Noncentral Generalized F Distributions with Applications to Joint Outlier Detection

Donald E. Ramirez
Department of Mathematics
University of Virginia
Charlottesville, VA 22904-4137
der@virginia.edu

Abstract

An algorithm for computing the noncentral generalized F distribution is discussed. This distribution is required to compute the power of tests of multiple outliers in linear models based on shifts of location and scale. Applications to joint outlier detection are given.

1 Special Distributions

Designate by $\mathcal{L}(\boldsymbol{Y})$ the distribution of $\boldsymbol{Y} \in \Re^n$, and by *pdf* and *cdf* its probability density and cumulative distribution functions respectively. In particular, $N_n(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ is the Gaussian distribution on $\Re^n$ having $E(\boldsymbol{Y}) = \boldsymbol{\mu}$, dispersion matrix $V(\boldsymbol{Y}) = \boldsymbol{\Sigma}$, and variance $Var(Y) = \sigma^2$ on $\Re^1$. The standard distributions on $\Re^1$ and $\Re^1_+$ are $t(\nu, \lambda)$, $\chi^2(\nu, \lambda)$, and $F(\nu_1, \nu_2, \lambda)$ - the noncentral Student's, chi-squared and Fisher distributions, having degrees of freedom $\{\nu, (\nu_1, \nu_2)\}$ and noncentrality parameter λ. The *cdf*'s corresponding to the chi-squared and Fisher distributions are denoted by $G(\cdot; \nu, \lambda)$ and $F(\cdot; \nu_1, \nu_2, \lambda)$, whereas $t(\nu)$, $\chi^2(\nu)$, and $F(\nu_1, \nu_2)$ will denote the central distributions.

Let V be a central chi-squared random variable with degrees of freedom ν having the *cdf* $G(\cdot, v)$. Let $\{U_1, \ldots, U_s\}$ be mutually independent $\{N_1(\omega_i, 1); 1 \leq i \leq s\}$ noncentral normal random scalars, assumed to be independent of V; consider $\boldsymbol{\alpha}' = [\alpha_1, \ldots, \alpha_s]$ as fixed positive weights; let $\boldsymbol{\omega}' = [\omega_1, \ldots, \omega_s]$; and let

$$W = \frac{(\alpha_1 U_1^2 + \cdots + \alpha_s U_s^2)/s}{V/\nu}, \tag{1}$$

having the *cdf* $F_s(w; \alpha_1, \ldots, \alpha_s; \omega_1, \ldots, \omega_s; \nu) = F_s(w; \boldsymbol{\alpha}'; \boldsymbol{\omega}'; \nu)$. Series expansions for the latter and its density, and bounds for errors accrued on truncating the series, are found in Ramirez and Jensen (1991) for the central case where $\boldsymbol{\omega} = \mathbf{0}$.

This paper discusses the corresponding expansions and error bounds for the noncentral distributions with $\boldsymbol{\omega} \neq \mathbf{0}$. The case $\alpha_1 = \ldots = \alpha_s = \alpha$ gives $F_s(w; \alpha, \ldots, \alpha; \omega_1, \ldots, \omega_s; \nu) = F(w/\alpha; s, \nu, \lambda)$ with $\lambda = \boldsymbol{\omega}'\boldsymbol{\omega}$, the noncentral Fisher F distribution. The case $\alpha_1 = \ldots = \alpha_s = 1$ with $\lambda = \boldsymbol{\omega}'\boldsymbol{\omega} = \mathbf{0}$ gives the central Fisher F distribution.

