The Annals of Statistics
1985, Vol. 13, No. 3, 1114–1128

ESTIMATING A QUANTILE OF A SYMMETRIC DISTRIBUTION

BY ARTHUR COHEN,[1] SHAW-HWA LO, AND KESAR SINGH[2]

Rutgers University

The problem is to estimate a quantile of a symmetric distribution. The cases of known and unknown center are studied for small and large samples. The estimators for known center are the sample quantile, the symmetrized sample quantile, the sample quantile from flipped over data, the Rao-Blackwellized sample quantile, and a Bayes estimator using a Dirichlet prior. For center unknown, we study the analogues of the first four estimators listed above. For small samples and center known, the Rao-Blackwellized sample quantile performs very well for normal and double exponential distributions while for the Cauchy distribution the flipped over estimator did well. In the center known case, the latter four estimators are asymptotically equivalent, asymptotically optimal in the sense of Hajek's convolution, and asymptotically minimax in the Hajek-LeCam sense. For center unknown, those properties remain true if one uses an adaptive estimator of the center for the symmetrized sample quantile, the flipped over estimator, and the Rao-Blackwell estimator.

1. Introduction and summary.

We study the problem of estimating quantiles of a symmetric distribution. Sometimes we assume the center of symmetry is known and sometimes we do not assume the center of symmetry is known. Both situations seem to arise often enough in the real world. Dalal (1979) cites the case of a properly adjusted chemical balance as a situation where observations can be safely assumed to have a symmetric distribution with the true weight being the known center of symmetry. Quality control data oftentimes entails repeated measurements with instruments. Frequently these data collected regularly (weekly say) over many years reflect symmetry.

The following is a further example of a case where symmetry with known center is a reasonable assumption. The example is a prototype of situations encountered in the pharmaceutical industry and other businesses.

A distributor of small symmetrically shaped buttons sells them in packages containing a minimum number. The minimum number is large so he would not want to actually count the number in every package formed. Instead he counts out the minimum number (plus some extra) and puts them in a cannister and he observes a marker on the cannister where the buttons reach. (This is analogous to measuring cups of flour.) These buttons form the first package. For future packages he simply fills the cannister to the same marked height. In examples of this type, oftentimes the number of buttons in a package as a random variable

Received April 1984; revised December 1984.

[1] Research supported by NSF-MCS-81-18564-01.

[2] Research supported by NSF-MCS-83-01082.

AMS 1980 *subject classifications.* Primary 62G05; secondary 62F12.

Key words and phrases. Quantile, symmetry, Bahadur representation of quantiles, Rao-Blackwell estimator, Bayes estimator, Dirichlet prior, Hajek's convolution theorems, Hajek-LeCam minimaxity, bootstrap.

has a symmetric distribution about the known counted number in the first package. The symmetry assumption here would be even more compelling if it were validated from past or current data derived under similar conditions.

Statistical tests for symmetry with known center and unknown center exist in the literature. Hollander and Wolfe (1973) discuss distribution-free tests of symmetry and cite four references. Gastwrith (1971) offers an additional test for symmetry. Many researchers have studied inference problems under the symmetry model. For example, see Hannum and Hollander (1983), Schuster (1973), and Hodges and Lehmann (1956).

In Huber (1981), page 95, it is stated that most robustness literature adopts the assumption of symmetry. Huber points out however that such an assumption "violates the very spirit of robustness." Furthermore the assumption should not be made if the model distribution is asymmetric. The model of this paper is appropriate when one believes that it is reasonable to assume a symmetric distribution based on empirical evidence of the past, based on the physical considerations under which the data is collected, or for both of the above reasons. Intuitively the assumption of symmetry leads to substantially improved estimators.

Let $X_1, X_2, \cdots, X_n$ be a random sample of size n from a distribution with c.d.f. F. Let ξ_λ denote the λth quantile, with $.5 < \lambda < 1$, which is uniquely defined under the conditions imposed on F later.

We assume F is symmetric with θ as the center of symmetry. In the case of known center, $\theta = \theta_0$. The order statistics are $X_{(1)} \le X_{(2)} \le \cdots \le X_{(n)}$, and $Y_{(1)} \le \cdots \le Y_{(n)}$ represent the ordered values of $Y_j = |X_j - \theta_0|$, $j = 1, 2, \cdots, n$. When the center of symmetry is assumed known, we study the following five estimators (in the known center case, without loss of generality $\theta_0 = 0$): $T_1(\lambda)$ the sample λth quantile which we define as

$$T_1(\lambda) = \alpha X_{([n\lambda]+1)} + (1 - \alpha) X_{([n\lambda])}, \tag{1.1}$$

for $\alpha = n\lambda - [n\lambda] + 1/2$ and $[a]$ is the largest integer less than or equal to a; the quantity $T_1^+ = (T_1 - \theta_0 \vee 0)$.

$$T_2(\lambda) = (T_1^+(\lambda) - T_1^-(1 - \lambda))/2, \tag{1.2}$$

where $T_1^- = (T_1 - \theta_0 \wedge 0)$. $T_3(\lambda)$ is the $(2\lambda - 1)$th sample quantile using the Y sample. We call $T_3(\lambda)$ the flipped-over estimator. $T_4(\lambda)$ is the Rao-Blackwellized version of T_1^+ and it depends on the Y sample. (In Section 2 we give $T_4(\lambda)$ explicitly.) $T_5(\lambda)$ is a nonparametric Bayes estimator using a Dirichlet prior and a symmetric linear loss function.

Estimators T_2 and T_3 are intuitive and obvious choices as competitors to T_1 under the symmetry assumption. The choice of T_4 is based on the fact that the Rao-Blackwellized version of T_1^+ must be superior in mean square error to T_1^+ for any possible symmetric distribution and for all sample sizes. T_5 is studied out of interest as a nonparametric Bayes estimator for a prior that should not be too informative. Typically such nonparametric estimators have been shown to be reasonable. See, for example, Ferguson (1973), Section 5.

We verify that estimators T_2, T_3, T_4 and T_5 differ by $O(n^{-3/4}\log n)$ a.s., and hence all have the same asymptotic variance. The proof of the first-order asymptotic equivalence of T_5 and T_3 does not require the symmetry model of the rest of the paper and so this gives a result of independent interest. That is, the Bayes estimator for a Dirichlet prior for the symmetric absolute error loss function is asymptotically equivalent to the sample quantile.

For all sample sizes, it is clear that the mean square error for T_4 is smaller than the mean square error for T_1^+ and T_2. Computer work for finite samples indicates superiority of T_4 in the normal and exponential cases. For the Cauchy distributions, however, T_3, appears best.

For the case where the center of symmetry is unknown, analogues of T_1, T_2, T_3 and T_4 are suitably defined and represented asymptotically as a sum of a location component and a shape component. These representations help one visualize the asymptotic nature of these estimators. The problem of asymptotic optimality and asymptotic minimaxity is studied in both cases of center of symmetry known and unknown.

