SCHOOL OF OPERATIONS RESEARCH
AND INDUSTRIAL ENGINEERING
COLLEGE OF ENGINEERING
CORNELL UNIVERSITY
ITHACA, NY 14853

TECHNICAL REPORT NO. 888

January 1990

# EXACT CALCULATIONS FOR SEQUENTIAL $t$, $\chi^2$ AND $F$ TESTS

by

Christopher Jennison[1] and Bruce W. Turnbull

[1]Present address: School of Mathematical Sciences, University of Bath, Bath, BA2 7AY, U.K.

This research was supported in part by grant GM 28364 from the U.S. National Institutes of Health.

# Exact calculations for sequential $t$, $\chi^2$ and $F$ tests.

BY CHRISTOPHER JENNISON

*School of Mathematical Sciences, University of Bath, Bath, BA2 7AY, U.K.*

AND BRUCE W. TURNBULL

*School of Operations Research and Industrial Engineering, Cornell University, Ithaca, New York, NY 14853, U.S.A.*

## SUMMARY

Sequential and group sequential procedures are proposed for monitoring repeated $t$, $\chi^2$ or $F$ statistics. These can be used to construct hypothesis tests or repeated confidence intervals when the parameter of interest is a normal mean with unknown variance or a multivariate normal mean with variance matrix known or known up to a scale factor. Exact methods for calculating error probabilities and sample size distributions are described and tables of critical values needed to implement the procedures are provided.

*Some key words*: Group sequential test; Repeated significance test; Repeated confidence interval; Sequential chi-square test; Sequential $t$-test; Sequential $F$-test.

## 1. INTRODUCTION

We consider group sequential designs for independent identically distributed normal observations. We suppose observations are taken in groups with $g_k$ observations in the $k$th group; important special cases are equally sized groups, $g_k \equiv g$, and the fully sequential case, $g_k \equiv 1$. Tests for the mean of univariate observations with known variance have been treated extensively. Wald (1947), Anderson (1960) and Siegmund (1985) derived accurate analytic approximations to performance measures of certain sequential tests; more general numerically exact methods have been used by Aroian (1968, 1976), Armitage *et al.* (1969) and many subsequent authors. For more detailed surveys see the books by Armitage (1975), Whitehead (1983) and Wetherill and Glazebrook (1986).

In this paper we present numerically exact methods for sequential and group sequential designs for univariate normal response with unknown variance, the sequential $t$-test, and for multivariate normal response with known or unknown variance, the sequential $\chi^2$ and $F$ tests.

## 2. UNIVARIATE NORMAL OBSERVATIONS WITH UNKNOWN VARIANCE

Let $X_1, X_2, \ldots$ be independent normal observations with mean $\mu$ and unknown variance $\sigma^2$. We consider the problem of testing $H_0 : \mu = \mu_0$ against the two-sided alternative $H_1 : \mu \neq \mu_0$ with a specified power at $\mu = \mu_0 \pm \delta\sigma$ for given $\delta$. Allowing the value of $\mu$ at which the power condition is satisfied to depend on the unknown $\sigma$ is quite reasonable in some applications: Rushton (1950) shows that this is a natural requirement in an acceptance sampling problem and Whitehead (1983, p. 64) describes a similar problem in a two period crossover trial.

Early work on the sequential $t$-test by Wald (1947), Rushton (1950, 1952), Barnard (1952) and Hajnal (1961) used analytic approximations to calculate error probabilities for sequential $t$-tests with constant likelihood ratio boundaries. Schneiderman and Armitage (1962) and Myers, Schneiderman and Armitage (1966) presented closed sequential $t$-tests derived heuristically from sequential tests for observations with known variance. Suich and Iglewicz (1970) and Alexander and Suich (1973) proposed an approximate modification of Anderson's (1960) method to the case of unknown variance. Siegmund (1985, p. 116) derived analytic approximations for a modified repeated significance test based on the $t$-statistic.

Schmee (1974) has calculated exact values for the error probabilities of sequential $t$-tests. He used a different recursive method from that which we shall propose and his computed results were limited to tests with a maximum of 5 observations.

All the above methods concern the fully sequential approach, i.e., $g_k \equiv 1$. Pocock (1977, p. 195-6) has suggested an approximate group sequential $t$-test based on the nominal significance levels appropriate to the case of known variance. Our exact calculations show that this approximate method is reasonably accurate, as is the analogous approximation for sequential $\chi^2$ tests.

We define $n_k = g_1 + \ldots + g_k$, the cumulative sample size at stage $k$, and let $X_{n_{k-1}+1}, \ldots, X_{n_k}$ denote the $g_k$ observed responses in the $k$th group. We define

$$\bar{X}_k = \frac{1}{n_k} \sum_{i=1}^{n_k} X_i$$

and

$$s_k^2 = \frac{\sum_{i=1}^{n_k} (X_i - \bar{X}_k^2)}{n_k - 1} ,$$

the sample mean and standard estimate of $\sigma^2$ at the $k$th analysis. We now form the usual $t$-statistic,

$$T_k = \sqrt{n_k}\ (\bar{X}_k - \mu_0)\ / s_k .$$

A group sequential test is specified by a maximum number of groups, $K$, and a partition $\{A_k, B_k, C_k\}$ of the real line for each stage $k=1,\ldots,K$. The test stops at the first stage that $T_k \in A_k \cup B_k$; $H_0$ is chosen if $T_k \in A_k$ and $H_1$ if $T_k \in B_k$. As long as $T_k \in C_k$ the test continues. The final set $C_K$ is necessarily empty. The type I error probability of this test is given by

$$\alpha = \sum_{k=1}^{K} \pi_k \tag{2.1}$$

where

$$\pi_k = P_{\mu_0}(T_1 \in C_1, \ldots, T_{k-1} \in C_{k-1}, T_k \in B_k) \tag{2.2}$$

for $k \geq 1$. Thus $\pi_k$ is the probability under $H_0$ of stopping at stage $k$ and choosing $H_1$; this quantity can be viewed as the error probability 'spent' at the $k$th interim analysis. The nominal significance level at the $k$th analysis is defined to be the marginal probability $\alpha_k = P_{\mu_0}(T_k \in B_k)$ in the absence of an early stopping rule. The quantities $\{\pi_k; k \geq 1\}$ and $\{\alpha_k : k \geq 1\}$ should not be confused.

