Is the t Test Really Conservative When the Parent Distribution is Long-Tailed?

Yoav Benjamin

Journal of the American Statistical Association, Vol. 78, No. 383 (Sep., 1983), 645-654.

Is the *t* Test Really Conservative When the Parent Distribution is Long-Tailed?

YOAV BENJAMINI*

It is generally believed that the *t* test is conservative for a sample from a long-tailed symmetric distribution. Yet the probability inequalities expressing this property have not been proved. The inequalities are explored here using various criteria for long-tailedness and leaning on the geometrical interpretation of the *t* test. It is proved that the *t* test is conservative but only for large enough critical values. Examples of a liberal *t* test for lower values are given. The results are used to explain some curiosities in the asymptotic distribution of the *t* statistic and to study its behavior when the parent distribution is skewed.

KEY WORDS: *t* test; Long-tailed distributions; Probability inequality; Conservatism; Scale mixture of normals.

1. INTRODUCTION

It has long been part of statistical folklore that the *t* test is conservative for a sample from a long-tailed symmetric distribution. This conservatism means the type I error of the *t* test is smaller under a long-tailed symmetric distribution than it would be under normality. Formally, this statement translates to the following probability inequality:

Let $>_t$ be a (partial) ordering of distribution functions that conveys the notion of having longer tails. Let $\mathbf{x} = (x_1, \ldots, x_n)$ be a random sample, and let $T_n(\mathbf{x})$ be its *t* statistic. Let F be a distribution symmetric about 0, and let Φ be a standard normal distribution.

If $F >_t \Phi$, then $\alpha = P_\Phi(T_n(\mathbf{x}) > t_\alpha) \geq P_F(T_n(\mathbf{x}) > t_\alpha)$.

Knowledge of the distribution of $T_n(\mathbf{x})$ for nonnormal F would allow comparison of the F-based with the Φ-based tail probabilities. Student himself, in the same paper in which the *t* distribution was introduced, first considered this problem in 1908. Since then much effort has been devoted to determining the exact and approximate distribution of $T_n(\mathbf{x})$ under nonnormal distributions. The literature describing this effort has been reviewed by Johnson (1978) and Cressie (1980). The works of Hotelling (1961) and Efron (1969) differ markedly from others in their use of geometrical arguments. Since my results shed light on those of Hotelling and Efron, their results are discussed in Section 9.

The evidence collected over the years indicates the following (Johnson 1978):

1. A long-tailed parent distribution causes $T_n(\mathbf{x})$ to be less dispersed.
2. A distribution that is skewed to the right causes $T_n(\mathbf{x})$ to be skewed to the left, and vice versa.

This evidence enabled Yuen and Murthy (1974) to write: "It is also *well known* that the usual Student's *t*-test is conservative and hence less powerful when the underlying distribution is long-tailed" [italics mine]. See also Kafadar (1982). Yet, despite a long history of research, the results are neither comprehensive nor complete. Exact results are few and pertain mainly to samples of two and three. Approximations rarely carry bounds, and simulations are usually limited in their scope. Even Efron (1969), who presents an exact result, is cautious in his statement, "We *suggest* that the size of Student's one-sample *t* test is in *most cases* conservative under the null hypothesis of orthant symmetry" [italics mine]. Efron also finds some apparent contradiction between his results and those of Hotelling.

This article formulates some explicit statements about the conservatism of the *t* test and proves them using geometrical arguments. It focuses on the direction of deviation and not on its size. Departure from normality implying long-tailedness is described in Section 3, using the concept of stretching. A probability inequality is proved in Section 4 and is applied to show that the *t* test is conservative for stretched distributions. However, it is assured only for critical values larger than the number of degrees of freedom, that is, for small α levels. Liberal examples of the *t* test for lower critical values are given in Section 5 using simulations. Section 6 examines the conservatism of the *t* test for lower critical values using conditions stricter than stretching and suggests a much stronger result again using geometrical arguments. In particular, when the parent distribution is a scale mixture of normals, the *t* test is conservative for critical values larger than 1.8. The extension of the previous results to asymmetric distributions is explored in Section 7. In Sec-

---

* Yoav Benjamini is a Postdoctoral Fellow, Department of Statistics, Tel Aviv University, Tel Aviv, Israel. This article presents part of the author's Ph.D. thesis, done at Princeton University under the guidance of P. Bloomfield, to whom the author owes many thanks. Professor Bloomfield's example in 4.2, first given in a lecture at Princeton, was the motivation for this entire work. Thanks are also extended to H. Braun for some helpful comments on an earlier draft and to a referee whose detailed comments have greatly contributed to the organization and clarity of the presentation. A part of this article was presented under the same title at the 1980 ASA meeting in Houston, Texas. The research was supported in part by D.O.E. contract EY-76-S-02-2310, awarded to the Department of Statistics at Princeton University, and by The Analysis Center at the Wharton School at the University of Pennsylvania.

© Journal of the American Statistical Association
September 1983, Volume 78, Number 383
Theory and Methods Section

tions 8 and 9 the results are applied to explain the asymptotic behavior of the t statistic and to resolve the contradiction between the works of Hotelling and Efron. The article concludes with a discussion of the practical implications and of some open problems.

2. THE GEOMETRY OF THE *t* TEST

Because the ideas presented in this article depend on the geometric interpretation of the t statistic and its distribution, we review it here. Discussions of the geometric approach can be found in Hotelling (1961) and Efron (1969).

The vector defined by the observations $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ is a point in E^n. Let $\mathbf{e}$ be the unit vector having equal coordinates: $\mathbf{e} = (n^{-1/2}, \ldots, n^{-1/2})$. The projection of $\mathbf{x}$ on $\mathbf{e}$ is $\mathbf{p} = (\bar{x}, \bar{x}, \ldots, \bar{x})$, $\bar{x}$ being the sample mean. Its length is $\| \mathbf{p} \| = n^{1/2} \mid \bar{x} \mid$. The difference $\mathbf{x} - \mathbf{p} = \mathbf{s}$ is perpendicular to $\mathbf{p}$ and its length is $\| \mathbf{s} \| = (\sum (x_i - \bar{x})^2)^{1/2}$. Thus, the t statistic for the sample is

$$T_n(\mathbf{x}) = \frac{n^{1/2}\bar{x}}{(\sum (x_i - \bar{x})^2/(n - 1))^{1/2}}$$

$$= (n - 1)^{1/2} \frac{\| \mathbf{p} \|}{\| \mathbf{s} \|} = (n - 1)^{1/2} \cot(\theta), \quad (2.1)$$

