A Path Length Inequality for the Multivariate-*t* Distribution, With Applications to Multiple Comparisons

Melinda McCann and Don Edwards

This article presents a new inequality for the multivariate-t distribution, which implies a new method for multiple comparisons whose foundation rests on a recent inequality due to Naiman. The new method is promising in view of the fact that it utilizes information (estimator intercorrelations) ignored by the most widely used multiple comparison methods yet is not computationally prohibitive, requiring only the numerical evaluation of a single one-dimensional integral. In this article the validity of the new method in the normal-theoretic general linear model is established, and efficiency studies relative to the methods of Scheffé, Bonferroni, Šidák, and Hunter-Worsley are presented. The new method is shown to always improve on Scheffé's method. The new method is also shown to perform well; that is, to lead to a smaller critical point than its competitors, with low degrees of freedom. But the method is not as efficient as the Hunter-Worsley method for high degrees of freedom. In addition, the method appears to increase in relative efficiency as the number of comparisons increases relative to the rank of the correlation matrix of the estimators.

KEY WORDS: Confidence intervals; Family-wise error; Naiman's inequality; Simultaneous inference.

1. INTRODUCTION

Consider the general problem of multiple comparisons where there are k unknown parameters of interest, $\boldsymbol{\beta}' = (\beta_1, \beta_2, \ldots, \beta_k)$. Let $\mathbf{b}' = (b_1, b_2, \ldots, b_k)$ be an estimator of $\boldsymbol{\beta}'$ where the distribution of $\mathbf{b}'$ is assumed to be multivariate normal with mean $\boldsymbol{\beta}$ and covariance matrix $\sigma^2\mathbf{V}$, where $\mathbf{V}$ is known and without loss of generality full rank. Let s^2 be an estimator of σ^2 such that $\nu s^2/\sigma^2$ has a χ^2_ν distribution independent of $\mathbf{b}$, where ν is a known integer. Often it is of interest to estimate p linear combinations of the k parameters $\mathbf{c}_j'\boldsymbol{\beta}, j = 1, 2, \ldots p$, using intervals with a joint confidence level of $(1-\alpha) \times 100\%$. Here $\mathbf{c}_j' = (c_{j1}, c_{j2}, \ldots, c_{jk})$ is an arbitrary vector of constants, $j = 1, 2, \ldots, p$, chosen by the experimenter *a priori*.

These intervals are usually of the form

$$\mathbf{c}_j'\mathbf{b} \pm ds_j, \qquad j = 1, \ldots, p, \tag{1}$$

where d is a probability point to be determined and

$$s_j = s(\mathbf{c}_j'\mathbf{V}\mathbf{c}_j)^{1/2}$$

is the standard error of $\mathbf{c}_j'\mathbf{b}$. In the sequel we refer to $\mathbf{C}$, the $p \times k$ matrix of rank r whose rows are $\mathbf{c}_1', \mathbf{c}_2', \ldots, \mathbf{c}_p'$, and $\mathbf{R} = [r_{ij}]$, the correlation matrix of rank r derived from $\sigma^2\mathbf{CVC}'$, the covariance matrix of $\mathbf{b}$.

The exact critical point for (1) is a quantile of the p-dimensional multivariate-t distribution (Dunnett and Sobel 1954) with ν degrees of freedom and underlying correlation matrix $\mathbf{R}$, denoted by $t_{\alpha,\nu,\mathbf{R}}$. The probability point $t_{\alpha,\nu,\mathbf{R}}$ is currently difficult to obtain or unobtainable for many situations. Thus researchers must often settle for conservative solutions, which waste data. The main result of this article is a new inequality that is of interest in and of itself and will in some situations outperform existing conservative methods for multiple comparisons.

Section 2 states the main result, which is proved in the Appendix. Section 3 discusses the most accessible existing conservative solutions to the general multiple comparisons problem. Finally, Section 4 reports on numerical comparisons of the new method with existing conservative solutions, and Section 5 gives some conclusions.

2. A PATH LENGTH INEQUALITY

The result can be considered a corollary to Naiman's inequality (Naiman 1986).

Theorem. Let $\mathbf{T}$ have a p-dimensional multivariate-t distribution with degrees of freedom ν and underlying correlation matrix $\mathbf{R}$ of rank r. The probability

$$P(|T_j| \le d, \quad j = 1, 2, \ldots, p)$$

is bounded below by the expression

$$1 - \int_0^{1/d} \min(F_{r-2,2}[(2((dt)^{-2} - 1))/(r-2)] \times (\Lambda/\pi) + F_{r-1,1}[((dt)^{-2} - 1)/(r-1)], 1) f_T(t)\, dt, \tag{2}$$

with

$$\Lambda = \sum_{j=1}^{p-1} \cos^{-1}(|r_{j,j+1}|), \tag{3}$$

where $F_{m,n}$ is the distribution function of an F random variable with m and n degrees of freedom and f_T is the density function of a random variable T such that $rT^2 \sim F_{\nu,r}$. If d is such that the foregoing expression is at least $1-\alpha$, then the intervals (1) will be conservative simultaneous $(1-\alpha) \times 100\%$ confidence intervals.