A symmetric group sequential test without early stopping to accept $H_0$ is given by

$$A_k = \varnothing \ \ (1 \leq k \leq K-1), \qquad A_K = (-c_K, c_K)$$
$$B_k = (-\infty, -c_k] \cup [c_k, \infty) \quad 1 \leq k \leq K$$
$$C_k = (-c_k, c_k) \ \ (1 \leq k \leq K-1), \quad C_K = \varnothing$$

where $\varnothing$ denotes the empty set. Thus, constants $c_1, c_2, \ldots, c_K$ are specified and the test stops to accept $H_1$ if $|T_k| \geq c_k$ at any $1 \leq k \leq K$. The test is truncated at stage $K$ and $H_0$ is accepted if $|T_K| < c_K$. For this stopping rule,

$$\pi_k = P_{\mu_0} \{ |T_1| < c_1, \ldots, |T_{k-1}| < c_{k-1}, |T_k| \geq c_k \}. \tag{2.3}$$

For given boundary values $c_1, c_2, \ldots, c_K$ or, equivalently, nominal levels $\alpha_1, \alpha_2, \ldots, \alpha_K$, the exact probabilities $\{\pi_k; k=1,\ldots,K\}$ can be calculated recursively as described below. Let

$$\bar{X}(k) = \frac{1}{g_k} \sum_{i=n_{k-1}+1}^{n_k} X_i$$

be the mean of the $k$th group. By straightforward algebra,

$$\begin{aligned}(n_{k+1}-1)s_{k+1}^2 &= (n_k-1)s_k^2 + \sum_{i=n_k+1}^{n_{k+1}} \{X_i - \bar{X}(k+1)\}^2 \\ &+ \frac{n_k n_{k+1}}{g_{k+1}} (\bar{X}_{k+1} - \bar{X}_k)^2 .\end{aligned} \tag{2.4}$$

Define the scale invariant quantities

$$Z_k = \frac{n_k(\bar{X}_k - \mu_0)}{\sigma}$$

and

$$R_k = \frac{(n_k-1)}{\sigma^2} s_k^2 .$$

The joint distribution of $\{T_1, T_2, \ldots\}$ when observations have mean $\mu$ and variance $\sigma^2$ can be obtained from the joint distribution of $\{Z_1, R_1, Z_2, R_2, \ldots\}$ which, in turn, can be constructed from successive conditional distributions. Firstly, $Z_1$ and $R_1$ are independent with $Z_1 \sim N(n_1(\mu-\mu_0)/\sigma, n_1)$ and $R_1 \sim \chi^2_{n_1-1}$. The conditional distributions of $Z_2$ given $Z_1$ and $R_1$ and of $R_2$ given $Z_1$, $R_1$ and $Z_2$ are

$$\mathcal{P}(Z_2 | Z_1, R_1) \sim N(Z_1 + g_2 \frac{\mu-\mu_0}{\sigma}, g_2)$$

and, using identity (2.4),

$$\mathcal{P}(R_2 | Z_1, R_1, Z_2) \sim R_1 + \frac{n_1 n_2}{g_2} \left[ \frac{Z_1}{n_1} - \frac{Z_2}{n_2} \right]^2 + \chi^2_{g_2-1} .$$

In general

$$\begin{aligned}\mathcal{P}(Z_{k+1} | Z_1, R_1, Z_2, R_2, \ldots, Z_k, R_k) &= \mathcal{P}(Z_{k+1} | Z_k, R_k) \\ &\sim N(Z_k + g_{k+1} \frac{\mu-\mu_0}{\sigma}, g_{k+1})\end{aligned} \tag{2.5}$$

and

$$\mathcal{P}(R_{k+1} \mid Z_1, R_1, Z_2, R_2, \ldots, Z_k, R_k, Z_{k+1}) = \mathcal{P}(R_{k+1} \mid Z_k, R_k, Z_{k+1})$$

$$\sim R_k + \frac{n_k n_{k+1}}{g_{k+1}} \left[ \frac{Z_k}{n_k} - \frac{Z_{k+1}}{n_{k+1}} \right]^2 + \chi^2_{g_{k+1}-1} \cdot \qquad (2.6)$$

Note that when the group size $g_{k+1}$ is equal to 1, the value of $R_{k+1}$ is completely determined by $Z_k$, $R_k$ and $Z_{k+1}$. Combining (2.5) and (2.6) for $k = 0,1,\ldots,K-1$ we can obtain the joint density of $(Z_1, R_1, Z_2, R_2, \ldots, Z_K, R_K)$. Since $T_k = \sqrt{(n_k-1)}\, Z_k / \sqrt{(n_k R_k)}$, this determines the joint density of $\{T_k; k=1,\ldots,K\}$. Precise details of this recursion are given in the Appendix.

These joint densities depend on the parameters $\mu$, $\mu_0$ and $\sigma$ only through $(\mu-\mu_0)/\sigma$. However, the probabilities $\pi_k (k \geq 1)$ in (2.3) are always calculated with $\mu = \mu_0$, and hence they can be computed for specified $\{c_k; k \geq 1\}$ without knowledge of $\mu_0$ or $\sigma$. Conversely, if the $\{\pi_k; k \geq 1\}$ are given, values of $\{c_k; k \geq 1\}$ satisfying (2.3) can be found successively. If $\mu = \mu_0$, $T_1$ has a central $t$-distribution with $g_1 - 1$ degrees of freedom and $c_1$ can be obtained from standard tables. Values of $c_k$ for $k \geq 2$ can be obtained using numerical integration to evaluate the right hand side of (2.3).

Nominal two-sided significance levels of the $t$-statistic are defined by

$$\alpha_k = 2\{1 - \Psi(c_k; n_k - 1)\} \qquad (k \geq 1)$$

where $\Psi(\cdot\,;\,\nu)$ denotes the cumulative distribution function of the $t$-distribution with $\nu$ degrees of freedom. In particular, $\alpha_1 = \pi_1$.

There have been several suggestions in the literature for the known variance case as to how to choose these constants subject to the constraint (2.1). A constant nominal significance level $\alpha_1 = \ldots = \alpha_K = \alpha'$, say, was used by Armitage *et al.* (1969) and, in the group sequential case, by Pocock (1977). Adapting this approach to the group sequential $t$-test with $K$ groups of size $g$, we reject $H_0$ at stage $k$ $(1 \leq k \leq K)$ if

$$2\,\{1 - \Psi(|T_k|\,;\, n_k - 1)\} \leq \alpha'$$

where $\alpha'$ is chosen so that (2.1) is satisfied. We then define $Z_P(K,g,\alpha) = \Phi^{-1}(1 - \alpha'/2)$ where $\Phi$ is the standard normal cumulative distribution function. Table 1 gives values of $Z_P(K,g,\alpha)$ for $\alpha = 0.01$, 0.05 and 0.10, $g = 1$, 2, 3, 5 and 10 and $K = 2, \ldots, 10$. If $g=1$, it is assumed that no analysis is performed at the first stage, $k=1$, which explains the different pattern of the entries in the rows for $g=1$. The critical values for the case of known variance, $Z_P(K,\infty,\alpha)$, which equal the values $z$ in Table 1 of Pocock (1977), are shown for comparison. The entries were calculated by use of numerical integration and the recursive formulae of the Appendix; the error probabilities of tests computed using these constants are within $10^{-4}$ or less of the specified values of $\alpha$.

