33rd European Study Group with Industry
14 – 18 September 1998
pp. 1 – 10

To t or not to t?

Pieta C. Boon
Faculty of Mathematical Sciences, University of Twente
P.O. Box 217, 7500 AE Enschede, The Netherlands
e-mail: p.c.boon@math.utwente.nl

Harrie Hendriks
Department of Mathematics, University of Nijmegen
P.O. Box 9010, 6500 GL Nijmegen, The Netherlands
e-mail: hhendr@sci.kun.nl

Chris A.J. Klaassen [1]
Korteweg - de Vries Institute for Mathematics
Institute for Business and Industrial Statistics (IBIS UvA BV), University of Amsterdam
Plantage Muidergracht 24, 1018 TV Amsterdam, The Netherlands
e-mail: chrisk@wins.uva.nl

Rolf Muijlwijk
NMi Van Swinden Laboratorium BV
P.O. Box 654, 2600 AR Delft, The Netherlands
e-mail: rmuijlwijk@nmi.nl

Consider an unknown distribution with a symmetric unimodal density and the induced location-scale family. We study confidence intervals for the location parameter based on Student's t-statistic, and we conjecture that the uniform distribution is least favorable in that it leads to confidence intervals that are largest given their coverage probability. This conjecture is supported by an argument based on second order asymptotics in the sample size and on asymptotics in the length of the confidence interval, and by simulation results. Consequently, our answer to the question in the title of this paper is: "Yes, provided the t-statistic confidence interval is based on the assumption of uniformity of the random variables generating the observations."

1. Student's t-statistic

Consider a location-scale family of distributions with the mean as location and the standard deviation as scale parameter. We are interested in confidence intervals for the location parameter based on observations that will be viewed as

[1] Corresponding author

realisations of independent and identically distributed (i.i.d.) random variables from this location-scale family. We will adopt the following notation.

Let $X_1, \dots, X_n$ be i.i.d. random variables with distribution function

$$P(X_i \le x) = F_{\mu,\sigma}(x) = F(\frac{x-\mu}{\sigma}), \ x \in \mathbb{R}, \ i = 1, \dots, n. \tag{1.1}$$

Here μ is the unknown location parameter and σ denotes the unknown scale parameter. The distribution function $F = F_{0,1}$ is standardized such that

$$E_F X_1 = 0, \quad \mathrm{var}_F X_1 = E_F X_1^2 = 1, \tag{1.2}$$

i.e. the standard distribution function F has mean 0 and variance 1. Consequently, we have

$$E_{F_{\mu,\sigma}} X_1 = \mu, \ \mathrm{var}_{F_{\mu,\sigma}} X_1 = \sigma^2. \tag{1.3}$$

The realizations $x_1, \dots, x_n$ of $X_1, \dots, X_n$ are called observations. A confidence interval for μ is an interval $I(x_1, \dots, x_n)$ based on the observations, to which the following statement is attached: the value of the location parameter is contained in the interval, i.e.

$$\mu \in I(x_1, \dots, x_n).$$

This statement may be true or false. The stochastic version of this statement has confidence level $1 - \alpha$, if for all $\mu \in \mathbb{R}, \ \sigma > 0$, we have

$$P_{F_{\mu,\sigma}}(\mu \in I(X_1, \dots, X_n)) \ge 1 - \alpha. \tag{1.4}$$

One also phrases (1.4) as "the coverage probability equals at least $1 - \alpha$."

We will denote the sample mean $n^{-1}\Sigma_{i=1}^n X_i$ by $\bar{X}_n$ and the sample variance $(n-1)^{-1}\Sigma_{i=1}^n (X_i - \bar{X}_n)^2$ by S_n^2 with $S_n \ge 0$. Note that the statistic

$$T_n(\mu) = \frac{\sqrt{n}(\bar{X}_n - \mu)}{S_n} \tag{1.5}$$

is location and scale invariant, i.e. its distribution under $F_{\mu,\sigma}$ is the same as the distribution of $T_n = T_n(0) = \sqrt{n}\bar{X}_n / S_n$ under F. It is called Student's t-statistic. If F equals the standard normal distribution function Φ, then under $\Phi_{\mu,\sigma}(\cdot) = \Phi((\cdot - \mu)/\sigma)$ the t-statistic $T_n(\mu)$ has a Student t-distribution with $n-1$ degrees of freedom. This distribution is called after Student, pseudonym of William Sealy Gosset, who was chemist and Brewer-in-Charge of the Experimental Brewery of Guinness' Brewery. Gosset determined the density of the t-statistic from (1.5) under normality as a function of $t \in \mathbb{R}$ as

