Econometric Theory, 7, 1991, 253–263. Printed in the United States of America.

NONUNIFORM BOUNDS FOR NONPARAMETRIC t-TESTS

JEAN-MARIE DUFOUR
Université de Montréal

MARC HALLIN
Université Libre de Bruxelles

This paper gives simple nonuniform bounds on the tail areas of the permutation distribution of the usual Student's t-statistic when the observations are independent with symmetric distributions. As opposed to uniform bounds, nonuniform bounds depend on the observed sample. It is shown that the nonuniform bounds proposed are always tighter than uniform exponential bounds previously suggested. The use of the bounds to perform nonparametric t-tests is discussed and numerical examples are presented. Further, the bounds are extended to t-tests in the context of a simple linear regression.

1. INTRODUCTION

Let $X_1, \ldots, X_n$ be independent random variables with symmetric distributions and consider the problem of testing $H_0: \text{med}(X_i) = 0,\ i = 1, \ldots, n$. The most widely used test statistic for a problem of this type is $T_n = \sqrt{n}\bar{X}/s$, where $\bar{X} = \sum_{i=1}^{n} X_i/n$ and $s^2 = \sum_{i=1}^{n} (X_i - \bar{X})^2/(n-1)$; when $X_1 = \cdots = X_n = 0$, we set $T_n = 0$. The finite-sample distribution of T_n depends heavily on the probability laws of $X_1, \ldots, X_n$. It is usually assumed that $X_1, \ldots, X_n$ are independent with identical normal distributions. However, even though the distribution of T_n could be derived under other special assumptions, the distributions of the observations are usually unknown and, in particular, may not be identical, for example, heteroskedastic observations. In this context, a result that can be especially useful is the inequality

$$P(T_n > y) \le \exp\{-g_n(y)^2/2\} \tag{1}$$

where $y \ge 0$ and

$$g_n(y) = n^{1/2}y/[(n-1) + y^2]^{1/2}; \tag{2}$$

see Edelman [10]. This inequality holds whenever $X_1, \ldots, X_n$ are independent and symmetric about zero, even if the observations are not identically distributed or do not have finite moments. It provides a simple way of find-

This work was supported by the Natural Sciences and Engineering Research Council of Canada, the Social Sciences and Humanities Research Council of Canada, and the Fonds FCAR (Government of Quebec). The authors thank Bryan Campbell, Marcel Dagenais, David Edelman, Constance van Eeden, Pierre Perron, Peter Robinson, Michael Wickens, and two anonymous referees for several useful comments.

 0266-4666/91 $5.00 + .00

ing conservative critical values under weak distributional assumptions (i.e., critical values ensuring that the probability of rejection under the null hypothesis is not larger than the nominal level of the test; for further discussion of conservative and liberal statistical procedures, see Dufour [8,9]). $X_1, \ldots, X_n$ may even be discrete random variables. Further, if we replace X_i by $X_i - \mu$ where $\mu \neq 0$, we can consider in the same way any hypothesis of the form $H_0': \mathrm{med}(X_i) = \mu$, $i = 1, \ldots, n$. For related work on the distribution of t-statistics under non-normality, see Benjamini [2] and the survey of Cressie [7].

Many testing problems in statistics and econometrics can be written in the form $H_0: \mathrm{med}(X_i) = 0$, $i = 1, \ldots, n$, where $X_1, \ldots, X_n$ are independent random variables with symmetric distributions. (a) For example, in the random walk model

