DETECTING SHIFTS IN LOCATION AND SCALE IN REGRESSION

D.R. Jensen
Department of Statistics, Virginia Polytechnic Institute and State University
Blacksburg, VA 24061
djensen@vt.edu
D.E. Ramirez
Department of Mathematics, University of Virginia, Charlottesville, VA
22904–4137
der@virginia.edu

Key words: Canonical leverages; generalized noncentral F distributions; regression diagnostics; subset, outlier and influence diagnostics.

Summary: Complete stochastic properties are given for regression diagnostics on a comprehensive list under shifts in location and scale at designated design points under Gaussian errors. These are enabled by a unified approach to deletion diagnostics through recent studies to be cited. Findings are reported under single-case and subset deletions, and these are quantified through selected case studies from the literature.

1. Introduction

Regression diagnostics are concerned with validating the assumptions central to models in linear statistical inference. Prominent among these are diagnostics for identifying influential and outlying observations. Influence refers to disturbances in an estimator rendered through perturbations in the data. Specifically, an *influence measure* $\Im\left(\widehat{\theta},\widetilde{\theta}\right)$, as a deletion diagnostic, is

STMA (2001) SUBJECT CLASSIFICATION: 07:020(03:100)

intended to quantify disturbances in estimating some parameter θ before $\left(\widehat{\theta}\right)$ and after $\left(\widetilde{\theta}\right)$ deleting individual records or subsets of records from the data. References include books by Belsley, Kuh and Welsch (1980), Cook and Weisberg (1982), Barnett and Lewis (1984), Atkinson (1985), Rousseeuw and Leroy (1987), Chatterjee and Hadi (1988), Myers (1990) and Fox (1991). For a recent survey see Fung (1995).

Here we investigate a comprehensive list of deletion diagnostics, to include single-case and subset deletions, under outliers resulting from shifts in location and scale at designated design points. Earlier studies considered translation shifts only. Under such disturbances a complete assessment of their stochastic properties is given for many diagnostics on the list. These findings are enabled by recent work, to be cited, serving to unify the study of diverse diagnostic tools in current usage. An outline follows.

Preliminary developments in Section 2 include a survey of nonstandard distributions and a computational supplement to cover noncentral cases. Section 3 consolidates recent work unifying the study of diverse deletion diagnostics. The principal findings are reported in Section 4. These in turn are quantified through selected case studies from the literature as reported in Section 5. Section 6 offers a brief summary and conclusions, and proofs are deferred to an Appendix.

2. Preliminaries

2.1 Notation

Symbols identify $\Re^n$ as Euclidean n-space, $\Re^n_+$ as its positive orthant, and $\mathbb{S}_n$ as the real symmetric $(n \times n)$ matrices. Vectors and matrices are set in bold type; the transpose, inverse, trace, and determinant of A are A', A^{-1}, $tr(A)$, and $|A|$; $\mathbf{I}_n$ is the $(n \times n)$ identity; the unit vector on $\Re^n$ is $\mathbf{1}_n = [1, ..., 1]'$; $Diag(A_1, ..., A_k)$ is a blockdiagonal array; and the spectral decomposition of A in $\mathbb{S}_n$ is $A = QD_\alpha Q' = \sum_{i=1}^n \alpha_i \boldsymbol{q}_i \boldsymbol{q}_i'$, with $\boldsymbol{D}_\alpha = Diag(\alpha_1, ..., \alpha_n)$. A *reflexive g-inverse* A^- of A satisfies both $AA^-A = A$ and $A^-AA^- = A^-$. Given an array $\boldsymbol{a} \in \Re^n$ ordered as $\{a_1 \geq ... \geq a_n\}$, and similarly for $\boldsymbol{b} \in \Re^n$, then $\boldsymbol{a}$ is said to be *majorized* by $\boldsymbol{b}$, i.e., $\boldsymbol{a} \prec b$, if and only if, for each $1 \leqslant k < n$, $\sum_{i=1}^k a_i \leqslant \sum_{i=1}^k b_i$, whereas $\sum_{i=1}^n a_i = \sum_{i=1}^n b_i$.

2.2 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. In particular, $N_n(\boldsymbol{\mu}, \Sigma)$ is Gaussian on $\Re^n$ having $E(\boldsymbol{Y}) = \boldsymbol{\mu}$, dispersion matrix $V(\boldsymbol{Y}) = \Sigma$, and variance $Var(Y) = \sigma^2$ on $\Re^1$. Distributions of note on $\Re^1$ and $\Re^1_+$ are $t(\nu, \lambda)$, $\chi^2(\nu, \lambda)$, and $F(\nu_1, \nu_2, \lambda)$ as the noncentral Student's, chi-squared and Fisher distributions, having degrees of freedom $\{\nu, (\nu_1, \nu_2)\}$ and noncentrality λ. The *cdf*'s corresponding to the latter are $G(\cdot; \nu, \lambda)$ and $F(\cdot; \nu_1, \nu_2, \lambda)$, whereas $t(\nu), \chi^2(\nu)$, and $F(\nu_1, \nu_2)$ refer to central distributions.

Let $\{U_1, ..., U_s\}$ be mutually independent $\{N_1(\omega_i, 1); 1 \leq i \leq s\}$ random scalars, assumed to be independent of V having the *cdf* $G(v; \gamma)$; consider $\alpha' = [\alpha_1, ..., \alpha_s]$ as fixed positive weights; let $\boldsymbol{\omega}' = [\omega_1, ..., \omega_s]$; and let $W = \gamma\left(\alpha_1 U_1^2 + ... + \alpha_s U_s^2\right)/sV$, having the *cdf* $F_s(w; \alpha_1, ..., \alpha_s; \omega_1, ..., \omega_s; \gamma) = F_s(w; \alpha'; \boldsymbol{\omega}'; \gamma)$. 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 $\boldsymbol{\omega} = \mathbf{0}$. Corresponding expansions and error bounds are given next to account for the noncentral distributions to be encountered subsequently. The case $\alpha_1 = ... = \alpha_s = \alpha$ gives $F_s(w;\ \alpha, ..., \alpha;\ \omega_1, ..., \omega_s;\ \gamma) = F(w/\alpha;\ s, \gamma, \lambda)$ with $\lambda = \omega'\omega$. Moreover, W increases stochastically in each α_i and $|\omega_i|$ as other parameters are held fixed.

