n X_i(t)\, dt$. We assume that the estimation processes satisfy the following CLT:

Assumption 1 *The estimation processes $\widehat{\theta}_1, \ldots, \widehat{\theta}_k$ are independent, and there exists a finite positive constant η such that for each i,*

$$n^{\eta} \left[\widehat{\theta}_i(n) - \theta_i \right] \Rightarrow N(0, \sigma_i^2) \tag{1}$$

as $n \to \infty$, where $\sigma_i > 0$ is a finite constant.

In most applications, η has the canonical value of 1/2, but we do not require this. In our previous example when θ_i represents a steady-state mean of a stochastic process X_i, Assumption 1 holds with $\eta = 1/2$ in a wide variety of contexts; e.g., see Asmussen (2003) for many examples.

We call σ_i^2 the *variance parameter* of alternative i. We assume that there is a variance-estimation process $V_i = [V_i(n) : n > 0]$, where $V_i(n) \geq 0$ is the estimator of σ_i^2 based on a sample size of n from alternative i. We assume V_i is consistent:

Assumption 2 *For each alternative i, $V_i(n) \Rightarrow \sigma_i^2$ as $n \to \infty$.*

For our previous example when θ_i represents a steady-state mean, if we further assume that the process X_i is regenerative (e.g., see Glynn and Iglehart (1993)), then we can use the regenerative method to construct an estimator $V_i(n)$ satisfying Assumption 2. In Section 4 we provide further details on this and other examples of settings satisfying our two assumptions.

3 Multiple-Comparison Procedures

We now present the MCPs for comparing the parameters θ_i, $i = 1, 2, \ldots, k$, of independent alternatives. For each one, we let $1 - \alpha$, $0 < \alpha < 1$, denote the desired joint confidence level of the simultaneous confidence intervals constructed. Also, let $\boldsymbol{n} = (n_1, n_2, \ldots, n_k)$, where n_i is the sample size taken from alternative i. Thus, for each alternative i, we have the estimators $\widehat{\theta}_i(n_i)$ and $V_i(n_i)$ of θ_i and σ_i^2, respectively. Also, define

$$W_{i,j}(\boldsymbol{n}) = \sqrt{\frac{V_i(n_i)}{n_i^{2\eta}} + \frac{V_j(n_j)}{n_j^{2\eta}}}$$

for each pair $i, j = 1, 2, \ldots, k$, where η is defined in Assumption 1.

3.1 MCB Procedure

To construct the joint MCB intervals, define the constant γ to satisfy $\Phi(\gamma) = (1 - \alpha)^{1/(k-1)}$, where Φ is the distribution function of a standard (mean 0 and variance 1) normal distribution. For each i, define

$$\begin{aligned} D_i^+(\boldsymbol{n}) &= \left(\min_{j \neq i} \left\{ \widehat{\theta}_i(n_i) - \widehat{\theta}_j(n_j) + \gamma W_{i,j}(\boldsymbol{n}) \right\} \right)^+, \\ A(\boldsymbol{n}) &= \{ i : D_i^+(\boldsymbol{n}) > 0 \}, \\ D_i^-(\boldsymbol{n}) &= \begin{cases} 0 & \text{if } A(\boldsymbol{n}) = \{i\}, \\ \min_{j \in A(\boldsymbol{n}), j \neq i} \{ \widehat{\theta}_i(n_i) - \widehat{\theta}_j(n_j) - \gamma W_{i,j}(\boldsymbol{n}) \} & \text{otherwise,} \end{cases} \end{aligned}$$

and define the MCB intervals

$$I_i(\boldsymbol{n}) = [\, D_i^-(\boldsymbol{n}),\ D_i^+(\boldsymbol{n}) \,] \tag{2}$$

for $\theta_i - \max_{j \neq i} \theta_j$, $i = 1, \ldots, k$. The following result, whose proof is in Section 6.1, establishes the asymptotic validity of the MCB intervals in (2).

Theorem 1 *Suppose $n_i = \zeta_i n$ for $i = 1, 2, \ldots, k$, where $\zeta_i > 0$ is any constant. If Assumptions 1 and 2 hold, then*

$$\lim_{n \to \infty} P\left\{ \theta_i - \max_{j \neq i} \theta_j \in I_i(\boldsymbol{n}),\ i = 1, 2, \ldots, k \right\} \geq 1 - \alpha.$$

The proof of Theorem 1 relies on Slepian's (1962) inequality, which is sharper than Bonferroni's inequality but nevertheless still results in conservative intervals. We can reduce the conservatism by avoiding both inequalities and instead applying an asymptotic version of a method developed by Dunnett (1955) for multiple comparisons with a control. Piegorsch (1991) develops this idea to yield an asymptotic version of an MCB method of Hsu (1996, Theorem 4.1.2) to compare treatments with Bernoulli responses; i.e., when we sample i.i.d. Bernoulli random variables $X_{i,1}, X_{i,2}, \ldots$ from each alternative i, and we compare $\theta_i = E[X_{i,1}]$, $i = 1, \ldots, k$, using estimators $\widehat{\theta}_i(n) = (1/n) \sum_{j=1}^{n} X_{i,j}$ and $V_i(n) = \widehat{\theta}_i(n)(1 - \widehat{\theta}_i(n))$. This approach requires replacing the constant γ with other quantities $\widetilde{\gamma}_i$, $i = 1, \ldots, k$, that depend on all the $V_i(n_i)$, $i = 1, \ldots, k$, so $\widetilde{\gamma}_i$, $i = 1, \ldots, k$, must be numerically evaluated after all samples are collected. However, we do not develop this method for our general setting because of its added computational burden.