$$Y_i - Y_{i-1} = \delta + u_i, \qquad i = 1, \ldots, n,$$

where $u_1, \ldots, u_n$ are independent (but not identically distributed) and follow distributions symmetric about zero that are more heavy tailed than the normal (e.g., Cauchy distributions), we may wish to test whether the drift coefficient is zero ($H_0: \delta = 0$, with $X_i \equiv Y_i - Y_{i-1}$, $i = 1, \ldots, n$). (b) If $\hat{Y}_i$ is a forecast of Y_i based on survey data (e.g., as in Friedman [13]), we may wish to test whether the forecast errors $X_i \equiv Y_i - \hat{Y}_i$ are unbiased (as an implication of rational expectations) without assuming that they are identically or normally distributed. (c) If $Y_1, \ldots, Y_n$ is a time series of independent random variables with distributions symmetric about zero while $Z_1, \ldots, Z_n$ is any other (possibly autocorrelated) time series with median zero, we can test independence between the two series against an alternative of dependence at lag k by taking $X_i = Y_i Z_{i+k}$ (where $i = 1, \ldots, n - k$, when $k \geq 0$, and $i = |k| + 1, \ldots, n$, when $k \leq -1$). This is easy to see by considering the conditional distribution of the X_i's given $(Z_1, \ldots, Z_n)'$: under the assumption of independence between $(Y_1, \ldots, Y_n)'$ and $Z = (Z_1, \ldots, Z_n)'$, the X_i's are mutually independent conditional on Z, and $\mathrm{med}(X_i | Z) = 0$ for each i. (d) Finally, it is possible to show that bounds similar to (1) can also be applied to t-statistics obtained in the more general context of a simple linear regression with one regressor; see Section 4.

The inequality (1) is proved by considering the sign-permutation (or sign-randomization) distribution of T_n, that is, the distribution of T_n conditional on $|X| = (|X_1|, |X_2|, \ldots, |X_n|)'$. From the theorem given by Edelman [10] one sees easily that

$$P(T_n > y \mid |X|) \leq \exp\{-g_n(y)^2/2\} \tag{3}$$

for any $y \geq 0$ and $|X| \neq 0$. Clearly, the inequality (1) follows from (3). Thus tests based on inequality (1) can be viewed as sign-permutation tests (Fisher [11]) where the critical value does not depend on $|X|$: one replaces the critical value based on the conditional distribution of T_n given $|X|$—a point which is a function of $|X|$ — with a *uniform upper bound* that does not depend on the observed sample. Any test with level α when $X_1, \ldots, X_n$ are independent with distributions symmetric about zero must be conditional on $|X|$; see Pratt and Gibbons [19, p. 218]. Further, among the conditional tests, the one-sided sign-permutation t-test $T_n > c(|X|)$ is most powerful against a normal alternative with $\mu > 0$ [and similarly for $T_n < c(|X|)$ against $\mu < 0$], while the corresponding two-sided test is most powerful unbiased as well as most stringent similar against $\mu \neq 0$; see Lehmann and Stein [18], Hoeffding [15, p. 176], Pratt and Gibbons [19, pp. 233–234], and Lehmann [17, chapters 5 and 6]. The sign-permutation t-test is a nonparametric test that takes into account the information contained in the signs of the observations as well as their absolute values $|X|$, while other available nonparametric tests of H_0 only consider the signs, for example, the sign test, and (possibly) information about the ranking of $|X_1|, \ldots, |X_n|$, for example, linear signed rank tests. Among nonparametric tests of H_0, the sign-permutation t-test has the best power when $X_1, \ldots, X_n$ are i.i.d. $N(\mu, \sigma^2)$.

The permutation distribution of T_n depends on $|X|$, that is, it varies with each sample studied, and thus no useful tabulation is possible. This distribution is typically established by enumeration. Since the number of cases to consider grows like 2^n, it becomes rapidly too expensive to establish. Consequently, a bound like (3) can simplify considerably the execution of the test. Further, even though asymptotic normal approximations are available for such tests (see Hoeffding [15], Fraser [12, pp. 280–283], Albers, Bickel, and van Zwet [1]), these require regularity conditions that are not met in the setup considered here. Unlike asymptotic approximations, Edgeworth expansions (see Albers, Bickel, and van Zwet [1]) or saddlepoint approximations (Robinson [20]), which can be too liberal (i.e., can lead to tests that reject too often), a bound like (3) is *certainly* conservative and thus tests based on it never reject too often under the null hypothesis.

