View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by Osuva

Working Papers of The University of Vaasa,
Department of Mathematics and Statistics,
16

The joint distribution of a linear transformation of internally studentized least squares residuals

Seppo Pynnönen

Preprint, September 2010

University of Vaasa
Department of Mathematics and Statistics
P.O.Box 700, FI-95101 Vaasa, Finland
Preprints available at: http://lipas.uwasa.fi/julkaisu/ewp.html

The joint distribution of an arbitrary linear transformation of internally studentized least squares residuals of a linear regression

Seppo Pynnonen, Department of Mathematics and Statistics, University of Vaasa, P.O.Box 700, FI-65101, Vaasa, Finland, Email address: sjp@uwasa.fi, Office phone: +358-6-3248259, Fax: +358-6-3248557

Abstract. The term 'studentization' is commonly used to describe a scale parameter dependent quantity U by a scale estimate S such that the resulting ratio, U/S, has a distribution that is free of from the nuisance unknown scale parameter. *External* studentization refers to a ratio in which the nominator and denominator are independent, while *internal* studentization refers to a ratio in which they are dependent. The advantage of the internal studentization is that typically one can use a single common scale estimator, while in the external studentization every single residual is scaled by different scale estimator to gain the independence. For normal regression errors the joint distribution of and arbitrary (linearly independent) subset of internally studentized residuals is well known. However, in some application a linear combination of internally studentized residuals may be useful. The boundedness of an arbitrary linear combination these residuals is well documented, but the distribution of a linear combination seems not have been derived in literature. This paper contributes the existing literature by deriving the joint distribution of a general linear transformation of internally studentized residuals. The distribution of a univariate linear combination and joint distributions of all major versions of commonly utilized internally studentized residuals are obtained as special cases. The paper shows that the joint distribution of all linear combinations (including an arbitrary linearly independent subset of residuals) belong to the same family of distributions. Some test statistics for inference purposes are also derived.

Key words: Linear combination of residuals; normed residuals; distribution of residuals; bounds for linear combination of studentized residuals; t-statistic for studentized residuals; F-statistic for studentized residuals; regression diagnostics

MSC: 62E15; 62J05

1. Introduction

Consider the usual regression model with n observations

$$\boldsymbol{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{u}, \tag{1}$$

where $\mathbf{X}$ is an $n \times p'$ nonstochastic matrix of rank $p \leq p' < n$, $\boldsymbol{y}$ is an n-vector of observable responses, $\boldsymbol{\beta}$ is a p'-vector of slope parameters, and $\boldsymbol{u} \sim N(\mathbf{0}, \boldsymbol{\sigma^2}\mathbf{I})$ is an n-vector of unobservable homoscedastic independent normally distributed errors with (unkown) variance parameter σ^2 and $\mathbf{I}$ is an $n \times n$ identity matrix.

The least squares residuals are given by

$$\begin{aligned} \hat{\boldsymbol{u}} &= \mathbf{Q}\boldsymbol{y} \\ &= \mathbf{Q}\boldsymbol{u}, \end{aligned} \tag{2}$$

where

$$\mathbf{Q} = \mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-}\mathbf{X}' \tag{3}$$

is an $n \times n$ idempotent matrix with rank $n - p$ in which the prime denotes the matrix transposition and $(\mathbf{X}'\mathbf{X})^{-}$ is a generalized inverse of $\mathbf{X}'\mathbf{X}$. The error sum of squares $\boldsymbol{u}'\boldsymbol{u}$ can be estimated by the residual sum of squares

$$S^2 = \hat{\boldsymbol{u}}'\hat{\boldsymbol{u}} = \boldsymbol{u}'\mathbf{Q}\boldsymbol{u}. \tag{4}$$

By equation (2) the least squares residuals are linear combinations of normal random variables, which implies that $\hat{\boldsymbol{u}}$ follows a (singular) normal distribution, such that

$$\hat{\boldsymbol{u}} \sim N\left(0, \sigma^2\mathbf{Q}\right). \tag{5}$$

Studentization is a common term used to describe division of a scale parameter dependent statistic, say U, by a scale estimate S such that the distribution of the resulting ratio U/S is free from the nuisance scale parameters [see e.g. Margolin (1977)]. Typically U and S are derived from the same data, in which case the ratio U/S is called *internally* studentized if U and S are dependent and *externally* studentized if they are independent [see, Cook and Weisberg (1982, p.18)].