2.3 Computations: Generalized F

To continue, let $T = \alpha_1 U_1^2 + ... + \alpha_s U_s^2$, with weights satisfying $\{\alpha_s \geq \alpha_{s-1} \geq ... \geq \alpha_1 > 0\}$ in subsequent applications. For the central case with $\omega = 0$, the *pdf* for $F_s(w;\ \alpha_1, ..., \alpha_s;\ 0, ..., 0;\ \gamma)$, 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 $F_s(w;\ \alpha_1, ..., \alpha_s;\ 0, ..., 0;\ \gamma)$ is given in Ramirez (2000).

For the general case with $\omega \neq 0$, we may expand $\mathcal{L}(T)$ as a mixture of central chi-squared *pdf*s as in Ruben (1962) and Kotz, Johnson and Boyd (1967b). To these ends recursively define the sequences

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

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 $\{c_i > 0;\ i = 0, 1, ...\}$. In a manner similar to arguments supporting Theorem 3.1 of Ramirez and Jensen

(1991), we establish the following series expansions for the *pdf* for $F_s(w;\ \alpha_1, \ldots, \alpha_s;\ \omega_1, \ldots, \omega_s;\ \gamma)$ generally as follows.

Theorem 1:

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

$$h_F(w) = \sum_{i=0}^{\infty} \frac{c_i}{\beta} \frac{s}{\gamma} \frac{\Gamma[(s+2i+\gamma)/2]}{\Gamma[(s+2i)/2]\,\Gamma(\gamma/2)} \frac{\left(\frac{s}{\gamma}\frac{w}{\beta}\right)^{(s+2i-2)/2}}{\left(1+\frac{s}{\gamma}\frac{w}{\beta}\right)^{(s+2i+\gamma)/2}}, \tag{2.1}$$

and in terms of *pdf*s $f_F(\cdot;\ \nu.\gamma)$ for central F distributions as

$$h_F(w) = \sum_{i=0}^{\infty} \frac{c_i}{\beta} \frac{s}{s+2i} f_F\left(\frac{s}{s+2i}\frac{w}{\beta};\ s+2i,\ \gamma\right). \tag{2.2}$$

A global error bound for truncating at the τth partial sum of the *pdf* for W at both (2.1) and (2.2) is given by

$$\begin{aligned} e_\tau(w) &= \sum_{i=\tau+1}^{\infty} \frac{c_i}{\beta} \frac{s}{s+2i} f_F\left(\frac{s}{s+2i}\frac{w}{\beta};\ s+2i,\ \gamma\right) \\ &\le \frac{s}{\beta[s+2(\tau+1)]}\left[1-(c_0+\ldots+c_\tau)\right] = e_\tau. \end{aligned}$$

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

To compute the *cdf* of W, the series in Equation (2.1) is truncated and then integrated numerically. The probabilities to be reported subsequently were computed using this procedure, where the error tolerance was set as 10^{-5} and τ was increased until e_τ was less than the prescribed tolerance.

2.4 The models

To fix ideas, consider the model

$$\{Y_i = \beta_0 + \beta_1 X_{i1} + \ldots + \beta_k X_{ik} + \varepsilon_i;\ 1 \le i \le n\} \tag{2.3}$$

relating a response Y_i to k nonrandom regressors $\{X_{i1}, X_{i2}, \ldots, X_{ik}\}$ through $p = k+1$ unknown parameters $\beta = [\beta_0, \beta_1, \ldots, \beta_k]'$ Written as $Y_0 = X_0\beta + \epsilon_0$, each row x_i' of X_0 is a point in the design space. The reduced model $Y = X\beta + \epsilon$ follows on deleting from 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 $Z(s \times p)$, and deleting corresponding elements from Y_0. Here the matrices X_0, X, and Z are fixed nonrandom matrices. Then $\left(\widehat{\beta}, S^2\right)$ and $\left(\widehat{\beta}_I, S_I^2\right)$ are Gauss-Markov estimators and residual mean squares from the full and reduced data. Deleting x_i' from X_0 and Y_i from Y_0 gives $\left(\widehat{\beta}_i, S_i^2\right)$ for the case $s = 1$, together with values $\widehat{Y}_i$ and $\widehat{Y}_{i(i)}$ predicted at x_i' as elements of $\widehat{Y}_0 = X_0\widehat{\beta}$ and $\widehat{Y}_{(i)} = X_0\widehat{\beta}_i$, respectively, where $\widehat{\beta} = \left[\widehat{\beta}_0, \widehat{\beta}_1, \ldots, \widehat{\beta}_k\right]'$ and $\widehat{\beta}_i = \left[\widehat{\beta}_{0(i)}, \widehat{\beta}_{1(i)}, \widehat{\beta}_{k(i)}\right]'$.

The *leverage* of an individual design point x_i' of the fixed design X_0 is the element (h_{ii}) on the diagonal of the matrix $H = X_0 (X_0'X_0)^{-1}X_0'$. For subset deletions, the natural leverages emerge on letting $H_{II} = Z(X_0'X_0)^{-1}Z'$, then taking Q from the spectral decomposition $Q'H_{II}Q = D_\lambda = Diag(\lambda_1, \ldots, \lambda_s)$, to find that $Q'\widehat{Y}_I$ serve as *principal predictors*, $Q'Y_I$ as *principal responses*, and $\{\lambda_1 \geq \ldots \geq \lambda_s > 0\}$ as *canonical leverages*, the *canonical coleverages* having vanished, as shown in Jensen (2001). To assess the joint influence of points in Z, and to identify outliers at those points, consider the row-partitioned forms $Y_0' = [Y', Y_I']$, $X_0' = [X', Z']$, and $\epsilon_0' = [\epsilon', \epsilon_I']$. The full data give $\left(\widehat{\beta}, S^2\right)$ as before, as well as $\widehat{Y}_0 = X_0\widehat{\beta}$, and $e_0 = \left(Y_0 - \widehat{Y}_0\right)$, to be partitioned as $\widehat{Y}_0' = \left[\widehat{Y}', \widehat{Y}_I'\right]$ and $e_0' = [e', e_I']$, where $S^2 = e_0'e_0/(n-p)$. Corresponding values from the reduced data are $\left(\widehat{\beta}_I, S_I^2\right)$ as before, and $\widehat{Y}_{I(I)} = Z\widehat{\beta}_I$, where $S_I^2 = \left(Y - X\widehat{\beta}_I\right)'\left(Y - X\widehat{\beta}_I\right)/(n-p-s)$. 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 identified.

Gauss-Markov assumptions on moments of errors, then their distributions, are modified here as follows, where $\Xi(\sigma^2, \sigma_I^2) = Diag(\sigma^2 \boldsymbol{I}_r, \sigma_1^2 \boldsymbol{I}_s)$, *with* $r + s = n$.