Pocock (1977, p. 195-6) recommended the use of the same nominal significance level for a repeated $t$-test as is needed for the case of known variance and presented simulation results to support this suggestion. It is clear from Table 1 that only a slight adjustment to this approximation is needed; in fact, the standard errors of Pocock's simulation results do not do justice to the accuracy of his proposal. For example, for 5 groups of 3 observations and an intended $\alpha = 0.05$, using $Z_P = 2.413$ in place of the correct value 2.457 gives an actual error probability of 0.056. The discrepancy here is quite typical for groups of $g = 2$ or more observations but the case $g = 1$ is somewhat different since a $t$-test cannot be performed after the first group of just 1 observation.

[Tables 1 and 2 about here]

There is no restriction to constant nominal significance levels; any of the methods proposed for known variance can be adapted to allow unknown variance. In the case of the O'Brien and Fleming (1979) test for equal group sizes, it is natural to define the $\{c_k;k\geq 1\}$ in terms of significance levels by setting

$$\Psi(c_k;n_k-1) = \Phi\,(Z_{OBF}\,\sqrt{(K/k)}\,)$$

where $Z_{OBF} = Z_{OBF}(K,g,\alpha)$ depends on the number of analyses, $K$, the number of observations per group, $g$, and the error rate, $\alpha$. Values of $Z_{OBF}$ for $K = 2, \ldots, 10$, $g = 1$, 2, 3, 5 and 10 and $\alpha = 0.01$, 0.05 and 0.10 are shown in Table 2, the values of $Z_{OBF}(K,\infty,\alpha)$ corresponding to the case of known variance being included for comparison. Again, if $g$=1, it is assumed that no analysis is performed at the first stage. Broadly speaking, arguments concerning the relative merits of ways to choose the $\{c_k;k\geq 1\}$ will be the same as in the known variance case; see the discussions by Pocock (1982) and Jennison and Turnbull (1989).

Repeated confidence intervals for $\mu$ can also be constructed using the same constants $\{c_k\}$. The repeated confidence interval at the $k$th stage is given by $I_k = (\bar{X}_k - c_k s_k/\sqrt{n_k},\ \bar{X}_k + c_k s_k/\sqrt{n_k})$. For a full discussion of the motivation and applications of repeated confidence intervals, see Jennison and Turnbull (1989).

The same method of numerical calculation could be used for other types of stopping regions, such as the triangular or double triangular regions described by Whitehead (1983), the 'wedge' regions proposed by Schneiderman and Armitage (1962), Myers, Schneiderman and Armitage (1966) and Gould and Pecore (1982) or the modified repeated significance tests of Siegmund (1985). Alternatively, stopping regions for one-sided tests can be constructed from repeated confidence intervals in the manner described by Jennison and Turnbull (1984, 1989).

The power functions of these tests can be obtained by repeating the above calculations with $\mu \neq \mu_0$. In particular, the group size needed to ensure power $1-\beta$ at $\mu = \mu_0 \pm \delta\sigma$ for given $\beta$ and $\delta$ can be found by a simple one-dimensional search. As noted earlier, it is natural in some applications to specify power at a value of $\mu$ a fixed multiple of $\sigma$ away from $\mu_0$. If, instead, power is to be guaranteed at a fixed value of $\mu$, a sequential test with predetermined boundary does not exist; see Dantzig (1940). The two-sample test of Stein (1945) might be adapted to this setting. Alternatively, an adaptive approach in which group sizes are determined by the current estimate of $\sigma^2$ should give at least an approximate procedure.

Finally we note that our methodology can be extended to procedures in which the average range method is used for estimating $\sigma^2$. Such procedures have applications in multiple sampling inspection plans by variables.

## 3. MULTIVARIATE OBSERVATIONS

We assume that multivariate normal observations $X_1, X_2, \ldots$ of dimension $p \geq 1$ with mean vector $\mu$ and covariance matrix $\sigma^2 \Sigma$ are available sequentially. Here $\Sigma$ is known and the scale factor $\sigma^2$ may be known or unknown. By applying the linear transformation $X_i \rightarrow \Sigma^{-1/2} X_i$, we can assume $\Sigma = I_p$ without loss of generality. We shall describe the use of sequentially computed $\chi^2$ and $F$ statistics to test a null hypothesis of the form $H_0$: $\mu = \mu_0$ and to construct repeated confidence sets for $\mu$.

Data of this form might arise in a medical trial with multiple endpoints: Whitehead (1986) describes several examples including a trial concerned with both length and weight of newborn babies. Another application is to industrial sampling inspection where acceptance of batches is is based on several variables; see Jackson and Bradley (1961) for an example. By applying the Helmert transformation (Stuart and Ord, 1987, p. 350) to eliminate location effects, the same methods can also be used to test equality of $p$+1 means in a one-way analysis of variance layout.

Wetherill and Glazebrook (1986, Chapter 5) have surveyed the literature on sequential $\chi^2$ and $F$ tests. Most papers have been concerned with asymptotic or empirical results for procedures with constant likelihood ratio boundaries which are generalizations of Wald's (1947) sequential probability ratio test. Of particular note is the paper by Siegmund (1980), who derived analytic approximations for operating characteristics and sample sizes of sequential $\chi^2$ and $F$ tests, using a generalization of repeated significance test boundaries; see also Siegmund (1985, p. 111).

In the next section, we present numerically exact calculations and tables of constants for sequential and group sequential tests of the null hypothesis $H_0$: $\mu = \mu_0$, the repeated $\chi^2$ test. Repeated confidence ellipsoids for $\mu$ can also be constructed using the same results. In Section 3.2, we indicate how these methods can be extended to give a sequential $F$-test in the case of unknown $\sigma^2$.