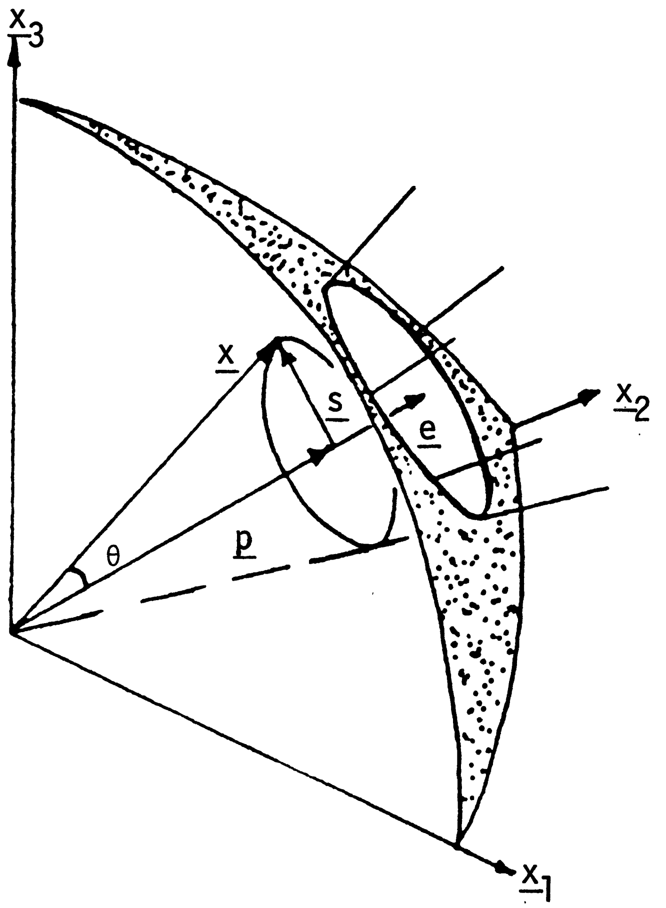

Figure 1. The geometry of the t statistic. $\mathbf{x}_1$, $\mathbf{x}_2$, $\mathbf{x}_3$, *and* $\mathbf{e}$ *are the vectors (1, 0, 0), (0, 1, 0), (0, 0, 1), and* $(1/\sqrt{3}, 1/\sqrt{3}, 1/\sqrt{3})$, *respectively.* $\mathbf{x}$ *is a vector of observations; its projection on* $\mathbf{e}$ *is* $\mathbf{p}$.

where θ is the angle between $\mathbf{x}$ and $\mathbf{e}$ (see Figure 1). For fixed n, $T_n(\mathbf{x})$ is a decreasing function of the angle θ only. Some conclusions are as follows:

1. $T_n(\mathbf{x}) = T_n(c\mathbf{x})$ for $c > 0$ (i.e., T_n is homogeneous of degree 0), because θ is constant along rays.
2. Define $S_n(\mathbf{x}) = \sum x_i/(\sum x_i^2)^{1/2}$, called "self normalized sum" by Logan et al. (1973). Results about the distribution of T_n may be translated into results about S_n and vice versa, because S_n is also a decreasing function of the angle θ:

$$S_n(\mathbf{x}) = n^{1/2} \| \mathbf{p} \| / \| \mathbf{x} \| = n^{1/2} \cos \theta. \quad (2.2)$$

3. The rays of angle θ with $\mathbf{e}$ trace the envelope of a cone. The interior of this cone contains all those points with angle less than θ. Thus, the tail probability of the t statistic $\Pr(T_n(\mathbf{x}) > t)$ is the probability content of such a cone with the appropriate θ.

An important consequence of conclusion 3 is the following: Because the vertex of this cone is always at 0, even the extreme tails of the t statistic contain probability mass from the central part of the parent distribution. Thus, the behavior of the t test is affected by the global shape of the parent distribution and not only by its tails. Moreover, no criteria involving only the tails can be sufficient to ensure conservatism for high critical values.

The cone of constant $T_n(\mathbf{x})$ cuts a circle centered at $\mathbf{e}$ on the unit sphere $B_n = \{\mathbf{x} \mid \|\mathbf{x}\| = 1\}$. Because of conclusion 1 the density along a ray can be projected to where the ray intersects B_n. The tail probability of the t statistic is thus the marginal probability content of an appropriate circular cap. In Figure 2, B_n and some circular caps are seen from a point along $\mathbf{e}$. The figure displays the following properties proved by Hotelling (1961) for $n = 3$:

1. The circular caps of $\{\mathbf{x} \mid T_n(\mathbf{x}) > t\}$ will be confined to the projection of the positive orthant onto B only when $t \geq n - 1$.
2. As long as $t > 1$, the cap does not entirely contain the positive orthant.

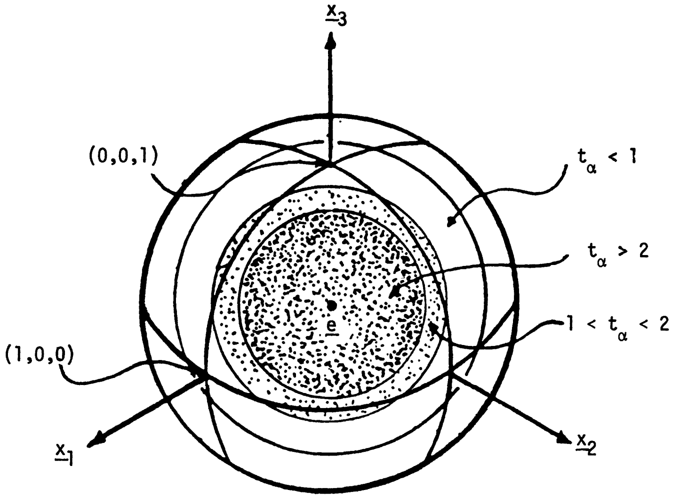

Figure 2. The unit ball B_3. $\mathbf{x}_1$, $\mathbf{x}_2$, $\mathbf{x}_3$, *and* $\mathbf{e}$ *are as in Figure 1. The circles are the spherical caps corresponding to different critical values* t_α.

For a standard normal parent distribution, the density at **x** depends only on $\| \mathbf{x} \|$, so once projected onto B_n the marginal density on B_n is uniform. The tail probability of the t distribution is the $(n - 1)$-dimensional area of such a spherical cap. It remains so whenever the sample **x** has a spherical symmetric distribution (see Efron 1969 and Jensen 1979).

3. ORDERING TAILS BY STRETCHING

In this section a partial ordering on distribution functions is used for the formulation of long-tailedness. For a cumulative distribution function F, the inverse is $F^{-1}(p) = \inf(x \mid F(x) \geq p)$ for $0 < p < 1$.

Definition 1. Let F and G be two symmetric distributions. F is *stretched* with respect to $G (F >_{st} G)$ if $(F^{-1}(p) - F^{-1}(\frac{1}{2}))/(G^{-1}(p) - G^{-1}(\frac{1}{2}))$ is increasing in p for $\frac{1}{2} < p < 1$. F is *squeezed* with respect to G if $(F^{-1}(p) - F^{-1}(\frac{1}{2}))/(G^{-1}(p) - G^{-1}(\frac{1}{2}))$ is decreasing in p for $\frac{1}{2} < p < 1$.

(Increasing and decreasing are used hereafter in the weak sense.)

Note that this partial ordering of distributions is location and scale invariant and thus orders their shapes in the sense of Mosteller and Tukey (1977 pp. 7–10). Without loss of generality let $F^{-1}(\frac{1}{2}) = G^{-1}(\frac{1}{2}) = 0$. Observe that,

$$F^{-1}(p) = (F^{-1}(p)/G^{-1}(p))\, G^{-1}(p) \qquad (3.1)$$

Thus F^{-1} is derived from G^{-1} through multiplication by an increasing function: The shape of G is stretched to get the shape of F. Definition 1 is a natural comparison of length/heaviness of the tails: If $F >_{st} G$, and F is scaled to match G at some $x_0 > 0$ (i.e., $1 - F(x_0) = 1 - G(x_0)$), then farther out the tail of F is heavier than the tail of G (i.e., $1 - F(x) \geq 1 - G(x)$ for all $x > x_0$). In Definition 1 this property is extended to the central part of the distribution because of the comment after conclusion 3 of Section 2. Even so, most of the distributions commonly thought of as being long-tailed are stretched with respect to the normal. Important examples are scale mixtures of normals, used in Andrews et al. (1972).

Remark. Stretching is closely related to, but more general than, the star ordering of distributions as presented in Barlow and Proschan (1966). It allows a comparison when G is discrete and this is explicitly used in the sequel.

4. PROVING THE CONSERVATISM

The discussion in this article is based on the following theorems.