In least squares regression, the internally studentized residuals are defined as

$$r_i = \frac{\hat{u}_i}{s\sqrt{q_{ii}}}, \tag{6}$$

where $\hat{u}_i$ is the ith component of the vector $\hat{\boldsymbol{u}}$, $s = \sqrt{\hat{\boldsymbol{u}}'\hat{\boldsymbol{u}}/(n-p)}$, and q_{ii} is the ith diagonal element of the matrix $\mathbf{Q}$. Although r_i is the familiar ratio of a normally distributed random variable and a square root of a scaled Chi-square random variable, the end result is not a t-distributed random variable. The reason is that the nominator and the denominator are not independent due to the fact that $\hat{u}_i^2 \leq \hat{\boldsymbol{u}}'\hat{\boldsymbol{u}}$ for all $i = 1, \ldots, n$. Furthermore, because $q_{ii} \geq 1 + 1/n$ [Cook and Weisberg (1982, p. 12)], it follows that unlike a t-distributed random variable that can assume all real values, $r_i^2 \leq n - p$.

Stefansky (1972), Ellenberg (1973), and Díaz-García and Gutiérrez-Jáimez (2007) derived the joint distribution of an arbitrary (nonsingular) subset of the internally studentized residuals defined in (6). Beckman and Trussel (1974) derive the distribution for a t-statistic for a single r_i. The present paper derives the joint distribution of an arbitrary non-singular linear transformation of internally studentized residuals.

2. The Main Result

Definition 1 *Let* $\mathbf{M}$ *be an* $m \times n$ *matrix. We call*

$$\boldsymbol{r}_M = \frac{1}{s}\mathbf{M}\hat{\boldsymbol{u}} \tag{7}$$

a studentized linear transformation of least squares residuals $\hat{\boldsymbol{u}}$*, where* $s = \sqrt{\hat{\boldsymbol{u}}'\hat{\boldsymbol{u}}/(n-p)}$*. Furthermore, we call the transformation nonsingular if the matrix* $\mathbf{MQM}'$ *is positive definite, in which case* $m \leq n - p$*.*

All the major classes of residuals defined in literature can be obtained as special cases of (7). For example, consider in addition to the internally studentized residuals defined in equation (6), also the following alternatives [see, e.g., Chatterjee and Hadi (1988) and Lloynes (1979)]:

Normalized residuals:

$$\hat{u}_i/\sqrt{\hat{\boldsymbol{u}}'\hat{\boldsymbol{u}}}. \tag{8}$$

Standardized residuals:

$$\hat{u}_i/s. \tag{9}$$

Abrahamse-Koerts residuals (normalized):

$$\mathbf{B}'\hat{\boldsymbol{u}}/\sqrt{\hat{\boldsymbol{u}}'\hat{\boldsymbol{u}}}, \tag{10}$$

where $\mathbf{B}$ is an $n \times n$ matrix defined in Abrahamse and Koerts (1971), satisfying $\hat{\boldsymbol{u}}'\mathbf{B}\mathbf{B}'\hat{\boldsymbol{u}} = \hat{\boldsymbol{u}}'\hat{\boldsymbol{u}}$.

Consider an arbitrary linearly independent subset $I_m = \{i_1, i_2, \ldots, i_m\} \subset \{1, \ldots n\}$, $m \leq n - p$ of the above classes of residuals. It is easily seen that each of these is a special case of the linear transformation defined in equation (7). For the purpose, define $\mathbf{M}_I$ as a $m \times n$ matrix in which each row $j = 1, \ldots, m$ is a $1 \times n$ vector with element $i_j = 1$ and zeros elsewhere, $i_j \in I_m$. Furthermore, let $\mathbf{D}^{-1/2}$ denote an $n \times n$ diagonal matrix with elements $(q_{11})^{-1/2}, \ldots, (q_{nn})^{-1/2}$. Then a set, I_m, of internally studentized residuals, defined in equation (6), is obtained by defining in equation (7), $\mathbf{M} = \mathbf{M}_I\mathbf{D}^{-1/2}$, a set of normalized residuals, defined in equation (8), is obtained by defining $\mathbf{M} = (n-p)^{-1/2}\mathbf{M}_I$, a set of standardized residuals, defined in equation (9), is obtained by defining $\mathbf{M} = \mathbf{M}_I$, and a set of Abrahamse-Koerts residuals are obtained, defined in equation (10), by defining $\mathbf{M} = (n-p)^{-1/2}\mathbf{M}_I\mathbf{B}'$.

