The Annals of Mathematical Statistics
1970, Vol. 41, No. 2, 667–672

THE ORDER OF THE MINIMUM VARIANCE IN A NON-REGULAR CASE

BY THOMAS POLFELDT

University of Lund

1. Introduction and summary. Let $f(y)$ be a probability density function on the real line, and $F(y)$ the corresponding distribution function. It is assumed that

$$F(y) = 0 \quad \text{for} \quad y \leqq 0, \qquad F(y) > 0 \quad \text{for} \quad y > 0. \tag{1}$$

Let θ be a location parameter, and let $X = (x_1, \cdots, x_n)$ denote a sample of n independent observations, with each x_i distributed according to $F(x-\theta)$. In this paper, we study the minimum variance of unbiased estimators $t = t(X)$ of θ, with special reference to the order, in n, of that variance. For example, if $F(y) = 1-e^{-y}(y > 0)$, the minimum variance is n^{-2} rather than of order n^{-1} as in regular cases.

We shall state conditions on $f(y)$ which determine this order. One of these is that $f(y)$ varies regularly at zero with exponent $c-1$ $(c > 0)$ (cf. [3], chapter 8, sect. 8–9). Under the conditions imposed, the smallest attainable variance order is n^{-1} if $c > 2$, but $(F^{-1}(n^{-1}))^2$ if $0 < c < 2$. The case $c = 2$ has special features. Since $F(y)$ varies regularly with exponent c, the minimum variance order will be $n^{-2/c}L(n^{-1})$ with slowly varying L $(0 < c < 2$; also true for $c = 2)$.

When $c > \frac{1}{2}$, the Chapman and Robbins inequality [2] is used to obtain a lower bound for the minimum variance. For $0 < c \leqq \frac{1}{2}$, a new inequality is used; we then restrict slightly the class of unbiased estimators.

The results carry over, of course, to distributions with $F(y) < 1$ for $y < 0$, $F(y) = 1$ for $y \geqq 0$. A generalization to biased estimators (or to mean square error) is straightforward, but some conditions on the bias function will be necessary.

Some questions recently raised by Blischke *et al.* [1] are answered by the theorems.

The conditions imposed here may probably be relaxed to some extent.

Notation. K and K' denote positive, finite constants.

If there exist K and K' such that $K < a(x)/b(x) < K'(|x| < x_0)$, we shall write $a(x) = \Omega(b(x))$ $(x \to 0)$. The qualification $(x \to 0)$ will often be omitted.

2. Theorem 1. *Let $t = t(X)$ be an unbiased estimator of θ. If*

(i) *there is an h_0 such that $H = \{h \mid 0 < h < h_0\} \subset \{h \mid h \neq 0, f(y) = 0$ implies $f(y-h) = 0\} \cap \{h \mid \theta + h$ is a possible parameter value$\} = \mathscr{H}$*

(ii) $f(ky)/f(y) \to k^{c-1}$ $(y \downarrow 0$, *all* $k > 0)$. *This defines the constant c.*

(iii) *$f(y)$ is continuous in $h \leqq y \leqq \eta$ ($\eta > 0$ is fixed, all $h \in H$)*

(iv) $\{f(y-h)/f(y)-1\}(y/h) = O(1)$ $(c \neq 2)$, $= \Omega(1)$ $(c = 2)$ (*all* y, $\lambda h \leqq y \leqq \eta$, *all* $h \in H$; $\lambda > 1$ *is fixed*)

(v) $0 < \lim_{h\to 0} h^{-2} \int_\eta^\infty \{f(y-h)/f(y)-1\}^2 f(y)\, dy < \infty$

(vi) $-\int_0^\infty y^2\, d(1-F(y))^m < \infty$ *for m large enough*

Received December 4, 1967; revised September 24, 1968.