3.2 All Pairwise Comparisons

We now consider making all pairwise comparisons. To do this, first choose the constant $\gamma' = q_k/\sqrt{2}$, where q_k is the upper α point of the range of k i.i.d. standard normals; i.e., $P\{\max_{1 \leq i < j \leq k} |Z_i - Z_j| \leq q_k\} = 1 - \alpha$, where $Z_1, \ldots, Z_k$ are i.i.d. $N(0, 1)$ random variables. To obtain appropriate values for q_k, we can consult Table 8 of Hochberg and Tamhane (1987) or use a statistical package such as R (R Development Core Team, 2008). Then we construct the (two-sided) simultaneous confidence intervals

$$I'_{i,j}(\boldsymbol{n}) = \left[\widehat{\theta}_i(n_i) - \widehat{\theta}_j(n_j) \ \pm \ \gamma' W_{i,j}(\boldsymbol{n}) \right] \tag{3}$$

for $\theta_i - \theta_j$, $1 \le i < j \le k$. The following result, whose proof is given in Section 6.2, is an asymptotic version of the Tukey-Kramer method (Tukey, 1953; Kramer, 1956) generalized to apply to our framework.

Theorem 2 *Suppose $n_i = \zeta_i n$ for $i = 1, 2, \ldots, k$, where $\zeta_i > 0$ is any constant. If Assumptions 1 and 2 hold, then*

$$\lim_{n \to \infty} P \left\{ \theta_i - \theta_j \in I'_{i,j}(\boldsymbol{n}), \ \forall \ i < j \right\} \ge 1 - \alpha.$$

Hochberg and Tamhane (1987, Section 10.1.1) present the special case of (3) for Bernoulli response. Also, Piegorsch (1991) empirically studies the small-sample performance of (3) in the case of Bernoulli response.

The proof of Theorem 2 makes use of the inequality of Hayter (1984), which is sharper than the Bonferroni inequality. Thus, Theorem 2 yields shorter confidence intervals than those based on the Bonferroni inequality.

Theorem 2 shows that the joint intervals in (3) are conservative. One could obtain exact intervals when the ratios of the variances are known by developing an asymptotic version of a procedure by Spurrier and Isham (1985) for comparing $k = 3$ normal means (also see pp. 86–88 of Hochberg and Tamhane (1987) for an extension by Hayter (1985)). However, we do not develop this idea further.

3.3 All Contrasts

Now we examine all contrasts $\sum_{i=1}^{k} c_i \theta_i$ with $\boldsymbol{c} = (c_1, c_2, \ldots, c_k) \in C$, where $C = \{\boldsymbol{c} \in \Re^k : \sum_{i=1}^{k} c_i = 0\}$ is the k-dimensional contrast space. This can be used to analyze a weighted average of the parameters, such as $\theta_1 - (\theta_2 + \theta_3)/2$. For each $\boldsymbol{c} = (c_1, c_2, \ldots, c_k) \in C$, we now define the confidence interval

$$I_{\boldsymbol{c}}(\boldsymbol{n}) \;=\; \left[\sum_{i=1}^{k} c_i \widehat{\theta}_i(n_i) \;\pm\; \frac{2}{\sum_{l=1}^{k} |c_l|} \sum_{i=1}^{k} \sum_{j=1}^{k} c_i^{+} c_j^{-} \gamma' W_{i,j}(\boldsymbol{n}) \right]$$

for $\sum_{i=1}^{k} c_i\theta_i$, where γ' is defined as in Section 3.2.

Theorem 3 *Suppose $n_i = \zeta_i n$ for $i = 1, 2, \ldots, k$, where $\zeta_i > 0$ is any constant. If Assumptions 1 and 2 hold, then*

$$\lim_{n\to\infty} P\left\{ \sum_{i=1}^{k} c_i\theta_i \in I_{\boldsymbol{c}}(\boldsymbol{n}),\ \forall\, \boldsymbol{c} = (c_1, c_2, \ldots, c_k) \in C \right\} \geq 1 - \alpha.$$

Theorem 3 is intended for constructing only those confidence intervals with $\boldsymbol{c} \in C$ that are of interest.

3.4 All Linear Combinations

Now we consider all linear combinations $\sum_{i=1}^{k} a_i\theta_i$ with $\boldsymbol{a} = (a_1, a_2, \ldots, a_k) \in \Re^k$. This can be used to analyze simultaneously parameters and a weighted average of the parameters, such as θ_1, θ_2, θ_3 and $\theta_1 - (\theta_2 + \theta_3)/2$. For each $\boldsymbol{a} = (a_1, a_2, \ldots, a_k) \in \Re^k$ and $p \geq 1$, we define the confidence interval

$$\bar{I}_{\boldsymbol{a},p}(\boldsymbol{n}) \;=\; \left[\sum_{i=1}^{k} a_i \widehat{\theta}_i(n_i) \;\pm\; \bar{\gamma}_p \left\{ \sum_{i=1}^{k} \left(|a_i| \sqrt{\frac{V_i(n_i)}{n_i^{2\eta}}} \right)^{p} \right\}^{1/p} \right]$$

for $\sum_{i=1}^{k} a_i\theta_i$, where $\bar{\gamma}_p$ satisfies $P\{(\sum_{i=1}^{k} Z_i^q)^{1/q} \leq \bar{\gamma}_p\} = 1-\alpha$, with $Z_1, \ldots, Z_k$ i.i.d. $N(0,1)$ and $q \geq 1$ satisfying $1/p + 1/q = 1$. In the special case when $p = 1$ and $q = \infty$, we have $\bar{\gamma}_1$ satisfies $\Phi(\bar{\gamma}_1) = (1-\alpha)^{1/k}$. When $p = q = 2$, then $\bar{\gamma}_2$ satisfies $P\left\{\sqrt{\chi_k^2} \leq \bar{\gamma}_2\right\} = 1-\alpha$, where χ_k^2 is a chi-squared random variable with k degrees of freedom. Then we have the following, which is an asymptotic generalization of a method of Dalal (1978), originally developed for comparing means of normal populations, to handle our general framework.