The following lemma is useful.

Lemma 1 *Let* $\mathbf{M}$ *be a non-singular transformation as given in Definition 1, then*

$$V_M = \hat{\boldsymbol{u}}'\hat{\boldsymbol{u}} - \hat{\boldsymbol{u}}'\mathbf{M}'(\mathbf{MQM}')^{-}\mathbf{M}\hat{\boldsymbol{u}} \tag{11}$$

is distributed as $\sigma^2\chi^2(n-p-q)$*, where* $q = \mathrm{rank}(\mathbf{MQ})$ *and*

$$U_M = \hat{\boldsymbol{u}}'\mathbf{M}'(\mathbf{MQM}')^{-}\mathbf{M}\hat{\boldsymbol{u}} \tag{12}$$

is distributed as $\sigma^2\chi^2(q)$*.*

Proof. By equation (2), $\hat{\boldsymbol{u}} = \mathbf{Q}\boldsymbol{u}$, where $\mathbf{Q}$ is an indempotent matrix with $\mathrm{tr}\,\mathbf{Q} = n-p$. Thus V_M in (11) can be written as

$$V_M = \boldsymbol{u}'\left(\mathbf{Q} - \mathbf{QM}'(\mathbf{MQM}')^{-}\mathbf{MQ}\right)\boldsymbol{u}. \tag{13}$$

Because $\boldsymbol{u} \sim N(\mathbf{0}, \sigma^2\mathbf{I})$, V_M is $\sigma^2\chi^2(n-p-q)$ distributed if and only if the matrix, $\mathbf{Q} - \mathbf{QM}'(\mathbf{MQM}')^{-}\mathbf{MQ}$, in the quadratic form in (13) is idempotent of rank $n-p-q$. Direct multiplication of the matrix $\mathbf{Q} - \mathbf{QM}'(\mathbf{MQM}')^{-}\mathbf{MQ}$ by itself shows that it is idempotent (the only trick is to show that the matrix $\mathbf{QM}'(\mathbf{MQM}')^{-}\mathbf{MQM}'(\mathbf{MQM}')^{-}\mathbf{MQ}$ equals $\mathbf{QM}'(\mathbf{MQM}')^{-}\mathbf{MQ}$, which can be easily shown by taking the difference of these matrices and multiplying by itself to find out that the product is zero). The rank of an idempotent matrix is the same as its trace, which in this case is

$$\begin{aligned}
\mathrm{tr}\left(\mathbf{Q} - \mathbf{QM}'(\mathbf{MQM}')^{-}\mathbf{MQ}\right) &= \mathrm{tr}\,\mathbf{Q} - \mathrm{tr}\left(\mathbf{QM}'(\mathbf{MQM})^{-}\mathbf{MQ}\right) \\
&= n - p - \mathrm{tr}\left((\mathbf{MQM}')^{-}\mathbf{MQM}'\right) \\
&= n - p - q,
\end{aligned}$$

because $(\mathbf{MQM}')^{-}\mathbf{MQM}'$ is idempotent with trace equalling the rank of the matrix. Thus, the proof that $V_M \sim \sigma^2\chi^2(n-p-q)$ is complete.

In the same manner, writing $U_M = \boldsymbol{u}'\mathbf{QM}'(\mathbf{MQM}')^{-}\mathbf{MQ}\boldsymbol{u}$, we find immediately that $\mathbf{QM}'(\mathbf{MQM}')^{-}\mathbf{MQ}$ is idempotent and $\mathrm{tr}\,(\mathbf{QM}'(\mathbf{MQM}')^{-}\mathbf{MQ}) =$

$\operatorname{tr}((\mathbf{MQM}')^{-}(\mathbf{MQM}')) = q$, which together with $\boldsymbol{u} \sim N(\mathbf{0}, \sigma^2\mathbf{I})$ imply that $U_M \sim \sigma^2\chi^2(q)$, completing the proof.