Melinda McCann is Assistant Professor, Department of Statistics, Oklahoma State University, Stillwater, OK 74078. Don Edwards is Associate Professor, Department of Statistics, University of South Carolina, Columbia, SC 29208. The authors would like to thank the associate editor and referees for helpful comments.

Journal of the American Statistical Association
March 1996, Vol. 91, No. 433, Theory and Methods

The inequality requires only the evaluation of a one-dimensional integral and depends on $\mathbf{R}$ through its rank r and also through the "path length" constant Λ. If there are enough large correlations in $\mathbf{R}$, then Λ will be small and so will the critical constant d. Obviously, the value of Λ depends on the ordering of the indices 1, 2, ..., p. The problem of determining the optimal ordering is equivalent to the traveling salesman problem. Currently, we use a heuristic known as the nearest-neighbor algorithm (Townsend 1987), with branch lengths defined by $\cos^{-1}|r_{ij}|$.

The path length method can also be applied to one-sided confidence bounds. Here the coverage probability would be bounded below by the expression (2) with the first term inside the $\min(\quad)$ divided by 2 and with Λ given by (3) with r_{ij} replacing $|r_{ij}|$. Besides applications to simultaneous confidence bounds for arbitrary linear combinations of $\boldsymbol{\beta}$, the one-sided critical points could be very useful in ranking and selection settings, such as Hsu's (1984) MCB or Bofinger's (1985) complete ordering formulations, or in order-restricted simultaneous inference as discussed by Marcus and Peritz (1976), Marcus (1978, 1982), and Hayter (1990).

If we write $\mathbf{R} = \mathbf{A}\mathbf{A}'$ for $\mathbf{A}$ a $p \times r$ matrix of rank r with rows $\mathbf{a}_j'$, then it is interesting to note that the path length terms $\cos^{-1}(|r_{j,j+1}|)$ in (3) are the angles between consecutive $\mathbf{a}_j$ vectors, which are points on the r-dimensional sphere. Also, it is straightforward to show that the d that sets (2) equal to $1-\alpha$ is strictly increasing in Λ, and

$$\begin{aligned}\lim_{\Lambda\to\infty} \int_0^{1/d} & \min(F_{r-2,2}[(2((dt)^{-2}-1))/(r-2)] \\ & \times (\Lambda/\pi) + F_{r-1,1}[((dt)^{-2}-1)/(r-1)], 1) f_T(t)\, dt \\ = \int_0^{1/d} & f_T(t)\, dt.\end{aligned}$$

From this, as $\Lambda \to \infty, d \to (rF_{\alpha,r,\nu})^{1/2}$, which is Scheffé's critical point for the intervals given by (1). Thus the new method always outperforms Scheffé's method. Finally, as $\Lambda \to 0, d \to t_{\alpha/2,\nu}$, the upper-alpha point of the univariate-t distribution and the exact point for a single interval. This is intuitively pleasing because $\Lambda \to 0$ implies a path of correlations in $\mathbf{R}$ approaching 1, in which case the p-dimensional distribution is for all practical purposes one-dimensional.

3. EXISTING CONSERVATIVE SOLUTIONS

The Bonferroni, Scheffé, and Šidák (1967) methods are all widely accessible conservative methods that provide upper bounds for $t_{\alpha,\nu,\mathbf{R}}$ for the general case. Bonferroni's method makes use of the inequality

$$t_{\alpha,\nu,\mathbf{R}} \le t_{\alpha/2p,\nu}.$$

Scheffé's method makes use of the inequality

$$t_{\alpha,\nu,\mathbf{R}} \le (rF_\alpha(r,\nu))^{1/2}.$$

Šidák's method makes use of the inequality

$$t_{\alpha,\nu,\mathbf{R}} \le t_{\alpha,\nu,\mathbf{I}}.$$

Tables of $t_{\alpha,\nu,\mathbf{I}}$ are widely available (for example, in Hahn and Hendrickson 1971). To apply any of the aforementioned conservative methods, one uses the right side of the appropriate inequality in place of d in (1).

A more computationally intensive conservative method for the general case is an improved Bonferroni method (Hunter 1976; Worsley 1982). The method of Hunter and Worsley uses the inequality for the probability of a p-event union,

$$P\left(\bigcup_{i=1}^{p} A_i\right) \le \sum_{i=1}^{p} P(A_i) - \max_{\tau\in T} \sum_{\tau} P(A_i A_j), \qquad (4)$$

where T is the set of all spanning trees of the indices $\{1, 2, \ldots, p\}$. In our setting A_i is the event that interval i does not contain its target $\mathbf{c}_i'\boldsymbol{\beta}$. For any given critical point d, the family-wise error probability $P(\bigcup A_i)$ can be bounded by the right side using the univariate and bivariate-t distributions, requiring $p-1$ one-dimensional numerical integrals for the bivariate-t cases. In this case the maximal spanning tree that uses $P(A_iA_j)$ as the lengths of its branches depends only on the relative magnitudes of the $P(A_iA_j)$, which can be determined from the correlation matrix $\mathbf{R}$. Specifically, the maximal spanning tree τ for (4) is also the maximal spanning tree for $\sum_\tau |r_{ij}|$. To apply this method to the particular problem of interest, one would find the maximal spanning tree once using the algorithm of Kruskal (1956) and put the upper bound (4) into an iterative method to obtain the point d where the bound equals α.