In the next section we give the estimators and their asymptotic variances for the cases where the center of symmetry is known and unknown. In Section 3 we show that T_2, T_3, T_4 and T_5 are asymptotically equivalent. Section 4 is devoted to asymptotic representations of the estimators when the center of symmetry is unknown. Asymptotic optimality is discussed in Section 5.

2. Estimators and their asymptotic distributions. Throughout all sections we will assume that F is twice differentiable in a neighborhood of ξ_λ; that $F'(x) = f(x)$ is bounded away from 0 in the neighborhood and that $|F''|$ is bounded in the neighborhood.

Estimators T_1, T_1^+ and T_2 were given explicitly in Section 1. The estimator $T_3(\lambda)$ may be expressed as

$$T_3(\lambda) = \alpha Y_{([n(2\lambda-1)]+1)} + (1-\alpha) Y_{([n(2\lambda-1)])}, \tag{2.1}$$

where $\alpha = n(2\lambda - 1) - [n(2\lambda - 1)] + ½$. In order to express $T_4(\lambda)$ explicitly, note that for any integer r, $1 \le r \le n$,

$$E(X_{(r)} \vee 0 \mid Y_{(1)}, Y_{(2)}, \cdots, Y_{(n)}) = (½)^n \textstyle\sum_{j=1}^{r} 2^{j-1}\binom{n-j}{r-j} Y_{(j)}. \tag{2.2}$$

If the α in expression (1.1) is 0 or 1, (2.2) suffices to define $T_4(\lambda)$ explicitly. Also in studying $T_4(\lambda)$ asymptotically, it suffices to look at (2.2) with $r = [n\lambda]$ and $[n\lambda] + 1$. See Lemma 3.1. If the α in expression (2.2) is not 0 or 1 then, although we cannot write one explicit expression for $T_4(\lambda)$ in all cases, the idea used in deriving (2.2) can be used in particular cases to derive

$$E\{[\alpha X_{([n\lambda]+1)} + (1-\alpha) X_{([n\lambda])}] \vee 0 \mid Y_{(1)}, Y_{(2)}, \cdots, Y_{(n)}\}.$$

The Bayes estimator $T_5(\lambda)$ is derived for the loss function $L(a, \xi_\lambda) = |a - \xi_\lambda|$ and for a Dirichlet prior whose support is the class of distributions symmetric about 0. For details on such priors, see Dalal (1979). The Bayes estimator is any median of the posterior distribution of ξ_λ. It follows from Ferguson (1973), page 225, and Dalal (1979) that the Bayes estimator is the $u(\frac{1}{2}, 2\lambda - 1, \alpha(R) + n)$th quantile of $\hat{F}_n$, where $\hat{F}_n(t) = (M/(M+n))F^*(t) + (1 - M/(M+n))F_{n,0}(t)$, where $F_{n,0}(t)$ is the empirical c.d.f. symmetrized around 0, $F^*(t) = \alpha(-\infty, t)/M$, $M = \alpha(R)$, α is the parameter of the Dirichlet prior, and u is the solution of the equation:

$$\int_{2\lambda-1}^{1} \Gamma(M)/[\Gamma(uM)\Gamma((1-u)M)]z^{uM-1}(1-z)^{(1-u)M-1}\,dz = \frac{1}{2}. \tag{2.3}$$

It follows from the central limit theorem for quantiles that $T_1(\lambda)$ is asymptotically normal with mean ξ_λ and variance $\lambda(1-\lambda)/nf^2(\xi_\lambda)$. (See, for example, Serfling, 1980, page 80.) The multivariate central limit theorem for quantiles can be used to prove that $T_2(\lambda)$ is asymptotically normal with mean ξ_λ and variance $(2\lambda-1)(1-\lambda)/2nf^2(\xi_\lambda)$. (Again see Serfling, 1980, page 80.) Asymptotically T_2, T_3, T_4 and T_5 are equivalent (see Section 3) and so the variances of the asymptotic distributions of T_3, T_4 and T_5 are also $(2\lambda-1)(1-\lambda)/2nf^2(\xi_\lambda)$. The ratio of asymptotic variances for T_1 and T_2 is $e = 2\lambda/(2\lambda - 1)$ which is a measure of the improvement in estimating a quantile under the symmetry assumption. Clearly $2 \le e \le \infty$ since $\frac{1}{2} \le \lambda \le 1$.

We next determine the analogue of T_2, T_3 and T_4 when the center of symmetry is unknown. Let θ denote the center of symmetry and let $\hat{\theta}$ be an estimator of it. The analogue of T_2 is $T_2(\hat{\theta}, \lambda) = T_2^*(\lambda) + \hat{\theta}$, where $T_2^*(\lambda)$ is the analogue of $T_2(\lambda)$ but now θ_0 is replaced by $\hat{\theta}$. The analogue of T_3 is

$$T_3(\hat{\theta}, \lambda) = \alpha Z_{([n(2\lambda-1)]+1)}(\hat{\theta}) + (1-\alpha)Z_{([n(2\lambda-1)])}(\hat{\theta}), \tag{2.4}$$

where $Z_i(\hat{\theta}) = X_i$ if $X_i > \hat{\theta}$ and $Z_i(\hat{\theta}) = 2\hat{\theta} - X_i$ otherwise; and $Z_{(i)}(\hat{\theta})$ are the ordered Z's. Similarly the analogue of $T_4(\lambda)$ namely $T_4(\hat{\theta}, \lambda)$ is obtained as $T_4(\lambda)$ except that the $Y_{(i)}$ are replaced by $Z_{(i)}(\hat{\theta})$. The asymptotic variances of these analogues are discussed in Sections 4 and 5.

REMARK 2.1. Suppose we define the λth population quantiles $\xi_\lambda(F) = \inf\{x: F(x) \ge \lambda\}$, where F lies in the class of all distributions symmetric about 0. Then the nonparametric maximum likelihood estimator (m.l.e.) of ξ_λ is $\inf\{x: F_{n,0}(x) \ge \lambda\}$ where $F_{n,0}(\cdot)$ is the symmetrized empirical c.d.f. around 0. It is clear that the m.l.e. is virtually the same as $T_3(\lambda)$.

3. Asymptotic equivalence of T_2, T_3, T_4 and T_5. In this section we will repeatedly use Bahadur's asymptotic representation of quantiles. See Bahadur (1966) or Serfling (1980), page 91.

THEOREM 3.1. *With probability one,* $|T_2 - T_3| = O(n^{-3/4}(\log n)^{3/4})$, $n \to \infty$.