$$\frac{\Gamma(n/2)}{\sqrt{\pi(n-1)}\,\Gamma((n-1)/2)} \left(1 + \frac{t^2}{n-1}\right)^{-n/2} \tag{1.6}$$

in Section III of his paper Student (1908).

Let $t_{n-1}(p)$ denote the p-th quantile of Student's t-distribution with $n-1$ degrees of freedom, i.e.

$$P_\Phi(T_n \le t_{n-1}(p)) = p,\ 0 < p < 1. \tag{1.7}$$

Via the equivalence

$$T_n \le t \Leftrightarrow \mu \ge \bar{X}_n - t\tfrac{S_n}{\sqrt{n}}$$

we obtain

$$I_\alpha = \left[\bar{X}_n - t_{n-1}(1-\tfrac{1}{2}\alpha)\tfrac{S_n}{\sqrt{n}},\ \bar{X}_n - t_{n-1}(\tfrac{1}{2}\alpha)\tfrac{S_n}{\sqrt{n}}\right]. \tag{1.8}$$

as a confidence interval, based on the t-statistic, for the location parameter μ with coverage probability $1-\alpha$, provided the underlying distribution of the random variables generating the observations is normal. Since T_n inherits the symmetry of the underlying distribution Φ we have $t_{n-1}(\frac{1}{2}\alpha) = t_{n-1}$ $(1-\frac{1}{2}\alpha)$ resulting in the interval

$$I_t(X_1,\dots,X_n) = [\bar{X}_n - t\tfrac{S_n}{\sqrt{n}},\ \bar{X}_n + t\tfrac{S_n}{\sqrt{n}}] \tag{1.9}$$

with $t = t_{n-1}(1-\frac{1}{2}\alpha)$. Note that this interval is symmetric around $\bar{X}_n$.

In daily practice, this confidence interval is used extremely frequently. It is the classical, standard method for constructing confidence intervals for location parameters. The NMi Van Swinden Laboratorium (Nederlands Meetinstituut) also uses it routinely, since it is agreed upon as an international standard. However, quite often the observations may not be viewed as stemming from normal random variables and consequently, one is not sure then about the true coverage probability

$$P_{F_{\mu,\sigma}}(\mu \in I_t(X_1,\dots,X_n)) = P_F(0 \in I_t(X_1,\dots,X_n)) \tag{1.10}$$

and one is not even sure if $1-\alpha$ is still a valid confidence level, cf. (1.4). This possible non-normality of the underlying random variables has been noticed also in the practice of the NMi Van Swinden Laboratorium. To the 33rd European Study Group with Industry, Leiden, September 14–18, 1998, a naturally related question was suggested: "How should t be chosen such that the interval $I_t(X_1,\dots,X_n)$ has confidence level $1-\alpha$ under F". This question is paraphrased by the title of this paper after Shakespeare (1601). We will discuss this question under the simplifying assumption that F be symmetric, i.e. $F(-x) = 1-F(x)$, and under the complicating assumption that the sample size n be small. In practice, sample sizes as small as 3 are not rare. Nevertheless, we will study asymptotics, as $n \to \infty$, first.

2. Asymptotics

By the Central Limit Theorem the standardized sample mean converges in distribution to a standard normal random variable, i.e.

$$\sqrt{n}(\bar{X}_n - \mu)/\sigma \xrightarrow{\mathcal{D}} \mathcal{N}(0,1) \tag{2.1}$$

as $n \to \infty$. By the Law of Large Numbers the sample variance converges in probability to σ^2, namely

$$S_n^2 = \tfrac{1}{n-1}\sum_{i=1}^{n}(X_i-\mu)^2 - \tfrac{n}{n-1}(\bar{X}_n-\mu)^2 \xrightarrow{P} \sigma^2. \tag{2.2}$$

Combining (2.1) and (2.2) we arrive at

$$T_n \xrightarrow{\mathcal{D}} \mathcal{N}(0,1) \tag{2.3}$$

as $n \to \infty$, whatever the underlying distribution function F (with finite variance). In particular, by the choice $F = \Phi$ this implies that the t-distribution converges to the standard normal distribution as the degrees of freedom converge to infinity.

