THE t-STATISTIC FOR UNDERLYING SKEW/STRETCHED-TAIL DISTRIBUTIONS

Susan Peterson Arthur

July 1980
PP-80-5

INTERNATIONAL INSTITUTE FOR APPLIED SYSTEMS ANALYSIS
A-2361 Laxenburg, Austria

ACKNOWLEDGEMENTS

The results reported in this study form part of my Ph.D. dissertation, Princeton University, 1979. I would like to thank my advisor, John Tukey, and the International Institute for Applied Systems Analysis for its generous support.

ABSTRACT

For underlying skew distributions, Student-t confidence intervals about the mean have unequal tail probabilities — the interval does not cover the mean in a "balanced" way. This study uses Monte-Carlo methods to estimate, for a class of highly skew, stretched-tail distributions, the population characteristic covered by the t-interval with symmetric loss. Results indicate that this "balanced" population characteristic depends on the degree of skewness and stretch, the desired significance level, and the sample size.

Estimates of the balanced population characteristic can be used to modify Student-t confidence intervals about the mean to achieve symmetric loss. The resulting tail probabilities are estimated and are found to be reasonably close to desired levels for many cases. Most of the discrepancy between true tail probabilities and tabled Student-t values is corrected, for the distributions of this study, by this simple modification.

The reults of this study are applicable to a family of underlying distributions that are more skew and stretched-tail than generally considered in robustness studies of the t-statistic. Furthermore, the approach does not require large samples — results are given for small to moderate sample sizes.

The t-Statistic for Underlying Skew/Stretched-tail Distributions

Susan Peterson Arthur

1. INTRODUCTION

For non-gaussian underlying distributions, the distribution of sample t values is in general unknown. Nevertheless, critical values of the Student-t distribution are frequently used to form approximate confidence intervals about the population mean. When the underlying distribution is skew, this produces unequal tail probabilities. If $x_1, \ldots, x_n$ are a sample from a gaussian underlying distribution, then

$$t=(\bar{x}-\mu)/s_n\ ,\quad s^2=\sum_{i=1}^{n}(x_i-\bar{x})^2/(n-1),\quad s_n=sn^{-1/2}$$

has a Student-t distribution with n-1 degrees of freedom and we can form confidence intervals

$$P(\bar{x}-\tau_\alpha s_n\leq\mu\leq\bar{x}+\tau_\alpha s_n)=1-2\alpha,$$
$$P(\mu<\bar{x}-\tau_\alpha s_n)=P(\mu>\bar{x}+\tau_\alpha s_n)=\alpha,$$

where τ_α is the Student-t critical point corresponding to one-sided probability α. For a skew underlying distribution,

$$P(\bar{x}-\tau_\alpha s_n\leq\mu\leq\bar{x}+\tau_\alpha s_n)=1-\alpha_R-\alpha_L\ ,$$
$$P(\mu<\bar{x}-\tau_\alpha s_n)=\alpha_R,\quad P(\mu>\bar{x}+\tau_\alpha s_n)=\alpha_L\ ,$$

where in general $\alpha_R\neq\alpha_L$, and $\alpha_R+\alpha_L$ is not necessarily equal to 2α. The interval is doing a poor job of capturing the population mean, μ, since μ is more likely to lie outside the interval on one side than the other (asymmetric loss). We may say, in this case, that the t-interval does not cover the mean in a balanced way.

This study uses Monte Carlo methods to find, for sample sizes n= 5,10,20 and a class of highly skew/stretched-tail underlying distributions, the population characteristic which the t confidence interval does cover in a balanced way. That is, the population characteristic, c, for which $t=(\bar{x}-c)/s_n$ has equal tail probabilities $P(c < \bar{x}-\tau_\alpha s_n) = P(c > \bar{x}+\tau_\alpha s_n) = \alpha^*$. In a sense, this is the population characteristic which is captured by Student-t intervals.

Once the balanced population characteristic is known, it can be used to modify t confidence intervals about the mean to achieve equal tail probabilities. After the t-statistic has been thus modified, tail probabilities are estimated and compared to the nominal Student-t probability, α. Tail probabilities for the modified intervals are found to be reasonably close to α in many cases, and are much closer than are the tail probabilities associated with unmodified Student-t intervals. Most of the discrepancy between true tail probabilities for t-statistics based on the underlying distributions of this study and tabled Student-t values is corrected by this simple modification.

Further modifications in the choice of t-distribution critical values to achieve desired significance levels are developed for a class of symmetric, stretched-tail distributions. These modifications can also be used for some of the skew distributions, applied to t-statistics previously modified for equal tails (i.e. balanced).

Robustness of t with respect to non-normality has been of considerable interest to statisticians. For a comprehensive review of the literature, see Hatch and Posten (1966). In contrast to this paper, however, most studies are primarily interested in how close $\alpha_R+\alpha_L$ is to 2α, not in the asymmetry of loss found for underlying skew distributions. Furthermore, suggested modifications of the Student-t procedure are based almost entirely on series approximation techniques – requiring either that the underlying distribution be relatively close to gaussian, or that the sample size be large. Neither is required for the modifications proposed in this study.

The next section of the paper describes the underlying distributions to be considered (a class of distributions suggested by Tukey – the so-called g/h family). In Section 3 the population characteristic covered by Student-t confidence intervals with equal tail probability is estimated for several of the skew g/h distributions, for sample sizes n= 5, 10, 20, and one-sided significance levels $\alpha =$ 0.01, 0.025, 0.05, and 0.10. In Section 4, tail probabilities are estimated — for t-statistics from a class of symmetric, stretched-tail distributions and for modified t-statistics from the skew distributions of Section 3. Tail probabilities are compared to tabled Student-t values and modifications of critical values are developed for the symmetric h-distributions.

2. The Distributions

The distributions used in this study belong to a four-parameter family suggested by Tukey (1976), the g/h class of distributions. This family is particularly well-suited for robustness studies, as it covers a range of shapes much wider than the distributions usually encountered in empirical work.

The g/h family is based on simple monotonic functions of a unit gaussian variate, making the generation of samples and the derivation of densities and moments easy. One function induces skewness (and some stretch), while another stretches the tails symmetrically.

The transformation for skewness is:

$$y = (e^{gz} - 1)/g, \quad g \neq 0, \tag{2.1}$$

where z is a unit gaussian variate, and skewness is indexed by g. Re-expressions of this type form the familiar family of log-normal distributions. The -1 in the numerator and g in the denominator cause $y(0)$ to be zero and $y(\epsilon) \approx \epsilon$ for ϵ close to zero.

The transformation for symmetric stretching of tails is:

$$y = z \exp(hz^2/2), \quad h \geq 0, \tag{2.2}$$

where z is again a unit gaussian variate, and h is the index of stretch. Large values of h imply more stretching. Re-expressions of this type form a family we will refer to as the class of h-distributions. Squeezed-tail distributions cannot be generated by taking h< 0, as (2.2) is then no longer monotonic.

The two transformations can be combined,

$$y = (e^{gz} - 1)g^{-1}\exp(hz^2/2), \quad g \neq 0, \quad h \geq 0 \tag{2.3}$$

to generate distributions which are skew and more stretched than the log-normals. Distributions generated by these transformations form the g/h family.

In this study we consider values of g from 0.1 to 2.5 and values of h from 0.05 to 0.9. Although we look only at positive skewness, the development and results are analogous for negative skewness (g< 0).

Densities can be derived by simple transformation-of-variable techniques. Letting f(y) denote the density of y, we have for the log-normals (2.1),

$$f(y) = (2\pi)^{-1/2}(gy+1)^{-1}\exp(-(\ln(gy+1))^2/2g^2), \quad y > -1/g. \tag{2.4}$$

For the h-distributions (2.2),

$$f(y) = (2\pi)^{-1/2}(hz^2+1)^{-1}\exp(-z^2(1+h)/2), \tag{2.5}$$

where z is such that $y = z\exp(hz^2/2)$. Finally, the density for the combined g/h distributions is:

$$f(y) = (2\pi)^{-1/2}(e^{gz} + hz(e^{gz}-1)/g)^{-1}\exp(-z^2(1+h)/2), \tag{2.6}$$

where z is such that $y = (e^{gz}-1)g^{-1}\exp(hz^2/2)$. Some of the log-normal densities are plotted in Figure A. In Figure B the square root of the density is plotted for a few of the g/h-distributions, to magnify tail behaviour.

Clearly, these are highly non-normal distributions. To see just how non-normal, we compare them to the well-known Karl Pearson curves. The Pearson curves are often used in empirical work, and form the basis of E.S. Pearson's two extensive robustness studies of the t-statistic (1929 and 1975).

Pearson curves rely heavily on moments for measuring characteristics of shape. For the g/h family moments do not play such a central role — in fact, moments often do not exist. However, since the log-normal, h-, and g/h-distributions are simple re-expressions of a unit gaussian, moments can be derived in a straightforward manner. Detailed derivations are given in Arthur (1979). Table 1 lists population moments for various log-normal, h-, and g/h-distributions. Note that for the h- and g/h-distributions, μ_1 exists only for h< 1, μ_2 for h< 1/2, μ_3 for h< 1/3, and μ_4 for h< 1/4. Of course, since the h-distributions are symmetric