Assumptions A:

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

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

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

This model allows for shifts in both location and scale at design points in $\boldsymbol{Z}$.

2.5 Studentized Statistics

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 $\boldsymbol{x}_i'$. For subsets, the R-Fisher statistic $F_I = \boldsymbol{e}_I' (\boldsymbol{I}_s - \boldsymbol{H}_{II})^{-1} \boldsymbol{e}_I / sS_I^2$ 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, it is known that $\mathcal{L}(F_I) = F(s, n-p-s, \lambda)$ with $\lambda = \boldsymbol{\delta}' (\boldsymbol{I}_s - \boldsymbol{H}_{II}) \boldsymbol{\delta}/\sigma^2$, thus supporting exact α-level tests for $H_0 : \boldsymbol{\delta} = \mathbf{0}$ against $H_1 : \boldsymbol{\delta} \neq \mathbf{0}$. This issue is reexamined subsequently under Assumptions A as given.

The foregoing assertions apply strictly to a designated design point, or to designated subsets of design points. Their broader implications in regression diagnostics are outlined in Section 6.

3. Unified diagnostics

Single-case and subset deletion diagnostics as currently used are listed in the accompanying tables. The diagnostics are arranged by type; their essential properties are summarized; and subscripts distinguish between single-case (i) and subset *(I)* deletions. Group I features the R-Student statistic t_i, and influence diagnostics $DFT_i = \Im\left(\widehat{Y}_i, \widehat{Y}_{i(i)}\right)$ and $DFB_{ij} = \Im\left(\widehat{\beta}_j, \widehat{\beta}_{j(i)}\right)$, known also as $DIFFITs$ and $DFBETAs$, as in Table 1. The latter serve to assess disturbances in predicted values, and in estimates for the individual beta coefficients, on deleting Y_i at design point $\boldsymbol{x}_i'$. Observe from Table 1 that both DFT_i and DFB_{ij} consist of a product of two components, namely, a scaling factor based on the fixed design matrix $\boldsymbol{X}_0$, and a random component based on the observed responses.

Group II outlier diagnostics are the 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), as defined in Table 2. Corresponding diagnostics $\{t_i^2, OUT_i, AP_i, CR_i, FV_i\}$ refer to the case $s = 1$.

Group III influence diagnostics encompass $\{C_I, WK_I, W_I, D_I\}$ as in Cook (1977), Welsch and Kuh (1977), Welsch (1982), and Jensen and Ramirez (1998), for gauging disturbances in the vector $\widehat{\beta}$ under deletions, as listed in Table 4. These are intended to metrize the vector-valued *sample influence curve* $SIC_I = (n-s)\left(\widehat{\beta} - \widehat{\beta}_I\right)$ for β, as in Chapter 3 of Cook and Weisberg (1982). Corresponding diagnostics $\{C_i, WK_i, W_i, D_i\}$ are listed separately in Table 3 for the case $s = 1$, where more complete properties are available than for $s > 1$. The forms for D_i and D_I use any reflexive g-inverse V^- of $V = V\left(\widehat{\beta} - \widehat{\beta}_I\right)$. A detailed explanation follows.

These diagnostics have been reexamined in response to anomalies in their usage. The principal issues are redundancies among the several diagnostics, and inconsistencies among conventional benchmarks for their use. It is now known that all single-case deletion diagnostics of Tables 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; and their joint distribution thus is singular of unit rank. 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 meanshift outlier at x_i' using t_i. Similar conclusions were reported for $\{DFT_i, DFB_{ij}, CR_i, C_i\}$ by LaMotte (1999) using different methods. Examples of inconsistencies among *ad hoc* contemporary rules were documented in Jensen (1998) and LaMotte (1999). To ensure consistency across diagnostics, Jensen (2000) gave revised benchmarks for the several diagnostics as listed in Tables 1, 2 and 3. Here $t_{\alpha/2}$ represents the $100(1-\alpha/2)$ percentile of $t(n-p-1)$, whereas $c_\alpha = t^2_{\alpha/2}$ in Table 2 under single-case deletions.

Subset deletion diagnostics were studied further as listed in Tables 2 and 4. The diagnostics $\{F_I, OUT_I, AP_I, CR_I, FV_I, D_I\}$ from Groups II and III are now known to correspond one-to-one, their joint distribution being singular of unit rank. 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. Conventional benchmarks again apply inconsistently across diagnostics, and thus may be supplanted by those appearing in Tables 2 and 4, where c_α is the $100(1-\alpha)$ percentile of $F(s, n-p-s)$, and α may be chosen at user discretion. Moreover, for the diagnostics $\{C_I, WK_I, W_I\}$, bounds on their upper cutoff values k_α are as listed in Table 4. For further details and proofs see Jensen (2001).

Regarding the Group III diagnostics $\{C_I, WK_I, W_I\}$, we first note that the distribution of C_I is clouded by dependence between its numerator and denominator. In Jensen and Ramirez (2000), we determined the subset influence for selected pairs of observations $(I = \{i, j\})$ in the *BOQ* data (see Section 5) arising through shifts in location only. In this case, WK_I and W_I are both noncentral generalized F variates. Our numerical studies there showed that the computed p-values, in testing $H_0 : \delta = 0$, were nearly identical using WK_I and W_I. Additionally, fewer terms were required for W_I than for WK_I on truncating central versions of the series (2.1), and thus we recommended using W_I. The p-values using the diagnostic F_I, equivalently D_I, were generally slightly larger than for W_I. In this case where there is no shift in scale, the diagnostic F_I has a central F-distribution under $H_0 : \delta = 0$.

4. The principal findings

Our goal is to establish stochastic properties for many of the diagnostics appearing in Tables 1–4. Our studies encompass all single-case deletion diagnostics as listed. Under subset deletions our findings include $\{F_I, OUT_I, AP_I, CR_I, FV_I, D_I\}$ from Groups II and III. As these are all equivalent, it suffices to establish properties of t_i and F_I under Assumptions A. The latter statistics have been studied heretofore under outliers arising through shifts in location only. Here we go beyond those earlier studies to include possible shifts in scale as well as location at designated design points comprising the matrix Z. Throughout we suppose that scale outliers are typified by $\sigma_1^2 \geq \sigma^2$, although findings parallel to these may be given for the case $\sigma_1^2 \leq \sigma^2$ as well. We next consider properties of F_I under Assumptions A; properties of t_i then follow as the special case where $s = 1$. In what follows we identify $\kappa = \sigma_1^2/\sigma^2$ and $\Xi(\kappa) = Diag(\boldsymbol{I}_r, \kappa \boldsymbol{I}_s)$, in which case $V(\varepsilon_0) = \Xi(\sigma^2, \sigma_1^2) = \sigma^2 \Xi(\kappa)$ at Assumptions A2 and A3.