2 Generalized *F* Distribution

Let $T = \alpha_1 U_1^2 + \cdots + \alpha_s U_s^2$, with weights satisfying $\{\alpha_s \geq \alpha_{s-1} \geq \cdots \geq \alpha_1 > 0\}$. Note that α_1 is the minimum of the set $\{\alpha_s \geq \alpha_{s-1} \geq \cdots \geq \alpha_1 > 0\}$. For the

central case with $\boldsymbol{\omega} = \mathbf{0}$, the *pdf* for $F_s(w; \alpha_1, \ldots, \alpha_s; 0, \cdots, 0; \nu)$, expanded as a weighted series of central F *pdf*s, is found in Ramirez and Jensen (1991), together with bounds on truncation errors. The series is based on expanding $\mathcal{L}(T)$ as a mixture of central chi-squared *pdf*s, as given in Ruben (1962) and Kotz, Johnson, and Boyd (1967a). The Fortran code to evaluate $F_s(w; \alpha_1, \ldots, \alpha_s; 0, \cdots, 0; \nu)$ is given in Ramirez (2000). This Fortran code requires the IMSL Fortran library. To make the algorithm discussed in this paper more accessible, this paper will present the companion code in Appendix 1 for the noncentral generalized F distribution using Maple.

For the general case with $\boldsymbol{\omega} \neq \mathbf{0}$, we may expand $\mathcal{L}(T)$ as a mixture in terms of either noncentral or central chi-squared *pdf*s as in Ruben (1962) and Kotz, Johnson, and Boyd (1967b). We will use the expansion in terms of central chi-squared *pdf*s since these distributions are available in Maple. To these ends recursively define the sequences

$$\begin{aligned} c_0 &= e^{-\frac{1}{2}\lambda} \prod_{i=1}^{s} (\beta/\alpha_i)^{\frac{1}{2}}, \\ c_j &= \frac{1}{2j} \sum_{i=0}^{j-1} d_{j-i} c_i, \ j = 1, 2, \cdots, \\ d_k &= \sum_{i=1}^{s} (1 - \beta/\alpha_i)^k + k\beta \sum_{i=1}^{s} \frac{\omega_i^2}{\alpha_i} (1 - \beta/\alpha_i)^{k-1}, \ k = 1, 2, \cdots. \end{aligned} \tag{2}$$

Here β satisfies $0 < \beta < \alpha_1$; this assures that $\{0 < 1 - \beta/\alpha_i < 1; 1 \leq i \leq s\}$ and that $\sum_{i=0}^{\infty} c_i = 1$ with $\{c_i > 0; \ i = 0, 1, \cdots\}$. In a manner similar to arguments supporting Theorem 3.1 of Ramirez and Jensen (1991), we can establish the following series expansions for the *pdf* for $F_s(w; \alpha_1, \ldots, \alpha_s; \omega_1, \ldots, \omega_s; \nu)$ generally as follows.

Theorem 1 *With the foregoing notation, the pdf for* $F_s(w; \alpha_1, \ldots, \alpha_s; \omega_1, \ldots, \omega_s; \nu)$ *has the representation in terms of gamma functions as*

$$h_F(w) = \sum_{i=0}^{\infty} \frac{c_i}{\beta} \frac{s}{\nu} \frac{\Gamma[(s + 2i + \nu)/2]}{\Gamma[(s + 2i)/2]\Gamma(\nu/2)} \frac{(\frac{s}{\nu}\frac{w}{\beta})^{(s+2i-2)/2}}{(1 + \frac{s}{\nu}\frac{w}{\beta})^{(s+2i+\nu)/2}}, \tag{3}$$

and in terms of pdfs $f_F(\cdot; \nu, \gamma)$ *for central* F *distributions as*

$$h_F(w) = \sum_{i=0}^{\infty} \frac{c_i}{\beta} \frac{s}{s + 2i} f_F(\frac{s}{s + 2i} \frac{w}{\beta}; s + 2i, \nu). \tag{4}$$

A global error bound for truncating at the τ*th partial sum of the pdf for* W *from (3) and (4) is given by*

$$\begin{aligned} e_\tau(w) &= \sum_{i=\tau+1}^{\infty} \frac{c_i}{\beta} \frac{s}{s + 2i} f_F(\frac{s}{s + 2i} \frac{w}{\beta}; s + 2i, \nu) \\ &\leq \frac{s}{\beta[s + 2(\tau + 1)]} [1 - (c_0 + \cdots + c_\tau)] = e_\tau. \end{aligned} \tag{5}$$

A global error bound for truncating at the τ*th partial sum of the cdf for* W *is also given by* e_τ.