In this paper, we derive *nonuniform bounds* that can be implemented easily. Nonuniform bounds take $|X|$ into account and thus can be much tighter than uniform bounds. In particular, we describe simple nonuniform bounds that dominate in all cases the uniform bound based on inequality (1). While they still have size not greater than the nominal level of the test, both conditional on $|X|$ and unconditionally, the critical regions yielded by the nonuniform bounds contain the critical region based on the uniform bound. One reason for this is that nonuniform bounds make adjustments for the form of the underlying distribution. We show also that the improvement obtained in this way can be considerable, especially when the observations follow heavy-tailed distributions. In addition, we show that similar bounds are applicable to t-tests in the more general context of linear regressions with one regressor.

In Section 2, we give the basic propositions on which these results are based. In Section 3, we discuss their use and give some numerical examples illustrating the advantages of the bounds suggested. Finally, in Section 4, we describe the generalization to simple linear regressions.

2. BASIC PROPOSITIONS

The following theorem is obtained by improving a result given by Edelman [10] and Hoeffding [16, Theorem 2].

THEOREM 1. *Let* $X_1, \ldots, X_n$ *be independent random variables with distributions symmetric about zero,*

$$S_n = \frac{\sum_{j=1}^{n} X_j}{\left(\sum_{i=1}^{n} X_i^2\right)^{1/2}} \quad \text{and} \quad w_j = \frac{|X_j|}{\left(\sum_{i=1}^{n} X_i^2\right)^{1/2}}, \qquad j = 1, \ldots, n,$$

where we set $w_j = 0$ *and* $S_n = 0$ *when* $X_1 = \cdots = X_n = 0$. *Then, for all* $y \geq 0$,

$$P[S_n \geq y \mid |X|] \leq \overline{M}(y, |X|) \leq M(y, y, |X|) \leq e^{-y^2/2} \tag{4}$$

where

$$M(t, y, |X|) = e^{-ty} \prod_{j=1}^{n} \cosh(w_j t), \qquad \overline{M}(y, |X|) = \inf_{t \geq 0} M(t, y, |X|) \tag{5}$$

and $\cosh(x) = (e^x + e^{-x})/2$. *Further, for* $y > 0$,

$$\begin{aligned} \overline{M}(y, |X|) &= 0, && \text{if } \sum_{j=1}^{n} w_j < y, \\ &= \left(\frac{1}{2}\right)^n, && \text{if } \sum_{j=1}^{n} w_j = y, \\ &= M(t^*, y, |X|), && \text{if } \sum_{j=1}^{n} w_j > y, \end{aligned} \tag{6}$$

where t^* *is the unique value of* t *such that* $t^* > 0$ *and*

$$\sum_{j=1}^{n} w_j \left\{ \frac{[1 - \exp(-2w_j t^*)]}{[1 + \exp(2w_j t^*)]} \right\} = y. \tag{7}$$

Proof. Since the theorem holds trivially when $X_1 = \cdots = X_n = 0$, we will suppose that $X_j \neq 0$ for at least one j. We can write

$$S_n = \sum_{j=1}^{n} w_j s(X_j) = \sum_{j \in A(|X|)} w_j s(X_j),$$

where $s(x)$ is the sign function [$s(x) = -1$ if $x < 0$, $s(x) = 0$ if $x = 0$, and $s(x) = 1$ if $x > 0$] and $A(|X|) = \{j : |X_j| \neq 0,\ 1 \leq j \leq n\}$. Since each X_j has a distribution symmetric about zero, though possibly with a probability mass at zero, we have for $j \in A(|X|)$

$$P[s(X_j) = -1 \mid |X_j|] = P[s(X_j) = 1 \mid |X_j|] = \tfrac{1}{2}. \tag{8}$$

Hence, since $X_1, \ldots, X_n$ are mutually independent, the variables in the set $\{s(X_j) : j \in A(|X|)\}$ are mutually independent conditional on $|X|$ and satisfy (8) with $|X_j|$ replaced by $|X|$. Thus, conditional on $|X|$, the moment generating function of S_n is