Theorem 1. Let G be any symmetric distribution centered at 0, and $t \geq n - 1$. For any symmetric distribution F centered at 0; if $F >_{st} G$, then

$$P_F(T_n(\mathbf{x}) > t) \leq P_G(T_n(\mathbf{x}) > t). \qquad (4.1)$$

It is later shown that for $n = 2$ the condition $t \geq n - 1$ is also necessary. This is not true for larger samples. However, the next theorem shows that in this general set-up the condition cannot be relaxed to $t \geq t_0$ for any $t_0 < n - 1$.

Theorem 2. Let B be the symmetric distribution on ± 1. For any t satisfying $(n/2 - 1 \leq t < n - 1$, a symmetric F_t exists that satisfies $F_t >_{st} B$, yet $P_{F_t}(T_n(\mathbf{X}) > t) > P_B(T_n(\mathbf{X}) > t)$.

Theorem 1 can be used to prove the conservatism of the t test, by setting G to be the standard normal distribution. This yields the following result: For a sample from a symmetric distribution stretched with respect to the normal, the t test, using a critical value larger than the number of degrees of freedom, is conservative.

Theorem 1 proves the "folk theorem" under two restrictions. The first is that the behavior of the tail is extended towards the center; the second is the "high enough critical value" restriction where enough depends on the sample size. This latter condition is reasonable only for very small samples and is very restrictive for small to moderate sample sizes, as can be seen from Table 1.

Theorem 2 indicates the problems in relaxing this second restriction. These are explored further in Section 5.

4.1 Proof of Theorem 1

The idea of the proof is as follows: Using the criterion that F is stretched with respect to G, a sample from F is shown to have absolute values that tend to be less equal than those of a sample from G. This pushes a sample in the positive orthant away from vector **e**, the vector with equal components. But $\{\mathbf{x} \mid T_n(\mathbf{x}) > t\}$ is a cone centered about **e** and thus its probability content is less under F than under G. Majorization theory is used for the formal proof (Marshal and Olkin 1979).

Definition 2. For $\mathbf{x}, \mathbf{y} \in E^n$, **x** majorizes **y** $(\mathbf{x} > \mathbf{y})$ if, upon reordering **x** and **y** so that $x_1 \geq x_2 \geq \cdots \geq x_n$ and $y_1 \geq y_2 \geq \cdots \geq y_n$,

$$\sum_{i=1}^{k} x_i \geq \sum_{i=1}^{k} y_i \quad \text{for } k = 1, 2, \ldots, n - 1, \qquad (4.2)$$

and

$$\sum_{i=1}^{n} x_i = \sum_{i=1}^{n} y_i. \qquad (4.3)$$

The intuitive meaning of $\mathbf{x} > \mathbf{y}$ is that the components of

Table 1. The Translation of n − 1 to α-Level

Sample Size	Bound on t_α	α-Level (one-sided)	α-Level (two-sided)
3	2	.18350	.09175
4	3	.05768	.02884
5	4	.01614	.00807
6	5	.00103	.00205
7	6	.00096	.00048

$\mathbf{x}$ are more unequal than those of $\mathbf{y}$. The following characterization of majorization is used.

Theorem 3. (Hardy, Littlewood, and Polya 1934).

$$\mathbf{x} > \mathbf{y} \text{ iff } \sum_{i=1}^{n} g(x_i) \ge \sum_{i=1}^{n} g(y_i) \quad \text{for all convex } g.$$

The proof of Theorem 1 begins with the following two lemmas.

Lemma 1. Let $F >_{st} G$ and $p_i \in (0, 1)$ for $i = 1, 2, \ldots, n$. Then, for some constant $C > 0$,

$$(|F^{-1}(p_1)|, \ldots, |F^{-1}(p_n)|) > C \cdot (|G^{-1}(p_1)|, \ldots, |G^{-1}(p_n)|) \quad (4.4)$$

Proof. Define $p'_1 = |p_1 - \frac{1}{2}| + \frac{1}{2}$ and assume without loss of generality that $p'_1 \geqslant p'_2 \cdots \geqslant p'_n$. Because of the symmetry of both F and G about 0, $(F^{-1}(p'_1), \ldots, F^{-1}(p'_n))$ is a reordering of the left side of (4.4), and $(G^{-1}(p'_1), \ldots, G^{-1}(p'_n))$ is a reordering of the right side. Because $F >_{st} G$, $F^{-1}(p'_i)/G^{-1}(p'_i)$ is decreasing in i. Apply now proposition 5.B.2 in Marshal and Olkin (1979) to prove Lemma 1.

Lemma 2. If $\mathbf{x} > \mathbf{y}$ then $T_n(\mathbf{x}) \le T_n(\mathbf{y})$. ($T_n$ is monotonically decreasing in majorization, a property also called Schur concavity).

Proof. $\mathbf{x} > \mathbf{y}$ implies $\sum x_i/n = \sum y_i/n$ and thus,

$$\begin{aligned} T_n(\mathbf{x}) &= \frac{\sqrt{n}\,\bar{x}}{(\sum (x_i - \bar{x})^2/(n-1))^{1/2}} \\ &\leqslant \frac{\sqrt{n}\,\bar{x}}{(\sum (y_i - \bar{x})^2/(n-1))^{1/2}} \qquad (4.5) \\ &= \frac{\sqrt{n}\,\bar{y}}{(\sum (y_i - \bar{y})^2/(n-1))^{1/2}} = T_n(\mathbf{y}), \end{aligned}$$

where the inequality is from the majorization theorem.

Let J be the open interval $(\frac{1}{2}, 1)$, let $d\mathbf{p}$ be the Lebesgue measure on J^n, and let $I_{t_\alpha}(\cdot)$ be the indicator function of the set (t_α, ∞). The condition $t_\alpha > n - 1$ implies that $T_n(\mathbf{x}) \ge t_\alpha \Longrightarrow \mathbf{x} \in E_+{}^n$ and $F(x_i) \in J$ for $i = 1, 2, \ldots, n$. With the above notation we have to prove that

$$\int_{J^n} I_{t_\alpha}[T_n(F^{-1}(p_1), \ldots, F^{-1}(p_n))]d\mathbf{p} \le \int_{J^n} I_{t_\alpha}[T_n(G^{-1}(p_1), \ldots, G^{-1}(p_n))]d\mathbf{p} \quad (4.6)$$

Because I_{t_α} is monotone this will follow if we prove

$$\forall (p_1, \ldots, p_n) \in J^n$$
$$T_n(F^{-1}(p_1), \ldots, F^{-1}(p_n)) \le T_n(G^{-1}(p_1), \ldots, G^{-1}(p_n)) \quad (4.7)$$

But inequality (4.7) follows using Lemma 1 and then applying Lemma 2 together with the homogeneity of $T_n(\mathbf{x})$.

Remark. Theorem 1 proves a probability inequality for multivariate distributions that belong to a family parameterized by an infinite-dimensional parameter (F and G are the values of the parameter). Similar inequalities, with an emphasis on finite-dimensional parameter spaces, are treated in a recent review by Eaton (1982).

4.2 An Example: Sample of Size Two

This example (Bloomfield—see Acknowledgment footnote) demonstrates geometrically the conservatism of the t test and how it fails to hold for sample of size two.

When $t > 1$, $\{\mathbf{x} \mid T_2(\mathbf{x}) > t\}$ is a wedge completely contained within the positive orthant. As we move from distribution G to distribution F satisfying $F >_{st} G$, probability mass escapes outside of this wedge.

When $t = 1$, $\{\mathbf{x} \mid T_2(\mathbf{x}) > t\}$ is exactly the positive orthant. Under a symmetric distribution its probability is always $\frac{1}{4}$.

