The Annals of Statistics
1975, Vol. 3, No. 2, 429-444

NORMAL-THEORY APPROXIMATIONS TO TESTS FOR LINEAR HYPOTHESES

BY D. R. JENSEN[1] AND L. S. MAYER[2]

Virginia Polytechnic Institute and State University

Let $\{\mathbf{y}_1, \cdots, \mathbf{y}_N\}$ be N independent repetitions of an experiment for which $\mathscr{E}\mathbf{y} = \mathbf{X}\boldsymbol{\beta}$ and $\mathscr{E}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})' = \sigma^2\mathbf{I}$, where $\mathbf{X}$ is nonstochastic of full rank and $\boldsymbol{\beta}' = [\boldsymbol{\beta}_1', \cdots, \boldsymbol{\beta}_r']$ and σ^2 are unknown parameters. We investigate some large-sample properties of $N^{\frac{1}{2}}(\hat{\boldsymbol{\beta}}_N - \boldsymbol{\beta})$, of $V_N = N(\hat{\boldsymbol{\beta}}_N - \boldsymbol{\beta}_0)'\mathbf{T}^{-1}(\hat{\boldsymbol{\beta}}_N - \boldsymbol{\beta}_0)/\sigma^2$, and of

$$V_{Nj} = N(\hat{\boldsymbol{\beta}}_{Nj} - \boldsymbol{\beta}_{j0})'\mathbf{T}_{jj}^{-1}(\hat{\boldsymbol{\beta}}_{Nj} - \boldsymbol{\beta}_{j0})/\sigma^2\,; \qquad 1 \leqq j \leqq r$$

where $\hat{\boldsymbol{\beta}}_N = N^{-1}\mathbf{T}\mathbf{X}'(\mathbf{y}_1 + \cdots + \mathbf{y}_N)$ and $\mathbf{T} = [\mathbf{T}_{ij}] = (\mathbf{X}'\mathbf{X})^{-1}$. Our conclusions thus apply to problems of inference regarding $\boldsymbol{\beta}$ and $\{\boldsymbol{\beta}_1, \cdots, \boldsymbol{\beta}_r\}$. Given certain higher-order moments of $\mathbf{y}$, we provide bounds of the Berry-Esséen type on the rates of convergence to their limiting forms of the distributions of $N^{\frac{1}{2}}(\hat{\boldsymbol{\beta}}_N - \boldsymbol{\beta})$, of V_N, of $\{V_{N1}, \cdots, V_{Nr}\}$, and of the variance ratios $U_N = \sigma^2 V_N/\hat{\sigma}_N^2$ and $\{U_{N1}, \cdots, U_{Nr}\}$, where $U_{Nj} = \sigma^2 V_{Nj}/\hat{\sigma}_N^2$ and $\hat{\sigma}_N^2$ is the sample variance.

1. Introduction. The following aspects of linear models are standard. Let $\mathbf{X}(m \times p)$ be a nonstochastic matrix of rank $p \leqq m$, σ^2 an unknown scalar and $\boldsymbol{\beta}(p \times 1)$ a vector of unknown parameters, and $\mathbf{y}(m \times 1)$ a vector-valued random element such that $\mathscr{E}\mathbf{y} = \mathbf{X}\boldsymbol{\beta}$ and $\mathscr{E}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})' = \sigma^2\mathbf{I}_m$. Define $\mathbf{T} = [t_{ij}] = (\mathbf{X}'\mathbf{X})^{-1}$. Least-squares estimators for $\boldsymbol{\beta}$ are $\tilde{\boldsymbol{\beta}} = \mathbf{T}\mathbf{X}'\mathbf{y}$, in which case $\mathscr{E}\tilde{\boldsymbol{\beta}} = \boldsymbol{\beta}$ and $\mathscr{E}(\tilde{\boldsymbol{\beta}} - \boldsymbol{\beta})(\tilde{\boldsymbol{\beta}} - \boldsymbol{\beta})' = \sigma^2\mathbf{T}$, and the corresponding unbiased estimator for σ^2 is $\tilde{\sigma}^2 = (\mathbf{y} - \mathbf{X}\tilde{\boldsymbol{\beta}})'(\mathbf{y} - \mathbf{X}\tilde{\boldsymbol{\beta}})/(m - p)$. The variance ratio test for the hypothesis $H\colon \boldsymbol{\beta} = \boldsymbol{\beta}_0$ utilizes the statistic

$$U = (\tilde{\boldsymbol{\beta}} - \boldsymbol{\beta}_0)'\mathbf{T}^{-1}(\tilde{\boldsymbol{\beta}} - \boldsymbol{\beta}_0)/p\tilde{\sigma}^2 \tag{1.1}$$

together with a rejection region in the upper tail of its null distribution. If in addition the parameters $\boldsymbol{\beta}$ are partitioned into the natural subsets $\boldsymbol{\beta}' = [\boldsymbol{\beta}_1', \cdots, \boldsymbol{\beta}_r']$ about which inferences are to be made separately, where $\boldsymbol{\beta}_j$ is of order $(p_j \times 1)$ and $p_1 + \cdots + p_r = p$, then the variance ratio tests for the r hypotheses H_j: $\boldsymbol{\beta}_j = \boldsymbol{\beta}_{j0}$, $1 \leqq j \leqq r$, utilize the statistics

$$U_j = (\tilde{\boldsymbol{\beta}}_j - \boldsymbol{\beta}_{j0})'\mathbf{T}_{jj}^{-1}(\tilde{\boldsymbol{\beta}}_j - \boldsymbol{\beta}_{j0})/p_j\tilde{\sigma}^2\,, \qquad 1 \leqq j \leqq r \tag{1.2}$$

Received January 1972; revised April 1974.

[1] Supported in part by the National Institute of General Medical Sciences under Research Career Development Award 5-K03-GM37209-05, and in part by Contract N00014-72-A-0136-0003 from the Office of Naval Research. Reproduction in whole or in part is permitted for any purpose of the United States Government.

[2] Work partially supported by the National Institute of Mental Health grant No. NIH-1R03-MH21454-01.

AMS 1970 *subject classifications.* Primary 62J05; Secondary 60F05.

Key words and phrases. Linear models; joint distribution of variance ratios; central limit theory; Berry-Esséen bounds.

together with upper-tail rejection regions, where the arrays $\mathbf{X}$, $\mathbf{T} = [\mathbf{T}_{ij}]$, $\tilde{\boldsymbol{\beta}}$ and $\boldsymbol{\beta}_0$ are partitioned conformably with $\boldsymbol{\beta}$.

For later reference define

$$V_0 = \mathbf{y}'(\mathbf{I}_m - \mathbf{X}\mathbf{T}\mathbf{X}')\mathbf{y}/\sigma^2 \tag{1.3}$$

$$V = (\tilde{\boldsymbol{\beta}} - \boldsymbol{\beta}_0)'\mathbf{T}^{-1}(\tilde{\boldsymbol{\beta}} - \boldsymbol{\beta}_0)/\sigma^2 \tag{1.4}$$

and

$$V_j = (\tilde{\boldsymbol{\beta}}_j - \boldsymbol{\beta}_{j0})'\mathbf{T}_{jj}^{-1}(\tilde{\boldsymbol{\beta}}_j - \boldsymbol{\beta}_{j0})/\sigma^2\,, \qquad 1 \leqq j \leqq r\,. \tag{1.5}$$

When $\mathbf{y}$ is Gaussian the estimators $\tilde{\boldsymbol{\beta}}$ are Gaussian; V and V_0 are independent chi-squared (χ^2) variates having p and $(m - p)$ degrees of freedom; the variance ratio test is equivalent to the likelihood ratio test; and the statistic U has the Snedecor-Fisher F distribution with p and $(m - p)$ degrees of freedom. Moreover, the joint null distributions of $\{V_1, \cdots, V_r\}$ and $\{U_1, \cdots, U_r\}$ are known multivariate χ^2 and F distributions [7]; these depend on the off-diagonal blocks of $\mathbf{T}$ in such a way that $\{V_1, \cdots, V_r\}$ are mutually independent when $\mathbf{X}'\mathbf{X} = \operatorname{Diag}(\mathbf{X}_1'\mathbf{X}_1, \cdots, \mathbf{X}_r'\mathbf{X}_r)$, a block-diagonal matrix.

Under suitable moment conditions it is known that normal-theory approximate procedures can be justified in large samples for non-normal data by virtue of central limit theory, the approximations tending to improve with increasing sample size. In the present study such notions are made precise. Using a construction which enables us to treat at once the central and noncentral cases, we study rates of convergence of the distributions of $\tilde{\boldsymbol{\beta}}$, V, U, $\{V_1, \cdots, V_r\}$ and $\{U_1, \cdots, U_r\}$ to their limiting normal-theory forms. Following some preliminary developments in Section 2, our main findings are given in Section 3 for $\tilde{\boldsymbol{\beta}}$, V and $\{V_1, \cdots, V_r\}$, and in Section 4 for U and $\{U_1, \cdots, U_r\}$.

2. Preliminaries. Bold-faced characters represent arrays with elements from the field $\mathbb{R}^1$ of real numbers, lower case for vectors and upper case for matrices; $\mathbb{R}_+^m$ is the positive orthant of the m-dimensional Euclidean space $\mathbb{R}^m$. Let $\mathscr{A}$, $\mathscr{A}_X$ and $\mathscr{A}_Y$ be separable metric spaces and $\{X, X_N; N = 1, 2, \cdots\}$ a stochastic sequence in $\mathscr{A}_X$ having the sequence $\{P(\cdot), P_N(\cdot); N = 1, 2, \cdots\}$ of probability measures. If X_N converges in probability to X we write $X_N \to_P X$; if X_N converges in distribution to X so that $\lim_{N\to\infty} P_N(\cdot) = P(\cdot)$ at every continuity set of the latter, we write $\mathscr{L}_\infty(X_N) = \mathscr{L}(X)$. Now combining Theorems 4.4 and 5.1 of Billingsley [5], we have

LEMMA 1. *Let* $\{(X, c), (X_N, Y_N); N = 1, 2, \cdots\}$ *be a stochastic sequence in* $\mathscr{A}_X \times \mathscr{A}_Y$, *and* $g : \mathscr{A}_X \times \mathscr{A}_Y \to \mathscr{A}$ *a continuous mapping, such that*

(i) $\mathscr{L}_\infty(X_N) = \mathscr{L}(X)$,

(ii) $Y_N \to_P c$, *a point in* $\mathscr{A}_Y$.