$$E[\exp(tS_n) \mid |X|] = \prod_{j \in A(|X|)} \cosh(w_j t) = \prod_{j=1}^{n} \cosh(w_j t),$$

where the last identity follows by observing that $\cosh(w_j t) = \cosh(0) = 1$ for $j \notin A(|X|)$. By Markov's inequality,

$$\begin{aligned} P[S_n \geq y \mid |X|] &\leq P[\exp(tS_n) \geq \exp(ty) \mid |X|] \leq E[\exp(tS_n) \mid |X|]/\exp(ty) \\ &= \left[\prod_{j=1}^{n} \cosh(w_j t)\right] \Big/ e^{ty} = M(t, y, |X|) \end{aligned} \tag{9}$$

for all $t \geq 0$ and for all y. Consequently, for all $y \geq 0$,

$$\begin{aligned} P[S_n \geq y \mid |X|] &\leq \inf_{t \geq 0} \left[e^{-ty} \prod_{j=1}^{n} \cosh(w_j t)\right] \\ &\leq e^{-y^2} \left[\prod_{j=1}^{n} \cosh(w_j y)\right] \leq e^{-y^2/2}, \end{aligned}$$

where the second inequality follows by taking $t = y$ in (9), while the third one follows from the inequality $\cosh(x) \leq \exp(x^2/2)$; see Edelman [10]. Finally, to derive expression (6), consider the function $h(t) \equiv \ln[M(t, y, |X|)]$ and its derivative

$$h'(t) = \left\{\sum_{j=1}^{n} w_j \frac{1 - \exp(-2w_j t)}{1 + \exp(2w_j t)}\right\} - y \tag{10}$$

where $y > 0$ and $w_j \neq 0$ for at least one j. It is easy to see that the function $h'(t)$ is continuous everywhere, strictly increasing for $t > 0$, with $h'(0) = -y < 0$ and $h'(t) \to \sum_{j=1}^{n} w_j - y$ as $t \to \infty$. Consequently, if $\sum_{j=1}^{n} w_j > y$, the equation $h'(t^*) = 0$ has a unique solution such that $t^* > 0$, with $h'(t) < 0$ for $0 \leq t < t^*$ and $h'(t) > 0$ for $t > t^*$. Thus, $h(t^*)$ is the minimum of $h(t)$ over the interval $t \geq 0$ or, equivalently, $\overline{M}(y, |X|) = M(t^*, y, |X|)$ where t^* is the solution of equation (7). If $\sum_{j=1}^{n} w_j \leq y$, we have $h'(t) < 0$ for all $t > 0$, so that $h(t)$ and $M(t, y, |X|)$ have no minimum. In this case, it is easy to verify directly that $\overline{M}(y, |X|) \equiv \lim_{t \to \infty} M(t, y, |X|) = 0$ if $\sum_{j=1}^{n} w_j < y$, and $\overline{M}(y, |X|) = (1/2)^n$ if $\sum_{j=1}^{n} w_j = y$. ∎

For the case considered here, the two nonuniform bounds $\overline{M}(y,|X|)$ and $M(y,y,|X|)$ improve the uniform bound $e^{-y^2/2}$ given by Edelman [10] and Hoeffding [16, Theorem 2]. Note also that the latter improves, in turn, other inequalities given by Bernstein ([4], [5, pp. 159–165]) (see also Uspenky [21, pp. 204–206] and Godwin [14, p. 936]), Craig [6], Bennett [3, (7a) and (8a)] and Hoeffding [16, (2.2) and (2.3)].

Using the transformation $S_n = n^{1/2}T_n/[(n-1) + T_n^2]^{1/2}$, we get the following important corollary.

COROLLARY 1. *Under the assumptions of Theorem* 1,

$$P[T_n \geq y \mid |X|] \leq \overline{M}[g_n(y),|X|] \leq M[g_n(y),g_n(y),|X|] \leq \exp[-g_n(y)^2/2] \tag{11}$$

for all $y \geq 0$, *where the function* $g_n(y)$ *is defined by* (2). ■

Clearly, the inequality (11) implies the inequalities (1) and (3). Note here that, even though the condition $y \geq 0$ is not explicitly stated by Edelman [10], it is straightforward to see that (1) is not generally true for $y < 0$. Corollary 1 provides two improvements over the inequalities (1) and (3). We now discuss their use.