Theorem 2:

Consider the R-Fisher diagnostic F_I under Assumptions A; let $\{\lambda_1 \geq ... \geq \lambda_s > 0\}$ comprise the canonical subset leverages; set $\kappa = \sigma_1^2/\sigma^2 \geq 1$; let $\boldsymbol{\theta} = \boldsymbol{Q}'\boldsymbol{\delta}$, such that $\boldsymbol{Q}'\boldsymbol{H}_{II}\boldsymbol{Q} = Diag(\lambda_1, ..., \lambda_s)$; and identify $\lambda = n - p - s$.

(i) The *cdf* of F_I is given by $F_s(w; \boldsymbol{\alpha}'; \boldsymbol{\omega}'; \gamma)$, with weights $\{\alpha_i = \kappa - (\kappa - 1)\lambda_i;\ 1 \leq i \leq s\}$ satisfying $\{\alpha_s \geq ... \geq \alpha_1 \geq 1\}$, and with location parameters $\left\{\omega_i = \theta_i/\sigma\,[\kappa + \lambda_i/(1-\lambda_i)]^{\frac{1}{2}};\ 1 \leq i \leq s\right\}$.

(ii) Bounds for the *cdf*'s in terms of Fisher's distribution are given by $F(w/\kappa;\ s, \gamma, \lambda) \leq F(w/\alpha_s;\ s, \gamma, \lambda) \leq F_s(w;\ \boldsymbol{\alpha}';\ \boldsymbol{\omega}';\ \gamma) \leq F(w/\alpha^*;\ s, \gamma, \lambda) \leq \boldsymbol{F}(\boldsymbol{w};\ \boldsymbol{s}, \gamma, \boldsymbol{\lambda})$ where $\lambda = \Sigma_{i=1}^s \theta_i^2/\sigma^2\,[\kappa + \lambda_i/(1-\lambda_i)]$ and $\alpha^* = (\alpha_1 ... \alpha_s)^{1/s}$ is the geometric mean.

(iii) $F(w/\alpha^*;\ s, \gamma, \lambda)$, when considered as a function of $\{\alpha_1, ..., \alpha_s\}$, is Schur-convex, i.e., if $\boldsymbol{\alpha} \prec \boldsymbol{\tau}$ on reordering elements, then $F(w/\alpha^*;\ s, \gamma, \lambda) \leq F(w/\tau^*;\ s, \gamma, \lambda)$.

Proof:

A proof is given in an Appendix.

Without further proof the foregoing results now specialize to include the nonstandard distribution of the R -Student diagnostic t_i under the nonstandard Assumptions A.

Corollary 1:

Consider the R -Student diagnostic t_i under Assumptions A with $s = 1$ having location-scale shifts $(\triangle, \sigma_1^2)$ at $\boldsymbol{x}_i'$; set $\kappa = \sigma_1^2/\sigma^2 \geq 1$; and identify $\gamma = n - p - 1$. Then

(i) $\mathcal{L}(t_i) = \mathcal{L}\left(\sqrt{\alpha_i}t'\right)$, where $\alpha_i = [\kappa - (\kappa - 1)\, h_{ii}]$ and t' designates a noncentral $t\,(\gamma, \delta)$ random variate having noncentrality $\delta = \triangle/\sigma\,[\kappa + h_{ii}/\,(1 - h_{ii})]^{\frac{1}{2}}$.

(ii) Under a scale shift at $\boldsymbol{x}_i'$ with $\triangle = 0$, the nonnull distribution of t_i is the scaled $\mathcal{L}(t_i) = \mathcal{L}\left(\sqrt{\alpha_i}t\right)$, where $\alpha_i = [\kappa - (\kappa - 1)\, h_{ii}]$ and t designates a central $t(\gamma)$ random variate.

Several points deserve notice. Theorem 2 (iii) asserts that the inner bounds of conclusion (ii) become tighter as variation among the canonical leverages $\{\lambda_1, ..., \lambda_s\}$, and thus among $\{\alpha_1, ..., \alpha_s\}$, diminishes in the sense of majorization. Moreover, when testing $H_0 : \left(\delta = 0, \sigma_1^2 = \sigma^2\right)$ at design points in $\boldsymbol{Z}$ at level α using F_I, we are assured that the test is unbiased owing to the stochastically increasing character of $F_s\,(\cdot;\, \alpha';\, \omega';\, \gamma)$ in each $\{\alpha_i \mid \omega_i \mid\}$ as other parameters are held fixed. In addition, the power increases monotonically with $\kappa \geq 1$ at $\delta = 0$. The most striking feature, to be examined numerically in case studies to follow, is that increasingly large canonical leverages serve increasingly to mask evidence of outliers arising through a shift in location, or a shift in scale, or both. In many applications this phenomenon in effect may negate any reasonable expectations of discerning outliers of either type at subsets of points having large canonical leverages. Parallel assessments apply in the case of single-case deletion diagnostics using the R-Student t_i.

5. Numerical studies: The BOQ data

5.1 The study

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.

5.2 Subset diagnostics

We focus on sites $\{15, 20, 21, 23, 24\}$, having individual leverages $\{0.5576, 0.3663, 0.0704, 0.9885, 0.8762\}$, their individual R-Student values exceeding the widely used ± 2 rule. For further study we select subsets of sizes two, three and four. Pairs of sites selected are $S_1 = \{20, 21\}$, $S_2 = \{15, 20\}$ and $S_3 = \{23, 24\}$, reflecting smaller, intermediate, and larger individual leverages. Subsets of three and four sites are $S_4 = \{20, 23, 24\}$ and $S_5 = \{15, 20, 21, 24\}$, reflecting a wide range of leverages. Details are reported in Tables 5 and 6 in testing for outliers using F_I at the 0.05 level. Specifically, we consider power of the test against a standardized location shift of $\triangle/\sigma \in \{1, 2, 3, 4\}$ units and a scale ratio $\kappa = \sigma_1^2/\sigma^2 \in \{1, 2, 3\}$ for observations at each of the several subsets of design points, where $\delta = \triangle \mathbf{1}_s$. Computations utilize the *SAS* and *Maple* software packages.

