The Annals of Statistics
1982, Vol. 10, No. 1, 197–204

NONPARAMETRIC INTERVAL AND POINT PREDICTION USING DATA TRIMMED BY A GRUBBS-TYPE OUTLIER RULE

BY RONALD W. BUTLER

The University of Texas

For a fixed probability $0 < \gamma < 1$, the "most outlying" $100(1-\gamma)\%$ subset of the data from a location model may be located with a Grubbs outlier subset test statistic. This subset is essentially located in terms of its complement, which is the connected $100\gamma\%$ span of the data which supports the smallest sample variance. We show that this range of the data may be characterized approximately as the $100\gamma\%$ span such that its midpoint is equal to the trimmed mean averaged over the span. Such a range forms a tolerance interval for predicting a future observation from the location model, and the asymptotic laws for its location, coverage, and center are presented.

1. Introduction and summary. A randomly-located tolerance interval is proposed for predicting a future observation from a location model. The tolerance interval is chosen to span a fixed proportion $0 < \gamma < 1$ of the data and represents the connected $100\gamma\%$ span which supports the smallest sample variance. Such a span remains following the arbitrary removal of the $100(1-\gamma)\%$ "most outlying" subset of data as determined by a smoothed version of a Grubbs (1950) multiple outlier test statistic. The asymptotic properties of this interval and the trimmed mean that it supports are studied.

On the basis of a random sample $X_1, \cdots, X_n$ from an absolutely continuous population distribution F, we wish to predict the next independent observation X_{n+1} from the same population using a $100\gamma\%$ tolerance interval. If F is strictly increasing over interval support and $q = F^{-1}$ denotes the quantile function, then the class of $100\gamma\%$ tolerance intervals is

$$\{I(\delta) = [q(\delta), q(\delta+\gamma)] : 0 \le \delta \le 1-\gamma\}. \tag{1.1}$$

We consider predicting X_{n+1} with the interval in (1.1) which supports the smallest trimmed variance; i.e., we use $I(\delta^*)$ where $\delta^* = \delta^*(\gamma)$ is the value of δ which minimizes

$$\begin{aligned}\sigma^2(\delta) &= \gamma^{-1}\int_{q(\delta)}^{q(\delta+\gamma)} x^2\,dF(x) - \left\{\gamma^{-1}\int_{q(\delta)}^{q(\delta+\gamma)} x\,dF(x)\right\}^2 \\ &= \gamma^{-1}\int_{\delta}^{\delta+\gamma} q^2(t)\,dt - \mu^2(\delta),\end{aligned} \tag{1.2}$$

where

$$\mu(\delta) = \gamma^{-1}\int_{\delta}^{\delta+\gamma} q(t)\,dt. \tag{1.3}$$

An estimator of δ^*, $\hat{\delta}^*$ say, is obtained when q in (1.2) is replaced by q_n, a piecewise linear variant of the empirical quantile process. (Sample analogues of population functionals will be denoted with a "$\hat{}$" throughout.)

The sample quantile function assumed here has been recommended by Parzen (1979) because of its accuracy in small samples. Let q_n be piecewise linear between the order statistics $\{X_{(i)} : i = 1, \cdots, n\}$ such that

Received August 1979; revised April 1981.

AMS 1970 *subject classifications.* Primary 62G15, 62G35; secondary 62G05, 62M20.

Key words and phrases. Prediction, tolerance intervals, nonparametric prediction, outliers.

$$q_n\{(i-\tfrac{1}{2})/n\} = X_{(i)} \qquad i = 1, \cdots, n$$
$$q_n(t) = X_{(1)} \qquad 0 \le t \le (2n)^{-1}$$
$$q_n(t) = X_{(n)} \qquad 1-(2n)^{-1} \le t \le 1.$$

With this notational convention, then, $\hat{\delta}^*$ locates the minimum of $\hat{\sigma}^2(\cdot)$. The consistency and asymptotic normality of $\hat{\delta}^*$ are shown in Section 2. The random interval $\hat{I}(\hat{\delta}^*) = [q_n(\hat{\delta}^*), q_n(\hat{\delta}^* + \gamma)]$ may be viewed as an estimator of $I(\delta^*)$. In Section 3, its coverage, P_n say, is shown to satisfy $\sqrt{n}(P_n - \gamma) \to_D N(0, \gamma(1-\gamma))$ as $n \to \infty$. Using this result, an approximate coverage assurance for $\hat{I}(\hat{\delta}^*)$ is easily set. With $\hat{\mu}(\cdot)$ based on (1.3), $\hat{\mu}(\hat{\delta}^*)$ provides an estimator of $\mu(\delta^*)$ which is considered in Section 4.