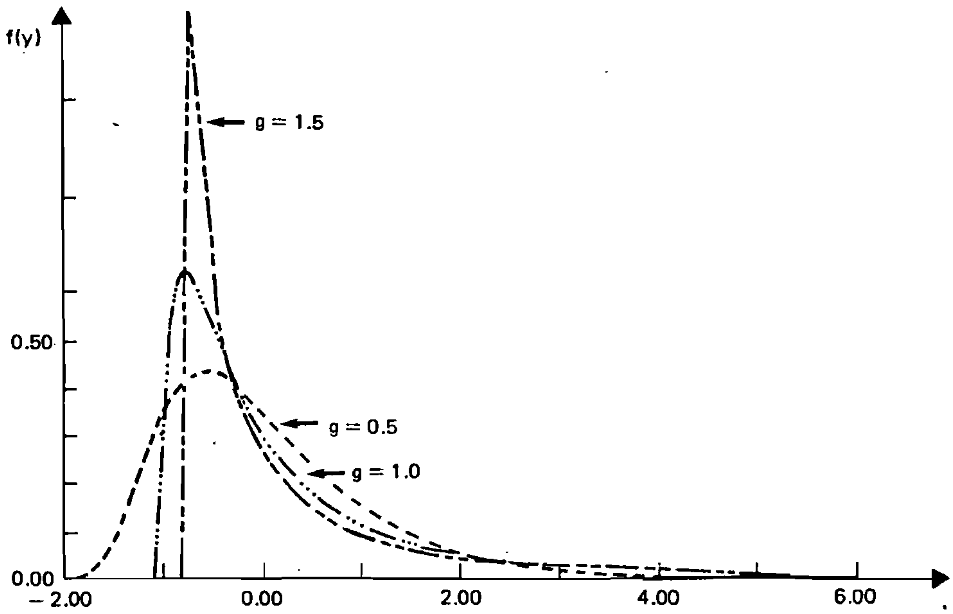

FIGURE A. Densities of the Log-normals

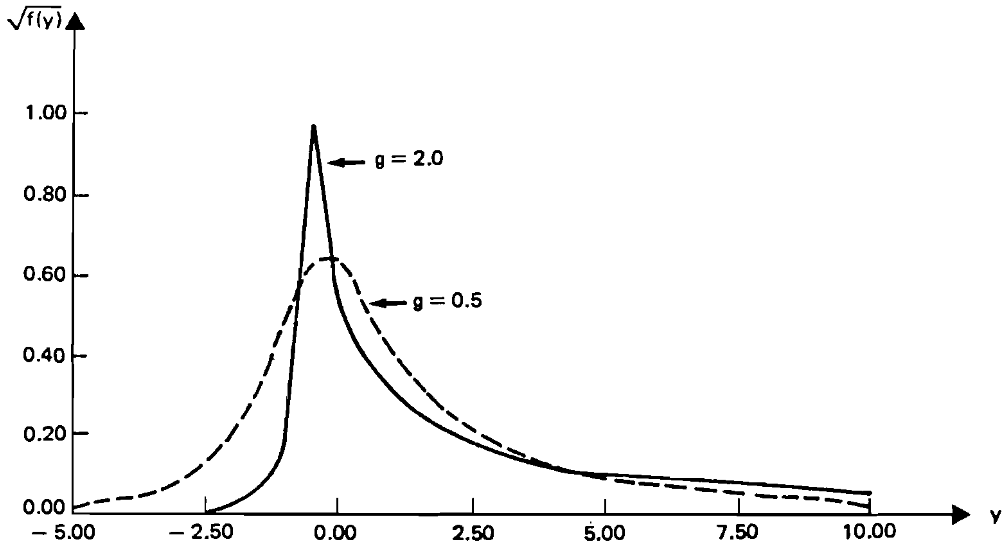

FIGURE B. Square Root of Density for g/h-Distributions, h = 0.2

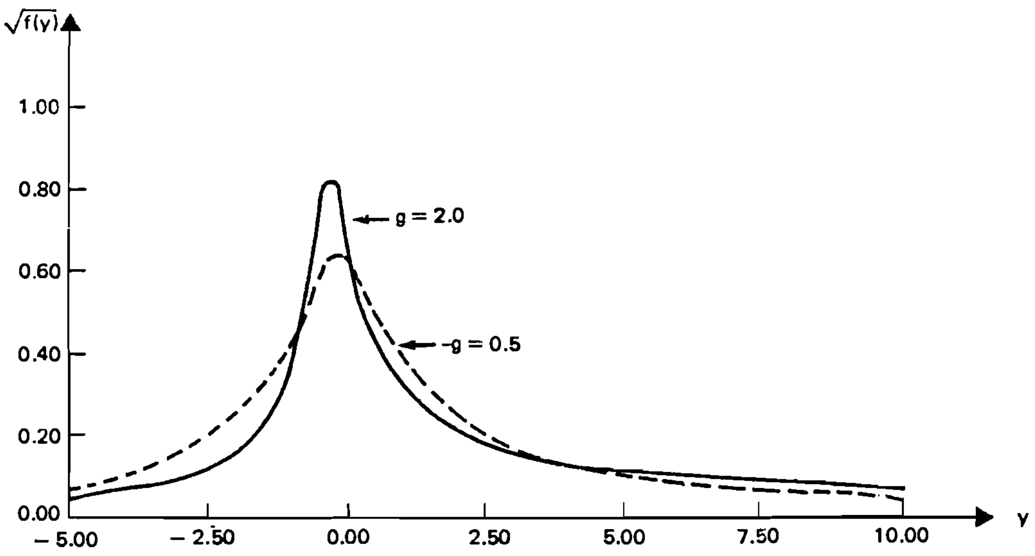

FIGURE B. Square Root of Density for g/h-Distributions, h = 0.8

Table 1. Moments of the Distributions

a. Log-normal Distributions

g=	μ	μ_2	$\beta_1 1/2$	β_2
0.1	0.050	1.015	0.302	3.162
0.2	0.101	1.062	0.614	3.678
0.3	0.153	1.145	0.950	4.645
0.4	0.208	1.273	1.322	6.260
0.5	0.266	1.459	1.750	8.898
0.7	0.397	2.106	2.888	20.790
1.0	0.649	4.671	6.185	113.900
1.5	1.387	35.790	33.470	1.008×10^4
2.0	3.194	731.600	414.400	9.221×10^6
2.5	8.704	4.285×10^4	1.182×10^4	$>1.00\times10^{10}$

b. h-Distributions

h=	μ	μ_2	$\beta_1 1/2$	β_2
0.05	0	1.171	0	3.820
0.10	0	1.398	0	5.508
0.15	0	1.708	0	10.169
0.20	0	2.152	0	36.224
0.40	0	11.180	0	-
0.60	0	-		-
0.80	0	-		-
0.90	0	-		-

Table 1. (cont'd) Moments of the Distributions

c. g/h-Distributions

	μ	μ_2	$\beta_1^{1/2}$	β_2
h=0.1				
g=0.1	0.059	1.425	0.521	6.192
g=0.3	0.180	1.664	1.702	13.410
g=0.5	0.314	2.272	3.407	44.240
g=0.7	0.471	3.628	6.517	220.100
g=1.0	0.783	9.948	21.040	7.942×10^{3}
g=1.5	1.750	131.200	450.200	1.584×10^{8}
g=2.0	4.336	6.131×10^{3}	4.577×10^{4}	$>1.00\times10^{10}$
g=2.5	13.160	1.092×10^{6}	1.886×10^{7}	$>1.00\times10^{10}$
h=0.2				
g=0.1	0.070	2.210	1.299	56.035
g=0.3	0.216	2.738	4.840	625.100
g=0.5	0.378	4.183	13.160	4.290×10^{4}
g=0.7	0.572	7.872	44.780	4.883×10^{7}
g=1.0	0.971	30.600	713.600	$>1.00\times10^{10}$
g=1.5	2.296	1.025×10^{3}	1.405×10^{6}	$>1.00\times10^{10}$
g=2.0	6.251	1.992×10^{5}	$>1.00\times10^{10}$	$>1.00\times10^{10}$
g=2.5	21.780	$>1.00\times10^{10}$	$>1.00\times10^{10}$	$>1.00\times10^{10}$
h=0.4				
g=0.1	0.108	12.180		
g=0.3	0.335	23.610		
g=0.5	0.598	84.130		
g=0.7	0.930	585.400		
g=1.0	1.680	4.920×10^{4}		
g=1.5	4.752	5.874×10^{9}		
g=2.0	17.450	$>1.00\times10^{10}$		
g=2.5	93.880	$>1.00\times10^{10}$		
h=0.9				
g=0.1	1.621			
g=0.3	5.991			
g=0.5	15.750			
g=0.7	47.830			
g=1.0	466.200			
g=1.5	1.621×10^{5}			
g=2.0	7.671×10^{8}			
g=2.5	$>1.00\times10^{10}$			

and centered on zero, $\mu_i = 0$, i odd, for these distributions. Also, since the three types of distribution have y= 0 at z= 0, the population median is zero in all cases.

Figure C is adapted from a diagram presented in the 1963 paper of Johnson, Nixon, Amos, and E.S. Pearson on percentage points of Pearson curves. It plots values of $\beta_1^{1/2}(=\mu_3/\mu_2^{3/2})$ and $\beta_2(=\mu_4/\mu_2^2)$ for various types of Pearson curves. Values of $\beta_1^{1/2}$ and β_2 have been added for several g/h-distributions. In contrast to the Pearson curves, the g/h-distributions have only one sharp boundary, at h= 0, and cover the entire $(\beta_1^{1/2},\beta_2)$-plane below the h= 0 line (log-normals) in a smooth way. However, they do not include any distributions with tails less stretched than the log-normals.

The shaded area in Figure C shows the boundaries of the distributions included by E.S. Pearson and Please in their 1975 study of the effects of non-normality on t, the most extensive empirical study to date. The g/h-distributions to be used in this study include only about one-third of this area, but extend to distributions with much more extreme skewness and stretch than those of the Pearson and Please study.

In sum, the g/h-distributions are a family particularly well suited to studies of extreme distributions. Their simple relation to the gaussian makes densities, percent points and moments easy to find, for any values of g and h. Extensive tables such as those required by the Pearson curves are unnecessary. Using this family allows us to look at the robustness of t with respect to underlying distributions with much more extreme skewness and stretch than has been done before.

3. The Balanced Center for Underlying Skew Distributions

In this section we estimate the population characteristic covered in a balanced way for several of the g/h distributions, several sample sizes and significance levels. This means finding c such that $P(c < \bar{x} - \tau_\alpha s_n) = P(c > \bar{x} + \tau_\alpha s_n) = \alpha^*$. We do not require that α^* be close to the nominal significance level, α, only that the right- and left-hand tails be equal. The value of α^* is, of course, also important and will be discussed in the next section. The population charac-

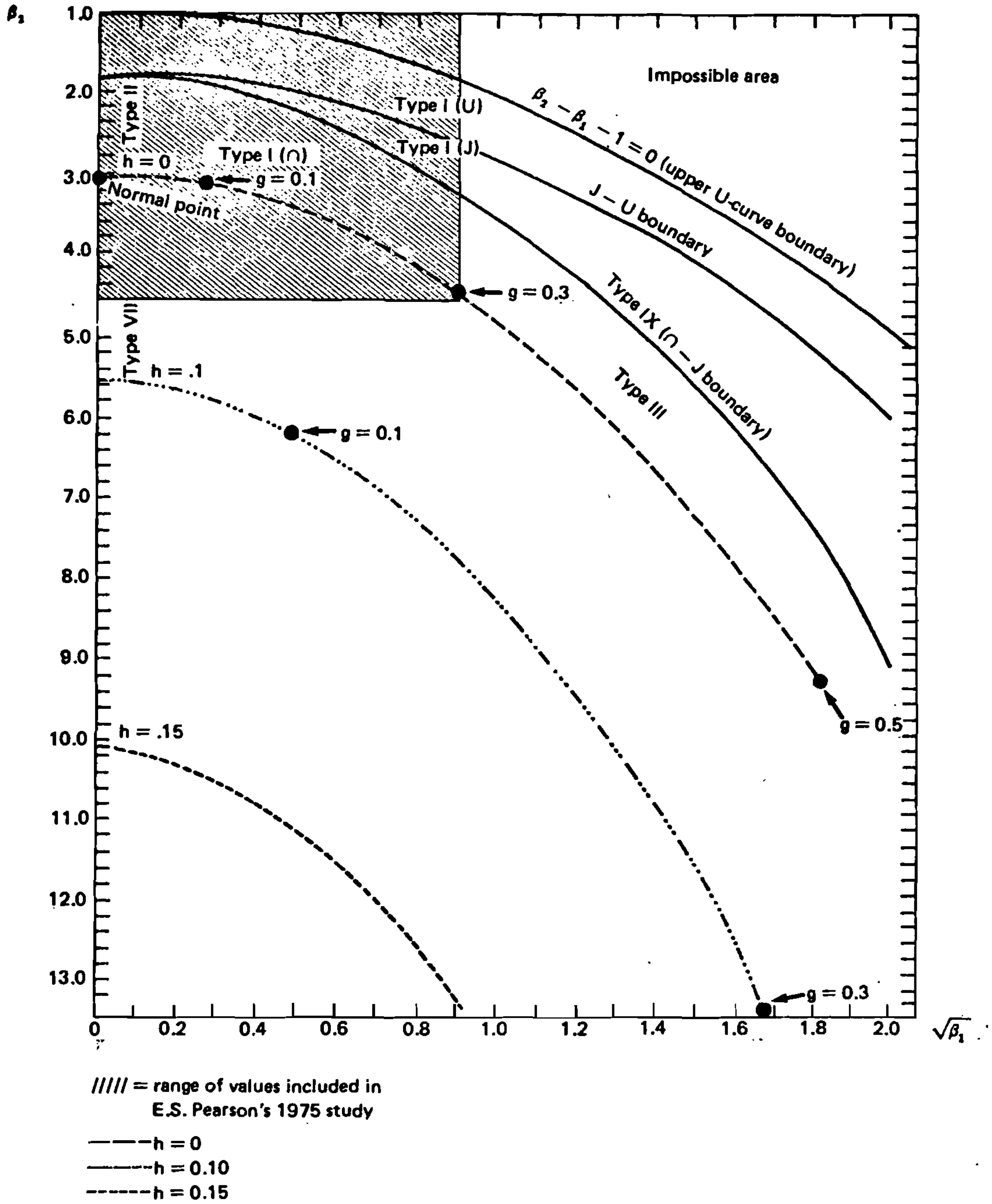

FIGURE C. Values of $(\sqrt{\beta_1}, \beta_2)$ for several g/h-Distributions teristic c will be referred to as the balanced center.

The actual numerical value of c is probably of little interest. Other aspects of the balanced center are more important. In this study we look at:

— the location of the balanced center relative to the mean (μ) and median (m) of the underlying distribution, denoted by a where $c = a\mu + (1-a)m$; and

— the percent point of the underlying distribution at which the balanced center lies, denoted $pct(c)$.

The first of these is the most important for this study because knowing a allows us to modify t-intervals about μ to achieve equal tails. Let

$$t_a = (\bar{x} - (a\mu + (1-a)m)) \,/\, s_n,$$

where $c = a\mu + (1-a)m$ is the balanced center. Then we know

$$P(-\tau_\alpha \le t_a \le \tau_\alpha) = 1-2\alpha^*,$$
$$P(t_a < -\tau_\alpha) = P(t_a > \tau_\alpha) = \alpha^*.$$

Re-arranging terms, we can find:

$$P(\ (\bar{x}-\tau_\alpha s_n-(1-a)m)/a \le \mu \le (\bar{x}+\tau_\alpha s_n-(1-a)m)/a\ ) = 1-2\alpha^*,$$
$$P(\mu < (\bar{x}-\tau_\alpha s_n-(1-a)m)/a\ ) = P(\mu > (\bar{x}+\tau_\alpha s_n-(1-a)m)/a\ ) = \alpha^*.$$

In essence, the μ confidence interval is no longer symmetric about $\bar{x}$. The modified confidence interval for underlying skew distributions has one boundary farther from $\bar{x}$ than the other. However, the probability that μ lies outside this confidence interval to the left now equals the probability that it lies outside the confidence interval to the right. The modified confidence interval has symmetric loss — it is balanced.