Theorem 4 *Suppose $n_i = \zeta_i n$ for $i = 1, 2, \ldots, k$, where $\zeta_i > 0$ is any constant. If Assumptions 1 and 2 hold, then for any $p \geq 1$,*

$$\lim_{n\to\infty} P\left\{ \sum_{i=1}^{k} a_i\theta_i \in \bar{I}_{\boldsymbol{a},p}(\boldsymbol{n}),\ \forall\, \boldsymbol{a} = (a_1, a_2, \ldots, a_k) \in \Re^k \right\} = 1 - \alpha.$$

Hjort (1988) establishes the same result for the case that $p = q = 2$ and $\eta = 1/2$ but allows $\widehat{\theta}_1, \ldots, \widehat{\theta}_k$ to be dependent.

4 Examples

We now discuss examples of estimation settings satisfying Assumptions 1 and 2. Most of the examples below are taken from Glynn and Whitt (1992a; 1992b). As we shall see, Assumptions 1 and 2 hold in a wide spectrum of contexts arising in practice.

Example 1 (Means) Consider k independent populations, and let $X_{i,1}, X_{i,2}, \ldots$ be i.i.d. samples from the ith population. We want to compare the populations in terms of their means $\theta_i = E[X_{i,1}]$, $i = 1, \ldots, k$. If we define an estimator of θ_i to be $\widehat{\theta}_i(n) = (1/n)\sum_{j=1}^{n} X_{i,j}$, then Assumption 1 holds with $\eta = 1/2$ when $\sigma_i^2 \equiv E[(X_{i,1} - \theta_i)^2] < \infty$. A consistent estimator of σ_i^2 is $V_i(n) = (1/(n-1))\sum_{j=1}^{n}[X_{i,j} - \widehat{\theta}_i(n)]^2$, so Assumption 2 holds. Thus, the MCPs in Section 3 are asymptotically valid when comparing means of independent populations using these definitions of $\widehat{\theta}_i(n)$ and $V_i(n)$. Note that we did not assume normality, and this example covers the problem of comparing alternative stochastic systems relative to a terminating performance measure by simulating independent replications. ∎

Example 2 (Function of means) Consider k independent populations, and let $\boldsymbol{X}_{i,1}, \boldsymbol{X}_{i,2}, \ldots$ be i.i.d. samples from the ith population, where $\boldsymbol{X}_{i,j} = (X_{i,j}^{(1)}, X_{i,j}^{(2)}, \ldots, X_{i,j}^{(d)}) \in \Re^d$. Let $\boldsymbol{\mu}_i = E[\boldsymbol{X}_{i,1}]$, and let $g_i : \Re^d \to \Re$. We now want to compare the populations in terms of

$\theta_i = g_i(\boldsymbol{\mu}_i)$, $i = 1, \ldots, k$. An example is comparing the coefficients of variation of k populations. In this case, we take $d = 2$ and $g_i(x_1, x_2) = \sqrt{x_2 - x_1^2}/x_1$, and let $\boldsymbol{X}_{i,j} = (Y_{i,j}, Y_{i,j}^2)$, where $Y_{i,j}$ is the jth i.i.d. sample from population i. Also, this framework encompasses comparing the rth moments, $r \geq 1$, of independent populations by taking $d = 1$ and $g_i(x) = x^r$.

Assume that g_i is continuously differentiable in a neighborhood of $\boldsymbol{\mu}_i$. Let ∇g_i denote the gradient of g_i, and assume that $\nabla g_i(\boldsymbol{\mu}_i) \neq 0$. We define an estimator of θ_i to be $\widehat{\theta}_i(n) = g_i(\bar{\boldsymbol{X}}_i(n))$, where $\bar{\boldsymbol{X}}_i(n) = (1/n) \sum_{j=1}^n \boldsymbol{X}_{i,j}$. Let $\| \cdot \|$ denote the Euclidean norm on $\Re^d$, and assume $E[\|\boldsymbol{X}_{i,1}\|^2] < \infty$ for each i. Let $\boldsymbol{\Sigma}_i = E[(\boldsymbol{X}_{i,1} - \boldsymbol{\mu}_i)(\boldsymbol{X}_{i,1} - \boldsymbol{\mu}_i)^\top]$ denote the covariance matrix of $\boldsymbol{X}_{i,1}$, where superscript $\top$ denotes transpose, and assume each $\boldsymbol{\Sigma}_i$ is positive definite. Then Assumption 1 holds with $\eta = 1/2$ and $\sigma_i^2 = \nabla g_i(\boldsymbol{\mu}_i)^\top \boldsymbol{\Sigma}_i \nabla g_i(\boldsymbol{\mu}_i)$; e.g., see p. 124 of Serfling (1980). We can define a consistent estimator of σ_i^2 as follows. Let $\widehat{\boldsymbol{\Sigma}}_i(n) = (1/(n-1)) \sum_{j=1}^n (\boldsymbol{X}_{i,j} - \bar{\boldsymbol{X}}_i(n))(\boldsymbol{X}_{i,j} - \bar{\boldsymbol{X}}_i(n))^\top$, and then $V_i(n) = \nabla g_i(\bar{\boldsymbol{X}}_i(n))^\top \widehat{\boldsymbol{\Sigma}}_i(n) \nabla g_i(\bar{\boldsymbol{X}}_i(n))$ satisfies Assumption 2. Thus, the MCPs in Section 3 are asymptotically valid when comparing functions of means using these definitions of $\widehat{\theta}_i(n)$ and $V_i(n)$. ■

