Mathematical
Inequalities
& Applications
Volume 12, Number 3 (2009), 455–472

COMPARISON OF LOCATION ESTIMATORS USING BANKS' CRITERION

H. SUSITHA I. KARUNARATNE AND PETROS HADJICOSTAS

(*Communicated by N. Elezović*)

Abstract. In this paper, we analyze further Banks' (1997) closeness criterion for estimators, which is an alternative to Pitman's (1937) closeness criterion. We mainly concentrate our analysis on location estimation, and justify a conjecture by Banks (1997) that for heavy tail distributions the sample median is better than the sample mean when estimating a location parameter. The conlusion is reversed for distributions with lighter tails. To achieve this, we use asymptotics and exact probability calculations.

1. Introduction

The purpose of this paper is to analyze further Banks' (1997) closeness criterion for estimators, which is an alternative to Pitman's (1937) closeness criterion. Even though Banks' criterion can be used to compare estimators that estimate any kind of a parameter, we will mainly concentrate our analysis on location estimation. We begin our discussion with some definitions.

All random variables and vectors are assumed to be defined on a common measurable space $(\Omega, \mathscr{F})$, unless otherwise specified. The data is described by a random vector $\mathbf{X} : \Omega \to \mathbb{R}^p$, where p is a positive integer. The set of all possible values of $\mathbf{X}$ is denoted by $\mathscr{X}$, and we assume $\mathscr{X}$ is a Borel-measurable set, i.e., $\text{Range}(\mathbf{X}) = \mathscr{X} \in \mathscr{B}^{(p)}$. We denote by Θ the parameter space and by $\mathscr{A}$ the estimation space, and we are interested in estimating a function $\kappa(\theta)$ of the parameter $\theta \in \Theta$, where $\kappa : \Theta \to \mathscr{A}$. We assume the estimation space $\mathscr{A}$ is a normed vector space equipped with the norm $\|\cdot\| : \mathscr{A} \to [0, \infty)$.

Assume $\mathscr{C}$ is a σ-field on $\mathscr{A}$ such that $\{a\} \in \mathscr{C}$ for all $a \in \mathscr{A}$. An estimator δ is any measurable function from $\mathscr{X}$ into $\mathscr{A}$, and is a rule that specifies what estimate we have based on what we observed. If we observe $\mathbf{x}$, then the estimate according to the estimator δ is $\delta(\mathbf{x})$.

Mathematics subject classification (2000): 62F10, 62F12, 26D07, 26D15.
Keywords and phrases: Banks' criterion; Gautschi's inequality; location estimation; Pitman's criterion.

Parts of this paper are based on the Master Thesis of the first author at the Department of Mathematics and Statistics at Texas Tech University, cf. [8].

Consider the family of probability measures $(P_\theta : \theta \in \Theta)$, each defined on $(\Omega, \mathscr{F})$. When we have two possible estimators δ_1 and δ_2 for the parameter of interest $\kappa(\theta)$, one can use Pitman's (1937) criterion of "closeness". According to this criterion, if

$$P_\theta \left[\|\delta_1(\mathbf{X}) - \kappa(\theta)\| < \|\delta_2(\mathbf{X}) - \kappa(\theta)\| \right] > \frac{1}{2} \quad \forall \theta \in \Theta, \tag{1}$$

then we choose δ_1 over δ_2, and say that δ_1 is a *closer* estimate than δ_2 (with respect to the family of probability distributions $(P_\theta : \theta \in \Theta)$). This definition is more general than the one given by Pitman (1937), since he only considered $\mathscr{X}$ and Θ that are subsets of the real line. As Nayak (1990) notes, if for some $\theta_0 \in \Theta$ we have $P_{\theta_0} \left[\|\delta_1(\mathbf{X}) - \kappa(\theta_0)\| = \|\delta_2(\mathbf{X}) - \kappa(\theta_0)\| \right] > 0$, we should modify Pitman's criterion as follows: we choose δ_1 over δ_2 if the inequality

$$P_\theta \left[\|\delta_1(\mathbf{X}) - \kappa(\theta)\| < \|\delta_2(\mathbf{X}) - \kappa(\theta)\| \right] \geqslant P_\theta \left[\|\delta_2(\mathbf{X}) - \kappa(\theta)\| < \|\delta_1(\mathbf{X}) - \kappa(\theta)\| \right]$$

holds for all $\theta \in \Theta$ with a strict inequality for at least one $\theta \in \Theta$. This modification is necessary, because now the sum of the two sides of the above inequality is not equal to one.

The concept of Pitman's (1937) criterion for choosing an estimator is modified by Banks (1997). According to Banks, for fixed $\varepsilon > 0$ (whose value depends on the problem and the analyst), we say that we prefer δ_1 over δ_2 if

$$P_\theta[\ \|\delta_1(\mathbf{X}) - \kappa(\theta)\| < \varepsilon\ ] > P_\theta[\ \|\delta_2(\mathbf{X}) - \kappa(\theta)\| < \varepsilon\ ]$$

for all $\theta \in \Theta$. In other words, $\delta_1 \circ \mathbf{X} = \delta_1(\mathbf{X})$ has greater probability of being within ε of $\kappa(\theta)$ than $\delta_2 \circ \mathbf{X} = \delta_2(\mathbf{X})$.

In either case, whether we use Pitman's criterion of "closeness" or Banks' criterion of "closeness," we try to find an estimator that is nearer to the true value than other statistics.

It is well known that Pitman's closeness criterion is intransitive (cf. Keating *et al.* [9, pp. 66-74]). In other words, there are three estimators δ_1, δ_2, and δ_3 such that δ_1 is better than δ_2, δ_2 is better than δ_3, and δ_3 is better than δ_1 according to Pitman's criterion; i.e., for all $\theta \in \Theta$,

$$P_\theta[\|\delta_i(\mathbf{X}) - \kappa(\theta)\| < \|\delta_j(\mathbf{X}) - \kappa(\theta)\|] > 1/2 \quad \text{for} \quad (i,j) \in \{(1,2),(2,3),(3,1)\}.$$

Keating *et al.* [9, p. 128, Th. 4.6.1]) give sufficient conditions under which Pitman's criterion for a family of estimators is transitive.

According to Banks' criterion, for given (small) $\varepsilon > 0$, if we have possible estimators δ_1, δ_2 and δ_3 of a parameter of interest $\kappa(\theta)$, and we prefer δ_1 over δ_2, and δ_2 over δ_3, then for all $\theta \in \Theta$,

$$P_\theta[\|\delta_1(\mathbf{X}) - \kappa(\theta)\| < \varepsilon] > P_\theta[\|\delta_2(\mathbf{X}) - \kappa(\theta)\| < \varepsilon] > P_\theta[\|\delta_3(\mathbf{X}) - \kappa(\theta)\| < \varepsilon].$$

It follows that we prefer δ_1 over δ_3, i.e., Banks' criterion is transitive.

As Banks (1997) observes, there is a connection between risk functions in decision theory and his criterion. For an estimator δ, one possible loss function is the *absolute error loss function*, given by

$$L_1(\theta,\delta(\mathbf{x})) = \|\kappa(\theta)-\delta(\mathbf{x})\| \ \forall\theta\in\Theta \ \forall\mathbf{x}\in\mathscr{X}. \tag{2}$$

Another possible loss function is the *squared error loss function*, given by

$$L_2(\theta,\delta(\mathbf{x})) = \|\kappa(\theta)-\delta(\mathbf{x})\|^2 \ \forall\theta\in\Theta \ \forall\mathbf{x}\in\mathscr{X}. \tag{3}$$

Even though the above losses are widely used in practice, as Banks (1997) notes, there are situations where the use of either one does not make sense. For example, say X is the breaking strength of a roller bearing in a helicopter shaft, $\kappa(\theta)=\theta=E_\theta(X)=\mu$ is the expected value of X with respect to P_θ, and δ_1 and δ_2 are two estimators of θ, each of which is a disastrously bad answer for θ. Then it makes little sense in finding which one is less bad with respect to either loss L_1 or loss L_2. In such a case, it makes more sense to find which estimator is close to the correct answer with the highest probability. Therefore, Banks (1997) proposes using the following loss function:

$$L_3(\theta,\delta(\mathbf{x})) = \begin{cases} 0 \text{ if } \|\kappa(\theta)-\delta(\mathbf{x})\| < \varepsilon, \\ 1 \text{ otherwise,} \end{cases} \quad \forall\theta\in\Theta \ \forall\mathbf{x}\in\mathscr{X}. \tag{4}$$

The risk of an estimator δ with respect to the above loss is

$$R(\theta,\delta) = E_\theta[L_3(\theta,\delta(\mathbf{X}))] = P_\theta(\|\kappa(\theta)-\delta(\mathbf{X})\| \geqslant \varepsilon) = 1-P_\theta(\|\kappa(\theta)-\delta(\mathbf{X})\| < \varepsilon).$$

If we have two estimators δ_1 and δ_2, then for all $\theta\in\Theta$:

$$R(\theta,\delta_1) < R(\theta,\delta_2) \Leftrightarrow P_\theta(\|\kappa(\theta)-\delta_1(\mathbf{X})\| < \varepsilon) > P_\theta(\|\kappa(\theta)-\delta_2(\mathbf{X})\| < \varepsilon).$$

In other words, using loss (4), Banks' estimator between δ_1 and δ_2 is the one with the smaller risk (provided one risk is uniformly smaller than the other for all $\theta\in\Theta$).