When $0 < t < 1$, the wedge contains the positive orthant and can be described as the half plane $\{\mathbf{x} \mid T_2(\mathbf{x}) \ge 0\}$ minus two wedges. Because of symmetry, these two wedges correspond to a wedge centered at $\mathbf{e}$: $\{\mathbf{x} \mid T_2(\mathbf{x}) \ge t'\}$. Hence

$$\begin{aligned} &P_F(T_2(\mathbf{x}) > t) \\ &= \tfrac{1}{2} - P_F(0 \le T_2(\mathbf{x}) \le t) = \tfrac{1}{2} - P_F(T_2(\mathbf{x}) > t') \\ &> \tfrac{1}{2} - P_G(T_2(\mathbf{x}) > t') = P_G(T_2(\mathbf{x}) > t). \end{aligned}$$

For this example, the performance of the t test is completely solved:

For $t \ge 1$ the t test is conservative

For $0 \le t < 1$ the t test is liberal.

The equivalent result holds for the negative tail because of the symmetry.

Result. For $n = 2$, the condition $t \ge n - 1$ is not only sufficient but is also necessary.

4.3 Proof of Theorem 2

Define $F_{p,\delta}$ to be the symmetric distribution putting mass p on 1 and -1 and mass $q = \frac{1}{2} - p$ on δ and $-\delta$ with $\delta \ge 1$. $B^{-1}(p)$ is constant on $(\frac{1}{2}, 1)$ while $F_{p,\delta}{}^{-1}(p)$ is an increasing step function, so $F_{p,\delta} >_{st} B$.

For $\mathbf{y}$, which is a sample from B, $(|y_1|, \ldots, |y_n|) = (1, 1, \ldots, 1)$. Because $T_n(1, \ldots, 1, -1, -1) < T_n(1, \ldots, 1, -1) < T_n(1, \ldots, 1, 1)$ and because $T_n(1, \ldots, 1, -1) = (n/2) - 1$, we have, for $(n/2) - 1 < t < n - 1$, that $P_B(T_n(\mathbf{y}) > t) = P_B(\mathbf{y} = (1, \ldots, 1)) = 2^{-n}$. Define $\mathbf{z}(\delta) = (\delta, \delta, \ldots, \delta, -1)$. Then $T_n(z(\delta)) = (\delta(n-1) - 1)/(\delta + 1)$, which is increasing in δ. $T_n(z(\infty)) = n - 1$ and $T_n(z(1)) = n/2 - 1$, so for any $n/2 - 1 < t < n - 1$, a large enough $\delta(t)$ can be found such that $t <$

$T_n(z(\delta(t)))$ (see Figure 3). For this $\delta(t)$,

$$\begin{aligned} P_{F_{p,\delta(t)}}(T_n(\mathbf{x}) > t) &= P(\mathbf{x} = (\delta(t), \delta(t), \ldots, \delta(t))) \\ &\quad + nP(\mathbf{x} = (\delta(t), \ldots, \delta(t), 1)) \\ &\quad + \cdots + P(\mathbf{x} = (1, 1, \ldots, 1))) \\ &\quad + nP(\mathbf{x} = (\delta(t), \ldots, \delta(t), -1)) + \cdots \\ &\geq P(\mathbf{x} = (\delta(t), \delta(t), \ldots, \delta(t))) \\ &\quad + 2nP(\mathbf{x} = (\delta(t), \ldots, \delta(t), 1)) \\ &= q^n + 2nq^{n-1}p \end{aligned} \tag{4.8}$$

Now choose p small enough so that

$$nq^{n-1}p > \binom{n}{2} q^{n-2}p^2 + \cdots + p^n. \tag{4.9}$$

This p depends only on n. Finally because $(p + q)^n = 2^{-n}$,

$$P_{F_{p,\delta(t)}}(T_n(\mathbf{x}) > t) > 2^{-n} = P_B(T_n(\mathbf{x}) > t). \tag{4.10}$$

Figure 3 demonstrates the proof for the case of $n = 3$: The points • are placed by B, each with probability $\frac{1}{8}$. For a given $t < 2$, the spherical cap, whose content is the tail of $T_n(\mathbf{x})$, exceeds the boundary of the spherical triangle that is the projection onto the sphere of the positive orthant. For δ sufficiently large, F places the points x so that the spherical cap also contains points ⊗ from outside the positive orthant. For $\frac{1}{2} < t$ the cap contains under B only one point—the one from the positive orthant.

Remark. Let H_1 and H_2 be two symmetric, zero-centered distributions having equal variances. When all higher (even) moments of H_1 are larger than the corresponding moments of H_2, H_1 is usually regarded as being more dispersed about the origin than H_2. For example, define H_F (H_B) to be the distributions of $S_n(\mathbf{x})$ when $\mathbf{x}$ is a sample from some symmetric F (B of the previous section) appropriately. Efron (1969) shows that all moments of H_B are larger or equal to those of H_F and cautiously interprets this as an indication of tail-probability inequality. Theorem 3 shows this approach to be valid only in the extreme tails for $S_n^2(\mathbf{x}) \geq n - 1$. But let F be standard normal and $s_0 = 4/\sqrt{6} \simeq 1.633$; then

$$\begin{aligned} P_F(S_n(\mathbf{x}) > s_0) &= 1 - H_F(s_0) = .051 \\ &> .015 = 1 - H_B(s_0) = P_B(S_n(\mathbf{x}) > s_0). \end{aligned}$$

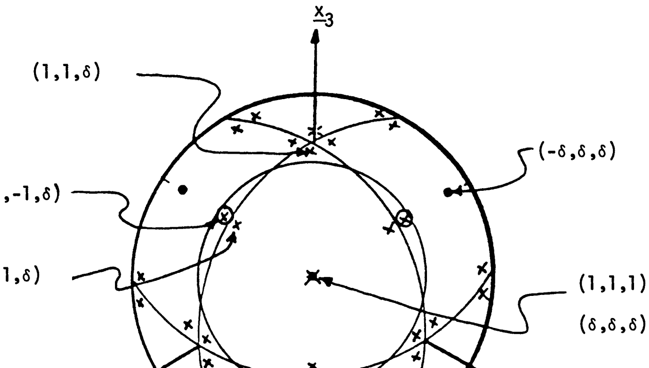

*Figure 3. Liberal example for the binomial distribution as a reference. • are points on which B is supported; x are points on which $F_{p,\delta}$ is supported; **x** are points x that do not belong to the positive orthant yet are included in the spherical cap for small enough δ.*

5. LIBERAL EXAMPLES FOR DISTRIBUTIONS STRETCHED WITH RESPECT TO NORMAL

It is clear from Table 1 that the condition $t \geq n - 1$ in Theorem 1, used to prove the conservatism of the t test, is quite restrictive. For $n = 2$ it has been shown to be not only a sufficient condition but also necessary, and for larger n it has been shown to be the weakest possible condition for *some* reference distribution G but not for all. This section explores the extent to which the condition $t \geq n - 1$ is the weakest possible also when the reference distribution is normal.

The family of inverses of distributions stretched with respect to the normal, can be described as a cone of generalized convex functions $C(\Phi^{-1})$ (Karlin and Studden 1969) in the set of increasing functions on $(\frac{1}{2}, 1)$. The tail probability of the t statistic $P_F(T_n(\mathbf{x}) > t)$ can be viewed as a nonlinear functional on this cone. For fixed $t > n - 1$ this functional is maximized on Φ^{-1} (Theorem 1), which is an extreme ray of $C(\Phi^{-1})$; that is, for $t > n - 1$,

$$\sup_{F^{-1} \in C(\Phi^{-1})} P_F(T_n(\mathbf{x}) > t) = P_\Phi(T_n(\mathbf{x}) > t). \tag{5.1}$$

Thus it is natural to search for liberal examples on extreme rays, when $t < n - 1$ as well. (In fact, the distributions $F_{p,\delta}{}^{-1}$ of the previous section were extreme rays of the cone generated by B^{-1}).