3. APPLICATION

For a two-sided test of H_0 with level α $(0 < \alpha < 1)$, using the critical values suggested by Edelman [10] is equivalent to rejecting H_0 when

$$\exp[-g_n(|T_n|)^2/2] \leq \alpha/2. \tag{12}$$

Similar equivalences hold for one-sided tests. However, by Corollary 1, tighter bounds are available. In particular, one can say that the sign-permutation t-test of H_0 is significant at level α when

$$M[g_n(|T_n|),g_n(|T_n|),|X|] \leq \alpha/2. \tag{13}$$

The critical region (13) has size not greater than α, both conditional on $|X|$ and unconditionally. If we view the critical regions defined by (12) and (13) as subsets of the sample space $\mathbb{R}^n$ for X, the critical region (12) is a subset of the one given by (13), so that (13) is less conservative at any given level and the effective power of the test based on $M[g_n(|T_n|),g_n(|T_n|),|X|]$ is greater than the one based on $\exp[-g_n(|T_n|)^2/2]$. An intuitive reason for this is that the weights w_j take into account the form of the underlying distribution. For example, if the sample contains one large outlier, say X_1, then $w_1 \simeq 1$, $w_j \simeq 0$ for $j \neq 1$ and $\prod_{j=1}^n \cosh(w_j y) \simeq \cosh(y)$. This is especially likely when the observations follow heavy-tailed distributions or when a few observations have big variances. As will be seen below, the function $\prod_{j=1}^n \cosh(w_j y)$ and the bound $M(y,y,|X|)$ are then relatively small for any given y. The opposite happens when the weights are approximately equal. An even larger critical region of size not greater than α is given by

$$\inf_{t\geq 0} M[t, g_n(|T_n|), |X|] \leq \alpha/2. \tag{14}$$

The condition $M[t, g_n(|T_n|), |X|] \leq \alpha/2$, for some $t \geq 0$, also defines a critical region of size not greater than α. For $0 < g_n(|T_n|) < \sum_{j=1}^n w_j$, which is the most probable case (because $\sum_{j=1}^n w_j$ is the largest possible value of S_n), it is difficult to find a closed-form expression for the infimum of $M[t, g_n(|T_n|), |X|]$. The latter, however, can be minimized easily by numerical methods. This can be done by solving equation (7) for t^*, with y replaced by $g_n(|T_n|)$, or equivalently by finding $t^* > 0$ such that $h'(t^*) = 0$, where $h'(t)$ is given by (10) with $y = g_n(T_n)$. Since $h'(t) < 0$ for $0 \leq t < t^*$ and $h'(t) > 0$ for $t > t^*$, an especially simple way to proceed is to find first some $\bar{t}$ such that $h'(\bar{t}) > 0$ and then use a bisection algorithm over the interval $(0, \bar{t})$ to determine t^*.

To see that the p values used in (13) and (14) can be substantially smaller than the one in (12), consider the case where the sample contains one large outlier, say X_1. By letting $|X_1|$ become large relative to the other observations, that is, $w_1 \to 1$, we see that

$$M[g_n(y), g_n(y), |X|] \to LM[g_n(y)] \equiv \exp[-g_n(y)^2]\cosh[g_n(y)],$$

$$\overline{M}[g_n(y), |X|] \to \overline{LM}[g_n(y)] \equiv \inf_{t\geq 0} \{\exp[-tg_n(y)]\cosh(t)\}.$$

It is easy to check that $\overline{LM}[g_n(y)] = 0$ whenever $y > 1$: thus, by varying $|X|$, $\overline{M}[g_n(y), |X|]$ can take values arbitrarily close to zero for $y > 1$. Similarly, depending on the sample, $M[g_n(y), g_n(y), |X|]$ can take values arbitrarily close to $LM[g_n(y)]$. In Table 1 (part A), we report some numerical comparisons between $LM[g_n(y)]$ and $\exp[-g_n(y)^2/2]$ for $n = 20, 40$, and $y = 1.0, 1.5, 2.0, 2.5, 3.0, 3.5, 4.0$. It is clear from these numbers that $LM[g_n(y)]$ can be much smaller than $\exp[-g_n(y)^2/2]$.