It should be noted that Peddada [13], Rao *et al.* [16], Khattree [10], and Khattree and Peddada [11] generalize Pitman's (1937) criterion from the univariate situation to the multivariate case by using loss functions. Corresponding to the loss function $L(\cdot,\cdot):\Theta\times\mathscr{A}\to\mathbb{R}$, according to the generalized Pitman criterion, we prefer estimator δ_1 over estimator δ_2 if

$$P_\theta[L(\theta,\delta_1(\mathbf{X})) < L(\theta,\delta_2(\mathbf{X}))] > 1/2 \quad \forall\theta\in\Theta.$$

If in the above definition we use loss functions L_1 and L_2 given by equations (2) and (3), respectively, then in both cases we get definition (1).

For most of the paper, we assume that the data are independently and identically distributed coming from some univariate probability distribution (with a location parameter). The theoretical mean of the distribution does not necessarily have to exist. The organisation of the paper is as follows. In Section 2, we use asymptotics to compare the sample mean and the sample median of the data using Bank's criterion for various families of probability distributions for which the theoretical mean equals the theoretical median. In Section 3, we use exact probability calculations to compare the sample mean and the sample median using Banks' criterion when the data come from either a normal or a Cauchy distribution, and the sample size is a positive odd integer. Finally, Section 4 contains some concluding remarks.

In this paper we use the following two results (see Lemma 1.1 and Theorem 1.2 below) from Elezović, Giordano and Pečarić (2000). Recall that

$$\psi(z) = \frac{d\left[\ln\Gamma(z)\right]}{dz} = \frac{\Gamma'(z)}{\Gamma(z)}, \quad z > 0,$$

is the well-known digamma function. For $s,t \geqslant 0$ define $\tilde{\psi}_{s,t} : (-\min(s,t),\infty) \to \mathbb{R}$ by

$$\tilde{\psi}_{s,t}(x) = \begin{cases} \frac{\psi(x+t)-\psi(x+s)}{t-s} & \text{if } t \neq s, \\ \psi'(x+s) & \text{if } t = s, \end{cases}$$

for $x > -\min(s,t)$. In Elezović *et al.* (2000), the lemma below is stated for $s,t > 0$ and the domain of x is not stated. A careful examination of the proof shows that the lemma is valid for $x > -\min(r_2,0)$, and is true even when $s = 0$ or $t = 0$.

LEMMA 1.1. *Let $s,t \geqslant 0$ and β_0 be defined by*

$$\beta_0 = -\frac{1}{2} + \sqrt{st + \frac{1}{4}}.$$

Then, for $x > -\min(r_2,0)$,

$$\frac{1}{x+r_1} \leqslant \tilde{\psi}_{s,t}(x) \leqslant \frac{1}{x+r_2}, \tag{5}$$

where

$$r_1 := \max\left\{\frac{s+t-1}{2}, \beta_0\right\}, \quad r_2 := \min\left\{\frac{s+t-1}{2}, \beta_0\right\}.$$

Each equality in (5) holds if and only if $|t-s| = 1$.

For $s,t > 0$ with $r = \min(s,t)$ define the function $z_{s,t} : (-r,+\infty) \to \mathbb{R}$ by

$$z_{s,t}(x) = \begin{cases} \left(\frac{\Gamma(x+t)}{\Gamma(x+s)}\right)^{\frac{1}{t-s}} - x & \text{for } t \neq s, \\ e^{\psi(x+t)} - x & \text{for } t = s, \end{cases}$$

for $x > -r$. Note that $z_{s+\alpha,t+\alpha}(x-\alpha) = z_{s,t}(x) + \alpha$ for $x > \alpha - r > -r$. The following inequalities are known as Gautschi's inequalities, and the best bounds were obtained by Elezović *et al.* (2000).

THEOREM 1.2. *Let* $s,t>0$. *For all* $x>x_0>-\min(s,t)$ *the inequalities*

$$\frac{s+t-1}{2}<z_{s,t}(x)<z_{s,t}(x_0)$$

hold whenever $|t-s|<1$, *and with reversed signs whenever* $|t-s|>1$. *The bounds are the best possible.* (*When* $|t-s|=1$, *we have* $z_{s,t}(x)=(s+t-1)/2$ *for all* $x>-\min(s,t)$.)

In p. 245 of Elezović *et al.* (2000), Lemma 1.1 above is used in the proof of a theorem that leads to Theorem 1.2. It is used to prove that $\lim_{x\to\infty} z'_{s,t}(x)\leqslant 0$ for the case $|t-s|<1$. This follows from the fact that for $|t-s|<1$ the function $z'_{s,t}$ is (strictly) increasing on $(0,\infty)$, and the inequality

$$z'_{s,t}(x)<\left(\frac{\Gamma(x+t)}{\Gamma(x+s)}\right)^{1/(t-s)}\cdot\frac{1}{x+r_2}-1,$$

which is valid for $x>-\min(r_2,0)$ (even though this is not stated explicitly in p. 245 of the paper). Lemma 1.1 is also needed to prove similar results for the case $|t-s|>1$. For related results, see also Qi, Guo and Chen (2006).

2. Comparison of the sample mean and the sample median using asymptotics

In this section we use asymptotics to compare the sample mean and sample median using Banks' criterion. For the rest of the section, assume

$$\Theta\subseteq\mathbb{R}\times(0,\infty)=\{(\mu,\sigma^2):-\infty<\mu<\infty,0<\sigma^2<\infty\},$$

and $\mathscr{A}=\mathbb{R}=\{\mu:-\infty<\mu<\infty\}$. We are interested in estimating

$$\kappa(\theta)=\kappa(\mu,\sigma^2)=\mu.$$

Let $(P_\theta:\theta\in\Theta)$ be a family of probability measures defined on the measurable space $(\Omega,\mathscr{F})$, and let $(F_\theta:\theta\in\Theta)$ be a collection of univariate cumulative distribution functions (c.d.f.) such that for each $\theta=(\mu,\sigma^2)\in\Theta$:

(a) μ is the finite mean of F_θ (i.e., $\int|x|\,dF_\theta(x)<\infty$ and $\mu=\int x\,dF_\theta(x)$);

(b) σ^2 is the finite variance of F_θ (i.e., $0<\sigma^2=\int(x-\mu)^2\,dF_\theta(x)<\infty$).

We denote by $m=m(\theta)$ any median of F_θ.

All random variables are defined on $(\Omega,\mathscr{F})$, and as usual, "i.i.d." means "independently identically distributed" (with respect to the family $(P_\theta:\theta\in\Theta)$). Denote by $\Phi(\cdot)$ the cumulative distribution function of $N(0,1)$, the standard normal distribution. The main result of the section is about distributions F_θ for which all the medians equal μ. Define $\mathbb{N}^*=\{1,2,3,\ldots\}$.