PROOF. A standard argument shows that $P(\{T_1 = T_1^+\} \cup \{T_1 = T_1^-\})$ decays exponentially fast. In view of this, use Bahadur's representation of T_1 and note that with probability one

$$
(3.1)\quad
\begin{aligned}
T_2-\xi_\lambda &= [(T_1(\lambda)-\xi_\lambda)/2]-[(T_1(1-\lambda)-\xi_{1-\lambda})/2]+R_n\\
&= \{[\lambda-F_n(\xi_\lambda)]/2f(\xi_\lambda)\}+\{[F_n(\xi_{1-\lambda})-(1-\lambda)]/2f(\xi_{1-\lambda})\}+R_n\\
&= \{(2\lambda-1)-[F_n(\xi_\lambda)-F_n(\xi_{1-\lambda})]\}/2f(\xi_\lambda)+R_n,
\end{aligned}
$$

where F_n is the empirical c.d.f. of X_i's and $R_n = O(n^{-3/4}(\log n)^{3/4})$, a.s. as $n\to\infty$. Let f_0, F_0, denote the density and the c.d.f. for the distribution of $Y = |X|$, and let $F_{n,0}$ be the empirical c.d.f. of Y_i's ($Y_i = |X_i|$). Let ξ_t^* denote the tth quantile of F_0. Then by Bahadur representation with probability one

$$
(3.2)\quad T_3-\xi_\lambda = ((2\lambda-1)-F_{n,0}(\xi^*_{2\lambda-1}))/f(\xi^*_{2\lambda-1})+R_n \quad \text{as} \quad n\to\infty.
$$

The theorem follows now from (3.1) and (3.2) recognizing that $f(\xi^*_{2\lambda-1}) = 2f(\xi_\lambda)$ and $F_{n,0}(\xi^*_{2\lambda-1}) = F_n(\xi_\lambda)-F_n(\xi_{1-\lambda})+O(1/n)$. □

In Theorem 3.5 below we show the asymptotic equivalence of T_3 and $E(X_{([n\lambda])}\mid Y_{(1)}, Y_{(2)}, \cdots, Y_{(n)}) = \sum_{k=1}^r c_k Y_{(k)}$ (see (2.2)). The same proof yields the asymptotic equivalence of T_3 and $E(X_{([n\lambda]+1)}\mid Y_{(1)}, \cdots, Y_{(n)})$. Then the equivalence of T_3 and T_4 will follow from

LEMMA 3.2. *Let* $\mathbf{Y} = (Y_{(1)}, Y_{(2)}, \cdots, Y_{(n)})$. *Then*

$$
\begin{aligned}
&E\{(\alpha X_{([n\lambda]+1)}+(1-\alpha)X_{([n\lambda])})\vee 0\mid \mathbf{Y}\}\\
&\qquad = \alpha E(X_{([n\lambda]+1)}\vee 0\mid \mathbf{Y})+(1-\alpha)E(X_{([n\lambda])}\vee 0\mid\mathbf{Y})+\gamma_n,
\end{aligned}
$$

where $|\gamma_n|\le(1-\alpha)p_nY_{([n(1-\lambda)]+1)}$ *and* $p_n = \binom{n}{[n\lambda]}2^{-n}$.

PROOF. Note that $((1-\alpha)X_{([n\lambda])}+\alpha X_{([n\lambda]+1)})\vee 0$ differs from $[(1-\alpha)X_{([n\lambda])}\vee 0]+[\alpha X_{([n\lambda])}\vee 0]$ only when $X_{([n\lambda])}<0$ and $X_{([n\lambda]+1)}>0$. The conditional probability for this is $\binom{n}{[n\lambda]}2^{-n}$. When this does happen the difference between the two cannot exceed $(1-\alpha)Y_{([n(1-\lambda)]+1)}$. □

Note that p_n in Lemma 3.2 goes to zero exponentially fast. To prove Theorem 3.5 we need

LEMMA 3.3. *Let k be such that* $|k-(2r-n)|\ge c(n(\log n))^{1/2}$, *for a constant c to be determined. Then* $c_k = O(n^{-3})$.

PROOF. Use the version of Stirling's formula as given in Feller (1968), page 54, formula (9.15) to find that

$$
(3.3)\quad
\begin{aligned}
c_k &\le (k/\sqrt{2\pi})2^{-(n-k+1)}(n-k)^{(n-k+1/2)}(r-k)^{-(r-k+1/2)}(n-r)^{-(n-r+1/2)}\\
&\le Kn^{1/2}2^{-(n-k)}\exp[(n-k)\log(n-k)-(r-k)\log(r-k)-(n-r)\log(n-r)]
\end{aligned}
$$

where K is a constant. Let $s = k/n$, $\lambda = r/n$, rewrite the right-hand side of (3.3) in terms of s and t and it becomes

$$Kn^{1/2}\exp[-n(1-s)\log 2 + n(1-s)\log n(1-s) - n(\lambda - s)\log n(\lambda - s) - n(1-\lambda)\log n(1-\lambda)]. \tag{3.4}$$

The derivative of the exponent in (3.4) with respect to s is $n \log 2(1 - (1-\lambda)/(1-s))$. Hence the derivative is strictly decreasing as s varies from 0 to λ. It is positive for $s < 2\lambda - 1$, 0 for $s = 2\lambda - 1$ and negative for $s > 2\lambda - 1$. Thus it suffices to set

$$s = (2\lambda - 1) \pm c(\log n)^{1/2}/n^{1/2} \tag{3.5}$$

in (3.4) and verify that (3.4) $= O(n^{-3})$.

Using (3.5), (3.4) becomes

$$\begin{aligned} &Kn^{1/2}2^{-2n(1-\lambda)}2^{\pm nc(\log n/n)^{1/2}} \\ &\quad\cdot \exp[(2n(1-\lambda) \mp nc(\log n/n)^{1/2})\log(2n(1-\lambda) \mp nc(\log n/n)^{1/2}) \\ &\qquad -(n(1-\lambda) \mp nc(\log n/n)^{1/2})\log(n(1-\lambda) \mp nc(\log n/n)^{1/2}) \\ &\qquad - n(1-\lambda)\log n(1-\lambda)]. \end{aligned} \tag{3.6}$$