The asymptotic normality (2.3) means that asymptotically the coverage probability of $I_t(X_1,\ldots,X_n)$ equals $\Phi(t)-\Phi(-t)$, or

$$\lim_{n\to\infty} P_{F_{\mu,\sigma}}\left(\mu \in I_t(X_1,\ldots,X_n)\right) = 2\Phi(t)-1,\ \mu\in\mathbb{R},\ \sigma>0,\ t>0, \tag{2.4}$$

irrespectively of the underlying distribution function F with finite variance. Although this result implies that a confidence interval with approximate confidence level $1-\alpha$ may be constructed by choosing $t = \Phi^{-1}(1-\alpha/2)$ in $I_t(X_1,\ldots,X_n)$, the confidence level might be quite misleading for finite sample sizes n. A better approximate confidence interval might be obtained by application of Edgeworth expansions for the distribution function of T_n. In fact, this yields the following approximation, with φ denoting the standard normal density.

Theorem 2.1 *Let F be symmetric and non-singular. If F has finite fourth moment, then $\kappa(F) = E_F X^4 (E_F X^2)^{-2} - 3$, the deviation from normal kurtosis, is well defined and*

$$\begin{aligned} &P_{F_{\mu,\sigma}}(\mu\in I_t(X_1,\ldots,X_n)) = 2\Phi(t)-1 \\ +\ &\tfrac{t}{6n}\{\kappa(F)(t^2-3)-3t^2-3\}\varphi(t)+o(\tfrac{1}{n}) \end{aligned} \tag{2.5}$$

holds uniformly in t. In other words, if the t-statistic confidence interval $I_t(X_1,\ldots,X_n)$ has coverage probability

$$P_{F_{\mu,\sigma}}(\mu\in I_t(X_1,\ldots,X_n)) = 1-\alpha,\ \mu\in\mathbb{R},\ \sigma>0, \tag{2.6}$$

under F, then

$$\begin{aligned} t = \Phi^{-1}(1-\tfrac{\alpha}{2})\{1+\tfrac{1}{12n}[&-\kappa(F)(\{\Phi^{-1}(1-\tfrac{\alpha}{2})\}^2-3) \\ &+3\{\Phi^{-1}(1-\tfrac{\alpha}{2})\}^2+3]\}+o(\tfrac{1}{n}). \end{aligned} \tag{2.7}$$

Proof Straightforward application of the Theorem of Hall (1987) yields (2.5). Note that Hall's T_0 equals $\sqrt{1+1/(n-1)}\,T_n$ and hence his y has to be replaced by $t+(2n)^{-1}t$. Since this expansion (2.5) is uniform in t the second statement is implied by it. □

Let us assume

$$\alpha < 2(1 - \Phi(\sqrt{3})) = 0.0832. \tag{2.8}$$

Then, the right-hand side of (2.7) is decreasing in $\kappa(F)$. Consequently, the confidence interval $I_t(X_1, \ldots, X_n)$ has coverage probability at least $1 - \alpha$ for all F if t satisfies (2.7) with $\kappa(F)$ minimal. Under all distributions F, the deviation from normal kurtosis $\kappa(F)$ equals at least -2, i.e.

$$\kappa(F) = E_F X^4 (E_F X^2)^{-2} - 3 \geq -2, \tag{2.9}$$

in view of $(E_F X^2)^2 \leq E_F X^4$. Equality is attained in (2.9) if X^2 is degenerate. However, in many applications this is not a natural distribution. Moreover, if n is small, $n < 1 - \log\alpha / \log 2$, then a bounded confidence interval for μ based on Student's t-statistic does not exist for this distribution. In fact, if $P(X = 1) = P(X = -1) = 1/2$ holds, then $P(T_n = \infty) = P(X_1 = \cdots = X_n) = 2^{1-n} > \alpha$.

In the next Section we will restrict attention to distributions F with unimodal density and we will determine the (distribution with) minimal value of $\kappa(F)$ within this class of unimodal distributions.