3.1. Estimation

To estimate a, samples are generated for given sample size and given parameters g and h. The value of t_i, $t_i = (\bar{x}-(i\mu+(1-i)m))/s_n$, is calculated for several values of i. Tail probabilities $P(t_i < -\tau_\alpha)$ and $P(t_i > \tau_\alpha)$ are estimated for the nominal significance level, α, and are used to form ratios

$$r_i = P(t_i < -\tau_\alpha)/P(t_i > \tau_\alpha).$$

Interpolation is then done on the r_i to find the value a for which $r_a = 1$. This is our estimate $\hat{a}$ of a. Accuracy will depend on the step size used in interpolation. Estimate $\hat{a}$ will be said to be interpolated to grid size 0.05 if interpolation is done between r_i and $r_{i+0.05}$. The variance of $\hat{a}$ is estimated by jackknife techniques — we generate ten groups of 1000 samples each and estimate a ten times, leaving out one of the groups of samples each time. This procedure allows jackknife estimation of both a and its variance.

In estimating tail probabilities $P(t_a < -\tau_\alpha)$ and $P(t_a > \tau_\alpha)$, accuracy can be increased by taking advantage of the known gaussian tail probabilities, using an estimation technique due to Fieller and Hartley (1954). Suppose we generate a gaussian sample, $x_1, \ldots, x_n$ and transform it to a g/h sample $y_1, \ldots, y_n$. We can calculate two t-statistics, t_x based on the gaussian sample, and t_y based on the g/h sample obtained by transformation. Now suppose we generate N gaussian samples of size n, divide the range of t into cells, and form cell counts. Let

n_{ij} = the number of samples for which t_y falls in cell i, and t_x falls in cell j,

p_{ij} = probability that t_y falls in cell i, t_x in cell j.

Then the $p_{.j}$, where dot denotes summation, are known, and the Fieller-Hartley estimate of $p_{i.}$ is

$$\hat{p}_{i.} = \sum_j p_{.j} u_{ij}, \quad u_{ij} = n_{ij}/n_{.j} \quad \text{if } n_{.j} > 0$$
$$u_{ij} = n_{i.}/N \quad \text{if } n_{.j} = 0.$$

An estimate of the approximate variance is

$$\hat{var}(\hat{p}_{i.}) = \hat{p}_{i.}(1-\hat{p}_{i.})/N - (\sum_j (\hat{p}_{ij} - \hat{p}_{i.} p_{.j})^2 / N p_{.j} \; .$$

Estimates $\hat{p}_{i.}$ may be summed to form tail probability estimates.

3.2. Results

The value of a has been estimated for sample sizes n = 5, 10, 20, α = 0.01, 0.025, 0.05, 0.10, and several values of g and h. These estimates, and estimates of the standard error, are given in Table 2. All unstarred estimates in Table 2 were interpolated to grid size 0.05 and were jackknifed using ten groups of of 1000 samples each. Where estimates required more

TABLE 2. Estimates of a, Log-normals

	n=5 â	n=5 s.e.	n=10 â	n=10 s.e.	n=20 â	n=20 s.e.
g=0.1						
α=0.010	-0.377	0.056	-0.019	0.024	0.449	0.046
=0.025	-0.169	0.050	0.226	0.025	0.626	0.032
=0.050	-0.056	0.025	0.410	0.024	0.671	0.017
=0.100	0.181	0.015	0.577	0.037	0.793	0.018
g=0.3						
α=0.010	-0.394	0.038	0.021	0.026	0.418	0.033
=0.025	-0.218	0.022	0.227	0.018	0.578	0.010
=0.050	-0.032	0.013	0.381	0.012	0.668	0.008
=0.100	0.204	0.009	0.568	0.018	0.772	0.008
g=0.5						
α=0.010	-0.378	0.027	0.015	0.016	0.409	0.025
=0.025	-0.203	0.014	0.221	0.013	0.564	0.010
=0.050	-0.039	0.010	0.363	0.010	0.643	0.005
=0.100	0.188	0.009	0.552	0.011	0.752	0.007
g=0.7						
α=0.010	-0.342	0.015	0.035	0.012	0.395	0.013
=0.025	-0.196	0.013	0.211	0.011	0.517	0.011
=0.050	-0.036	0.011	0.347	0.008	0.609	0.005
=0.100	0.180	0.009	0.520	0.006	0.722	0.007
g=1.0						
α=0.010	-0.282	0.008	0.023	0.011	0.335	0.007
=0.025	-0.159	0.012	0.176	0.010	0.446	0.006
=0.050	-0.034	0.007	0.294	0.005	0.548	0.002
=0.100	0.153	0.006	0.465	0.005	0.664	0.007
g=1.5						
α=0.010	-0.172	0.005*	0.009	0.009*	0.227	0.006*
=0.025	-0.096	0.007*	0.112	0.005*	0.325	0.005*
=0.050	-0.024	0.006*	0.198	0.003*	0.411	0.003*
=0.100	0.101	0.004*	0.341	0.003*	0.528	0.005*
g=2.0						
α=0.010	-0.079	0.004*	0.002	0.003*	0.133	0.003*
=0.025	-0.054	0.001*	0.055	0.003*	0.200	0.004*
=0.050	-0.016	0.002*	0.114	0.003*	0.267	0.002*
=0.100	0.054	0.002*	0.209	0.002*	0.366	0.005*
g=2.5						
α=0.010	-0.0294	0.0009*	0.0002	0.0010*	0.0621	0.0023*
=0.025	-0.0217	0.0005*	0.0223	0.0017*	0.0995	0.0019*
=0.050	-0.0074	0.0009*	0.0515	0.0015*	0.1410	0.0015*
=0.100	0.0238	0.0009*	0.1062	0.0015*	0.2094	0.0036*

*: interpolation grid size 0.01

TABLE 2 (cont'd). Estimates of a, h=0.10

	n=5		n=10		n=20	
	â	s.e.	â	s.e.	â	s.e.
g=0.1						
α=0.010	-0.2962	0.0537	0.1094	0.0720	0.1878	0.0711
=0.025	-0.1682	0.0332	0.0722	0.0225	0.5368	0.0456
=0.050	-0.0299	0.0245	0.2879	0.0232	0.5772	0.0230
=0.100	0.1249	0.0196	0.5068	0.0297	0.7165	0.0184
g=0.3						
α=0.010	-0.3115	0.0438	0.0153	0.0218	0.3687	0.0215
=0.025	-0.1846	0.0166	0.1664	0.0195	0.4816	0.0107
=0.050	-0.0566	0.0129	0.2950	0.0120	0.5762	0.0083
=0.100	0.1474	0.0120	0.4824	0.0150	0.6943	0.0075
g=0.5						
α=0.010	-0.3054	0.0505	-0.0183	0.0169	0.3315	0.0145
=0.025	-0.1892	0.0199	0.1612	0.0093	0.4641	0.0108
=0.050	-0.0404	0.0148	0.2847	0.0090	0.5603	0.0036
=0.100	0.1395	0.0092	0.4737	0.0114	0.6790	0.0064
g=0.7						
α=0.010	-0.2838	0.0219	0.0033	0.0118	0.3194	0.0096
=0.025	-0.1626	0.0125	0.1525	0.0103	0.4232	0.0074
=0.05	-0.0398	0.0079	0.2700	0.0069	0.5275	0.0036
=0.10	0.1304	0.0071	0.4393	0.0055	0.6427	0.0070
g=1.0						
α=0.010	-0.2310	0.0114	-0.0070	0.0098	0.2649	0.0102
=0.025	-0.1321	0.0092	0.1273	0.0101	0.3713	0.0050
=0.050	-0.0409	0.0049	0.2304	0.0047	0.4662	0.0025
=0.100	0.1133	0.0067	0.3896	0.0051	0.5834	0.0062
g=1.5						
α=0.01	-0.1312	0.0052*	-0.0035	0.0031*	0.1732	0.0057*
=0.025	-0.0810	0.0053*	0.0778	0.0040*	0.2597	0.0038*
=0.05	-0.0266	0.0032*	0.1527	0.0037*	0.3363	0.0029*
=0.10	0.0731	0.0033*	0.2734	0.0034*	0.4475	0.0057*
g=2.0						
α=0.01	-0.0571	0.0026*	-0.0013	0.0020*	0.0953	0.0021*
=0.025	-0.0410	0.0012*	0.0368	0.0026*	0.1506	0.0023*
=0.05	-0.0146	0.0018*	0.0814	0.0025*	0.2016	0.0018*
=0.10	0.0369	0.0017*	0.1568	0.0017*	0.2868	0.0037*
g=2.5						
α=0.01	-0.0194	0.0006*	-0.0008	0.0007*	0.0405	0.0017*
=0.025	-0.0143	0.0004*	0.0132	0.0009*	0.0667	0.0013*
=0.05	-0.0054	0.0005*	0.0338	0.0009*	0.0960	0.0012*
=0.10	0.0147	0.0006*	0.0713	0.0009*	0.1480	0.0025*

*: interpolation grid size 0.01

TABLE 2 (cont'd). Estimates of a, h=0.20

	n=5 â	n=5 s.e.	n=10 â	n=10 s.e.	n=20 â	n=20 s.e.
g=0.1						
α=0.010	-0.1215	0.0844m	0.0384	0.0380m	0.2759	0.0367m
=0.025	-0.2416	0.0724m	0.1449	0.0421m	0.3919	0.0301m
=0.050	-0.0361	0.0239m	0.2121	0.0212m	0.5003	0.0193m
=0.100	0.0765	0.0223m	0.4218	0.0124m	0.6008	0.0163m
g=0.3						
α=0.010	-0.3194	0.0582	-0.0081	0.0435	0.2438	0.0119
=0.025	-0.1606	0.0340	0.1137	0.0121	0.3868	0.0113
=0.050	-0.0466	0.0126	0.2396	0.0127	0.4859	0.0074
=0.100	0.1051	0.0120	0.4013	0.0144	0.6102	0.0066
g=0.5						
α=0.010	-0.2383	0.0146	-0.0097	0.0210	0.2524	0.0134
=0.025	-0.1403	0.0155	0.1156	0.0109	0.3697	0.0112
=0.050	-0.0403	0.0105	0.2237	0.0076	0.4654	0.0037
=0.100	0.1042	0.0092	0.3870	0.0108	0.5943	0.0085
g=0.7						
α=0.010	-0.2195	0.0143	-0.0061	0.0111	0.2538	0.0117
=0.025	-0.1316	0.0131	0.1068	0.0107	0.3397	0.0050
=0.050	-0.0436	0.0068	0.2067	0.0073	0.4417	0.0037
=0.100	0.0958	0.0070	0.3635	0.0059	0.5592	0.0063
g=1.0						
α=0.010	-0.1741	0.0081	-0.0116	0.0054	0.2080	0.0103
=0.025	-0.1039	0.0081	0.0894	0.0087	0.2936	0.0040
=0.050	-0.0411	0.0062	0.1774	0.0035	0.3824	0.0029
=0.100	0.0819	0.0055	0.3158	0.0042	0.4980	0.0047
g=1.5						
α=0.010	-0.0947	0.0047*	-0.0065	0.0024*	0.1292	0.0035*
=0.025	-0.0620	0.0025*	0.0536	0.0031*	0.1983	0.0042*
=0.050	-0.0243	0.0033*	0.1113	0.0028*	0.2623	0.0015*
=0.100	0.0503	0.0028*	0.2103	0.0022*	0.3631	0.0047*
g=2.0						
α=0.010	-0.0377	0.0017*	-0.0017	0.0014*	0.0651	0.0015*
=0.025	-0.0276	0.0007*	0.0226	0.0018*	0.1041	0.0022*
=0.050	-0.0107	0.0013*	0.0540	0.0016*	0.1442	0.0017*
=0.100	0.0241	0.0014*	0.1111	0.0014*	0.2120	0.0033*
g=2.5						
α=0.010	-0.0112	0.0003*	-0.0012	0.0005*	0.0248	0.0007*
=0.025	-0.0087	0.0003*	0.0076	0.0005*	0.0405	0.0010*
=0.050	-0.0034	0.0003*	0.0197	0.0006*	0.0600	0.0009*
=0.100	0.0086	0.0004*	0.0443	0.0007*	0.0952	0.0016*