In (3.6) write $\log(2n(1-\lambda) \mp nc(\log n/n)^{1/2})$ as $\log 2n(1-\lambda) + \log(1 - c(\log n/n)^{1/2}/2(1-\lambda))$ and other terms similarly. Then we use the fact that $\log(1 + z) = z - z^2/2 + e_n$ with $z = c(\log n/n)^{1/2}/2(1-\lambda)$ and $e_n = O(\log n/n)$. After some algebra (3.6) becomes

$$Kn^{1/2}\exp\{(-c^2\log n/4(1-\lambda)) + o(1)\}. \tag{3.7}$$

Choose $c = (10(1-\lambda))^{1/2}$ and the lemma holds. □

We also need

LEMMA 3.4. *Let $\sum^* c_k$ denote the sum of c_k over all k such that $|k - (2r - n)| \le c(n \log n)^{1/2}$, where c is as in Lemma 3.2. Let $x_k = (k - (2r - n))/n^{1/2}$. Then*

$$\textstyle\sum^* c_k = 1 + O((\log n)^{1/2}/n) \tag{3.8}$$

$$\textstyle\sum^* x_k c_k = O((\log n)^{1/2}/\sqrt{n}). \tag{3.9}$$

PROOF. Use Stirling's formula again and proceed as in the proof of Lemma 3.2 to find that, for $|k - (2r - n)| \le c(n \log n)^{1/2}$,

$$\begin{aligned} c_k &= \{[1 + O((\log n)^{1/2}/n)]/n^{1/2}(4\pi(1-\lambda))^{1/2}\} \\ &\quad \cdot \exp[-(x_k^2/4(1-\lambda)) + (\tfrac14\sqrt{n}(1-\lambda))x_k - (\tfrac14(1-\lambda)^2\sqrt{n})x_k^3] \\ &= [1 + O((\log n)^3/n)][1 + (\tfrac14\sqrt{n}(1-\lambda))x_k - (\tfrac14\sqrt{n}(1-\lambda)^2)x_k^3] \\ &\quad \cdot \exp^{(-x_k^2}/4(1-\lambda)). \end{aligned} \tag{3.10}$$

Now (3.10) and some calculations reveal that $\sum^* c_k$ as a Riemann sum over subintervals of width $(1/\sqrt{n})$ approximates

$$(1/\sqrt{2\pi\, 2(1-\lambda)})\int_{-a_n}^{a_n} e^{-x^2/4(1-\lambda)}[1 + x/4\sqrt{n}(1-\lambda) - x^3/4\sqrt{n}(1-\lambda)^2]\,dx,$$

where $a_n = c(\log n)^{1/2}$, up to $O((\log n)^3/n)$. With the help of integration by parts, this latter integral is seen to be $1 + O(1/n)$. Similar arguments yield (3.9). □

THEOREM 3.5. *With probability one, as* $n \to \infty$ $|E(X_{([n\lambda])} \vee 0 \mid \mathbf{Y}) - Y_{(2n-[n\lambda])}| = O(n^{-3/4}\log n)$

PROOF. Consider

$$\begin{aligned}\sum{}^* c_k Y_{(k)} &= \sum{}^* c_k(Y_{(k)} - Y_{(2r-n)} + Y_{(2r-n)}) \\ &= \sum{}^* c_k(Y_{(k)} - Y_{(2r-n)} + Y_{(2r-n)}) \sum{}^* c_k.\end{aligned} \tag{3.11}$$

Lemmas 3.3 and 3.4 imply the conclusion of the theorem provided it can be shown that

$$\sum{}^* c_k(Y_{(k)} - Y_{(2r-n)}) = O(n^{-3/4}(\log n)^{3/4}), \tag{3.12}$$

with probability 1 as $n \to \infty$. We may rewrite (3.12) as the sum of

$$\sum{}^* c_k[F_{n,0}^{-1}(k/n) - F_0^{-1}(k/n) - F_{n,0}^{-1}((2r-n)/n) + F_0^{-1}((2r-n)/n)] \tag{3.13}$$

and

$$\sum{}^* c_k[F_0^{-1}(k/n) - F_0^{-1}((2r-n)/n)]. \tag{3.14}$$

To deal with (3.14) use Taylor's theorem for F_0 to find that (3.14) is

$$\sum{}^* c_k[x_k/\sqrt{n}f_0((2r-n)/n) + x_k^2/nf_0'(\nu)], \tag{3.15}$$

where ν is near $(2r-n)/n$. Since $|x_k| \le c(\log n/n)^{1/2}$ for those indices k in the $\sum^*$, it follows that (3.15) is $O((\log n)/n)$.

From the uniform version of quantile representation (see Kiefer, 1970), it follows that (3.13) is bounded above by

$$\begin{aligned}&K \sup_{|a| \le c(\log n/n)^{1/2}} |F_{n,0}(\xi_{2\lambda-1}^*) - F_0(\xi_{2\lambda-1}^*) \\ &\quad - F_{n,0}(\xi_{2\lambda-1}^* + a) + F_0(\xi_{2\lambda-1}^* + a)| + O(n^{-3/4}(\log n)).\end{aligned} \tag{3.16}$$

Now by Lemma 1 of Bahadur (1966) (3.16) is $O(n^{-3/4}\log n)$ with probability 1. □

Next we demonstrate the asymptotic equivalence of T_3 and T_5. To accomplish this we consider the more general problem of estimating a quantile of a distribution without imposing the symmetry condition. We will show that the Bayes estimator, say T_B, for symmetric absolute error loss and Dirichlet prior with parameter α is asymptotically equivalent to the sample quantile T_1 in the sense that $|T_B - T_1| = O(\log n/n)$, with probability one as $n \to \infty$. This same proof along with the definition of T_5 given in (2.3) of Section 2, will imply the asymptotic equivalence of T_3 and T_5 in the symmetric case. This result is of interest in its own right.

The model and notation is as in previous sections, only now we do not require that F is symmetric.

THEOREM 3.6. *The estimators T_B and T_1 are such that with probability one, $|T_B - T_1| = O((\log n)/n)$, $n \to \infty$.*

PROOF. It follows from Ferguson (1973), page 225, that T_B is a solution in t of

$$P([Z/(Z + W)] > \lambda) = \tfrac{1}{2}, \tag{3.17}$$

where Z and W are independent gamma random variables with shape parameters $[\alpha(-\infty, t) + nF_n(t)]$ and $[\alpha(t, \infty) + n(1 - F_n(t))]$, respectively. Represent Z as $\sum_{i=1}^n \nu_i$, where ν_i are i.i.d. gamma variables with parameter $[(\alpha(-\infty, t)/n) + F_n(t)]$. Similarly represent W as $\sum_{i=1}^n \nu_i'$, where ν_i' are i.i.d. gamma variables with parameters $[(\alpha(t, \infty)/n) + (1 - F_n(t))]$. Let $\eta_i = (1 - \lambda)\nu_i$ and $\eta_i' = -\lambda\nu'$. Then (3.17) may be written as

$$P(\textstyle\sum_{i=1}^n (\eta_i + \eta_i') > 0) = \tfrac{1}{2}, \tag{3.18}$$

and so the left-hand side of (3.18) can be written as

$$\begin{aligned} &P([\textstyle\sum_{i=1}^n (\eta_i + \eta_i') - E(\sum_{i=1}^n (\eta_i + \eta_i'))]/\sigma(\sum_{i=1}^n (\eta_i + \eta_i')) \\ &\qquad \geq -E([\textstyle\sum(\eta_i + \eta_i')]/\sigma(\sum(\eta_i + \eta_i'))), \end{aligned} \tag{3.19}$$

where $\sigma(\sum(\eta_i + \eta_i'))$ is the standard deviation of $(\sum (\eta_i + \eta_i'))$. We may apply a Berry-Esseen Theorem to (3.19). See Feller (1971), page 544, Theorem 2. In applying the theorem it is easily verified, using the moment properties of ν_i and ν_i', that the remainder term is $O(1/\sqrt{n})$. Thus from the Berry-Esseen Theorem we have that (3.19) is

$$\Phi(E(\textstyle\sum (\eta_i + \eta_i'))/\sigma(\sum (\eta_i + \eta_i'))) + O(1/\sqrt{n}), \tag{3.20}$$

as $n \to \infty$, uniformly in t. Hence from (3.17) we seek the solution of

$$E(\textstyle\sum (\eta_i + \eta_i')/\sigma(\sum (\eta_i + \eta_i'))) = O(1/\sqrt{n}). \tag{3.21}$$