Institute of Mathematical Statistics is collaborating with JSTOR to digitize, preserve, and extend access to *The Annals of Mathematical Statistics.*
www.jstor.org

and if

(2) $$G(x) = x^2 \int_{\lambda x}^{\eta} y^{-2} f(y)\, dy \qquad \textit{then}$$

$$\begin{aligned} \inf_t V_\theta(t) &= \Omega(n^{-1}) && c > 2 \\ &= \Omega((G^{-1}(n^{-1}))^2) && c = 2 \\ &= \Omega((F^{-1}(n^{-1}))^2) && \tfrac{1}{2} < c < 2. \end{aligned}$$

The proof relies on the following

LEMMA 1. *If* $f(y)$ *satisfies* (1) *and* (i)–(v) *of Theorem* 1, *then for* $h > 0$,

$$\begin{aligned} \phi(h) = \int_h^\infty (f(y-h))^2/f(y)\, dy &= 1+\Omega(h^2) && c > 2 \\ &= 1+\Omega(G(h)) && c = 2 \\ &= 1+\Omega(F(h)) && \tfrac{1}{2} < c < 2. \end{aligned}$$

The proof of Lemma 1 is deferred to the next section.

PROOF OF THEOREM 1. From Chapman and Robbins results in [2], it follows that

(3) $$V_\theta(t) \geqq h^2/\{(\phi(h))^n - 1\} \qquad (\text{any } h \in \mathscr{H}).$$

This is true for all t, so we can write $\inf_t V_\theta(t)$, and for all $h \in H$ because of (i). We insert the following $h \in H$ in (3):

(4) $$h \to 0 \,(c > 2), \qquad h = G^{-1}(n^{-1}) \quad (c = 2), \qquad h = F^{-1}(n^{-1}) (\tfrac{1}{2} < c < 2)$$

and apply Lemma 1. For $c > 2$, we then have $\inf V_\theta(t) \geqq Kn^{-1}$, while for $\frac{1}{2} < c \leqq 2$, the denominator of (3) is $\Omega(1)$, and so

$$\begin{aligned} \inf_t V_\theta(t) \geqq Kh^2 &= K(G^{-1}(n^{-1}))^2 && c = 2 \\ &= K(F^{-1}(n^{-1}))^2 && \tfrac{1}{2} < c < 2. \end{aligned}$$

Next, we find unbiased estimators t_0 with $V(t_0)$ of the desired order. Since $\inf V_\theta(t) \leqq V_\theta(t_0)$, the theorem will then be proved. For $c > 2$, (1) and (iii) imply the existence of y_1, $0 < y_1 < \eta$, with $f(y_1) > 0$ and $f(y)$ continuous at y_1. If $F(y_1) = p$, the 100pth percentile ζ_p of the sample is asymptotically normally distributed, with variance $p(1-p)\{f(y_1)n\}^{-1}$. Since $F(y) = F(x-\theta)$, $E(\zeta_p) = \theta + b_n$, and $t_0 = \zeta_p - b_n$ is unbiased with $V(t_0) = \Omega(n^{-1})$—the existence of mean and variance of t_0 and their convergence to θ and $p(1-p)\{f(y_1)n\}^{-1}$ follows from (vi) (cf. [7]).

In [10], the author has constructed an estimator that has the desired variance order, $(G^{-1}(n^{-1}))^2$, when $c = 2$.

For the case $\frac{1}{2} < c < 2$, we first note that (ii) implies $\lim F(ky)/F(y) = k^c$ ($y \to 0$, all $k > 0$) ([3] chapter 8, sect. 9, Theorem). When this condition and (1) are satisfied, Gnedenko [5] has shown that if $t_0 = \min(x_1, \cdots, x_n) - B_n$, where B_n is determined by $E(t_0) = \theta$, and if $a_n = F^{-1}(n^{-1})$, then, as $n \to \infty$, $\lim P(a_n^{-1}(t_0 + B_n - \theta) \leqq y) = 1 - \exp(-y^c)$ $(y > 0)$. Using (vi) and [7], we get

$$\lim a_n^{-2} V_\theta(t_0) = \lim V_\theta(a_n^{-1}(t_0 + B_n - \theta))$$
$$= \textstyle\int_0^\infty y^2\, d(1 - \exp(-y^c)) - (\int_0^\infty y\, d(1 - \exp(-y^c)))^2,$$

or $V_\theta(t_0) = \Omega(a_n^2) = \Omega((F^{-1}(n^{-1}))^2)$. The proof is complete.