To compute the *cdf* of W, the series in Equation (3) is truncated and then integrated numerically. The probabilities to be reported subsequently were computed using this procedure, where the error tolerance was set to be 10^{-5} and τ was increased until e_τ was less than the prescribed error tolerance.

Appendix 1 shows the Maple code with

$$\begin{aligned} s &= 2 \\ \nu &= 15 \\ \alpha &= [1.6335, 1.9300]' \\ \omega &= [1.2757, 1.3537]' \\ y_0 &= 3.6823. \end{aligned} \tag{6}$$

These values are from the outlier example that appears in the applications section. The value chosen for y_0 corresponds to the 95-percentile value for the central F distribution. The program uses the error tolerance $e_\tau = 10^{-5}$. The output includes the number of terms N required to achieve this error tolerance and the p-value $\Pr[W > y_0]$ as follows

$$\begin{aligned} N &= 10 \\ \Pr[W &> y_0] = 0.5518. \end{aligned}$$

A plot of the corresponding noncentral generalized F distribution also is shown. In the applications section, we will interpret the p-value as the power of a test for misspecification of a linear model due to shifts in location and scale.

3 Misspecified Linear Models

To fix ideas, consider the model

$$\{Y_i = \beta_0 + \beta_1 X_{i1} + \cdots + \beta_k X_{ik} + \varepsilon_i; 1 \le i \le n\} \tag{7}$$

relating a response Y_i to k nonrandom regressors $\{X_{i1}, X_{i2}, \ldots, X_{ik}\}$ through $p = k + 1$ unknown parameters $\boldsymbol{\beta} = [\beta_0, \beta_1, \ldots, \beta_k]'$. Written as $\boldsymbol{Y}_0 = \boldsymbol{X}_0\boldsymbol{\beta} + \boldsymbol{\varepsilon}_0$, each row $\boldsymbol{x}_i'$ of $\boldsymbol{X}_0$ is a point in the design space. The experimental design matrix $\boldsymbol{X}_0$ may be assumed to have full rank p. The reduced model $\boldsymbol{Y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$ follows on deleting from $\boldsymbol{X}_0(n \times p)$, s rows to be indexed by $I = \{i_1, \ldots, i_s\} \subset \{1, 2, \ldots, n\}$, to comprise the matrix $\boldsymbol{Z}(s \times p)$, and deleting corresponding elements from $\boldsymbol{Y}_0$. Here the matrices $\boldsymbol{X}_0$, $\boldsymbol{X}$, and $\boldsymbol{Z}$ are fixed nonrandom matrices. Then $(\widehat{\boldsymbol{\beta}}, S^2)$ and $(\widehat{\boldsymbol{\beta}}_I, S_I^2)$ are Gauss-Markov estimators and residual mean squares from the full and reduced data.

The *leverage* of an individual design point $\boldsymbol{x}_i'$ of the fixed design $\boldsymbol{X}_0$ is the element (h_{ii}) on the diagonal of the matrix $\boldsymbol{H} = \boldsymbol{X}_0(\boldsymbol{X}_0'\boldsymbol{X}_0)^{-1}\boldsymbol{X}_0'$. For subset deletions, the natural leverages emerge on letting $\boldsymbol{H}_{II} = \boldsymbol{Z}(\boldsymbol{X}_0'\boldsymbol{X}_0)^{-1}\boldsymbol{Z}'$. Apply the spectral decomposition $\boldsymbol{Q}'\boldsymbol{H}_{II}\boldsymbol{Q} = \boldsymbol{D}_\lambda = Diag(\lambda_1, \ldots, \lambda_s)$. We will denote