In the same table (part B), we also present some illustrative computations of p values based on samples generated by Monte Carlo methods. Two sample sizes ($n = 20, 40$) and three probability laws were considered: the $N(0.5, 1)$ distribution, the Cauchy distribution with median 1.5, and a normal distribution with mean 1.5, variance 1 for one-half of the sample and variance 16 for the other half. For each of these cases, the t statistics of two different samples are presented. It is clear from these examples that the nonuniform bounds can provide substantial improvements over the uniform bound $\exp[-g_n(|T_n|)^2/2]$. The difference is especially important for the Cauchy samples, which have heavy tails.

To illustrate further the advantage of using the improved bounds suggested in this paper, we present in part C of Table 1 some power comparisons for

TABLE 1. Comparisons between alternative bounds

A	n	$y = 1.0$	$y = 1.5$	$y = 2.0$	$y = 2.5$	$y = 3.0$	$y = 3.5$	$y = 4.0$
$\exp[-g_n(y)^2/2]$	20	0.6075	0.3469	0.1757	0.0841	0.0402	0.0198	0.0103
	40	0.6075	0.3359	0.1556	0.0631	0.0235	0.0084	0.0030
$LM[g_n(y)]$	20	0.5677	0.2718	0.1020	0.0331	0.0103	0.0033	0.0011
	40	0.5677	0.2600	0.0851	0.0211	0.0043	0.0008	0.0001

B p-values for selected samples	Normal[a] $\mu = 0.5$, $\sigma = 1$				Cauchy[a] $\mu = 1.5$				Normal with heteroskedasticity[a] $\mu = 1.0$			
	$n = 20$		$n = 40$		$n = 20$		$n = 40$		$n = 20$		$n = 40$	
	Sample 1	Sample 2	Sample 1	Sample 2	Sample 1	Sample 2	Sample 1	Sample 2	Sample 1	Sample 2	Sample 1	Sample 2
T_n	3.071	3.414	4.106	3.796	2.142	2.140	2.249	3.040	2.374	1.898	3.521	4.038
$\exp[-g_n(\|T_n\|)^2/2]$	0.0363	0.0223	0.0024	0.0045	0.1429	0.1434	0.1006	0.0217	0.1015	0.2033	0.0080	0.0028
$M[g_n(\|T_n\|), g_n(\|T_n\|), \|X\|]$	0.0245	0.0124	0.0012	0.0026	0.0958	0.0821	0.0799	0.0107	0.0782	0.1840	0.0050	0.0013
$M[g_n(\|T_n\|), \|X\|]$	0.0171	0.0049	0.0006	0.0018	0.0139	0.00003	0.0689	0.0020	0.0646	0.1781	0.0036	0.0006

C Effective power comparisons[b]	$\mu = 0.25$			$\mu = 0.50$			$\mu = 0.75$		
Test based on	$n = 20$	$n = 40$	$n = 80$	$n = 20$	$n = 40$	$n = 80$	$n = 20$	$n = 40$	$n = 80$
$\exp[-g_n(\|T_n\|)^2/2]$	0.023	0.089	0.289	0.161	0.614	0.946	0.535	0.957	1.00
$\underline{M}[g_n(\|T_n\|), g_n(\|T_n\|), \|X\|]$	0.041	0.109	0.304	0.224	0.654	0.953	0.626	0.965	1.00
$\overline{M}[g_n(\|T_n\|), \|X\|]$	0.052	0.115	0.305	0.274	0.667	0.953	0.683	0.966	1.00

[a] The normal samples have mean 0.5 and variance 1. The Cauchy samples have median 1.5. The heteroskedastic normal samples have mean 1.0, variance 1 for observations $j = 1, \ldots, n/2$ and variance 16 for observations $j = (n/2) + 1, \ldots, n$. For each n, two different samples are considered.