m: est. based on 10 groups of 2000 samples each
*: interpolation grid size 0.01

TABLE 2 (cont'd). Estimates of a, h=0.40

	n=5		n=10		n=20	
	â	s.e.	â	s.e.	â	s.e.
g=0.1						
α=0.010	-0.2371	0.0636	0.0323	0.0636	0.1880	0.0422
=0.025	-0.0574	0.0546	0.0966	0.0716	0.2836	0.0578
=0.050	-0.0474	0.0274	0.1423	0.0200	0.3428	0.0243
=0.100	0.0627	0.0181	0.2673	0.0156	0.4174	0.0223
g=0.3						
α=0.010	-0.1648	0.0256	-0.0228	0.0177	0.1475	0.0222
=0.025	-0.0954	0.0237	0.0639	0.0134	0.2408	0.0175
=0.050	-0.0354	0.0104	0.1508	0.0103	0.3239	0.0107
=0.100	0.0580	0.0107	0.2509	0.0078	0.4241	0.0068
g=0.5						
α=0.010	-0.1539	0.0136	-0.0209	0.0128	0.1542	0.0157
=0.025	-0.0868	0.0086	0.0568	0.0136	0.2253	0.0090
=0.050	-0.0342	0.0085	0.1371	0.0055	0.3031	0.0053
=0.100	0.0496	0.0079	0.2415	0.0077	0.4110	0.0059
g=0.7						
α=0.010	-0.1312	0.0068	-0.0068	0.0078	0.1573	0.0129
=0.025	-0.0786	0.0059	0.0585	0.0080	0.2068	0.0043
=0.050	-0.0306	0.0055	0.1188	0.0050	0.2773	0.0038
=0.100	0.0477	0.0055	0.2208	0.0047	0.3791	0.0054
g=1.0						
α=0.010	-0.0928	0.0044	0.0017	0.0038	0.1092	0.0068
=0.025	-0.0590	0.0038	0.0404	0.0055	0.1679	0.0033
=0.050	-0.0254	0.0029	0.0924	0.0025	0.2284	0.0016
=0.100	0.0381	0.0034	0.1838	0.0029	0.3194	0.0039
g=1.5						
α=0.010	-0.0411	0.0025*	-0.0040	0.0014*	0.0596	0.0014*
=0.025	-0.0307	0.0006*	0.0207	0.0023*	0.0968	0.0020*
=0.050	-0.0129	0.0011*	0.0495	0.0017*	0.1314	0.0010*
=0.100	0.0202	0.0018*	0.1034	0.0016*	0.1941	0.0034*
g=2.0						
α=0.010	-0.0131	0.0006*	-0.0013	0.0007*	0.0229	0.0009*
=0.025	-0.0098	0.0003*	0.0071	0.0007*	0.0377	0.0008*
=0.050	-0.0040	0.0004*	0.0184	0.0006*	0.0545	0.0007*
=0.100	0.0078	0.0004*	0.0410	0.0006*	0.0854	0.0015*
g=2.5						
α=0.010	-0.0025	0.0001**	-0.0003	0.0002**	0.0053	0.0003**
=0.025	-0.0019	0.0001**	0.0014	0.0001**	0.0096	0.0003**
=0.050	-0.0009	0.0001**	0.0043	0.0001**	0.0146	0.0002**
=0.100	0.0017	0.0001**	0.0107	0.0002**	0.0249	0.0005**

*: interpolation grid size 0.01
**:interpolation grid size 0.001

TABLE 2 (cont'd). Estimates of a, h=0.90

	n=5		n=10		n=20	
	â	s.e.	â	s.e.	â	s.e.
g=0.1						
α=0.010	-0.0141	0.0102*	0.0044	0.0038*	0.0237	0.0093*
=0.025	-0.0047	0.0081*	0.0081	0.0035*	0.0180	0.0043*
=0.050	-0.0029	0.0044*	0.0090	0.0042*	0.0320	0.0035*
=0.100	0.0019	0.0034*	0.0176	0.0022*	0.0333	0.0047*
g=0.3						
α=0.010	-0.0067	0.0032*	0.0020	0.0015*	0.0103	0.0039*
=0.025	-0.0046	0.0015*	0.0034	0.0012*	0.0145	0.0012*
=0.050	-0.0018	0.0013*	0.0078	0.0012*	0.0218	0.0012*
=0.100	0.0024	0.0008*	0.0149	0.0008*	0.0290	0.0014*
g=0.5						
α=0.010	-0.0046	0.0015*	0.0002	0.0005*	0.0057	0.0012*
=0.025	-0.0033	0.0004*	0.0014	0.0005*	0.0090	0.0006*
=0.050	-0.0011	0.0006*	0.0049	0.0003*	0.0131	0.0004*
=0.100	0.0017	0.0003*	0.0094	0.0003*	0.0189	0.0006*
g=0.7						
α=0.010	-0.0018	0.0004**	-0.0001	0.0003**	0.0027	0.0003**
=0.025	-0.0016	0.0001**	0.0008	0.0002**	0.0042	0.0002**
=0.050	-0.0005	0.0002**	0.0022	0.0001**	0.0061	0.0002**
=0.100	0.0008	0.0001**	0.0045	0.0001**	0.0092	0.0002**
g=1.0						
α=0.010	-0.0002	0.0000***	0.0000	0.0000***	0.0004	0.0000***
=0.025	-0.0002	0.0000***	0.0001	0.0000***	0.0006	0.0000***
=0.500	-0.0001	0.0000***	0.0003	0.0000***	0.0009	0.0000***
=0.100	0.0001	0.0000***	0.0007	0.0000***	0.0015	0.0000***

*: interpolation grid size 0.01
**: interpolation grid size 0.001
***: interpolation grid size 0.0001

samples or a narrower interpolation grid to achieve a reasonable standard error, this is indicated in the table.

How does $\hat{a}$ depend on n, α, and g? Figure D plots $\hat{a}$ as a function of $-10\log(\alpha)$ for n= 5, 10, 20, for the log-normal distributions with g= 0.5 to 2.5. Since a is a simple weighting of the population mean and median, plots of $\hat{a}$ show where the balanced center falls relative to these two population characteristics.

Looking first at the effect of sample size on $\hat{a}$, note that for sample size n= 20 the balanced center always falls between the mean and median. As n decreases, holding all other factors constant, the balanced center moves, in a relative sense, towards the median, and even past it (a< 0) for n= 5 and some α. Roughly speaking, the balanced center is closer to the median than to the mean for n= 5 and all combinations of g and α considered. For n= 10 it is closer in most cases, for n= 20 in some cases.

Unfortunately, the balanced center depends on α as well as the sample size and population parameters g and h. The value of a decreases as α decreases. That is, the balanced center is closer to the median when we are interested in the extremes of the distribution.

When $0 \leq a \leq 1$, a depends on g in the way we would expect, decreasing as g increases. That is, the balanced center is closer to the median as skewness increases. Once the balanced center lies to the left of the median ($a < 0$), however, there is what might be called a cross-over effect, and a increases (becomes less negative) as g increases. This may be due to the fact that as g increases the left-hand tail of the underlying distribution is being pushed rapidly in toward zero (see Figure A). Values of a cannot be too negative as g increases or the balanced center would lie outside the range of the distribution. If the cross-over effect is in some sense an artifact of the limits of the distribution we might expect some measure of the balanced center to be a monotonic function of n, α, and g, showing no cross-over effect. It can be shown that this is true for the distance from the balanced center to the mean, which always increases as n decreases, as α decreases, and as g increases.

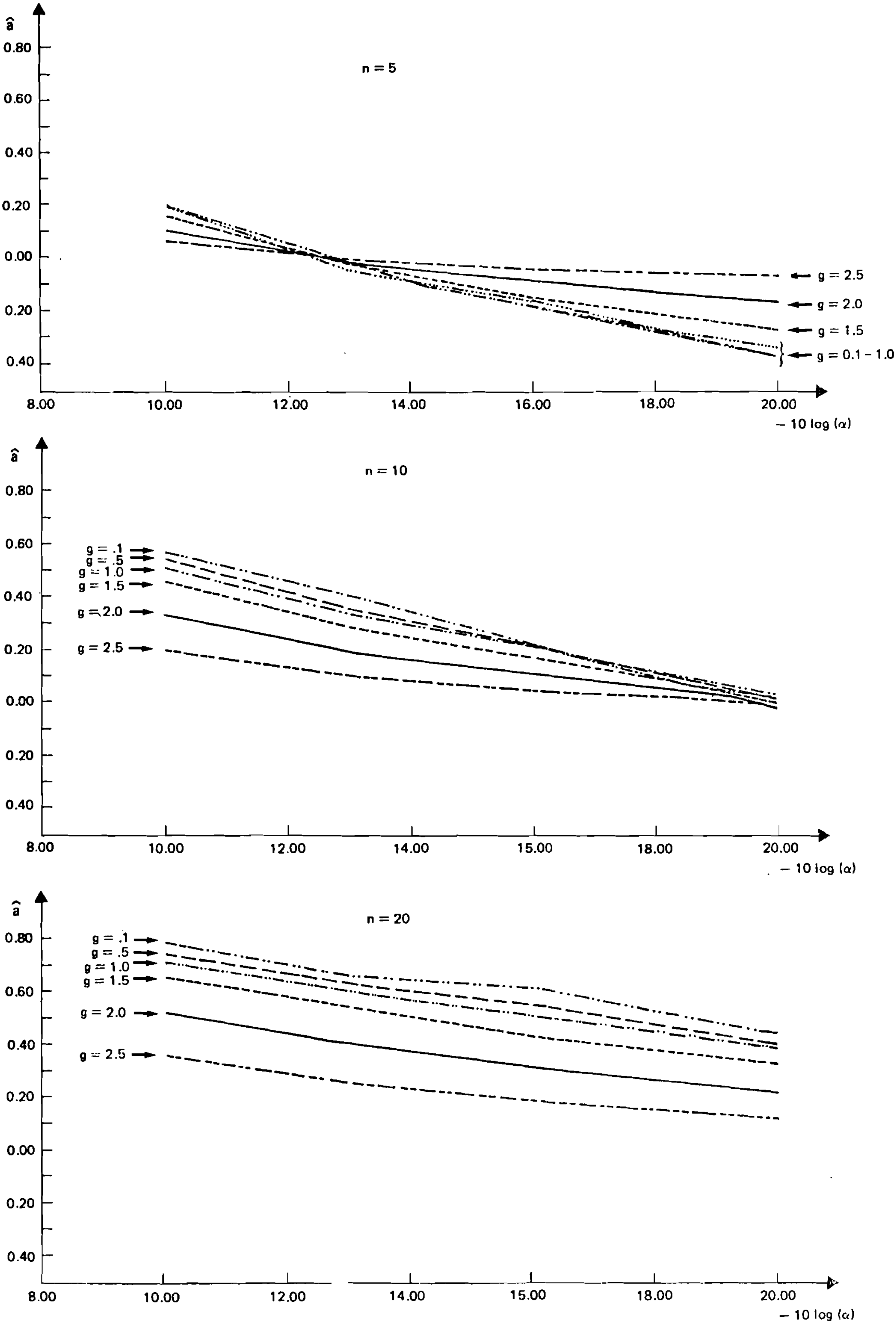

FIGURE D. Plots of $\hat{a}$ against $-10 \log(\alpha)$

The effect of h on the balanced center can be seen most easily from Figures E and F, which show plots of $pct(\hat{c})$ and $\hat{a}$ against h for a few values of n, α, and g. On the one hand, since h stretches the tails of the underlying distribution symmetrically we might expect that increasing h will not disturb the balanced center. On the other hand, since μ moves farther into the tail of the underlying distribution as h increases, this should have some effect. Figure E shows that the balanced center as a percent point of the underlying distribution, $pct(\hat{c})$, is fairly stable over h, particularly for only moderately non-normal distributions. However, the position of the balanced center relative to the mean and median, measured by a, changes dramatically (see Figure F). As h increases, $\hat{a}$ moves rapidly towards zero. We can say that the balanced center lies at about the same percent point for underlying distributions with differing amounts of stretch. However, as stretch increases, the balanced center moves, in a relative sense, closer to the median.

This section provides tables of the balanced center in terms of a for several values of n, α, g and h. If the balanced center is needed for values of g and h not tabled, interpolation will be necessary. It is possible to interpolate in any of the tables presented, but for increased accuracy, Table 3 should be used. This gives values of $(pct(\hat{c})-pct(m))/g$, a quantity which is fairly constant over a reasonable range of g $(0.1 \leq g \leq 1.0)$ and over h for the higher values of α. Using this table, interpolation will give the population characteristic covered in a balanced way by t-intervals for h between 0 and 0.9, g between 0. and 2.5, for three sample sizes and four significance levels. These estimates of a can then be used to make the appropriate modifications of t-intervals about the mean to achieve equal tail probabilities.