THEOREM 2.1. *For each* $\theta=(\mu,\sigma^2)\in\Theta$ *and each positive integer* n, *assume* $X_{n1},X_{n2},\ldots,X_{nn}$ *are i.i.d. random variables with a c.d.f.* F_θ *that satisfies properties* (a) *and* (b) *above, i.e., for* $i=1,2,\ldots,n$, $P_\theta(X_{ni}\leqslant x)=F_\theta(x)$ *for all* $x\in\mathbb{R}$. *For* *each $\theta \in \Theta$, assume F_θ has a unique median $m = m(\theta)$ and is differentiable in a neighborhood of m, say $(m-\eta, m+\eta)$ for some $\eta = \eta(\theta) > 0$, with derivative $f_\theta : (m-\eta, m+\eta) \to \mathbb{R}$ satisfying $0 < f_\theta(m) < \infty$. Let $\overline{X}_n$ be the sample mean, and $\widetilde{X}_n$ be a sample median of the observations. Let $\varepsilon = (\varepsilon_n : n \in \mathbb{N}^*)$ be a sequence of positive constants such that $\lim_{n\to\infty} \sqrt{n}\,\varepsilon_n = c$ for some constant $c > 0$. Define*

$$S_\theta(\varepsilon) = \lim_{n\to\infty} P_\theta(|\overline{X}_n - \mu| < \varepsilon_n) - \lim_{n\to\infty} P_\theta(|\widetilde{X}_n - \mu| < \varepsilon_n),$$

if both limits exist. If $\mu = m(\theta)$, then $S_\theta(\varepsilon)$ exists, and $S_\theta(\varepsilon) > 0$ if and only if $\sigma^2 < 1/(4f_\theta^2(m))$.

Proof. By the Central Limit Theorem, as $n \to \infty$, $\sqrt{n}(\overline{X}_n - \mu)/\sigma$ converges in distribution (under P_θ) to $N(0,1)$. Let Z be a random variable with a $N(0,1)$ distribution. Since the absolute value is a continuous function, $\sqrt{n}|(\overline{X}_n - \mu)/\sigma|$ converges in distribution (under P_θ) to $|Z|$. By Lemma 2.11 in [18, p. 12],

$$\lim_{n\to\infty} \left(\sup_{z\in\mathbb{R}} \left| P_\theta(\sqrt{n}|\overline{X}_n - \mu|/\sigma < z) - P_\theta(|Z| < z) \right| \right) = 0.$$

Using the above equation, it is not difficult to show that

$$\lim_{n\to\infty} |P_\theta(|\overline{X}_n - \mu| < \varepsilon_n) - (2\Phi(\varepsilon_n\sqrt{n}/\sigma) - 1)| = 0.$$

Since also $\lim_{n\to\infty}(2\Phi(\varepsilon_n\sqrt{n}/\sigma) - 1) = 2\Phi(c/\sigma) - 1$, we have

$$\lim_{n\to\infty} P_\theta(|\overline{X}_n - \mu| < \varepsilon_n) = 2\Phi(c/\sigma) - 1.$$

By Theorem 7.25 in Schervish [17, p. 405], as $n \to \infty$, $2f_\theta(m)\sqrt{n}\left(\widetilde{X}_n - m\right)$ converges in distribution (under P_θ) to $N(0,1)$. Using this fact, Lemma 2.11 in [18, p. 12]), and the assumptions $\mu = m$ and $\lim_{n\to\infty} \sqrt{n}\varepsilon_n = c > 0$, one can show that

$$\lim_{n\to\infty} P_\theta(|\widetilde{X}_n - \mu| < \varepsilon_n) = 2\Phi(2f_\theta(m)c) - 1.$$

It follows $S_\theta(\varepsilon)$ exists and equals $2[\Phi(c/\sigma) - \Phi(2f_\theta(m)c)]$. Since Φ is strictly increasing, it is easy to prove the conclusion of the theorem. □

The previous theorem says that, if $\mu = m$ and we have a large random sample, the sample mean is a better estimator than the sample median for estimating the theoretical mean if and only if $\sigma^2 < 1/(4f_\theta^2(m))$. This condition does not depend on ε. It is equivalent to $\sigma^2/n < 1/(4nf_\theta^2(m))$, which says that the variance of the sample mean is less than the asymptotic variance of the sample median.

When the data come from $N(\mu, \sigma^2)$, a normal distribution with mean μ and variance σ^2 (with $0 < \sigma^2 < \infty$), or when they come from $U(a,b)$, a uniform distribution with parameters a and b (where $a < b$), the assumptions of Theorem 2.1 hold. In both cases the theoretical mean equals the theoretical median, and $\sigma^2 < 1/(4f_\theta^2(m))$.

Indeed for the $N(\mu,\sigma^2)$ case, $f_\theta(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$ for $-\infty < x < \infty$, and so $f_\theta(m) = \frac{1}{\sqrt{2\pi}\sigma}$. It follows $f_\theta^2(m) = \left(\frac{1}{\sqrt{2\pi}\sigma}\right)^2 < \frac{1}{4\sigma^2}$, i.e., $\sigma^2 < 1/(4f_\theta^2(m))$. For the $U(a,b)$ case, $f_\theta(x) = \frac{1}{b-a}$ for $a < x < b$, and $m = \mu = \frac{a+b}{2}$. It is then easy to show $f_\theta^2(m) = (\frac{1}{b-a})^2 < \frac{1}{4\sigma^2}$, where $\sigma^2 = \frac{(b-a)^2}{12}$.

If F_θ is the c.d.f. of a beta(α,β) distribution with $\alpha,\beta > 0$, then $\mu = \frac{\alpha}{\alpha+\beta}$ and $\sigma^2 = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$. If $\alpha = \beta$, we let

$$\Theta = \left\{(\mu,\sigma^2) = \left(\frac{1}{2}, \frac{1}{4(2\alpha+1)}\right) : \alpha > 0\right\},$$

and we have the following result.

THEOREM 2.2. *Suppose f_θ is the probability density function of beta(α,β) distribution with $\alpha = \beta > 0$. Then $m = \mu = 1/2$ and $\sigma^2 < 1/(4f^2(m))$.*

Proof. We have

$$f_\theta(x) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1}, \qquad 0 < x < 1,$$

and therefore

$$f_\theta(m) = \frac{\Gamma(2\alpha)}{(\Gamma(\alpha))^2}\left(\frac{1}{2}\right)^{2(\alpha-1)}.$$

We need to show that $f_\theta^2(m) < \frac{1}{4\sigma^2}$, which is equivalent to:

$$\frac{\Gamma^2(2\alpha)}{(\Gamma(\alpha))^4}\left(\frac{1}{2}\right)^{4(\alpha-1)} < 2\alpha+1. \tag{6}$$

Taking logarithms on both sides, the previous inequality is equivalent to:

$$S(\alpha) = 2\ln\Gamma(2\alpha) - 4\ln\Gamma(\alpha) + 4(1-\alpha)\ln 2 - \ln(2\alpha+1) < 0.$$

Then

$$\frac{dS}{d\alpha} = 4\psi(2\alpha) - 4\psi(\alpha) - 4\ln 2 - \frac{2}{2\alpha+1} \text{ for } \alpha > 0.$$

Since $\psi(2\alpha) = \frac{1}{2}\psi(\alpha) + \frac{1}{2}\psi\left(\alpha+\frac{1}{2}\right) + \ln 2$ (see [1, p. 259, formula 6.3.8]), we have:

$$\frac{dS}{d\alpha} = 2\left[\psi\left(\alpha+\frac{1}{2}\right) - \psi(\alpha) - \frac{1}{2\alpha+1}\right].$$

Let $T(\alpha) = \psi\left(\alpha+\frac{1}{2}\right) - \psi(\alpha) - \frac{1}{2\alpha+1}$. Then:

$$T(\alpha+1) - T(\alpha) = \psi\left(\alpha+\frac{3}{2}\right) - \psi\left(\alpha+\frac{1}{2}\right) - \left[\psi(\alpha+1) - \psi(\alpha)\right] - \frac{1}{2\alpha+3} + \frac{1}{2\alpha+1}.$$

Since $\psi(\alpha+1) = \psi(\alpha) + \frac{1}{\alpha}$ (see [1, p. 258, formula 6.3.5]), we have:

$$T(\alpha+1) - T(\alpha) = -\frac{3}{(2\alpha+1)\alpha(2\alpha+3)} < 0 \quad \text{for } \alpha > 0.$$

Hence, for fixed $\alpha > 0$ we have: $T(\alpha) > T(\alpha+1) > T(\alpha+2) > \ldots > T(\alpha+n)$ for all positive integers n.