It is easily verified that $\sigma(\sum (\eta_i + \eta_i)) = O(\sqrt{n})$ and so (3.21) reduces to

$$E(\textstyle\sum (\eta_i + \eta_i')) = O(1). \tag{3.22}$$

We may rewrite (3.22) as

$$(1 - \lambda)[(\alpha(-\infty, t)/n) + F_n(t)] - \lambda[(\alpha(t, \infty)/n) + (1 - F_n(t))] = O(1/n), \tag{3.23}$$

which yields the equation

$$F_n(t) = \lambda + O(1/n). \tag{3.24}$$

From (3.24) it is not hard to show that the solution in t, namely T_B, is such that $|T_B - T_1| = O((\log n)/n)$ with probability one. □

REMARK 3.7. For estimating a quantile, Ferguson (1973) considers the following class of loss functions:

$$L(\xi_\lambda, \hat{\xi}_\lambda) = p(\xi_\lambda - \hat{\xi}_\lambda) \quad \text{if} \quad \xi_\lambda \geq \hat{\xi}_\lambda$$
$$= (1-p)(\hat{\xi}_\lambda - \xi_\lambda) \quad \text{if} \quad \xi_\lambda \leq \hat{\xi}_\lambda. \tag{3.25}$$

For a loss function in this class, the Bayes estimator is a solution (in t) of the following equation:

$$P(\beta(a, b) > \lambda) = p,$$

where $\beta(a, b)$ is a Beta variable with $a = \alpha(-\infty, t) + [\text{number of } X_i \leq t]$ and $b = \alpha(t, \infty) + [\text{number of } X_i > t]$; α is the measure of the Dirichlet prior. The method of proof in Theorem 3.6 yields the following representation for the Bayes estimators:

$$\hat{\xi}_\lambda = F_n^{-1}(t) + g(p, \lambda)/\sqrt{n} + O(n^{-1}(\log n) \quad \text{a.s.}, \tag{3.26}$$

where $g(p, \lambda)$ is a nonzero real except for $p = ½$ when $g(½, \lambda) = 0$. Using (3.26) one notes that the limiting distribution of $\sqrt{n}(\hat{\xi}_\lambda - \xi_\lambda)$ is $N(g(p, \lambda), \lambda(1-\lambda)/f^2(\xi_\lambda))$. Hence the Bayes estimator is asymptotically inferior in a bias and mean square sense to $F_n^{-1}(\lambda)$ except when $p = ½$.

4. Asymptotic representation of estimators when center is unknown. Let $\hat{\theta}$ be an estimator of the center θ such that $|\hat{\theta} - \theta| = O(n^{-1/2}(\log n)^{1/2})$ a.s. Define $q(a, t)$ to be the tth population quantile when the population is flipped to the right around a. That is, the tth quantile of the distribution whose c.d.f. is

$$F_a(x) \begin{cases} = F(x) - F(2a - x) & \text{for} \quad x > a \\ = 0 & \text{otherwise.} \end{cases} \tag{4.1}$$

Also define $\hat{q}(a, t)$ to be the tth sample quantile corresponding to the data flipped to the right around a. So $\hat{q}(a, t)$ is the tth quantile of $F_{n,a}(x)$, which is the empirical c.d.f. of the data flipped to the right of a. We need the following two lemmas:

LEMMA 4.1. *The function $q(a, 2\lambda - 1)$ is partially differentiable with respect to a in a neighborhood of θ and*

$$\partial q(a, 2\lambda - 1)/\partial a\,|_{a=\theta} = 1,$$
$$\partial^2 q(a, 2\lambda - 1)/\partial^2 a\,| \quad \textit{is bounded in the neighborhood.} \tag{4.2}$$

PROOF. From (4.1), $q = q(a, 2\lambda - 1)$ is the solution of the equation

$$F(q) - F(2a - q) = (2\lambda - 1). \tag{4.3}$$

Differentiating both sides of (4.3) with respect to a yields

$$F'(q)(\partial q/\partial a) - F'(2a - q)(2 - (\partial q/\partial a)) = 0. \tag{4.4}$$

In (4.4) substitute θ for a, use the fact that $F'(q) = F'(2\theta - q)$ and solve for $(\partial q/\partial a)\,|_{a=\theta}$ to establish (4.2). The second part of the lemma follows by differentiating the left-hand side of (4.4) and using the fact that $F'(\cdot)$ is bounded away from zero in a neighborhood of ξ_λ. □

LEMMA 4.2. *There exists a neighborhood I of θ and a neighborhood J of $(2\lambda - 1)$ such that*

$$\begin{aligned}(4.5)\qquad &\sup_{a\in I, t\in J} |\hat{q}(a, t) - q(a, t) + (F_{n,a}(q(a, t)) - t)/F_a'(q(a, t))| \\ &\qquad = O(n^{-3/4}(\log n)^{3/4}) \quad a.s.\end{aligned}$$

PROOF. Make a choice of the neighborhoods I and J small enough such that for all $a \in I$ and $t \in J$, $F_a(x)$ is twice differentiable for $x \in [q(a, t) - \varepsilon, q(a, t) + \varepsilon]$, $\varepsilon > 0$, and on this interval $F_a'(\cdot)$ is bounded away from zero and $|F_a''|$ is bounded. That such a choice of intervals is possible is seen by doing a two-term Taylor expansion of $q(a, t)$ around $(\theta, 2\lambda - 1)$ and using the properties of F, F', F'' assumed throughout.