4. Tail Probabilities

The previous section estimated the balanced center and indicated how the t-intervals about the mean must be modified to achieve balanced tail probabilities for a class of skew distributions. The actual values of the balanced tail probabilities were not considered. In this section we estimate the balanced tail probabilities for the distributions of the last section, as well as for several of the symmetric h-distributions (for which the balanced center is, of course, the

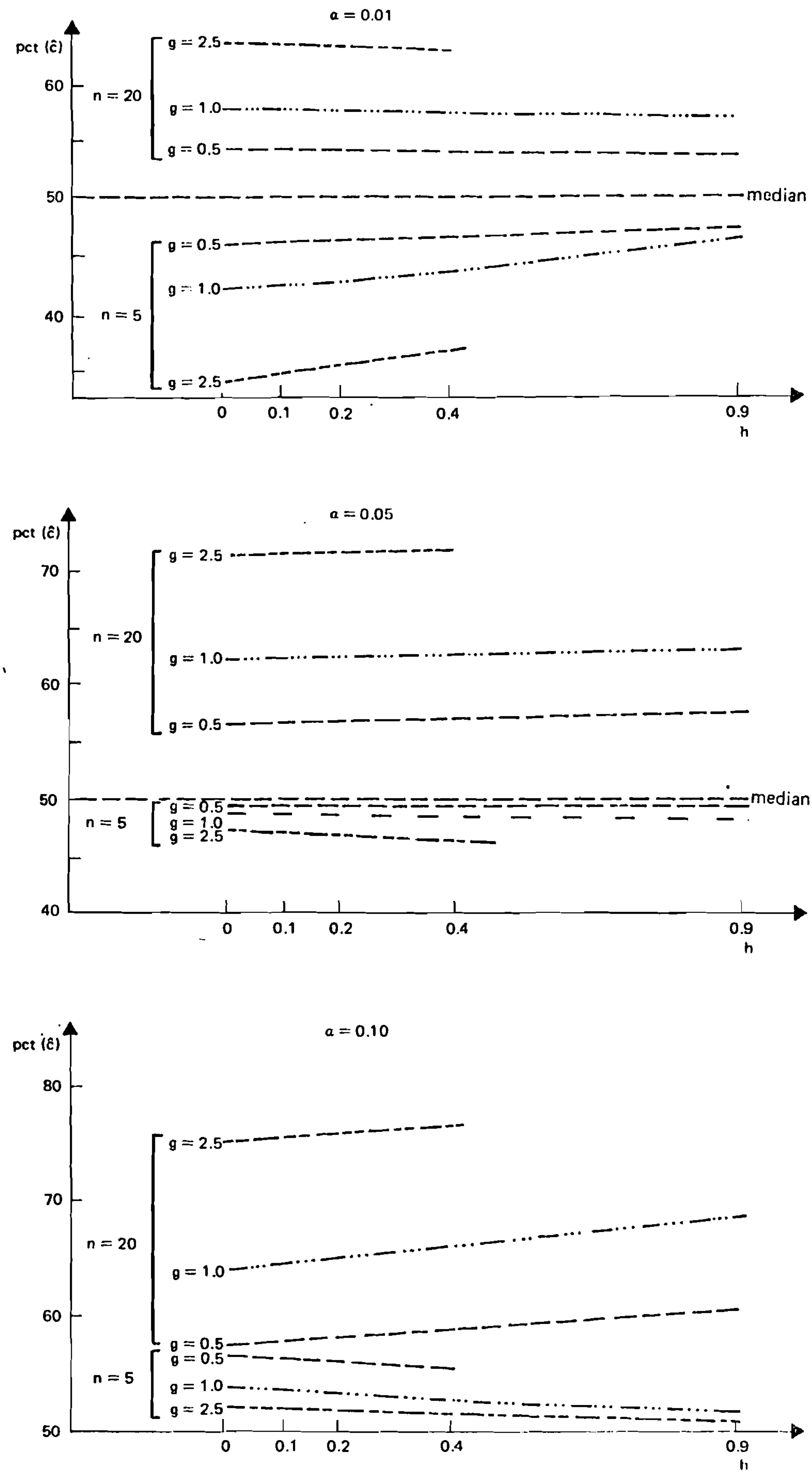

FIGURE E. Plots of pct($\hat{c}$) against h

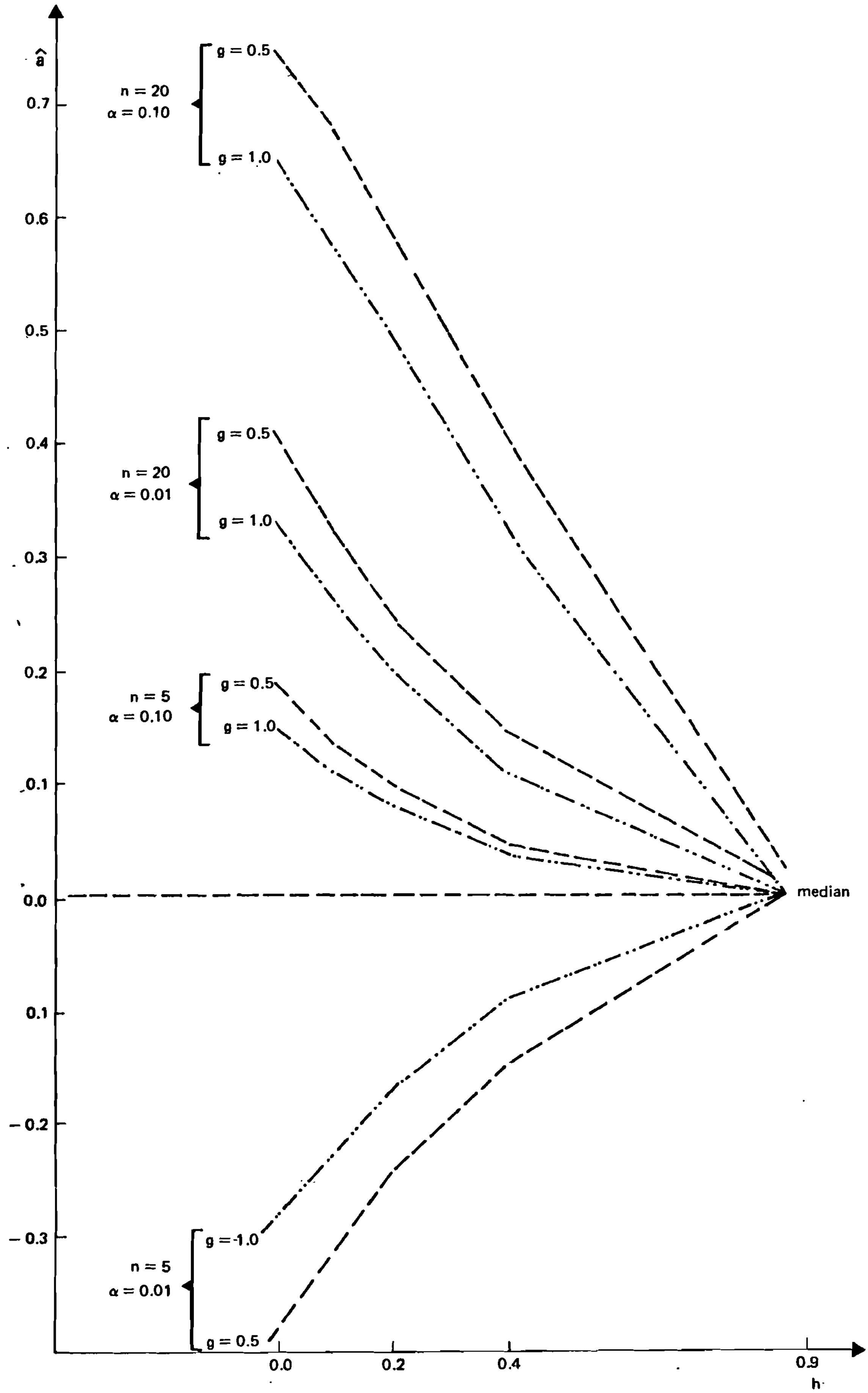

FIGURE F. Plots of â against h

TABLE 3. Interpolation Table of (pct($\hat{c}$)-pct($\hat{m}$))/g, n=5

	h=0.0	h=0.1	h=0.2	h=0.4	h=0.9
α =0.010					
g=0.1	-7.50	-6.90	-3.40	-10.20	-9.10
g=0.3	-8.10	-7.53	-9.23	-7.37	-5.37
g=0.5	-8.22	-7.82	-7.34	-7.42	-5.86
g=0.7	-8.10	-7.97	-7.46	-7.09	-5.04
g=1.0	-7.99	-7.89	-7.32	-6.41	-3.88
g=1.5	-7.75	-7.35	-6.87	-5.49	
g=2.0	-6.84	-6.64	-6.17	-5.05	
g=2.5	-6.34	-6.23	-5.75	-4.46	
α =0.025					
g=0.1	-3.40	-3.90	-6.80	-2.50	-3.00
g=0.3	-4.47	-4.43	-4.63	-4.27	-3.67
g=0.5	-4.38	-4.80	-4.28	-4.18	-4.20
g=0.7	-4.56	-4.49	-4.40	-4.21	-4.47
g=1.0	-4.34	-4.35	-4.23	-4.04	-3.88
g=1.5	-3.95	-4.21	-4.23	-4.05	
g=2.0	-4.18	-4.33	-4.17	-3.69	
g=2.5	-4.04	-4.00	-4.02	-3.26	
α =0.050					
g=0.1	-1.10	-0.70	-1.00	-2.00	-1.90
g=0.3	-0.67	-1.37	-1.33	-1.57	-1.43
g=0.5	-0.84	-1.02	-1.22	-1.64	-1.38
g=0.7	-0.80	-1.07	-1.43	-1.63	-1.37
g=1.0	-0.90	-1.30	-1.62	-1.72	-1.90
g=1.5	-0.92	-1.28	-1.55	-1.67	
g=2.0	-1.09	-1.35	-1.43	-1.44	
g=2.5	-1.12	-1.25	-1.30	-1.44	
α =0.100					
g=0.1	3.60	2.90	2.10	2.70	1.20
g=0.3	4.13	3.53	3.00	2.57	1.90
g=0.5	3.94	3.46	3.12	2.36	2.12
g=0.7	3.97	3.43	3.06	2.50	2.16
g=1.0	3.77	3.39	3.05	2.51	1.81
g=1.5	3.37	3.10	2.83	2.46	
g=2.0	2.96	2.76	2.61	2.53	
g=2.5	2.65	2.50	2.44	2.27	

TABLE 3. Interpolation Table of (pct($\hat{c}$)-pct(m))/g, n=10
(cont'd)

	h=0.0	h=0.1	h=0.2	h=0.4	h=0.9
α=0.010					
g=0.1	-0.40	2.60	1.10	1.40	2.80
g=0.3	0.43	0.37	-0.23	-1.03	1.60
g=0.5	0.32	-0.46	-0.30	-1.00	0.26
g=0.7	0.79	0.09	-0.20	-0.36	-0.27
g=1.0	0.58	-0.22	-0.45	0.11	0.00
g=1.5	0.33	-0.17	-0.40	-0.51	
g=2.0	0.12	-0.11	-0.22	-0.46	
g=2.5	0.03	-0.17	-0.43	-0.46	
α=0.025					
g=0.1	4.50	1.70	4.00	4.20	5.20
g=0.3	4.60	3.97	3.23	2.83	2.70
g=0.5	4.64	3.98	3.44	2.70	1.74
g=0.7	4.63	3.99	3.40	3.06	2.16
g=1.0	4.30	3.78	3.31	2.66	1.81
g=1.5	3.71	3.29	2.99	2.52	
g=2.0	2.98	2.75	2.47	2.32	
g=2.5	2.51	2.29	2.20	1.90	
α=0.050					
g=0.1	8.20	6.70	5.90	6.10	5.80
g=0.3	7.70	7.00	6.80	6.67	6.17
g=0.5	7.52	6.96	6.60	6.44	6.02
g=0.7	7.47	6.93	6.46	6.13	5.76
g=1.0	6.92	6.58	6.30	5.92	5.17
g=1.5	6.07	5.91	5.69	5.69	
g=2.0	5.40	5.25	5.07	5.41	
g=2.5	4.73	4.69	4.56	4.88	
α=0.100					
g=0.1	11.50	11.90	11.80	11.50	11.40
g=0.3	11.43	11.40	11.33	11.07	11.67
g=0.5	11.28	11.40	11.22	11.20	11.26
g=0.7	10.90	10.93	10.96	11.10	11.17
g=1.0	10.39	10.47	10.48	11.13	10.80
g=1.5	9.30	9.34	9.36	10.61	
g=2.0	8.22	8.26	8.32	9.88	
g=2.5	7.36	7.37	7.44	8.84	