Corollary 1 *The studentized residuals defined in equation* (7) *are bounded such that*

$$\boldsymbol{r}_M' \left(\mathbf{MQM}'\right)^{-} \boldsymbol{r}_M \leq n - p. \tag{14}$$

Proof. Let $s^2 = \hat{\boldsymbol{u}}'\hat{\boldsymbol{u}}/(n-p)$. Then by Lemma 1 and by the definition of $\boldsymbol{r}_M$ in equation (7), $0 \leq V_M = \hat{\boldsymbol{u}}'\hat{\boldsymbol{u}} - \hat{\boldsymbol{u}}'\mathbf{M}'(\mathbf{MQM}')^{-}\mathbf{M}\hat{\boldsymbol{u}} = (n-p)s^2 - s^2\boldsymbol{r}_M' \left(\mathbf{MQM}'\right)^{-} \boldsymbol{r}_M$, from which the result follows immediately, completing the proof.

Lemma 2 *Under the assumptions of Lemma* 1, V_M *and* $\mathbf{M}\hat{\boldsymbol{u}}$ *are independent.*

Proof. Because $\mathbf{M}\hat{\boldsymbol{u}} = \mathbf{MQ}\boldsymbol{u}$ is a normal random vector and V_M is the quadratic form of the normal random vector

$$\boldsymbol{u}_M = \left(\mathbf{Q} - \mathbf{QM}'(\mathbf{MQM}')^{-}\mathbf{MQ}\right)\boldsymbol{u}, \tag{15}$$

a sufficient condition for the independence of $\mathbf{M}\hat{\boldsymbol{u}}$ and V_M is that $\mathbf{MQ}\boldsymbol{u}$ and $\boldsymbol{u}_M$, defined in equation(15), are uncorrelated. This is the case if the defining matrices, $\mathbf{MQ}$ and $\mathbf{Q} - \mathbf{QM}'(\mathbf{MQM}')^{-}\mathbf{MQ}$, of these random vectors are orthogonal. Now

$$\mathbf{MQ}\left(\mathbf{Q} - \mathbf{QM}'(\mathbf{MQM}')^{-}\mathbf{MQ}\right) = \mathbf{MQ} - \mathbf{MQM}'(\mathbf{MQM}')^{-}\mathbf{MQ} = \mathbf{P}.$$

It is straightforward to show that $\mathbf{PP}' = \mathbf{0}$, which implies that $\mathbf{P} = \mathbf{0}$, from which the orthogonality and hence uncorrelation follow, completing the proof of the lemma.

Lemma 2 implies immediately:

Corollary 2 *Under the assumptions of Lemma* 2 *the chi-square random variables* V_M *and* U_M*, defined in equations* (11) *and* (12)*, respectively, are independent.*

With these results we can derive the main result of this paper.

Theorem 1 *Under the assumptions of the linear regression model in* (1)*, the joint distribution of the nonsingular studentized linear transform of residuals, defined in equation* (7) *when* $m < n - p$*, is*

$$\begin{aligned} f_{r_M}(\boldsymbol{r}) &= \frac{\Gamma\left[(n-p)/2\right]\left|\mathbf{MQM}'\right|^{-1/2}}{(\pi(n-p))^{m/2}\,\Gamma\left[(n-p-m)/2\right]} \\ &\quad \times \left(1 - \frac{\boldsymbol{r}'\left(\mathbf{MQM}'\right)^{-1}\boldsymbol{r}}{n-p}\right)^{\frac{1}{2}(n-p-m)-1}, \quad \boldsymbol{r}'\left(\mathbf{MQM}'\right)^{-1}\boldsymbol{r} \leq n-p, \end{aligned} \tag{16}$$

where $\Gamma(\cdot)$ *is the Gamma function.*

Proof. Under the nonsingularity assumption $\mathbf{MQM}'$ is positive definite, the inverse $(\mathbf{MQM}')^{-1}$ exists and that $r = m$, i.e., the rank of the matrix is m. The bounds, $\boldsymbol{r}'\left(\mathbf{MQM}'\right)^{-1}\boldsymbol{r} \leq n-p$, follow from Corollary 1. The rest can be proceeded analogously to Ellenberg (1973). That is, due to the χ^2-distribution result of V_M in Lemma 1, normality of $\boldsymbol{u}_M = \mathbf{M}\hat{\boldsymbol{u}}$, and independence of V_M and $\boldsymbol{u}_M$ by Lemma 2, their joint density is the product of their appropriate densities, resulting to

$$\begin{aligned} f_{\boldsymbol{u}_M,V_M}(\boldsymbol{x}, v) &= \frac{1}{\pi^{\frac{m}{2}}(2\sigma^2)^{\frac{1}{2}(n-p)}|\mathbf{MQM}'|^{\frac{1}{2}}\Gamma[(n-p-m)/2]} v^{\frac{1}{2}(n-p-m)-1} \\ &\quad \times \exp\left\{-\frac{1}{2\sigma^2}\left[\boldsymbol{x}'(\mathbf{MQM}')^{-1}\boldsymbol{x} + v\right]\right\}. \end{aligned} \tag{17}$$