Let

$$(4.6)\qquad R_n(a, t) = \hat{q}(a, t) - q(a, t) + (F_n(q(a, t)) - t))/F_a'(q(a, t)).$$

Note that $\hat{q}(a, t)$ is nondecreasing in each argument. This fact and the Lipschitz property of $(F_n - F)$ (see Lemma 1 of Bahadur, 1966) imply

$$\begin{aligned}&\sup_{a\in I, t\in J} |R_n(a, t)| \\ &\quad = \max\{R_n(a_i, t_j)\colon a_i = \theta + in^{-3/4}(\log n)^{3/4}, \\ (4.7)\qquad &\qquad\qquad t_j = (2\lambda - 1) + jn^{-3/4}(\log n)^{3/4}, \\ &\qquad \text{over all integers } i \text{ and } j \text{ such that } a_i \in I \text{ and } t_j \in J\} \\ &\qquad + O(n^{-3/4}(\log n)^{3/4}) \quad \text{a.s.}\end{aligned}$$

For any fixed (i, j) such that $a_i \in I$, $t_j \in J$, one obtains the following probability bound by mimicking and slightly modifying the proofs in Bahadur (1966):

$$(4.8)\qquad P(|R_n(a_i, t_j)| > C_1 n^{-3/4}(\log n)^{3/4}) \le C_2 n^{-3},$$

where C_1, C_2 do not depend on (i, j). Now the proof of the lemma is concluded using Bonferroni's inequality and the Borel-Cantelli Lemma. □

THEOREM 4.3. *The difference*

$$(4.9)\qquad T_3(\hat{\theta}, \lambda) - T_3(\theta, \lambda) = (\hat{\theta} - \theta) + O(n^{-3/4}(\log n)^{3/4}) \quad a.s.$$

The significance of the theorem is that it yields the representation $T_3(\hat{\theta}, \lambda) - \xi_\lambda = (T_3(\theta, \lambda) - \xi_\lambda) + (\hat{\theta} - \theta) + O(n^{-3/4}(\log n)^{3/4})$. Typically $T_3(\theta, \lambda)$ and $(\hat{\theta} - \theta)$ are asymptotically uncorrelated. See Remark 4.6 below.

PROOF OF THEOREM 4.3. The following two statements yield the theorem:

$$\sup_{|\theta-a|\le cn^{-1/2}(\log n)^{1/2}} \cdot |\hat{q}(a, 2\lambda-1) - q(a, 2\lambda-1) - \hat{q}(\theta, 2\lambda-1) + q(\theta, 2\lambda-1)| = O(n^{-3/4}(\log n)^{3/4}) \quad \text{a.s.} \tag{4.10}$$

$$q(a, 2\lambda-1) = q(\theta, 2\lambda-1) + (a-\theta) + O((a-\theta)^2) \quad \text{as } |\theta-a| \to 0. \tag{4.11}$$

Clearly (4.11) follows from Lemma 4.1 and Taylor's expansion. Lemma 4.2 implies that the left-hand side of (4.10) cannot exceed

$$\begin{aligned} \sup_{|\theta-a|\le cn^{-1/2}(\log n)^{1/2}} |&(F_{n,a}(q(a, 2\lambda-1)) - (2\lambda-1))/F_a'(q(a, 2\lambda-1)) \\ &- (F_{n,a}(q(\theta, 2\lambda-1)) - (2\lambda-1))/F_a'(q(\theta, 2\lambda-1))| \\ &+ O(n^{-3/4}(\log n)^{3/4}) \quad \text{a.s.} \end{aligned} \tag{4.12}$$

Using Lemma 1 of Bahadur (1966), (4.12) is easily seen to be $O(n^{-3/4}(\log n)^{3/4})$; this establishes (4.10). □

THEOREM 4.4. *The difference*

$$T_4(\hat{\theta}, \lambda) - T_4(\theta, \lambda) = (\hat{\theta} - \theta) + O(n^{-3/4}(\log n)^{3/4}) \quad a.s. \tag{4.13}$$

PROOF. We may write $T_4(\hat{\theta}, \lambda) = \sum_{k=1}^{[n\lambda]} c_k Z_{(k)}(\hat{\theta})$, where $Z_{(k)}(\hat{\theta})$ are defined in Section 2. We use the properties of c_k from Lemmas 3.2 and 3.3. Since $c_k = O(n^{-3})$ for $|k - n(2\lambda-1)| \ge cn^{1/2}(\log n)^{1/2}$, we need only show

$$\textstyle\sum^* c_k Z_{(k)}(\hat{\theta}) = T_4(\theta, \lambda) = (\hat{\theta} - \theta) + O(n^{-3/4}(\log n)^{3/4}) \quad \text{a.s.}, \tag{4.14}$$

where $\sum^*$ runs over k such that $|k - (2\lambda-1)n| \le cn^{1/2}(\log n)^{1/2}$. Write

$$\textstyle\sum^* c_k Z_k(\hat{\theta}) = \sum^* Z_{([n(2\lambda-1)])}(\hat{\theta})\, c_k + \sum^* (Z_{(k)}(\theta) - Z_{([n(2\lambda-1)])}(\hat{\theta}))c_k. \tag{4.15}$$

From (3.8) of Lemma (3.4), the fact that $T_3(\hat{\theta}, \lambda)$ may be taken as $Z_{([n(2\lambda-1)])}(\hat{\theta})$ up to appropriate order, an application of Theorem 4.3 yields that (4.15) may be written as

$$\begin{aligned} &\hat{q}(\theta, 2\lambda-1) + (\hat{\theta} - \theta) \\ &\quad + \textstyle\sum^* [Z_{(k)}(\hat{\theta}) - Z_{([n(2\lambda-1)])}(\hat{\theta})]c_k + O(n^{-3/4}(\log n)^{3/4}) \quad \text{a.s.} \\ &= T_4(\theta, \lambda) + (\hat{\theta} - \theta) + Q_n + O(n^{-3/4}(\log n)^{3/4}) \quad \text{a.s.} \end{aligned} \tag{4.16}$$

where

$$Q_n = \textstyle\sum^* [Z_{(k)}(\hat{\theta}) - Z_{([n(2\lambda-1)])}(\hat{\theta})]c_k. \tag{4.17}$$

Now (4.17) is

$$\textstyle\sum^* [\hat{q}(\hat{\theta}, k/n) - \hat{q}(\hat{\theta}, 2\lambda-1)]c_k,$$

which may be expressed as the sum of

$$\sum{}^{*} [\hat{q}(\hat{\theta}, k/n) - q(\hat{\theta}, k/n) - \hat{q}(\hat{\theta}, 2\lambda - 1) + q(\hat{\theta}, 2\lambda - 1)]c_k, \tag{4.18}$$

and

$$\sum{}^{*} [q(\hat{\theta}, k/n) - q(\hat{\theta}, 2\lambda - 1)]c_k. \tag{4.19}$$

It follows from Lemma 4.2 that (4.18) is $O(n^{-3}/(\log n)^{3/4})$ a.s. Study (4.19) by doing a Taylor expansion in k/n about $(2\lambda - 1)$ and use the fact that

$$\sum{}^{*} [(k/n) - (2\lambda - 1)]c_k$$

is zero + $O(\log n/n)$. Thus (4.18) is $O(n^{-3/4}(\log n)^{3/4})$ which in turn implies (4.14) and the theorem is proved. □

REMARK 4.5. Clearly the estimator $T_2(\hat{\theta}, \lambda)$ can be represented as $T_2(\theta, \lambda) + (\hat{\theta} - \theta)$.