Of practical importance for the interval $\hat{I}(\hat{\delta}^*)$ is its correspondence with the span of the remaining data after the arbitrary trimming of the "most outlying" $100(1-\gamma)\%$ data subset as determined by $\hat{\sigma}^2(\cdot)$. Since $\hat{\sigma}^2(\cdot)$ is a smoothed version of the Grubbs (1950) multiple outlier test statistic, the two trimming procedures should be approximately the same and asymptotically equivalent. Allowing $\hat{\delta}^*$ to assume a continuous range of values, however, lends greater flexibility in the location of the tolerance interval.

A similar but asymptotically less successful approach to tolerance interval selection is based on the $100(1-\gamma)/2\%$ shown in Andrews et. al. (1972). This method estimates the shortest member of (1.1) indexed by δ^0. Since $\hat{\delta}^0 - \delta^0 = O_p(n^{-1/3})$, the properties of $\hat{I}(\hat{\delta}^0)$ have the same slow convergence rate with the main asymptotic variation due to $\hat{\delta}^0$.

Our proposal to fix γ and adaptively determine the trimming location $\delta^*(\gamma)$ is the opposite of Jaeckel's (1971) idea, which assumes symmetrical trimming $\delta^* = (1-\gamma)/2$ and adaptively determines the amount of trimming γ.

2. Large sample properties of $\hat{\delta}^*$. Sufficient conditions for the uniqueness of δ^* are specified below. From this, the consistency of $\hat{\delta}^*$ may be shown.

First, however, we must define a function Ψ as follows. Let $y_\delta(\cdot)$ denote the secant line segment to $q(\cdot)$, as shown in Figure 1, which connects points $(\delta, q(\delta))$ and $(\delta+\gamma, q(\delta+\gamma))$ for $\delta \in (0, 1-\gamma)$. Then let

$$\Psi(\delta) = \gamma^{-1}\int_{\delta}^{\delta+\gamma}\{y_\delta(t) - q(t)\}\,dt = \tfrac{1}{2}\{q(\delta) + q(\delta+\gamma)\} - \mu(\delta), \tag{2.1}$$

so that $\gamma\Psi(\delta)$ is the signed area between the graphs of y_δ and q. Then it is clear from the geometrical interpretation of Ψ that it admits a root in $(0, 1-\gamma)$ when F has unbounded support. Also, if Ψ has no root then it must be the result of F having bounded support.

LEMMA 2.1. *Let F be continuous and strictly increasing on interval support with differentiable density f that is unimodal. Then,*

(i) *$\sigma^2(\cdot)$ has a unique minimum at δ^*.*
(ii) *If (Case I) $\delta^* \in (0, 1-\gamma)$ then*
$$\Psi(\delta^*) = \tfrac{1}{2}[q(\delta^*) + q(\delta^*+\gamma)] - \mu(\delta^*) = 0.$$
Otherwise (Case II), $\delta^* = 0 (= 1-\gamma)$ and $\Psi(\delta) \ge 0 (\le 0)\ \forall\ \delta \in [0, 1-\gamma]$.
(iii) *In Case I, $\Psi'(\delta^*) > 0$.*
(iv) *$I(\delta^*) = [q(\delta^*), q(\delta^*+\gamma)]$ contains a neighborhood of the mode.*

Note that in case I, $I(\delta^*)$ is characterized as centered at the trimmed mean that it supports. Roughly speaking then, δ^* picks out the "most symmetric" $100\gamma\%$ of F by centering the interval mean.

PROOF. Since

$$\partial\sigma^2(\delta)/\partial\delta = 2\gamma^{-1}\{q(\delta+\gamma) - q(\delta)\}\Psi(\delta),$$

then $\sigma^2(\cdot)$ is monotonic when Ψ has no roots and (ii) follows. Results (i) and (iii) follow if