[b] Based on normal samples with mean μ and variance 1. One thousand replications were used.

tests of level 0.05 against normal alternatives with three different means ($\mu = 0.25, 0.50, 0.75$) and three sample sizes ($n = 20, 40, 80$). The frequencies of rejection were computed by a Monte-Carlo experiment with 1000 replications. As expected from our theoretical results, we see that the frequencies of rejection are always larger when the nonuniform bounds are used (as opposed to the uniform bounds), sometimes by sizable margins (e.g., for $\mu = 0.50$ and $n = 20, 40$).

4. GENERALIZATION

Finally, it is interesting to note that the bounds described above can also be applied to simple linear regressions. Consider the model

$$y_i = x_i\beta + u_i, \qquad i = 1, \ldots, n \tag{15}$$

where $x_1, \ldots, x_n$ are (fixed) scalars not all equal to zero, β is an unknown parameter, and $u_1, \ldots, u_n$ are independent random disturbances with distributions symmetric about zero ($n \geq 2$). We wish to test $H_0: \beta = \beta_0$.

The standard t statistic for H_0 is

$$t_n = \frac{\hat{\beta} - \beta_0}{s\left(\sum_{i=1}^{n} x_i^2\right)^{-1/2}} \tag{16}$$

where

$$\hat{\beta} = \left(\sum_{i=1}^{n} x_i^2\right)^{-1} \sum_{i=1}^{n} x_i y_i, \qquad s^2 = \frac{\sum_{i=1}^{n} \hat{u}_i^2}{n-1} \quad \text{and}$$

$$\hat{u}_i = y_i - x_i\hat{\beta}, \qquad i = 1, \ldots, n;$$

when $s = 0$, we set $t_n = 0$. Under H_0, we have

$$\left(\sum_{i=1}^{n} x_i^2\right)^{1/2} (\hat{\beta} - \beta_0) = \left(\sum_{i=1}^{n} x_i^2\right)^{-1/2} \left(\sum_{i=1}^{n} x_i u_i\right) = \left(\sum_{i=1}^{n} x_i^2\right)^{-1/2} \left(\sum_{i=1}^{n} x_i^2 u_i^2\right)^{1/2} S_n \tag{17}$$

where

$$S_n = \left(\sum_{i=1}^{n} x_i^2 u_i^2\right)^{-1/2} \left(\sum_{i=1}^{n} x_i u_i\right) \tag{18}$$

and $u_i = y_i - x_i\beta_0$, $i = 1, \ldots, n$, with $S_n \equiv 0$ when $x_1 u_1 = \cdots = x_n u_n = 0$. Using the identity

$$\sum_{i=1}^{n} u_i^2 = \sum_{i=1}^{n} \hat{u}_i^2 + \left(\sum_{i=1}^{n} x_i^2\right) (\hat{\beta} - \beta_0)^2,$$

we get from (16) and (17)

$$t_n = \frac{\sqrt{n-1}\left(\sum_{i=1}^{n} x_i^2\right)^{-1/2}\left(\sum_{i=1}^{n} x_i^2 u_i^2\right)^{1/2} S_n}{\left[\sum_{i=1}^{n} u_i^2 - \left(\sum_{i=1}^{n} x_i^2\right)^{-1}\left(\sum_{i=1}^{n} x_i^2 u_i^2\right) S_n^2\right]^{1/2}} \tag{19}$$

and

$$S_n = Ct_n/[(n-1) + t_n^2]^{1/2} \equiv h_n(t_n, C) \tag{20}$$

where

$$C = \left[\left(\sum_{i=1}^{n} x_i^2\right)\left(\sum_{i=1}^{n} u_i^2\right)\Big/\left(\sum_{i=1}^{n} x_i^2 u_i^2\right)\right]^{1/2},$$