REMARK 4.6. We remark that $\hat{\theta}$ and $\hat{q}(\theta, 2\lambda - 1)$ are typically uncorrelated. Let $\hat{\theta} = \hat{\theta}(x_1, x_2, \cdots, x_n)$ be an odd location functional as defined in Hogg (1960). That is, $\hat{\theta}(x_1 + k, x_2 + k, \cdots, x_n + k) = \hat{\theta}(x_1, x_2, \cdots, x_n) + k$, and $\hat{\theta}(-x_1, -x_2, \cdots, -x_n) = -\hat{\theta}(x_1, x_2, \cdots, x_n)$. Our scale statistic $S(x_1, x_2, \cdots, x_n) = \hat{q}(\theta, 2\lambda - 1)$ is not an even scale statistic in the sense of Hogg (1960). Nevertheless, it is invariant under the transformation $w_i = 2\theta - x_i$, which amounts to rotating the data 180 degrees around θ. If $\sqrt{n}[\hat{\theta} - \theta, \hat{s} - q(\theta, \lambda)]$ has a limiting bivariate normal limit (U, V), it follows from the invariance properties of $\hat{\theta}$ and $\hat{s}$ that $\sqrt{n}[\hat{\theta}(2X_1 - \theta, \cdots, 2X_n - \theta) - \theta, \hat{s}(2X_1 - \theta, \cdots, 2X_n - \theta) - q(\theta, \lambda)]$ has its weak limit $(-U, V)$. However, in light of the symmetry of F about θ, the latter weak limit remains (U, V). The conclusion is that $(U, V) = (-U, V)$ in law. This implies that $\operatorname{cov}(U, V) = 0$.

REMARK 4.7. Stone (1975) studies an adaptive estimator which is location and scale invariant.

5. Asymptotic optimalities. In this section we need to add the further assumption that $F'(x) = f(x)$ exists everywhere. We retain our other assumptions given in Section 2. We start by assuming the center θ is known and without loss of generality $\theta = 0$ and $\lambda \geq .5$ so that $\xi_\lambda \geq 0$. Note that ξ_λ can be defined implicitly by the M-functional

$$\int \psi(x - \xi_\lambda)\, dF(x) = 0, \tag{5.1}$$

where $\psi(x) = \lambda/(1 - \lambda)$ if $x \geq 0$, $\psi(x) = -1$ if $x < 0$. The quantile ξ_λ as a functional whose domain is $\mathscr{F} = \{F\colon F' \text{ exists everywhere}\}$ is Hellinger differentiable. (See Beran, 1977.) Let $\mathscr{F}_S = \{F\colon f = F' \text{ exists everywhere}, f \text{ is symmetric about } 0, f \text{ satisfies the assumptions of Section } 2\}$.

LEMMA 5.1. *Let $F \in S$. A Hellinger derivative of ξ_λ at F with respect to S is*

$$(\tfrac{1}{2})[\rho_f(x) + \rho_f(-x)], \tag{5.2}$$

where

$$\rho_f(x) = 2\psi(x - \xi_\lambda)f^{1/2}(x)(1 - \lambda)/f(\xi_\lambda), \tag{5.3}$$

and the derivative is a symmetric function around zero.

PROOF. From the argument of Beran (1977), Theorem 4, (5.3) is a Hellinger derivative of ξ_λ with respect to the class of all densities in $\mathscr{F}_S$. Decompose ρ_f as follows:

$$\rho_f = (\rho_f(x) + \rho_f(-x))/2 + (\rho_f(x) - \rho_f(-x))/2 = \rho_f^+ + \rho_f^- \text{(say)}. \tag{5.4}$$

If $f_n \in \mathscr{F}_S$ then

$$\langle(\rho_f(x) - \rho_f(-x))/2, f_n^{1/2} - f^{1/2}\rangle = 0. \tag{5.5}$$

Therefore it is clear from the definition of Hellinger differentiability (see Beran, 1977, (2.1)) that ρ_f^+ is a Hellinger derivative with respect to $\mathscr{F}_S$, which is a symmetric function around zero. □

THEOREM 5.2. *For any regular estimator $\hat{\xi}_\lambda$ of ξ_λ, the asymptotic distribution of $\sqrt{n}(\hat{\xi}_\lambda - \xi_\lambda)$ can be represented as the convolution of the $N(0, \|\rho_f^+\|^2/4)$ distribution with $\mathscr{D}(f)$, a distribution depending on f.* (Note that the definition of regular is in Beran, 1977, page 438.)

PROOF. The result follows from the arguments in Theorem 6 of Beran (1977). The symmetry of ρ_f^+ is crucial in the argument if the domain is $\mathscr{F}_S$. □

The quantity

$$\|\rho_f^+\|^2/4 = (2\lambda - 1)(1 - \lambda)/2f^2(\xi_\lambda). \tag{5.6}$$

Recall from Section 2 that (5.6) is the variance of the asymptotic distribution of T_2, T_3, T_4 and T_5. Hence by virtue of Theorem 5.2, T_2, T_3, T_4 and T_5 are asymptotically optimal in the sense of minimum asymptotic variance in the class of regular estimates. See Remark 5.4 where the regular property is discussed for T_2, T_3, T_4 and T_5.

Next we discuss asymptotic efficiency from the point of view of the Hajék-LeCam minimax bound. We prove below that the asymptotic variance of T_2, T_3, T_4 and T_5 is the asymptotic minimax bound. We need to add to our assumptions that $f'(x)$ exists and

$$\int [(-f(x) - xf'(x))/f^{1/2}(x)]^2\, dx < \infty. \tag{5.7}$$

THEOREM 5.3. *Let F be fixed. Let $\hat{\xi}_\lambda$ be any estimator of ξ_λ. Then*

$$\lim_{c\to\infty}\lim_{n\to\infty}\inf_{\hat{\xi}_\lambda}\sup_{G\in B(F,c/\sqrt{n})}E_G[\sqrt{n}(\hat{\xi}_\lambda - \xi_\lambda)]^2 \geq (2\lambda - 1)(1-\lambda)/2f^2(\xi_\lambda) \tag{5.8}$$

where $B(F, c/\sqrt{n})$ is the Hellinger ball, of radius $c/\sqrt{n}$ around F.