Table 5 lists power comparisons for subsets $\{S_1, S_2, S_3\}$ comprising pairs of design points, where exact values for $\kappa = 1$ derive from standard noncentral F-distributions. For other cases, values are listed in triples, where the top and bottom entries are lower and upper bounds as determined using the three inner inequalities of Theorem 2(ii). The middle entry gives exact power as outlined following Theorem 1, where parentheses contain the number of terms required to achieve an error bound at 10^{-5}. For example, in subset S_1 at $\triangle/\sigma = 3$, the exact power at $\kappa = 1$ is 0.8631, whereas at $\kappa = 2$ the exact power of 0.8273, requiring 15 terms, is seen to lie between 0.8262 and 0.8472.

Evidence presented in Table 5 is rather striking. Recall that individual leverages for subsets $\{S_1, S_2, S_3\}$ are $\{0.3663, 0.0704\}$, $\{0.5576, 0.3663\}$ and $\{0.9885, 0.8762\}$, whereas their corresponding canonical leverages are $\{0.3665, 0.0702\}$, $\{0.5582, 0.3657\}$ and $\{0.9946, 0.8700\}$, respectively. Within each subset the power tends to increase with $\triangle/\sigma$ for each fixed κ. On the other hand, the power tends to increase with κ for smaller values of $\triangle/\sigma$, whereas this trend reverses for larger values of $\triangle/\sigma$ within the ranges studied, with the exception of subset S_3. What is most striking, however, is the daunting extent to which effects of both translation and scale shifts are masked by increasing leverages. For example, in Table 5 the exact power to detect the extreme shifts $(\triangle/\sigma, \kappa) = (4, 3)$ decreases from 0.9426 at subset S_1, through 0.8826 at subset S_2, to 0.1613 at S_3, as their canonical leverages increase. These trends hold even for $\kappa = 1$, where the power to detect a location shift of four standard deviations decreases from 0.9852 at S_1, to 0.9319 at S_2, to the diminutive 0.1388 at S_3. Large canonical leverages in effect are seen to negate any reasonable expectations for detecting shifts in either location, or scale, or both.

Similar trends are seen in Table 6 for subsets S_4 and S_5, having individual leverages $\{0.3663, 0.9885, 0.8762\}$ and $\{0.5576, 0.3663, 0.0704, 0.8762\}$, and canonical leverages as given by $\{0.9949, 0.9211, 0.3149\}$ and $\{0.9605, 0.5364, 0.3123, 0.0613\}$, respectively. In particular, the extent of masking is less severe in the case of subset S_5, where the canonical leverages are dominated by those for subset S_4. Note in Tables 5 and 6 that lower and upper bounds on power are tightened as the range of the canonical leverages diminishes, as is evident also from the scale structure of the inner inequalities of Theorem 2(ii).

6. Conclusions

The single-case deletion diagnostics of Groups I, II and III, as in Tables 1, 2 and 3, are all equivalent. Moreover, the Group II subset deletion diagnostics in Table 2, and the Group III diagnostic D_I from Table 4, are all equivalent to the R-Fisher diagnostic F_I. These subset deletion diagnostics have been studied here under outliers arising from shifts in both location and scale. In these circumstances the distribution of F_I has been shown to be a noncentral generalized F distribution. We have given series expansions for these distributions, as well as global error bounds for their partial sums. Bounds for the *cdf*s of noncentral generalized F distributions also have been given in terms of the widely supported noncentral Fisher distributions.

The noncentral generalized F distributions have been used to compute the power for the diagnostic F_I, and thus for all equivalent diagnostics, under shifts in location and scale at selected subsets of the *BOQ* data. These case studies demonstrate that subsets with large canonical leverages tend to associate with tests having low power, so that shifts in both location and scale may be masked at points of high leverages.

Our experience as consultants is that really knowledgeable researchers often can identify prospective problem data from field or laboratory notes prior to data analyses. Our findings clearly apply where subsets of design points can be so designated. In practice, however, it is often encumbent on the analyst to identify anomalies from the data. Under single-case deletions it is standard practice to apply benchmarks in multiple inference based on a Bonferroni inequality, where each of several R-Student statistics has the requisite marginal distribution. These bounds carry over directly to include use of the R-Fisher statistics, each having the requisite marginal distribution under the usual validating assumptions. We thank one of the referees for requesting this clarification.

Fung (1995) focused on the daunting computational demands in evaluating numerous subset diagnostics currently used in conventional practice. Our studies show that these demands may be decimated when equivalent diagnostics are represented by a single diagnostic to be chosen at the discretion of the user. Moreover, our studies support a complete probabilistic assessment regarding the operating characteristics of numerous diagnostics in current usage.

Appendix

To prove Theorem 2, we proceed in several steps as follows, often invoking Assumptions A without further comment. Write $\boldsymbol{H}$ in partitioned form as $\boldsymbol{H} = [\boldsymbol{H}_{ij};\ i, j \in \{0, I\}]$ with $\boldsymbol{H}_{00} = \boldsymbol{X}(\boldsymbol{X}_0'\boldsymbol{X}_0)^{-1}\boldsymbol{X}'$, $\boldsymbol{H}_{I0}' = \boldsymbol{H}_{0I} = \boldsymbol{X}(\boldsymbol{X}_0'\boldsymbol{X}_0)^{-1}\boldsymbol{Z}$, and $\boldsymbol{H}_{II} = \boldsymbol{Z}(\boldsymbol{X}_0'\boldsymbol{X}_0 X)^{-1}\boldsymbol{Z}'$ as before. To show independence of the linear form, $\boldsymbol{e}_I = \boldsymbol{B}\boldsymbol{Y}_0 = -\boldsymbol{H}_{0I}\boldsymbol{Y} + (\boldsymbol{I}_s - \boldsymbol{H}_{II})\boldsymbol{Y}_I$, and the quadratic form $(n-p-s)\,S_I^2 = \left(\boldsymbol{Y} - \boldsymbol{X}\widehat{\boldsymbol{\beta}}_I\right)'\left(\boldsymbol{Y} - \boldsymbol{X}\widehat{\boldsymbol{\beta}}_I\right) = \boldsymbol{Y}_0'\boldsymbol{A}\boldsymbol{Y}_0$ under Assumptions A,with $\boldsymbol{A} = Diag\left(\left(\boldsymbol{I}_r - \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\right), \boldsymbol{0}\right)$, the reader may verify directly that $\boldsymbol{B}\boldsymbol{\Xi}(\kappa)\boldsymbol{A} = \boldsymbol{0}$. Independence under Gaussian errors now follows from standard theory, as required in Section 2.2 in regard to $F_s(w;\ \boldsymbol{\alpha}';\ \boldsymbol{\omega}';\ \gamma)$.