We need to show that $\lim_{\alpha\to\infty} T(\alpha) = 0$. Since the function $\psi(\alpha)$ is increasing for $\alpha > 0$, $\psi(\alpha) \leqslant \psi(\alpha+\frac{1}{2}) \leqslant \psi(\alpha+1)$. Therefore for $\alpha > 0$:

$$-\frac{1}{2\alpha+1} \leqslant T(\alpha) = \psi\Big(\alpha+\frac{1}{2}\Big) - \psi(\alpha) - \frac{1}{2\alpha+1} \leqslant \psi\Big(\alpha+1\Big) - \psi(\alpha) - \frac{1}{2\alpha+1}.$$

Since $\psi\Big(\alpha+1\Big) - \psi(\alpha) - \frac{1}{2\alpha+1} = \frac{1}{\alpha} - \frac{1}{2\alpha+1}$, we can rewrite the above as:

$$-\frac{1}{2\alpha+1} \leqslant T(\alpha) \leqslant \frac{1}{\alpha} - \frac{1}{2\alpha+1}.$$

It follows that $\lim_{\alpha\to\infty} T(\alpha) = 0$. Therefore for a fixed $\alpha > 0, \lim_{n\to\infty} T(\alpha+n) = 0$, and so $T(\alpha) > T(\alpha+1) \geqslant \lim_{n\to\infty} T(\alpha+n) = 0$. Hence, $\frac{dS}{d\alpha} = 2T(\alpha) > 0$ for $\alpha > 0$. Therefore, $S(\alpha)$ is strictly increasing on $(0,\infty)$. By Lemma A.1 in the appendix, $\lim_{\alpha\to\infty} S(\alpha) = \ln\left(\frac{2}{\pi}\right) < \ln(1) = 0$, and so $S(\alpha) < 0$. This means that $f_\theta^2(m) < \frac{1}{4\sigma^2}$, and the proof of the theorem is complete. □

It is well-known that the mean of a t-distribution with $\alpha > 0$ degrees of freedom exists and is finite if and only if $\alpha > 1$. In addition, for the same probability distribution, the variance exists and is finite if and only if $\alpha > 2$. We have the next theorem.

THEOREM 2.3. *Consider the t distribution* $\mathscr{J}(\alpha,\mu,\tau^2)$ *with* $\alpha > 2$ *degrees of freedom, location parameter* $m = \mu$, *and scale parameter* τ^2, *where* $-\infty < \mu < \infty$ *and* $0 < \tau^2 < \infty$. *Then there is a unique* $\alpha_0 > 2$ *such that* $\sigma^2 < 1/(4f_\theta^2(m))$ *if and only if* $\alpha_0 < \alpha < \infty$.

Proof. We have (see [3, p. 561])

$$f_\theta(x) = \frac{\Gamma[(\alpha+1)/2]}{\tau(\alpha\pi)^{1/2}\Gamma(\alpha/2)}\Big(1+\frac{(x-\mu)^2}{\alpha\tau^2}\Big)^{-(\alpha+1)/2} \quad \text{for } -\infty < x < \infty,$$

with mean μ and variance $\sigma^2 = \frac{\alpha\tau^2}{\alpha-2}$. We have

$$f_\theta(m) = \frac{\Gamma[(\alpha+1)/2]}{\tau(\alpha\pi)^{1/2}\Gamma(\alpha/2)}.$$

The inequality $\sigma^2 < 1/(4f_\theta^2(m))$ is equivalent to:

$$\frac{\Gamma^2[(\alpha+1)/2]}{\Gamma^2(\alpha/2)} < \frac{(\alpha-2)\pi}{4}.$$

Taking logarithms on both sides, we see that the previous inequality is equivalent to:

$$M(\alpha) := 2\ln\Gamma\Big(\frac{\alpha+1}{2}\Big) - 2\ln\Gamma\Big(\frac{\alpha}{2}\Big) - \ln(\alpha-2) - \ln\Big(\frac{\pi}{4}\Big) < 0.$$

Taking the derivative of M with respect to α, we have:

$$\frac{dM}{d\alpha} = \psi\Big(\frac{\alpha+1}{2}\Big) - \psi\Big(\frac{\alpha}{2}\Big) - \frac{1}{\alpha-2} \text{ for } \alpha > 2.$$

Applying Lemma 1.1 with $x = \alpha/2$, $t = 1/2$ and $s = 0$, we get

$$\psi\left(\frac{\alpha+1}{2}\right) - \psi\left(\frac{\alpha}{2}\right) < \frac{1}{\alpha-\frac{1}{2}} < \frac{1}{\alpha-2}$$

because $\beta_0 = 0$, $r_1 = \max(-1/4, 0) = 0$, $r_2 = \min(-1/4, 0) = -1/4$, and

$$x = \frac{\alpha}{2} > -\min(r_2, 0) = \frac{1}{4}$$

(since $\alpha > 2$). Hence, $\frac{dM}{d\alpha} = T(\alpha) < 0$ for $\alpha > 2$. Therefore, M is strictly decreasing on $(2,\infty)$. Note that $\lim_{\alpha\to 2^+} M(\alpha) = +\infty$. By Lemma A.2 in the appendix, $\lim_{\alpha\to\infty} M(\alpha) = \ln(2) - \ln(\pi) < 0$. Since M is strictly decreasing and continuous in $(2,\infty)$, and $M(2^+) = \infty > 0 > M(\infty)$, by the Intermediate Value Theorem, there is a unique $\alpha_0 > 2$ such that $M(\alpha) > 0$ if and only if $2 < \alpha < \alpha_0$, and $M(\alpha) < 0$ if and only if $\alpha_0 < \alpha < \infty$. Since $M(\alpha) < 0$ if and only if $\sigma^2 < 1/(4f_\theta^2(m))$, the proof of the theorem is complete. □

Using the bisection method and a Matlab program provided by P. Seshaiyer, we estimated $\alpha_0 \approx 4.67879$. (This can also be checked in Maple using the *fsolve* command.)

COROLLARY 2.4. *Let* α_0 *be the number defined in Theorem* 2.3. *Then*

$$2 + \frac{3}{\pi-2} < \alpha_0 < 2 + \frac{2+\frac{4}{\pi}}{\pi-2}.$$

These estimates give $4.62790 < \alpha_0 < 4.86726$.

Proof. By Theorem 1.2 with $t = \alpha_0/2$, $s = (\alpha_0-1)/2$, $x = 1/2$ and $x_0 = 1 - (\alpha_0/2)$ we obtain

$$\frac{\frac{\alpha_0-1}{2} + \frac{\alpha_0}{2} - 1}{2} < \left[\frac{\Gamma\left(\frac{\alpha_0+1}{2}\right)}{\Gamma\left(\frac{\alpha_0}{2}\right)}\right]^2 - \frac{1}{2} < \left[\frac{\Gamma(1)}{\Gamma\left(\frac{1}{2}\right)}\right]^2 - \left(1 - \frac{\alpha_0}{2}\right). \tag{7}$$

It follows from the proof of Theorem 2.3 that

$$\left[\frac{\Gamma\left(\frac{\alpha_0+1}{2}\right)}{\Gamma\left(\frac{\alpha_0}{2}\right)}\right]^2 = \frac{\pi(\alpha_0-2)}{4}. \tag{8}$$

Using (7) and (8), we can easily prove the corollary. □

3. Comparison of the sample mean and the sample median using exact probability calculations

In the previous section we proved some asymptotic results that allow us to use Banks' criterion to compare the sample mean and the sample median as estimators of a location parameter when the sample size is large. In this section we do the same for finite samples using exact probability calculations when the sample size n is an *odd* positive integer. We comment about even sample sizes at the end of the section because their treatment is much more difficult.