PROOF. Note that the problem of estimating ξ_λ, a quantile of f, can be viewed as a problem of estimating ξ_λ, a scale parameter of g where $g(x) = \xi_\lambda f(x\xi_\lambda)$. We consider this equivalent problem of estimating the scale parameter for g when X is observed from $g\xi_\lambda(x) = (1/\xi_\lambda)g(x/\xi_\lambda)$. We want to apply Theorem 3.2 of Begun et al. (1983). Therefore we compute

$$\begin{aligned}\tau(g_{\xi_\lambda}) &= \partial\sqrt{g_{\xi_\lambda}}(x)/\partial\xi_\lambda \\ &= (½\ \xi_\lambda g_{\xi_\lambda}^{1/2}(x))[-g_{\xi_\lambda}(x) - xg'_{\xi_\lambda}(x)].\end{aligned} \tag{5.9}$$

We also find that $\beta^*(g_{\xi_\lambda})$, where β^* is defined in Begun et al. (1983), equals

$$\tau(g_{\xi_\lambda}) - \rho^+_{g_{\xi_\lambda}}/\|\rho^+_{g_{\xi_\lambda}}\|^2. \tag{5.10}$$

Now (5.9) and (5.10) enable application of the Begun et al. theorem. The asymptotic lower bound is $\|\rho_f^+\|^2/4$ which is the right side of (5.8). □

Now we treat the case of θ unknown. Define

$$\rho_f^*(x) = \rho_f^+(x) + 4f'_\theta(x)/I(f)f_\theta^{1/2}(x), \tag{5.11}$$

where $I(f) = \int [f'(x)/f(x)]^2 f(x)\,dx$. It can be shown that ρ_f^* is a Hellinger derivative in this case. Again Beran (1977), Theorem 6, implies that for any regular estimator $\hat{\xi}_\lambda$ of ξ_λ, the asymptotic distribution of $\sqrt{n}(\hat{\xi}_\lambda - \xi_\lambda)$ can be represented as a convolution of the $N(0, \|\rho_{f_\theta}^*\|^2/4)$ distribution with a distribution depending on f_θ only. Note that

$$\|\rho_{f_\theta}^*\|^2/4 = [(2\lambda - 1)(1-\lambda)/2f_\theta^2(\xi_\lambda)] + (1/I(f_\theta)). \tag{5.12}$$

Suppose then that in estimating ξ_λ we construct $T_2(\hat{\theta}, \lambda)$, $T_3(\hat{\theta}, \lambda)$, $T_4(\hat{\theta}, \lambda)$ with $\hat{\theta}$ an adaptive estimator of θ which is asymptotically uncorrelated with $T_2(\theta, \lambda)$, $T_3(\theta, \lambda)$, $T_4(\theta, \lambda)$. It follows then from the asymptotic representations of $T_j(\hat{\theta}, \lambda)$, $j = 2, 3, 4$, given in Section 4 and from (5.12) that such $T_j(\hat{\theta}, \lambda)$ are asymptotically optimal.

For θ unknown, the asymptotic efficiency in Hajék-LeCam sense would follow as in the case of θ known by again transforming the problem. This time the problem is to estimate ξ_λ, the scale parameter of $g_{\theta,\xi_\lambda} = g((x-\theta)/\xi_\lambda)$. It turns out that the asymptotic minimax bound is $\|\rho_f^*\|^2/4$. Certainly, the estimators $T_2(\hat{\theta}, \lambda)$, $T_3(\hat{\theta}, \lambda)$ and $T_4(\hat{\theta}, \lambda)$ are all optimal in this sense if $\hat{\theta}$ is adaptive.

REMARK 5.4. It can be shown that the sample quantile is a regular estimator. This in turn implies that T_3 is regular and in light of the asymptotic equivalence of T_3 with T_2, T_4 and T_5 it follows, using a contiguity argument, that T_2, T_3, T_4 and T_5 are regular.

To show that the sample quantile is regular, we use the fact that the remainder term times $\sqrt{n}$ in the Bahadur representation of quantiles converges to zero in probability under $\prod_{i=1}^{n} f(x_i)$. This in turn implies the same term converges to zero in probability under $\prod^{n} f_n(x_i)$, where f_n is such that $\| f_n^{1/2} - f^{1/2} - \beta/\sqrt{n} \| = o(1/\sqrt{n})$ for β orthogonal to $f^{1/2}$. Contiguity of $\prod_{i=1}^{n} f_n(x_i)$ to $\prod f(x_i)$ is used here.

Acknowledgement. The authors would like to thank Professor Charles Stone for a conversation which precipitated this study.

REFERENCES

BAHADUR, R. R. (1966). A note on quantiles in large samples. *Ann. Math. Statist.* **37** 577–580.

BEGUN, J. M., HALL, W. J., HWANG, W. M. and WELLNER, J. A. (1983). Information and asymptotic efficiency in parametric-nonparametric models. *Ann. Statist.* **11** 432–452.

BERAN, R. (1977). Robust location estimates. *Ann. Statist.* **6** 431–444.

DALAL, S. R. (1979). Dirichlet invariant processes and applications to nonparametric estimation of symmetric distribution functions. *Stochastic Process. Appl.* **9** 99–107.

FELLER, W. (1968). *An Introduction to Probability Theory and Its Applications* I, 3rd ed. Wiley, New York.

FELLER, W. (1971). *An Introduction to Probability Theory and Its Applications* II, 2nd ed. Wiley, New York.

FERGUSON, T. S. (1973). A Bayesian analysis of some nonparametric problems. *Ann. Statist.* **1** 209–280.

GASTWIRTH, J. L. (1971). On the sign test for symmetry. *J. Amer. Statist. Assoc.* **66** 821–823.

HANNUM, R. and HOLLANDER, M. (1983). Robustness of Ferguson's Bayes estimator of a distribution function. *Ann. Statist.* **11** 632–639, 1267.

HODGES, J. L., JR. and LEHMANN, E. L. (1956). The efficiency of some nonparametric competitors of the *t*-test. *Ann. Math. Statist.* **27** 324–335.

HOGG, R. V. (1960). Certain uncorrelated statistics. *J. Amer. Statist. Assoc.* **55** 265–267.

HOLLANDER, M. and WOLFE, D. A. (1973). *Nonparametric Statistical Methods.* Wiley, New York.

HUBER, P. J. (1981). *Robust Statistics.* Wiley, New York.

KIEFER, J. (1970). Deviations between the sample quantile process and the sample d.f. *Proc. Conference on Nonparametric Techniques in Statistical Inference, Bloomington* (M. L. Puri, ed.) Cambridge University Press, 299–319.

SCHUSTER, E. F. (1973). On the goodness of fit problem for continuous symmetric distribution. *J. Amer. Statist. Assoc.* **68** 713–715.

SERFLING, R. J. (1980). *Approximation Theorems of Mathematical Statistics*, Wiley, New York.

STONE, C. J. (1975). Adaptive maximum likelihood estimators of a location parameter. *Ann. Statist.* **3** 267–284.

DEPARTMENT OF STATISTICS
RUTGERS UNIVERSITY
NEW BRUNSWICK, NEW JERSEY 08903