### 3.1 Multivariate normal mean with known covariance matrix

We suppose that observations $X_i$, $i = 1, 2, \ldots$, are multivariate normal with mean $\mu$ and covariance matrix $\sigma^2 I_p$. Observations are collected in groups and, as before, we denote the $i$th group size by $g_i$ and let $n_k = g_1 + \ldots + g_k$ be the cumulative sample size at the $k$th analysis ($k \geq 1$). We define $Y_k = X_1 + \ldots + X_{n_k}$. The test of $H_0 : \mu = \mu_0$ will be based on successive values of the statistic

$$S_k = S_k(\mu_0) = \frac{1}{n_k\, \sigma^2} \; \| \; Y_k - n_k \mu_0 \; \|^2 \; . \tag{3.1}$$

We consider a test that stops to reject $H_0$ the first time that $S_k$ exceeds a critical value $c_k$. The hypothesis $H_0$ is accepted at the final stage $K$ if $S_K < c_K$. The critical values $\{ c_k \; ; \; 1 \leq k \leq K \}$ are to be constructed so that the size of the test is equal to some prespecified level $\alpha$. Note that this will also imply that the sequence of repeated confidence sets $\{ \mu : S_k(\mu) \leq c_k \}$ for $k = 1, \ldots, K$ has level $1 - \alpha$ .

The marginal distribution of $S_k$ is chi-squared with $p$ degrees of freedom and non-centrality parameter $n_k \| \mu - \mu_0 \|^2 / \sigma^2$, denoted by $\chi_p^2(n_k \| \mu - \mu_0 \|^2 / \sigma^2)$. In particular we use this to obtain the distribution of $S_1$. We proceed to show how to construct recursively the joint distribution of $\{ S_k ; \; k = 1, \ldots, K \}$; from this distribution with $\mu = \mu_0$, appropriate critical values $\{ c_k ; k = 1, \ldots, K \}$ can then be found. To derive the joint distribution of the $\{ S_k ; k \geq 1 \}$ under $\mu = \mu_0$ we note the identity

$$S_{k+1} = \frac{1}{n_{k+1} \sigma^2} \; \| (Y_k - n_k \mu_0) + (Y_{k+1} - Y_k - g_{k+1} \mu_0) \|^2$$

where $Y_{k+1} - Y_k - g_{k+1} \mu_0 \sim N_p(0, g_{k+1} \sigma^2 I_p)$ is independent of $Y_k$ and, hence, of $S_k$. Thus, conditionally on $Y_1, \ldots, Y_k$,

$$S_{k+1} \sim \frac{g_{k+1}}{n_{k+1}} \; \chi_p^2 \left[ \frac{\| Y_k - n_k \mu_0 \|^2}{g_{k+1} \sigma^2} \right] = \frac{g_{k+1}}{n_{k+1}} \; \chi_p^2 \left[ \frac{n_k}{g_{k+1}} \; S_k \right] . \tag{3.2}$$

Therefore the sequence $\{ S_k ; k \geq 1 \}$ is Markov and the joint distribution of $\{ S_k ; k \geq 1 \}$ under $\mu_0$ can be constructed by multiplying together the conditional densities of $S_{k+1}$ given $S_k$ for $k \geq 1$. Critical values $\{ c_k ; k \geq 1 \}$ based on exit probabilities $\{ \pi_k ; k \geq 1 \}$ or nominal significance levels $\{ \alpha_k ; k \geq 1 \}$ can be calculated in the same way as for the univariate normal, known variance case but with non-central $\chi^2$ densities replacing the normal densities.

For a Pocock (1977) type boundary with constant nominal significance levels, we set $c_1 = \ldots = c_K = C_P(p, K, \alpha)$, say. For boundaries analogous to those of O'Brien and

Fleming (1979) we set $c_k = (K/k)C_{OBF}(p,K,\alpha)$. For the case of equal group sizes, the required constants $C_P(p,K,\alpha)$ and $C_{OBF}(p,K,\alpha)$ are tabulated in Tables 3 and 4, respectively, for values of $\alpha = 0.01$, $0.05$ and $0.10$, $K = 1, \ldots, 10$, 20 and 50 and $p = 1, \ldots, 5$. The entries were calculated using numerical integration and the recursive formulae described below. Note that when $p = 1$, $C_P(1,K,\alpha) = Z_P^2(K,\infty,\alpha)$ and $C_{OBF}(1,K,\alpha) = Z_{OBF}(K,\infty,\alpha)^2$. When $K = 1$, $c_1 = C_P(p,1,\alpha) = C_{OBF}(p,1,\alpha)$, the usual percentage point of the $\chi_p^2$ distribution.

[Tables 3 and 4 about here]

As for the sequential $t$-test, an approximate method for obtaining critical values is to use percentage points of the $\chi_p^2$ distribution with the same nominal significance levels as is required for univariate normal observations with known variance. This approximation also works reasonably well. For example, for $\alpha = 0.05$ and 5 groups of observations in a $p = 5$ dimensional problem the actual Type I error probability of a Pocock-type boundary constructed in this way is 0.053.

For power calculations it is necessary to evaluate the exit probabilities for the sequence $\{S_k; k \geq 1\}$ under the non null case $\mu \neq \mu_0$. Without loss of generality, we can take $\mu - \mu_0 = \|\mu - \mu_0\| (1,0,0,\ldots,0)$. We write $Y_k - n_k\mu_0 = (Y_k^{(1)}, Y_k^{(2)})$, where the scalar $Y_k^{(1)}$ is the first element of $Y_k - n_k\mu_0$ and $Y_k^{(2)}$ is a $(p-1)$-vector denoting the remaining elements. Then, given $Y_k^{(1)}$ and $W_k = \|Y_k^{(2)}\|^2$, the conditional distributions of $Y_{k+1}^{(1)}$ and $W_{k+1} = \|Y_{k+1}^{(2)}\|^2$ are independent and given by

$$\mathcal{P}(Y_{k+1}^{(1)} \mid Y_k^{(1)}, W_k) \sim N(Y_k^{(1)} + g_{k+1}\|\mu - \mu_0\|, \; g_{k+1}\sigma^2)$$

and

$$\mathcal{P}(W_{k+1} \mid Y_k^{(1)}, W_k) \sim g_{k+1}\sigma^2 \chi_{p-1}^2(W_k / g_{k+1}\sigma^2)\,.$$

The conditional distribution of $S_{k+1}$ then follows from the relation

$$S_{k+1} = \frac{1}{n_{k+1}\sigma^2}\{(Y_{k+1}^{(1)})^2 + W_{k+1}\}$$

As in Section 2, the recursive formulae to determine exact probabilities now involve double rather than single integrals. Further details of their evaluation are omitted.