3. As $n \to \infty$, the uniform distribution is least favorable

Our discussion will be based on the following inequality.

Theorem 3.1 *Let F be a distribution with a symmetric unimodal density. If F has finite fourth moment, then the kurtosis of F equals at least $9/5$, which implies*

$$\kappa(F) \geq -6/5 \tag{3.1}$$

with equality iff F is uniform.

Proof Let f be the symmetric unimodal density of F. Then f may be written as a mixture of symmetric uniform densities. This means that there exists a probability distribution G on $(0, \infty)$ such that

$$f(x) = \int_0^\infty \tfrac{1}{2y} \mathbf{1}_{(-y,y)}(x) dG(y) \tag{3.2}$$

holds. By this representation of f we obtain for $k = 1$ or 2

$$\begin{aligned} E_F X^{2k} &= 2 \int_0^\infty x^{2k} \int_0^\infty \tfrac{1}{2y} \mathbf{1}_{(-y,y)}(x) dG(y) dx \\ &= \int_0^\infty \int_0^y \tfrac{1}{y} x^{2k} dx dG(y) = \tfrac{1}{2k+1} E_G Y^{2k}, \end{aligned} \tag{3.3}$$

and hence

$$E_F X^4 (E_F X^2)^{-2} = \tfrac{9}{5} E_G Y^4 (E_G Y^2)^{-2} \geq \tfrac{9}{5} \tag{3.4}$$

by Cauchy-Schwarz as in (2.9). Note that equality holds iff Y^2 is degenerate, i.e. iff X is uniformly distributed. □

Together with (2.7) inequality (3.1) shows that asymptotically to second order, the uniform distribution is least favorable in the class of symmetric unimodal distributions for constructing confidence intervals based on Student's t-statistic. In other words, at fixed n and $\alpha < 0.0832$ such a confidence interval is largest if the underlying distribution is uniform.

At first sight, this might seem surprising since e.g. the normal distribution itself has heavier tails than the uniform distribution. However, (2.7) seems to imply that in some sense, heavier tails in the underlying distribution of the observations result in less heavy tails for the distribution of the t-statistic and vice versa. In fact, (2.7) itself is an asymptotic version of this conjecture. For finite sample size, the situation that interests us most, this conjectured phenomenon has been formulated precisely and proved for $t \to \infty$ by Van Zwet (1964a,b). We will prove in the next section that the uniform distribution is least favorable under the unimodal distributions in that sense too.

4. As $t \to \infty$, the uniform distribution is least favorable

In this section the relative tailbehavior of $P_F(0 \notin I_t(X_1, \ldots, X_n))$ with respect to $P_H(0 \notin I_t(X_1, \ldots, X_n))$ as $t \to \infty$, will be discussed for H being the uniform distribution function.

THEOREM 4.1 *Let F and H be unimodal distributions symmetric about* 0. *If H is uniform, then*

$$\lim_{t\to\infty} \frac{P_F(0 \notin I_t(X_1, \ldots, X_n))}{P_H(0 \notin I_t(X_1, \ldots, X_n))} \leq 1 \tag{4.1}$$

holds.

PROOF By symmetry, the left-hand side of (4.1) equals

$$\lim_{t\to\infty} \frac{P_F(T_n \geq t)}{P_H(T_n \geq t)} = \frac{R_n(F)}{R_n(H)} = 2^n n \int_0^\infty x^{n-1} f^n(x)dx \tag{4.2}$$

with

$$R_n(F) = \lim_{t\to\infty} \frac{P_F(T_n \geq t)}{P_\Phi(T_n \geq t)} \tag{4.3}$$

introduced and studied by Hotelling (1961); cf. his (3.2) and (3.6). According to Theorem 6.2.1 of Van Zwet (1964a) $R_n(H) \geq R_n(F)$ holds for F and H symmetric about 0, if $F^{-1}(H(x))$ is convex in $x \in (0, \infty)$. This proves the theorem, since for H uniform, this convexity of $F^{-1}(H)$ is equivalent to convexity of $F^{-1}(u)$ in $u \in (\frac{1}{2}, 1)$ which in turn is equivalent to unimodality of f.