Then $\mathscr{L}_\infty[g(X_N, Y_N)] = \mathscr{L}[g(X, c)]$.

We review some known bounds of the Berry–Esséen type on rates of convergence for sequences of independent identically distributed (i.i.d.) random vectors in $\mathbb{R}^m$. Included are bounds due to Esséen [6] on probabilities assigned to the ball of radius t; bounds on multidimensional distribution functions; and bounds due to various authors on probabilities assigned to more general convex sets. Let $\Psi_\nu(\cdot)$ be the cdf of the central χ^2 distribution having ν degrees of freedom, and denote by $\theta_{hj} = \mathscr{E}|x_j|^h$ the hth absolute moment of the jth component of $\mathbf{x}' = [x_1, \cdots, x_m] \in \mathbb{R}^m$. From Esséen [6] we have

LEMMA 2. *Let* $\{\mathbf{y}_1, \cdots, \mathbf{y}_N\}$ *be a sequence of* i.i.d. *random vectors in* $\mathbb{R}^m$ *whose elements have zero means and unit variances, are uncorrelated, and have finite moments* θ_{4j}, $1 \leqq j \leqq m$. *Define* $\mathbf{y}(N) = N^{-\frac{1}{2}}(\mathbf{y}_1 + \cdots + \mathbf{y}_N)$ and $G_N(z) = P_N(\mathbf{y}'(N)\mathbf{y}(N) \leqq z)$, *where* $P_N(\cdot)$ *is the measure associated with* $\mathbf{y}(N)$. *Then for all* $N = 1, 2, \cdots$,

$$\sup_z |G_N(z) - \Psi_m(z)| \leqq \frac{c(m)\theta_4^{\frac{3}{2}}}{N^{m/(m+1)}}$$

where $\theta_4 = \theta_{41} + \cdots + \theta_{4m}$ *and* $c(m)$ *is a finite positive constant depending only on* m.

Bounds on the rate of convergence of the cdf of $\mathbf{y}(N)$ itself also are available (cf. Bergström [2] and Sazonov [8]) as follows.

LEMMA 3. *Let* $\{\mathbf{y}_1, \cdots, \mathbf{y}_N\}$ *be an* i.i.d. *sequence with typical member* $\mathbf{y} \in \mathbb{R}^m$ *such that* $\mathscr{E}\mathbf{y} = \mathbf{0}$, $\mathscr{E}\mathbf{y}\mathbf{y}' = \mathbf{\Sigma}$ *(a definite matrix), and* $\theta_{3j} = \mathscr{E}|\mathbf{y}_j|^3 < \infty$ *for* $1 \leqq j \leqq m$. *Let* $F_N(\cdot)$ *be the* cdf *of* $\mathbf{y}(N) = N^{-\frac{1}{2}}(\mathbf{y}_1 + \cdots + \mathbf{y}_N)$ *and* $\Phi(\cdot)$ *the* m-*dimensional Gaussian* cdf *having zero means and the second-moment matrix* Σ. *Then for all* $N = 1, 2, \cdots$,

$$\sup_{\mathbf{x} \in \mathbb{R}^m} |F_N(\mathbf{x}) - \Phi(\mathbf{x})| \leqq \frac{c_0(m) \sum_{i=1}^m \gamma_{ii}^{\frac{3}{2}} \theta_{3i}}{N^{\frac{1}{2}}}$$

where $\mathbf{\Gamma} = [\gamma_{ij}] = \mathbf{\Sigma}^{-1}$ *and* $c_0(m)$ *is a finite positive constant depending only on* m.

More general results are available. Let $\mathscr{C}^m$ be the class of all measurable convex sets in $\mathbb{R}^m$. The following result is in a form due to Sazonov [8], [9]; it also follows upon modifying a proof due to Bergström [3], who assumed that $\theta_{3j} < \infty$, $1 \leqq j \leqq r$, or upon specializing some findings of Bhattacharya [4], who assumed that $\theta_{3+\delta, j} < \infty$ for some positive δ and $1 \leqq j \leqq r$.

LEMMA 4. *Let* $\{\mathbf{y}_1, \cdots, \mathbf{y}_N\}$ *be an* i.i.d. *sequence in* $\mathbb{R}^m$ *having zero means, the nonsingular second-moment matrix* $\mathbf{\Sigma}$, *and finite absolute third moments* θ_{3j}, $1 \leqq j \leqq m$. *Let* $P_N(\cdot)$ *be the probability measure associated with* $N^{-\frac{1}{2}}(\mathbf{y}_1 + \cdots + \mathbf{y}_N)$ *and let* $P(\cdot)$ *be the limiting Gaussian measure having the parameters* $\mathbf{0}$ *and* $\mathbf{\Sigma}$. *Then for each* $N = 1, 2, \cdots$,

$$\sup_{A \in \mathscr{C}^m} |P_N(A) - P(A)| \leqq \frac{c_1(m) \sum_{i=1}^m \gamma_{ii}^{\frac{3}{2}} \theta_{3i}}{N^{\frac{1}{2}}}$$

where $\mathbf{\Gamma} = [\gamma_{ij}] = \mathbf{\Sigma}^{-1}$ *and* $c_1(m)$ *is a finite positive constant depending only on* m.

REMARK. It follows from Bergström's [3] development that $c_1(m)$ can be replaced by a function of m and $\mathbf{\Sigma}$ and a constant not depending on m as follows.

Let $\{\lambda_1 \geqq \lambda_2 \geqq \cdots \geqq \lambda_m > 0\}$ be the eigenvalues of $\mathbf{\Sigma}$ and define $\delta^2 = \lambda_m/\lambda_1$. Then for each $N = 1, 2, \cdots$,

$$\sup_{A \in \mathscr{C}^m} |P_N(A) - P(A)| \leqq \frac{c_1^* m^3 \sum_{i=1}^m \gamma_{ii}^{\frac{3}{2}} \theta_{3i}}{\delta N^{\frac{1}{2}}}$$

where c_1^* is a finite positive constant not depending on m, $\mathbf{\Sigma}$, θ_{3i} or N.

We recall some notions of convexity. Following Berge [1], let C be an open convex set in $\mathbb{R}^n$. A real-valued function $f(\cdot)$ is said to be convex in C if, for each pair of points $\mathbf{x}, \mathbf{y} \in C$ and each $\alpha = 1 - \bar{\alpha} \in [0, 1]$, we have $f(\alpha\mathbf{x} + \bar{\alpha}\mathbf{y}) \leqq \alpha f(\mathbf{x}) + \bar{\alpha} f(\mathbf{y})$. Moreover, if $f(\cdot)$ is bi-differentiable in C, it is convex in C if and only if the matrix $[f_{ij}'']$ of second partial derivatives is positive definite. If f is convex, then f is *quasi convex*, i.e. the set $A(z) = \{\mathbf{x} \,|\, f(\mathbf{x}) \leqq z\}$ is convex for each $z \in \mathbb{R}^1$.

DEFINITION. Let $\mathbf{x} \in \mathbb{R}^m$, $u \in \mathbb{R}_+^1$, and $s \in \mathbb{R}_+^1 - u$; then $\mathscr{X}$ is the natural product space $\mathscr{X} = \mathbb{R}^m \times (\mathbb{R}_+^1 - u)$ in $\mathbb{R}^{m+1}$.

LEMMA 5. *Let $\mathbf{x} \in \mathbb{R}^m$ and $s \in \mathbb{R}_+^1 - u$. If $\mathbf{M}(m \times m)$ is a fixed positive semi-definite matrix of rank $m' \leqq m$ and $\boldsymbol{\tau}$ a fixed element of $\mathbb{R}^m$, then the function $f(\mathbf{x}, s) = (\mathbf{x} - \boldsymbol{\tau})'\mathbf{M}(\mathbf{x} - \boldsymbol{\tau})/(s + u)$ is convex in $\mathscr{X}$.*

PROOF. From invariance considerations we inquire whether $(x_1, \cdots, x_m, y) \to [(x_1 - \tau_1)^2 + \cdots + (x_m - \tau_m)^2]/y$ is convex in $\mathbb{R}^m \times \mathbb{R}_+^1$. Evidently it suffices to determine that $(x, y) \to (x - \tau)^2/y$ is convex in $\mathbb{R}^1 \times \mathbb{R}_+^1$. Upon differentiating twice, we find that convexity follows from the nonnegativity of the matrix

$$\begin{bmatrix} 2/y & -2(x - \tau)/y^2 \\ -2(x - \tau)/y^2 & 2(x - \tau)^2/y^3 \end{bmatrix}$$

of second partial derivatives, which completes the proof.

LEMMA 6. *Let $\mathbf{x} \in \mathbb{R}^m$, $\boldsymbol{\tau} \in \mathbb{R}^m$, $u \in \mathbb{R}_+^1$ and $s \in \mathbb{R}_+^1 - u$, and partition $\mathbf{x}' = [\mathbf{x}_1', \cdots, \mathbf{x}_r']$ and $\boldsymbol{\tau}' = [\boldsymbol{\tau}_1', \cdots, \boldsymbol{\tau}_r']$ such that $\mathbf{x}_j$ and $\boldsymbol{\tau}_j$ are of order $(m_j \times 1)$ and $m_1 + \cdots + m_r = m$. Let $\{\mathbf{M}_i(m_i \times m_i); 1 \leqq i \leqq r\}$ be an arbitrary collection of positive semidefinite matrices. Then the set*

$$A(z_1, \cdots, z_r) = \{(\mathbf{x}, s) \,|\, (\mathbf{x}_i - \boldsymbol{\tau}_i)'\mathbf{M}_i(\mathbf{x}_i - \boldsymbol{\tau}_i)/(s + u) \leqq z_i;\ 1 \leqq i \leqq r\}$$

is convex in $\mathscr{X}$ for each $\mathbf{z} = [z_1, \cdots, z_r]' \in \mathbb{R}_+^r$.