Other methods that are not widely accessible and will not be further discussed here include large-scale numerical integration programs such as those discussed by Schervish (1984) and Wang and Kennedy (1992). We also do not discuss methods that utilize, at least in part, Monte Carlo to approximate a critical point; for example, those of Foutz (1981), Edwards and Berry (1987), and Naiman and Wynn (1992). Simulation-based methods provide a useful approximate working solution to the general multiple comparisons problem, but we feel that the statistical community should still strive to find hard numerical solutions to these problems.

4. RELATIVE EFFICIENCIES

For our study, we considered the relative efficiency of one method to another as the ratio of the methods' squared critical points; this ratio is approximately the ratio of sample sizes needed to achieve intervals of equal length. Critical points were computed using FORTRAN with IMSL (1987) routines. More specifically, the integrals of interest were calculated using an IMSL (1987) numerical integration routine with relative and absolute errors of 10^{-4}. The critical points were then calculated using the secant method with an error of 10^{-4}. A systematic investigation of the effects of various factors on relative efficiency was conducted. The factors were as follows:

Factor 1: V matrix. The levels chosen were

a. AR(1) process with parameter φ; that is, $\mathbf{V} = [v_{ij}]$ for $v_{ij} = \varphi^{|i-j|}$. $\varphi = 0$, .5, .8, .9, .95, and .99.

b. Random **V** generated by a Wishart distribution with high (.8) and low (.3) average correlations, and high (5 df) and low (60 df) variability in the correlations.

Factor 2: C matrix. The levels chosen were **C** matrices to estimate

a. All pairwise comparisons of $\beta_1, \ldots, \beta_k$ (MCA). Thus $p = k(k-1)/2$.
b. All comparisons of one parameter to the other $k-1$ (MCC). Thus $p = k-1$.
c. All of the parameters individually (SMM, the studentized maximum modulus), $p = k$.

Factor 3: Number of base parameters, *k*. The levels chosen were 4 and 10.

Factor 4: α. The levels chosen were .05 and .10.

Factor 5: Error degrees of freedom. The levels chosen were 5, 15, and 60. Figures 1–3 graph the relative efficiencies of the path length inequality method to the Šidák and Hunter-Worsley methods as functions of φ for the AR1(φ) **V** matrices, for $\alpha = .05$. The results for $\alpha = .10$ are very similar, with a slight increase in all relative efficiencies.

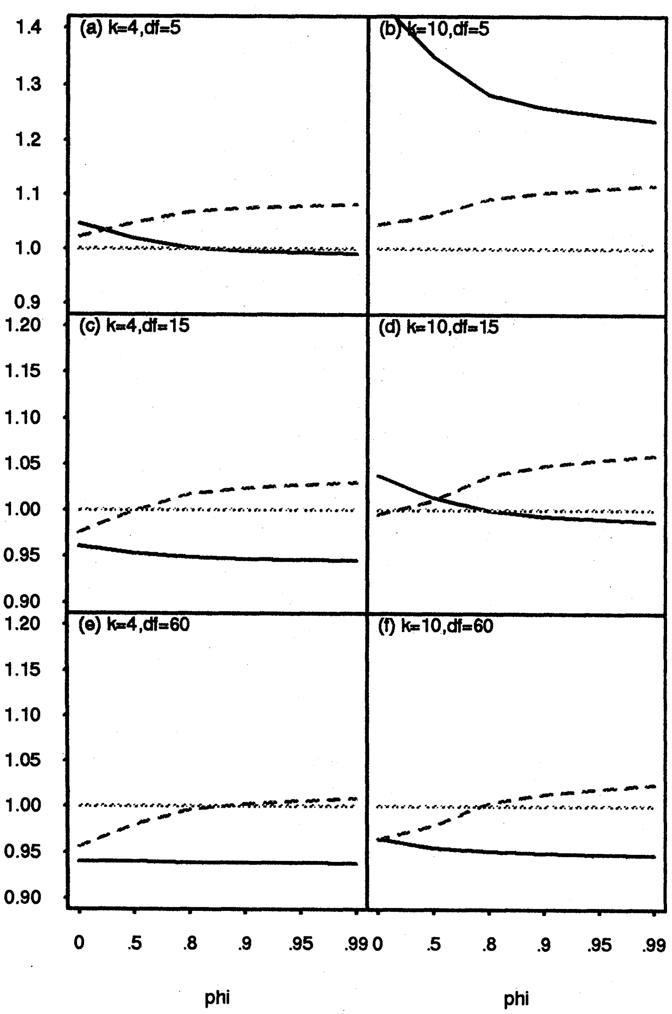

Figure 1. Relative Efficiency of the Path Length Inequality Method to the Hunter-Worsley (——) and Šidák (– – –) Methods for AR1(phi) Estimator Covariance Structure. All pairwise comparisons (MCA), k treatments, $\alpha = .05$. An additional dashed line is drawn through 1.