As before, let $(P_\theta : \theta \in \Theta)$ be a family of probability measures defined on $(\Omega, \mathscr{F})$. For each $\theta \in \Theta$ and each positive integer n, assume $X_{n1}, X_{n2}, \ldots, X_{nn}$ are i.i.d. random variables with c.d.f. F_θ and a unique theoretical median $m = m(\theta)$ (i.e., $P_\theta(X_{ni} \leqslant x) = F_\theta(x)$ and $F_\theta(m) = 1/2$). If $X_{n(1)}, X_{n(2)}, \cdots, X_{n(n)}$ are the ordered statistics for the random sample $X_{n1}, X_{n2}, \ldots, X_{nn}$, then for odd integer n, the sample median $\widetilde{X}_n$ is the middle ordered statistic $X_{((n+1)/2)}$. In such a case, if F_θ possesses a probability density function (p.d.f.) f_θ with respect to (w.r.t.) the Lebesgue measure, then it is well-known that the p.d.f. of the sample median $\widetilde{X}_n$ (w.r.t. the Lebesgue measure) is

$$g_{n,\theta}(y) = \frac{\Gamma(2\beta_n)}{\Gamma(\beta_n)^2}[F_\theta(y)(1 - F_\theta(y))]^{\beta_n - 1} f_\theta(y), \quad -\infty < y < \infty, \tag{9}$$

where $\beta_n = (n+1)/2$; e.g. see Billingsley [4, p. 200, Ex. 14.7]. It follows that for $\varepsilon > 0$:

$$P_\theta(|\widetilde{X}_n - m| < \varepsilon) = \int_{F_\theta(m-\varepsilon)}^{F_\theta(m+\varepsilon)} \frac{\Gamma(2\beta_n)}{\Gamma(\beta_n)^2}(u(1-u))^{\beta_n - 1}\,\mathrm{d}u.$$

If F_θ is locally symmetric around the point $m = m(\theta)$, i.e., there is $\eta = \eta(\theta) \in (0, \infty]$ such that $F_\theta(m-y) + F_\theta(m+y) = 1$ for all $y \in (-\eta, \eta)$, then for all $\varepsilon \in (0, \eta)$,

$$P_\theta(|\widetilde{X}_n - m| < \varepsilon) = 2\int_0^{F_\theta(m+\varepsilon)} \frac{\Gamma(2\beta_n)}{\Gamma(\beta_n)^2}(u(1-u))^{\beta_n - 1}\,\mathrm{d}u - 1. \tag{10}$$

Finding a similar expression for the sample mean is in general more difficult. The most notable exceptions to this occur when F_θ is the c.d.f. of a normal, a Cauchy, or a gamma distribution. We will deal only with the first two cases.

THEOREM 3.1. *Let $\Theta = \{(\mu, \sigma^2) : -\infty < \mu < \infty, 0 < \sigma^2 < \infty\}$, and for each $\theta = (\mu, \sigma^2) \in \Theta$, let F_θ be the c.d.f. of a $N(\mu, \sigma^2)$ distribution. For each $\theta \in \Theta$ and each positive odd integer $n \geqslant 3$, assume $X_{n1}, X_{n2}, \ldots, X_{nn}$ are i.i.d. random variables with c.d.f. F_θ. Let $\overline{X}_n$ be the sample mean, and $\widetilde{X}_n$ be the sample median of the observations. For all $\theta \in \Theta$ and all $\varepsilon > 0$,*

$$P_\theta(|\overline{X}_n - \mu| < \varepsilon) > P_\theta(|\widetilde{X}_n - \mu| < \varepsilon). \tag{11}$$

Proof. Using equation (10) and properties of the normal distribution, it is easy to see that inequality (11) will follow from the following stronger inequality:

$$w(\alpha, \beta) := \Phi(\alpha\sqrt{2\beta - 1}) - \int_0^{\Phi(\alpha)} \frac{\Gamma(2\beta)}{\Gamma(\beta)^2}(u(1-u))^{\beta - 1}\mathrm{d}u > 0, \tag{12}$$

where $\alpha > 0$ and $\beta > 1$. We have

$$\frac{\partial w}{\partial \alpha}(\alpha,\beta) = \sqrt{2\beta-1}\Phi'(\alpha\sqrt{2\beta-1}) - \Phi'(\alpha)\frac{\Gamma(2\beta)}{\Gamma(\beta)^2}[\Phi(\alpha)(1-\Phi(\alpha))]^{\beta-1},$$

and $\Phi'(\alpha) = \frac{1}{\sqrt{2\pi}}e^{-\alpha^2/2}$.

We first show that for each $\beta > 1$, the equation $\frac{\partial w}{\partial \alpha}(\alpha_0,\beta) = 0$ holds for at most one $\alpha_0 \in (0,\infty)$. The equation is equivalent to

$$g(\alpha_0) = \left[\sqrt{2\beta-1}\frac{\Gamma(\beta)^2}{\Gamma(2\beta)}\right]^{1/(\beta-1)}, \tag{13}$$

where $g(\alpha) = e^{\alpha^2}\Phi(\alpha)(1-\Phi(\alpha))$. Note that $g(0) = 1/4$, and

$$\frac{\partial(\ln g)(\alpha)}{\partial \alpha} = \Phi'(\alpha)\left[\frac{2\alpha}{\Phi'(\alpha)} - \frac{1}{1-\Phi(\alpha)} + \frac{1}{\Phi(\alpha)}\right].$$

By Lemma A.3 in the appendix, $\frac{\partial(\ln g)(\alpha)}{\partial \alpha} > 0$, so the function $g(\alpha)$ is strictly increasing for $\alpha > 0$. Therefore, equation (13) has at most one solution $\alpha_0 \in (0,\infty)$ (for fixed $\beta > 1$).

By Lemma A.4 in the appendix, for each $\beta > 1$,

$$\frac{\partial w}{\partial \alpha}(0,\beta) = \frac{1}{\sqrt{2\pi}}\left[\sqrt{2\beta-1} - \frac{\Gamma(2\beta)}{\Gamma(\beta)^2}\left(\frac{1}{4}\right)^{\beta-1}\right] > 0.$$

Also, for each $\beta > 1$, $\lim_{\alpha\to\infty}\frac{\partial w}{\partial \alpha}(\alpha,\beta) = 0$. Since for each $\beta > 1$, $\lim_{\alpha\to 0^+} w(\alpha,\beta)$ $= 0 =$ $\lim_{\alpha\to+\infty} w(\alpha,\beta)$, and since for each $\beta > 1$, both $w(\alpha,\beta)$ and $\frac{\partial w}{\partial \alpha}(\alpha,\beta)$ are continuous functions of $\alpha \in (0,\infty)$, we conclude that $\frac{\partial w}{\partial \alpha}(\alpha_0,\beta) = 0$ for exactly one $\alpha_0 \in (0,\infty)$, $w(\alpha,\beta)$ has a maximum at $\alpha = \alpha_0$ and minimum at $\alpha = 0$ and $\alpha = \infty$. Therefore, for each $\beta > 1$, $w(\alpha,\beta) > 0$ for $0 < \alpha < \infty$, and the proof of the theorem is complete. □

The next theorem deals with the Cauchy distribution, which is a special case of the t distribution with 1 degree of freedom (see the statement and the proof of Theorem 2.3).

THEOREM 3.2. *Let* $\Theta = \{(m,\tau^2) : -\infty < m < \infty, 0 < \tau^2 < \infty\}$, *and for each* $\theta = (m,\tau^2) \in \Theta$, *let* F_θ *be the c.d.f. of a* t*-distribution* $\mathscr{T}(1,m,\tau^2)$ *with* 1 *degree of freedom, location parameter* m, *and scale parameter* τ^2. *For each* $\theta \in \Theta$ *and each positive odd integer* $n \geqslant 3$, *assume* $X_{n1}, X_{n2}, \ldots, X_{nn}$ *are i.i.d. random variables with c.d.f.* F_θ. *Let* $\overline{X}_n$ *be the sample mean, and* $\widetilde{X}_n$ *be the sample median of the observations. Then for each* $\theta \in \Theta$ *and each* $\varepsilon > 0$,

$$P_\theta(|\overline{X}_n - m| < \varepsilon) < P_\theta(|\widetilde{X}_n - m| < \varepsilon). \tag{14}$$