### 3.2 The case of unknown scale factor, $\sigma$

We consider the same situation as in Section 3.1 where the multivariate normal observations $X_1, X_2, \ldots$ have covariance matrix $\sigma^2\Sigma$, $\Sigma$ is a known $p \times p$ positive definite matrix but now the scalar $\sigma^2$ is unknown. Again we may take $\Sigma = I_p$ without loss of generality. We proceed as in Section 3.1, but replace $\sigma^2$ by the estimator

$$s_k^2 = \frac{\sum_{i=1}^{n_k} \|X_i - \bar{X}_k\|^2}{p(n_k - 1)}$$

where $\bar{X}_k = Y_k / n_k$.

Group sequential tests and repeated confidence sets are therefore based on the sequentially computed $F$-statistics $\{F_k; k \ge 1\}$ given by

$$F_k = F_k(\mu_0) = \frac{\|Y_k - n_k \mu_0\|^2 / n_k}{\sum_{i=1}^{n_k} \|X_i - \bar{X}_k\|^2 / (n_k - 1)} . \tag{3.3}$$

Dividing numerator and denominator by $p$ we see that marginally $F_k \sim F_{p, p(n_k - 1)}$ under $\mu = \mu_0$.

We now derive the joint distribution of $\{F_k; k = 1, \ldots, K\}$ when $\mu = \mu_0$. Write

$$U_k = \|Y_k - n_k \mu_0\|^2$$

so

$$F_k = \frac{U_k}{p n_k s_k^2} .$$

We proceed recursively. First, $U_1$ and $s_1^2$ are independent with $n_1 \sigma^2 \chi_p^2$ and $\sigma^2 \chi^2_{p(n_1 - 1)} / \{p(n_1 - 1)\}$ distributions respectively. For $k \ge 1$, let $e = (Y_k - n_k \mu_0) / \|Y_k - n_k \mu_0\|$ be a unit vector in the direction $Y_k - n_k \mu_0$ and let the scalar random variable $A$ and vector random variable $B$ be defined by the orthogonal decomposition.

$$Y_{k+1} - Y_k - g_{k+1} \mu_0 = Ae + B$$

where $B^T e = 0$.

Now, $Ae + B$ is distributed as $N_p(0, g_{k+1} \sigma^2 I_p)$ and by the spherical symmetry of this distribution, $A$ and $B$ are independent of each other, as well as of $Y_k$. Using these facts and the multivariate analogue of (2.4),

$$p(n_{k+1} - 1) s_{k+1}^2 =$$

$$p(n_k - 1) s_k^2 + \frac{n_k n_{k+1}}{g_{k+1}} \left\| \frac{Y_{k+1}}{n_{k+1}} - \frac{Y_k}{n_k} \right\|^2 + \sigma^2 \chi^2_{p(g_{k+1} - 1)} , \tag{3.4}$$

we can derive the conditional joint distribution of $U_{k+1}$ and $s_{k+1}^2$, given $U_i$ and $s_i^2$ for $1 \le i \le k$, from the relations

$$U_{k+1} = (\sqrt{U_k} + A)^2 + \|B\|^2 \tag{3.5}$$

and

$$p(n_{k+1}-1)s_{k+1}^2 =$$

$$p(n_k-1)s_k^2 + \frac{n_k n_{k+1}}{g_{k+1}} \left\{ \left[ \frac{n_k A - g_{k+1}\sqrt{U_k}}{n_k n_{k+1}} \right]^2 + \frac{\|B\|^2}{n_{k+1}^2} \right\} + C \qquad (3.6)$$

where $A$, $\|B\|^2$ and $C$ are independent scalar random variables with $A \sim N(0, g_{k+1}\sigma^2)$, $\|B\|^2 \sim g_{k+1}\sigma^2\ \chi_{p-1}^2$ and $C \sim \sigma^2\ \chi_{p(g_{k+1}-1)}^2$.

Note that $U_{k+1}$ and $s_{k+1}^2$ depend on the past history of the bivariate process $\{(U_i, s_i^2); i=1,\ldots,k\}$, only through $U_k$ and $s_k^2$. Using techniques similar to those used in Section 2, boundary values $\{c_k; k\geq 1\}$ can be computed so that the test has the required size $\alpha$. Since the distribution of the sequence $\{F_k; k\geq 1\}$ does not depend on $\sigma^2$, these values can be calculated using any convenient value, $\sigma^2=1$, say.

### 3.3. Extensions to other multivariate models

If, in Section 3.2, the covariance matrix $\Sigma$ were completely unknown it could be estimated by the sample covariance matrix $\hat{\Sigma}$. The natural statistic on which to base group sequential tests and repeated confidence sets would then be Hotelling's $T^2$ statistic (Jackson and Bradley, 1961). If the initial group sizes are large, a test with *approximate* level $1-\alpha$ might be constructed using the methods of Sections 3.1 and 3.2. Nominal significance levels $\{\alpha_k; k\geq 1\}$ corresponding to critical values calculated using the methods of Section 3.1 or 3.2 could be converted to critical values $\{c_k; k\geq 1\}$ for the $T^2$ statistic using percentage points of the Hotelling distribution with the appropriate degrees of freedom. Exact methods would involve joint distributions of repeated Wishart variables!

### 3.4. Survival data and contingency tables

The univariate normal known variance case can serve as a basis for survival data and contingency table data methods as described, for example, in Jennison and Turnbull (1989). The methods of Section 3 could analogously be developed to handle survival data from trials with three or more treatments or discrete data that can be expressed as $2\times k$ contingency tables, either stratified or unstratified.

## ACKNOWLEDGEMENT

This research was supported in part by grant GM 28364 from the US National Institutes of Health.