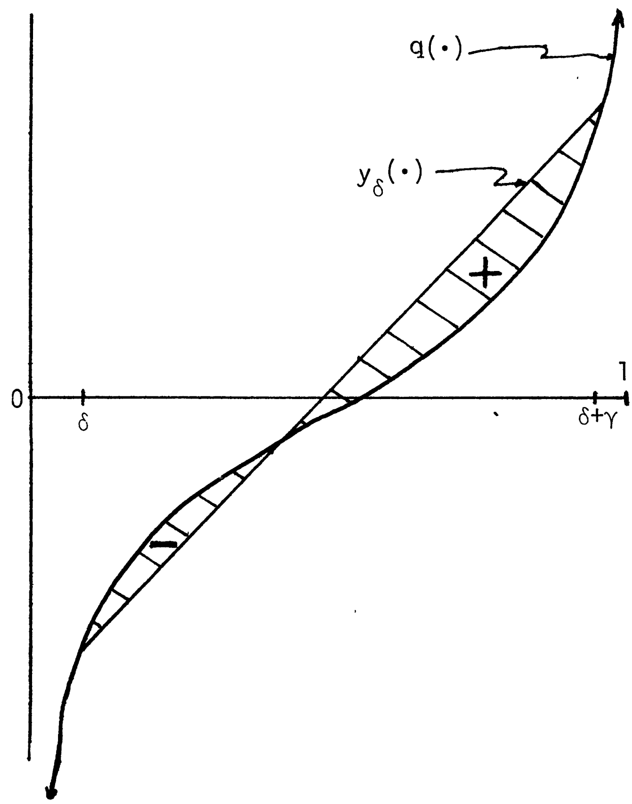

FIG. 1. *Geometrical interpretation of* $\gamma\Psi(\delta)$ *defined in equation* (2.1).

it can be shown that for δ^* an arbitrary root of Ψ, then

$$\partial^2\sigma^2(\delta^*)/\partial\delta^{*2} = 2\gamma^{-1}\{q(\delta^* + \gamma) - q(\delta^*)\}\Psi'(\delta^*) > 0.$$

Firstly,

$$\Psi'(\delta^*) \geq \min\{q'(\delta^*), q'(\delta^* + \gamma)\} - \gamma^{-1}\{q(\delta^* + \gamma) - q(\delta^*)\}. \tag{2.2}$$

From Figure 1 it is also clear that

$$\gamma^{-1}\{q(\delta^* + \gamma) - q(\delta^*)\} = q'(\delta_1) = q'(\delta_2) \tag{2.3}$$

by the mean value theorem, where $\delta^* < \delta_1 < m < \delta_2 < \delta^* + \gamma$ and m locates the inflection point of q, i.e. $q(m)$ = mode of f. Now since $\delta^* < \delta_1 < m$ then

$$q'(\delta^*) = \{f(q(\delta^*))\}^{-1} > \{f(q(\delta_1))\}^{-1} = q'(\delta_1)$$

and since $m < \delta_2 < \delta^* + \gamma$ then $q'(\delta^* + \gamma) > q'(\delta_2)$. From (2.2) and (2.3) then $\Psi'(\delta^*) > 0$. From Figure 1, $\delta^* < m < \delta^* + \gamma$ so (iv) follows.

The next lemma is necessary for deriving the large sample properties of $\hat{I}(\hat{\delta}^*)$.

LEMMA 2.2. *Let F be as in Lemma* 2.1 *and define* $Q_n(t) = \sqrt{n}\{q_n(t) - q(t)\}$ *for* $0 \leq t \leq 1$. *Then there exists a probability space on which a Brownian bridge* $\{B(t) : 0 \leq t \leq 1\}$ *is defined such that an equivalently distributed version of* $Q_n(t)$ *satisfies*

$$\sup_{c\leq t\leq d}|Q_n(t) - q'(t)B(t)| \to_P 0$$

as $n \to \infty$ *for any* $[c, d] \subset (0, 1)$.

PROOF. An immediate consequence of Skorokhod (1956) and Bickel (1967).

This lemma allows us to prove convergence in distribution of estimators based on q_n in terms of the stronger convergence of their equivalently-distributed versions based on the specially constructed quantile function. Hereafter, no notational distinction will be made between an estimator and its equivalent version based on the special process of Lemma 2.2.

LEMMA 2.3. *Assume the conditions of Lemma* 2.1 *and let F have a finite mean. Then, for Case I,*

$$\lim_{n\to\infty} P\{\hat{\Psi}(\hat{\delta}^*) = 0\} = 1. \tag{2.4}$$

NOTE. This states that with asymptotic certainty, $\hat{I}(\hat{\delta}^*)$ is centered at the sample mean it supports.

PROOF. From (1.2),

$$\partial\hat{\sigma}^2(\delta)/\partial\delta = 2\gamma^{-1}\{q_n(\delta+\gamma) - q_n(\delta)\}\hat{\Psi}(\delta). \tag{2.5}$$

Therefore if $\hat{\delta}^*$ does not occur at a root of $\hat{\Psi}$, then it must be 0 or $1-\gamma$. That $P\{\hat{\delta}^* = 0\} \to 0$ is easily seen by first noting that $\{\hat{\delta}^* = 0\} \subseteq \{\hat{\Psi}(0) \geq 0\}$. Since $\hat{\Psi}(0) \to_P \Psi(0) < 0$, the result follows. Similarly $P\{\hat{\delta}^* = 1-\gamma\} \to 0$.