with $Ct_n \equiv 0$ when $x_1 u_1 = \cdots = x_n u_n = 0$. Under H_0 and conditional on $|u| = (|u_1|, \ldots, |u_n|)'$, t_n is a monotonically increasing transformation of S_n. Further, S_n satisfies the conditions of Theorem 1 with $X_i = x_i u_i$, $i = 1, \ldots, n$. Thus

$$\begin{aligned} P[t_n \geq y \mid |X|] &\leq \overline{M}[h_n(y,C), |X|] \leq M[h_n(y,C), h_n(y,C), |X|] \\ &\leq \exp[-h_n(y,C)^2/2], \end{aligned} \tag{21}$$

where $|X| = (|x_1 u_1|, \ldots, |x_n u_n|)'$ and $u_i = y_i - x_i \beta_0$.

When $x_1 = \cdots = x_n = 1$, $h_n(y, C) = g_n(y)$ and (21) reduces to (11). Note also that all three bounds in (21), including the largest one, are now nonuniform because C is a function of $|X|$. It is clear that the bounds in (21) can be used in a way similar to those in (11).

REFERENCES

1. Albers, W., P.J. Bickel & W.R. van Zwet. Asymptotic expansions for the power of distribution free tests in the one-sample problem. *The Annals of Statistics* 4 (1976): 108–156.
2. Benjamini, Y. Is the t test really conservative when the parent distribution is long-tailed? *Journal of the American Statistical Association* 78 (1983): 645–654.
3. Bennett, G. Probability inequalities for the sum of independent random variables. *Journal of the American Statistical Association* 57 (1962): 33–45.
4. Bernstein, S. Sur une modification de l'inégalité de Tchebichef. *Ann. Sc. Instit. Sav. Ukraine, Sect. Math.* I (1924): 38–49 (Russian, French summary).
5. Bernstein, S. *Theory of Probability*. Moscow, 1927.
6. Craig, C.E. On the Tchebychef inequality. *Annals of Mathematical Statistics* 4 (1933): 94–102.
7. Cressie, N. Relaxing assumptions in the one sample t-test. *Australian Journal of Statistics* 22 (1980): 143–153.
8. Dufour, J.-M. Nonlinear hypotheses, inequality restrictions, and non-nested hypotheses: exact simultaneous tests in linear regressions. *Econometrica* 57 (1989): 335–355.
9. Dufour, J.-M. Exact tests and confidence sets in linear regressions with autocorrelated errors. *Econometrica* 58 (1990): 475–494.
10. Edelman, D. Bounds for a nonparametric t table. *Biometrika* 73 (1986): 242–243.

11. Fisher, R. *The Design of Experiments.* Edinburgh: Oliver and Boyd, 1935.
12. Fraser, D.A.S. *Nonparametric Methods in Statistics.* New York: Wiley, 1957.
13. Friedman, B.M. Survey evidence on the 'rationality' of interest rate expectations. *Journal of Monetary Economics* 6 (1979): 453–465.
14. Godwin, H.J. On generalizations of Tchebychef's inequality. *Journal of the American Statistical Association* 50 (1955): 923–945.
15. Hoeffding, W. The large-sample power of tests based on permutations of observations. *Annals of Mathematical Statistics* 23 (1952): 169–192.
16. Hoeffding, W. Probability inequalities for sums of bounded random variables. *Journal of the American Statistical Association* 58 (1963): 13–30.
17. Lehmann, E.L. *Testing Statistical Hypotheses, Second Edition.* New York: Wiley, 1986.
18. Lehmann, E.L. & C. Stein. On the theory of some non-parametric hypotheses. *Annals of Mathematical Statistics* 20 (1949): 28–45.
19. Pratt, J.W. & J.D. Gibbons. *Concepts of Nonparametric Theory.* New York: Springer-Verlag, 1981.
20. Robinson, J. Saddlepoint approximations for permutation tests and confidence intervals. *Journal of the Royal Statistical Society, Series B* 44 (1982): 91–101.
21. Uspenky, J. *Introduction to Mathematical Probability.* New York: McGraw-Hill, 1937.