## APPENDIX

*The recursive formula for the exit probabilities of the repeated t-statistics.*

Suppose boundary values $c_1,\ldots,c_K$ or, equivalently, nominal levels $\alpha_1,\ldots,\alpha_K$ are given. We wish to determine the exit probabilities $\{\pi_k; k=1,\ldots,K\}$ as defined in (2.3). Maintaining the notation of Section 2, for $k\leq 2$ let

$$F_k(z,r) = P(Z_k\leq z,\ R_k\leq r \text{ and } |T_i|<c_i \text{ for all } 1\leq i\leq k-1)$$

and let

$$f_k(z,r) = \frac{\partial^2 F_k}{\partial z\partial r}.$$

For $k = 1$, we define

$$f_1(z,r) = q_1(z)h_1(r)$$

where $q_1$ is the normal density with mean $n_1(\mu-\mu_0)/\sigma$ and variance $n_1$ and $h_1$ is a $\chi^2_{n_1-1}$ density function. For $k \geq 1$ we can recursively construct

$$f_{k+1}(z,r) = \iint_{C_k} f_k(u,v)q_{k+1}(z\,|\,u,v)h_{k+1}(r\,|\,u,v,z)du\ dv$$

where

$$C_k = \{(u,v):\ v>0,\ |u| < c_k\surd(\frac{n_k v}{n_k-1})\ \},$$

$q_{k+1}(z\,|\,u,v)$ is the conditional density of $Z_{k+1}$ given $Z_k = u$ and $R_k = v$, and $h_{k+1}(r\,|\,u,v,z)$ is the conditional density of $R_{k+1}$ given $Z_k = u$, $R_k = v$ and $Z_{k+1} = z$. The conditional densities $q_{k+1}$ and $h_{k+1}$ are given by Equations (2.5) and (2.6) respectively. Finally, for $k \geq 1$ the exit probabilities are given by

$$\pi_k = \iint_{D_k} f_k(u,v)du\ dv \qquad (k\geq 1)$$

where

$$D_k = \{(u,v):\ v>0,\ |u| \geq c_k\surd(\frac{n_k v}{n_k-1})\}.$$

## REFERENCES

Alexander, R. and Suich, R.C. (1973) A truncated sequential $t$-test for general $\alpha$ and $\beta$. *Technometrics*, **15**, 79-86.

Anderson, T. W. (1960) A modification of the sequential probability ratio test to reduce the sample size. *Ann. Math. Statist.*, **31**, 165-197.

Armitage, P. (1975) *Sequential Medical Trials, Second edition.* Oxford: Blackwell.

Armitage, P., McPherson, C. K. and Rowe, B. C. (1969) Repeated significance tests on accumulating data. *J. R. Statist. Soc.* A, **132**, 235-244.

Aroian, L. A. (1968) Sequential analysis, direct method. *Technometrics*, **10**, 125-132.

Aroian, L. A. (1976) Applications of the direct method in sequential analysis. *Technometrics*, **18**, 301-306.

Barnard, G. A. (1952) The frequency justification of certain sequential tests. *Biometrika*, **39**, 144-150.

Dantzig G. B. (1940) On the non-existence of tests of "Student's" hypothesis having power functions independent of $\sigma$. *Ann. Math. Statist.*, **11**, 186-192.

Gould, A. L. and Pecore, V. J. (1982) Group sequential methods for clinical trials allowing early acceptance of $H_0$ and incorporating costs. *Biometrika*, **69**, 75-80.

Hajnal, J. (1961) A two-sample sequential $t$-test. *Biometrika*, **48**, 65-75.

Jackson, J.E. and Bradley, R.A. (1961) Sequential $\chi^2$ and $T^2$ tests and their application to an acceptance sampling problem. *Technometrics*, **3**, 519-534.

Jennison, C. and Turnbull, B. W. (1984) Repeated confidence intervals for group sequential clinical trials. *Controlled Clinical Trials*, **5**, 33-45.

Jennison, C. and Turnbull, B. W. (1989) Interim analyses: the repeated confidence interval approach (with Discussion). *J. R. Statist. Soc.* B, **51**, 305-361.

Myers, M. H., Schneiderman, M. A. and Armitage, P. (1966) Boundaries for closed (wedge) sequential $t$ test plans. *Biometrika*, **53**, 431-437.

O'Brien, P. C. and Fleming, T. R. (1979) A multiple testing procedure for clinical trials. *Biometrics*, **35**, 549-556.

Pocock, S. J. (1977) Group sequential methods in the design and analysis of clinical trials. *Biometrika*, **64**, 191-199.

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

Rushton, S. (1950) On a sequential $t$-test. *Biometrika*, **37**, 326-333.

Rushton, S. (1952) On a two-sided sequential $t$-test. *Biometrika*, **39**, 302-308.

Schmee, J. (1974) Exact solution for the sequential $t$-test and a new method of sequential estimation. Unpublished Ph.D. thesis. Union College, New York.

Schneiderman, M. A. and Armitage, P. (1962) Closed sequential $t$-tests. *Biometrika*, **49**, 359-366.

Siegmund, D. (1980) Sequential chi-square and $F$ tests and the related confidence intervals. *Biometrika*, **67**, 389-402.

Siegmund, D. (1985) *Sequential Analysis*. New York: Springer-Verlag.

Stein, C. (1945) A two-sample test for a linear hypothesis whose power is independent of the variance. *Ann. Math. Statist.*, **16**, 243-258.

Stuart, A. and Ord, J. K. (1987) *Kendall's Advanced Theory of Statistics*, Vol. 1, 5th edition.

Suich, R. and Iglewicz, B. (1970) A truncated sequential $t$-test. *Technometrics*, **12**, 789-798.

Wald, A. (1947) *Sequential Analysis*. New York: Wiley.

Wetherill, G.B. and Glazebrook, K.D. (1986) *Sequential Methods in Statistics, Third Edition*. London: Chapman and Hall.

Whitehead, J. (1983) *The Design and Analysis of Sequential Clinical Trials*. Chichester: Ellis Horwood.

Whitehead, J. (1986) Supplementary analysis at the conclusion of a sequential clinical trial. *Biometrics*, **42**, 461-471.

**Table 1.** *Constants $Z_P(K,g,\alpha)$ for Pocock type repeated t-tests with constant nominal significance level and overall error rate $\alpha$.*