LEMMA 2.4. *Assume the conditions of Lemma* 2.1 *and let F have a finite mean. Then* $\hat{\delta}^* \to_P \delta^*$ *as* $n \to \infty$. *Moreover, for Case II,*

$$P\{\hat{\delta}^* = 0\} \to 1 \qquad \text{if } \delta^* = 0 \text{ and } \Psi(\delta) > 0\ \forall\delta,$$

$$\text{and } P\{\hat{\delta}^* = 1-\gamma\} \to 1 \qquad \text{if } \delta^* = 1-\gamma \text{ and } \Psi(\delta) < 0\ \forall\delta.$$

PROOF. *Case I.* By Lemma 2.3, it suffices to show that an arbitrary sequence of roots for $\hat{\Psi}$ is consistent for δ^*. Let $\hat{\delta}^*$ denote such a sequence.

We first find neighborhoods of 0 and $1-\gamma$, $[0, \varepsilon_1]$ and $[1-\gamma-\varepsilon_2, 1-\gamma]$ say for $\varepsilon_1, \varepsilon_2 > 0$, such that $P\{\hat{\delta}^* \in [0, \varepsilon_1]\} \to 0$ and $P\{\hat{\delta}^* \in [1-\gamma-\varepsilon_2, 1-\gamma]\} \to 0$. This is possible by choosing $\varepsilon_1 < \delta^*$ small enough so that

$$\Pi_1 = ½\,\{q(\varepsilon_1) + q(\varepsilon_1+\gamma)\} - \gamma^{-1}\int_0^\gamma q(t)\,dt < 0,$$

and ε_2 such that $\delta^* < 1-\gamma-\varepsilon_2$ and

$$\Pi_2 = ½\,\{q(1-\gamma-\varepsilon_2) + q(1-\varepsilon_2)\} - \gamma^{-1}\int_{1-\gamma}^1 q(t)\,dt > 0. \tag{2.6}$$

Then

$$\{0 \leq \hat{\delta}^* \leq \varepsilon_1\} \subseteq \{\min_{0\leq\delta\leq\varepsilon_1}\hat{\Psi}(\delta) \leq 0 \leq \max_{0\leq\delta\leq\varepsilon_1}\hat{\Psi}(\delta)\} \subseteq \{0 \leq \max_{0\leq\delta\leq\varepsilon_1}\hat{\Psi}(\delta) < \hat{\Pi}_1\}. \tag{2.7}$$

Since $\hat{\Pi}_1 \to_P \Pi_1 < 0$, the event on the right hand side of (2.7) has asymptotic probability zero. Also

$$\{1-\gamma-\varepsilon_2 \leq \hat{\delta}^* \leq 1-\gamma\} \subseteq \{\hat{\Pi}_2 \leq \min_{0\leq\delta\leq\varepsilon_1}\hat{\Psi}(\delta) \leq 0\}. \tag{2.8}$$

Since $\hat{\Pi}_2 \to_P \Pi_2 > 0$, the events in (2.8) have asymptotic probabilities zero.

Therefore, estimator $\hat{\delta}^*$ is asymptotically certain to appear in $[\varepsilon_1, 1-\gamma-\varepsilon_2]$ and we restrict attention to this interval. If we can assume that

$$\max_{\varepsilon_1\leq\delta\leq 1-\gamma-\varepsilon_2}|\hat{\Psi}(\delta) - \Psi(\delta)| \to_P 0 \tag{2.9}$$

and show that for any $\xi > 0$, there exists an $\eta > 0$ such that

$$\{\max_{\varepsilon_1 \leq \delta \leq 1-\gamma-\varepsilon_2} |\hat{\Psi}(\delta) - \Psi(\delta)| < \eta\} \subseteq \{|\hat{\delta}^* - \delta^*| < \xi\}, \tag{2.10}$$

then consistency will have been shown.

Result (2.9) follows since

$$\max_{\varepsilon_1 \leq \delta \leq 1-\gamma-\varepsilon_2} |\hat{\Psi}(\delta) - \Psi(\delta)| \leq 2 \max_{\varepsilon_1 \leq \delta \leq 1-\gamma-\varepsilon_2} |q_n(\delta) - q(\delta)| \to_P 0$$

by Lemma 2.2.