Example 3 (Quantiles) For a distribution function G on $\Re$ and $0 < y < 1$, let $G^{-1}(y) = \inf\{x : G(x) \geq y\}$ be the yth quantile of G. Let F_i be the distribution function of the ith population, and for a fixed $0 < p < 1$, let $\theta_i = F_i^{-1}(p)$, so we are comparing the k independent populations in terms of their pth quantiles. Let $X_{i,1}, X_{i,2}, \ldots, X_{i,n}$ be i.i.d. samples from the ith population, where $X_{i,j} \in \Re$. Define the empirical distribution function $F_{i,n}(x) = (1/n) \sum_{j=1}^n 1\{X_{i,j} \leq x\}$, where $1\{A\}$ denotes the indicator function of the event $\{A\}$. An estimator of θ_i is then

$$\widehat{\theta}_i(n) = F_{i,n}^{-1}(p). \tag{4}$$

If each F_i has a density function f_i (with respect to Lebesgue measure) such that $f_i(\theta_i) > 0$, then Assumption 1 holds with $\eta = 1/2$ and

$$\sigma_i^2 = \frac{p(1-p)}{f_i^2(\theta_i)}; \tag{5}$$

e.g., see p. 77 of Serfling (1980). We can construct a consistent estimator of σ_i^2, which is *not* the variance of F_i, as follows. For each $n > 0$, define a constant $q_{i,n}$ such that

$$q_{i,n} = p + \sqrt{\frac{p(1-p)}{n}} + o\left(\frac{1}{n^{1/2}}\right)$$

as $n \to \infty$, where a function $g(n) = o(h(n))$ means $g(n)/h(n) \to 0$ as $n \to \infty$. Then the estimator

$$V_i(n) = n\left[F_{i,n}^{-1}\left(q_{i,n}\right) - F_{i,n}^{-1}(p)\right]^2 \tag{6}$$

satisfies Assumption 2; e.g., see p. 94 of Serfling (1980). Thus, the MCPs in Section 3 are asymptotically valid when comparing quantiles using these definitions of $\widehat{\theta}_i(n)$ and $V_i(n)$. ■

Example 4 (Steady-state mean rewards of stochastic processes) Let $X_1, \ldots, X_k$ be k independent stochastic processes, where $X_i = [X_i(t) : t \geq 0]$ lives on a state space S_i. Let f_i be a real-valued "reward" function on S_i. Assume that each X_i has a steady-state in the sense that there exists a finite constant θ_i such that $(1/t)\int_0^t f_i(X_i(s))\,ds \Rightarrow \theta_i$ as $t \to \infty$. Under great generality, $\widehat{\theta}_i(t) = (1/t)\int_0^t f_i(X_i(s))\,ds$ satisfies Assumption 1 with $\eta = 1/2$ and some σ_i^2; e.g., see Asmussen (2003) for examples of processes satisfying such a CLT. Autocorrelations typically present in X_i make constructing a consistent estimator of σ_i^2 a delicate undertaking. Examples of such estimators for which Assumption 2 holds under various conditions include spectral estimators (Damerdji, 1991), regenerative estimators (Glynn and Iglehart, 1993), autoregressive estimators (Fishman, 1978, p. 252), and batch means and batched area estimators in which the number of batches $m \to \infty$ at an appropriate rate as the run length n increases (Damerdji, 1994).

We now provide details on the regenerative estimator of σ_i^2. Assume that there exist nonnegative times $A_{i,0} < A_{i,1} < A_{i,2} < \cdots$ such that X_i is regenerative with respect to the sequence $(A_{i,j} : j \geq 0)$. For $j \geq 1$, define $\tau_{i,j} = A_{i,j} - A_{i,j-1}$ and $Y_{i,j} = \int_{A_{i,j-1}}^{A_{i,j}} f_i(X_i(s))\, ds$, and let $N_i(t) = \sup\{A_{i,j} : A_{i,j} \leq t\}$. If $E[\tau_{i,1}] < \infty$ and $0 < E[(Y_{i,1} - \theta_i \tau_{i,1})^2] < \infty$, then $\theta_i = E[Y_{i,1}]/E[\tau_{i,1}]$ and $\sigma_i^2 = E[(Y_{i,1} - \theta_i \tau_{i,1})^2]/E[\tau_{i,1}]$. We then define $V_i(t) = (1/t) \sum_{j=1}^{N_i(t)} [Y_{i,j} - \theta_i(t)\tau_{i,j}]^2$, which satisfies Assumption 2; see Glynn and Iglehart (1993). ∎

Example 5 (Kiefer-Wolfowitz stochastic approximation) Suppose there are k systems, and for each system i, let $Z_i(\rho_i)$ be a random variable denoting the (random) performance of system i under parameter value $\rho_i \in \Re$. Let $\beta_i(\rho_i) = E[Z_i(\rho_i)]$, and assume that β_i is three-times differentiable on $\Re$. Let the minimizer of β_i be $\theta_i = \rho_i^*$, and assume that ρ_i^* is the unique solution to $\beta_i'(\rho_i) = 0$, where prime denotes derivative. The goal is to compare the k systems in terms of $\theta_1, \ldots, \theta_k$. To estimate each θ_i, we apply the Kiefer-Wolfowitz (1952) stochastic approximation algorithm, which generates a sequence $\rho_{i,1}, \rho_{i,2}, \ldots$ that converges a.s. to ρ_i^*. Specifically, for each system i, let $(c_{i,n} : n \geq 0)$ and $(h_{i,n} : n \geq 0)$ be deterministic sequences of nonnegative constants, and the algorithm produces successive estimates of θ_i as $\rho_{i,n+1} = \rho_{i,n} - c_{i,n} X_{i,n+1}$, where $X_{i,n+1}$ satisfies for each (measurable) set A of real numbers,