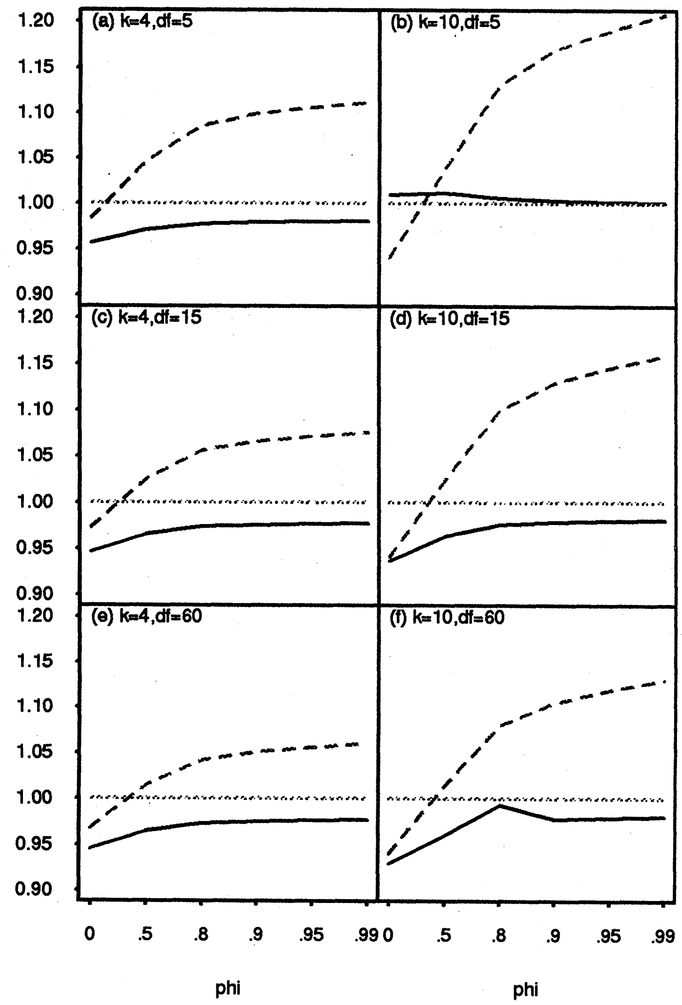

Figure 2. Relative Efficiency of the Path Length Inequality Method to the Hunter-Worsley (——) and Šidák (– – –) Methods for AR1(phi) Estimator Covariance Structure. All-to-one comparisons (MCC), k treatments, $\alpha = .05$. An additional dashed line is drawn through 1.

Relative efficiencies to the Scheffé and Bonferroni methods are not presented, as the new method always outperforms the Scheffé and the Hunter-Worsley always outperforms the Bonferroni. In addition, because the Scheffé and Bonferroni critical constants do not change as φ changes, the shape of the graphs of the relative efficiency to these methods would parallel that of Šidák's.

The graphs show that the new method's relative performance improves as the number of comparisons relative to the rank of **R** increases, implying that the method would perform relatively well for large problems with many comparisons of interest. The new method does poorly versus the Hunter-Worsley method for large degrees of freedom (60) but performs well compared to the Hunter-Worsley in most cases for low degrees of freedom (5). At 15 df, the method performs well compared to the Hunter-Worsley for the case with the largest p/r ratio (MCA, $k = 10$) but poorly versus the Hunter-Worsley for the other cases. When the new method is compared to the Šidák method, it performs better as φ increases, outperforming the Šidák method for all cases of interest when $\varphi \geq .8$.