Suppose the event on the left of (2.10) is true for sufficiently small η. Then

$$\max_{\varepsilon_1 \leq \delta \leq 1-\gamma-\varepsilon_2} \Big| |\hat{\Psi}(\delta)| - |\Psi(\delta)| \Big| < \eta. \tag{2.11}$$

Let $I_\xi = \{\delta \in [\varepsilon_1, 1 - \gamma - \varepsilon_2] : |\delta - \delta^*| \geq \xi\}$ and suppose that η is chosen such that

$$0 < \eta < \tfrac{1}{2} \cdot \min_{\delta \in I_\xi} |\Psi(\delta)|. \tag{2.12}$$

Then for any $\delta \in I_\xi$,

$$|\hat{\Psi}(\delta)| - |\hat{\Psi}(\delta^*)| = |\hat{\Psi}(\delta)| - |\Psi(\delta)| + |\Psi(\delta)| - |\Psi(\delta^*)| + |\Psi(\delta^*)| - |\hat{\Psi}(\delta^*)|$$
$$> -\eta + \min_{\delta \in I_\xi} |\Psi(\delta)| - \eta > 0$$

by (2.11) and (2.12). Therefore, the left side of (2.10) is

$$\subseteq \{|\hat{\Psi}(\delta)| > |\hat{\Psi}(\delta^*)| \,\forall\, \delta \in I_\xi\}$$
$$\subseteq \{\hat{\delta}^* \notin I_\xi\} = \{|\hat{\delta}^* - \delta^*| < \xi\},$$

when η satisfies (2.12).

PROOF. *Case II* ($\delta^* = 0$). We proceed as in Case I. Let $[0, \varepsilon_1]$, $[\varepsilon_1, 1 - \gamma - \varepsilon_2]$, and $[1 - \gamma - \varepsilon_2, 1 - \gamma]$ partition the range of $\hat{\delta}^*$ so that

$$\Pi_3 = \tfrac{1}{2}\{q(0) + q(\gamma)\} - \gamma^{-1} \int_{\varepsilon_1}^{\varepsilon_1+\gamma} q(t)\, dt > 0$$

and $\Pi_2 > 0$ for Π_2 as in (2.6). Then because

$$\{\min_{0 \leq \delta \leq 1-\gamma} \hat{\Psi}(\delta) > 0\} \subseteq \{\hat{\delta}^* = 0\}, \tag{2.13}$$

it is sufficient to show that min $\hat{\Psi}$ assumes a positive value with asymptotic probability one for each of the three intervals in the partition. For the first two,

$$\min_{0 \leq \delta \leq \varepsilon_1} \hat{\Psi}(\delta) \geq \hat{\Pi}_3 \to_P \Pi_3 > 0, \quad \text{and} \quad \min_{1-\gamma-\varepsilon_2 \leq \delta \leq 1-\gamma} \hat{\Psi}(\delta) \geq \hat{\Pi}_2 \to_P \Pi_2 > 0.$$

Hence, the events in (2.13) have asymptotic probability one if it can be shown that

$$\min_{\varepsilon_1 \leq \delta \leq 1-\gamma-\varepsilon_2} \hat{\Psi}(\delta) \to_P \min_{\varepsilon_1 \leq \delta \leq 1-\gamma-\varepsilon_2} \Psi(\delta) > 0. \tag{2.14}$$

To show this, let $\tilde{\delta}_n(\tilde{\delta})$ locate the minimum of $\hat{\Psi}$ (Ψ) over $[\varepsilon_1, 1 - \gamma - \varepsilon_2]$. Then by (2.9), $\hat{\Psi}(\tilde{\delta}_n) - \Psi(\tilde{\delta}_n) \to_P 0$. We would like to show that $\Psi(\tilde{\delta}_n) \to_P \Psi(\tilde{\delta})$, and it suffices to show $\tilde{\delta}_n \to_P \tilde{\delta}$. This follows from the argument used in Case I by replacing $|\Psi|$ with Ψ in the proof. An analogous proof holds when $\delta^* = 1 - \gamma$ and $\Psi(\delta) < 0 \,\forall\, \delta$.

The large sample distribution of $\hat{\delta}^*$ is degenerate at δ^* for Case II as shown in Lemma 2.4. The next theorem considers its asymptotic distribution for Case I.

THEOREM 2.1. *Let F be continuous, strictly increasing on its interval support, and with a finite mean and a differentiable density which is unimodal. Then for Case I,*

$$\sqrt{n}\,(\hat{\delta}^* - \delta^*) \to_D \{\Psi'(\delta^*)\}^{-1} \tag{2.15}$$
$$\cdot\left[\gamma^{-1}\int_{\delta^*}^{\delta^*+\gamma} B(t)q'(t)\,dt - \tfrac{1}{2}\{B(\delta^*)q'(\delta^*) + B(\delta^*+\gamma)q'(\delta^*+\gamma)\}\right]$$

as $n \to \infty$ where $\{B(t): 0 \le t \le 1\}$ denotes a Brownian bridge.