PROOF. Let $\{A_1, \cdots, A_r\}$ be cylinder sets in $\mathscr{X}$ defined as

$$A_i(z_i) = \{(\mathbf{x}, s) \,|\, (\mathbf{x}_i - \boldsymbol{\tau}_i)'\mathbf{M}_i(\mathbf{x}_i - \boldsymbol{\tau}_i)/(s + u) \leqq z_i\}, \qquad 1 \leqq i \leqq r.$$

These sets clearly are convex in cross section: Let $\mathbf{M} = \mathrm{Diag}\,(\mathbf{0}, \mathbf{M}_i, \mathbf{0})$; observe that $f_i(\mathbf{x}, s) = (\mathbf{x}_i - \boldsymbol{\tau}_i)'\mathbf{M}_i(\mathbf{x}_i - \boldsymbol{\tau}_i)/(s + u)$ is convex in $\mathscr{X}$ by Lemma 5 and thus quasiconvex. The proof is complete upon noting that $A(z_1, \cdots, z_r)$ is the intersection $A(z_1, \cdots, z_r) = \bigcap_{i=1}^r A_i(z_i)$ of convex bodies in $\mathscr{X}$ and thus is convex.

COROLLARY 1. *The set*

$$B(z_1, \cdots, z_r) = \{\mathbf{x} \mid (\mathbf{x}_i - \boldsymbol{\tau}_i\}'\mathbf{M}_i(\mathbf{x}_i - \boldsymbol{\tau}_i) \leqq z_i;\ 1 \leqq i \leqq r\}$$

is convex in $\mathbb{R}^m$ *for each* $[z_1, \cdots, z_r] \in \mathbb{R}_+{}^r$.

For later reference we introduce special notation as follows. Let $\mathbf{u} = [\mathbf{u}_1', \cdots, \mathbf{u}_r']'$ be a partitioned Gaussian vector of order $(\nu \times 1)$ having the means $\boldsymbol{\mu} = [\boldsymbol{\mu}_1', \cdots, \boldsymbol{\mu}_r']'$ and the second-moment matrix $\boldsymbol{\Sigma} = [\boldsymbol{\Sigma}_{ij}]$, where $\mathbf{u}_i$ and $\boldsymbol{\mu}_i$ are of order $(\nu_i \times 1)$ and $\boldsymbol{\Sigma}_{ij}$ is of order $(\nu_i \times \nu_j)$ such that $\nu_1 + \cdots + \nu_r = \nu$. We take $\boldsymbol{\Sigma}$ to be definite. Upon defining $W_j = \mathbf{u}_j'\boldsymbol{\Sigma}_{jj}^{-1}u_j$, $1 \leqq j \leqq r$, we denote by $\Psi(\cdot, \cdots, \cdot; \boldsymbol{\nu})$ the r-dimensional cdf of $\{W_1, \cdots, W_r\}$, where $\boldsymbol{\nu} = [\nu_1, \cdots, \nu_r]$. As each one-dimensional marginal distribution of W_i is χ^2 having ν_i degrees of freedom, central or noncentral according as $\boldsymbol{\mu}_i = \mathbf{0}$ or $\boldsymbol{\mu}_i \neq \mathbf{0}$, we refer to $\Psi(\cdot, \cdots, \cdot; \boldsymbol{\nu})$ as a multivariate χ^2 distribution. An expression for its probability density function is available in the central case as a series in Laguerre polynomials of vector argument (see [7]) apart from scale. Under Gaussian theory the joint cdf of $\{V_1, \cdots, V_r\}$ is $\Psi(\cdot, \cdots, \cdot; \mathbf{p})$, where $\mathbf{p} = [p_1, \cdots, p_r]$.

3. Some large-sample results. We now examine large-sample properties of (i) the least-squares estimators for $\boldsymbol{\beta}$ and (ii) the quadratic forms V and $\{V_1, \cdots, V_r\}$. Although more general developments can be formulated, considerable simplification follows upon assuming N independent repetitions of the same experiment. This assumption is natural when designs are subject to experimental control, e.g. randomized complete block experiments having equal numbers of observations per cell. Accordingly, let $\{\mathbf{y}_1, \cdots, \mathbf{y}_N\}$ be a sequence of i.i.d. random vectors with typical member $\mathbf{y} = [y_1, \cdots, y_m]' \in \mathbb{R}^m$. Let $\{\tilde{\boldsymbol{\beta}}_1, \cdots, \tilde{\boldsymbol{\beta}}_N\}$ be the corresponding sequence of least-squares estimators, one from each experiment, where $\{\tilde{\boldsymbol{\beta}}_i = \mathbf{TX}'\mathbf{y}_i;\ 1 \leqq i \leqq N\}$. A typical member of the latter is designated as $\tilde{\boldsymbol{\beta}} = \mathbf{TX}'\mathbf{y}$. Upon defining

$$\tilde{\sigma}_i{}^2 = \mathbf{y}_i'(\mathbf{I}_m - \mathbf{XTX}')\mathbf{y}_i/(m - p)\,, \qquad 1 \leqq i \leqq N$$

we find that, based on N repetitions, the least-squares estimator for $\boldsymbol{\beta}$ is

$$\begin{aligned} \hat{\boldsymbol{\beta}}_N &= N^{-1}\mathbf{TX}'(\mathbf{y}_1 + \cdots + \mathbf{y}_N) \\ &= N^{-1}(\tilde{\boldsymbol{\beta}}_1 + \cdots + \tilde{\boldsymbol{\beta}}_N) \end{aligned} \tag{3.1}$$

and the corresponding estimator for σ^2 is

$$\hat{\sigma}_N{}^2 = N^{-1}(\tilde{\sigma}_1{}^2 + \cdots + \tilde{\sigma}_N{}^2)\,. \tag{3.2}$$

Throughout this section, however, we consider σ^2 to be known.

For later reference we itemize assumptions as follows. We invariably assume A1 and A2, and at times we adopt one or more of A3, A4 and A5.

ASSUMPTIONS.

A1. $\mathscr{E}\mathbf{y} = \mathbf{X}\boldsymbol{\beta}$, where $\mathbf{X}' = [\mathbf{x}_1, \cdots, \mathbf{x}_m]$ $(p \times m)$ is a non-stochastic matrix of rank $p \leqq m$ and $\boldsymbol{\beta}(p \times 1)$ is a vector of unknown parameters;

A2. $\mathscr{E}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})' = \sigma^2\mathbf{I}_m$, where σ^2 is a positive constant;
A3. $\mathscr{E}|y_i - \mathbf{x}_i'\boldsymbol{\beta}|^3 < \infty$, $1 \leqq i \leqq m$;
A4. $\mathscr{E}(y_i - \mathbf{x}_i'\boldsymbol{\beta})^4 < \infty$, $1 \leqq i \leqq m$;
A5. $\mathscr{E}(y_i - \mathbf{x}_i'\boldsymbol{\beta})^6 < \infty$, $1 \leqq i \leqq m$.

We first consider the large-sample properties of the standardized variables $N^{\frac{1}{2}}(\hat{\boldsymbol{\beta}}_N - \boldsymbol{\beta})$. From (3.1), together with the assumptions A1 and A2, it follows that $\mathscr{E}N^{\frac{1}{2}}(\hat{\boldsymbol{\beta}}_N - \boldsymbol{\beta}) = \mathbf{0}$ and $\mathscr{E}N(\hat{\boldsymbol{\beta}}_N - \boldsymbol{\beta})(\hat{\boldsymbol{\beta}}_N - \boldsymbol{\beta})' = \sigma^2\mathbf{T}$. The limiting distribution of $N^{\frac{1}{2}}(\hat{\boldsymbol{\beta}}_N - \boldsymbol{\beta})$, and a bound on the rate of convergence to this limit, are given in the following theorem under the assumptions A1, A2 and A3. We have

THEOREM 3.1. *Let* $\{\mathbf{y}_1, \cdots, \mathbf{y}_N\}$ *be a sequence of* i.i.d. *random variables with typical element* $\mathbf{y} \in \mathbb{R}^m$ *satisfying* A1, A2 *and* A3. *Define* $\tilde{\boldsymbol{\beta}} = \mathbf{TX}'\mathbf{y}$ *and* $\hat{\boldsymbol{\beta}}_N = N^{-1}\mathbf{TX}'(\mathbf{y}_1 + \cdots + \mathbf{y}_N)$. *Let* $F_N(\cdot)$ *be the* cdf *of* $N^{\frac{1}{2}}(\hat{\boldsymbol{\beta}}_N - \boldsymbol{\beta})$ *and* $\Phi(\cdot)$ *the p-dimensional Gaussian* cdf *having zero means and covariance matrix* $\sigma^2\mathbf{T}$. *Then*

$$\text{(i)} \qquad \lim_{N\to\infty} F_N(\cdot) = \Phi(\cdot)$$

and, for each $N = 1, 2, \cdots$,

$$\text{(ii)} \qquad \sup_{\mathbf{z}\in\mathbb{R}^p} |F_N(\mathbf{z}) - \Phi(\mathbf{z})| \leqq \frac{c_0(p)\sum_{i=1}^p \gamma_{ii}^{\frac{3}{2}}\theta_{3i}}{\sigma^3 N^{\frac{1}{2}}}$$

where $\theta_{3i} = \mathscr{E}|\tilde{\beta}_i - \beta_i|^3 < \infty$, $1 \leqq i \leqq p$; $\boldsymbol{\Gamma} = [\gamma_{ij}] = \mathbf{T}^{-1}$; *and* $c_0(p)$ *is a finite positive constant depending only on p.*

PROOF. Under assumptions A1 and A2, conclusion (i) follows immediately from a multidimensional version of the Central Limit Theorem (cf. Varadarajan [10]). Conclusion (ii) follows directly from assumption A3 and Lemma 3, together with the fact that A3 implies $\theta_{3i} < \infty$, $1 \leqq i \leqq p$.