$$\{\lambda_1 \ge \ldots \ge \lambda_s > 0\} \tag{8}$$

as *canonical leverages* following the notation of Jensen (2001). To assess the joint influence of points in $\boldsymbol{Z}$, and to identify outliers at those points, consider the row-partitioned forms $\boldsymbol{Y}_0' = [\boldsymbol{Y}', \boldsymbol{Y}_I']$, $\boldsymbol{X}_0' = [\boldsymbol{X}', \boldsymbol{Z}']$, and $\boldsymbol{\varepsilon}_0' = [\boldsymbol{\varepsilon}', \boldsymbol{\varepsilon}_I']$. The full data give $(\widehat{\boldsymbol{\beta}}, S^2)$ as before, as well as $\widehat{\boldsymbol{Y}}_0 = \boldsymbol{X}_0\widehat{\boldsymbol{\beta}}$, and $\boldsymbol{e}_0 = (\boldsymbol{Y}_0 - \widehat{\boldsymbol{Y}}_0)$, to be partitioned

as $\boldsymbol{e}_0' = [\boldsymbol{e}', \boldsymbol{e}_I']$, where $S^2 = \boldsymbol{e}_0'\boldsymbol{e}_0/(n-p)$. Corresponding values from the reduced data are $(\widehat{\boldsymbol{\beta}}_I, S_I^2)$ as before, and $\widehat{\boldsymbol{Y}}_{I(I)} = \boldsymbol{Z}\widehat{\boldsymbol{\beta}}_I$, where

$$S_I^2 = \frac{(\boldsymbol{Y} - \boldsymbol{X}\widehat{\boldsymbol{\beta}}_I)'(\boldsymbol{Y} - \boldsymbol{X}\widehat{\boldsymbol{\beta}}_I)}{n-p-s}. \tag{9}$$

The connection between S_I^2 and S^2 is given by $S_I^2/S^2 = (n-p)/[sF_I + (n-p-s)]$, with F_I as the R-Fisher statistic to be defined in the next section.

To consider misspecified scale, we use the notation $\boldsymbol{\Xi}(\sigma^2, \sigma_1^2) = Diag(\sigma^2\boldsymbol{I}_r, \sigma_1^2\boldsymbol{I}_s)$, with $r+s=n$. Assumption A1 below allows for misspecification due to a shift δ in location, and A2 allows for misspecification due to a shift in scale at the design points in Z.

Assumptions A.

A1. $E(\boldsymbol{\varepsilon}) = \boldsymbol{0} \in \Re^r$ *and* $E(\boldsymbol{\varepsilon}_I) = \boldsymbol{\delta} \in \Re^s$;

A2. $V(\boldsymbol{\varepsilon}_0) = \boldsymbol{\Xi}(\sigma^2, \sigma_1^2)$; *and*

A3. $\mathcal{L}(\boldsymbol{\varepsilon}, \boldsymbol{\varepsilon}_I - \boldsymbol{\delta}) = N_n(\boldsymbol{0}, \boldsymbol{\Xi}(\sigma^2, \sigma_1^2))$.

4 *R*-Fisher Statistic F_I

The R-Student ratio $t_i = e_i/S_i\sqrt{1-h_{ii}}$ is pivotal in single case deletion diagnostics in testing for a mean shift of δ units in Y_i at design point $\boldsymbol{x}_i'$. This statistics is often called the studentized deleted residual or the externally studentized residual. For subsets, we use the generalization of the R-Student statistic, which we call the R-Fisher statistic, with

$$F_I = \frac{\boldsymbol{e}_I'(\boldsymbol{I}_s - \boldsymbol{H}_{II})^{-1}\boldsymbol{e}_I}{S_I^2}. \tag{10}$$

This statistic will be used for our tests for a mean shift of $\boldsymbol{\delta}$ units in $\boldsymbol{Y}_I$ at design points in $\boldsymbol{Z}$. Under standard assumptions, where $\sigma_1^2 = \sigma^2$ in *A2* and *A3*, $\mathcal{L}(F_I) = F(s, n-p-s, \lambda)$ the noncentral F distribution with noncentrality parameter $\lambda = \boldsymbol{\delta}'(\boldsymbol{I}_s - \boldsymbol{H}_{II})\boldsymbol{\delta}/\sigma^2$, yielding exact α-level tests for $H_0 : \boldsymbol{\delta} = \boldsymbol{0}$ against $H_1 : \boldsymbol{\delta} \neq \boldsymbol{0}$.

