AD-A093 566 WISCONSIN UNIV-MADISON MATHEMATICS RESEARCH CENTER F/G 12/1
BOUNDS FOR THE T-TAIL AREA FOR ARBITRARY DEGREES OF FREEDOM.(U)
AUG 80 A P SOMS DAAG29-80-C-0041
UNCLASSIFIED MRC-TSR-2105 NL

1 OF 1
AD
A093566

END
DATE
FILMED
2-84
DTIC

MRC Technical Summary Report #2105

BOUNDS FOR THE t-TAIL AREA FOR ARBITRARY DEGREES OF FREEDOM

Andrew P. Soms

Mathematics Research Center
University of Wisconsin—Madison
610 Walnut Street
Madison, Wisconsin 53706

August 1980

(Received June 24, 1980)

Sponsored by

U. S. Army Research Office
P. O. Box 12211
Research Triangle Park
North Carolina 27709

UNIVERSITY OF WISCONSIN - MADISON
MATHEMATICS RESEARCH CENTER

BOUNDS FOR THE t-TAIL AREA FOR ARBITRARY DEGREES OF FREEDOM

Andrew P. Soms

Technical Summary Report #2105
August 1980

ABSTRACT

The bounds of Birnbaum (1942) and Sampford (1953) for the upper tail area of the normal distribution are extended to the upper tail of the t-distribution for arbitrary real degrees of freedom. This generalizes the results of Soms (1977a) for integral degrees of freedom. Numerical and theoretical comparisons are made with the bounds of Peizer and Pratt (1968), Wallace (1959) and Soms (1977b).

AMS (MOS) Subject Classification: 60E05

Key Words: Bounds, t-Tail area, t-Distribution

Work Unit Number 4 (Statistics and Probability)

Sponsored by the United States Army under Contract No. DAAG29-80-C-0041 and the University of Wisconsin-Milwaukee Graduate School.

SIGNIFICANCE AND EXPLANATION

Real degrees of freedom arise in the modification of the two-sample t-test when the variances cannot be assumed to be equal. For small degrees of freedom both linear interpolation and standard computer routines may be unsatisfactory. The present paper provides simple techniques for satisfactory estimates of small tail probabilities.

The responsibility for the wording and views expressed in this descriptive summary lies with MRC, and not with the author of this report.

BOUNDS FOR THE t-TAIL AREA FOR ARBITRARY DEGREES OF FREEDOM

Andrew P. Soms

1. Introduction

The purpose of this paper is to extend to the tail area of the t-distribution bounds given by Birnbaum (1942) and Sampford (1953) for the tail area of the normal distribution based on Mills' ratio. This extension of Mills' ratio type bounds is completed in Soms (1977b) and (1979) and the basis for these extensions is the asymptotic expansion obtained in Soms (1976). An excellent summary of the normal results can be found in Chapter 33 of Johnson and Kotz (1970). The organization of this paper is as follows. In Section 2, we introduce the notation and give needed previous results. The proofs are given in Section 3, in Section 4 theoretical comparisons between the bounds are made, and Section 5 consists of numerical comparisons.

2. Statement of Results

Let $\phi(x) = (2\pi)^{-1/2} e^{-x^2/2}$, $\bar{F}_k(x) = \int_x^\infty \phi(t)dt$, and $R_x = \bar{F}_k(x)/\phi(x)$, $x > 0$. Then Birnbaum (1942) and Sampford (1953) gave the following lower and upper bounds, respectively, for R_x:

$$(x/2 + (1+x^2/4)^{1/2})^{-1} < R_x < (3x + (1/2+x^2/16)^{1/2})^{-1} \quad . \tag{2.1}$$

Now let $f_k(t) = c_k(1+t^2/k)^{-(k+1)/2}$, $c_k = \dfrac{\Gamma((k+1)/2)}{\Gamma(k/2)(\pi k)^{1/2}}$, $\bar{F}_k(x) = 1 - F_k(x) = \int_x^\infty f_k(t)dt$, and $R_x = \bar{F}_k(x)/((1+x^2/k)f_k(x))$, for real $k > 0$ and real $x > 0$ here and throughout the paper. Then we will show here that