PROOF. Because of Lemma 2.3, $\sqrt{n}\,\hat{\Psi}(\hat{\delta}^*) \to_P 0$. In terms of the quantile process Q_n, this states that

$$0 \;{}_P\!\leftarrow \sqrt{n}\,\hat{\Psi}(\hat{\delta}^*) = \tfrac{1}{2}\{Q_n(\hat{\delta}^*) + Q_n(\hat{\delta}^*+\gamma)\} - \gamma^{-1}\int_{\hat{\delta}^*}^{\hat{\delta}^*+\gamma} Q_n(t)\,dt + \sqrt{n}\,\{\Psi(\hat{\delta}^*) - \Psi(\delta^*)\}.$$

Using Lemma A.1 of the Appendix and the delta method, then

$$\tfrac{1}{2}\{B(\delta^*)q'(\delta^*) + B(\delta^*+\gamma)q'(\delta^*+\gamma)\} - \gamma^{-1}\int_{\delta^*}^{\delta^*+\gamma} B(t)q'(t)\,dt$$
$$+ \{\Psi'(\delta^*) + o_P(1)\}\,\sqrt{n}\,(\delta^* - \delta^*) \to_P 0$$

and (2.15) follows. □

Example 1. Suppose f is the $N(0, 1)$ density and $\gamma = .9$. Then $\sqrt{n}\,(\hat{\delta}^* - .05) \to_D N(0, .0584)$.

3. Tolerance Intervals. The $100\gamma\%$ span of untrimmed data $\hat{I}(\hat{\delta}^*)$ provides a tolerance interval predictor of the next observation from the location sample. The following result specifies the large sample distribution of its coverage P_n, defined as the conditional probability of covering X_{n+1} with $\hat{I}(\hat{\delta}^*)$ given $X_1, \cdots, X_n$.

THEOREM 3.1. *Suppose the same assumptions as in Theorem 2.1, and let $P_n(\gamma) = F\{q_n(\hat{\delta}^* + \gamma)\} - F\{q_n(\hat{\delta}^*)\}$ be the coverage of $\hat{I}(\hat{\delta}^*)$. Then as $n \to \infty$*

$$\sqrt{n}\,\{P_n(\gamma) - \gamma\} \to_D N(0, \gamma(1-\gamma)). \tag{3.1}$$

PROOF. *Case I.*

$$\sqrt{n}\,(P_n - \gamma) = \sqrt{n}\,[F\{q_n(\hat{\delta}^*+\gamma)\} - F\{q(\delta^*+\gamma)\}] - \sqrt{n}\,[F\{q_n(\hat{\delta}^*)\} - F\{q(\delta^*)\}]. \tag{3.2}$$

Since

$$\begin{aligned}\sqrt{n}\,\{q_n(\hat{\delta}^*) - q(\delta^*)\} &= \sqrt{n}\,\{q_n(\hat{\delta}^*) \pm q(\hat{\delta}^*) - q(\delta^*)\}\\ &= Q_n(\hat{\delta}^*) + \sqrt{n}\,\{q(\hat{\delta}^*) - q(\delta^*)\}\\ &= q'(\delta^*)\{B(\delta^*) + \sqrt{n}\,(\hat{\delta}^* - \delta^*)\} + o_P(1)\end{aligned} \tag{3.3}$$

by Lemma A.1 and the delta method, then (3.2) is

$$\begin{aligned}&= f\{q(\delta^*+\gamma)\}q'(\delta^*+\gamma)\{B(\delta^*+\gamma) + \sqrt{n}\,(\hat{\delta}^* - \delta^*)\}\\ &\quad - f\{q(\delta^*)\}q'(\delta^*)\{B(\delta^*) + \sqrt{n}\,(\hat{\delta}^* - \delta^*)\} + o_P(1)\\ &= B(\delta^*+\gamma) - B(\delta^*) + o_P(1) \to_D N(0, \gamma(1-\gamma)).\end{aligned}$$

PROOF. *Case II.* Suppose $\delta^* = 0$. Then, by Lemma 2.4, it suffices to find the asymptotic distribution of $F\{q_n(\gamma)\} - F\{q_n(0)\}$. Since $q_n(0) - q(0) = O_P(n^{-1})$, then $\sqrt{n}\,(P_n - \gamma) \sim \sqrt{n}\,[F\{q_n(\gamma)\} - \gamma] \to_D N(0, \gamma(1-\gamma))$. A similar argument holds when $\delta^* = 1 - \gamma$. □

The simplicity of this result makes it an attractive method for setting the coverage assurance of $\hat{I}(\hat{\delta}^*)$ in large samples. Suppose in a sample of size n, for example, a c-coverage tolerance interval ($c \geq \frac{1}{2}$) is sought with g-guarantee of attaining the c-coverage. Then the smallest γ^* is sought ($\gamma^* \geq c$) which satisfies

$$P\{P_n(\gamma^*) \geq c\} \geq g. \tag{3.4}$$

For sufficiently large n, the probability (3.4) may be based on (3.1), the asymptotic distribution of P_n.