$$P\{X_{i,n+1} \in A \mid (\rho_{i,j}, X_{i,j}), j \leq n\} = P\left\{ \frac{Z_i(\rho_{i,n} + h_{i,n+1}) - Z_i(\rho_{i,n} - h_{i,n+1})}{2h_{i,n+1}} \in A \right\},$$

with $Z_i(\rho_{i,n} + h_{i,n+1})$ and $Z_i(\rho_{i,n} - h_{i,n+1})$ generated independently. Then define the estimator $\widehat{\theta}_i(n) = \rho_{i,n}$ for $n \geq 0$. For given positive constants c_i and h_i, suppose we set $c_{i,n} = c_i n^{-1}$ and $h_{i,n} = h_i n^{-1/3}$. Assume that $c_i \beta_i''(\rho_i^*) > 1/3$. Then under suitable regularity conditions, it follows from Ruppert (1982) that as $n \to \infty$,

$$n^{1/3}(\widehat{\theta}_i(n) - \theta_i) \Rightarrow N(0, \sigma_i^2),$$

with $\sigma_i^2 = c_i^2 \mathrm{Var}[Z_i(\rho_i^*)]/((c_i \beta_i''(\rho_i^*) - 1/3)(4h_i^2))$. Thus, Assumption 1 holds with noncanonical $\eta = 1/3$. For directions on constructing a variance estimator $V_i(n)$ satisfying Assumption 2, see Ventner (1967), p. 189. ∎

Glynn and Whitt (1992a; 1992b) provide additional examples, including others with noncanonical η, satisfying our assumptions. Actually, they show that stronger results than Assumptions 1 and 2 hold, namely, that $\widehat{\theta}_i(n)$ satisfies a functional central limit theorem (Billingsley, 1999) and $V_i(n)$ is strongly consistent.

4.1 When Assumptions 1 and 2 Are Violated

While Assumptions 1 and 2 hold for many situations arising in practice, we now discuss cases where they are not satisfied. Assumption 1 requires that the k parameters $\theta_1, \ldots, \theta_k$ are estimated independently, but there are practical settings when the estimation processes $\widehat{\theta}_1, \ldots, \widehat{\theta}_k$ may be dependent. For example, suppose there are d populations, and we perform i.i.d. sampling within each population. Suppose we want to simultaneously compare the means and compare the variances (and possibly also compare functions of these parameters, e.g., coefficients of variation) of the populations, using the same samples within a population to estimate all the parameters. When the populations are not normally distributed, the sample mean and sample variance are dependent (Lukacs, 1956), so our MCPs do not apply in this case. Instead, one could use the asymptotic procedure of Hjort (1988), which allows the estimators $\widehat{\theta}_1, \ldots, \widehat{\theta}_k$ to be dependent (and constructs simultaneous intervals for all suitably smooth functions of the estimators), but at the cost of wider intervals.

Also, Assumption 1 requires a CLT with scaling n^η and normal limit. In the setting of i.i.d. sampling to compare means of independent (not necessarily normal) populations, as in Example 1, this corresponds to the domain of normal attraction of the normal distribution, for which a necessary and sufficient condition is that the variance is finite; see p. 181 of Gnedenko and Kolmogorov (1968). Thus, Assumption 1 is violated for populations having infinite variance.

Assumption 2 requires consistent estimation of the parameter σ_i^2. If we are comparing steady-state mean rewards of stochastic processes (Example 4), one approach to estimate

σ_i^2 is to use a standardized time series (STS) method (Schruben, 1983). However, STS estimators are not consistent (Glynn and Iglehart (1990), Proposition 4.26), so Assumption 2 does not hold and our MCPs do not apply. As an alternative, we could employ the MCPs of Nakayama (1997), which are designed to use STS estimators.

5 Numerical Results

We now present results from a Monte Carlo experiment comparing independent normally distributed populations in terms of their pth quantiles for various values of $0 < p < 1$. There are $k = 4$ independent populations, where the ith population has a $N(\mu_i, 1)$ distribution. In all our experiments, we fixed $\mu_1 = \mu_2 = \mu_3 = 1$ and $\mu_4 = 1.1$. Let θ_i denote the pth quantile of the ith population, and we estimate each θ_i using (4). Our goal is to construct MCB intervals for $\theta_1, \ldots, \theta_4$. For any fixed p, population 4 is the best since θ_4 is the largest among $\theta_1, \ldots, \theta_4$, which we assume is unknown. We used the same sample size $n_i = n$ for each population with the values $n = 20$, 80 and 320, and we varied the quantile level p between 0.1 and 0.9. We assume that σ_i^2 in (5) is unknown, and we estimate σ_i^2 using $V_i(n)$ in (6) with $q_{i,n} = p + \sqrt{p(1-p)/n}$. We ran 10^4 independent replications for each set of parameters, where we constructed MCB intervals having nominal joint confidence level $1 - \alpha = 0.9$ in each replication.

Table 1 gives the coverage results from our simulations. For a fixed quantile level p, coverage increases as the sample size n for each population increases, with the coverage levels for the largest sample size $n = 320$ reasonably close to the nominal level of 90%. This agrees with our asymptotic theory in Theorem 1. Also, comparing the coverages for the different quantile levels p for the same n, we see that the coverages are typically lower for $p = 0.1$ and 0.9 than for $p = 0.3$, 0.5 or 0.7 when n is either 20 or 80, which is when coverages are far below the nominal level of 0.9. This may indicate that extreme quantiles are harder to estimate than those closer to the median.