An alternative proof may be based on (4.2) and the representation (3.2). Indeed, we have

$$\begin{aligned} \int_0^\infty x^{n-1} f^n(x)dx &= \int_0^\infty x^{n-1} \left\{ \int_0^\infty \tfrac{1}{2y} \mathbf{1}_{(0,y)}(x) dG(y) \right\}^n dx \\ \leq \int_0^\infty \int_0^\infty & x^{n-1}(2y)^{-n} \mathbf{1}_{(0,y)}(x) dx dG(y) = 2^{-n} n^{-1}. \end{aligned} \tag{4.4}$$

□

Consequently, studying the limit behavior of the coverage probability of the Student t-interval $I_t(X_1, \ldots, X_n)$ as $t \to \infty$ with n fixed, we have seen that within the class of symmetric unimodal distributions the uniform is least favorable since its confidence is converging to 1 at the smallest possible speed. The asymptotics both as $n \to \infty$ for t fixed and as $t \to \infty$ for n fixed, support the claim that the uniform distribution is least favorable. The next section adds some numerical evidence to this claim.

5. Simulation results

Fix the sample size n and the underlying distribution function F of the observations $X_1, \ldots, X_n$. Simulate $X_1, \ldots, X_n$ and compute the t-statistic. This results in the simulation Y of T_n under F. Choose N big and simulate in the same way the i.i.d. replicates $Y_1, \ldots, Y_N$ of T_n under F. We denote the order statistics by $Y_{(i)}$ with $Y_{(1)} \leq \cdots \leq Y_{(N)}$. The empirical quantile $Y_{(\lceil pN \rceil)}$ is an appropriate estimator of the p-quantile of the distribution of T_n under F.

The accuracy of this estimator may be estimated too, as described e.g. in the contribution of J.E. Walsh in Sarhan and Greenberg (1962). Let $U_{(1)} \leq \cdots \leq U_{(N)}$ be the order statistics of N i.i.d. uniform $(0, 1)$ random variables. Fix $0 < \beta < 1$ and determine the maximum value of r such that $P(U_{(r)} > p) \leq \frac{1}{2}\beta$ and the minimum value of s such that $P(U_{(s)} < p) \leq \frac{1}{2}\beta$, using the beta-distribution of the order statistics of a uniform sample. Since $F(Y_{(j)})$ and $U_{(j)}$ have the same distribution, $[Y_{(r)}, Y_{(s)}]$ is a $(1-\beta)$-confidence interval for the p-quantile of the distribution of T_n under F. In our simulations we have used $N = 10^8$ and $p = 0.9$, 0.95, and 0.975, i.e. $\alpha = 0.2$, 0.1, and 0.05. We exploited the symmetry of the distribution by actually determining the p-th quantiles with $p = 0.8$, 0.9 and 0.95 using the absolute values $|Y_j|$. The results for a given value of n are tabulated in the row indicated by $\mathrm{df} = n - 1$.

Furthermore, we have chosen $\beta = 2(1 - \Phi(1)) \approx 0.32$ and in our simulation results below we have reported half the length $(Y_{(s)} - Y_{(r)})/2$ of the $\Phi(1) - \Phi(-1) \approx 0.68$-confidence interval $[Y_{(r)}, Y_{(s)}]$ as the value between brackets. We used this convention in analogy with the usual way of reporting a sample mean together with the standard error. The interpretation of the given value between brackets as a standard error is justified also by the fact that the point estimates $Y_{(\lceil pN \rceil)}$ are approximately in the middle of these confidence intervals as well as in the 95% confidence intervals constructed in the same way. Moreover, the 95% confidence intervals are approximately 1.96 times wider than the $(\Phi(1) - \Phi(-1))$-confidence intervals.

We have estimated the different quantiles of the same Student-t distribution via the same Monte Carlo sample of size N. Therefore the estimators in the rows of the following tables are dependent, whereas there should be no dependence between the rows.

These results show that the third digit in our simulated quantiles is very reliable. The unimodal distributions used in these simulations, are ordered according to $\kappa(F)$ as follows

$$\begin{array}{rl} \text{uniform} & -6/5 \\ \text{triangular} & -3/5 \\ \text{normal} & 0 \\ \text{Laplace} & 3 \\ \text{Cauchy} & \infty \end{array} \tag{5.1}$$

Indeed, these results support our claim that the uniform distribution is least favorable in the class of unimodal distributions and that heavier tails (in terms of $\kappa(F)$) result in smaller quantiles for the t-statistic.