---

Sponsored by the United States Army under Contract No. DAAG29-80-C-0041 and the University of Wisconsin-Milwaukee Graduate School.

$$\left[\frac{(k-1)x}{2k} + \left[1+(\frac{(k+1)x}{2k})^2\right]^{\frac{1}{2}}\right]^{-1} < R_x < \left[\frac{(3k-1)x}{4k} + \left[\frac{k-1}{2k} + (\frac{(k+1)x}{4k})^2\right]^{\frac{1}{2}}\right]^{-1} , \tag{2.2}$$

where the lower bound is valid for $k \geq 1$ and the upper for $k \geq 2$. Note that if for fixed x, $k \to \infty$ in (2.2) we obtain (2.1) with "$\leq$" replacing "$<$". A result from Soms (1976) which will be used subsequently is that

$$1/x - k/((k+2)x^3) < R_x < 1/x , \tag{2.3}$$

for real $k > 0$ (actually the proof assumed k a positive integer but this fact was nowhere used).

3. Proofs

Let $\nu(x) = 1/R_x$. Then

$$\lambda(x) = \nu'(x) = \nu(x)(\nu(x) - \frac{k-1}{k} x)/(1+x^2/k) \tag{3.1}$$

and the first result is that $0 < \lambda(x) < 1$ for $k \geq 1$ (here and throughout, " ' " means the derivative).

<u>Theorem 3.1</u>: Let $\lambda(x)$ be given by (3.1). Then $0 < \lambda(x) < 1$ for $k \geq 1$.

<u>Proof</u>: For convenience, $\lambda(x)$ and $\nu(x)$ will be denoted by λ and ν. From (2.3) $\nu > x$, and hence $\lambda > 0$. Suppose $\lambda \geq 1$ for some x. Then there exists an x such that $\lambda \geq 1$ and $\lambda' = 0$, since from (2.3) $\lim_{x\to\infty} \lambda = 1$ and it will be shown in Lemma 3.1 that $\lambda(0) = 4c_k^2 < 1$. Now

$$\lambda' = \frac{\nu}{(1+x^2/k)^2}\,(\nu - \frac{k+1}{k}\,x)(\nu - \frac{k-1}{k}\,x) + \frac{\nu}{1+x^2/k}\left[\frac{\nu}{1+x^2/k}\,(\nu - \frac{k-1}{k}\,x) - \frac{k-1}{k}\right]$$

$$= \frac{\nu}{1+x^2/k}\left[\frac{(\nu-(k-1)x/k)(2\nu-(k+1)x/k)}{1+x^2/k} - \frac{k-1}{k}\right]$$

$$= \frac{\nu}{1+x^2/k}\left[\lambda(2 - \frac{k+1}{k}\,\frac{x}{\nu}) - \frac{k-1}{k}\right] > \frac{\nu}{1+x^2/k}\,(\frac{k-1}{k})(\lambda-1) \geq 0 \ ,$$

and hence $\lambda' > 0$, a contradiction, giving the conclusion.

From Theorem 3.1,

$$\nu(\nu - (k-1)x/k) < 1 + x^2/k \ . \tag{3.2}$$

Completing the square in (3.2) gives

$$(\nu - (k-1)x/2k)^2 < 1 + ((k+1)x/2k)^2 \ ,$$

or since $\nu - (k-1)x/2k > 0$ from (2.3),

$$\left[\frac{(k-1)x}{2k} + \left(1 + (\frac{(k+1)x}{2k})^2\right)^{\frac{1}{2}}\right]^{-1} < R_x , \tag{3.3}$$

the desired result.