Table 1: Coverage results, in percents, for MCB intervals for the pth quantiles, with sample size n.

p	n		
	20	80	320
0.1	71.8	74.0	84.9
0.3	83.6	85.4	88.2
0.5	82.4	84.2	85.7
0.7	85.6	87.9	89.3
0.9	74.3	77.3	87.9

For comparison, Table 2 contains results from experiments constructing MCB intervals for the means, i.e., $\theta_i = \mu_i$. In this case the coverage levels are all greater than the nominal level 0.9.

Nakayama (2007) presents results from constructing MCB intervals for quantiles and means of exponentially distributed populations, and the results there are similar. Thus, it appears that MCB intervals for quantiles require larger sample sizes than for means for the asymptotics to take effect.

Table 2: Coverage results, in percents, for MCB intervals for the means, with sample size n.

n		
20	80	320
91.3	91.2	91.5

6 Proofs

6.1 Proof of Theorem 1

Recall that each $n_i = \zeta_i n$, where $\zeta_i > 0$ is any constant, so $\boldsymbol{n} = (\zeta_1 n, \ldots, \zeta_k n)$. Define $(1), (2), \ldots, (k)$ such that $\theta_{(1)} \leq \theta_{(2)} \leq \cdots \leq \theta_{(k)}$; i.e., alternative (j) has the jth smallest parameter, where ties can be broken arbitrarily. Then, define the events

$$\begin{aligned} G(n) &= \left\{\theta_i - \theta_{(k)} \geq \widehat{\theta}_i(n_i) - \widehat{\theta}_{(k)}(n_{(k)}) - \gamma W_{i,(k)}(\boldsymbol{n}),\ \forall\ i \neq (k)\right\}, \\ G_1(n) &= \left\{\theta_i - \max_{j\neq i}\theta_j \leq \left[\min_{j\neq i}\left(\widehat{\theta}_i(n_i) - \widehat{\theta}_j(n_j) + \gamma W_{j,i}(\boldsymbol{n})\right)\right]^+,\ \forall\ i\right\}, \\ G_2(n) &= \left\{\theta_i - \max_{j\neq i}\theta_j \geq -\left[\min_{j\in A(\boldsymbol{n}),\ j\neq i}\left(\widehat{\theta}_i(n_i) - \widehat{\theta}_j(n_j) - \gamma W_{i,j}(\boldsymbol{n})\right)\right]^-,\ \forall\ i\right\}. \end{aligned}$$

Note that $\{\theta_i - \max_{j\neq i}\theta_j \in I_i(\boldsymbol{n}),\ i = 1, 2, \ldots, k\} = G_1(n) \cap G_2(n)$, and by slightly modifying the argument used by Hsu (1996, pp. 91–92), we can show that $G(n) \subseteq G_1(n) \bigcap G_2(n)$ for all n. We now show that

$$\lim_{n\to\infty} P(G(n)) \geq 1 - \alpha, \tag{7}$$

which will then establish Theorem 1.

Note that

$$W_{i,j}(\boldsymbol{n}) = \frac{1}{n^\eta}\sqrt{\frac{V_i(n_i)}{\zeta_i^{2\eta}} + \frac{V_j(n_j)}{\zeta_j^{2\eta}}},$$

so

$$P(G(n)) = P\left\{n^\eta \frac{(\widehat{\theta}_i(n_i) - \theta_i) - (\widehat{\theta}_{(k)}(n_{(k)}) - \theta_{(k)})}{\sqrt{(V_i(n_i)/\zeta_i^{2\eta}) + (V_{(k)}(n_{(k)})/\zeta_{(k)}^{2\eta})}} \leq \gamma,\ \forall\ i \neq (k)\right\}.$$

Assumption 1 implies

$$(n^\eta(\widehat{\theta}_i(n_i) - \theta_i), i = 1, \ldots, k) = \left(\frac{n_i^\eta}{\zeta_i^\eta}(\widehat{\theta}_i(n_i) - \theta_i), i = 1, \ldots, k\right) \Rightarrow (Y_i, i = 1, \ldots, k) \tag{8}$$

as $n \to \infty$, where $Y_1, \ldots, Y_k$ are independent with each $Y_i \sim N(0, \sigma_i^2/\zeta_i^{2\eta})$. It then follows from Assumption 2 and Theorem 3.9 of Billingsley (1999) that

$$(n^\eta(\widehat{\theta}_i(n_i) - \theta_i), V_i(n_i) : i = 1, \ldots, k) \Rightarrow (Y_i, \sigma_i^2 : i = 1, \ldots, k) \tag{9}$$

as $n \to \infty$. Thus, the continuous mapping theorem (e.g., Billingsley, 1999, Theorem 2.7) yields

$$\left(n^\eta \frac{(\widehat{\theta}_i(n_i) - \theta_i) - (\widehat{\theta}_{(k)}(n_{(k)}) - \theta_{(k)})}{\sqrt{(V_i(n_i)/\zeta_i^{2\eta}) + (V_{(k)}(n_{(k)})/\zeta_{(k)}^{2\eta})}} : i \neq (k) \right) \Rightarrow (Z_i : i \neq (k)) \tag{10}$$

as $n \to \infty$, where

$$Z_i = \frac{Y_i - Y_{(k)}}{\sqrt{(\sigma_i^2/\zeta_i^{2\eta}) + (\sigma_{(k)}^2/\zeta_{(k)}^{2\eta})}} \sim N(0,1).$$

Consequently, the absolute continuity of the distributions of Z_i, $i \neq (k)$, guarantees

$$\lim_{n\to\infty} P(G(n)) = P\{Z_i \leq \gamma,\, \forall\, i \neq (k)\} \tag{11}$$

by the Portmanteau Theorem (e.g., Billingsley, 1999, Theorem 2.1). Note that $\mathrm{Cov}(Z_i, Z_j) = \lambda_i\lambda_j > 0$ for $i \neq j$, where $\lambda_i = ((\sigma_{(k)}^2/\zeta_{(k)}^{2\eta})/((\sigma_i^2/\zeta_i^{2\eta}) + (\sigma_{(k)}^2/\zeta_{(k)}^{2\eta})))^{1/2}$, so Slepian's (1962) inequality implies

$$P\{Z_i \leq \gamma,\, \forall\, i \neq (k)\} \;>\; \prod_{i\neq(k)} P\{Z_i \leq \gamma\} \;=\; \prod_{i\neq(k)} (1-\alpha)^{1/(k-1)} \;=\; 1-\alpha \tag{12}$$

since γ satisfies $\Phi(\gamma) = (1-\alpha)^{1/(k-1)}$ and each $Z_i \sim N(0,1)$. Thus, (11) and (12) establish (7), completing the proof.