| Number of obs. per group, $g$ | Number of groups, $K$: 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|
| | | | | | $\alpha = 0.01$ | | | | |
| 1 | 2.576 | 2.792 | 2.903 | 2.977 | 3.032 | 3.068 | 3.099 | 3.121 | 3.146 |
| 2 | 2.801 | 2.914 | 2.986 | 3.036 | 3.074 | 3.107 | 3.133 | 3.153 | 3.169 |
| 3 | 2.797 | 2.907 | 2.977 | 3.026 | 3.063 | 3.093 | 3.118 | 3.139 | 3.157 |
| 5 | 2.790 | 2.897 | 2.965 | 3.014 | 3.051 | 3.081 | 3.105 | 3.126 | 3.144 |
| 10 | 2.782 | 2.887 | 2.954 | 3.002 | 3.039 | 3.068 | 3.093 | 3.114 | 3.132 |
| limit as $g \to \infty$ | 2.772 | 2.873 | 2.939 | 2.986 | 3.023 | 3.053 | 3.078 | 3.099 | 3.117 |
| | | | | | $\alpha = 0.05$ | | | | |
| 1 | 1.960 | 2.203 | 2.323 | 2.400 | 2.455 | 2.495 | 2.532 | 2.558 | 2.582 |
| 2 | 2.220 | 2.344 | 2.418 | 2.472 | 2.512 | 2.543 | 2.570 | 2.591 | 2.611 |
| 3 | 2.210 | 2.329 | 2.404 | 2.457 | 2.498 | 2.530 | 2.556 | 2.579 | 2.598 |
| 5 | 2.198 | 2.315 | 2.389 | 2.442 | 2.482 | 2.514 | 2.541 | 2.563 | 2.583 |
| 10 | 2.188 | 2.303 | 2.376 | 2.428 | 2.468 | 2.500 | 2.527 | 2.550 | 2.569 |
| limit as $g \to \infty$ | 2.178 | 2.289 | 2.361 | 2.413 | 2.453 | 2.485 | 2.512 | 2.535 | 2.555 |
| | | | | | $\alpha = 0.1$ | | | | |
| 1 | 1.645 | 1.896 | 2.021 | 2.100 | 2.159 | 2.204 | 2.235 | 2.266 | 2.290 |
| 2 | 1.922 | 2.050 | 2.128 | 2.184 | 2.225 | 2.258 | 2.285 | 2.308 | 2.328 |
| 3 | 1.908 | 2.033 | 2.111 | 2.166 | 2.208 | 2.242 | 2.269 | 2.293 | 2.313 |
| 5 | 1.894 | 2.017 | 2.094 | 2.149 | 2.191 | 2.225 | 2.253 | 2.276 | 2.297 |
| 10 | 1.884 | 2.004 | 2.080 | 2.135 | 2.177 | 2.211 | 2.239 | 2.263 | 2.283 |
| limit as $g \to \infty$ | 1.875 | 1.992 | 2.067 | 2.122 | 2.164 | 2.197 | 2.225 | 2.249 | 2.270 |

Observations are taken in up to $K$ groups of size $g$ and at each analysis the null hypothesis is rejected if the two-sided significance level of the Student's $t$-statistic, without adjustment for repeated looks, is less than $2\{1-\Phi(Z_P(K,g,\alpha))\}$.

**Table 2.** *Constants $Z_{OBF}(K,g,\alpha)$ for O'Brien & Fleming type repeated t-tests with overall error rate $\alpha$.*

| *Number of obs. per group, g* | *Number of groups, K* 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|
| | | | | | $\alpha = 0.01$ | | | | |
| 1 | 2.576 | 2.619 | 2.646 | 2.663 | 2.679 | 2.689 | 2.695 | 2.701 | 2.706 |
| 2 | 2.584 | 2.614 | 2.635 | 2.652 | 2.660 | 2.667 | 2.673 | 2.678 | 2.682 |
| 3 | 2.584 | 2.610 | 2.629 | 2.643 | 2.654 | 2.663 | 2.669 | 2.675 | 2.680 |
| 5 | 2.583 | 2.606 | 2.622 | 2.635 | 2.646 | 2.654 | 2.661 | 2.667 | 2.672 |
| 10 | 2.582 | 2.601 | 2.616 | 2.629 | 2.639 | 2.647 | 2.655 | 2.661 | 2.666 |
| limit as $g \to \infty$ | 2.580 | 2.595 | 2.609 | 2.621 | 2.631 | 2.640 | 2.648 | 2.654 | 2.660 |
| | | | | | $\alpha = 0.05$ | | | | |
| 1 | 1.960 | 2.047 | 2.077 | 2.097 | 2.112 | 2.117 | 2.128 | 2.133 | 2.138 |
| 2 | 1.995 | 2.034 | 2.058 | 2.075 | 2.084 | 2.094 | 2.100 | 2.105 | 2.109 |
| 3 | 1.993 | 2.027 | 2.050 | 2.065 | 2.077 | 2.088 | 2.096 | 2.102 | 2.108 |
| 5 | 1.989 | 2.019 | 2.041 | 2.056 | 2.068 | 2.078 | 2.086 | 2.093 | 2.099 |
| 10 | 1.984 | 2.012 | 2.033 | 2.048 | 2.061 | 2.071 | 2.079 | 2.087 | 2.093 |
| limit as $g \to \infty$ | 1.977 | 2.004 | 2.024 | 2.040 | 2.053 | 2.063 | 2.072 | 2.080 | 2.086 |
| | | | | | $\alpha = 0.1$ | | | | |
| 1 | 1.645 | 1.759 | 1.795 | 1.812 | 1.823 | 1.835 | 1.844 | 1.848 | 1.853 |
| 2 | 1.707 | 1.746 | 1.770 | 1.786 | 1.798 | 1.807 | 1.813 | 1.819 | 1.824 |
| 3 | 1.702 | 1.736 | 1.760 | 1.777 | 1.790 | 1.800 | 1.808 | 1.816 | 1.822 |
| 5 | 1.694 | 1.727 | 1.750 | 1.767 | 1.781 | 1.791 | 1.800 | 1.807 | 1.814 |
| 10 | 1.687 | 1.719 | 1.742 | 1.759 | 1.773 | 1.784 | 1.793 | 1.801 | 1.807 |
| limit as $g \to \infty$ | 1.678 | 1.710 | 1.733 | 1.751 | 1.765 | 1.776 | 1.786 | 1.794 | 1.801 |

Observations are taken in up to $K$ groups of size $g$ and the null hypothesis is rejected at the $k$th analysis if the two-sided significance level of the Student's $t$-statistic, without adjustment for repeated looks, is less than $2\{1-\Phi(Z_{OBF}(K,g,\alpha)\sqrt{(K/k)})\}$.