We next consider the large-sample properties of particular quadratic forms in the elements of $\hat{\boldsymbol{\beta}}_N$. Define

$$V_N = N(\hat{\boldsymbol{\beta}}_N - \boldsymbol{\beta}_0)'\mathbf{T}^{-1}(\hat{\boldsymbol{\beta}}_N - \boldsymbol{\beta}_0)/\sigma^2 \tag{3.3}$$

and

$$V_{Nj} = N(\hat{\boldsymbol{\beta}}_{Nj} - \boldsymbol{\beta}_{j0})'\mathbf{T}_{jj}^{-1}(\hat{\boldsymbol{\beta}}_{Nj} - \boldsymbol{\beta}_{j0})/\sigma^2\,, \qquad 1 \leqq j \leqq r \tag{3.4}$$

in terms of the partitioned vector $\hat{\boldsymbol{\beta}}_N' = [\hat{\boldsymbol{\beta}}_{N1}', \cdots, \hat{\boldsymbol{\beta}}_{Nr}']$. The limiting central distribution of V_N under assumptions A1 and A2, and a bound on the rate of convergence under the further assumption A4, are given in

THEOREM 3.2. *Let* $\{\mathbf{y}_1, \cdots, \mathbf{y}_N\}$ *be a sequence of* i.i.d. *random vectors with typical member* $\mathbf{y} \in \mathbb{R}^m$ *satisfying* A1, A2 *and* A4. *Define* $\tilde{\boldsymbol{\beta}} = \mathbf{TX}'\mathbf{y}$, $\hat{\boldsymbol{\beta}}_N = N^{-1}\mathbf{TX}'(\mathbf{y}_1 + \cdots + \mathbf{y}_N)$, *and* $V_N = N(\hat{\boldsymbol{\beta}}_N - \boldsymbol{\beta})'\mathbf{T}^{-1}(\hat{\boldsymbol{\beta}}_N - \boldsymbol{\beta})/\sigma^2$. *Let* $G_N(\cdot)$ *be the* cdf *of* V_N *and* $\Psi_\nu(\cdot)$ *the* cdf *of the central* χ^2 *distribution having* ν *degrees of freedom. Then*

$$\text{(i)} \qquad \lim_{N\to\infty} G_N(\cdot) = \Psi_p(\cdot)$$

and, for each $N = 1, 2, \cdots$,

$$\text{(ii)} \qquad \sup_{z \in \mathbb{R}_+^1} |G_N(z) - \Psi_p(z)| \leq \frac{c(p)\theta_4^{\frac{3}{4}}}{(\sigma^2)^3 N^{p/(p+1)}}$$

where $\theta_4 = \sum_{i=1}^p \theta_{4i}$, $\theta_{4i} = \mathscr{E}[\mathbf{b}_i'(\tilde{\boldsymbol{\beta}} - \boldsymbol{\beta})]^4 < \infty$, $\mathbf{B}' = [\mathbf{b}_1, \cdots, \mathbf{b}_p]$ *is a nonsingular* $(p \times p)$ *matrix such that* $\mathbf{BTB}' = \mathbf{I}_p$, *and* $c(p)$ *is a finite positive constant depending only on* p.

PROOF. Clearly V_N is a definite quadratic form of the type $V_N = \mathbf{z}_N'\boldsymbol{\Sigma}^{-1}\mathbf{z}_N$ in the elements of $\mathbf{z}_N$, where $\mathscr{E}\mathbf{z}_N = \mathbf{0}$, $\mathscr{E}\mathbf{z}_N\mathbf{z}_N' = \boldsymbol{\Sigma}$, and $\mathbf{z}_N$ is asymptotically Gaussian by Theorem 3.1. Conclusion (i) of the present theorem now follows upon applying Lemma 1 together with the continuity of V_N as a function of $\mathbf{z}_N$. Choose $\mathbf{B}(p \times p)$ to be nonsingular such that $\mathbf{BTB}' = \mathbf{I}_p$, and define $\mathbf{w}_N = N^{\frac{1}{2}}\mathbf{B}(\hat{\boldsymbol{\beta}}_N - \boldsymbol{\beta})$. Then $\mathscr{E}\mathbf{w}_N = \mathbf{0}$, $\mathscr{E}\mathbf{w}_N\mathbf{w}_N' = \sigma^2\mathbf{I}_p$, and $V_N = \mathbf{w}_N'\mathbf{w}_N/\sigma^2$. Assumptions A1, A2 and A4 assure that the conditions of Lemma 2 are met, and conclusion (ii) now follows from that lemma.

The next theorem and its corollary are concerned with the central and noncentral distributions of V_N and $\{V_{N1}, \cdots, V_{Nr}\}$. Let $\Psi_\nu(\cdot; \lambda)$ be the cdf of the noncentral χ^2 distribution having ν degrees of freedom and noncentrality parameter λ; this arises as the limiting nonnull distribution of V_N under a sequence of local alternatives, i.e. a sequence $\{\boldsymbol{\beta}_N; N = 1, 2, \cdots\}$ such that $\boldsymbol{\beta}_N = \boldsymbol{\beta}_0 + O(N^{\frac{1}{2}})$. For fixed N, however, λ is bounded without recourse to local alternatives, which fact we exploit for computing bounds. In what follows we refer to the noncentral distributions $\Psi_\nu(\cdot; \lambda)$ and $\Psi(\cdot, \cdots, \cdot; \boldsymbol{\nu})$, the particular assumptions regarding their noncentrality parameters becoming clear in the context.

THEOREM 3.3. *Let* $\{\mathbf{y}_1, \cdots, \mathbf{y}_N\}$ *be an* i.i.d. *sequence with typical member* $\mathbf{y} \in \mathbb{R}^m$ *satisfying* A1, A2 *and* A3. *Define* $\tilde{\boldsymbol{\beta}} = \mathbf{TX}'\mathbf{y}$, $\hat{\boldsymbol{\beta}}_N = N^{-1}\mathbf{TX}'(\mathbf{y}_1 + \cdots + \mathbf{y}_N)$, *and*

$$V_{Nj} = N(\hat{\boldsymbol{\beta}}_{Nj} - \boldsymbol{\beta}_{j0})'\mathbf{T}_{jj}^{-1}(\hat{\boldsymbol{\beta}}_{Nj} - \boldsymbol{\beta}_{j0})/\sigma^2\,, \qquad 1 \leq j \leq r$$

and let $G_N(\cdot, \cdots, \cdot)$ *be the joint* cdf *of* $\{V_{N1}, \cdots, V_{Nr}\}$. *Then under suitable local alternatives we have*

$$\text{(i)} \qquad \lim_{N\to\infty} G_N(\cdot, \cdots, \cdot) = \Psi(\cdot, \cdots, \cdot; \mathbf{p})$$

where $\mathbf{p} = [p_1, \cdots, p_r]'$ *and, for each* $N = 1, 2, \cdots$,

$$\text{(ii)} \qquad \sup_{\mathbf{z} \in \mathbb{R}_+^r} |G_N(\mathbf{z}) - \Psi(\mathbf{z}; \mathbf{p})| \leq \frac{c_1(p) \sum_{i=1}^p \gamma_{ii}^{\frac{3}{2}}\theta_{3i}}{\sigma^3 N^{\frac{1}{2}}}$$

where $\boldsymbol{\Gamma} = [\gamma_{ij}] = \mathbf{T}^{-1}$, $\theta_{3i} = \mathscr{E}|\tilde{\beta}_i - \beta_i|^3$, *and* $c_1(p)$ *is a finite positive constant depending only on* p. *Alternatively let* $\{\lambda_1 \geq \lambda_2 \cdots \geq \lambda_p > 0\}$ *be the eigenvalues of* $\mathbf{T}$ *and define* $\delta^2 = \lambda_p/\lambda_1$. *Then for each* $N = 1, 2, \cdots$,

$$\text{(iii)} \qquad \sup_{\mathbf{z} \in \mathbb{R}_+^r} |G_N(\mathbf{z}) - \Psi(\mathbf{z}; \mathbf{p})| \leq \frac{c_1^* p^3 \sum_{i=1}^p \gamma_{ii}^{\frac{3}{2}}\theta_{3i}}{\sigma^3\delta N^{\frac{1}{2}}}$$

where c_1^* *is a finite positive constant not depending on* p, $\mathbf{T}$, θ_{3i} *or* N.

PROOF. Define $\boldsymbol{\tau}_N = N^{\frac{1}{2}}(\boldsymbol{\beta}_0 - \boldsymbol{\beta})$; let $\mathbf{w}_N$ be the normalized sum $\mathbf{w}_N = N^{\frac{1}{2}}(\hat{\boldsymbol{\beta}}_N - \boldsymbol{\beta})$ and $P_N(\cdot)$ its probability measure. Let $P(\cdot)$ be the p-dimensional Gaussian measure having zero means and the second-moment matrix $\sigma^2\mathbf{T}$. Clearly the forms $\{V_{N1}, \cdots, V_{Nr}\}$ are quadratic, thus continuous, functions of $\mathbf{w}_N$, and the latter is asymptotically Gaussian by Theorem 3.1. Lemma 1 establishes under local alternatives that $\lim_{N\to\infty} G_N(\cdot, \cdots, \cdot)$ is the joint distribution of quadratic forms in Gaussian variables whose joint cdf is $\Psi(\cdot, \cdots, \cdot; \mathbf{p})$, and conclusion (i) now follows. Conformably partition $\mathbf{w}_N' = [\mathbf{w}_{N1}', \cdots, \mathbf{w}_{Nr}']$ and $\boldsymbol{\tau}_N' = [\boldsymbol{\tau}_{N1}', \cdots, \boldsymbol{\tau}_{Nr}']$. In terms of $P_N(\cdot)$ and $P(\cdot)$ we write the cdf of $\{V_{N1}, \cdots, V_{Nr}\}$ and its normal-theory approximation as $G_N(z_1, \cdots, z_r) = P_N(A(z_1, \cdots, z_r))$ and $\Psi(z_1, \cdots, z_r; \mathbf{p}) = P(A(z_1, \cdots, z_r))$, respectively, where

$$A(z_1, \cdots, z_r) = \{\mathbf{w}_N \mid (\mathbf{w}_{Ni} - \boldsymbol{\tau}_{Ni})'\mathbf{T}_{ii}^{-1}(\mathbf{w}_{Ni} - \boldsymbol{\tau}_{Ni}) \leqq z_i;\ 1 \leqq i \leqq r\}\,.$$

The set $A(\mathbf{z})$ is convex in $\mathbb{R}^p$ by Corollary 1; the assumptions A1, A2 and A3 assure that the conditions of Lemma 4 are met; and conclusions (ii) and (iii) now follow from Lemma 4 and the remarks immediately following it.