TABLE 3. Interpolation Table of (pct($\hat{c}$)-pct(m))/g, n=20 (cont'd)

	h=0.0	h=0.1	h=0.2	h=0.4	h=0.9
α=0.010					
g=0.1	9.00	4.40	7.70	8.10	15.30
g=0.3	8.47	8.73	6.93	6.53	8.10
g=0.5	8.44	8.08	7.42	7.24	6.98
g=0.7	8.44	8.11	7.84	8.04	6.99
g=1.0	7.80	7.46	7.27	6.92	6.71
g=1.5	6.78	6.56	6.43	6.71	
g=2.0	6.04	5.90	5.82	6.47	
g=2.5	5.35	5.28	5.31	5.68	
α=0.025					
g=0.1	12.50	12.60	10.90	12.20	11.60
g=0.3	11.63	11.37	10.93	10.63	11.37
g=0.5	11.52	11.18	10.72	10.48	10.80
g=0.7	10.86	10.56	10.30	10.43	10.50
g=1.0	10.03	10.04	9.85	10.27	9.52
g=1.5	8.97	8.99	8.97	10.07	
g=2.0	7.98	8.05	7.99	9.34	
g=2.5	7.10	7.11	7.09	8.32	
α=0.050					
g=0.1	13.40	13.50	14.00	14.70	20.60
g=0.3	13.40	13.57	13.67	14.23	16.87
g=0.5	13.04	13.38	13.36	13.92	15.32
g=0.7	12.61	12.91	13.09	13.70	14.56
g=1.0	11.95	12.17	12.29	13.44	13.14
g=1.5	10.66	10.83	10.93	12.69	
g=2.0	9.53	9.60	9.69	11.79	
g=2.5	8.51	8.56	8.63	10.34	
α=0.100					
g=0.1	15.80	16.70	16.80	17.90	21.40
g=0.3	15.43	16.27	17.10	18.50	22.13
g=0.5	15.14	16.02	16.80	18.56	21.22
g=0.7	14.70	15.40	16.11	18.11	20.30
g=1.0	13.98	14.58	15.15	17.64	18.70
g=1.5	12.63	13.07	13.47	16.43	
g=2.0	11.34	11.58	11.83	14.86	
g=2.5	10.14	10.32	10.47	12.83	

mean).

Estimation is done as described in Section 3, using Fieller-Hartley and jackknife techniques. In all cases considered in this section, right- and left-hand tail probabilities are equal (for underlying log-normal and g/h-distributions we assume t-statistics have been modified using the balanced center to achieve equal tails), so we look at the average of the two tail probability estimates. For h-distributions, the estimates will be:

$$\hat{\alpha}^* = (\hat{P}(\mu < \bar{x} - \tau_\alpha s_n) + \hat{P}(\mu > \bar{x} + \tau_\alpha s_n))/2,$$
$$= (\hat{P}(t > \tau_\alpha) + \hat{P}(t < -\tau_\alpha))/2.$$

For log-normal and g/h-distributions, the estimates will be:

$$\hat{\alpha}^* = \hat{P}(\mu < (\bar{x} - \tau_\alpha s_n - (1-a)m)/a) + \hat{P}(\mu > (\bar{x} + \tau_\alpha s_n - (1-a)m)/a)/2.$$
$$= (\hat{P}(t_a > \tau_\alpha) + \hat{P}(t_a < -\tau_\alpha))/2, \quad t_a = (\bar{x} - (a\mu + (1-a)m))/s_n.$$

The log-normal, h-, and g/h-distributions all have stretched tails. For such distributions, t is conservative — tails are squeezed-in compared to the Student-t distribution. That is, $\alpha^* < \alpha$, where α is the desired one-sided tail probability. For a review of the many studies demonstrating this phenomenon, see Hatch and Posten (1966).

Tables 4, 5, and 6 present estimates $\hat{\alpha}^*$ together with their estimated standard errors, for several of the symmetric h-distributions, and for several of the skew log-normal and g/h-distributions, respectively. As expected, the t-statistic is conservative in the tails for these stretched-tail distributions, and intervals will be too wide. The difference between α^* and α is negligible in many cases, but increases as stretch increases (i.e. as h and/or g increases).

Although t-intervals are still too wide, modification using the balanced center has made a large improvement in tail probabilities for the skew distributions. The estimated summed tail probability after balancing, $2\hat{\alpha}^*$, is much closer to the desired significance level, 2α, than is the tail probability $\hat{\alpha}_L + \hat{\alpha}_R$ associated with the unmodified Student-t interval. In Table 7, $\hat{\alpha}_L + \hat{\alpha}_R$ and $2\hat{\alpha}^*$ are compared to $2\alpha = 0.05$ for a few of the underlying g/h-distributions, for sample size n= 10. For very small degrees of skewness (g= 0.1), $\hat{\alpha}_L + \hat{\alpha}_R$ is reasonably close to 2α. However, as skewness increases, $\hat{\alpha}_L + \hat{\alpha}_R$ becomes rapidly much larger than 2α. For h= 0.2,

TABLE 4. Tail Probability Estimates for the h-Distributions

	n=5 $\hat{\alpha}^*$	s.e.	n=10 $\hat{\alpha}^*$	s.e.	n=20 $\hat{\alpha}^*$	s.e.
α=0.005						
h=0.05	0.0045	0.0001	0.0041	0.0001	0.0045	0.0001
h=0.10	0.0042	0.0001	0.0039	0.0001	0.0043	0.0002
h=0.15	0.0037	0.0002	0.0034	0.0002	0.0040	0.0002
h=0.20	0.0033	0.0002	0.0029	0.0002	0.0035	0.0002
h=0.40	0.0021	0.0002	0.0018	0.0003	0.0024	0.0003
h=0.60	0.0018	0.0002	0.0012	0.0003	0.0016	0.0003
h=0.80	0.0015	0.0002	0.0006	0.0002	0.0012	0.0003
h=0.90	0.0012	0.0002	0.0006	0.0002	0.0011	0.0003
α=0.010						
h=0.05	0.0092	0.0001	0.0092	0.0002	0.0092	0.0001
h=0.10	0.0084	0.0002	0.0084	0.0002	0.0087	0.0001
h=0.15	0.0074	0.0002	0.0072	0.0002	0.0084	0.0001
h=0.20	0.0068	0.0002	0.0068	0.0002	0.0079	0.0002
h=0.40	0.0049	0.0001	0.0052	0.0003	0.0056	0.0004
h=0.60	0.0038	0.0002	0.0037	0.0003	0.0036	0.0004
h=0.80	0.0031	0.0003	0.0027	0.0003	0.0030	0.0003
h=0.90	0.0029	0.0003	0.0021	0.0002	0.0026	0.0003
α=0.025						
h=0.05	0.0236	0.0002	0.0239	0.0003	0.0248	0.0002
h=0.10	0.0223	0.0003	0.0228	0.0003	0.0235	0.0004
h=0.15	0.0213	0.0003	0.0210	0.0004	0.0230	0.0005
h=0.20	0.0197	0.0004	0.0201	0.0005	0.0216	0.0005
h=0.40	0.0151	0.0004	0.0158	0.0005	0.0177	0.0005
h=0.60	0.0122	0.0002	0.0126	0.0006	0.0143	0.0008
h=0.80	0.0099	0.0003	0.0108	0.0006	0.0115	0.0008
h=0.90	0.0089	0.0002	0.0095	0.0006	0.0103	0.0007
α=0.050						
h=0.05	0.0480	0.0002	0.0497	0.0003	0.0498	0.0004
h=0.10	0.0464	0.0003	0.0487	0.0005	0.0502	0.0006
h=0.15	0.0448	0.0004	0.0477	0.0005	0.0497	0.0005
h=0.20	0.0427	0.0004	0.0464	0.0004	0.0481	0.0006
h=0.40	0.0368	0.0009	0.0403	0.0004	0.0428	0.0007
h=0.60	0.0312	0.0008	0.0348	0.0005	0.0370	0.0010
h=0.80	0.0270	0.0010	0.0295	0.0004	0.0320	0.0010
h=0.90	0.0252	0.0009	0.0279	0.0005	0.0306	0.0011
α=0.100						
h=0.05	0.0995	0.0003	0.1007	0.0002	0.1000	0.0004
h=0.10	0.0990	0.0004	0.1008	0.0005	0.0999	0.0007
h=0.15	0.0973	0.0005	0.1008	0.0007	0.0994	0.0011
h=0.20	0.0953	0.0007	0.1002	0.0010	0.1001	0.0013
h=0.40	0.0884	0.0010	0.0971	0.0012	0.1000	0.0012
h=0.60	0.0822	0.0015	0.0922	0.0014	0.0962	0.0014
h=0.80	0.0757	0.0018	0.0867	0.0010	0.0925	0.0016
h=0.90	0.0730	0.0019	0.0843	0.0011	0.0901	0.0014

Table 5. Balanced Tail Probability Estimates for the Log-normals

	n=5		n=10		n=20	
	$\hat{\alpha}*$	s.e.	$\hat{\alpha}*$	s.e.	$\hat{\alpha}*$	s.e.
α=0.01						
g=0.1	0.0100	0.0001	0.0102	0.0002	0.0094	0.0003
g=0.3	0.0099	0.0003	0.0106	0.0003	0.0099	0.0003
g=0.5	0.0090	0.0004	0.0103	0.0005	0.0095	0.0004
g=0.7	0.0079	0.0003	0.0100	0.0006	0.0099	0.0005
g=1.0	0.0072	0.0004	0.0090	0.0005	0.0090	0.0005
g=1.5	0.0056	0.0004	0.0077	0.0004	0.0080	0.0006
g=2.0	0.0045	0.0004	0.0057	0.0004	0.0069	0.0006
g=2.5	0.0033	0.0004	0.0047	0.0005	0.0058	0.0006
α =0.025						
g=0.1	0.0251	0.0002	0.0251	0.0005	0.0252	0.0002
g=0.3	0.0249	0.0003	0.0252	0.0006	0.0264	0.0007
g=0.5	0.0244	0.0004	0.0249	0.0006	0.0263	0.0008
g=0.7	0.0227	0.0004	0.0249	0.0005	0.0257	0.0009
g=1.0	0.0210	0.0006	0.0237	0.0005	0.0247	0.0012
g=1.5	0.0185	0.0005	0.0220	0.0005	0.0237	0.0013
g=2.0	0.0146	0.0006	0.0187	0.0007	0.0223	0.0011
g=2.5	0.0111	0.0006	0.0170	0.0010	0.0203	0.0011
α =0.50						
g=0.1	0.0498	0.0005	0.0498	0.0004	0.0494	0.0007
g=0.3	0.0490	0.0005	0.0498	0.0004	0.0491	0.0012
g=0.5	0.0485	0.0006	0.0504	0.0006	0.0498	0.0012
g=0.7	0.0479	0.0007	0.0508	0.0006	0.0500	0.0013
g=1.0	0.0463	0.0008	0.0515	0.0007	0.0527	0.0015
g=1.5	0.0431	0.0010	0.0497	0.0008	0.0532	0.0015
g=2.0	0.0395	0.0009	0.0477	0.0010	0.0516	0.0015
g=2.5	0.0365	0.0010	0.0439	0.0010	0.0510	0.0014
α =0.10						
g=0.1	0.1003	0.0002	0.0992	0.0004	0.0999	0.0007
g=0.3	0.1010	0.0005	0.1013	0.0006	0.0998	0.0013
g=0.5	0.1020	0.0009	0.1015	0.0012	0.1019	0.0014
g=0.7	0.1030	0.0013	0.1045	0.0013	0.1045	0.0014
g=1.0	0.1043	0.0019	0.1074	0.0018	0.1083	0.0019
g=1.5	0.1031	0.0018	0.1127	0.0021	0.1161	0.0022
g=2.0	0.1020	0.0022	0.1135	0.0023	0.1204	0.0020
g=2.5	0.1003	0.0020	0.1146	0.0024	0.1228	0.0021