6.2 Proof of Theorem 2.

Since $n_i = \zeta_i n$ with $\zeta_i > 0$ fixed, we have by (9) and arguing as in (11) that

$$\begin{aligned}
&P\{\theta_i - \theta_j \in I'_{i,j}(\boldsymbol{n}),\ \forall\, i<j\} \\
&\quad = \;\; P\left\{ \left| n^\eta(\widehat{\theta}_i(n_i) - \theta_i) - n^\eta(\widehat{\theta}_j(n_j) - \theta_j) \right| \leq \gamma' \sqrt{(V_i(n_i)/\zeta_i^{2\eta}) + (V_j(n_j)/\zeta_j^{2\eta})},\ \forall\, i<j \right\} \\
&\quad \to \;\; P\left\{ |Y_i - Y_j| \leq \gamma' \sqrt{(\sigma_i^2/\zeta_i^{2\eta}) + (\sigma_j^2/\zeta_j^{2\eta})},\ \forall\, i<j \right\} \\
&\quad \geq \;\; P\left\{ |Y'_i - Y'_j| \leq \gamma'\sqrt{2},\ \forall\, i<j \right\} \;=\; 1-\alpha,
\end{aligned}$$

as $n \to \infty$ since $\gamma' = q_k/\sqrt{2}$, where the inequality follows from Hayter (1984) and $Y_1', \ldots, Y_k'$ are i.i.d. $N(0,1)$ random variables.

6.3 Proof of Theorem 3

The result immediately follows from Theorem 2 and the following lemma, due to Tukey (1953); also see pages 81–82 of Hochberg and Tamhane (1987).

Lemma 1 *Let $x = (x_1, x_2, \ldots, x_k) \in \Re^k$, and let $\xi_{i,j}$, $1 \le i < j \le k$, be nonnegative real numbers. Then $|x_i - x_j| \le \xi_{i,j}$ for all $i < j$ if and only if*

$$\left|\sum_{i=1}^{k} c_i x_i\right| \le \frac{2}{\sum_{l=1}^{k} |c_l|} \sum_{i=1}^{k} \sum_{j=1}^{k} c_i^+ c_j^- \xi_{i,j} \quad \forall\, \boldsymbol{c} = (c_1, c_2, \ldots, c_k) \in C.$$

6.4 Proof of Theorem 4

We need the following variation of Hölder's inequality.

Lemma 2 *Let $p, q \ge 1$ be such that $1/p + 1/q = 1$. Let $x = (x_1, x_2, \ldots, x_k) \in \Re^k$, and let $\xi \ge 0$ be a real number. Then $|\sum_{i=1}^{k} a_i x_i| \le \xi (\sum_{i=1}^{k} |a_i|^p)^{1/p}$ for all $\boldsymbol{a} = (a_1, \ldots, a_k) \in \Re^k$ if and only if $\left(\sum_{i=1}^{k} x_i^q\right)^{1/q} \le \xi$.*

Let $b_i = a_i \sqrt{V_i(n_i)/n_i^{2\eta}}$ and $\boldsymbol{b} = (b_1, \ldots, b_k)$. Then since $n_i = \zeta_i n$,

$$\begin{aligned}
& P\left\{\sum_{i=1}^{k} a_i \theta_i \in \bar{I}_{\boldsymbol{a},p}(\boldsymbol{n}),\ \forall\, \boldsymbol{a} \in \Re^k\right\} \\
&= P\left\{\left|\sum_{i=1}^{k} b_i \left(\frac{\widehat{\theta}_i(\zeta_i n) - \theta_i}{\sqrt{V_i(\zeta_i n)/(\zeta_i n)^{2\eta}}}\right)\right| \le \bar{\gamma}_p \left(\sum_{i=1}^{k} |b_i|^p\right)^{1/p},\ \forall\, \boldsymbol{b} \in \Re^k\right\} \\
&= P\left\{\left[\sum_{i=1}^{k} \left(\frac{\widehat{\theta}_i(\zeta_i n) - \theta_i}{\sqrt{V_i(\zeta_i n)/(\zeta_i n)^{2\eta}}}\right)^q\right]^{1/q} \le \bar{\gamma}_p\right\}
\end{aligned}$$

by Lemma 2. Using arguments similar to those applied to establish (10), we can show that

$$\left(\frac{\widehat{\theta}_i(\zeta_i n) - \theta_i}{\sqrt{V_i(\zeta_i n)/(\zeta_i n)^{2\eta}}},\ i = 1, \dots, k \right) \Rightarrow (R_i, i = 1, \dots, k)$$

as $n \to \infty$, where $R_1, \dots, R_k$ are i.i.d. $N(0,1)$. The continuous mapping theorem then implies

$$\left[\sum_{i=1}^{k} \left(\frac{\widehat{\theta}_i(\zeta_i n) - \theta_i}{\sqrt{V_i(\zeta_i n)/(\zeta_i n)^{2\eta}}} \right)^q \right]^{1/q} \Rightarrow \left[\sum_{i=1}^{k} R_i^q \right]^{1/q}$$

as $n \to \infty$. Hence, since each R_i has an absolutely continuous distribution, the Portmanteau theorem implies

$$\lim_{n\to\infty} P\left\{ \sum_{i=1}^{k} a_i\theta_i \in \bar{I}_{\boldsymbol{a},p}(\boldsymbol{n}),\ \forall\, \boldsymbol{a} \in \Re^k \right\} = P\left\{ \left[\sum_{i=1}^{k} R_i^q \right]^{1/q} \le \bar{\gamma}_p \right\},$$

and the result follows by the definition of $\bar{\gamma}_p$.