Recall that

$$\lambda' = \frac{\nu}{1+x^2/k}\left[\frac{(\nu-(k-1)x/k)(2\nu-(k+1)x/k)}{1+x^2/k} - \frac{k-1}{k}\right] =$$

$$= \frac{\nu}{1+x^2/k}\,(\phi - \frac{k-1}{k}),\ \phi = \frac{(\nu-(k-1)x/k)(2\nu-(k+1)x/k)}{1+x^2/k} \ . \tag{3.4}$$

Then the second result which will give the upper bound in (2.2) is

<u>Theorem 3.2</u>: Let λ' and ϕ be as in (3.4). Then, for $k \geq 2$, $\lambda' > 0$ or, equivalently, $\phi > (k-1)/k$.

<u>Proof</u>: Suppose $\lambda' \leq 0$ for some x. Then there exists an x such that $\phi \leq (k-1)/k$ and $\phi' = 0$, since $\phi(0) = 8c_k^2 > (k-1)/k$ by Lemma 3.1 and $\lim_{x\to\infty} \phi(x) = (k-1)/k$, since $\lim_{x\to\infty} \nu/x = 1$ by (2.3). Now using (3.1),

$$\phi' = \frac{[(\lambda-(k-1)/k)(2\nu-(k+1)x/k) + (\nu-(k-1)x/k)(2\lambda-(k+1)/k)](1+x^2/k)-2(1+x^2/k)x\phi/k}{(1+x^2/k)^2}$$

$$= \frac{(\nu-2x/k)\phi-((k-1)/k)(2\nu-(k+1)x/k) + (\nu-(k-1)x/k)(2\lambda-(k+1)/k)}{(1+x^2/k)} \quad . \tag{3.5}$$

Suppose $\nu \leq 2x/k$. Then $x < \nu \leq 2x/k$, since $\nu > x$ by (2.3), and so $kx < 2x$, which is impossible since $k \geq 2$, and so $\nu > 2x/k$. Adding and subtracting $(\nu-2x/k)(k-1)/k$ to the numerator of (3.5), gives, after some simple algebra, $\phi' = [(\nu-2x/k)(\phi-(k-1)/k) + 2(\nu-(k-1)x/k)(\lambda-1)]/(1+x^2/k) < 0$, since $\phi \leq (k-1)/k$ by assumption and $\lambda < 1$ by Theorem 3.1, which is a contradiction and completes the proof.

From Theorem 3.2,

$$(\nu-(k-1)x/k)(2\nu-(k+1)x/k) > (k-1)(1+x^2/k)/k \ , \tag{3.6}$$

and completing the square in (3.6) gives

$$(\nu-(3k-1)x/4k)^2 > (8(k-1)/k + ((k+1)x/k)^2)/16 \ , \tag{3.7}$$

and since $\nu > (3k-1)x/4k$, (3.7) gives, after some algebra,

$$R_x < \left[\frac{(3k-1)x}{4k} + \left(\frac{k-1}{2k} + (\frac{(k+1)x}{4k})^2\right)^{\frac{1}{2}}\right]^{-1} \ , \tag{3.8}$$

which is the remaining part of (2.2).

To complete the proofs of Theorems 3.1 and 3.2 we need

Lemma 3.1: $1-4c_k^2 > 0$ for $k > 0$ and $8c_k^2 > (k-1)/k$ for $k \geq 2$.

Proof: The proof is an immediate consequence of the fact shown in Soms (1979) that for $k > 0$ the c_k are a strictly increasing function of k with limit $1/\sqrt{2\pi}$. For the sake of completeness we provide a brief outline of that proof. Let $g(k) = \frac{d}{dk} \ln c_k = \frac{1}{2}\Psi(\frac{k+1}{2}) - \frac{1}{2}\Psi(\frac{k}{2}) - \frac{1}{2k}$, where $\Psi(x)$ is the digamma function (see, e.g., Abramowitz and Stegun, 1965, p. 258). From Artin (1964,

p. 17), $h(k) = \frac{1}{2}\,\Psi(\frac{k+1}{2}) - \frac{1}{2}\,\Psi(\frac{k}{2}) = \sum_{i=0}^{\infty} \frac{1}{(k+2i)(k+2i+1)}$. By the Euler-MacLaurin formula,

$$\frac{1}{2}\,\ln\frac{k+1}{k} + \frac{1}{2}\,\frac{1}{k(k+1)} \le h(k) \ ,$$

and so $h(k) > \frac{1}{2k}$, giving the conclusion.