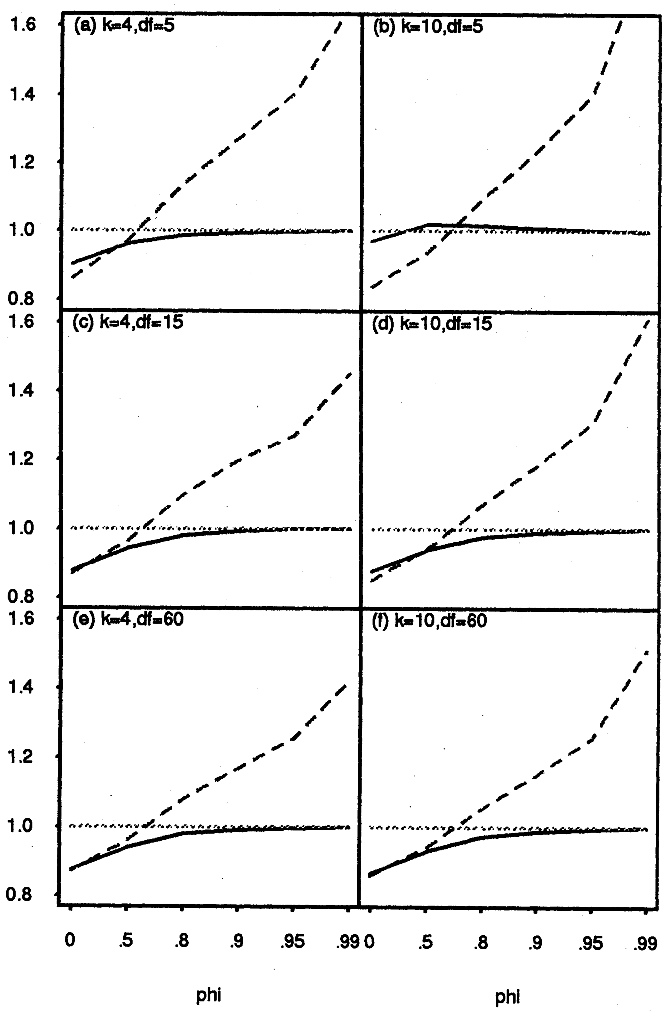

Figure 3. Relative Efficiency of the Path Length Inequality Method to the Hunter-Worsley (——) and Šidák (– – –) Methods for AR1(phi) Estimator Covariance Structure. Intervals for each treatment mean (SMM), k treatments, $\alpha = .05$. An additional dashed line is drawn through 1.

The pattern of increasing relative efficiency with φ versus the Šidák method is to be expected; the Šidák method makes no use of the information in $\mathbf{R}$, assuming the worst case, $\mathbf{R} = \mathbf{I}$. When φ is far from zero, the general level of estimator intercorrelation is also far from what it would be under $\mathbf{R} = \mathbf{I}$; hence we should not expect good performance from the Šidák method.

The patterns versus the Hunter-Worsley procedure are more difficult to rationalize. As pointed out by Naiman and Wynn (1992), for any fixed critical point d, the exact familywise error probability of the intervals is given by the integral (2) with the $\min(\quad)$ term replaced by the surface area of a certain set on the r-dimensional sphere. The set is the union of p equal-sized disks centered at the $\mathbf{a}_j, j = 1, \ldots, p$, whose radius is inversely related to the variable of integration t. The Hunter-Worsley method and the method of this article could be regarded as substituting easily computed upper bounds for the surface area of this disk-union set into the integral. Each method first orders the disks such that the length of the path connecting their centers is minimized. The method of this article then bounds the surface area by the minimum of 1 and the surface area of a tube about this path whose diameter is just large enough to enclose each disk. Intuition suggests that this bound will not work well when large weight (in the second term f_T) is given to values of t for which there are few intersections of these disks. The Hunter-Worsley method bounds the area by the sum of the individual disk areas, minus the areas of the $p-1$ intersections of consecutive disks in the ordering. When t is small, the disks are large, and there are many important disk intersections other than those removed by the subtraction. This may explain the poor performance of the Hunter-Worsley method (and the Bonferroni method) when the error degrees of freedom are small, in which case the second term of the integrand of (2) gives large weight to small values of t (large disks). Also, when there are relatively many disks, there are more possibilities of high-order intersections of the disks, hence the increasing relative efficiency of the new method to Hunter-Worsley as p/r increases.

For the random $\mathbf{V}$ matrices, similar patterns were noted with respect to degrees of freedom, alpha, and the number of comparisons relative to the rank of $\mathbf{R}$. There were no clear patterns in relative efficiencies for a change in the average correlations or in variability of the correlations, perhaps due to the randomness of the matrices generated. In addition, the relative efficiencies for the random $\mathbf{V}$ matrices were slightly lower than those for the AR(1) $\mathbf{V}$ matrices.

In a few cases of $\mathbf{V}$ and $\mathbf{C}$, exact critical points are available. For example, for MCA or MCC when $\mathbf{V}$ is the identity, Tukey's method and Dunnett's (1964) method, respectively, are exact. Though the inequality of this article was not meant to be used in these cases, it may be useful to examine relative efficiency to the exact methods. Tables 1 and 2 provide these relative efficiencies for some of the cases examined in this section. Apparently, for these cases, the path length inequality can be improved on substantially by the exact critical points.