COROLLARY 2. *Let* $\{\mathbf{y}_1, \cdots, \mathbf{y}_N\}$ *be an* i.i.d. *sequence satisfying* A1, A2 *and* A3. *Define* $\tilde{\boldsymbol{\beta}}$, $\hat{\boldsymbol{\beta}}_N$, *and* $V_N = N(\hat{\boldsymbol{\beta}}_N - \boldsymbol{\beta}_0)'\mathbf{T}^{-1}(\hat{\boldsymbol{\beta}}_N - \boldsymbol{\beta}_0)/\sigma^2$ *as before, and let* $F_N(\cdot)$ *be the* cdf *of* V_N. *Then under suitable local alternatives we have*

(i) $$\lim_{N\to\infty} F_N(\cdot) = \Psi_\nu(\cdot; \lambda)$$

and, for each $N = 1, 2, \cdots,$

(ii) $$\sup_{z\in\mathbb{R}_+^1} |G_N(z) - \Psi_\nu(z; \lambda_N)| \leqq \frac{c_1(p)\sum_{i=1}^p \gamma_{ii}^{\frac{3}{2}}\theta_{3i}}{\sigma^3 N^{\frac{1}{2}}}$$

where $\lambda_N = N(\boldsymbol{\beta}_0 - \boldsymbol{\beta})'\mathbf{T}^{-1}(\boldsymbol{\beta}_0 - \boldsymbol{\beta})/\sigma^2$ *and the bound on the right is defined as in Theorem* 3.3.

REMARKS. Theorem 3.2 and Corollary 2 provide bounds on the rate of convergence of the distribution of V_N to its limit under various moment conditions. Observe that bounds of order $O(N^{-p/(p+1)})$ are provided in Theorem 3.2 for the central case under a fourth-moment condition, as compared to bounds of order $O(N^{-\frac{1}{2}})$ under finite absolute third moments in Corollary 2. The referee has pointed out that Esséen-type bounds of order $O(N^{-p/(p+1)})$ cannot, in general, hold in the noncentral case. For it is necessary (but by no means sufficient) that the second term in the formal Cramér–Edgeworth expansion vanish in order that the error of the Gaussian approximation to the probabilility of some set B be of order smaller than $N^{-\frac{1}{2}}$. When evaluated over a ball whose center is not the origin, such expansions do not have a vanishing second term unless the underlying distribution exhibits further special symmetries, e.g. the vanishing of all moments of third order.

4. Convergence rates for variance ratios. We relax the requirement that σ^2 be known. From assumptions A1 and A2 we have $\mathscr{E}\tilde{\sigma}_i^2 = \sigma^2$, where $\tilde{\sigma}_i^2 = \mathbf{y}_i'(\mathbf{I}_m - \mathbf{X}\mathbf{T}\mathbf{X}')\mathbf{y}_i/(m-p)$, $1 \leqq i \leqq N$. Under the further assumption that

$\{\mathbf{y}_1, \cdots, \mathbf{y}_N\}$ are i.i.d. random vectors in $\mathbb{R}^m$, it follows that $\{\tilde{\sigma}_1^2, \cdots, \tilde{\sigma}_N^2\}$ are i.i.d. random variables in $\mathbb{R}_+^1$. Thus by one of Kolmogorov's strong laws of large numbers we infer that $\hat{\sigma}_N^2 = N^{-1}(\tilde{\sigma}_1^2 + \cdots + \tilde{\sigma}_N^2)$ converges almost surely (a.s.) to σ^2 and we write

$$\hat{\sigma}_N^2 \to_{\text{a.s.}} \sigma^2 \qquad \text{as} \quad N \to \infty . \tag{4.1}$$

We specifically consider the variance ratios U_N and $\{U_{N1}, \cdots, U_{Nr}\}$ given in

$$U_N = N(\hat{\boldsymbol{\beta}}_N - \boldsymbol{\beta}_0)'\mathbf{T}^{-1}(\hat{\boldsymbol{\beta}}_N - \boldsymbol{\beta}_0)/\hat{\sigma}_N^2 \tag{4.2}$$

$$= N(\hat{\boldsymbol{\beta}}_{Nj} - \boldsymbol{\beta}_{j0})'\mathbf{T}_{jj}^{-1}(\hat{\boldsymbol{\beta}}_{Nj} - \boldsymbol{\beta}_{j0})/\hat{\sigma}_N^2 , \qquad 1 \leqq j \leqq r . \tag{4.3}$$

Their limiting distributions are $\Psi_p(\cdot; \lambda)$ and $\Psi(\cdot, \cdots, \cdot; \mathbf{p})$, respectively. These limits, however, can be considered the end results of two limiting processes, namely, the convergence in distribution of V_N and $(V_{N1}, \cdots, V_{Nr}\}$, and the a.s. convergence to unity of $\hat{\sigma}_N^2/\sigma^2$. In order to study these processes together, our construction employs Berry–Esséen type bounds on the rate of convergence of the joint probability measure associated with $N^{\frac{1}{2}}(\hat{\boldsymbol{\beta}}_N - \boldsymbol{\beta})$ and $N^{\frac{1}{2}}(\hat{\sigma}_N^2 - \sigma^2)$. This construction is valid under further assumptions, for the existence of sixth-order moments of $\mathbf{y}$ assures the third absolute moments of $\tilde{\sigma}^2 = \mathbf{y}'(\mathbf{I}_m - \mathbf{XTX}')\mathbf{y}/(m - p)$.

We first consider U_N. In terms of the typical experiment $\mathbf{y}$ we write

$$\tilde{\boldsymbol{\xi}}' = [(\tilde{\boldsymbol{\beta}} - \boldsymbol{\beta})', (\tilde{\sigma}^2 - \sigma^2)] \tag{4.4}$$

and $\mathscr{E}\tilde{\boldsymbol{\xi}}\tilde{\boldsymbol{\xi}}' = \boldsymbol{\Sigma}$, of order $(p + 1) \times (p + 1)$, where

$$\boldsymbol{\Sigma} = \begin{bmatrix} \sigma^2\mathbf{T} & \mathbf{c} \\ \mathbf{c}' & \omega^2 \end{bmatrix} . \tag{4.5}$$

Define

$$\boldsymbol{\xi}_N{}^* = [(\hat{\boldsymbol{\beta}}_N - \boldsymbol{\beta})', (\hat{\sigma}_N^2 - \sigma^2)]' \tag{4.6}$$

$$\hat{\boldsymbol{\xi}}_N = [(\hat{\boldsymbol{\beta}}_N - \boldsymbol{\beta})', \hat{\sigma}_N^2]' \tag{4.7}$$

i.e. $\hat{\boldsymbol{\xi}}_N = \boldsymbol{\xi}_N{}^* + [\mathbf{0}', \sigma^2]'$, and observe that $\boldsymbol{\xi}_N{}^* \in \mathbb{R}^p \times (\mathbb{R}_+^1 - \sigma^2)$, while $\hat{\boldsymbol{\xi}}_N \in \mathbb{R}^p \times \mathbb{R}_+^1$.

We identify probability measures as follows. With the standardized sums $N^{\frac{1}{2}}\boldsymbol{\xi}_N{}^*$ we associate $P_N{}^*(\cdot)$ and, for $N < \infty$, we associate $P_N(\cdot)$ with $N^{\frac{1}{2}}\hat{\boldsymbol{\xi}}_N$. Clearly $P_N{}^*$ and P_N are identical apart from shift. Let $\Phi_{p+1}^*(\cdot)$ be the $(p + 1)$-dimensional Gaussian cdf, and $Q_{p+1}^*(\cdot)$ the corresponding Gaussian measure, having zero means and the covariance matrix $\boldsymbol{\Sigma}$ at (4.5). Similarly let $\Phi_p{}^*(\cdot)$ be the joint marginal cdf of the first p components and $Q_p{}^*(\cdot)$ the corresponding Gaussian measure.

Let $\mathbf{M}(p \times p)$ be positive definite and symmetric, let $\boldsymbol{\tau} \in \mathbb{R}^p$, and suppose $\gamma > 0$. Define

$$A^*(z) = \{(\mathbf{w}, s) \in \mathbb{R}^p \times (\mathbb{R}_+^1 - \gamma) \mid (\mathbf{w} - \boldsymbol{\tau})'\mathbf{M}(\mathbf{w} - \boldsymbol{\tau})/(s + \gamma) \leqq N^{-\frac{1}{2}}z\} \tag{4.8}$$

$$A(z) = \{(\mathbf{w}, u) \in \mathbb{R}^p \times \mathbb{R}_+^1 \mid (\mathbf{w} - \boldsymbol{\tau})'\mathbf{M}(\mathbf{w} - \boldsymbol{\tau})/u \leqq N^{-\frac{1}{2}}z\} \tag{4.9}$$

and

$$B(z) = \{\mathbf{w} \in \mathbb{R}^p \,|\, (\mathbf{w} - \boldsymbol{\tau})'\mathbf{M}(\mathbf{w} - \boldsymbol{\tau})/\sigma^2 \leqq z\} . \tag{4.10}$$