The extreme rays of the $C(\Phi^{-1})$ are of the form

$$\begin{aligned} F^{-1}(p) &= \Phi^{-1}(p) & \tfrac{1}{2} < p < p_0 \\ &= \sigma\Phi^{-1}(p) & p_0 \leqslant p \end{aligned} \tag{5.2}$$

for $\sigma > 1$ and $p_0 \in (\frac{1}{2}, 1)$.

For some F in the family (5.2), using different values of σ and p_0 the cumulative distribution function of $T_n(\mathbf{x})$ is calculated using experimental sampling and is compared with the theoretical one under normality, for different values of σ and p_0. Ten thousand samples were used for each example. This was done on a PDP 11 computer, using the Box-Muller-Knuth Algorithm to generate normal samples. From these, samples of F were derived using the stretching transformation. The critical values were taken at steps of .01. Figure 4 describes these results when $n = 3$. In this figure (and in the following ones) conservatism is evidenced by a positive difference—hence for $t > 2$, due to theorem 1, the graphs are all positive.

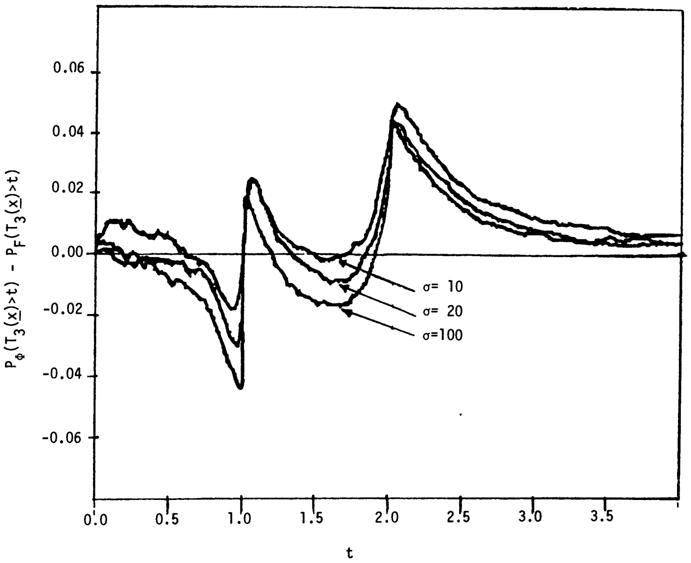

Figure 4. The tail of $T_3(X)$ when the parent is normal minus the tail when the parent corresponds to an extreme ray of $C(\Phi^{-1})$, $\Phi^{-1}(p_0) = .2$. Note: The maximal standard error of the simulated difference is .005. The standard error is less than .004 at $t = 2$.

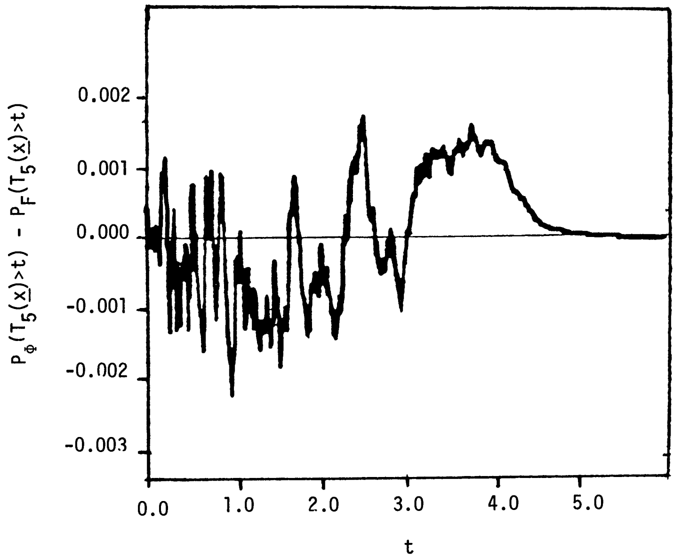

Figure 6. The tail of $T_5(x)$ when the parent is normal minus the tail when the parent corresponds to an extreme ray of $C(\Phi^{-1})$, $\Phi^{-1}(p_0) = .15$. Note: The maximal standard error of the simulated difference is .0018. The standard error is .0009 at $t = 3$.

It is clear that as σ increases we get liberal inequalities for critical values closer to two. These get even closer when p_0 of (5.2) gets closer to $\frac{1}{2}$ (compare Figures 4 and 5). As n increases the simulation become noisier because very small probabilities are compared. We thus initially let $\sigma \longrightarrow \infty$, which can be done by rescaling since $T_n(\mathbf{x})$ is scale invariant, and increase the number of samples to 30,000. Then $P_\Phi(T_n(\mathbf{x}) > t) - P_F(T_n(\mathbf{x}) > t)$ is evaluated by experimental sampling. Liberal examples still exist quite far into the tail (see e.g., Figure 6).

One very noticeable phenomenon in Figures 4 and 5 are the bumps at $t = 1$ and $t = 2$. The reason for these bumps is that the parent distributions used for the liberal examples put high probability near the lower-dimensional subspaces, while for a normal parent the probability is uniformly spread over spheres. As t decreases, the cone $\{\mathbf{x} \mid T_n(\mathbf{x}) > t\}$ hits the k-dimensional subspace for the first time at the point $t_k = T((1, \ldots 1, 0, \ldots 0))$. For $n = 3$ these points are $t_1 = T((1, 0, 0)) = 1$ and $t_2 = T((1, 1, 0)) = 2$. For $n = 5$ these points are $t_1 = 1$, $t_2 = 1.43$, $t_3 = 2.45$, and $t_4 = 4$. The bumps at these points in Figure 6 are again very clear.

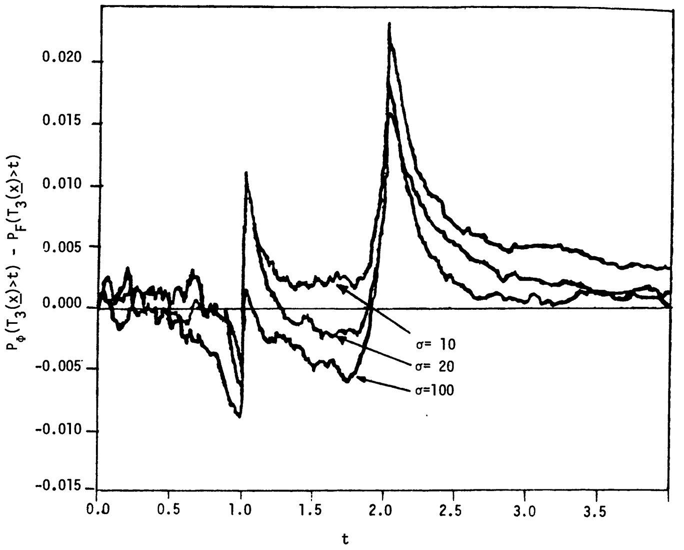

Figure 5. The tail of $T_3(x)$ when the parent is normal minus the tail when the parent corresponds to an extreme ray of $C(\Phi^{-1})$, $\Phi^{-1}(p_0) = .1$. Note: The maximal standard error of the simulated difference is .005. The standard error is less than .004 at $t = 2$.