It is reasonable to ask whether consideration of derivatives of R_x would also lead to bounds. In fact it does lead to analogues of the bounds of Gordon (1941), but these bounds, just as in the normal case, are inferior to (2.2). For the details of these and some other second-best bounds the reader is referred to Soms (1977a).

4. Theoretical Comparisons

We give here the three types of other bounds with which either numerical, theoretical, or both comparisons will be made. The bounds of Peizer and Pratt (1968) are

$$\frac{1}{x}\,\frac{x^2(k+2)}{x^2(k+2)+k} < R_x < \frac{1}{x}\,\frac{(k+2)(k+4)x^2+2(k+1)k}{(k+2)(k+4)x^2+2(k+1)k+k(k+4)} \ , \tag{4.1}$$

the bounds of Wallace (1959)

$$\frac{1-\Phi[(k\,\ln(1+x^2/k))^{\frac{1}{2}}]}{(1+x^2/k)f_k(x)} < R_x < \frac{1-\Phi[(1-1/2k)^{\frac{1}{2}}(k\,\ln(1+x^2/k))^{\frac{1}{2}}]}{(1+x^2/k)f_k(x)}, \tag{4.2}$$

$k > \frac{1}{2}$ for the upper bound, and of Soms (1977b)

$$p(x,\gamma_{min}) < R_x < p(x,\gamma_{max}) \ , \tag{4.3}$$

where $p(x,\gamma) = \frac{1+\gamma}{(x^2+4c_k{}^2(1+\gamma)^2)^{\frac{1}{2}}+\gamma x}$, for $k > 2$, $\gamma_{min} = \frac{k}{2(k+2)c_k{}^2} - 1$, $\gamma_{max} = 4c_k{}^2/(1-4c_k{}^2)$, for $k < 2$ the definitions of γ_{min} and γ_{max} are interchanged and for $k = 2$, $\gamma_{min} = \gamma_{max}$ and $R_x = p(x,\gamma_{max})$.

Let us denote the lower bounds of (2.2), (4.1), (4.2), and (4.3) by SL, PL, WL, and BL, respectively, and the upper bounds by SU, PU, WU, and BU.

Then, after some algebra, comparison of (2.2) with (4.1) yields the following results. SL is better (bigger) than PL for $x < k/(k+2)^{\frac{1}{2}}$. SU is better (smaller) than PU for $x < -\frac{b}{2a} + (\frac{b^2-4ac}{4a^2})^{\frac{1}{2}}$, where

$$a = -2(k+2)(k+4)^2 ,$$

$$b = -2k(k+4)(2k^2+3k+4) ,$$

and $$c = 4(k-1)k^2(k+1)^2 .$$

Analytical comparisons with Wallace (1959) appear impossible.

BL is better than SL for $k > k_o$, where $1 < k_o < 2$, and for sufficiently large x and $1 \leq k \leq k_o$, SL is better. For sufficiently large x, SU is better than BU. Further analytical comparisons here are prohibitively difficult.

We note here that for $k = 1$, we have the Cauchy distribution and (2.2) and (2.3) combine for the interesting simple and sharp result

$$\frac{1}{\pi(1+x^2)^{\frac{1}{2}}} < \frac{1}{\pi} \int_x^\infty \frac{1}{1+t^2} dt < \frac{1}{\pi x} .$$

5. Numerical Comparisons

Let $g(x) = (1+x^2/k)f_k(x)$. Then (2.2) gives

$$g(x)SL < \bar{F}_k(x) < g(x)SU , \tag{5.1}$$

and we can approximate $\bar{F}_k(x)$ by $L = g(x)SL$ and $U = g(x)SU$. This is done in Table 1 for selected degrees of freedom and known t percentiles (Cramer, 1946, p. 560).