3. Proof of Lemma 1. A regularly varying function (at zero) with exponent $c-1$ is defined by (ii). It can also be characterized by $f(y) = y^{c-1}R(y)$ (cf. [3] chapter 8, sect. 8), where $R(y)$ varies slowly, i.e. (ii) holds with $c - 1 = 0$. The following property (5) of slowly varying functions follows immediately from the representation $R(y) = a(y)\exp(\int_y^1 \varepsilon(t)/t\, dt)$, where $a(y) \to A \neq 0$, and $\varepsilon(y) \to 0$ $(y \to 0)$ ([3] chapter 8, sect. 9, Corollary). For each $\kappa > 0$, there is a $K > 0$ such that

$$(u/v)^\kappa < R(hu)/R(hv) < (u/v)^{-\kappa} \qquad (0 < u < v < K,\ \text{all } h \in H). \tag{5}$$

Since $(f(y-h))^2 = -(f(y))^2 + 2f(y-h)f(y) + (f(y-h) - f(y))^2$, we have

$$\begin{aligned}\phi(h) &= \textstyle\int_h^\infty (f(y-h))^2/f(y)\, dy\\ &= -(1 - F(h)) + 2 + \textstyle\int_h^\infty \{f(y-h)/f(y) - 1\}^2 f(y)\, dy\\ &= 1 + F(h) + I.\end{aligned} \tag{6}$$

The integral I is divided into

$$I = \textstyle\int_h^{\lambda h} + \int_{\lambda h}^{\eta} + \int_\eta^\infty = I_1 + I_2 + I_3; \tag{7}$$

$\lambda > 1$ and $\eta > 0$ must not depend on h. For I_1, we write $f(y) = y^{c-1}R(y)$. Applying (5), we find that I_1 takes on values in the interval $\int_h^{\lambda h}\{(1 - h/y)^{c-1\pm\kappa} - 1\}^2 f(y)\, dy$. Using the mean value theorem (relying on (iii)) and substituting $1 - h/y = t$, we find that this is $(1 < \lambda' < \lambda)$

$$f(\lambda' h) \textstyle\int_0^{1-\lambda^{-1}} (t^{c-1\pm\kappa} - 1)^2 h(1-t)^{-2}\, dt. \tag{8}$$

Because of (ii) and Theorem 1 in chapter 8, sect. 9 of [3], we have, for $c > \frac{1}{2}$, $I_1 = O(hf(\lambda' h)) = O(hf(h)) = O(F(h))$ (if $c \neq 1$, $(t^{c-1\pm\kappa} - 1)^2$ is strictly positive over an interval of positive length, and so we get $\Omega(F(h))$ instead of $O(F(h))$).

For I_2, we apply (iv) and the mean value theorem. Then,

$$\begin{aligned}I_2 &= \textstyle\int_{\lambda h}^{\eta} \{f(y-h)/f(y) - 1\}^2 (y/h)^2 h^2 y^{-2} f(y)\, dy\\ &= O(1) h^2 \textstyle\int_{\lambda h}^{\eta} y^{-2} f(y)\, dy = O(1) I_{20}\\ &= O(1)(h^2\{[y^{-2}F(y)]_{\lambda h}^{\eta} + 2\textstyle\int_{\lambda h}^{\eta} y^{-3}F(y)\, dy\}).\end{aligned} \tag{9}$$

If $c = 2$, condition (iv) gives $\Omega(1)$ instead of $O(1)$. But $F(y)\{yf(y)\}^{-1} \to c$, and so $2\int_{\lambda h}^{\eta} y^{-3}F(y)\, dy = 2C^{-1}\int_{\lambda h}^{\eta} y^{-2} f(y)\, dy$, with C near c if η is small enough. Thus,

$I_{20} = h^2[y^{-2}F(y)]_{\lambda h}^{\eta} + 2C^{-1}I_{20}$. If $c \neq 2$, the order of I_{20} follows from this equation, and so, from (9) and the definition (2),

$$I_2 = O(1)\Omega(F(h)+h^2) \qquad c \neq 2 \tag{10}$$
$$= \Omega(1)\Omega(G(h)) \qquad c = 2.$$

As to I_3, (v) implies $I_3 = \Omega(h^2)$.