6. LIBERAL EXAMPLES UNDER STRONGER CONDITIONS

A possible way to confirm the folk theorem with a weaker condition on the critical values allowed is to formulate long-tailedness using a stronger condition. Rather than requiring $F >_{st} \Phi$ one might suggest other orderings of distribution, say, $\gg$, which are location and scale invariant, and such that $F \gg \Phi$ implies $F >_{st} \Phi$. Here it might turn out that in the set $\{F \mid F \gg \Phi\}$, distributions yielding liberal examples at commonly used critical values cannot be found.

A variety of increasingly stronger orderings have been examined. These are (a) requiring $F^{-1}(\Phi(x))$ to be convex (see Van Zwet 1964); (b) requiring the logarithm of the density of F to be concave with respect to that of Φ (or equivalently $\log f(\sqrt{x})$ convex); and (c) requiring F to be a scale mixture of normal distributions (see Rogers and Tukey 1972).

The possibility that a lower critical value is sufficient when using ordering (a) or (b) is explored using experimental sampling as presented in Section 5. Similar considerations point at parent distributions of the same nature—extreme rays of appropriate cones of functions.

The results concerning the behavior of the t test are also similar: liberal examples exist when the test is done using critical values that are insufficiently high. Here, the stronger the condition is and the larger n is, the more difficult it is to get close to $n - 1$, the bound beyond which conservatism is assured by Theorem 1. (This requires distributions that differ very slightly from the normal.) There is therefore no assurance that a lower bound does not exist. Yet the liberal examples cover the range of critical values at which testing is commonly done. For example, for $n = 10$ there exists a unimodal distribution, convex with respect to the normal such that the t test using a critical value of three is liberal. (See Figure 7).

6.1 Scale Mixture of Normals

We now suggest that for the class of scale mixtures of normals the conservatism of the t test holds over a wide range of critical values. First, geometry is used to argue that certain distributions are the most likely to yield liberal examples of t tests using high critical values. Second, a bound is presented for these distributions beyond which the t test is conservative. This bound is different for each sample size n, but the first dozen or so calculated numerically suggest the critical value of 1.8 as a global bound for all n beyond which conservatism is guaranteed.

Throughout this section the statistic $S_n(\mathbf{x})$ is used as defined in conclusion 2 of Section 2. The following theorem is needed for density f, which is a scale mixture of normals. We prove it under the weaker ordering (b) of 6.1.

Theorem 4. Let f be a density function such that $\log f(\sqrt{x})$ is convex. Let f_n be the marginal density induced by $\prod f(x_i)$ on B_n. If $\mathbf{x}, \mathbf{y} \in B_n$, and $(x_1^2, \ldots, x_n^2) > (y_1^2, \ldots, y_n^2)$, then $f_n(\mathbf{x}) \geq f_n(\mathbf{y})$.

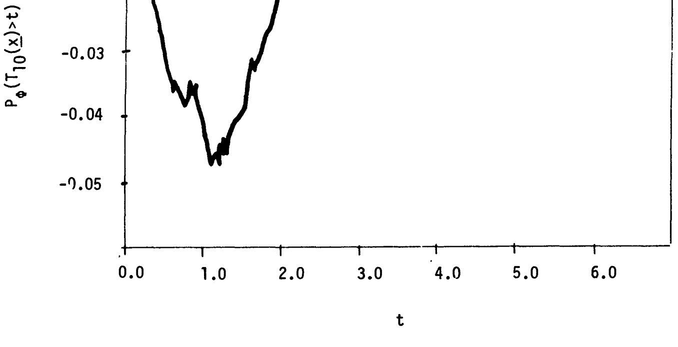

Figure 7. The tail of $T_{10}(x)$ when the parent is normal minus the tail when the parent corresponds to an extreme ray of the distributions convex with respect to Φ, $\Phi^{-1}(p_0) = .2$. Note: The maximal standard error of the simulated difference is .005. The standard error is less than .0014 at $t = 3$.

Proof. For each t, $0 < t < \infty$,

$$(t^2x_1^2, \ldots, t^2x_n^2) > (t_1^2y_1^2, \ldots, t^2y_n) \tag{6.1}$$

and

$$\begin{aligned}\prod_{i=1}^{i=n} f(tx_i) &= \exp\left(\sum_{i=1}^{i=n} \log f(tx_i)\right)\\ &= \exp\left(\sum_{i=1}^{i=n} \log f((t^2x_i^2)^{1/2})\right)\\ &\geq \exp\left(\sum_{i=1}^{i=n} \log f((t^2y_i^2)^{1/2})\right) = \prod_{i=1}^{i=n} f(ty_i),\end{aligned} \tag{6.2}$$

where the inequality is from Theorem 3. Finally,

$$\begin{aligned} f_n(\mathbf{x}) &= \int_0^\infty \prod f(tx_i)t^{n-1}dt \geq \int_0^\infty \prod f(ty_i)t^{n-1}dt\\ &= f_n(\mathbf{y}).\end{aligned} \tag{6.3}$$

Theorem 4 implies that for scale mixtures of normal distributions the density increases as we approach the lower dimensional subspaces. (Let E^k be the lowest dimensional subspace to which $\mathbf{x}$ belongs. If $k > 1$ then for each $1 \leq l < k$ there exist $y \in E^l$ and a path on B_n from $\mathbf{x}$ to $\mathbf{y}$ along which majorization increases).

Recall from the geometrical discussion at the end of Section 5 that the conservatism is assured until $\{\mathbf{x} \mid s_n(\mathbf{x}) > t\} \cap B_n$ hits E^{n-1} for the first time. Thus liberal examples are generated by distributions that induce high marginal densities on $E^{n-1} \cap B_n$. However, for scale mixtures of normals, Theorem 4 limits the behavior of the marginal density induced on B. Within this limit, mixing the scales with the discrete distribution used in Theorem 2 to yield liberal examples, generates distributions of the form

$$F(x) = (1 - \epsilon)\Phi(x) + \epsilon\Phi\left(\frac{x}{\sigma}\right) \tag{6.4}$$

with ϵ small and σ small. Choosing σ small moves the density closer to the lower-dimensional subspaces, and in the limit σ can be set to 0. Choosing ϵ small generates the highest possible marginal density on $E^{n-1} \cap B_n$. For F of this form, using the fact that, for $\mathbf{x}^* \in E^{n-1}$,

$$P_F(S_n(\mathbf{x}^*, x_n) > s \mid x_n = 0) = P_F(S_{n-1}(\mathbf{x}^*) > s),$$

we get

$$\begin{aligned}P_F(S_n(\mathbf{x}) > s) \approx (1 - n\epsilon)&P_\Phi(S_n(\mathbf{x}) > s)\\ &+ n\epsilon P_\Phi(S_{n-1}(\mathbf{x}) > s),\end{aligned} \tag{6.5}$$

where this approximation is of order ϵ^2. Thus $P_F(S_n(\mathbf{x}) > s) > P_\Phi(S_n(\mathbf{x}) > s)$ only if $P_\Phi(S_{n-1}(\mathbf{x}) > s) > P_\Phi(S_n(\mathbf{x}) > s)$.

Lemma 3 describes where this happens.

Table 2. Bounds Beyond Which the t-Test is Conservative

n	s_n	t_n	α-Level (one-sided)
2	1.000	1.000	.250
3	1.314	1.647	.121
4	1.428	1.767	.088
5	1.492	1.792	.074
6	1.534	1.795	.066
7	1.563	1.793	.063
8	1.584	1.789	.058
9	1.601	1.785	.056
10	1.615	1.781	.054
11	1.625	1.778	.053
12	1.634	1.775	.052
13	1.642	1.772	.051
14	1.649	1.770	.050
15	1.654	1.768	.049
16	1.659	1.766	.049
17	1.664	1.764	.048
18	1.668	1.763	.048

Lemma 3. There exists s_n such that for $0 < s < s_n$, $P_\Phi(S_{n-1}(\mathbf{x}) > s) > P_\Phi(S_n(\mathbf{x}) > s)$ and for $s > s_n$, $P_\Phi(S_{n-1}(\mathbf{x}) > s) \le P_\Phi(S_n(\mathbf{x}) > s)$.