Table 1. Relative Efficiencies of the New Method to Tukey's (Where Exact)

	$k = 4$				$k = 10$		
α	$\nu = 5$	$\nu = 15$	$\nu = 60$	α	$\nu = 5$	$\nu = 15$	$\nu = 60$
.05	.90	.91	.91	.05	.88	.89	.90
.10	.90	.90	.90	.10	.88	.88	.89

Table 2. Relative Efficiencies of the New Method to Dunnett's (Where Exact)

	$k = 4$				$k = 10$		
α	$\nu = 5$	$\nu = 15$	$\nu = 60$	α	$\nu = 5$	$\nu = 15$	$\nu = 60$
.05	.92	.93	.93	.05	.83	.86	.88

5. CONCLUSIONS

A general inequality for the multivariate-t distribution has been established. It implies a general multiple comparisons method. The new method seems to outperform its competitors for very low values of error degrees of freedom ν. As the number of comparisons p increases, the path length method also seems to perform better than its competitors. For cases where the number of comparisons is large, the proposed method also has an advantage over the Hunter-Worsley method in that computer time stays relatively small, requiring only a single one-dimensional numerical integral for each iteration. In comparison, the Hunter-Worsley method, which requires evaluation of $p-1$ one-dimensional integrals, becomes more computationally intensive as p increases. The proposed method seems to improve over all the competitors in certain situations, and in those situations it will be less data-wasteful.

APPENDIX: NAIMAN'S INEQUALITY AND PROOF OF THE MAIN RESULT

The path length inequality makes use of an inequality derived by Naiman (1986); his notation is modified slightly to fit our setting. He considered a regression model for y in terms of known functions $f_1, f_2, \ldots, f_r$ of a single regressor x:

$$y_i = \sum_{j=1}^{r} \omega_j f_j(x_i) + \varepsilon_i, \qquad i = 1, 2, \ldots, n, \qquad x \in I,$$

where I is a fixed, closed interval of the real line. Here the ω_j are unknown constants, and the ε_i are independent $N(0, \sigma^2)$ random variables. Let $\mathbf{f}$ be the vector $(f_1(x), f_2(x), \ldots, f_r(x))'$. $\mathbf{f}$ is assumed to be a continuous function from I to the k-dimensional reals, which is bounded away from the origin and piecewise differentiable, with

$$\int_I \|\partial \mathbf{f}(x)/\partial x\|^2 \, dx$$

finite. Naiman's goal was to construct simultaneous confidence bands for $E(y) = \boldsymbol{\omega}'\mathbf{f}(x)$ for all $x \in I$. Let $\mathbf{w}$ and s^2 be the usual unbiased estimators of $\boldsymbol{\omega}$ and σ^2, where $\mathbf{w}$ is distributed as multivariate normal with mean $\boldsymbol{\omega}$ and covariance matrix $\sigma^2\mathbf{B}$ of full-rank r.

Naiman's solution utilizes a representation due to Uusipaikka (1984) for the coverage probability of the bands

$$\mathbf{f}(x)'\mathbf{w} \pm ds(\mathbf{f}(x)'\mathbf{B}\mathbf{f}(x))^{1/2}, \qquad x \in I \qquad \text{(A.1)}$$

as weighted averages of the volumes of certain "tubes" in S^{r-1}, the unit sphere in r-dimensional Euclidean space. A tube surrounds a "path" in S^{r-1}, which is a piecewise differentiable function γ mapping I into S^{r-1} such that

$$\Lambda(\gamma) = \int_I \|\partial\gamma(x)/\partial x\| \, dx$$

is finite. $\Lambda(\gamma)$, henceforth abbreviated as Λ, is called the length of γ. For the regression confidence band (A.1), let

$$\gamma(x) = \|\mathbf{P}\mathbf{f}(x)\|^{-1}\mathbf{P}\mathbf{f}(x), \qquad x \in I, \qquad \text{(A.2)}$$

where $\mathbf{P}$ is a square matrix such that $\mathbf{P}'\mathbf{P} = \mathbf{B}$. If the assumptions on $\mathbf{f}$ hold, then it can be seen that γ and $-\gamma$ are paths in S^{r-1}. Naiman's (1986) theorem 4.2, which we call Naiman's inequality, states that the coverage probability of the bands (A.1) is bounded above by expression (2) from Section 2 of this article.

The main result of Section 2 is established as follows: Let $\mathbf{T} = (T_1, T_2, \ldots, T_p)'$, where

$$T_j = (\mathbf{c}_j'\mathbf{b} - \mathbf{c}_j'\beta)/(s(\mathbf{c}_j'\mathbf{V}\mathbf{c}_j)^{1/2}), \qquad j = 1, 2, \ldots, p.$$

Then $\mathbf{T}$ has a multivariate-t distribution with degrees of freedom ν and underlying correlation matrix $\mathbf{R}$ of rank r. Here

$$r_{ij} = (\mathbf{c}_i'\mathbf{V}\mathbf{c}_j)/((\mathbf{c}_i'\mathbf{V}\mathbf{c}_i)(\mathbf{c}_j'\mathbf{V}\mathbf{c}_j))^{1/2}, \qquad 1 \le i \ne j \le p.$$

