Sample extreme values are biased estimators of the end-points of a distribution, and hence, jackknifing is useful. However, the properties of jackknifing in such a case differ considerably from those in the regular case. These are studied here. Along with a modification of jackknifing, some applications are also considered.


Key Words & Phrases: bias; extreme values; jackknifing; mean square; order of terminal contact; studentized form; Tukey-estimator of variance.

1. INTRODUCTION

Let \( \{X_i, i \geq 1\} \) be a sequence of independent and identically distributed random variables (i.i.d.r.v.) with a distribution function (df) \( F \), defined on \( (-\infty, \infty) \). It is assumed that \( F \) has a finite (unknown) lower end-point \( \theta \), that is

\[
-\infty < \theta = \sup \{x : F(x) = 0\} < \infty
\]

and that \( F(x) \) is continuous and monotonic in \( x \in (\theta, \theta + \delta) \), for some \( \delta > 0 \). A natural estimator of \( \theta \) is the sample minimum i.e.,

\[
\hat{\theta}_n = \min\{X_1, \ldots, X_n\} = X_{n,1} \quad (n \geq 1),
\]

where \( X_{n,1} \leq \ldots \leq X_{n,n} \) stand for the ordered variables corresponding to \( X_1, \ldots, X_n; n \geq 1 \). \( \hat{\theta}_n \) is a (strongly) consistent estimator of \( \theta \), but it is not an unbiased one; the nature of its bias depends on the order of terminal contact of \( F \) (at \( \theta \)). It may therefore be appealing to use the jackknife estimator corresponding to \( \hat{\theta}_n \).

Under quite general regularity conditions (viz., [1,2,4]), jackknifing meets three objectives: (a) Bias reduction. If \( \theta^* \) be the jackknife estimator then \( nE(\theta^*_n - \theta) \to 0 \) as \( n \to \infty \). (b) Asymptotic normality. If \( n^{1/2}(\hat{\theta}_n - \theta) \) is asymptotically normal, then the same limit law holds for \( n^{1/2}(\theta^*_n - \theta) \). (c) The Tukey estimator \( \nu_n^2 \) [defined by (2.5)] is a (strongly) consistent estimator of the variance of \( n^{1/2}(\theta^*_n - \theta) \).

Since the asymptotic distributions of sample extrema are non-normal and, depending on the order of terminal contact, the bias of \( \hat{\theta}_n \) is \( O(n^{-a}) \).
for some $0 < a \leq 1$, the effectiveness of jackknifing in regard to (a) and (b) remains to be examined carefully. Further, in this case, $\nu_n^2$ does not converge (stochastically). Along with the preliminary notions, expressions for $\theta^*$ and $\nu_n^2$ are considered in Section 2. The main results are studied in Section 3. Section 4 deals with a modification of jackknifing appropriate for the case of the bias of $O(n^{-a})$ for some $a < 1$. Some general remarks are made in the concluding section.

2. PRELIMINARY NOTIONS

We assume that for some non-negative integer $m$, $F(x)$ has continuous $j$th derivative $F^{(j)}(x) = f^{(j-1)}(x)$ for all $x \in (\theta, \theta + \delta)$, $\delta > 0$, $1 \leq j \leq m + 1$. We denote the (right hand) derivatives at $\theta$ by $F^{(j)}(\theta) = f^{(j-1)}(\theta)$, $1 \leq j \leq m + 1$ and $F^{(0)}(\theta) = 0$, $f^{(0)}(\theta) = f_{+}(\theta)$. Then, a terminal contact of order $m$ is defined by

\[(2.1) \quad F^{(j)}(\theta) = 0, \quad 0 \leq j \leq m \quad \text{and} \quad 0 < f^{(m)}_{+}(\theta) < \infty.\]

Also, for the study of the bias, we assume that

\[(2.2) \quad \nu = \int_{\theta}^{\infty} |x|^\alpha dF(x) < \infty \quad \text{for some} \quad \alpha > 0.\]

To define $\theta_n^*$, we let for each $i$: $1 \leq i \leq n$,

\[(2.3) \quad \delta_{n-1}^i = \min\{x_{i-1}, \ldots, x_{i-2}, x_{i+1}, \ldots, x_n\}, \quad \hat{\delta}_{n,1}^i = n\delta_n^* - (n-1)\delta_{n-1}^i.\]

Then, $\hat{\delta}_{n-1}^i$ is equal to $x_{n,1}$ or $x_{n,2}$ according as $X_i$ is $\neq$ or $= x_{n,1}$, $1 \leq i \leq n$. Also,
The Tukey estimator $\nu^2_n$, defined by
\begin{equation}
\nu^2_n = \frac{1}{(n-1)} \sum_{i=1}^{n} (\hat{\theta}_{n,i} - \theta^*)^2 = (n-1) \sum_{i=1}^{n} (\hat{\theta}_{n-1,i} - \theta^*)^2 ,
\end{equation}

reduces in our case to

\begin{equation}
\nu^2_n = (X_{n,2} - X_{n,1})^2 (n-1)(n^2+n-1)/n \quad (\sim \{n(X_{n,2} - X_{n,1})\})^2 .
\end{equation}