The foregoing construction supports the following conclusions. From the invariance of the measure of a set to translation of both the measure and the set, it follows that

$$P_N{}^*(A^*(z)) = P_N(A(z)) . \tag{4.11}$$

Upon identifying $\mathbf{w} = N^{\frac{1}{2}}(\hat{\boldsymbol{\beta}}_N - \boldsymbol{\beta})$, $\boldsymbol{\tau} = N^{\frac{1}{2}}(\boldsymbol{\beta}_0 - \boldsymbol{\beta})$ and $\mathbf{M} = \mathbf{T}^{-1}$, we find that

$$\begin{aligned} Q_p{}^*(B(z)) &= Q_p{}^*(N(\hat{\boldsymbol{\beta}}_N - \boldsymbol{\beta}_0)'\mathbf{T}^{-1}(\hat{\boldsymbol{\beta}}_N - \boldsymbol{\beta}_0)/\sigma^2 \leqq z) \\ &= \Psi_p(z; \lambda_N) \end{aligned} \tag{4.12}$$

where $\lambda_N = N(\boldsymbol{\beta}_0 - \boldsymbol{\beta})'\mathbf{T}^{-1}(\boldsymbol{\beta}_0 - \boldsymbol{\beta})/\sigma^2$. Upon making the further identification $u = N^{\frac{1}{2}}\hat{\sigma}_N{}^2$, we find that

$$\begin{aligned} P_N(A(z)) &= P_N(N(\hat{\boldsymbol{\beta}}_N - \boldsymbol{\beta}_0)'\mathbf{T}^{-1}(\hat{\boldsymbol{\beta}}_N - \boldsymbol{\beta}_0)/N^{\frac{1}{2}}\hat{\sigma}_N{}^2 \leqq N^{-\frac{1}{2}}z) \\ &= P_N(U_N \leqq z) = G_N{}^*(z) \end{aligned} \tag{4.13}$$

where $G_N{}^*(\cdot)$ is the cdf of U_N. Finally let $\gamma = N^{\frac{1}{2}}\sigma^2$ and define

$$H_N(z) = Q_{p+1}^*(A^*(z)) . \tag{4.14}$$

Then we have

THEOREM 4.1. *Let* $\{\mathbf{y}_1, \cdots, \mathbf{y}_N\}$ *be an* i.i.d. *sequence in* $\mathbb{R}^m$ *satisfying* A1, A2 *and* A5. *Define* $\tilde{\boldsymbol{\xi}}' = [(\tilde{\boldsymbol{\beta}} - \boldsymbol{\beta})', (\tilde{\sigma}^2 - \sigma^2)]$; *let* $U_N = N(\hat{\boldsymbol{\beta}}_N - \boldsymbol{\beta}_0)'\mathbf{T}^{-1}(\hat{\boldsymbol{\beta}}_N - \boldsymbol{\beta}_0)/\hat{\sigma}_N{}^2$ *and denote its* cdf *by* $G_N{}^*(\cdot)$. *Then under suitable local alternatives we have*

$$\text{(i)} \qquad \lim_{N\to\infty} G_N{}^*(\cdot) = \Psi_p(\cdot; \lambda)$$

and, for each $z \in \mathbb{R}_+{}^1$ *and for each* $N = 1, 2, \cdots,$

$$\text{(ii)} \qquad |G_N{}^*(z) - \Psi_p(z; \lambda_N)| \leqq \frac{c_1(p+1)\sum_{i=1}^{p+1}\gamma_{ii}^{\frac{3}{2}}\theta_{3i}}{N^{\frac{1}{2}}} + |H_N(z) - \Psi_p(z; \lambda_N)|$$

where $\lambda_N = N(\boldsymbol{\beta}_0 - \boldsymbol{\beta})'\mathbf{T}^{-1}(\boldsymbol{\beta}_0 - \boldsymbol{\beta})/\sigma^2$, $\boldsymbol{\Gamma} = [\gamma_{ij}] = \boldsymbol{\Sigma}^{-1}$, $\theta_{3i} = \mathscr{E}|\xi_i|^3 < \infty$, $c_1(p+1)$ *is a finite positive constant depending only on* p, *and* $H_N(z)$ *is defined at* (4.14). *Alternatively, if* $\{\lambda_1 \geqq \lambda_2 \geqq \cdots \geqq \lambda_{p+1} > 0\}$ *are the eigenvalues of* $\boldsymbol{\Sigma}$ *and if* $\delta^2 = \lambda_{p+1}/\lambda_1$, *then*

$$\text{(iii)} \qquad |G_N{}^*(z) - \Psi_p(z; \lambda_N)| \leqq \frac{c_1{}^* \cdot (p+1)^3 \sum_{i=1}^{p+1}\gamma_{ii}^{\frac{3}{2}}\theta_{3i}}{\delta N^{\frac{1}{2}}} + |H_N(z) - \Psi_p(z; \lambda_N)|$$

where $c_1{}^*$ *is a finite positive constant not depending on any of the remaining parameters.*

PROOF. Use (3.3) and (4.2) to write $U_N = \sigma^2 V_N/\hat{\sigma}_N{}^2$. Because $\hat{\sigma}_N{}^2/\sigma^2 \to_{\text{a.s.}} 1$ as $N \to \infty$ it follows from Lemma 1 that $\mathscr{L}_\infty(U_N) = \mathscr{L}(U)$, where U has the distribution $\Psi_p(\cdot; \lambda)$ under appropriate local alternatives. Upon applying (4.13), (4.12) and (4.11), then the triangle inequality, we write

$$\begin{aligned}
&|G_N{}^*(z) - \Psi_p(z; \lambda_N)| \\
&\quad = |P_N(A(z)) - Q_p{}^*(B(z))| \\
&\quad = |P_N{}^*(A^*(z)) - Q_{p+1}^*(A^*(z)) + Q_{p+1}^*(A^*(z)) - Q_p{}^*(B(z))| \\
&\quad \leqq |P_N{}^*(A^*(z)) - Q_{p+1}^*(A^*(z))| + |Q_{p+1}^*(A^*(z)) - Q_p{}^*(B(z))| \,.
\end{aligned}$$

Now the second term following the inequality is

$$|Q_{p+1}^*(A^*(z)) - Q_p{}^*(B(z))| = |H_N(z) - \Psi_p(z; \lambda_N)|$$

from (4.12) and (4.13). Moreover, $P_N{}^*(\cdot)$ converges to $Q_{p+1}^*(\cdot)$ from assumptions A1, A2 and A4 (the latter implied by A5). We now invoke the assumption A5 directly, together with the fact that, for all z, $A^*(z)$ is convex on $\mathbb{R}^p \times (\mathbb{R}_+{}^1 - N^{\frac{1}{2}}\sigma^2)$, the domain of $N^{\frac{1}{2}}\boldsymbol{\xi}_N{}^*$ as defined at (4.6). This fact was established in Lemma 5. Clearly for each $z \in \mathbb{R}_+{}^1$ we have

$$|P_N{}^*(A^*(z)) - Q_{p+1}^*(A^*(z))| \leqq \sup_{C \in \mathscr{C}^{p+1}} |P_N{}^*(C) - Q_{p+1}^*(C)|$$

and the right member in turn is bounded as in Lemma 4 and the remarks following it. The proof is now complete.

Observe that $|H_N(z) - \Psi_p(z; \lambda_N)|$ is the difference between definite integrals over multivariate Gaussian densities depending on N. Some attempts toward bounding this difference are given after the next theorem.

Our next developments treat the joint distribution of $\{U_{N1}, \cdots, U_{Nr}\}$ using arguments parallel to those employed in the proof of Theorem 4.1. Let $\{\mathbf{M}_i(p_i \times p_i); 1 \leqq i \leqq r\}$ be positive definite symmetric matrices; let $\boldsymbol{\tau} \in \mathbb{R}^p$ and $\gamma > 0$. Partition $\mathbf{w}$ and $\boldsymbol{\tau}$ as $\mathbf{w}' = [\mathbf{w}_1', \cdots, \mathbf{w}_r']$ and $\boldsymbol{\tau}' = [\boldsymbol{\tau}_1', \cdots, \boldsymbol{\tau}_r']$ such that $\mathbf{w}_i$ and $\boldsymbol{\tau}_i$ are $(p_i \times 1)$ and $p_1 + \cdots + p_r = p$, and let $\mathscr{X}_\gamma = \mathbb{R}^p \times (\mathbb{R}_+{}^1 - \gamma)$ and $\mathscr{X}_0 = \mathbb{R}^p \times \mathbb{R}_+{}^1$. Define $A^*(\mathbf{z}) = A^*(z_1, \cdots, z_r)$ and $A(\mathbf{z}) = A(z_1, \cdots, z_r)$ as

$$A^*(\mathbf{z}) = \{(\mathbf{w}, s) \in \mathscr{X}_\gamma \,|\, (\mathbf{w}_j - \boldsymbol{\tau}_j)'\mathbf{M}_j(\mathbf{w}_j - \boldsymbol{\tau}_j)/(s + \gamma) \leqq N^{-\frac{1}{2}}z_j; \; 1 \leqq j \leqq r\} \tag{4.15}$$

$$A(\mathbf{z}) = \{(\mathbf{w}, u) \in \mathscr{X}_0 \,|\, (\mathbf{w}_j - \boldsymbol{\tau}_j)'\mathbf{M}_j(\mathbf{w}_j - \boldsymbol{\tau}_j)/u \leqq N^{-\frac{1}{2}}z_j; \; 1 \leqq j \leqq r\}\,, \tag{4.16}$$

respectively, and $B(z_1, \cdots, z_r)$ as

$$B(\mathbf{z}) = \{\mathbf{w} \in \mathbb{R}^p \,|\, (\mathbf{w}_j - \boldsymbol{\tau}_j)'\mathbf{M}_j(\mathbf{w}_j - \boldsymbol{\tau}_j)/\sigma^2 \leqq N^{-\frac{1}{2}}z_j; \; 1 \leqq j \leqq r\}\,. \tag{4.17}$$

From the foregoing construction we conclude immediately that

$$P_N{}^*(A^*(z_1, \cdots, z_r)) = P_N(A(z_1, \cdots, z_r))\,. \tag{4.18}$$

Upon identifying $\mathbf{w}_j = N^{\frac{1}{2}}(\hat{\boldsymbol{\beta}}_{Nj} - \boldsymbol{\beta}_j)$, $\boldsymbol{\tau}_j = N^{\frac{1}{2}}(\boldsymbol{\beta}_{0j} - \boldsymbol{\beta}_j)$ and $\mathbf{M}_j = \mathbf{T}_{jj}^{-1}$, we further conclude (compare (4.12)) that

$$Q_p{}^*(B(z_1, \cdots, z_r)) = \Psi(z_1, \cdots, z_r; \mathbf{p})\,. \tag{4.19}$$

Corresponding to (4.13) where $u = N^{\frac{1}{2}}\hat{\sigma}_N^2$, we have

$$P_N(A(z_1, \cdots, z_r)) = G_N^*(z_1, \cdots, z_r) \tag{4.20}$$

where $G_N^*(\cdot, \cdots, \cdot)$ is the joint cdf of $\{U_{N1}, \cdots, U_{Nr}\}$. Finally let $\gamma = N^{\frac{1}{2}}\sigma^2$ at (4.15) and define

$$H_N(z_1, \cdots, z_r) = Q_{p+1}^*(A^*(z_1, \cdots, z_r)) \,. \tag{4.21}$$

The limiting form of the joint distribution of $\{U_{N1}, \cdots, U_{Nr}\}$, and a bound on the rate of convergence to this limit, are given in the following theorem.