Proof. Let Φ_C be the c.d.f. of a Cauchy distribution with location parameter 0 and scale parameter 1. Then $F_\theta(a) = \Phi_C((a-m)/\tau)$. Also, using characteristic functions, one can show that $\overline{X}_n$ has a $\mathscr{J}(1,m,\tau^2)$ distribution under P_θ (see also [4, Problem 20.20, p. 278]). Using these facts and equation (10), it is easy to see that inequality (14) will follow from the next stronger inequality:

$$g(\alpha,\beta) := \Phi_C(\alpha) - \int_0^{\Phi_C(\alpha)} \frac{\Gamma(2\beta)}{\Gamma(\beta)^2}(u(1-u))^{\beta-1}\,\mathrm{d}u < 0, \tag{15}$$

where $\alpha > 0$ and $\beta > 1$. Since for each $\beta > 1$, $g(\alpha,\beta)$ and $\frac{\partial g(\alpha,\beta)}{\partial \alpha}$ are continuous functions of $\alpha \in (0,\infty)$, to show the inequality, it is enough to show that: (a) $\frac{\partial g(\alpha,\beta)}{\partial \alpha} = 0$ has at most one solution in $\alpha \in (0,\infty)$ for each $\beta > 1$; (b) $\frac{\partial g(0,\beta)}{\partial \alpha} < 0$ for each $\beta > 1$; and (c) $g(0,\beta) = 0 = \lim_{\alpha\to\infty} g(\alpha,\beta)$ for each $\beta > 1$. We have:

$$\frac{\partial g(\alpha,\beta)}{\partial \alpha} = \Phi_C'(\alpha)\left\{1 - \frac{\Gamma(2\beta)}{\Gamma(\beta)^2}\left[\Phi_C(\alpha)(1-\Phi_C(\alpha))\right]^{\beta-1}\right\},$$

with

$$\Phi_C(\alpha) = \frac{1}{2} + \frac{\arctan(\alpha)}{\pi} \quad \text{and} \quad \Phi_C'(\alpha) = \frac{1}{\pi(1+\alpha^2)}.$$

(We assume $-\frac{\pi}{2} < \arctan(\alpha) < \frac{\pi}{2}$ for real α.)

(a) If $\alpha > 0$ and $\beta > 1$, the equation $\frac{\partial g(\alpha,\beta)}{\partial \alpha} = 0$ is equivalent to

$$\Phi_C(\alpha)(1-\Phi_C(\alpha)) = \left[\frac{\Gamma(\beta)^2}{\Gamma(2\beta)}\right]^{1/(\beta-1)}. \tag{16}$$

Note that, if k is a constant, the equation $x^2 - x + k = 0$ has at most two distinct real solutions, of which at most one is greater than or equal to $1/2$. Since $1/2 \leqslant \Phi_C(\alpha) \leqslant 1$ for $\alpha > 0$, we conclude that for fixed $\beta > 1$, equation (16) has at most one solution in $\alpha > 0$.

(b) Since

$$\frac{\partial g(0,\beta)}{\partial \alpha} = \frac{1}{\pi}\left[1 - \frac{\Gamma(2\beta)}{\Gamma(\beta)^2}\left(\frac{1}{4}\right)^{\beta-1}\right],$$

it follows from Lemma A.4 in the appendix that $\frac{\partial g(0,\beta)}{\partial \alpha} < 0$ for each $\beta > 1$.

(c) Since $\Phi_C(0) = 1/2$ and $\lim_{\alpha\to\infty}\Phi_C(\alpha) = 1$, and since

$$\int_0^{1/2} \frac{\Gamma(2\beta)}{\Gamma(\beta)^2}(u(1-u))^{\beta-1}\,\mathrm{d}u = \frac{1}{2}\int_0^1 \frac{\Gamma(2\beta)}{\Gamma(\beta)^2}(u(1-u))^{\beta-1}\,\mathrm{d}u = \frac{1}{2},$$

we conclude that $g(0,\beta) = 0 = \lim_{\alpha\to\infty} g(\alpha,\beta)$ for each $\beta > 1$.

This completes the proof of the theorem. □

Unfortunately, for even sample sizes n, formula (9) is not useful anymore. A sample median $\widetilde{X}_n$ is any convex combination of the middle two ordered statistics:

$\widetilde{X}_n(\alpha) = \alpha X_{n(n/2)} + (1-\alpha)X_{n((n/2)+1)}$, where $0 \leqslant \alpha \leqslant 1$. In most elementary Statistics books, various authors use $\alpha = 1/2$, i.e.,

$$\widetilde{X}_n = \widetilde{X}_n(1/2) = \frac{X_{n(n/2)} + X_{n((n/2)+1)}}{2}.$$

The next theorem gives a formula for $P_\theta(|\widetilde{X}_n(\alpha) - m| < \varepsilon)$ when n is even. For $i = 1,2,\ldots,n$, we denote by $X_{n(i)}$ the i^{th} ordered statistics of the random sample $X_{n1},\ldots,X_{nn}$.

THEOREM 3.3. *Let n be an even positive integer and $0 \leqslant \alpha \leqslant 1$. For each $\theta \in \Theta$, assume $X_{n1}, X_{n2}, \ldots, X_{nn}$ are i.i.d. random variables with c.d.f. F_θ and a unique theoretical median $m = m(\theta)$. Let $X_{n(1)}, X_{n(2)}, \ldots, X_{n(n)}$ be the ordered statistics and define $\widetilde{X}_n(\alpha) = \alpha X_{n(n/2)} + (1-\alpha)X_{n((n/2)+1)}$. If F_θ possesses a p.d.f. f_θ w.r.t. the Lebesgue measure, then for each $\theta \in \Theta$ and $\varepsilon > 0$,*

$$P_\theta(|\widetilde{X}_n(\alpha) - m| < \varepsilon) = \frac{2\Gamma(n)}{\Gamma\left(\frac{n}{2}\right)^2}(a - b + c), \tag{17}$$

where

$$a = \int_{-\infty}^{m-\varepsilon} F_\theta(y)^{\frac{n}{2}-1} f_\theta(y) \left[1 - F_\theta\left(\frac{m-\varepsilon-\alpha y}{1-\alpha}\right)\right]^{n/2} dy,$$
$$b = \int_{-\infty}^{m+\varepsilon} F_\theta(y)^{\frac{n}{2}-1} f_\theta(y) \left[1 - F_\theta\left(\frac{m+\varepsilon-\alpha y}{1-\alpha}\right)\right]^{n/2} dy,$$
$$c = \int_{m-\varepsilon}^{m+\varepsilon} F_\theta(y)^{\frac{n}{2}-1} f_\theta(y) \left[1 - F_\theta(y)\right]^{n/2} dy.$$

Proof. See Appendix A.4. □

In the case of a location-scale family of p.d.f.'s, Theorem 3.3 can be simplified. The proof of the corollary is easy and hence is omitted.

COROLLARY 3.4. *Let $\Theta = \{(m,\tau^2) : -\infty < m < \infty, 0 < \tau^2 < \infty\}$ and let $\tilde{\Phi}$ be an absolutely continuous c.d.f. with respect to the Lebesgue measure with a p.d.f. $\tilde{\phi}$. Assume that $\tilde{\Phi}$ has a unique theoretical median equal to 0. For each $\theta = (m,\tau^2) \in \Theta$ define F_θ to be a c.d.f. such that*

$$F_\theta(x) = \tilde{\Phi}\left(\frac{x-m}{\tau}\right)$$

for each $x \in (-\infty,\infty)$ (where $\tau = +\sqrt{\tau^2}$).