Observe further that $E(\boldsymbol{e}_I) = \boldsymbol{B}E(\boldsymbol{Y}_0) = (\boldsymbol{I}_s - \boldsymbol{H}_{II})\,\boldsymbol{\delta}$ and
$V(\boldsymbol{e}_I) = \sigma^2\boldsymbol{B}\boldsymbol{\Xi}(\kappa)\boldsymbol{B}' = \sigma^2\left[\kappa(\boldsymbol{I}_s - \boldsymbol{H}_{II})^2 + \boldsymbol{H}_{I0}\boldsymbol{H}_{0I}\right] = \sigma_1^2\left[(\boldsymbol{I}_s - \boldsymbol{H}_{II})^2 + \boldsymbol{H}_{I0}\boldsymbol{H}_{0I}\right]$
$-\left(\sigma_1^2 - \sigma^2\right)\boldsymbol{H}_{I0}\boldsymbol{H}_{0I}$. Next expand the idempotent matrix $(\boldsymbol{I}_n - \boldsymbol{H})^2 =$
$(\boldsymbol{I}_n - \boldsymbol{H})$
in its block-partitioned form and identify expressions on both sides to infer that $(\boldsymbol{I}_s - \boldsymbol{H}_{II})^2 + \boldsymbol{H}_{I0}\boldsymbol{H}_{0I} = (\boldsymbol{I}_s - \boldsymbol{H}_{II})$ and $\boldsymbol{H}_{I0}\boldsymbol{H}_{0I} = (\boldsymbol{I}_s - \boldsymbol{H}_{II})\boldsymbol{H}_{II}$, so that $V(\boldsymbol{e}_I/\sigma) = [\kappa(\boldsymbol{I}_s - \boldsymbol{H}_{II}) - (\kappa - 1)(\boldsymbol{I}_s - \boldsymbol{H}_{II})\boldsymbol{H}_{II}]$.

We next take the dispersion matrix into diagonal form on transforming $\boldsymbol{e}_I/\sigma \to \boldsymbol{Q}'\boldsymbol{e}_I/\sigma$, such that $E(\boldsymbol{Q}'\boldsymbol{e}_I/\sigma) = \boldsymbol{Q}'(\boldsymbol{I}_s - \boldsymbol{H}_{II})\boldsymbol{Q}\boldsymbol{Q}'\boldsymbol{\delta}/\sigma = \boldsymbol{D}_{1-\lambda}\boldsymbol{\theta}/\sigma$ and $V(\boldsymbol{Q}'\boldsymbol{e}_I/\sigma) = \kappa\boldsymbol{D}_{1-\lambda} - (\kappa-1)\boldsymbol{D}_{1-\lambda}\boldsymbol{D}_\lambda$, where $\boldsymbol{\theta} = \boldsymbol{Q}'\boldsymbol{\delta}$, $\boldsymbol{D}_\lambda = Diag(\lambda_1, \ldots, \lambda_s)$, and $\boldsymbol{D}_{1-\lambda} = Diag((1-\lambda_1), \ldots (1-\lambda_s))$. The matrix of the quadratic form in the numerator of F_I, namely, $\boldsymbol{e}_I'(\boldsymbol{I}_s - \boldsymbol{H}_{II})^{-1}\boldsymbol{e}_I = \boldsymbol{e}_I'\boldsymbol{Q}\boldsymbol{Q}'(\boldsymbol{I}_s - \boldsymbol{H}_{II})^{-1}\boldsymbol{Q}\boldsymbol{Q}'\boldsymbol{e}_I$, becomes $\boldsymbol{D}_{1-\lambda}^{-1}$. Thus as in Section 2.2, the required weights in the *cdf* $F_s(w;\ \alpha_1, \ldots, \alpha_s;\ \omega_1, \ldots, \omega_s;\ \gamma)$ are the eigenvalues of $\boldsymbol{D}_{1-\lambda}^{-1}[\kappa\boldsymbol{D}_{1-\lambda} - (\kappa-1)\boldsymbol{D}_{1-\lambda}\boldsymbol{D}_\lambda] = \kappa\boldsymbol{I}_s - (\kappa-1)\boldsymbol{D}_\lambda$. We now may identify the elements of the vector $\boldsymbol{U}$ as in Section 2.2, on letting $\boldsymbol{U} = [\kappa\boldsymbol{D}_{1-\lambda} - (\kappa-1)\boldsymbol{D}_{1-\lambda}\boldsymbol{D}_\lambda]^{-\frac{1}{2}}\boldsymbol{Q}'\boldsymbol{e}_I/\sigma$, so that $E(\boldsymbol{U}) = [\kappa\boldsymbol{D}_{1-\lambda} - (\kappa-1)\boldsymbol{D}_{1-\lambda}\boldsymbol{D}_\lambda]^{-\frac{1}{2}}\boldsymbol{D}_{1-\lambda}\boldsymbol{\theta}/\sigma$. It follows that the required location parameters for $F_s(w;\ \boldsymbol{\alpha}';\ \boldsymbol{\omega}';\ \gamma)$ are given by $\left\{\omega_i = \theta_i/\sigma\,[\kappa + \lambda_i/(1-\lambda_i)]^{\frac{1}{2}};\ 1 \le i \le s\right\}$. Clearly $\gamma = n - p - s$, to complete our proof for Theorem 2 (i).

To continue, we suppose that $\sigma_1^2 \geq \sigma^2$ as before, so that $\kappa \geq 1$. The remaining steps consist in bounding $\{\alpha_1, ..., \alpha_s\}$; using the property that $F_{s.}(w;\ \alpha';\ \boldsymbol{\omega}';\ \gamma)$ increases stochastically in each $\{\alpha_i, |\omega_i|\}$ with remaining parameters held fixed; and invoking the identity $F_s(w;\ \alpha, ..., \alpha,\ \omega_1, ..., \omega_s;\ \gamma) = F(w/\alpha;\ s, \gamma,\ \lambda)$ with $\lambda = \boldsymbol{\omega}'\boldsymbol{\omega}$. Conclusion (ii) now follows provisionally from $\{\alpha_i = \kappa - (\kappa - 1)\lambda_i;\ 1 \leq i \leq s\}$, and $\{\kappa \geq \alpha_s \geq ... \geq \alpha_1 \geq 1\}$, to give $F(w/\kappa;\ s, \gamma, \lambda) \leq F(w/\alpha_s;\ s, \gamma, \lambda) \leq F_s(w;\ \boldsymbol{\alpha}';\ \boldsymbol{\omega}';\ \gamma) \leq F(w/\alpha_1;\ s, \gamma, \lambda) \leq F(w;\ s, \gamma, \lambda)$. However, we may sharpen the inner upper bound on applying a result of Okamoto (1960) conditionally, given S_I^2, to get $F_s\left(w;\ \boldsymbol{\alpha}';\ \boldsymbol{\omega}';\ \gamma\,|\,S_I^2\right) \leq F\left(w/\alpha^*;\ s, \gamma, \lambda\,|\,S_I^2\right)$, with α^* as the geometric mean, and then unconditionally on taking expectations, to establish conclusion (ii) as stated in the theorem. Conclusion (iii) follows on noting that the geometric mean $\phi(\boldsymbol{\alpha}) = (\alpha_1 ... \alpha_s)^{1/s}$ is Schur-concave (cf. Marshall and Olkin, 1979, p. 79), i.e., $\boldsymbol{\alpha} \prec \boldsymbol{\tau}$ implies that $\phi(\boldsymbol{\alpha}) \geq \phi(\boldsymbol{\tau})$, so that $F(w/\alpha^*;\ s, \gamma, \lambda) \leq F(w/\tau^*;\ s, \gamma, \lambda)$ as claimed, to complete our proof.