Let $\mathbf{R} = \mathbf{A}\mathbf{A}'$, where $\mathbf{A}$ is $p \times r$ of rank r. Denote the rows of $\mathbf{A}$ by $\mathbf{a}_j', j = 1, 2, \ldots, p$; note that $\mathbf{a}_j'\mathbf{a}_j = 1$ for each j and $\mathbf{a}_i'\mathbf{a}_j = r_{ij}$ for $1 \le i \ne j \le p$. Naiman's inequality can be applied to bound the multivariate-t probability

$$P(|T_j| \le d, \quad j = 1, 2, \ldots, p) \qquad \text{(A.3)}$$

by defining a function $\mathbf{f}$ mapping from $I = \{x\colon 1 \le x \le p\}$ to r-dimensional space such that $\mathbf{f}(1) = \mathbf{a}_1, \mathbf{f}(2) = \mathbf{a}_2, \ldots, \mathbf{f}(p) = \mathbf{a}_p$, and defining $\mathbf{f}(x)$ for noninteger x to connect these points; that is,

$$\mathbf{f}(x) = (x-j)(\mathbf{a}_{j+1} - \mathbf{a}_j) + \mathbf{a}_j, \qquad j \le x \le j+1, \qquad j = 1, 2, \ldots, p-1. \qquad \text{(A.4)}$$

It is straightforward to show that $\mathbf{f}$ satisfies Naiman's regularity conditions provided that no $\mathbf{a}_j = \mathbf{0}$ and no $\mathbf{a}_j$ is a multiple of $\mathbf{a}_{j+1}$. Naiman's $\mathbf{B}$ can be taken to be $\mathbf{I}_r$, the identity matrix of order r. Let $\mathbf{Z} = (Z_1, Z_2, \ldots, Z_p)$ be a random vector distributed as $N_p(\mathbf{0}, \mathbf{R})$ independent of s. Now the probability (A.3) can be written as

$$\begin{aligned} & P(|T_j| \le d, j = 1, 2, \ldots, p) \\ &= P(|Z_j| \le ds, j = 1, 2, \ldots, p) \\ &= P(|\mathbf{a}_j'(\mathbf{w} - \boldsymbol{\omega})| \le ds, j = 1, 2, \ldots, p) \\ &\le P(|\mathbf{f}(x)'(\mathbf{w} - \boldsymbol{\omega})| \le ds(\mathbf{f}(x)'\mathbf{f}(x))^{1/2} \ \forall\, x \in I), \qquad \text{(A.5)} \end{aligned}$$

the second equality following from the fact that $\mathbf{Z}$ distributed as $N_p(\mathbf{0}, \mathbf{R})$ is equal in distribution to $\mathbf{A}(\mathbf{w} - \boldsymbol{\omega})$ when $\mathbf{B} = \mathbf{I}_r$. The last expression in (A.5) is the coverage probability of the bands (A.1) when $\mathbf{B} = \mathbf{I}_r$. The main result follows using Naiman's inequality. It remains only to show that Λ for the choice of $\mathbf{f}$ in (A.4) is given by Equation (3).

Because $\mathbf{B}$ is the identity of order r, so is the matrix $\mathbf{P}$ in (A.2). For each fixed $j, 1 \le j \le p-1$, and $x \in [j, j+1]$, (A.2) then becomes

$$\begin{aligned} \gamma(x) &= \|\mathbf{f}(x)\|^{-1}\mathbf{f}(x) \\ &= (\mathbf{a}_j + (x-j)(\mathbf{a}_{j+1} - \mathbf{a}_j))/(2(x-j)^2(1 - r_{j,j+1}) \\ &\quad - 2(x-j)(1 - r_{j,j+1}) + 1)^{1/2}, \end{aligned}$$

using $\mathbf{a}_j'\mathbf{a}_j = 1$ and $\mathbf{a}_i'\mathbf{a}_j = r_{ij}$. The derivative $\partial\gamma(x)/\partial x$ can be seen to be

$$[2(x-j)^2(1 - r_{j,j+1}) - 2(x-j)(1 - r_{j,j+1}) + 1]^{-3/2} \; [-(x-j)(1 - r_{j,j+1})(\mathbf{a}_{j+1} + \mathbf{a}_j) - r_{j,j+1}\mathbf{a}_j + \mathbf{a}_{j+1}],$$

and its norm, after some algebra, reduces to

$$(1 - r_{j,j+1}^2)^{1/2}/\{2(x-j)^2(1 - r_{j,j+1}) - 2(x-j)(1 - r_{j,j+1}) + 1\}.$$

Integrating this on the interval $[j, j+1]$ gives

$$\int_j^{j+1} \|\partial\gamma(x)/\partial x\| = 2\tan^{-1}\{[(1 - r_{j,j+1})/(1 + r_{j,j+1})]^{1/2}\}.$$