Define next the transformations

$$\begin{aligned} \boldsymbol{r} &= \boldsymbol{x}/\sqrt{y/(n-p)} \\ y &= \boldsymbol{x}'(\mathbf{MQM}')^{-1}\boldsymbol{x} + v \end{aligned}$$

The Jacobian of this transformation is

$$\left(\frac{y}{n-p}\right)^{\frac{1}{2}m}. \tag{18}$$

Using these, the joint density of $\boldsymbol{r}_M$ and $S^2 = \hat{\boldsymbol{u}}'\hat{\boldsymbol{u}}$ becomes

$$\begin{aligned} f_{\boldsymbol{r}_M,S^2}(\boldsymbol{r}, y) &= \left(\frac{y}{n-p}\right)^{\frac{m}{2}} f_{\boldsymbol{u}_M,V_M}\left(\boldsymbol{r}\sqrt{y/(n-p)}, y - \boldsymbol{x}'(\mathbf{MQM}')^{-1}\boldsymbol{x}\right) \\ &= \frac{|\mathbf{MQM}'|^{-\frac{1}{2}}}{(\pi(n-p))^{\frac{m}{2}}\,\Gamma[(n-p-m)/2]} \\ &\quad \times \left(1 - \frac{\boldsymbol{r}'(\mathbf{MQM}')^{-1}\boldsymbol{r}}{n-p}\right)^{\frac{1}{2}(n-p-m)-1} \\ &\quad \times \frac{1}{2^{(n-p)/2}\sigma^2}\left(\frac{y}{\sigma^2}\right)^{\frac{1}{2}(n-p)-1} e^{-\frac{y}{2\sigma^2}}. \end{aligned}$$

Introducing a further transformation $z = y/\sigma^2$ with Jacobian σ^2 and integrating with respect to z yields finally the marginal density of $\boldsymbol{r}_M$ given in (16), which completes the proof of the theorem.

Theorem 1 gives the joint distribution of the studentized nonsingular transformation, $\boldsymbol{r}_M$, when $m < n-p$. The next theorem gives the distribution for a nonsingular studentized transformation with $m = n-p$.

Theorem 2 *Under the assumption of Theorem* 1*, when $m = n-p$,*

$$\boldsymbol{r}_M \stackrel{d}{=} \sqrt{n-p}\,\tilde{\mathbf{M}}\boldsymbol{U} \tag{19}$$

where $\boldsymbol{U}$ is uniformly distributed $(n-p)\times 1$ random vector on the unit sphere $\mathcal{S}_{n-p} = \{\boldsymbol{x} \in \mathbb{R}^{n-p} : \boldsymbol{x}'\boldsymbol{x} = 1\}$, $\stackrel{d}{=}$ means that the random variables have *the same distribution, and* $\tilde{\mathbf{M}} = \mathbf{M}\mathbf{H}_m$ *is an* $m \times m$ *matrix in which* $\mathbf{H}_m$ *are the* $m = n - p$ *columns of the eigenvectors of* $\mathbf{Q}$ *that correspond the unit eigenvalues, such that* $\tilde{\mathbf{M}}\tilde{\mathbf{M}}' = \mathbf{M}\mathbf{Q}\mathbf{M}'$.

Proof: A proof is given in Pynnönen (2010b).

3. Applications

In particular, selecting $m = 1$ such that $\mathbf{M}$ becomes a row vector vector, Theorem 1 implies that the density function of the distribution of an arbitrary (nonsingular) linear combination $r_m = \boldsymbol{m}'\hat{\boldsymbol{u}}$, where $\boldsymbol{m}$ is an $n \times 1$ vector of real numbers satisfying $\boldsymbol{m}'\mathbf{Q}\boldsymbol{m} > 0$, is

$$f_{r_m}(r) = \frac{\Gamma[(n-p)/2](\boldsymbol{m}'\mathbf{Q}\boldsymbol{m})^{-1/2}}{\Gamma[(n-p-1)/2]\Gamma[1/2]\sqrt{n-p}} \left(1 - \frac{r^2}{(n-p)\,\boldsymbol{m}'\mathbf{Q}\boldsymbol{m}}\right)^{\frac{1}{2}(n-p-1)-1} \tag{20}$$