References

ANDREWS, D.F. & PREGIBON, D. (1978). Finding outliers that matter. *J. Royal Statist. Soc. B* **40**: 85–93.

ATKINSON, A.C. (1985). *Plots, transformations and regression.* Oxford University Press, Oxford, U.K.

BARNETT, V. & LEWIS, T. (1984). *Outliers in statistical data, Second ed.* John Wiley & Sons, New York.

BELSLEY, D.A., KUH, E. & WELSCH, R.E. (1980). *Regression diagnostics: Identifying influential data and sources of collinearity.* John Wiley & Sons, New York.

CHATTERJEE, S. & HADI, A.S. (1988). *Sensitivity analysis in linear regression.* John Wiley & Sons, New York.

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

COOK, R.D. & WEISBERG, S. (1982). *Residuals and influence in regression.* Chapman & Hall, London.

FOX, J. (1991). *Regression diagnostics.* Sage Publications, Inc., Newbury Park, CA.

FUNG, W.K. (1995). A stepwise method for the identification of multiple outliers and influential observations. *South African Statist. J.* **29**: 51–64.

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.

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

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

JENSEN, D.R. & 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 & C.R. Rao, eds., pp. 387–402, Elsevier Science Publishers, B.V., North-Holland, Amsterdam.

JENSEN, D.R. & RAMIREZ, D.E. (2000). Detecting mean-shift outliers via distances. *J. Statist. Computation and Simulation* **58**: (to appear; 6 pages).

KOTZ, S., JOHNSON, N.L., & BOYD, D.W. (1967a). Series representations of distributions of quadratic forms in normal variables I. Central case. *Ann. Math. Statist.* **38**: 823–837.

KOTZ, S., JOHNSON, N.L., & BOYD, D.W. (1967b). Series representations of distributions of quadratic forms in normal variables II. Non-central case. *Ann. Math. Statist.* **38**: 838–848.

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

MARSHALL, A.W. & OLKIN, I. (1979). *Inequalities: Theory of majorization and its applications.* Academic Press, New York.

MYERS, R.H. (1990). *Classical and modern regression with applications, Second ed.* PWS-KENT Publishing Company, Boston, MA.

OKAMOTO, M. (1960). An inequality for the weighted sum of χ^2 variates. *Bull. Math. Statist.* **9**: 69–70.

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

RAMIREZ, D.E. & JENSEN, D.R. (1991). Misspecified T^2 tests. II. Series expansions. *Commun. Statist. Simula.* **20**: 97–108.

ROUSSEEUW, P.J. & LEROY, A.M. (1987). *Robust regression and outlier detection.* John Wiley and Sons, New York.

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.

WELSCH, R.E. (1982). Influence functions and regression diagnostics. In *Modern data analysis,* R.L. Launer & A.F. Siegel, eds., Academic Press, New York.

WELSCH, R.E. & KUH, E. (1977). Linear regression diagnostics. *Technical Report 923–77,* Sloan School of Management, Massachusetts Institute of Technology, Cambridge, MA.

Manuscript received, 2000.06, revised, 2001.06.

TABLE 1. Group I single-case deletion diagnostics, rules and benchmarks for their use, and their ranges, where $R = [r_{ji}] = (X'X)^{-1}X'$.

Diagnostic	Expression	Rule	Benchmark	Range
t_i	$\frac{(Y_i - \widehat{Y}_i)}{S_i\sqrt{1-h_{ii}}}$	$\pm$	$t_{\alpha/2}$	$(-\infty, \infty)$
DFT_i	$\frac{(\widehat{Y}_i - \widehat{Y}i(i))}{S_i\sqrt{h_{ii}}}$	$\pm$	$\frac{t_{\alpha/2}\sqrt{h_{ii}}}{\sqrt{1-h_{ii}}}$	$(-\infty, \infty)$
DFB_{ij}	$\frac{(\widehat{\beta}_j - \widehat{\beta}_{j(i)})}{S_i\sqrt{c_{jj}}}$	$\pm$	$\frac{t_{\alpha/2}r_{ji}}{\sqrt{c_{jj}(1-h_{ii})}}$	$(-\infty, \infty)$

TABLE 2. Group II subset deletion diagnostics, rules and benchmarks for their use, and their ranges, where $\eta = (n-p)/(n-p-s)$, $\Lambda = \prod_{i=1}^{s}(1-\lambda_i)^{-1}$, $GL_I = 1 - \prod_{i=1}^{s}(1-\lambda_i)$, and c_α becomes $t^2_{\alpha/2}$ for single-case deletions.

Diagnostic	Expression	Rule	Benchmark	Range
F_I	$\frac{e_I'(I_s - H_{II})^{-1}e_I}{sS_I^2}$	$>$	c_α	$(0, \infty)$
OUT_I	$1 - \frac{S_I^2}{S^2}$	$>$	$\frac{s(c_\alpha - 1)}{[sc_\alpha + (n-p-s)]}$	$\left[\frac{-s}{n-p-s}, 1\right]$
AP_I	$1 - \frac{(n-p-s)S_I^2\,\lvert X'X\rvert}{(n-p)S^2\,\lvert X_0'X_0\rvert}$	$>$	$\frac{[sc_\alpha + (n-p-s)GL_I]}{[sc_\alpha + (n-p-s)]}$	$[GL_I, 1]$
CR_I	$\frac{\lvert S_I^2(X'X)^{-1}\rvert}{\lvert S^2(X_0'X_0)^{-1}\rvert}$	$<$	$\left[\frac{(n-p)}{[sc_\alpha + (n-p-s)]}\right]^p \Lambda$	$[0, \eta^p\Lambda]$
FV_I	$\frac{\lvert S_I^2 Z(X'X)^{-1}Z'\rvert}{\lvert S^2 Z(X_0'X_0)^{-1}Z'\rvert}$	$<$	$\left[\frac{(n-p)}{[sc_\alpha + (n-p-s)]}\right]^s \Lambda$	$[0, \eta^s\Lambda]$

TABLE 3. Group III single-case deletion diagnostics, rules and benchmarks for their use, and their ranges, where $\gamma_i^2 = h_{ii}/(1-h_{ii})$.