Following the notation of Jensen and Ramirez (2001), we classify as Group I statistics the single case $s=1$ statistics: R-Student statistic t_i, and influence diagnostics DFT_i and DFB_{ij}, known also as $DIFFITs$ and $DFBETAs$.

Group II outlier diagnostics are the multiple case $s > 1$ statistics: R-Fisher statistic F_I and OUT_I of Barnett and Lewis (1984); subset influence diagnostics include AP_I of Andrews and Pregibon (1978), CR_I ($COVRATIO_I$) and FV_I ($FVARATIO_I$) of Belsley *et al.* (1980). Corresponding diagnostics $\{t_i^2, OUT_i, AP_i, CR_i, FV_i\}$ refer to the case $s=1$.

Group III influence diagnostics encompass multiple case $s > 1$ statistics: $\{C_I, WK_I, W_I, D_I\}$ as in Cook (1977), Welsch and Kuh (1977), Welsch (1982), and Jensen and Ramirez (1998), for gauging disturbances in the vector $\widehat{\boldsymbol{\beta}}$. Corresponding diagnostics $\{C_i, WK_i, W_i, D_i\}$ refer to the case $s=1$. The forms for D_i and D_I use any reflexive g-inverse $\boldsymbol{V}^-$ of $\boldsymbol{V} = V(\widehat{\boldsymbol{\beta}} - \widehat{\boldsymbol{\beta}}_I)$ as

$$D_I = \frac{(\widehat{\beta} - \widehat{\beta}_I)'\boldsymbol{V}^-(\widehat{\beta} - \widehat{\beta}_I)/s}{S_I^2}. \tag{11}$$

It is now known that all single case deletion diagnostics of Groups 1, 2, and 3 are equivalent; each represents either a scaling of the R-Student t_i or corresponds one-to-one with its square. These facts were established in Jensen (1998), with an

abridged version appearing as Jensen (2000). Each diagnostic thus supports a test equivalent to testing for a mean-shift outlier at $\boldsymbol{x}_i'$ using t_i. Similar conclusions were reported for $\{DFT_i, DFB_{ij}, CR_i, C_i\}$ by LaMotte (1999) using different methods.

The multiple case diagnostics $\{F_I, OUT_I, AP_I, CR_I, FV_I, D_I\}$ from Groups II and III also yield equivalent tests for outliers. Each diagnostic thus supports a test equivalent to testing for a mean-shifted vector outlier at design points in $\boldsymbol{Z}$ using the R-Fisher statistic F_I. For further details and proofs see Jensen (2001).

5 Applications of the Noncentral Generalized F

In what follows we identify $\kappa = \sigma_1^2/\sigma^2$ as the misspecified shift in scale with $\boldsymbol{\Xi}(\kappa)$ $= Diag(\boldsymbol{I}_r, \kappa \boldsymbol{I}_s)$, in which case $V(\boldsymbol{\varepsilon}_0) = \boldsymbol{\Xi}(\sigma^2, \sigma_1^2) = \sigma^2 \boldsymbol{\Xi}(\kappa)$ as in *Assumptions A2* and *A3*. The misspecified shift is location will be $\delta \neq 0$. The following is the principal result from Jensen and Ramirez (2001).