Table 6. Balanced Tail Probability Estimates for the g/h-Distributions, h=0.1

	n=5 $\hat{\alpha}^*$	s.e.	n=10 $\hat{\alpha}^*$	s.e.	n=20 $\hat{\alpha}^*$	s.e.
α=0.01						
h=0.1,						
symmetric	0.0084	0.0002	0.0084	0.0002	0.0087	0.0001
g=0.1	0.0082	0.0002	0.0083	0.0003	0.0087	0.0003
g=0.3	0.0077	0.0002	0.0089	0.0002	0.0085	0.0003
g=0.5	0.0072	0.0002	0.0088	0.0004	0.0089	0.0005
g=0.7	0.0066	0.0003	0.0088	0.0004	0.0085	0.0005
g=1.0	0.0063	0.0004	0.0079	0.0004	0.0085	0.0005
g=1.5	0.0051	0.0004	0.0066	0.0004	0.0074	0.0006
g=2.0	0.0038	0.0004	0.0053	0.0004	0.0065	0.0007
g=2.5	0.0031	0.0004	0.0042	0.0004	0.0055	0.0006
α=0.025						
h=0.1,						
symmetric	0.0223	0.0003	0.0228	0.0003	0.0235	0.0004
g=0.1	0.0225	0.0003	0.0229	0.0004	0.0236	0.0004
g=0.3	0.0218	0.0003	0.0227	0.0006	0.0244	0.0005
g=0.5	0.0210	0.0004	0.0231	0.0005	0.0250	0.0008
g=0.7	0.0202	0.0004	0.0227	0.0005	0.0244	0.0008
g=1.0	0.0186	0.0005	0.0223	0.0006	0.0236	0.0012
g=1.5	0.0171	0.0005	0.0203	0.0006	0.0225	0.0014
g=2.0	0.0132	0.0006	0.0177	0.0008	0.0214	0.0012
g=2.5	0.0106	0.0006	0.0160	0.0009	0.0198	0.0011
α=0.50						
h=0.1,						
symmetric	0.0464	0.0003	0.0487	0.0005	0.0502	0.0006
g=0.1	0.0462	0.0003	0.0481	0.0003	0.0503	0.0008
g=0.3	0.0457	0.0005	0.0477	0.0003	0.0489	0.0012
g=0.5	0.0452	0.0006	0.0485	0.0005	0.0494	0.0011
g=0.7	0.0445	0.0008	0.0488	0.0005	0.0502	0.0012
g=1.0	0.0434	0.0008	0.0492	0.0005	0.0514	0.0014
g=1.5	0.0411	0.0010	0.0483	0.0009	0.0520	0.0016
g=2.0	0.0374	0.0009	0.0462	0.0009	0.0503	0.0014
g=2.5	0.0342	0.0009	0.0429	0.0010	0.0492	0.0016
α=0.10						
h=0.1,						
symmetric	0.0990	0.0004	0.1008	0.0005	0.0999	0.0007
g=0.1	0.0991	0.0002	0.0996	0.0006	0.0995	0.0008
g=0.3	0.0989	0.0008	0.1005	0.0007	0.1013	0.0011
g=0.5	0.1001	0.0011	0.1017	0.0013	0.1035	0.0015
g=0.7	0.1005	0.0014	0.1040	0.0014	0.1051	0.0013
g=1.0	0.1015	0.0018	0.1070	0.0017	0.1086	0.0017
g=1.5	0.1008	0.0016	0.1114	0.0021	0.1166	0.0021
g=2.0	0.0990	0.0021	0.1127	0.0021	0.1198	0.0018
g=2.5	0.0980	0.0019	0.1134	0.0023	0.1221	0.0023

Table 6. (cont'd) Balanced Tail Probability Estimates for the g/h/Distributions, h=0.2

	n=5		n=10		n=20	
	$\hat{\alpha}*$	s.e.	$\hat{\alpha}*$	s.e.	$\hat{\alpha}*$	s.e.
α=0.01						
h=0.2,						
symmetric	0.0068	0.0002	0.0068	0.0002	0.0079	0.0002
g=0.1	0.0067	0.0002	0.0067	0.0003	0.0078	0.0002
g=0.3	0.0063	0.0002	0.0073	0.0004	0.0078	0.0004
g=0.5	0.0062	0.0002	0.0072	0.0004	0.0075	0.0004
g=0.7	0.0057	0.0003	0.0070	0.0004	0.0076	0.0004
g=1.0	0.0052	0.0003	0.0067	0.0004	0.0075	0.0005
g=1.5	0.0046	0.0003	0.0055	0.0003	0.0068	0.0005
g=2.0	0.0035	0.0004	0.0050	0.0004	0.0060	0.0007
g=2.5	0.0028	0.0004	0.0037	0.0004	0.0051	0.0007
α =0.025						
h=0.2,						
symmetric	0.0197	0.0004	0.0201	0.0005	0.0216	0.0005
g=0.1	0.0193	0.0003	0.0202	0.0004	0.0212	0.0005
g=0.3	0.0192	0.0004	0.0206	0.0005	0.0214	0.0007
g=0.5	0.0189	0.0005	0.0207	0.0004	0.0225	0.0007
g=0.7	0.0180	0.0003	0.0204	0.0005	0.0230	0.0008
g=1.0	0.0169	0.0003	0.0199	0.0007	0.0227	0.0011
g=1.5	0.0150	0.0004	0.0185	0.0007	0.0214	0.0013
g=2.0	0.0115	0.0006	0.0167	0.0008	0.0201	0.0012
g=2.5	0.0096	0.0006	0.0152	0.0008	0.0189	0.0010
α =0.50						
h=0.2,						
symmetric	0.0427	0.0004	0.0464	0.0004	0.0481	0.0006
g=0.1	0.0428	0.0003	0.0457	0.0005	0.0488	0.0007
g=0.3	0.0424	0.0006	0.0454	0.0007	0.0477	0.0010
g=0.5	0.0417	0.0008	0.0461	0.0005	0.0485	0.0012
g=0.7	0.0410	0.0008	0.0467	0.0006	0.0492	0.0013
g=1.0	0.0402	0.0008	0.0472	0.0004	0.0500	0.0016
g=1.5	0.0385	0.0009	0.0467	0.0010	0.0508	0.0015
g=2.0	0.0359	0.0008	0.0438	0.0009	0.0495	0.0014
g=2.5	0.0326	0.0008	0.0420	0.0011	0.0485	0.0015
α =0.10						
h=0.2,						
symmetric	0.0953	0.0007	0.1002	0.0010	0.1001	0.0013
g=0.1	0.0963	0.0006	0.0996	0.0010	0.0995	0.0012
g=0.3	0.0969	0.0011	0.0999	0.0010	0.1012	0.0014
g=0.5	0.0967	0.0012	0.1011	0.0014	0.1045	0.0015
g=0.7	0.0981	0.0014	0.1030	0.0016	0.1054	0.0012
g=1.0	0.0991	0.0019	0.1058	0.0017	0.1100	0.0018
g=1.5	0.0984	0.0019	0.1099	0.0018	0.1170	0.0020
g=2.0	0.0973	0.0021	0.1119	0.0022	0.1195	0.0018
g=2.5	0.0962	0.0018	0.1123	0.0024	0.1212	0.0021

Table 6. (cont'd) Balanced Tail Probability Estimates for the g/h-Distributions, h=0.4

	n=5 $\hat{\alpha}*$	s.e.	n=10 $\hat{\alpha}*$	s.e.	n=20 $\hat{\alpha}*$	s.e.
α=0.01						
h=0.4,						
symmetric	0.0049	0.0001	0.0052	0.0003	0.0056	0.0004
g=0.1	0.0050	0.0002	0.0052	0.0003	0.0057	0.0003
g=0.3	0.0049	0.0002	0.0050	0.0003	0.0056	0.0003
g=0.5	0.0046	0.0002	0.0049	0.0003	0.0058	0.0005
g=0.7	0.0042	0.0002	0.0048	0.0003	0.0061	0.0005
g=1.0	0.0041	0.0002	0.0048	0.0004	0.0060	0.0003
g=1.5	0.0038	0.0003	0.0041	0.0004	0.0053	0.0004
g=2.0	0.0029	0.0003	0.0037	0.0004	0.0049	0.0006
g=2.5	0.0025	0.0003	0.0030	0.0002	0.0045	0.0007
α=0.025						
h=0.4,						
symmetric	0.0151	0.0004	0.0158	0.0005	0.0177	0.0005
g=0.1	0.0147	0.0005	0.0160	0.0005	0.0177	0.0007
g=0.3	0.0147	0.0004	0.0161	0.0005	0.0176	0.0007
g=0.5	0.0145	0.0004	0.0160	0.0004	0.0179	0.0009
g=0.7	0.0144	0.0005	0.0163	0.0005	0.0186	0.0010
g=1.0	0.0137	0.0006	0.0165	0.0008	0.0197	0.0009
g=1.5	0.0119	0.0004	0.0159	0.0007	0.0192	0.0011
g=2.0	0.0098	0.0005	0.0146	0.0008	0.0185	0.0010
g=2.5	0.0079	0.0004	0.0134	0.0006	0.0177	0.0011
α=0.50						
h=0.4,						
symmetric	0.0368	0.0009	0.0403	0.0004	0.0428	0.0007
g=0.1	0.0369	0.0008	0.0397	0.0006	0.0431	0.0008
g=0.3	0.0357	0.0009	0.0406	0.0005	0.0437	0.0011
g=0.5	0.0354	0.0009	0.0412	0.0004	0.0435	0.0013
g=0.7	0.0358	0.0009	0.0416	0.0007	0.0456	0.0016
g=1.0	0.0353	0.0009	0.0416	0.0005	0.0468	0.0016
g=1.5	0.0335	0.0009	0.0418	0.0009	0.0471	0.0016
g=2.0	0.0314	0.0009	0.0410	0.0009	0.0476	0.0015
g=2.5	0.0296	0.0009	0.0389	0.0009	0.0467	0.0013
α=0.10						
h=0.4,						
symmetric	0.0884	0.0010	0.0971	0.0012	0.1000	0.0012
g=0.1	0.0893	0.0010	0.0959	0.0012	0.0998	0.0013
g=0.3	0.0904	0.0012	0.0967	0.0014	0.1004	0.0015
g=0.5	0.0900	0.0019	0.0983	0.0016	0.1030	0.0016
g=0.7	0.0906	0.0019	0.0999	0.0014	0.1063	0.0014
g=1.0	0.0923	0.0020	0.1025	0.0016	0.1112	0.0015
g=1.5	0.0933	0.0019	0.1079	0.0022	0.1155	0.0018
g=2.0	0.0938	0.0021	0.1100	0.0022	0.1186	0.0018
g=2.5	0.0925	0.0019	0.1109	0.0022	0.1198	0.0020

Table 6. (cont'd) Balanced Tail Probability Estimates for the g/h-Distributions, h=0.9

	n=5 $\hat{\alpha}*$	s.e.	n=10 $\hat{\alpha}*$	s.e.	n=20 $\hat{\alpha}*$	s.e.
α =0.01						
h=0.9,						
symmetric	0.0029	0.0003	0.0021	0.0002	0.0026	0.0003
g=0.1	0.0028	0.0003	0.0021	0.0003	0.0027	0.0003
g=0.3	0.0029	0.0003	0.0024	0.0004	0.0027	0.0003
g=0.5	0.0025	0.0003	0.0025	0.0004	0.0027	0.0003
g=0.7	0.0024	0.0003	0.0024	0.0004	0.0027	0.0003
g=1.0	0.0027	0.0002	0.0025	0.0004	0.0030	0.0004
α =0.025						
h=0.9,						
symmetric	0.0089	0.0002	0.0095	0.0006	0.0103	0.0007
g=0.1	0.0089	0.0003	0.0094	0.0005	0.0103	0.0007
g=0.3	0.0090	0.0004	0.0089	0.0006	0.0109	0.0008
g=0.5	0.0091	0.0005	0.0088	0.0006	0.0106	0.0007
g=0.7	0.0086	0.0005	0.0087	0.0006	0.0109	0.0007
g=1.0	0.0083	0.0005	0.0099	0.0006	0.0117	0.0006
α =0.50						
h=0.9,						
symmetric	0.0252	0.0009	0.0279	0.0005	0.0306	0.0011
g=0.1	0.0250	0.0009	0.0279	0.0006	0.0298	0.0012
g=0.3	0.0249	0.0009	0.0282	0.0008	0.0304	0.0014
g=0.5	0.0248	0.0010	0.0291	0.0009	0.0313	0.0015
g=0.7	0.0248	0.0009	0.0293	0.0008	0.0329	0.0015
g=1.0	0.0241	0.0009	0.0292	0.0009	0.0348	0.0012
α =0.10						
h=0.9,						
symmetric	0.0730	0.0019	0.0843	0.0011	0.0901	0.0014
g=0.1	0.0733	0.0017	0.0839	0.0012	0.0904	0.0015
g=0.3	0.0739	0.0017	0.0847	0.0020	0.0902	0.0015
g=0.5	0.0745	0.0018	0.0852	0.0020	0.0923	0.0017
g=0.7	0.0756	0.0018	0.0871	0.0022	0.0965	0.0017
g=1.0	0.0766	0.0017	0.0912	0.0016	0.1034	0.0015

Table 7. Tail Probabilities for Modified/Unmodified t - Intervals, n = 10

	α	$\hat{\alpha}_L$	$\hat{\alpha}_R$	$\hat{\alpha}^*$	$\partial\alpha$	$\hat{\alpha}_L+\hat{\alpha}_R$	$\partial\hat{\alpha}^*$
h=0							
g=0.1	0.025	0.032	0.020	0.025	0.050	0.051	0.050
0.5	0.025	0.072	0.006	0.025	0.050	0.078	0.050
1.0	0.025	0.157	0.001	0.024	0.050	0.158	0.047
1.5	0.025	0.282	0.000	0.022	0.050	0.282	0.044
2.0	0.025	0.434	0.000	0.019	0.050	0.434	0.037
2.5	0.025	0.586	0.000	0.017	0.050	0.586	0.034
h=0.2							
g=0.1	0.025	0.030	0.014	0.020	0.050	0.044	0.040
0.5	0.025	0.098	0.002	0.021	0.050	0.100	0.041
1.0	0.025	0.243	0.000	0.020	0.050	0.243	0.040
1.5	0.025	0.416	0.000	0.018	0.050	0.416	0.037
2.0	0.025	0.589	0.000	0.017	0.050	0.589	0.033
2.5	0.025	0.733	0.000	0.015	0.050	0.733	0.030

α : desired one-sided tail probability.