From the last result and from (6), (9) and (10), the lemma follows.

4. Extension to $0 < c \leqq \frac{1}{2}$. The above proof breaks down at (8) for $c \leqq \frac{1}{2}$. The inequalities by Kiefer [6] and Fraser and Guttman [4] also give $V(t) \geqq 0$ if $c < \frac{1}{2}$ (proved in [8]). Restricting slightly the unbiased estimates to the class T, including only those t where $V_\theta(t)$ and $V_{\theta+\delta}(t)$ have the same order of n for small δ (cf. (12)), we prove

THEOREM 2. *Under the conditions of Theorem* 1,

$$\inf_{t \in T} V_\theta(t) = \Omega((F^{-1}(n^{-1}))^2) \qquad 0 < c \leqq \tfrac{1}{2}.$$

The proof relies on

LEMMA 2. *For all h such that $\theta+h$ is a possible parameter value,*

$$(1-p)V_\theta(t) + pV_{\theta+h}(t) \geqq h^2(Q^{-1}-p) \tag{11}$$

where $0 \leqq p \leqq 1$, *and* $Q = \int_{R^n} \{f(X;\,\theta+h)\}^2\{(1-p)f(X;\,\theta)+pf(X;\,\theta+h)\}^{-1}\,d\mu$; *$f(X;\theta)$ denotes the n-dimensional density of $(x_1, \cdots, x_n)$, and μ is Lebesgue-measure in R^n.*

PROOF. Applying the Cauchy–Schwarz inequality to

$$\int (t-\theta)\{(1-p)f(X;\theta)+pf(X;\theta+h)\}^{+\frac{1}{2}-\frac{1}{2}}f(X;\theta+h)\,d\mu = h,$$

we obtain

$$\{(1-p)V_\theta(t)+pV_{\theta+h}(t)+ph^2\}Q \geqq h^2,$$

whence the lemma.

Various generalizations of this lemma, e.g. using Kiefer's idea [6], are possible.

COROLLARY. *If $t \in T$, that is, if*

$$K' \leqq V_{\theta+\delta}(t)/V_\theta(t) \leqq K(\theta,\delta_0) = K \quad (\text{all } n, \text{all } |\delta| < \delta_0), \textit{ then} \tag{12}$$

$$V_\theta(t) \geqq h^2(Q^{-1}-p)\{1+Kp\}^{-1}. \tag{13}$$

PROOF OF THEOREM 2. In Q, the integration is in effect over $\Xi = \{X \mid \text{all } x_i > \theta+h\} = \{Y \mid \text{all } y_i > h\} = A \cup (\Xi - A) = \{Y \mid \text{all } y_i > h\lambda\} \cup (\Xi - A)$. Integrating over A, we omit $pf(X;\,\theta+h)$ of the integrand of Q. Then,

$$\int_A \leqq (1-p)^{-1}\int_A (f(X;\theta+h))^2/f(X;\theta)\,d\mu = (1-p)^{-1}\{\textstyle\int_{\lambda h}^{\infty}(f(y-h))^2/f(y)\,dy\}^n.$$

Rewriting the integrand as before (6) and introducing I_2 and I_3 of (7),

$$\int_A \leqq (1-p)^{-1}\{1+I_2+I_3+F(\lambda h)-2F(\nu h)\}^n$$

(where $\nu = \lambda - 1 > 0$). Integrating over $\Xi - A$, we omit $(1-p)f(X;\theta)$ of the integrand. Then,

$$\begin{aligned}\textstyle\int_{\Xi - A} \leqq p^{-1} \int_{\Xi - A} f(X;\theta+h)\,d\mu &= p^{-1} - p^{-1}\{\textstyle\int_{\lambda h}^{\infty} f(y-h)\,dy\}^n \\ &= p^{-1}(1-(1-F(\nu h))^n).\end{aligned}$$

Now, take $h = \nu^{-1}F^{-1}(n^{-1})$. For $0 < c \leqq \frac{1}{2}$, we know from (10) and the line following it that $I_2 = \Omega(F(h)) = \Omega F(\nu h))$; $I_3 = \Omega(h^2)$. Then there is a constant C, positive or negative, such that

$$\begin{aligned}Q = \textstyle\int_A + \int_{\Xi - A} &\leqq (1-p)^{-1}\{1 + CF(\nu h)\}^n + p^{-1}\{1-(1-F(\nu h))^n\} \\ &\leqq (1-p)^{-1}e^C + p^{-1}(1-e^{-2}).\end{aligned}$$