TABLE 1. Uniform distribution, estimated p-th quantiles of the t-statistic.

	p=0.9	p=0.95	p=0.975
df= 1;	3.9998(.0010)	9.0009(.0030)	18.9982(.0085)
df= 2;	2.0731(.0004)	3.5881(.0008)	5.7400(.0016)
df= 3;	1.6707(.0003)	2.6314(.0005)	3.8531(.0009)
df= 4;	1.5199(.0002)	2.2577(.0004)	3.1465(.0006)
df= 5;	1.4554(.0002)	2.0737(.0003)	2.7913(.0005)
df=10;	1.3572(.0002)	1.8177(.0002)	2.2724(.0003)
df=20;	1.3171(.0002)	1.7244(.0002)	2.1002(.0003)

Straightforward computation shows that for $n = 2$ and $p \geq 3/4$ we have

$$t_{n-1}(p) = t_1(p) = \tfrac{1}{2}(1-p)^{-1} - 1 \tag{5.2}$$

under the uniform distribution, yielding in Table 1 the exact values 4, 9 and 19, respectively, for df=1. An analytic calculation as in Perlo (1933) shows that the exact values for df=2 are 2.073664, 3.589439 and 5.741739.

TABLE 2. Triangular distribution, estimated p-th quantiles of the t-statistic.

	p=0.9	p=0.95	p=0.975
df= 1;	3.1384(.0007)	6.5498(.0021)	13.2634(.0059)
df= 2;	1.9043(.0003)	2.9574(.0006)	4.3999(.0011)
df= 3;	1.6556(.0002)	2.3911(.0004)	3.2321(.0006)
df= 4;	1.5420(.0002)	2.1670(.0003)	2.8331(.0005)
df= 5;	1.4780(.0002)	2.0409(.0003)	2.6217(.0004)
df=10;	1.3678(.0002)	1.8183(.0002)	2.2492(.0003)
df=20;	1.3218(.0002)	1.7252(.0002)	2.0932(.0002)

TABLE 3. Laplace (double exponential) distribution, estimated p-th quantiles of the t-statistic.

	p=0.9	p=0.95	p=0.975
df= 1;	2.5011(.0005)	5.0027(.0016)	10.0066(.0045)
df= 2;	1.7299(.0002)	2.4583(.0004)	3.4764(.0008)
df= 3;	1.5834(.0002)	2.1321(.0003)	2.7317(.0005)
df= 4;	1.5135(.0002)	1.9981(.0003)	2.4902(.0004)
df= 5;	1.4717(.0002)	1.9248(.0002)	2.3657(.0003)
df=10;	1.3847(.0002)	1.7861(.0002)	2.1490(.0003)
df=20;	1.3371(.0002)	1.7180(.0002)	2.0529(.0002)

In Table II on page 337 of Hotelling (1961) $t_2(0.975)$ has been computed as 3.48, in accordance with our simulations.

TABLE 4. Cauchy distribution, estimated p-th quantiles of the t-statistic.

	p=0.9	p=0.95	p=0.975
df= 1;	2.0825(.0004)	4.0779(.0012)	8.1157(.0036)
df= 2;	1.5342(.0002)	2.1265(.0003)	2.9412(.0007)
df= 3;	1.4241(.0002)	1.8558(.0002)	2.3551(.0004)
df= 4;	1.3769(.0002)	1.7485(.0002)	2.1475(.0003)
df= 5;	1.3506(.0001)	1.6908(.0002)	2.0417(.0003)
df=10;	1.3018(.0001)	1.5881(.0002)	1.8618(.0002)
df=20;	1.2793(.0001)	1.5422(.0002)	1.7849(.0002)

Hotelling (1961) derived $t_2(p)$ for the Cauchy distribution as (cf. his (6.41))

$$t_2(p) = (1 - 3\mathrm{tg}^2(\pi(2p-1)/6))^{-1/2}, \ p \geq \tfrac{1}{2} + \tfrac{3}{\pi}\,\mathrm{arctg}\tfrac{1}{2}, \tag{5.3}$$

which leads to $t_2(0.975) = 2.9412$ in line with our results (note that $t_2(0.995) = 6.46$ and not 3.69 as in (6.42) and Table II of Hotelling (1961)).