$\hat{\alpha}_L$: estimated left-hand tail probability for unmodified student-t interval.

$\hat{\alpha}_R$: estimated right-hand tail probability for unmodified student-t interval.

$\hat{\alpha}^*$: estimated one-sided tail probability for balanced t-interval (i.e. interval modified using balanced center to achieve equal tails).

g= 0.5, $\hat{\alpha}_L + \hat{\alpha}_R$ is already twice the desired level $-\mu$ falls outside the Student-t interval twice as often as desired. The t-interval modified by using the balanced center is, on the other hand, conservative, and achieves a significance level reasonably close to the desired 0.05 level. For the worst case tabled, h= 0.2 and g= 2.5, $\hat{\alpha}_L + \hat{\alpha}_R = 0.73$, while $2\hat{\alpha}^*$ is 0.03. Most of the error in tail probabilities has been corrected simply by balancing the t-statistic.

For t-statistics based on underlying h-distributions, and for modified t-statistics based on underlying log-normal or g/h-distributions, Tables 4, 5, and 6 can be used to estimate the actual confidence level achieved for various nominal significance levels. They can also be used to determine approximately the critical values needed to achieve a given desired significance level. Suppose we want to find critical values $\pm\tau(\alpha_0)$ such that

$$P(-\tau(\alpha_0) \leq t \leq \tau(\alpha_0)) = 1 - 2\alpha_0^*,$$

where α_0^* is the desired one-sided significance level. We enter Table 4, 5, or 6 to find the α_0 associated with tail probability α_0^* (interpolation will be necessary). The appropriate critical values are the Student-t distribution critical values associated with one-sided probability α_0. For example, for n= 5, and the symmetric distribution with h= 0.9, if we want one-sided significance 0.025, we find from Table 4 that we must use critical values associated with one-sided probability 0.05, from the Student-t distribution with four degrees of freedom.

The accuracy of critical value modification can be greatly increased if we know how α^* and α are related functionally, i.e. if we know the function f, where $\alpha^* = f(\alpha)$. For the h-distributions, for small values of α ($\alpha \leq 0.1$), it is possible to approximate f by a linear function when α^* and α are transformed into their corresponding gaussian critical values. For $\alpha > 0.1$, the relationship is no longer linear. For the purposes of this study, let

$\Phi^{-1}(\alpha)$ = upper gaussian critical value corresponding to one-sided probability α.

Figure G shows $\Phi^{-1}(\hat{\alpha}^*)$ plotted against $\Phi^{-1}(\alpha)$ for h= 0.2, 0.4, 0.9, n= 5, 10, 20, and $\alpha \leq 0.1$ ($\Phi^{-1}(\alpha) \geq 1.282$). For $\alpha \leq 0.1$ the relationship is reasonably linear and appears to be unaffected by sample size. Suppose we assume

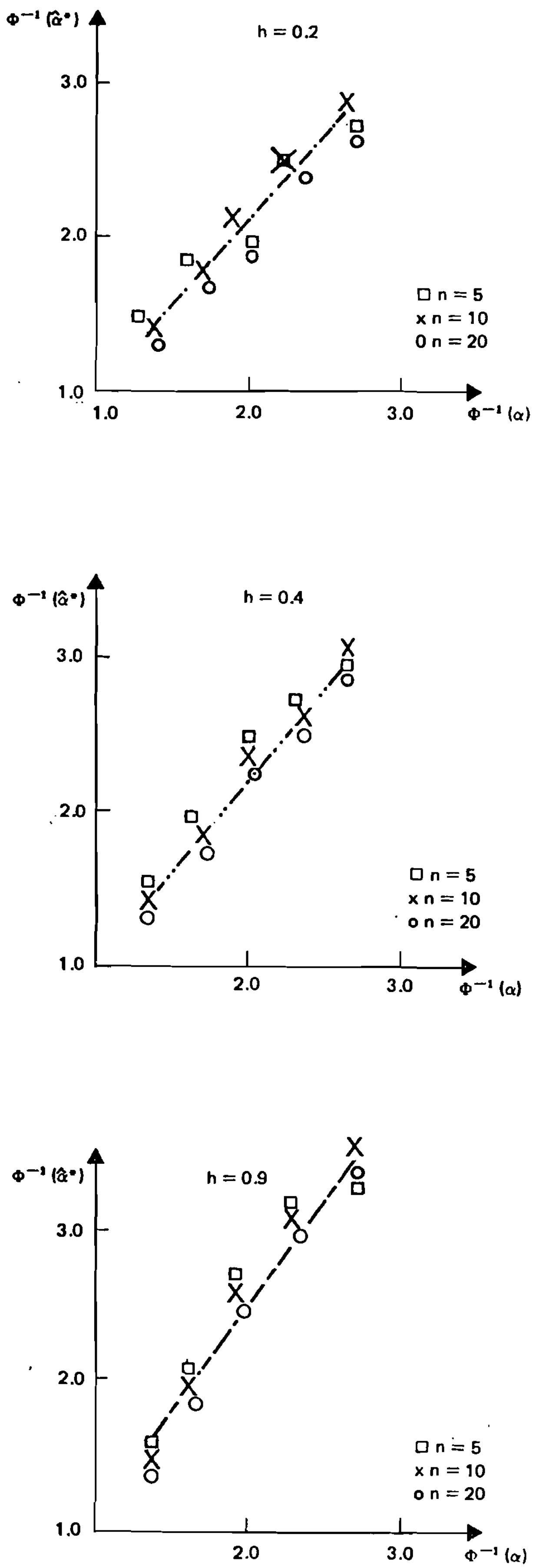

FIGURE G. Plots of $\Phi^{-1}(\hat{\alpha}*)$ against $\Phi^{-1}(\alpha)$, $\alpha \geq 0.1$

$$\Phi^{-1}(\alpha^*) = b_0(h) + b_1(h)\,\Phi^{-1}(\alpha)\,.$$

Now α is recoverable from

$$\Phi^{-1}(\alpha) = (\Phi^{-1}(\alpha^*) - b_0(h))\,/\,b_1(h) \tag{4.1}$$

as the one-sided probability associated with gaussian critical point $\Phi^{-1}(\alpha)$. To form t-intervals at significance level $2\alpha^*$, we use the Student-t critical values associated with α.

Table 8 presents least-square estimates $\hat{b}_0$ and $\hat{b}_1$ for each h. These equations fit very well for $h \leq 0.4$. Note that sample size, n, has been assumed to have no effect, a reasonable assumption for $h \leq 0.4$. Use of these equations allows us to determine the appropriate critical values to use in modifying t-intervals, much more accurately than by interpolation in Table 4. For values of h other than those tabled, interpolation should be done in Table 9, where b_0 and b_1 are treated as functions of h.

The simple relation (4.1) cannot be extended to the log-normal and g/h distributions. For these skew distributions the relationship between $\Phi^{-1}(\alpha)$ and $\Phi^{-1}(\alpha^*)$ is no longer linear, and is not independent of sample size. However, in some cases the tail probabilities for a given g/h distribution are quite close to those for the corresponding h-distribution. In these cases, the appropriate h-distribution modification can be used with reasonable success. This tends to be particularly true at high values of h, where the amount of stretch induced by g is negligible compared to that induced by h. The estimated tail probabilities for distributions with h= 0.9, g= 0.1 to 1.0, for example, are very close to those for the symmetric h-distribution, h= 0.9.

5. Conclusions

This study uses Monte-Carlo methods to estimate, for the g/h class of skew, stretched-tail distributions, the population characteristic covered by Student-t intervals with symmetric loss (equal tail probabilities). Results are presented in a series of tables. The "balanced center" is found to depend on sample size and desired significance level, as well as on the degree of skewness and stretch of the underlying distribution.

Estimates of the balanced center can be used to modify Student-t confidence intervals

Table 8. Regression Coefficients for $\Phi^{-1}(\alpha^*) = b_0 + b_1\, \Phi^{-1}(\alpha)$

h =	$\Phi^{-1}(\alpha^*)$ =
0.05	$1.036\, \Phi^{-1}(\alpha) - 0.050$
0.10	$1.054\, \Phi^{-1}(\alpha) - 0.070$
0.15	$1.083\, \Phi^{-1}(\alpha) - 0.104$
0.20	$1.110\, \Phi^{-1}(\alpha) - 0.135$
0.40	$1.198\, \Phi^{-1}(\alpha) - 0.221$
0.60	$1.257\, \Phi^{-1}(\alpha) - 0.256$
0.80	$1.310\, \Phi^{-1}(\alpha) - 0.285$
0.90	$1.328\, \Phi^{-1}(\alpha) - 0.287$

Table 9. $\Phi^{-1}(\alpha^*)$ as a Function of h and $\Phi^{-1}(\alpha)$

h =	$\Phi^{-1}(\alpha^*)$ =
0.05	$\Phi^{-1}(\alpha) + 0.55h\, (\Phi^{-1}(\alpha) + 0.4h - 0.009/h - 1.25)$
0.10	$\Phi^{-1}(\alpha) + 0.55h\, (\Phi^{-1}(\alpha) + 0.4h - 0.009/h - 1.25)$
0.15	$\Phi^{-1}(\alpha) + 0.55h\, (\Phi^{-1}(\alpha) + 0.4h - 0.009/h - 1.25)$
0.20	$\Phi^{-1}(\alpha) + 0.55h\, (\Phi^{-1}(\alpha) + 0.4h - 0.009/h - 1.25)$
0.40	$\Phi^{-1}(\alpha) + 0.50h\, (\Phi^{-1}(\alpha) + 0.4h - 0.009/h - 1.25)$
0.60	$\Phi^{-1}(\alpha) + 0.43h\, (\Phi^{-1}(\alpha) + 0.4h - 0.009/h - 1.23)$
0.80	$\Phi^{-1}(\alpha) + 0.39h\, (\Phi^{-1}(\alpha) + 0.4h - 0.009/h - 1.23)$
0.90	$\Phi^{-1}(\alpha) + 0.36h\, (\Phi^{-1}(\alpha) + 0.4h - 0.009/h - 1.23)$

about the mean to achieve symmetric loss. The resulting equal tail probabilities are estimated and tabled. Tail probabilities are reasonably close to desired significance levels in many cases. There is a large improvement over the tail probabilities associated with unmodified Student-t intervals.

Further modifications of critical values to give confidence intervals at the desired significance level are given in a general form for the h class of symmetric, stretched-tail distributions. These results can be used to determine approximate further modifications for some of the skew distributions.

The results of this study are applicable to a family of underlying distributions that are more skew and stretched-tail than previously considered in robustness studies of the t-statistic. Results are given for small to moderate sample sizes.

Throughout the study it has been assumed that the parameters g and h of the underlying distribution are known. In practice, of course, g and h may have to be estimated from a sample. For a discussion of sample estimates see Tukey (1976).

References

Arthur, S.P. (1979), "Skew/Stretched Distributions and the t-Statistic," unpublished Ph.D. thesis, Department of Statistics, Princeton University.

Fieller, E.C., and Hartley, H.O. (1954), "Sampling with Control Variables," *Biometrika*, 41, 494-501.

Hatch, L.O., and Posten, H.O. (1966), "Robustness of the Student-Procedure: A Survey," Research Report No. 24, Department of Statistics, University of Connecticut.

Johnson, N.L., Nixon, E., Amos, D.E., and Pearson, E.S. (1963), "Table of Percentage Points of Pearson Curves, for Given $\beta_1^{1/2}$ and β_2, Expressed in Standard Measure," *Biometrika*, 50, 459-498.

Pearson, E.S., and Adyanthaya, N.K. (1929), "The Distribution of Frequency Constants in Small Samples from Non-normal Symmetrical and Skew Populations, Part II," *Biometrika*, 21, 435-438.

Pearson, E.S., and Please, N.W. (1975), "Relation Between the Shape of Population Distribution and the Robustness of Four Simple Test Statistics," *Biometrika*, 62, 223-241.

Tukey, J.W. (1976), unpublished notes for statistics course, Princeton University.