1. S-Approximations to $\bar{F}_k(x)$

	Degrees of Freedom								
$\bar{F}_k(x)$	6			20			120		
	x	L	U	x	L	U	x	L	U
.1	1.440	.0956	.102	1.325	.0959	.101	1.289	.0959	.101
.05	1.943	.0487	.0505	1.725	.0486	.0503	1.658	.0487	.0502
.025	2.447	.0246	.0252	2.086	.0245	.0251	1.980	.0245	.0251
.010	3.143	.00989	.0100	2.528	.00988	.0100	2.358	.00988	.0100
.005	3.707	.00496	.00502	2.845	.00496	.00501	2.617	.00496	.00501
.0005	5.959	.000498	.000500	3.850	.000497	.000500	3.373	.000498	.000501

From Table 1 is appears that the approximation using the SL and SU-bounds gives good results if $\bar{F}_k(x)$ is less than or equal to .1.

In Table 2, SL, PL, WL, and BL are given and in Table 3, SU, PU, WU, and BU, for selected x and k. In the calculation of the WL bounds for extreme tails the Boyd (1959) approximation for the tail area of the normal distribution was used. A "-" indicates the bound is inapplicable.

While the superiority of SL to BL for $k = 1$ is not apparent from the tables, it is seen by direct calculation that, for example, for $x = 50$ SL is bigger than BL. In general, for lower bounds BL is best while the picture is mixed for upper bounds. For x larger than in the tables the performance of SL and SU would be better than for BL and BU.

2. Lower Bounds to R_x

	Degrees of Freedom											
x	1				10				20			
	SL	PL	WL	BL	SL	PL	WL	BL	SL	PL	WL	BL
.5	.8944	.8571	1.0000	1.1046	.7923	.4615	.8892	.8949	.7866	.4314	.8828	.8807
1.0	.7071	.7500	.6363	.7834	.6284	.5455	.6490	.6687	.6233	.5238	.6521	.6591
2.0	.4472	.4615	.3213	.4630	.4190	.4138	.3890	.4277	.4167	.4074	.4031	.4238
4.0	.2425	.2449	.1450	.2449	.2372	.2376	.1888	.2384	.2366	.2366	.2049	.2375
10.0	.0995	.0997	.0498	.0997	.0991	.0992	.0609	.0992	.0991	.0991	.0674	.0991
20.0	.0499	.0500	.0225	.0500	.0499	.0499	.0257	.0499	.0499	.0499	.0279	.0499

3. Upper Bounds to R_x

x	Degrees of Freedom											
	1				10				20			
	SU	PU	WU	BU	SU	PU	WU	BU	SU	PU	WU	BU
.5	-	1.2156	1.1598	1.1092	.9549	1.3034	.9020	.9045	.9343	1.3388	.8891	.8921
1.0	-	.7917	.8735	.7860	.6897	.7349	.6735	.6766	.6780	.7403	.6644	.6688
2.0	-	.4638	.5807	.4637	.4315	.4322	.4310	.4312	.4271	.4301	.4258	.4282
4.0	-	.2450	.3675	.2450	.2388	.2385	.2450	.2392	.2378	.2377	.2400	.2385
10.0	-	.0997	.2022	.0997	.0992	.0992	.1135	.0992	.0991	.0991	.1067	.0992
20.0	-	.0500	.1310	.0500	.0499	.0499	.0665	.0499	.0499	.0499	.0605	.0499

References

Abramowitz, Milton and Stegun, Irene A. (1965), Handbook of Mathematical Functions, Washington: National Bureau of Standards.

Artin, Emil (1964), The Gamma Function, New York: Holt, Rinehart and Winston.

Birnbaum, Z. W. (1942), "An Inequality for Mills' Ratio," Annals of Mathematical Statistics, 13, 245-6.