The idea of the proof is to differentiate the ratio of the densities of S_n and S_{n-1} and by a straightforward (though lengthy) calculation show that the two densities are equal on the support of S_n at exactly two points. As a result the distribution functions are equal at exactly three points. Two such points are 0 and $\sqrt{n}$ (the end point of the support of S_n) and the third is the desired s_n. The direction of the inequalities is finally determined by observing that the support of S_{n-1} is up to $(n-1)^{1/2}$.

As a result of Lemma 3 conservatism is assured for critical values larger than s_n. Table 2 presents the first few s_n, determined numerically (as well as the corresponding t_n's for the t test in the usual form). Each critical value t_n is much lower here than the bound $n - 1$, which holds for any stretched distribution. In fact, the one-sided level is above the range where testing is customarily done. As n increases, the α-level decreases and comes just below the conventional 5% level.

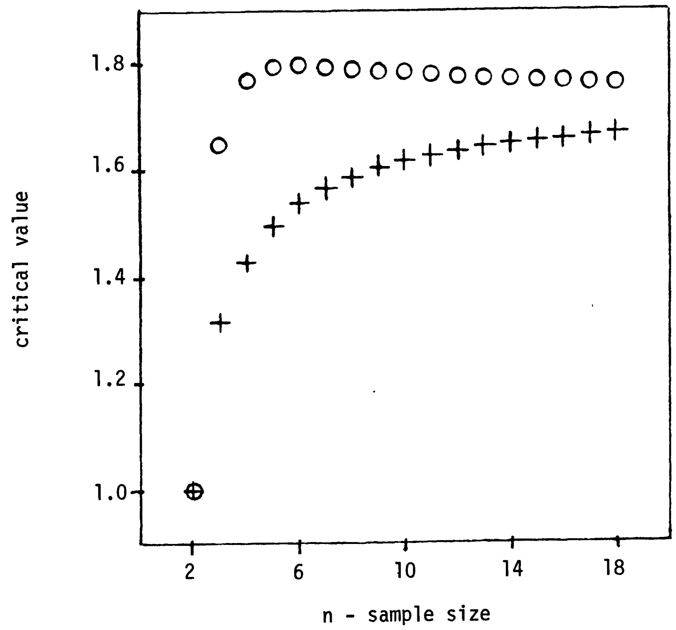

Figure 8. Bounds on critical values beyond which the t test is conservative for scale mixtures of normals. The bounds for the statistics $S_n(x)$ are denoted by +; for the statistic $T_n(x)$, by 0.

These numbers, also displayed in Figure 8, suggest there is a global bound for all n. First, from conclusion 2 of Section (2), it can easily be seen that if either s_n or the corresponding t_n converge to a limit as $n \to \infty$, the other one converges to the same limit. Now note in Figure 8, that as s_n increases, t_n decreases for $n > 6$ (with $t_6 = 1.795$). It seems that t_n and s_n converge to a point between 1.6 and 1.8 with $t_n < 1.8$ for all n.

7. THE *t* TEST WHEN THE PARENT DISTRIBUTION IS ASYMMETRIC

Skewness to the right can be formalized by requiring the right side of the distribution to be stretched with respect to the left side.

Definition 3. F is skewed to the right(left) if

$$(F^{-1}(p) - F^{-1}(\tfrac{1}{2}))/(F^{-1}(\tfrac{1}{2}) - F^{-1}(1-p)) \text{ increases (decreases) in } p \in (\tfrac{1}{2}, 1). \tag{7.1}$$

This condition is weaker than that suggested by Van Zwet (1964). Theorem 5 partially proves conclusion (2) of Section 1.

Theorem 5. Let F be a right-skewed distribution with median at 0. Then, for all $t_\alpha \ge n - 1$,

$$P_F(T_n(\mathbf{x}) > t_\alpha) \le P_F(T_n(\mathbf{x}) < -t_\alpha). \tag{7.2}$$

The opposite inequality holds when F is skewed to the left.

Proof. The proof is immediate, by applying Theorem 1 using both sides of the distribution.

Theorem 5 states that if the parent distribution is skewed to the right, the right tail probability of the t statistic is smaller than the corresponding left one, at least for high critical values. Whenever the right side is stretched with respect to the normal and the left side is squeezed with respect to the normal (implying skewness to the right), the condition $t_\alpha \ge n - 1$ might be relaxed along the lines suggested in Section 6.

8. GEOMETRICAL INTERPRETATION OF THE ASYMPTOTIC BEHAVIOR

Logan et al. (1973) study the limiting distribution of $S_n(\mathbf{x})$, where $\mathbf{x}$ is a sample from a stable distribution of parameter $\alpha < 2$. They show that $S_n(\mathbf{x})$ has a proper limiting distribution H, obtain an exact formula for its density h, and note two curiosities in this density. The first is that h has noticeable bumps at $\sqrt{1}$, $\sqrt{2}$, $\sqrt{3}$, and so on, at least for small α. The second is that when the parent is also asymmetric, h is still symmetric on $(-1, +1)$. The geometrical approach sheds some light on these two phenomena. In the symmetric case the parent distributions are mixtures of Gaussians, so Theorem 4 and the subsequent discussion apply. As discussed in the end of Section 5, the first time the cone $\{\mathbf{x} \mid S_n(\mathbf{x}) > s\}$ hits the k-dimensional subspace is when $\mathbf{x}_k = (1, \ldots, 1, 0, \ldots, 0)$, for which $s_n(\mathbf{x}_k) = \sqrt{k}$.

Together these explain the rises of the density at these points. The stable distributions of $\alpha < 2$ have the property that a few observations of larger modulus essentially determine the sums in the numerator and the denominator (see Darling 1952, also mentioned by Logan et al. 1973). This explains why the small sample properties studied in this article are still apparent in the limiting distribution.

In the asymmetric case, for $s > 1$, $\{\mathbf{x} \mid S_n(\mathbf{x}) \geq s\}$ includes part of the positive orthant, in which all x_i's come from the right side of the distribution. For the same reason $\{\mathbf{x} \mid S_n(\mathbf{x}) \leq -s\}$ contains the symmetric region in which all x_i's come from the left. The results from the asymmetric case explain why one tail is larger than the other. For $|s| < 1$, the cone is completely outside the positive and the negative orthants, where everywhere both tails contribute to the value of $S_n(\mathbf{x})$, and as $n \to \infty$ are mixed to yield symmetry on $(-1, 1)$. Thus the difference in the limiting behavior of $S_n(\mathbf{x})$, for $|s| < 1$ and for $|s| > 1$, has an intuitive explanation based on the geometric interpretation for small samples.

9. A COMPARISON OF HOTELLING'S AND EFRON'S RESULTS

As the tail probability of $T_n(\mathbf{x})$ is the integral of marginal density over the circular cap centered at $\mathbf{e}$, Hotelling (1961) suggested using the ratio of the nonnormal density f_n and the normal density ϕ_n, both evaluated at $\mathbf{e}$, to modify the size of the t test. Let this ratio be $R_n = f_n(\mathbf{e})/\phi_n(\mathbf{e})$. When $R_n < 1$, for some large enough critical value t_α,

$$P_f(T_n(\mathbf{x}) > t) < P_\Phi(T_n(\mathbf{x}) > t). \tag{9.1}$$

The opposite is true for $R_n > 1$.