Theorem 2 *Consider the R-Fisher diagnostic F_I under Assumptions A; let $\{\lambda_1 \geq \ldots \geq \lambda_s > 0\}$ comprise the canonical subset leverages; set $\kappa = \sigma_1^2/\sigma^2 \geq 1$; let $\boldsymbol{\theta} = \boldsymbol{Q}'\boldsymbol{\delta}$, such that $\boldsymbol{Q}'\boldsymbol{H}_{II}\boldsymbol{Q} = Diag(\lambda_1, \ldots, \lambda_s)$; and identify $\nu = n - p - s$. The cdf of F_I is given by $F_s(w; \boldsymbol{\alpha}'; \boldsymbol{\omega}'; \gamma)$, with weights*

$$\alpha_i = \kappa - (\kappa - 1)\lambda_i;\ 1 \leq i \leq s \tag{12}$$

satisfying $\{\alpha_s \geq \ldots \geq \alpha_1 \geq 1\}$, and with location parameters

$$\omega_i = \frac{\theta_i}{\sigma}[\kappa + \lambda_i/(1 - \lambda_i)]^{1/2}; 1 \leq i \leq s. \tag{13}$$

Appendix 2 will give the Maple code to implement Theorem 2.

6 The Data Set

Data regarding the administration of Bachelor Officers Quarters *(BOQ)* were reported for sites at $n = 25$ naval installations. Monthly man-hours (Y) were related linearly to average daily occupancy (X_1), monthly number of check-ins (X_2), weekly service desk operation in hours (X_3), size of common use area (X_4), number of building wings (X_5), operational berthing capacity (X_6), and number of rooms (X_7). The data are reported in Myers (1990), p. 218 ff, together with detailed analyses using single case deletion diagnostics. Subset deletion diagnostics are not reported there.

In Jensen and Ramirez (2002), we determined the subset influence for selected pairs of observations $(I = \{i, j\})$ in the BOQ data arising through shifts in location only. We focused on sites $\{15, 20, 21, 23, 24\}$, having individual leverages $\{0.5576,$ $0.3663, 0.0704, 0.9885, 0.8762\}$. For our example on computing power using F_I, we select the subset $S_1 = \{20, 21\}$ so $s = 2$. Appendix 2 shows how to compute the power of this test for outliers using F_I at the 0.05 level against a misspecification with $\delta = 2$ due to a location shift of $\delta\sigma = 2\sigma$ units with $\delta = 2\ \mathbf{1}_s = (2, 2)'$, and additionally against a misspecification due to a scale shift with ratio $\kappa = \sigma_1^2/\sigma^2 = 2$.

For this data set, the canonical leverages (the eigenvalues of $\mathbf{H}_{II}$) are computed to be

$$\{\lambda_1, \lambda_2\} = \{0.3665, 0.0702\}.$$

(The canonical leverages should be close in value to the corresponding leverages $0.3663, 0.0704$.)

Next the standardized shift in location is computed as

$$\mathbf{Q}'\mathbf{1}_s=[1.0243, 0.9751]'.$$

(Since $\mathbf{Q}$ is an orthogonal matrix, $1.0243^2 + 0.9751^2 = 2$.) We seek the power of the R-Fisher diagnostic F_I with scale shift $\kappa = \sigma_1^2/\sigma^2 = 2$ and with location shift

$$\theta = \mathbf{Q}'\delta = 2\mathbf{Q}'\mathbf{1}_s.$$

Using Theorem 2 and the Maple code from Appendix 2, we compute the associated weights and noncentrality parameters as

$$\begin{aligned} \alpha &= [1.6335, 1.9300]' \\ \omega &= [1.2757, 1.3537]'. \end{aligned}$$

These values are the values used in the first example from Equation 6 where we showed how to compute p-values for the noncentral generalized F distribution. To compute the power of the R-Fisher diagnostic F_I we are comparing the null hypothesis $H_0 : \kappa = 1$ and $\delta = \mathbf{0}$ against the alternative hypothesis $H_1 : \kappa > 1$ and $\delta \neq \mathbf{0}$. The power is

$$\Pr[W > 3.6823] = 0.5518,$$

where 3.6823 is the 95% critical value of the central F distribution with degrees of freedom $(s, n - p - s) = (2, 15)$. The values 3.6823 and 0.5518 are computed from the Maple code in Appendix 2.