For a terminal contact of order $m(\geq 0)$, we define

\begin{equation}
b_{n,m} = \{nf(m)(\theta)/(m+1)!\}^{1/(m+1)} , \quad a_m = 1/(m+1) .
\end{equation}

Then, the limiting distribution of $b_{n,m}(\hat{\theta}_n - \theta)$ is known to be

\begin{equation}
\Lambda_m(x) = \begin{cases} 0 , & x \leq 0 , \\
1 - \exp(-x^{m+1}) , & x > 0 . 
\end{cases}
\end{equation}

Also, by Theorem 3.1 of Sen (1961), as $n \to \infty$,

\begin{equation}
b_{n,m} E(X_{n,r} - \theta) = \frac{r+a_m}{r+o(1)} , \quad \text{for every (fixed) } r(\geq 1) .
\end{equation}

3. BASIC PROPERTIES OF JACKKNIFING

It follows from (2.4) that

\begin{equation}
n(\theta_n^* - \theta) = n(X_{n,1} - \theta) - (n-1)(X_{n,2} - X_{n,1})
= (2n-1)(X_{n,1} - \theta) - (n-1)(X_{n,2} - \theta) .
\end{equation}
Hence, from (2.9) and (3.1), we obtain that for a terminal contact of order \( m \),

\[
(3.2) \quad b_{n,m} E(\theta_n^* - \theta) = (1-a_m) \left[ \frac{1+a_m}{1+a_m} + o(1) \right]
= (1-a_m) \{ b_{n,m} E(\theta_n^* - \theta) \} + o(1) .
\]

For \( m = 0 \) i.e., \( a_m = 1 \), the right hand side (rhs) of (3.2) converges to 0, as \( n \to \infty \), while for \( m \geq 1 \) (i.e., \( a_m \leq \frac{1}{2} \)), jackknifing leads to effectively 100(1-a_m)% reduction in bias. Thus, the basic role of jackknifing is partially impaired for a terminal contact of order \( m(\geq 1) \).

**Theorem 1.** For a terminal contact of order \( m(\geq 0) \),

\[
A^*_n(x) = \lim_{n \to \infty} n^{\frac{1}{2}} \exp \left\{ \int_0^x \exp \left\{ \frac{a_m}{m+1} \right\} dy \right\}, \quad -\infty < x \leq 0 ,
\]

\[
= \begin{cases} 
1 - \exp \left( -x^{m+1} \right) + \int_x^\infty \exp \left\{ \frac{a_m}{m+1} \right\} dy, & x > 0,
\end{cases}
\]

where \( a_m \) and \( b_{n,m} \) are defined by (2.7).

**Proof.** Let \( Z_n = b_{n,m}(\theta_n^* - \theta) \) and let

\[
(3.4) \quad Y_n^{(1)} = nF(X_n,1) \quad \text{and} \quad Y_n^{(2)} = n[F(X_n,2) - F(X_n,1)] .
\]

Then, by (2.1), (2.2), (2.7), (3.1) and (3.4) and proceeding as in the proof of Theorem 3.1 of Sen (1961), we obtain that

\[
(3.5) \quad E[Z_n - 2Y_n^{(1)} + (Y_n^{(1)} + Y_n^{(2)})^{a_m}]^2 \to 0 \quad \text{as} \quad n \to \infty .
\]
and hence, by the Chebychev inequality, we have

\[(3.6) \quad \Lambda^*_m(x) = \lim_{n \to \infty} \Pr\left\{2Y_{n(1)}^m - (Y_{n(1)} + Y_{n(2)})^m \leq x \right\}, \quad \forall -\infty < x < \infty.\]

We may recall that \(Y_{n(1)}\) and \(Y_{n(2)}\) are asymptotically independently distributed according to a common simple exponential law and they are non-negative rv's. For \(x \leq 0\), \(\left[2Y_{n(1)}^m - (Y_{n(1)} + Y_{n(2)})^m \leq x \right] \iff \left[Y_{n(2)} \geq \left(2Y_{n(1)}^m - x\right)^{m+1} - Y_{n(1)}^m\right]\) and the first equation in (3.3) follows directly by finding the conditional probability given \(Y_{n(1)}\) and then integrating it out over \(Y_{n(1)}\). For \(x > 0\), if \(Y_{n(1)} \leq x^{m+1}\), then \(2Y_{n(1)}^m - (Y_{n(1)} + Y_{n(2)})^m \leq Y_{n(1)}^m\), while for \(Y_{n(1)} > x^{m+1}\), as before we need \(Y_{n(2)} \geq \left(2Y_{n(1)}^m - x\right)^{m+1} - Y_{n(1)}^m\), and hence, the last equation in (3.3) follows on parallel lines. Q.E.D.