Example 2. For $c = .85$, $g = .95$, and $n = 93$, then $\gamma^* = .906$ when the assurance statement (3.4) is based on (3.1).

A reverse problem involves determining the sample size (n) necessary for attaining prespecified g-guarantee of c-coverage using an interval with γ-span.

Example 3. For $g = .95$, $c = .85$, and $\gamma = .9$, then n must exceed 93 according to Theorem 3.1. If $[q_n(.05), q_n(.95)]$ were to be used, then $n \geq 82$ may be read from the charts of Murphy (1948) in order to attain the specified coverage guarantee.

4. Point Prediction. The center of $\hat{I}(\hat{\delta}^*)$ is a point predictor of X_{n+1} which arises in a very natural way. In the situation of Case I, Lemma 2.3 states that this center is asymptotically the mean of the data remaining after the elimination of the $100(1 - \gamma)\%$ "most outlying" subset of data as determined with a Grubbs-type statistic. The next theorem specifies the asymptotic distribution of this interval center.

THEOREM 4.1. *Suppose the conditions of Theorem* 2.1 *and let* $C_n(\gamma)$ *denote the center of* $\hat{I}(\hat{\delta}^*)$. *Then in Case I,*

$$\sqrt{n}\,\{\hat{\mu}(\hat{\delta}^*) - \mu(\delta^*)\} - \sqrt{n}\,[C_n(\gamma) - \tfrac{1}{2}\,\{q(\delta^*) + q(\delta^* + \gamma)\}] \to_P 0 \tag{4.1}$$

as $n \to \infty$ *with common asymptotic distribution given by*

$$\tfrac{1}{2}\,p(\delta^*)\{B(\delta^*)q'(\delta^*) + B(\delta^* + \gamma)q'(\delta^* + \gamma)\} + \{1 - p(\delta^*)\}\gamma^{-1}\int_{\delta^*}^{\delta^*+\gamma} B(t)q'(t)\,dt, \tag{4.2}$$

where $p(\delta^*) = \{\gamma\Psi'(\delta^*)\}^{-1}\{q(\delta^* + \gamma) - q(\delta^*)\}$. *This as a* $N(0, \rho^2(\delta^*))$ *distribution where*

$$\rho^2(\delta^*) = (\gamma - 2\,q\{\,(1-\gamma)/2\}f[q\{\,(1-\gamma)/2\}])^{-2}\int_{(1-\gamma)/2}^{(1+\gamma)/2} q^2(t)\,dt$$

for symmetric f. In Case II,

$$\sqrt{n}\,[C_n(\gamma) - \tfrac{1}{2}\{q(0) + q(\gamma)\}] \to_D N(0, \tfrac{1}{4}\gamma(1-\gamma)/f^2\{q(\gamma)\}). \tag{4.3}$$

Note that for large γ, $\rho^2(\delta^*) \simeq \gamma^{-1}\sigma^2(\delta^*)$.

PROOF. *Case I.* The equivalence of (4.1) follows since $\sqrt{n}\,\{\hat{\Psi}(\hat{\delta}^*) - \Psi(\delta^*)\} \to_P 0$. The common distribution (4.2) may be derived from the latter term of (4.1), which may be expressed as

$$\begin{aligned} &\tfrac{1}{2}\{Q_n(\hat{\delta}^*) + Q_n(\hat{\delta}^* + \gamma)\} + \tfrac{1}{2}\sqrt{n}\,[\{q(\hat{\delta}^*) - q(\delta^*)\} + \{q(\hat{\delta}^* + \gamma) - q(\delta^* + \gamma)\}] \\ &\quad = \tfrac{1}{2}\{B(\delta^*)q'(\delta^*) + B(\delta^* + \gamma)q'(\delta^* + \gamma)\} \\ &\qquad + \tfrac{1}{2}\{q'(\delta^*) + q'(\delta^* + \gamma)\}\sqrt{n}\,(\hat{\delta}^* - \delta^*) + o_P(1) \end{aligned} \tag{4.4}$$

by Lemma A.1 and the delta method. The probability limit of $\sqrt{n}\,(\hat{\delta}^* - \delta^*)$ is specified in (2.15) of Theorem 2.1; this proof really demonstrates the stronger convergence in probability that is required. Upon substitution of (2.15) into (4.4), then (4.2) follows.