7 Conclusions

This paper gives the computer codes using Maple for computing the noncentral generalized F distribution. Appendix 1 gives the code for computing the p-values for this distribution. The noncentral generalized F distributions have been used to compute the power for the R-Fisher outlier diagnostic F_I, and thus for all equivalent diagnostics, under shifts in location and shifts of scale at selected subsets of the BOQ data.

Consultants are urged to consider the use of the outlier diagnostic F_I for detecting subsets of outliers based on shifts of location and scale. Using the Maple codes contained in this paper, the p-values are easy to compute (Appendix 1) as well as the power of this test (Appendix 2).

References

[1] Andrews, D. F. and Pregibon, D. (1978). Finding outliers that matter. *J. Royal Statist. Soc. B* **40:** 85–93.

[2] Barnett, V. and Lewis, T. (1984). *Outliers in Statistical Data, Second ed.* John Wiley and Sons, New York.

[3] Belsley, D. A., Kuh, E. and Welsch, R. E. (1980). *Regression Diagnostics: Identifying Influential Data and Sources of Collinearity.* John Wiley and Sons, New York.

[4] Cook, R. D. (1977). Detection of influential observations in linear regression. *Technometrics* **19:** 15–18.

[5] Jensen, D. R. (1998). The use of standardized diagnostics in regression. *Research Report R079,* Department of Statistics, Virginia Polytechnic Institute and State University, October 1998.

[6] Jensen, D. R. (2000). The use of Studentized diagnostics in regression. *Metrika* **52:** 213–223.

[7] Jensen, D. R. (2001). Properties of selected subset diagnostics in regression. *Statist. Probab. Letters* **51:** 377–388.

[8] Jensen, D. R. and Ramirez, D. E. (1998). Some exact properties of Cook's D_I. Chapter 14 in *Handbook of Statistics, Volume 16, Order Statistics: Theory and Methods,* N. Balakrishnan and C. R. Rao, eds., pp. 387–402, Elsevier Science Publishers, B.V., North-Holland, Amsterdam.

[9] Jensen, D. R. and Ramirez, D. E. (2001). Detecting shifts in location and scale in regression, *South African Statis. J.* **35**, 181–204.

[10] Jensen, D. R. and Ramirez, D. E. (2002). Detecting mean-shift outliers via distances. *J. Statist. Computation and Simulation* **72(1):** 6–11.

[11] Kotz, S., Johnson, N. L., and Boyd, D. W. (1967a). Series representations of distributions of quadratic forms in normal variables I. Central case. *Ann. Math. Statist.* **38:** 823–837.

[12] Kotz, S., Johnson, N. L., and Boyd, D. W. (1967b). Series representations of distributions of quadratic forms in normal variables II. Non-central case. *Ann. Math. Statist.* **38:** 838–848.

[13] LaMotte, L. R. (1999). Collapsibility hypotheses and diagnostic bounds in regression analysis. *Metrika* **50:** 109–119.

[14] Myers, R. H. (1990). *Classical and Modern Regression with Applications, Second ed.* PWS-KENT Publishing Company, Boston, MA.

[15] Ramirez, D. E. (2000). The generalized F distribution. *J. Statistical Software* **5:** 1–13.

[16] Ramirez, D. E. and Jensen, D. R. (1991). Misspecified T^2 tests. II. Series expansions. *Commun. Statist. Simula.* **20:** 97–108.

[17] Ruben, H. (1962). Probability content of regions under spherical normal distributions, IV: The distribution of homogeneous and non-homogeneous quadratic functions of normal variables. *Ann. Math. Statist.* **34**: 542–570.

[18] Welsch, R. E. (1982). Influence functions and regression diagnostics. In *Modern Data Analysis,* R. L. Launer and A. F. Siegel, eds., Academic Press, New York.

[19] Welsch, R. E. and Kuh, E. (1977). Linear regression diagnostics. *Technical Report 923-77,* Sloan School of Management, Massachusetts Institute of Technology, Cambridge, MA.