This expression reduces to $\cos^{-1}(r_{j,j+1})$ using the trigonometric identity

$$\tan^2\left(\frac{\phi}{2}\right) = (1 - \cos(\phi))/(1 + \cos(\phi)).$$

To see that r_{ij} can be replaced by $|r_{ij}|$, consider the following: If $r_{12} < 0$, then replace $\mathbf{a}_2$ with $-\mathbf{a}_2$. If after this change $r_{23} < 0$, then replace $\mathbf{a}_3$ with $-\mathbf{a}_3$. Continue this procedure until $r_{(p-1)p}$ has been considered. Now all r_{ij} are positive, and we are bounding $\pm\, \mathbf{a}_j'(\mathbf{w} - \boldsymbol{\omega}), j = 1, 2, \ldots, p$. Because we were originally bounding these quantities in absolute value, multiplying some of these by -1 does not change the problem. Finally, the total path length is found by summing these segment-integral terms over $j = 1, 2, \ldots, p-1$.

[Received August 1994. Revised June 1995.]

REFERENCES

Bofinger, E. (1985), "Multiple Comparisons and Type III Errors," *Journal of the American Statistical Association*, 80, 433–437.

Dunnett, C. W. (1964), "New Tables for Multiple Comparisons With a Control," *Biometrics*, 20, 482–491.

Dunnett, C. W., and Sobel, M. (1954), "A Bivariate Generalization of Student's t Distribution, With Tables for Certain Special Cases," *Biometrika*, 31, 153–169.

Edwards, D., and Berry, J. J. (1987), "The Efficiency of Simulation-Based Multiple Comparisons," *Biometrics*, 43, 913–928.

Foutz, R. V. (1981), "Simultaneous Tests for Finite Families of Hypotheses," *Communications in Statistics, Series A—Theory and Methods*, 11, 1839–1853.

Hahn, G. J., and Hendrickson, R. W. (1971), "A Table of Percentage Points of the Distribution of the Largest Absolute Value of k Student t Variates and Its Applications," *Biometrika*, 58, 323–332.

Hayter, A. J. (1990), "A One-Sided Studentized Range Test for Testing Against a Simple Ordered Alternative," *Journal of the American Statistical Association*, 85, 778–785.

Hsu, J. C. (1984), "Constrained Two-Sided Simultaneous Confidence Intervals for Multiple Comparisons With the Best," *The Annals of Statistics*, 12, 1136–1144.

Hunter, D. (1976), "An Upper Bound for the Probability of a Union," *Journal of Applied Probability*, 13, 597–603.

International Mathematical and Statistical Libraries, Inc. (1987), *IMSL User's Manual: Math/Library: Fortran Subroutines for Mathematical Applications*, Houston: Author.

——— (1987), *IMSL User's Manual: Stat/Library: Fortran Subroutines for Statistical Applications*, Houston: Author.

Kruskal, J. B. (1956), "On the Shortest Spanning Subtree of a Graph and the Traveling Salesman Problem," *Proceedings of the American Mathematical Society*, 7, 48–50.

Marcus, R. (1978), "Further Results on Simultaneous Confidence Bounds in Normal Models With Restricted Alternatives," *Communications in Statistics, Part A—Theory and Methods*, A7, 573–590.

——— (1982), "Some Results on Simultaneous Confidence Intervals for Monotone Contrasts in One-Way ANOVA Model," *Communications in Statistics, Part A—Theory and Methods*, 11, 615–622.

Marcus, R., and Peritz, E. (1976), "Some Simultaneous Confidence Bounds in Normal Model With Restricted Alternatives," *Journal of the Royal Statistical Society*, Ser. B, 38, 157–165.

Naiman, D. Q. (1986), "Conservative Confidence Bands in Curvilinear Regression," *The Annals of Statistics*, 14, 896–906.

Naiman, D. Q., and Wynn, H. P. (1992), "Inclusion-Exclusion Bonferroni Identities and Inequalities for Discrete Tube-Like Problems via Euler Characteristics," *The Annals of Statistics*, 20, 43–76.

Scheffé, H. (1953), "A Method for Judging All Contrasts in the Analysis of Variance," *Biometrika*, 40, 87–104.

Schervish, M. J. (1984), "Multivariate Normal Probabilities With Error Bound, Algorithm AS 195," *Journal of the Royal Statistical Society, Applied Statistics*, 33, 81–94.

Šidák, Z. (1967), "Rectangular Confidence Regions for the Means of Multivariate Normal Distributions," *Journal of the American Statistical Association*, 62, 626–633.

Townsend, M. (1987), *Discrete Mathematics: Applied Combinatorics and Graph Theory*, Menlo Park, CA: Benjamin/Cummings.

Uusipaikka, E. (1984), "Exact Confidence Bands on Certain Restricted Sets in the General Linear Model," unpublished manuscript.

Wang, M., and Kennedy, W. J. (1992), "A Numerical Method for Accurately Approximating Multivariate Normal Probabilities," *Computational Statistics and Data Analysis*, 13, 197–210.

Worsley, K. J. (1982), "An Improved Bonferroni Inequality and Applications," *Biometrika*, 69, 297–302.