For \(m = 0\) (i.e., \(a_m = 1\)), \(\Lambda_m^0\) in (2.8) is the simple exponential while \(\Lambda_m^0\) in (3.3) is the double exponential df. For \(m \geq 0\), \(\Lambda_m^0\) and \(\Lambda_m^*\) are not the same df.

**Theorem 2.** For a terminal contact of order \(m(\geq 0)\),

\[(3.7) \quad \lim_{n \to \infty} \left\{E\left[\frac{b_n^2}{n,m}(\hat{\theta}_n - \theta)^2\right]\right\} = \left\{1 - \frac{2a_m(1-a_m)}{1+a_m}\right\}\lim_{n \to \infty} \left\{E\left[\frac{b_n^2}{n,m}(\hat{\theta}_n - \theta)^2\right]\right\} = \left(2a_m^2(1 - 2a_m(1-a_m)/(1+a_m))\right).\]

**Proof.** Since \(\hat{\theta}_n = \chi_{n,1}^\prime\) by an appeal to Theorem 3.1 of Sen (1961), we get that

\[(3.8) \quad b_n^2 E(\hat{\theta}_n - \theta)^2 + \frac{1+2a_m}{1+2a_m} = 2a_m(1-a_m) > 0.\]

Hence, to prove (3.7), by (3.5), it suffices to show that as \(n \to \infty\),
Towards this, we may note that

\[ E \left( \frac{2a_m}{n(1)} - Y_n(1) + Y_n(2) \right)^2 + 2a_m \left[ \frac{2a_m}{1 - 2a_m (1-a_m)/(1+a_m)} \right] . \]

For \( m = 0 \) (i.e., \( a_m = 1 \)), the second factor on the rhs of (3.7) is equal to 1, so that both \( \hat{\theta}_n \) and \( \theta_n^* \) have the same asymptotic variance, though their df's are not the same. For \( m \geq 1 \) (i.e., \( a_m \leq 1/2 \)), \( 2a_m (1-a_m)/(1+a_m) > 0 \) and is bounded from above by 1/3. Thus, from (3.2) and (3.7) we have that jackknifing reduces both the asymptotic bias and the asymptotic mean square to a fractional extent. This characteristic is different from the regular case where there is a complete reduction of asymptotic bias but no reduction of the asymptotic mean square.

From (2.6), (2.7) and (3.4), it follows that for a terminal contact of order \( m(\geq 0) \),

\[ (3.10) \quad \left| n^{-1} b_{n,m} v_n - \left\{ (Y_n(1) + Y_n(2))^a_m - Y_n(1) \right\} \right| \overset{p}{\to} 0, \quad \text{as} \quad n \to \infty . \]

Since \( (Y_n(1) + Y_n(2))^a_m - Y_n(1) + \left\{ (Y_1 + Y_2)^a_m - Y_1 \right\} \), where \( Y_1 \) and \( Y_2 \) are i.i.d.r.v. having the simple exponential df on \([0, \infty)\), \( n^{-1} b_{n,m} v_n \) either converges to a positive constant (when \( m = 0 \)) or goes to 0 (when \( m \geq 1 \)), it follows that either (for \( m = 0 \)) \( v_n \) has a non-degenerate asymptotic df.
or (for \( m \geq 1 \)) it goes to \(+\infty\), in probability as \( n \to \infty \). This characteristic is also different from the regular case where \( v_n \stackrel{P}{\to} \) a constant, as \( n \to \infty \). Nevertheless, for the studentized form, we have for a terminal contact of order \( m(\geq 0) \),

\[
T_n = n(\theta^*_n - \theta) / v_n = \frac{b_{n,m}(X_{n,1} - \theta)}{b_{n,m}(X_{n,2} - X_{n,1})} - (n-1)/n
\]

\[(3.11) \quad + \, o_p(1) \frac{a_m}{Y_1} \left\{ (Y_1 + Y_2) \frac{a_m}{Y_1} - Y_m \right\} - 1 , \]

so that noting that \( Y^* = Y_2 / Y_1 \) has the Fisher's variance-ratio distribution with degrees of freedom \((2,2)\), we have from (3.11) that

\[
(3.12) \quad \left[ 1 + (1 + T_n)^{-1} \right]^{m+1} - 1 \, \frac{a_m}{Y_1} Y^* = Y_2 / Y_1 .
\]

For \( m = 0 \), we have a simplified form

\[
(3.13) \quad T_n + 1 \, \frac{a_m}{Y_1} Y_2 = Y^* .
\]

Both (3.12) and (3.13) have important statistical applications.