It will then always be possible to find a p, $0 < p < 1$, such that $Q^{-1} - p > K' > 0$. Thus, for this choice of h and p, we find from (11) and (13)

$$\inf_{t \in T} V_\theta(t) \geqq \nu^{-2}(F^{-1}(n^{-1}))^2 K' \{1 + Kp\}^{-1}.$$

The upper bound for $\inf V_\theta(t)$ is obtained by means of Gnedenko's result, as in the proof of Theorem 1. The proof is complete.

5. Examples and remarks. The conditions of the theorems are easily seen to be satisfied by the Weibull distribution, $F(x-\theta) = 1 - \exp\{-(x-\theta)^c\}$ $(x > \theta)$, with known c, and the Pearson type III distribution, $F(x-\theta) = \int_0^{x-\theta}(\Gamma(c))^{-1}y^{c-1}e^{-y}\,dy$ $(x > \theta)$ (c known). Condition (iv) is perhaps best checked by a series expansion. Since in both cases $F(y) = \Omega(y^c)$, $F^{-1}(n^{-1}) = \Omega(n^{-1/c})$, and so $\inf V_\theta(t) = \Omega(n^{-2/c})$ for $0 < c < 2$ (if $0 < c \leqq \frac{1}{2}$: for $t \in T$). When $c = 2$, we have $G(y) = \Omega(-y^2 \log y)$, and $\inf V_\theta(t) = \Omega((n \log n)^{-1})$. For $c > 2$, $\inf V_\theta(t) = \Omega(n^{-1})$.

If $f(y) > 0$ only over $(0, a)$, condition (i) is not fulfilled. If a modification as indicated in [1] page 51 is carried out, the present techniques will give results similar to the above.

It is interesting to note that the values (4) can be found, approximately, as (multiples of) the h's that give the supremum in the right member of (3) (easy for $c \neq 2$).

A more discursive presentation of the material of this paper (including a short remark on $c = 0$) is found in [8]. In [9], the order of $G(h)$ is calculated. The author is preparing papers on various extensions of the present results.

Acknowledgment. I am grateful to Professor Gunnar Blom for valuable advice and encouragement.

REFERENCES

[1] Blischke, W. R., Truelove, A. J., Mundle, P. B., and Johns, M. V. (1966). Estimation of the location parameters of the Pearson type III and Weibull distributions in the non-regular case and other results in non-regular estimation. Office of Aerospace Research Report No. ARL 66–0233.

[2] Chapman, D. G. and Robbins, H. (1951). Minimum variance estimation without regularity assumptions. *Ann. Math. Statist.* **22** 581–586.

[3] FELLER, W. (1966). *An Introduction to Probability Theory and its Applications* **2**. Wiley, New York.

[4] FRASER, D. A. S. and GUTTMAN, I. (1952). Bhattacharyya bounds without regularity assumptions. *Ann. Math. Statist.* **23** 629–632.

[5] GNEDENKO, B. V. (1943). Sur la distribution limite du term maximum d'une série aléatoire. *Ann. of Math.* **44** 423–453.

[6] KIEFER, J. (1952). On minimum variance estimators. *Ann. Math. Statist.* **23** 627–629.

[7] PICKANDS, J, III (1968). Moment convergence of sample extremes. *Ann. Math. Statist.* **39** 881–889.

[8] POLFELDT, T. (1967). Determination of the order of the minimum variance attainable when estimating the location parameter of a one-sided distribution. Technical Report No. 1967: 1, Department of Mathematical Statistics, Lund Institute of Technology.

[9] POLFELDT, T. (1969a). Integrating regularly varying functions with exponent –1. *SIAM J. Appl. Math.* **17** 904–908.

[10] POLFELDT, T. (1969b). Asymptotically best systematic statistics in a non-regular case. Dept. of Statistics Technical Report No. 62, Stanford Univ.