Hotelling derived R_n for the Cauchy distribution and for the rectangular. In the first case, $R_n < 1$ and $R_n \to 0$ as $n \to \infty$; in the second case, $R_n > 1$ and $R_n \to \infty$ as $n \to \infty$. So as the number of observations increases the difference in the extreme t tails between f_n and ϕ_n can be made as large or as small as desired.

Van Zwet (1964) showed that ordering (a) of Section 6 can be used to determine whether $R_n \leq 1$ (when $F^{-1}(\Phi(x))$ is convex), or $R_n \geq 1$ (when $F^{-1}(\Phi(x))$ is concave). It is clear from Theorem 1 that the weaker condition of $F >_{st} \Phi$ is enough to assure this behavior when the limit R_n exists. Furthermore, from Theorem 4 one can deduce that under the stronger ordering (b) the correction suggested by Hotelling is liberal for $R_n < 1$ and conservative for $R_n > 1$.

Efron (1969) points at the contradiction that while he finds the t test to be generally conservative, Hotelling (1961) finds it both conservative and liberal. Efron attributes this to the fact that Hotelling's results are confined to the extreme tails. I suggest a different explanation: That Efron's results are based on the comparison of any symmetric distribution with the symmetric binomial B on $\pm C$. Every symmetric distribution is stretched with respect to B, because $B^{-1}(p)$ is constant for $p > \frac{1}{2}$, while $F^{-1}(p)$ is increasing. So the normal Φ is stretched with respect to B, with distributions such as the Cauchy distribution being more stretched then Φ, and distributions such as the rectangular being more squeezed than Φ. This confirms Hotelling's results about the limits as $t_\alpha \to \infty$. Yet, all of these are stretched with respect to the binomial B and so Efron's main theorem (see remark at the end of Section 4) holds also in the terms of probability inequality as long as $t_\alpha \geq n - 1$, the x_i's are also independent, and the behavior is compared to the t test with a parent binomial distribution. The conclusion derived from Efron's theorem about the general conservatism of the ordinary t test is too broad a statement, because the reference distribution is normal. This is the cause of the apparent contradiction.

10. PRACTICAL IMPLICATIONS AND SOME OPEN PROBLEMS

For very small samples the question raised in the title of this article can be answered positively at reasonable (i.e., low enough) levels. Also, the confidence interval based on the t test can be viewed as an "upper bound" confidence interval for the center (because of the generality of distributions stretched with respect to the normal). This overcomes a difficulty sometimes raised against the usage of the 95% confidence interval when $n \leq 6$, where the interval extends outside the range of the observations.

For samples of larger size, assuming that the parent distribution belongs to the family of scale mixtures of normals, the t test can be confidently used at the usual levels. Otherwise the parent distribution should be examined more closely. In the examples in which the t test is liberal, the parent distribution is strongly stretched near the center. Thus, if on top of the long-tailedness the center of the distribution is normal, or at least smoothly stretched, there is reason to believe that the conservatism will hold when the test is done at customary levels. It is comforting that such a check (possibly using *q-q* or *p-p* plots) is based on the central part of the distribution, for which most empirical information is available. Practically we might even confirm a limited version of Efron's statement in the following way: in view of Winsor's Principle that most distributions are normal in the middle and Tukey's Principle that most distributions have stretched tails, we might conclude that a t test, done at a critical value that is not too low, is conservative for most parent distributions.

The results in this article can be used indirectly to improve the robustness of efficiency of the t test. The departure from normality is described by a change in the shape, yielding general yet intuitively well-described alternatives. The price paid is the inability to give the magnitude of the change as Yuen and Murthy (1974) do for the t family. However the two approaches can be combined using Theorem 1: Bounding the shape of the parent distribution from inside and outside by distributions for which the t tails are tabulated gives upper and lower bounds on the tails of the t statistic.

Finally, some questions received only partial answers:

1. For the family $\{F \mid F >_{st} G\}$, the points beyond which the conservatism is guaranteed depend on n and on the reference distribution G. For the binomial reference these are $n - 1$, while for a Gaussian reference they increase in n but not so fast. Can these points be calculated? Is there a relationship between two such series of bounds that depends on a relation between their two reference distributions? The same questions might be asked for the more restrictive criteria.

2. In suggesting 1.8 as the global bound for the family of mixture of normals, two steps were justified only heuristically; a rigorous proof, possibly along the lines of Efron and Olshen (1978), would be desirable.

[*Received September 1981. Revised March 1983.*]

REFERENCES

ANDREWS, D.F., BICKEL, P.J., HAMPEL, F.R., HUBER, P.J., ROGERS, W.H., and TUKEY, J.W. (1972), *Robust Estimates of Location: Survey and Advances*, Princeton, N.J.: Princeton University Press.

BARLOW, R.E., and PROSCHAN, F. (1966), "Tolerance and Confidence Limits for Classes of Distributions Based on Failure Rates," *Annals of Mathematical Statistics*, 37, 76–81.

CRESSIE, N. (1980), "Relaxing Assumptions in the One Sample t-Test," *Australian Journal of Statistics*, 22, 143–153.

DARLING, D.A. (1952), "The Influence of the Maximum Term in the Addition of Independent Random Variables," *Transactions of the American Mathematical Society*, 73, 95–107.

EATON, M.L. (1982), "A Review of Selected Topics in Multivariate Probability Inequalities," *The Annals of Statistics*, 10, 11–43.

EFRON, B. (1969), "Student's t-Test Under Symmetry Conditions," *Journal of the American Statistical Association*, 63, 1278–1302.

EFRON, B., and OLSHEN, R.A. (1978), "How Broad is the Class of Normal Scale Mixtures?," *The Annals of Statistics*, 6, 1159–1164.

HARDY, G.H., LITTLEWOOD, J.E., and POLYA, G. (1934), *Inequalities*, London: Cambridge University Press.

HOTELLING, H. (1961), "The Behavior of Some Standard Statistical Tests Under Non-Standard Conditions," *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability*, 1, 319-360.

JENSEN, D.R. (1979), "Linear Models Without Moments," *Biometrika*, 66, 611-617.

JOHNSON, J.J. (1978), "Modified t Tests and Confidence Intervals for Asymmetrical Populations," *Journal of the American Statistical Association*, 73, 536–544.

KAFADAR, K. (1982), "A Biweight Approach to the One-Sample Problem," *Journal of the American Statistical Association*, 77, 416–424.

KARLIN, S., and STUDDEN, W.J. (1969), *Tchebycheff Systems, With Applications in Analysis and Statistics*, Wiley (Interscience) New York.

LOGAN, B.F., MALLOWS, C.L., RICE, S.O., and SHEPP, L.A. (1973), "Limit Distributions of Self Normalized Sums," *The Annals of Probability*, 1, 788–809.

MARSHALL, A.W., and OLKIN, I. (1979), *Inequalities: Theory of Majorization and Its Applications*, New York: Academic Press.

MOSTELLER, F., and TUKEY, J.W. (1977), *Data Analysis and Regression*, Reading, Mass.: Addison-Wesley.

ROGERS, W.H., and TUKEY, J.W. (1972), "Understanding Some Long-Tailed Distributions," *Statistica Nederlandica*, 26, 3.

VAN ZWET, W.R. (1964), *Convex Transformations of Random Variables*, Amsterdam: Mathematische Centrum.

YUEN, K.K., and MURTHY, V.K. (1974), "Percentage Points of the t-Statistic When the Parent is Student's t," *Technometrics*, 16, 495–497.