Diag-nostic	Expression	Rule	Benchmark	Range
C_i	$\frac{(\hat{\beta}-\hat{\beta}_i)' X_0'X_0 (\hat{\beta}-\hat{\beta}_i)}{pS^2}$	>	$\frac{(n-p)t^2_{\alpha/2}\gamma_i^2}{p(t^2_{\alpha/2}+n-p-1)}$	$[0, (n-p)\gamma_i^2/p)]$
WK_i	$\frac{(\hat{\beta}-\hat{\beta}_i)' X_0'X_0 (\hat{\beta}-\hat{\beta}_i)}{pS_i^2}$	>	$\frac{t^2_{\alpha/2}\gamma_i^2}{p}$	$(0, \infty)$
W_i	$\frac{(\hat{\beta}-\hat{\beta}_i)' X'X (\hat{\beta}-\hat{\beta}_i)}{pS_i^2}$	>	$\frac{h_{ii}t^2_{\alpha/2}}{p}$	$(0, \infty)$
D_i	$\frac{(\hat{\beta}-\hat{\beta}_i)' V^- (\hat{\beta}-\hat{\beta}_i)}{S_i^2}$	>	$t^2_{\alpha/2}$	$(0, \infty)$

TABLE 4. Definitions of Group III subset deletion diagnostics, bounds on their upper cutoff values k_α, and their ranges

Diag-nostic	Expression	Benchmark	Range
C_I	$\frac{(\hat{\beta}-\hat{\beta}_I)' X_0'X_0 (\hat{\beta}-\hat{\beta}_I)}{pS^2}$	$\frac{(n-p)sc_\alpha\alpha_s}{p(sc_\alpha+n-p-s)} \leq k_\alpha \leq \frac{(n-p)sc_\alpha\alpha_1}{p(sc_\alpha+n-p-s)}$	$[0, (n-p)\alpha_1/p)]$
WK_I	$\frac{(\hat{\beta}-\hat{\beta}_I)' X_0'X_0 (\hat{\beta}-\hat{\beta}_I)}{pS_I^2}$	$\frac{c_\alpha}{\alpha_1} \leq k_\alpha \leq \frac{c_\alpha}{\alpha_s}$	$[0, \infty]$
W_I	$\frac{(\hat{\beta}-\hat{\beta}_I)' X_0'X_0 (\hat{\beta}-\hat{\beta}_I)}{pS_I^2}$	$\frac{c_\alpha}{\lambda_1} \leq k_\alpha \leq \frac{c_\alpha}{\alpha_s}$	$[0, \infty]$
D_I	$\frac{(\hat{\beta}-\hat{\beta}_I)' V^- (\hat{\beta}-\hat{\beta}_I)}{sS_I^2}$	c_α	$[0, \infty]$

TABLE 5. Power of F_I under location shifts $\delta/\sigma = \Delta 1_s/\sigma$ for scale ratios $\kappa = \sigma_1^2/\sigma^2$ at design points comprising subsets S_1, S_2 and S_3 of the BOQ data, where pairs of entries give lower and upper bounds on the actual power from Theorem 1 (ii).

	Subset S_1			Subset S_2			Subset S_3		
	κ			κ			κ		
Δ/σ	1	2	3	1	2	3	1	2	3
1	0.1573	0.2703	0.3592	0.1240	0.2045	0.2759	0.0551	0.0634	0.0716
		0.2715(6)	0.3606(6)		0.2061(6)	0.2779(6)		0.0640(4)	0.0734(5)
		0.2987	0.3999		0.2237	0.3072		0.0723	0.0907
2	0.5088	0.5501	0.5836	0.3776	0.4293	0.4707	0.0707	0.0789	0.0869
		0.5518(10)	0.5848(9)		0.4327(9)	0.4743(9)		0.0805(5)	0.0904(6)
		0.5828	0.6231		0.4551	0.5058		0.0893	0.1085
3	0.8631	0.8262	0.8099	0.7199	0.7063	0.7041	0.0981	0.1057	0.1133
		0.8273(15)	0.8105(13)		0.7100(13)	0.7080(12)		0.1089(6)	0.1195(7)
		0.8472	0.8364		0.7290	0.7339		0.1185	0.1389
4	0.9852	0.9620	0.9424	0.9319	0.9007	0.8801	0.1388	0.1451	0.1514
		0.9623(21)	0.9426(18)		0.9028(18)	0.8826(16)		0.1504(7)	0.1613(8)
		0.9688	0.9534		0.9123	0.8971		0.1609	0.1823

TABLE 6. Power of F_I under location shifts $\delta = \Delta \mathbf{1}_s$ for scale ratios $\kappa = \sigma_1^2/\sigma^2$ at design points comprising subsets S_4 and S_5 of the BOQ data, where pairs of entries give lower and upper bounds on the actual power from Theorem 1 (ii).

	Subset S_4			Subset S_5		
	κ			κ		
Δ/σ	1	2	3	1	2	3
1	0.0728	0.1043	0.1334	0.1108	0.2005	0.2852
		0.1171(11)	0.1612(16)		0.2179(15)	0.3166(22)
		0.1946	0.3126		0.3075	0.4670
2	0.1517	0.1718	0.1949	0.3424	0.3977	0.4557
		0.2012(14)	0.2424(19)		0.4337(22)	0.5019(29)
		0.2893	0.4016		0.5316	0.6428
3	0.3018	0.2904	0.2988	0.6926	0.6674	0.6776
		0.3416(17)	0.3708(23)		0.7083(29)	0.7251(37)
		0.4371	0.5314		0.7831	0.8270
4	0.5090	0.4522	0.4380	0.9251	0.8785	0.8607
		0.5197(21)	0.5282(28)		0.9034(38)	0.8920(47)
		0.6107	0.6751		0.9375	0.9426