**Table 3.** *Constants $C_P(p,K,\alpha)$ for Pocock type repeated $\chi^2$ tests with constant nominal significance level and overall error rate $\alpha$.*

| | $\alpha = 0.01$ | | | | | $\alpha = 0.05$ | | | | | $\alpha = 0.10$ | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $p$ | | | | | $p$ | | | | | $p$ | | | | |
| $K$ | 1 | 2 | 3 | 4 | 5 | 1 | 2 | 3 | 4 | 5 | 1 | 2 | 3 | 4 | 5 |
| 1 | 6.63 | 9.21 | 11.34 | 13.28 | 15.09 | 3.84 | 5.99 | 7.81 | 9.49 | 11.07 | 2.71 | 4.61 | 6.25 | 7.78 | 9.24 |
| 2 | 7.68 | 10.40 | 12.64 | 14.66 | 16.55 | 4.74 | 7.08 | 9.04 | 10.82 | 12.49 | 3.52 | 5.63 | 7.42 | 9.07 | 10.63 |
| 3 | 8.25 | 11.05 | 13.34 | 15.41 | 17.33 | 5.24 | 7.67 | 9.69 | 11.53 | 13.25 | 3.97 | 6.18 | 8.05 | 9.76 | 11.37 |
| 4 | 8.64 | 11.48 | 13.81 | 15.90 | 17.86 | 5.58 | 8.06 | 10.13 | 12.00 | 13.75 | 4.27 | 6.55 | 8.47 | 10.22 | 11.86 |
| 5 | 8.92 | 11.79 | 14.15 | 16.27 | 18.24 | 5.82 | 8.35 | 10.44 | 12.35 | 14.12 | 4.50 | 6.83 | 8.77 | 10.55 | 12.22 |
| 6 | 9.14 | 12.04 | 14.41 | 16.55 | 18.54 | 6.02 | 8.58 | 10.69 | 12.62 | 14.41 | 4.68 | 7.04 | 9.01 | 10.81 | 12.50 |
| 7 | 9.32 | 12.24 | 14.63 | 16.78 | 18.78 | 6.18 | 8.77 | 10.90 | 12.84 | 14.65 | 4.83 | 7.22 | 9.21 | 11.03 | 12.73 |
| 8 | 9.47 | 12.41 | 14.81 | 16.98 | 18.99 | 6.31 | 8.92 | 11.07 | 13.02 | 14.84 | 4.95 | 7.37 | 9.37 | 11.21 | 12.92 |
| 9 | 9.60 | 12.56 | 14.97 | 17.14 | 19.16 | 6.43 | 9.06 | 11.21 | 13.18 | 15.01 | 5.06 | 7.50 | 9.51 | 11.36 | 13.09 |
| 10 | 9.71 | 12.69 | 15.10 | 17.29 | 19.31 | 6.53 | 9.17 | 11.34 | 13.32 | 15.16 | 5.15 | 7.61 | 9.63 | 11.49 | 13.23 |
| 20 | 10.40 | 13.45 | 15.94 | 18.16 | 20.23 | 7.14 | 9.88 | 12.13 | 14.15 | 16.04 | 5.72 | 8.28 | 10.39 | 12.30 | 14.09 |
| 50 | 11.15 | 14.28 | 16.83 | 19.10 | 21.21 | 7.82 | 10.66 | 12.98 | 15.06 | 17.00 | 6.37 | 9.03 | 11.22 | 13.20 | 15.04 |

Observations from a $p$-variate normal distribution are taken in up to $K$ equally sized groups. At each analysis, the null hypothesis is rejected if the standard $\chi^2$ statistic exceeds $C_P(p,K,\alpha)$. All entries were found by numerical integration and are accurate to two decimal places.

**Table 4.** *Constants $C_{OBF}(p,K,\alpha)$ for O'Brien and Fleming type repeated $\chi^2$ tests with constant nominal significance level and overall error rate $\alpha$.*

| | $\alpha = 0.01$ | | | | | $\alpha = 0.05$ | | | | | $\alpha = 0.10$ | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $p$ | | | | | $p$ | | | | | $p$ | | | | |
| $K$ | 1 | 2 | 3 | 4 | 5 | 1 | 2 | 3 | 4 | 5 | 1 | 2 | 3 | 4 | 5 |
| 1 | 6.63 | 9.21 | 11.34 | 13.28 | 15.09 | 3.84 | 5.99 | 7.81 | 9.49 | 11.07 | 2.71 | 4.61 | 6.25 | 7.78 | 9.24 |
| 2 | 6.65 | 9.22 | 11.35 | 13.28 | 15.09 | 3.91 | 6.02 | 7.83 | 9.50 | 11.08 | 2.82 | 4.67 | 6.29 | 7.80 | 9.25 |
| 3 | 6.73 | 9.27 | 11.39 | 13.31 | 15.11 | 4.02 | 6.12 | 7.92 | 9.57 | 11.14 | 2.92 | 4.78 | 6.39 | 7.90 | 9.33 |
| 4 | 6.81 | 9.34 | 11.45 | 13.36 | 15.16 | 4.10 | 6.20 | 7.99 | 9.64 | 11.21 | 3.00 | 4.86 | 6.48 | 7.98 | 9.42 |
| 5 | 6.87 | 9.40 | 11.51 | 13.42 | 15.21 | 4.16 | 6.27 | 8.06 | 9.71 | 11.27 | 3.07 | 4.93 | 6.54 | 8.05 | 9.49 |
| 6 | 6.92 | 9.45 | 11.56 | 13.47 | 15.26 | 4.21 | 6.33 | 8.11 | 9.77 | 11.33 | 3.11 | 4.99 | 6.60 | 8.11 | 9.54 |
| 7 | 6.97 | 9.50 | 11.60 | 13.52 | 15.31 | 4.26 | 6.37 | 8.16 | 9.82 | 11.38 | 3.16 | 5.03 | 6.65 | 8.16 | 9.60 |
| 8 | 7.01 | 9.55 | 11.64 | 13.56 | 15.35 | 4.29 | 6.41 | 8.20 | 9.86 | 11.42 | 3.19 | 5.07 | 6.69 | 8.20 | 9.64 |
| 9 | 7.04 | 9.58 | 11.67 | 13.60 | 15.39 | 4.33 | 6.45 | 8.23 | 9.90 | 11.46 | 3.22 | 5.10 | 6.72 | 8.24 | 9.68 |
| 10 | 7.08 | 9.61 | 11.70 | 13.63 | 15.42 | 4.35 | 6.48 | 8.26 | 9.93 | 11.50 | 3.24 | 5.13 | 6.75 | 8.27 | 9.71 |
| 20 | 7.26 | 9.81 | 11.93 | 13.85 | 15.64 | 4.52 | 6.67 | 8.48 | 10.14 | 11.71 | 3.39 | 5.31 | 6.96 | 8.48 | 9.93 |
| 50 | 7.46 | 10.04 | 12.17 | 14.10 | 15.91 | 4.69 | 6.86 | 8.69 | 10.37 | 11.95 | 3.54 | 5.49 | 7.15 | 8.69 | 10.16 |

Observations from a $p$-variate normal distribution are taken in up to $K$ equally sized groups. At the $k$th analysis, the null hypothesis is rejected if the standard $\chi^2$ statistic exceeds $(K/k)\,C_{OBF}(p,K,\alpha)$. All entries were found by numerical integration and are accurate to two decimal places.