for $|r| \le \sqrt{(n-p)\,\boldsymbol{m}'\mathbf{Q}\boldsymbol{m}}$. Thus, $r_m^2/[(n-p)\boldsymbol{m}'\mathbf{Q}\boldsymbol{m}]$ follows a Beta distribution with parameters $1/2$ and $(n-p-1)/2$. Setting the ith component in $\boldsymbol{m}$ equal to 1 and all others equal to zero gives the density function for a single internally studentized residual $t_i = \hat{u}_i/s$, where $s = \sqrt{\hat{\boldsymbol{u}}\hat{\boldsymbol{u}}/(n-p)}$,

$$f_{t_i}(r) = \frac{\Gamma[(n-p)/2](q_{ii})^{-1/2}}{\Gamma[(n-p-1)/2]\Gamma[1/2]\sqrt{n-p}} \left(1 - \frac{r^2}{(n-p)q_{ii}}\right)^{\frac{1}{2}(n-p-1)-1}, \tag{21}$$

$|r| \le \sqrt{(n-p)q_{ii}}$, where q_{ii} is the ith diagonal element of $\mathbf{Q}$.

Again, $r_i^2/[(n-p)q_{ii}]$ follows a Beta distribution with parameters $1/2$ and $(n-p-1)/2$. Thus, the distributions of single internally studentized residuals and their arbitrary linear combinations belong to the same family of distributions such that through simple transformations they are identically distributed as the Beta distribution with parameters $1/2$ and $(n-p-1)/2$. This obviously facilitates inference based on individual residuals or their linear combinations.

Moreover, as is shown below, the situation can be further facilitated by introducing a transformation that leads to a common t-distribution.

Individual residuals are typically used as diagnostic tools for model checking and testing for outliers and influential observations [see, Cook and Weisberg (1982)]. A potential field of applications of using linear combinations of internally studentized residuals is for example in testing for various seasonal patterns in the residuals of a time series regression. Typically seasonal effects are tested using appropriate dummy-variables, which can be shown to be equivalent of using a statistic based on an externally studentized sum of residuals. Perhaps a more concrete application, where the distribution results of this paper give additional insight is in the field of event studies in financial economics [an excellent review is in Campbell, Lo, and MacKinlay (1997, Ch. 4)]. The traditional parametric approach is based on external studentized residuals. However, a non-parametric approach based on Wilcoxon-type rank sums, suggested by Corrado (1989) for testing single period returns, and extended in Campbell and Wasley (1993) for testing cumulative multi-day rank sums, is gaining increasingly popularity. The used test statistics are obviously of the type of internally studentized residuals where the nominator and denominator are not independent. Thus, the distribution theory developed in this paper can be readily utilized in examination of the distributional properties of the test statistics in these cases.

Regarding externally studentized residuals in which the ratio of residual and the estimator of standard deviation are independent, we find that

$$t_m = r_m \left(\frac{n-p-1}{n-p-r_m^2} \right)^{\frac{1}{2}} = \frac{\boldsymbol{m}'\hat{\boldsymbol{u}}/\sqrt{\boldsymbol{m}'\mathbf{Q}\boldsymbol{m}}}{\sqrt{\frac{1}{n-p-1} V_m}}, \tag{22}$$

where $r_m = \boldsymbol{m}'\hat{\boldsymbol{u}}/s\sqrt{q_m}$, $q_m = \boldsymbol{m}'\mathbf{Q}\boldsymbol{m}$, and $V_m = \hat{\boldsymbol{u}}'\hat{\boldsymbol{u}} - (\boldsymbol{m}'\hat{\boldsymbol{u}})^2/(\boldsymbol{m}'\mathbf{Q}\boldsymbol{m})$, is a t-distributed random variable with $n-p-1$ degrees of freedom by virtue of Lemma 1, Lemma 2, and $\boldsymbol{m}'\hat{\boldsymbol{u}}/\sqrt{\boldsymbol{m}'\mathbf{Q}\boldsymbol{m}} \sim N(0,\sigma^2)$. Again, a special case of this is a single residual, in which case the right-most expression of (22) can be written as