THEOREM 4.2. *Let* $\{\mathbf{y}_1, \cdots, \mathbf{y}_N\}$ *be an* i.i.d. *sequence in* $\mathbb{R}^m$ *satisfying* A1, A2 *and* A5. *Define* $\tilde{\boldsymbol{\xi}}' = [(\tilde{\boldsymbol{\beta}} - \boldsymbol{\beta})', (\tilde{\sigma}^2 - \sigma^2)]$ *and let* $G_N^*(\cdot, \cdots, \cdot)$ *be the joint* cdf *of*

$$U_{Nj} = N(\hat{\boldsymbol{\beta}}_{Nj} - \boldsymbol{\beta}_{j0})'\mathbf{T}_{jj}^{-1}(\hat{\boldsymbol{\beta}}_{Nj} - \boldsymbol{\beta}_{j0})/\hat{\sigma}_N^2\,; \qquad 1 \leqq j \leqq r\,.$$

Then under suitable local alternatives we have

$$\text{(i)} \qquad \lim_{N\to\infty} G_N^*(\cdot, \cdots, \cdot) = \Psi(\cdot, \cdots, \cdot; \mathbf{p})$$

and, for each $\mathbf{z} \in \mathbb{R}_+^r$ *and for each* $N = 1, 2, \cdots,$

$$\text{(ii)} \qquad |G_N^*(\mathbf{z}) - \Psi(\mathbf{z}; \mathbf{p})| \leqq \frac{c_1(p+1)\sum_{i=1}^{p+1}\gamma_{ii}^{\frac{3}{2}}\theta_{3i}}{N^{\frac{1}{2}}} + |H_N(\mathbf{z}) - \Psi(\mathbf{z}; \mathbf{p})|$$

where $\boldsymbol{\Gamma} = [\gamma_{ij}] = \boldsymbol{\Sigma}^{-1}$, $\theta_{3i} = \mathscr{E}|\tilde{\xi}_i|^3 < \infty$, $c_1(p+1)$ *is a finite positive constant depending only on* p, *and* $H_N(\mathbf{z})$ *is defined at* (4.21).

PROOF. The proof parallels that of Theorem 4.1. Combine (3.4) and (4.3) as $U_{Nj} = \sigma^2 V_{Nj}/\hat{\sigma}_N^2$, $1 \leqq j \leqq r$. Conclusion (i) now follows from (4.1) and Theorem 3.3 by a standard argument. Proceeding as in the proof of Theorem 4.1, but now using (4.15)—(4.21), we have

$$|G_N^*(\mathbf{z}) - \Psi(\mathbf{z}; \mathbf{p})| \leqq |P_N^*(A^*(\mathbf{z})) - Q_{p+1}^*(A^*(\mathbf{z}))| + |H_N(\mathbf{z}) - \Psi(\mathbf{z}; \mathbf{p})|\,.$$

It follows from Lemma 6 that $A^*(z_1, \cdots, z_r)$ is convex on $\mathbb{R}^p \times (\mathbb{R}_+^1 - N^{\frac{1}{2}}\sigma^2)$, the domain of $N^{\frac{1}{2}}\boldsymbol{\xi}_N^*$, and thus Lemma 4 applies to the rate of convergence of the measure $P_N^*(\cdot)$ to $Q_{p+1}^*(\cdot)$ over convex sets. For each $\mathbf{z} \in \mathbb{R}_+^r$ we clearly have

$$|P_N^*(A^*(\mathbf{z})) - Q_{p+1}^*(A^*(\mathbf{z}))| \leqq \sup_{C \in \mathscr{C}^{p+1}} |P_N^*(C) - Q_{p+1}^*(C)|$$

and the expression on the right in turn is bounded as in Lemma 4.

REMARK. The first expression on the right of conclusion (ii) can be given the alternative form of conclusion (iii), Theorem 4.1. In both instances we use the work of Bergström [3] outlined earlier.

Our final undertaking is to study the difference

$$|H_N(z) - \Psi_p(z; \lambda_N)|$$

along lines suggested by the referee. We consider the central case for a single statistic only, as more general cases appear to require different methods. Assuming first that $\tilde{\boldsymbol{\beta}}$ and $\tilde{\sigma}^2$ are uncorrelated ($\mathbf{c} = \mathbf{0}$ at (4.5)), we have

THEOREM 4.3. *Let $\tilde{\boldsymbol{\beta}}$ and $\tilde{\sigma}^2$ be uncorrelated. Then*

$$|H_N(z) - \Psi_p(z)| \leqq c^+(p)/N^{\frac{1}{2}}$$

uniformly over $z \in R_+^1$, where $c^+(p)$ is a finite positive constant depending only on p.

PROOF. When $\mathbf{c} = \mathbf{0}$, the conditional distribution of

$$N(\hat{\boldsymbol{\beta}}_N - \boldsymbol{\beta})'\mathbf{T}^{-1}(\hat{\boldsymbol{\beta}}_N - \boldsymbol{\beta})/\sigma^2 ,$$

given $\hat{\sigma}_N^2$, is $\Psi_p(\cdot)$ under the probability law $Q_{p+1}^*(\cdot)$. Upon writing $\psi_p(x) = x^{\frac{1}{2}p-1}e^{-x}$ and

$$\Psi_p(z) = c_0(p) \int_0^{z/2} \psi_p(x)\, dx ,$$

we apply conditional arguments to obtain

$$H_N(z) = c_1(p) \int_{-N^{\frac{1}{2}}}^{\infty} e^{-\frac{1}{2}u^2}\Big[\int_0^{z(1+N^{-\frac{1}{2}}u)/2} \psi_p(x)\, dx\Big]\, du$$

and thus

$$|H_N(z) - \Psi_p(z)| \leqq c_1(p) \int_{-N^{\frac{1}{2}}}^{\infty} e^{-\frac{1}{2}u^2}\Big|\int_{z/2}^{z(1+N^{-\frac{1}{2}}u)/2} \psi_p(x)\, dx\Big|\, du + (2\pi)^{-\frac{1}{2}} \int_{-\infty}^{-N^{\frac{1}{2}}} e^{-\frac{1}{2}u^2}\, du .$$

Now partitioning $[-N^{\frac{1}{2}}, \infty) = [-N^{\frac{1}{2}}, 0) \cup [0] \cup (0, N^{\frac{1}{2}}] \cup (N^{\frac{1}{2}}, \infty)$ and taking signs into account in each interval, we find that

$$\int_{-N^{\frac{1}{2}}}^{0} e^{-\frac{1}{2}u^2}\Big[\int_{z(1+N^{-\frac{1}{2}}u)/2}^{z/2} \psi_p(x)\, dx\Big]\, du = \int_0^{N^{\frac{1}{2}}} e^{-\frac{1}{2}u^2}\Big[\int_{z(1-N^{-\frac{1}{2}}u)/2}^{z/2} \psi_p(x)\, dx\Big]\, du$$

from which it follows that

$$|H_N(z) - \Psi_p(z)| \leqq g_1(z) + g_2(z) + (2\pi)^{-\frac{1}{2}} \int_{-\infty}^{-N^{\frac{1}{2}}} e^{-\frac{1}{2}u^2}\, du \tag{4.22}$$

where

$$g_1(z) = c_1(p) \int_0^{N^{\frac{1}{2}}} e^{-\frac{1}{2}u^2}\Big[\int_{z(1-N^{-\frac{1}{2}}u)/2}^{z(1+N^{-\frac{1}{2}}u)/2} \psi_p(x)\, dx\Big]\, du$$

and

$$g_2(z) = c_1(p) \int_{N^{\frac{1}{2}}}^{\infty} e^{-\frac{1}{2}u^2}\Big[\int_{z/2}^{z(1+N^{-\frac{1}{2}}u)/2} \psi_p(x)\, dx\Big]\, du .$$

Clearly the latter expression satisfies

$$g_2(z) \leqq (2\pi)^{-\frac{1}{2}} \int_{N^{\frac{1}{2}}}^{\infty} e^{-\frac{1}{2}u^2}\, du$$

and upon taking $u \to -u$ we combine terms to obtain

$$|H_N(z) - \Psi_p(z)| \leqq g_1(z) + 2(2\pi)^{-\frac{1}{2}} \int_{-\infty}^{-N^{\frac{1}{2}}} e^{-\frac{1}{2}u^2}\, du . \tag{4.23}$$

Now considering z fixed and making the change of variables $y = x/z$, we have

$$g_1(z) = c_1(p) \int_0^{N^{\frac{1}{2}}} e^{-\frac{1}{2}u^2}\Big[\int_{(1-N^{-\frac{1}{2}}u)/2}^{(1+N^{-\frac{1}{2}}u)/2} (yz)^{\frac{1}{2}p}e^{-yz}y^{-1}\, dy\Big]\, du \leqq g_3(N)$$

where

$$g_3(N) = c_2(p) \int_0^{N^{\frac{1}{2}}} e^{-\frac{1}{2}u^2}\Big[\int_{(1-N^{-\frac{1}{2}}u)/2}^{(1+N^{-\frac{1}{2}}u)/2} y^{-1}\, dy\Big]\, du$$

and the inequality stems from the fact that $(yz)^{\frac{1}{2}p}e^{-yz}$ is bounded when $p \geqq 1$. Now letting $w = N^{-\frac{1}{2}}u$ in $g_3(N)$ and evaluating the inner integral, then letting

$w = -N^{-\frac{1}{2}}u$ in the final term of expression (4.23), we infer that

$$(4.24) \qquad |H_N(z) - \Psi_p(z)| \leq \frac{c_2(p)}{N^{\frac{1}{2}}} \left\{ \int_0^1 Ne^{-\frac{1}{2}Nw^2} \log\left[\frac{(1+w)}{(1-w)}\right] dw + c_3(p) \int_1^\infty Ne^{-\frac{1}{2}Nw^2}\, dw \right\} .$$