Let n be an even positive integer and $0 \leqslant \alpha \leqslant 1$. For each $\theta \in \Theta$, assume $X_{n1}, X_{n2}, \ldots, X_{nn}$ are i.i.d. random variables with c.d.f. F_θ. Then for each $\theta \in \Theta$ and $\varepsilon > 0$,

$$P_\theta(|\widetilde{X}_n(\alpha) - m| < \varepsilon) = \frac{2\Gamma(n)}{\Gamma\left(\frac{n}{2}\right)^2}(a - b + c), \tag{18}$$

where

$$a = \int_{-\infty}^{-\varepsilon/\tau} \tilde{\Phi}(z)^{\frac{n}{2}-1}\tilde{\phi}(z)\left[1-\tilde{\Phi}\left(-\frac{\varepsilon}{\tau(1-\alpha)}-\frac{\alpha z}{1-\alpha}\right)\right]^{n/2} dz,$$

$$b = \int_{-\infty}^{\varepsilon/\tau} \tilde{\Phi}(z)^{\frac{n}{2}-1}\tilde{\phi}(z)\left[1-\tilde{\Phi}\left(\frac{\varepsilon}{\tau(1-\alpha)}-\frac{\alpha z}{1-\alpha}\right)\right]^{n/2} dz,$$

$$c = \int_{-\varepsilon/\tau}^{\varepsilon/\tau} \tilde{\Phi}(z)^{\frac{n}{2}-1}\tilde{\phi}(z)\left[1-\tilde{\Phi}(z)\right]^{n/2} dz.$$

If $\tilde{\Phi} = \Phi$, the c.d.f. of a standard normal distribution, or $\tilde{\Phi} = \Phi_C$, the c.d.f. of a standard Cauchy distribution, then we can apply Corollary 3.4 to get an expression for $P_\theta(|\widetilde{X}_n(\alpha) - m| < \varepsilon)$ when n is even. Even though we suspect that (11) and (14) are still true even for an even sample size, we have not been able to prove our claim because of the complicated nature of the p.d.f. of the sample median when n is even. In Karunaratne (2004), Monte Carlo simulations show that for small n and small ε, $P_\theta(|\overline{X}_n - m| < \varepsilon)$ and $P_\theta(|\widetilde{X}_n - m| < \varepsilon)$ are very close to each other in the case of sampling from a normal distribution. This means that it is difficult to use numerical evidence to see which one is larger when n and ε are small and the data come from a symmetric distribution.

4. Conclusion

In this paper we carried out different comparisons to find out whether the sample mean or the sample median is a better estimator for the location parameter of a univariate distribution using Banks' criterion. In other words, we examined under what conditions the probability that the sample mean is within ε of the location parameter is greater than the probability that the sample median is within ε of the location parameter. Our results justify a conjecture by D. Banks (1997) that for heavy tail distributions the sample median is better than the sample mean when estimating a location parameter. The conlusion is reversed for distributions with lighter tails.

A. Some auxiliary results and a proof of a theorem

A.1. An auxiliary result for Theorem 2.2

Here we prove a result that is needed for the proof of Theorem 2.2. If $g(\alpha) > 0$ for large enough α, we write $f(\alpha) \sim g(\alpha)$ whenever $\lim_{\alpha\to\infty} \frac{f(\alpha)}{g(\alpha)} = 1$.

LEMMA A.1. *Let* $S(\alpha) = 2\ln\Gamma(2\alpha) - 4\ln\Gamma(\alpha) + 4(1-\alpha)\ln 2 - \ln(2\alpha+1)$ *for* $\alpha > 0$. *Then* $\lim_{\alpha\to\infty} S(\alpha) = \ln(2) - \ln(\pi)$.

Proof. Stirling's formula states that for real $x > 0$, we have

$$\Gamma(x) \sim \sqrt{2\pi}\, e^{-(x-1)}\, (x-1)^{x-1+\frac{1}{2}}$$

as $x \to \infty$. Therefore as $\alpha \to \infty$:

$$e^{S(\alpha)} \sim \frac{e^{-2}\,(2\alpha-1)^{4\alpha-1}\,2^{4(1-\alpha)}}{2\pi\,(\alpha-1)^{4\alpha-2}\,(2\alpha+1)} = \frac{2}{\pi}\cdot\frac{e^{-2}\left(1-\frac{1}{2\alpha}\right)^{4\alpha}\left(1-\frac{1}{2\alpha}\right)^{-1}}{\left(1+\frac{1}{2\alpha}\right)\left(1-\frac{1}{\alpha}\right)^{4\alpha}\left(1-\frac{1}{\alpha}\right)^{-2}}.$$

Since $\left(1+\frac{1}{\alpha}\right)^{\alpha} \to e$ as $\alpha \to \infty$, we have $e^{S(\alpha)} \sim \frac{2}{\pi}$, and the proof of the lemma is complete. □

A.2. An auxiliary result for Theorem 2.3

LEMMA A.2. *Let* $M(\alpha) = 2\ln\Gamma\left(\frac{\alpha+1}{2}\right) - 2\ln\Gamma\left(\frac{\alpha}{2}\right) - \ln(\alpha-2) - \ln(\frac{\pi}{4})$ *for* $\alpha > 2$. *Then* $\lim_{\alpha\to\infty} M(\alpha) = \ln(2) - \ln(\pi)$.

Proof. Using Stirling's formula, we obtain for $\alpha > 2$:

$$\begin{aligned} e^{M(\alpha)} &= \frac{\Gamma^2\left(\frac{\alpha+1}{2}\right)}{\Gamma^2\left(\frac{\alpha}{2}\right)(\alpha-2)\left(\frac{\pi}{4}\right)} \\ &\sim \frac{4(2\pi)\,e^{-2(\frac{\alpha+1}{2}-1)}\,(\frac{\alpha+1}{2}-1)^{2(\frac{\alpha+1}{2}-\frac{1}{2})}}{(2\pi)\,e^{-2(\frac{\alpha}{2}-1)}\,(\frac{\alpha}{2}-1)^{2(\frac{\alpha}{2}-\frac{1}{2})}\,(\alpha-2)\pi} \\ &\sim \frac{2\,e^{-1}\,(\alpha-1)^{\alpha}}{(\alpha-2)^{\alpha}\,\pi}. \end{aligned}$$

(The symbol $\sim$ is defined in Appendix A.1). We need to show $e^{M(\alpha)} \sim \frac{2}{\pi}$ as $\alpha \to \infty$. This follows from the fact that $\left(1+\frac{1}{\alpha}\right)^{\alpha} \to e$ as $\alpha \to \infty$. □

A.3. Auxiliary results for Theorems 3.1 and 3.2

LEMMA A.3. *For* $a > 0$,

$$\frac{2a}{\Phi'(a)} - \frac{1}{1-\Phi(a)} + \frac{1}{\Phi(a)} > 0, \tag{19}$$

where Φ *is the c.d.f. of a standard normal distribution.*

Proof. It is not difficult to show that (19) is equivalent to

$$q(a) := \Phi(a) - \left[\frac{1}{2} - \frac{\Phi'(a)}{2a} + \sqrt{\frac{1}{4} + \left(\frac{\Phi'(a)}{2a}\right)^2}\right] < 0.$$

We have:

$$q'(a) = \frac{a^2+1+(a^2-1)\sqrt{1+2\pi a^2\exp(a^2)}}{2a^2\sqrt{2\pi}\exp(a^2/2)\sqrt{1+2\pi a^2\exp(a^2)}}.$$

If $a \geqslant 1$, then obviously $q'(a) > 0$. Assume $0 < a < 1$.

Let

$$r(x) = \left(\frac{1+x}{1-x}\right)^2 - (1 + 2\pi x e^x).$$

Using Calculus, it can be shown that $r(x) = 0$ for exactly one $x \in (0,1)$, namely $x = x_0 \approx .3201$; $r(x) < 0$ for $0 < x < x_0$; and $r(x) > 0$ for $x_0 < x < 1$. Also for $0 < a < 1$, $q'(a) > 0$ if and only if $r(a^2) > 0$, and $q'(a) < 0$ if and only if $r(a^2) < 0$. Thus $q(a)$ strictly decreases for $0 < a < \sqrt{x_0}$, has a (local) minimum at $a = \sqrt{x_0} \approx .5658$, and strictly increases for $a > \sqrt{x_0}$. Since also $\lim_{a\to 0^+} q(a) = 0 = \lim_{a\to\infty} q(a)$, it follows that $q(a) < 0$ for $0 < a < \infty$. The proof of the lemma is complete. □

In the next lemma, the right inequality is similar to inequality (6) in the proof of Theorem 2.2, but unfortunately neither one implies the other. (Inequality (6) is defined for $\alpha > 0$, while the following lemma is for $\beta > 1$.)