$$t_i = \frac{\hat{u}_i}{s_{(i)}\sqrt{q_{ii}}} \tag{23}$$

where

$$s^2_{(i)} = \frac{(n-p)s^2 - \hat{u}_i^2/q_{ii}}{n-p-1} \tag{24}$$

is the residual means square from a sample with the ith observation removed from the regression [Beckman and Trussell (1974)]. By equation (22), the relationship between the externally studentized residual, t_i, and the internally studentized residual, r_i, given in equation (6), is

$$t_i = r_i \left(\frac{n-p-1}{n-p-r_i^2} \right)^{\frac{1}{2}} \tag{25}$$

[c.f. Cook and Weisberg (1980, p. 20)]. The end result is that an arbitrary (nonsingular) linear function of internally studentized residuals can be transformed to a t-statistic with $n-p-1$ degrees of freedom. Thus, utilizing the simple transformation in (22), rather than relying on the Beta distribution the familiar t-distribution can be used instead in the related statistical inference.

Finally, we note that for the multivariate applications Lemma 1 and Corollary 2 imply that

$$t^2_M = \frac{\hat{\boldsymbol{u}}'\mathbf{M}'(\mathbf{MQM}')^{-1}\mathbf{M}\hat{\boldsymbol{u}}/m}{\hat{\boldsymbol{u}}'\left(\mathbf{I} - \mathbf{M}'(\mathbf{MQM}')^{-1}\mathbf{M}\right)\hat{\boldsymbol{u}}/(n-p-m)} \tag{26}$$

follows the F-distribution with m and $n-p-m$ degrees of freedom. This is a generalization of the test statistic for testing multiple outlying cases in Cook Weisberg (1982, p. 30).

A multivariate analog to t_m, defined in (22), can be forumlated in terms of

$$r^2_M = \frac{\hat{\boldsymbol{u}}'\mathbf{M}'(\mathbf{MQM}')^{-1}\mathbf{M}\hat{\boldsymbol{u}}}{s}, \tag{27}$$

such that we can rewrite (26) as

$$t_M^2 = \frac{r_M^2(n-p-m)}{m(n-p-r_M^2)}. \tag{28}$$

4. Conclusions

This paper derives the joint distribution of a general linear transformation of internally studentized least squares residuals from a linear regression. Other types of scaled residuals, commonly used in practical applications, can be easily obtained as special cases by defining the linear transformation appropriately. The distributions of arbitrary subsets as well as marginal distributions of single residuals are obtained as special cases from the general distribution by defining the linear transformation in a suitable manner. The paper discusses also some potential applications in which the results can be readily applied.

References

Abrahamse, A.P.J. and J. Koerts (1971). New estimators in regresion analysis. *Journal of the American Statistical Associstion* 66, 71–74.

Beckman, R.J. and H.J. Trussell (1974). The distribution of an arbitrary studentized residual and the effects of updating in multiple regression. *Journal of the American Statistical Association* 69, 199–201.

Campbell, Cynthia, J. and Charles E. Wasley (1993). Measuring security price performance using NASDAQ returns. *Journal of Financial Economics* 33, 73–92.

Campbell, John Y., Andrew W. Lo, and A. Craig MacKinlay (1997). *The Econometrics of Financial Markets.* Princeton, NJ: Princeton University Press.

Chatterjee, Samprit and Ali S. Hadi (1988). *Sensitivity Analysis in Linear Regression*, Wiley, New York.

Corrado, Charles, J. (1989). A nonparametric test for abnormal security-price performance in event studies. *Journal of Financial Economics* 23, 385–395.

Cook, Dennis R. and Sanford Weisberg (1982). *Residuals and Influence in Regression*. Chapman and Hall.

Díaz-García, José A. and Ramón Gutiérrez-Jáimez (2007). The distribution of residuals from a general elliptical linear model. *Journal of Statistical Planning and Inference* 137, 2347–2354.

Ellenberg, Jonas K. (1973). The joint distribution of the standardized least squares residuals from a general linear regression. *Journal of the American Statistical Association* 68, 941–943.

Loynes, R.M. (1979). A note on Prescott's upper bound for normed residuals. *Biometrika* 66, 387–389.

Margolin, Barry H. (1977). The distribution of internally studentized statistics via Laplace transform inversion. *Biometrika* 64, 573–582.

Pynnönen, Seppo (2010). Joint distribution of a linear transformation of OLS regression residuals with general spherical error distributio. *Working Paper, department of Mathematics and Statistics, University of Vaasa*

Stefansky, Wilhelmine (1972). Rejecting outliers in fractional designs. *Technometrics* 14. 469–479.