PROOF. *Case II.* Because of Lemma 2.3, $\hat{I}(\hat{\delta}^*)$ behaves asymptotically like $[q_n(0), q_n(\gamma)]$. Therefore $C_n(\gamma)$ behaves like $\frac{1}{2}\{q_n(0) + q_n(\gamma)\}$. Since $q_n(0) - q(0) = O_p(n^{-1})$, (4.3) has the asymptotic distribution of $\frac{1}{2}\sqrt{n}\ \{q_n(\gamma) - q(\gamma)\}$ and is given in (4.3). □

Example 4. For f a $N(0, 1)$ density and $\gamma = .9$, then $\rho^2(.05) = 1.77$ and $C_n(.9)$ is 56% efficient relative to $\bar{X}$.

5. Acknowledgments. A preliminary version of this paper formed a part of the author's Ph.D. dissertation written at the University of Michigan under the supervision of Professor Edward D. Rothman. The author is grateful to Professor Michael Woodroofe of Ann Arbor for his help with the asymptotic distribution theory and to a referee and associate editor for their constructive criticism. General thanks goes to Shri B. K. S. Iyengar of Poona, India.

APPENDIX

LEMMA A.1. *Under the conditions of Theorem* 2.1,

$$Q_n(\hat{\delta}^*) \to_P B(\delta^*)q'(\delta^*), \qquad Q_n(\hat{\delta}^* + \gamma) \to_P B(\delta^* + \gamma)q'(\delta^* + \gamma), \tag{A.1}$$

and

$$\int_{\hat{\delta}^*}^{\hat{\delta}^*+\gamma} Q_n(t)\,dt \to_P \int_{\delta^*}^{\delta^*+\gamma} B(t)q'(t)\,dt$$

as $n \to \infty$.

PROOF. It suffices to assume $\hat{\delta}^* \in [\varepsilon_1, 1 - \gamma - \varepsilon_2]$ as in Lemma 2.4 since the complementary event has been shown to have a limiting probability of zero. Since

$$|Q_n(\hat{\delta}^*) - B(\delta^*)q'(\delta^*)| \le |Q_n(\hat{\delta}^*) - B(\hat{\delta}^*)q'(\hat{\delta}^*)| + |B(\hat{\delta}^*)q'(\hat{\delta}^*) - B(\delta^*)q'(\delta^*)|,$$

then (A.1) follows from Lemma 2.2 and the consistency of $\hat{\delta}^*$.

By adding and subtracting $\int_{\delta^*}^{\delta^*+\gamma} Q_n(t)\,dt$, then

$$\begin{aligned}\int_{\hat{\delta}_*}^{\hat{\delta}_*+\gamma} Q_n(t)\,dt &- \int_{\delta^*}^{\delta^*+\gamma} B(t)q'(t)\,dt \\ &= \int_{\hat{\delta}^*}^{\delta^*} Q_n(t)\,dt - \int_{\hat{\delta}^*+\gamma}^{\delta^*+\gamma} Q_n(t)\,dt + \int_{\delta^*}^{\delta^*+\gamma} \{Q_n(t) - B(t)q'(t)\}\,dt.\end{aligned} \tag{A.2}$$

Then (A.2) $\to_P 0$ because of Lemma 2.2 and the consistency of $\hat{\delta}^*$. □

REFERENCES

ANDREWS, D. F., BICKEL, P. J., HAMPEL, F. R., HUBER, P. J., ROGERS, W. H. and TUKEY, J. W. (1972). *Robust Estimates of Location: Survey and Advances.* Princeton Univ. Press, Princeton, New Jersey.

BICKEL, P. J. (1967). Some contributions to the theory of order statistics. *Proc. Fifth Berkeley Symp. Math. Statist. Prob.* **1** 575–591.

GRUBBS, F. E. (1950). Sample criteria for testing outlying observations. *Ann. Math. Statist.* **21** 27–58.

JAECKEL, L. A. (1971). Some flexible estimates of location. *Ann. Math. Statist.* **42** 1540–1552.

MURPHY, R. B. (1948). Non-parametric tolerance limits. *Ann. Math. Statist.* **19** 581–589.

PARZEN, E. (1979). Nonparametric statistical data modelling. *J. Amer. Statist. Assoc.* **74** 105–131.

SKOROKHOD, A. V. (1956). Limit theorems for stochastic processes. *Theory of Probab. and Appl.* **1** 261–290.

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF TEXAS
AUSTIN, TX 78712