Our proof is complete if we can demonstrate that each integral is bounded above uniformly in N, i.e. if positive constants γ_1, γ_2 and γ_3 can be found, not depending on N, such that

$$\text{(i)} \qquad \int_0^1 Ne^{-\frac{1}{2}Nw^2} \log(1+w)\, dw \leq \gamma_1$$

$$\text{(ii)} \qquad \int_0^1 Ne^{-\frac{1}{2}Nw^2}[-\log(1-w)]\, dw \leq \gamma_2$$

$$\text{(iii)} \qquad \int_1^\infty Ne^{-\frac{1}{2}Nw^2}\, dw \leq \gamma_3 .$$

The conclusion (iii) is immediate; to demonstrate (i) we use the bound $\log(1+w) \leq w$ together with the change of variables $v = Nw^2/2$ to obtain

$$\begin{aligned} \int_0^1 Ne^{-\frac{1}{2}Nw^2} \log(1+w)\, dw &\leq \int_0^1 Nwe^{-\frac{1}{2}Nw^2}\, dw \\ &\leq \int_0^{N/2} e^{-v}\, dv = 1 - e^{-\frac{1}{2}N} \\ &\leq 1 . \end{aligned}$$

To demonstrate (ii), let $\varepsilon \in (0, 1)$ be small enough to satisfy

(a) if $1 - \varepsilon < x \leq 1$, then $-\log(1-x) \leq (1-x)^{-\frac{1}{2}}$;
(b) if $0 \leq x < \varepsilon$, then $-\log(1-x) \leq x$;

and consider the integral (ii) in three parts. Using (a), then letting $y = 1 - w$, we have

$$\begin{aligned} \int_{1-\varepsilon}^1 Ne^{-\frac{1}{2}Nw^2}[-\log(1-w)]\, dw &\leq \int_{1-\varepsilon}^1 Ne^{-\frac{1}{2}Nw^2}(1-w)^{\frac{1}{2}}\, dw \\ &= Ne^{-\frac{1}{2}N} \int_0^\varepsilon e^{-\frac{1}{2}N(y^2-2y)} y^{-\frac{1}{2}}\, dy \leq Ne^{-\frac{1}{2}N} \int_0^\varepsilon e^{Ny} y^{-\frac{1}{2}}\, dy . \end{aligned}$$

Let $t = Ny$; then

$$\begin{aligned} Ne^{-\frac{1}{2}N} \int_0^\varepsilon e^{Ny} y^{-\frac{1}{2}}\, dy &= N^{\frac{1}{2}} e^{-\frac{1}{2}N} \int_0^{N\varepsilon} e^t t^{-\frac{1}{2}}\, dt \\ &\leq N^{\frac{1}{2}} e^{-\frac{1}{2}N} \left[\int_0^1 e^t t^{-\frac{1}{2}}\, dt + \int_1^{N\varepsilon} e^t\, dt\right] = K_N , \end{aligned}$$

say. Since K_N is finite for each N and $K_N \to 0$ as $N \to \infty$, there is a constant γ_{21} such that $K_N \leq \gamma_{21}$ for all N.

Similarly we have

$$\int_\varepsilon^{1-\varepsilon} Ne^{-\frac{1}{2}Nw^2}[-\log(1-w)]\, dw \leq Ne^{-\frac{1}{2}N\varepsilon^2}[-\log \varepsilon] = L_N ,$$

say, where L_N is finite for each N and $L_N \to 0$ as $N \to \infty$ assures a constant γ_{22} such that $L_N \leq \gamma_{22}$ for all N. Finally we have

$$\begin{aligned} \int_0^\varepsilon Ne^{-\frac{1}{2}Nw^2}[-\log(1-w)]\, dw &\leq \int_0^\varepsilon Ne^{-\frac{1}{2}Nw^2} w\, dw \\ &= \int_0^{\varepsilon^2/2} Ne^{-Nu}\, du = 1 - e^{-N\varepsilon^2/2} \leq 1 \end{aligned}$$

where we have used $u = w^2/2$ and condition (b). Now combining the results for case (ii), using $\gamma_2 = \gamma_{21} + \gamma_{22} + 1$, we have

$$|H_N(z) - \Psi_p(z)| \leqq \frac{c_2(p)}{N^{\frac{1}{2}}}(\gamma_1 + \gamma_2 + \gamma_3)$$

and our proof is complete.

Using simpler arguments of the type of Slutsky, the referee has removed the requirement that $\mathbf{c} = \mathbf{0}$ and has supplied the following bounds.

THEOREM 4.4. *For each* $N = 1, 2, \cdots$ *we have*

$$\sup_{z \in R_+^1} |H_N(z) - \Psi_p(z)| \leqq c(p, \omega) \log^{\frac{1}{2}} N/N^{\frac{1}{2}} \tag{4.25}$$

where $\omega^2 = \mathscr{E}(\tilde{\sigma}^2 - \sigma^2)^2$ *and* $c(p, \omega)$ *is a finite positive constant depending only on* p *and* ω.

PROOF. Let $\Pr(\cdot)$ be the measure appropriate in context, and write

$$H_N(z) = \Pr(\mathbf{Y}'\mathbf{T}^{-1}\mathbf{Y}/X \leqq z\sigma^2)$$
$$\Psi_p(z) = \Pr(\mathbf{Y}'\mathbf{T}^{-1}\mathbf{Y} \leqq z\sigma^2)$$

where $[Y_1, \cdots, Y_p, X]$ is a Gaussian vector having the means $[0, \cdots, 0, 1]$ and the dispersion matrix

$$\mathbf{\Sigma}^* = \begin{bmatrix} \sigma^2\mathbf{T} & N^{-\frac{1}{2}}\boldsymbol{\rho} \\ N^{-\frac{1}{2}}\boldsymbol{\rho}' & N^{-1}\kappa^2 \end{bmatrix}$$

where $\boldsymbol{\rho} = \mathbf{c}/\sigma^2$ and $\kappa^2 = \omega^2/\sigma^4$. For all $\varepsilon > 0$ it follows that

$$\begin{aligned} &\Psi_p(z(1-\varepsilon)) - \Pr(X < 1-\varepsilon) \\ &\qquad \leqq H_N(z) \leqq \Psi_p(z(1+\varepsilon)) + \Pr(X \leqq 0) + \Pr(X > 1+\varepsilon)\,. \end{aligned} \tag{4.26}$$

Now taking $\varepsilon = (2^{\frac{1}{2}}\omega/\sigma^2)\log^{\frac{1}{2}} N/N^{\frac{1}{2}}$ we have, for $p \geqq 2$,

$$\begin{aligned} \Psi_p(z(1+\varepsilon)) - \Psi_p(z) &\leqq c_0(p)e^{-z\sigma^2}[z\sigma^2(1+\varepsilon)]^{\frac{1}{2}p-1}z\sigma^2\varepsilon \\ &\leqq c_3(p)\varepsilon \end{aligned} \tag{4.27}$$

uniformly for $z \geqq 0$. In the case $p = 1$ we compute

$$\Psi_1(z(1+\varepsilon)) - \Psi_1(z) \leqq c_0(1)e^{-z\sigma^2}(z\sigma^2)^{-\frac{1}{2}}z\sigma^2\varepsilon \leqq c_3(1)\varepsilon \tag{4.28}$$

uniformly for $z > 0$. Moreover, we have

$$\begin{aligned} \Pr(|X-1| > \varepsilon) &= \Pr(|Z| > \varepsilon\sigma^2 N^{\frac{1}{2}}/\omega) \\ &= \Pr(|Z| > (2\log N)^{\frac{1}{2}}) < (2/\pi)^{\frac{1}{2}}(2\log N)^{-\frac{1}{2}}/N\,. \end{aligned} \tag{4.29}$$

The estimate (4.25) now follows from (4.26)—(4.29), and the proof is complete.

REMARKS. In the special case that $\mathbf{c} = \mathbf{0}$ the bounds are of smaller order than those otherwise available at present; compare $O(N^{-\frac{1}{2}})$ with $O(N^{-\frac{1}{2}}\log^{\frac{1}{2}} N)$ in Theorems 4.3 and 4.4, respectively. The condition $\mathbf{c} = \mathbf{0}$ does not follow from the fact that $\tilde{\boldsymbol{\beta}}$ and the residual errors are uncorrelated; nor is the stronger assumption that $\tilde{\boldsymbol{\beta}}$ and $\tilde{\sigma}^2$ be independent a tenable one, for this is tantamount to assuming a Gaussian model at the outset. However, as pointed out to the authors by Professor Bickel, it does follow that $\mathbf{c} = \mathbf{0}$ if $\mathbf{y} - \mathbf{X}\boldsymbol{\beta}$ is distributed symmetrically about $\mathbf{0}$, i.e. if $(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})$ and $-(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})$ have the same distribution.

Acknowledgment. The authors are grateful to Professor Peter Bickel and a referee for providing so generously their insights into this study. The present proof of Lemma 5 is due to Professor Bickel, and Theorem 4.4 and its proof are due entirely to the referee.

REFERENCES

[1] BERGE, C. (1963). *Topological Spaces.* MacMillan, New York.

[2] BERGSTRÖM, H. (1945). On the central limit theorem in the space R_k, $k > 1$. *Skand. Aktuarietidskr.* **28** 106–127.

[3] BERGSTRÖM, H. (1969). On the central limit theorem in R^k. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **14** 113–126.

[4] BHATTACHARYA, R. N. (1968). Berry-Esséen bounds for the multidimensional central limit theorem. *Bull. Amer. Math. Soc.* **74** 285–287.

[5] BILLINGSLEY, P. (1968). *Convergence of Probability Measures.* Wiley, New York.

[6] ESSÉEN, C. G. (1945). Fourier analysis of distribution functions. *Acta Math.* **77** 3–125.

[7] JENSEN, D. R. (1970). The joint distribution of traces of Wishart matrices and some applications. *Ann. Math. Statist.* **41** 133–145. Correction. *Ibid.* **41** 2186.

[8] SAZONOV, V. V. (1968). On the speed of convergence in the multidimensional central limit theorem. *Theor. Probability Appl.* **13** 188–191.

[9] SAZONOV, V. V. (1968). On the multi-dimensional central limit theorem. *Sankhyā Ser. A* **30** 181–204.

[10] VARADARAJAN, V. S. (1958). A useful convergence theorem. *Sankhyā Ser. A* **20** 221–222.

DEPARTMENT OF STATISTICS
VIRGINIA POLYTECHNIC INSTITUTE
BLACKSBURG, VIRGINIA 24061