For the sake of validation of the method we also compiled the next table, which might be compared to the last one.

TABLE 5. Normal distribution, estimated p-th quantiles of the t-statistic.

	p=0.9	p=0.95	p=0.975
df= 1;	3.0767(.0007)	6.3098(.0020)	12.7031(.0054)
df= 2;	1.8857(.0003)	2.9202(.0006)	4.3023(.0011)
df= 3;	1.6378(.0002)	2.3530(.0004)	3.1819(.0006)
df= 4;	1.5334(.0002)	2.1319(.0003)	2.7764(.0005)
df= 5;	1.4757(.0002)	2.0151(.0003)	2.5707(.0004)
df=10;	1.3721(.0002)	1.8125(.0002)	2.2281(.0003)
df=20;	1.3253(.0002)	1.7246(.0002)	2.0857(.0003)

TABLE 6. Normal distribution, exact p-th quantiles of the Student t-statistic.

	p=0.9	p=0.95	p=0.975
df= 1:	3.077684	6.313752	12.7062
df= 2:	1.885618	2.919986	4.302653
df= 3:	1.637744	2.353363	3.182446
df= 4:	1.533206	2.131847	2.776445
df= 5:	1.475884	2.015048	2.570582
df=10:	1.372184	1.812461	2.228139
df=20:	1.325341	1.724718	2.085963
df=∞:	1.281552	1.644854	1.959964

6. Conclusion

Consider the location problem with an unknown symmetric distribution. Confidence intervals for the location parameter may be based on Student's t-statistic. If the underlying distribution of the observations is known to have a unimodal density this can be done in a conservative way by assuming uniformity of the underlying distribution, if (2.8) holds. In other words, if the t-value in (1.9) is chosen according to the third column of Table 1 then the coverage probability will be at least $2p - 1 = 1 - \alpha = 0.95$, *whatever the unknown symmetric unimodal distribution of the observations.* This claim is supported by the Edgeworth expansion of Section 2 and its consequence of Section 3, by the finite sample analysis of the far tails in Section 4, and by the Monte Carlo results of Section 5.

Standard approaches like the bootstrap (i.e. estimating F) and empirical Edgeworth expansion (i.e. estimating $\kappa(F)$ in (2.5)) are based completely on asymptotic considerations as $n \to \infty$. Consequently, for small sample sizes these techniques are not reliable whereas for large sample sizes the gain in efficiency as compared to our recommendation of using the uniformity assumption, is not dramatic; compare Tables 1 and 6 and note that for $n = 21$ the relative difference in interval length is less than 0.7%.

Acknowledgement We would like to thank the referee for useful suggestions, especially for the reference to Van Zwet (1964a), which led to our Theorem 4.1.

References

1. Bowman, K.O., Beauchamp, J.J. and Shenton, L.R. (1977). The distribution of the t-statistic under non-normality. *Internat. Statist. Rev.* **45**, 233–242.
2. Hall, P. (1987). Edgeworth expansion for Student's t statistic under minimal moment conditions. *Ann. Probab.* **15**, 920–931.
3. Hotelling, H. (1961). The behavior of some standard statistical tests under nonstandard conditions. *Proc. Fourth Berkeley Symp. Math. Statist. Probab.* (Neyman, J. (ed.)) **I**, 319–359.
4. Perlo, V. (1933). On the distribution of Student's ratio for samples drawn from a rectangular distribution. *Biometrika* **25**, 203-204.
5. Sarhan, A.E. and Greenberg B.G. (1962). *Contributions to order statistics.* Wiley, New York.
6. Shakespeare, W. (1601). *Hamlet, Prince of Denmark,* Act III, Scene i, line 56.
7. Student (1908). The probable error of a mean. *Biometrika* **6**, 1–25.
8. Van Zwet, W.R. (1964a). *Convex Transformations of Random Variables,* Mathematical Centre Tracts 7, Mathematical Centre, Amsterdam.
9. Van Zwet, W.R. (1964b). Convex transformations: A new approach to skewness and kurtosis. *Statist. Neerlandica* **18**, 433–441.