Boyd, A. V. (1959), "Inequalities for Mills' Ratio," Reports of Statistical Applications Research, Japanese Union of Scientists and Engineers, 6, 44-6.

Cramer, Harald (1946), Mathematical Methods of Statistics, Princeton: Princeton University Press.

Gordon, Robert D. (1941), "Values of Mills' Ratio of Area to Bounding Ordinate and of the Normal Probability Integral for Large Values of the Argument," Annals of Mathematical Statistics, 12, 364-6.

Johnson, Norman L., and Kotz, Samuel (1970), Continuous Univariate Distributions-2, Boston: Houghton Mifflin Co.

Peizer, D. B., and Pratt, John W. (1968), "A Normal Approximation for Binomial, F, Beta, and Other Common, Related Tail Probabilities, I," Journal of the American Statistical Association, 63, 1416-56.

Sampford, M. R. (1953), "Some Inequalities on Mills' Ratio and Related Functions," Annals of Mathematical Statistics, 24, 130-2.

Soms, Andrew P. (1976), "An Asymptotic Expansion for the Tail Area of the t-Distribution," Journal of the American Statistical Association, 71, 728-30.

Soms, Andrew P. (1977a), "Bounds for the t-Tail Area," Technical Summary Report No. 1806, Mathematics Research Center, University of Wisconsin-Madison.

Soms, Andrew P. (1977b), "Rational Bounds for the t-Tail Area With an Application to Bonferroni t-Percentiles," Mathematics Research Center, University of Wisconsin-Madison, and to appear in the Journal of the American Statistical Association.

Soms, Andrew P. (1979), "A Note on an Extension of Rational Bounds for the t-Tail Area to Arbitrary Degrees of Freedom," submitted for publication.

Wallace, David L. (1959), "Bounds on Normal Approximation to Student's and the Chi-Square Distributions," Annals of Mathematical Statistics, 30, 1121-30.

(14) MRC-TSR-2105

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER 2105	2. GOVT ACCESSION NO. AD-A093566	3. RECIPIENT'S CATALOG NUMBER (9) Technical
4. TITLE (and Subtitle) BOUNDS FOR THE t-TAIL AREA FOR ARBITRARY DEGREES OF FREEDOM.		5. TYPE OF REPORT & PERIOD COVERED Summary Report - no specific reporting period
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Andrew P. Soms		8. CONTRACT OR GRANT NUMBER(s) DAAG29-80-C-0041
9. PERFORMING ORGANIZATION NAME AND ADDRESS Mathematics Research Center, University of Wisconsin, 610 Walnut Street, Madison, Wisconsin 53706		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS Work Unit Number 4 - Statistics and Probability
11. CONTROLLING OFFICE NAME AND ADDRESS U. S. Army Research Office, P.O. Box 12211, Research Triangle Park, North Carolina 27709 (11)		12. REPORT DATE August 1980
		13. NUMBER OF PAGES 11
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office) (12) 15		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Bounds, t-Tail area, t-Distribution

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

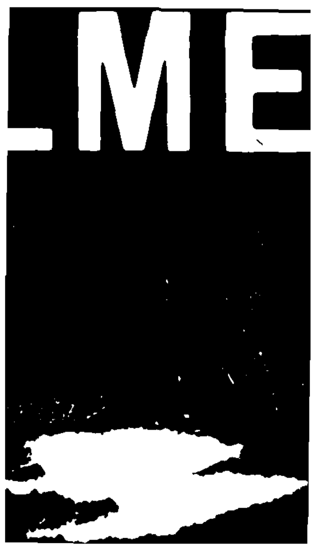
The bounds of Birnbaum (1942) and Sampford (1953) for the upper tail area of the normal distribution are extended to the upper tail of the t-distribution for arbitrary real degrees of freedom. This generalizes the results of Soms (1977a) for integral degrees of freedom. Numerical and theoretical comparisons are made with the bounds of Peizer and Pratt (1968), Wallace (1959) and Soms (1977b).

DD FORM 1473 1 JAN 73 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

221200

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)


LME