4. A MODIFICATION OF \( \theta^*_n \)

We have observed in (3.2) that for \( m \geq 1 \), \( b_{n,m} E(\theta^*_n - \theta) \) does not converge to 0 as \( n \to \infty \). Let \( C_n \) be the sigma-field generated by \( X_{n,1}, \ldots, X_{n,n} \) and by \( X_{n+j}, j \geq 1 \) (so that \( C_n \) is non-increasing in \( n \)). Then, in the regular case, [cf. (2.11) of Sen (1977)], we have

\[
(4.1) \quad \theta^*_n - \hat{\theta}_n = (n-1) E((\hat{\theta}_n - \hat{\theta}_{n-1}) | C_n) .
\]
In our case, for \( m \geq 1 \), \( \frac{b_{n,m}}{c_{n,m}} E(\hat{\theta}_n - \hat{\theta}_{n-1}) = -a_m \frac{1 + a_m}{1 + a_m} + o(1) \), where as
\( \frac{b_{n,m}}{c_{n,m}} E(\hat{\theta}_n - \theta) = \frac{1 + a_m}{1 + a_m} + o(1) \), and thereby, we get the resulting bias in (3.2). To eliminate the, we may consider the modified estimator

\[
(4.2) \quad \hat{\theta}_{n,m} = \frac{1}{a_m} E\{(\hat{\theta}_n - \hat{\theta}_{n-1})|c_n\}
\]

\[
= X_{n,1} - (m+1)n^{-1}(n-1)(X_{n,2} - X_{n,1})
\]

In that case, we have

\[
(4.3) \quad \frac{b_{n,m}}{c_{n,m}} E(\hat{\theta}_{n,m} - \theta) \to 0 \text{ as } n \to \infty.
\]

Also, following the same line as in the proof of Theorem 1, we obtain that

\[
A_{m}^{**}(x) = \lim_{n \to \infty} \{b_{n,m}(\hat{\theta}_{n,m} - \theta) \leq x\}
\]

\[
= \begin{cases} 
\int_{0}^{\infty} \exp\left\{ - \frac{a_m^m}{a_m^m} \right\} dy, & -\infty < x \leq 0 \\
1 - \exp\{-x^{-m+1}\} + \int_{x^{-m+1}}^{\infty} \exp\left\{ - \frac{a_m^m}{a_m^m} \right\} dy, & 0 < x < \infty 
\end{cases}
\]

Also, following the line of proof of Theorem 2, we have

\[
\lim_{n \to \infty} E\left\{b_{n,m}(\hat{\theta}_{n,m} - \theta)^2\right\} = \left(2a_m \frac{2a_m}{2a_m} \right) \left(1 - \frac{2a_m}{1 + a_m} (m+1)a_m - 1\right) = 2a_m \frac{2a_m}{2a_m}
\]

\[
\left(4.5\right) \quad \lim_{n \to \infty} E\left\{b_{n,m}(\hat{\theta}_n - \theta)^2\right\} \geq \lim_{n \to \infty} E\left\{b_{n,m}(\hat{\theta}_n - \theta)^2\right\}.
\]

Thus, whereas \( \hat{\theta}_{n,m}^{**} \) eliminates bias to the desired extent, it fails to reduce the mean square. In this sense, it is similar to the case of \( \theta_{n}^{*} \) in the regular case. [Though \( A_{m}^{**} \) and \( A_{m}^{*} \) are not the same.]
Finally, for the studentized case, in (3.11)-(3.13), the only changes we need to make is to replace $T_n$ by $T_n + m$; the rest remains the same.

5. SOME REMARKS

We have so far considered the case of the lower end-point. The case of the upper end-point (if finite) follows on parallel lines. Secondly, in practical applications, when the form of $F$ is not specified but the order of terminal contact is assumed to be known [viz., $m = 0$ when $F$ is U-shaped or inverted J-shaped, etc.], the studentized form in (3.11)-(3.13) may most conveniently be used to provide a jackknife test for a null hypothesis $H_0: \theta = \theta_0$ (specified) or a confidence interval for the unknown $\theta$. For a symmetric df with both end-points finite, jackknifing of the extreme mid-range (for estimating or testing for the location of the df) can be made — the jackknife estimator corresponding to the smallest and the largest order statistic are also asymptotically independent.

REFERENCES

Sample extreme values are biased estimators of the end-points of a distribution, and hence, jackknifing is useful. However, the properties of jackknifing in such a case differ considerably from those in the regular case. These are studied here. Along with modification of jackknifing, some applications are also considered.