Acknowledgments

The author thanks the anonymous reviewers for providing comments that helped improve the paper.

References

Asmussen, S. (2003), *Applied Probability and Queues*, New York: Springer, 2nd ed.

Billingsley, P. (1999), *Convergence of Probability Measures*, New York: John Wiley & Sons, 2nd ed.

Dalal, S. R. (1978), "Simultaneous confidence procedures for univariate and multivariate Behrens-Fisher types problems," *Biometrika*, 65, 221–224.

Damerdji, H. (1991), "Strong consistency and other properties of the spectral variance estimator," *Management Science*, 37, 1424–1440.

— (1994), "Strong consistency of the variance estimator in steady-state simulation output analysis," *Mathematics of Operations Research*, 19, 494–512.

Dunnett, C. W. (1955), "A multiple comparison procedure for comparing several treatments with a control," *Journal of the American Statistical Association*, 50, 1096–1121.

Fishman, G. S. (1978), *Principles of Discrete Event Simulation*, New York: Wiley.

Glynn, P. W. and Iglehart, D. L. (1990), "Simulation output analysis using standardized time series," *Mathematics of Operations Research*, 15, 1–16.

— (1993), "Conditions for the Applicability of the Regenerative Method," *Management Science*, 39, 1108–1111.

Glynn, P. W. and Whitt, W. (1992a), "The asymptotic efficiency of simulation estimators," *Operations Research*, 40, 505–520.

— (1992b), "The Asymptotic Validity of Sequential Stopping Rules in Stochastic Simulations," *Annals of Applied Probability*, 2, 180–198.

Gnedenko, B. V. and Kolmogorov, A. N. (1968), *Limit Distributions For Sums of Independent Random Variables*, Reading, Massachusetts: Addison-Wesley.

Hayter, A. J. (1984), "A proof of the conjecture that the Tukey-Kramer multiple comparisons procedure is conservative," *Annals of Statistics*, 12, 61–75.

— (1985), "A Study of the Tukey Multiple Comparisons Procedure Including a Proof of the Tukey Conjecture for Unequal Sample Sizes," Ph.D. thesis, Cornell University, Ithaca, NY.

Hjort, N. L. (1988), “On Large-Sample Multiple Comparison Methods,” *Scandinavian Journal of Statistics*, 15, 259–271.

Hochberg, Y. and Tamhane, A. C. (1987), *Multiple Comparison Procedures*, New York: Wiley.

Hsu, J. C. (1984), “Constrained simultaneous confidence intervals for multiple comparisons with the best,” *Annals of Statistics*, 12, 1136–1144.

— (1996), *Multiple Comparisons: Theory and Methods*, London: Chapman & Hall.

Kiefer, J. and Wolfowitz, J. (1952), “Stochastic estimation of the maximum of a regression function,” *Annals of Mathematical Statistics*, 23, 462–466.

Kramer, C. Y. (1956), “Extension of multiple range tests to group means with unequal numbers of replications,” *Biometrics*, 12, 307–310.

Lukacs, E. (1956), “Characterization of Populations by Properties of Suitable Statistics,” in *Proceedings of the 3rd Berkeley Symposium on Mathematical Statistics and Probability, Volume 2*, Los Angeles and Berkeley: University of California Press, pp. 195–214.

Nakayama, M. K. (1997), “Multiple-comparison procedures for steady-state simulations,” *Annals of Statistics*, 25, 2433–2450.

— (2007), “Single-Stage Multiple-Comparison Procedure for Quantiles and Other Parameters,” in *Proceedings of the 2007 Winter Simulation Conference*, eds. Henderson, S. G., Biller, B., Hsieh, M.-H., Shortle, J., Tew, J. D., and Barton, R. R., Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, p. to appear.

Piegorsch, W. W. (1991), “Multiple comparisons for analyzing dichotomous response,” *Biometrics*, 47, 45–52.

R Development Core Team (2008), *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, ISBN 3-900051-07-0.

Ruppert, D. (1982), "Almost sure approximations to the Robbins-Munro and Kiefer-Wolfowitz processes with dependent noise," *Annals of Probability*, 10, 178–187.

Scheffé, H. (1953), "A method for judging all contrasts in the analysis of variance," *Biometrika*, 40, 87–104.

Schruben, L. W. (1983), "Confidence interval estimation using standardized time series," *Operations Research*, 31, 1090–1108.

Serfling, R. J. (1980), *Approximation Theorems of Mathematical Statistics*, New York: John Wiley & Sons.

Slepian, D. (1962), "The one-sided barrier problem for Gaussian noise," *Bell System Technical Journal*, 41, 463–501.

Spurrier, J. D. and Isham, S. P. (1985), "Exact simultaneous confidence intervals for pairwise comparisons of three normal means," *Journal of the American Statistical Association*, 80, 438–432.

Tukey, J. W. (1953), "The Problem of Multiple Comparisons," Mimeographed monograph.

Ventner, J. H. (1967), "An extension of the Robbins-Munro procedure," *Annals of Mathematical Statistics*, 38, 181–190.