LEMMA A.4. *For all* $\beta > 1$,

$$1 < \frac{\Gamma(2\beta)}{\Gamma(\beta)^2}\left(\frac{1}{4}\right)^{\beta-1} < \sqrt{2\beta - 1}.$$

Proof. (a) To prove the right inequality, we need to show that for $\beta > 1$,

$$S(\beta) = \ln\Gamma(2\beta) - 2\ln\Gamma(\beta) + 2(1-\beta)\ln 2 - \frac{1}{2}\ln(2\beta - 1) < 0.$$

As in the proof of Theorem 2.2, we can show that

$$\frac{dS}{d\beta} = T(\beta) = \psi\left(\beta + \frac{1}{2}\right) - \psi(\beta) - \frac{1}{2\beta - 1}.$$

We can then easily prove that for $\beta > 1$:

$$T(\beta + 1) - T(\beta) = \frac{1}{\beta(2\beta+1)(2\beta-1)} > 0.$$

Thus $T(\beta) < T(\beta+1) < \ldots < T(\beta+\nu)$ for all positive integers ν. In a way similar to the proof of Theorem 2.2 we can show that $\lim_{\beta\to\infty} T(\beta) = 0$, and so $\lim_{\nu\to\infty} T(\beta + \nu) = 0$ for each $\beta > 1$. Therefore $\frac{dS(\beta)}{d\beta} = T(\beta) < 0$ for all $\beta > 1$. Thus $S(\cdot)$ is strictly decreasing in $(1,\infty)$, and so $S(\beta) < S(1^+) = 0$.

(b) To prove the left inequality, we need to show that

$$K(\beta) = \ln\Gamma(2\beta) - 2\ln\Gamma(\beta) + 2(1-\beta)\ln 2 > 0.$$

As in the proof of Theorem 2.2, we can then show that:

$$\frac{dK}{d\beta} = \psi\left(\beta + \frac{1}{2}\right) - \psi(\beta) > 0.$$

This means that $K(\cdot)$ is strictly increasing for $\beta > 1$, so $K(\beta) > K(1^+) = 0$, and the lemma has been proven. □

A.4. Proof of Theorem 3.3

In this section we prove Theorem 3.3. The joint p.d.f. of the ordered statistics $X_{n(n/2)}$ and $X_{n((n/2)+1)}$ is

$$g(y_1,y_2) = \frac{\Gamma(n+1)}{\Gamma\left(\frac{n}{2}\right)^2}\,[F_\theta(y_1)]^{\frac{n}{2}-1}\,[1-F_\theta(y_2)]^{\frac{n}{2}-1}\,f_\theta(y_1)f_\theta(y_2), \quad y_1 < y_2,$$

and zero otherwise (see Hogg and Craig (1995), Section 4.6). It follows that the p.d.f. of $\widetilde{X}_n(\alpha) = \alpha X_{n(n/2)} + (1-\alpha)X_{n((n/2)+1)}$ is

$$\tilde{g}(w) = \int_{-\infty}^{w} h(w,y)\,dy, \quad -\infty < w < \infty,$$

where

$$\begin{aligned} h(w,y) = {} & \frac{\Gamma(n+1)}{\Gamma\left(\frac{n}{2}\right)^2(1-\alpha)}\,[F_\theta(y)]^{\frac{n}{2}-1}\left[1-F_\theta\left(\frac{w}{1-\alpha}-\frac{\alpha y}{1-\alpha}\right)\right]^{\frac{n}{2}-1} \\ & \times f_\theta(y) f_\theta\left(\frac{w}{1-\alpha}-\frac{\alpha y}{1-\alpha}\right). \end{aligned}$$

It follows that, for $\varepsilon > 0$ and $\theta \in \Theta$,

$$P_\theta(|\widetilde{X}_n(\alpha)-m| < \varepsilon) = \int_{m-\varepsilon}^{m+\varepsilon}\int_{-\infty}^{w} h(w,y)\,dy\,dw.$$

Changing the order of integration, we get for $\varepsilon > 0$ and $\theta \in \Theta$:

$$P_\theta(|\widetilde{X}_n(\alpha)-m| < \varepsilon) = \int_{-\infty}^{m-\varepsilon}\int_{m-\varepsilon}^{m+\varepsilon} h(w,y)\,dw\,dy + \int_{m-\varepsilon}^{m+\varepsilon}\int_{y}^{m+\varepsilon} h(w,y)\,dw\,dy.$$

Evaluating the inner integrals by using

$$\frac{d\left[1-F_\theta\left(\frac{w}{1-\alpha}-\frac{\alpha y}{1-\alpha}\right)\right]}{dw} = -\frac{1}{1-\alpha} f_\theta\left(\frac{w}{1-\alpha}-\frac{\alpha y}{1-\alpha}\right),$$

we can easily prove equation (17). □

Acknowledgements. We would like to thank Roger W. Barnard for his help in proving Theorem 2.2, and Frits Ruymgaart and Kamal C. Chanda for their advice on Mathematical Statistics. We also like to thank an anonymous scientist who pointed out an error in the proof of Theorem 3.1 in an earlier draft of the paper; an anonymous referee whose useful suggestions greatly improved the presentation of the paper; and the Managing Editor for suggesting Corollary 2.4 and for suggesting a simplification of the proof of Theorem 2.3.

REFERENCES

[1] M. ABRAMOWITZ AND I. A. STEGUN, *Handbook of mathematical functions with formulas, graphs and mathematical tables*, National Bureau of Standards, Applied Mathematics Series (55), 1964.

[2] D. BANKS, *Some geometry for location estimators*, Classification Society of North American Newsletter, Issue #49 (April 1997); available at `http://www.public.iastate.edu/~larsen/csna/previous/csnanews49-97apr.htm`.

[3] J. O. BERGER, *Statistical decision theory and Bayesian analysis*, 2nd ed., Springer-Verlag, New York, 1985.

[4] P. BILLINGSLEY, *Probability and measure*, 2nd ed., John Wiley, New York, 1986.

[5] M. H. DEGROOT, *Probability and Statistics*, 2^{nd} ed., Addison Wesley, Reading, Massachusetts, 1986.

[6] N. ELEZOVIĆ, C. GIORDANO AND J. PEČARIĆ, *The bests bounds in Gautschi's inequality*, Math. Inequal. Appl. **3** (2000), 239-252.

[7] R.V. HOGG AND A. T. CRAIG, *Introduction to mathematical Statistics*, 5^{th} ed., Prentice Hall, Upper Saddle River, NJ, 1995.

[8] H. S. I. KARUNARATNE, *Comparison of location estimators using Banks' criterion*, Master Thesis, Department of Mathematics and Statistics, Texas Tech University, Lubbock, TX, USA, 2004; available at `http://etd.lib.ttu.edu/theses/available/etd-07312008-31295019601797/unrestricted/31295019601797.pdf`.

[9] J. P. KEATING, R. L. MASON AND P.K. SEN, *Pitman's measure of closeness: A comparison of statistical estimators*, SIAM, Philadelphia, 1993.

[10] R. KHATTREE, *On comparison of estimates of dispersion using generalized Pitman nearness criterion*, Comm. Statist. Theory Methods **16** (1987), 263-274.

[11] R. KHATTREE AND S.D. PEDDADA, *A short note on Pitman nearness for elliptically symmetric estimators*, J. Statist. Plann. Inference **16** (1987), 257-260.

[12] T. K. NAYAK, *Estimation of location and scale parameters using generalized Pitman nearness criterion*, J. Statist. Plann. Inference **24** (1990), 259-268.

[13] S. D. PEDDADA, *A short note on Pitman's measure of nearness*, Amer. Statist. **39** (1985), 298-299.

[14] E. J. G. PITMAN, *The 'closest' estimates of statistical parameters*, Proc. Cambridge Philos. Soc. **33** (1937), 212-222.

[15] F. QI, B.-N. GUO AND C.-P. CHEN, *The best bounds in Gautschi-Kershaw inequalities*, Math. Inequal. Appl. **9** (2006), 427-436.

[16] C. R. RAO, J. P. KEATING AND R. L. MASON, *The Pitman nearness criterion and its determination*, Comm. Statist. A – Theory Methods **15** (1986), 3173-3191.

[17] M. J. SCHERVISH, *Theory of Statistics*, Springer-Verlag, New York, 1995.

[18] A. W. VAN DER VAART, *Asymptotic Statistics*, Cambridge University Press, Cambridge, UK, 1998.

(Received December 20, 2006)

H. Susitha I. Karunaratne
Department of Mathematics, Statistics and Physics
Purdue University North Central
1401 South U.S. 421
Westville, IN 46391, USA

Petros Hadjicostas
Department of Mathematics and Statistics
Texas Tech University, Box 41042
Lubbock, TX 79409-1042
USA
e-mail: petros.hadjicostas@ttu.edu