COMMUN. STATIST.-THEORY METH., 18(3), 1023-1030 (1989)

UNBIASED ESTIMATION IN TYPE II CENSORED SAMPLES FROM A ONE-TRUNCATION PARAMETER DENSITY

K. Krishnamoorthy
Department of Statistics
Temple University
Philadelphia, PA 19122

and

Vijay K. Rohatgi and Josef Blass
Department of Mathematics and Statistics
Bowling Green State University
Bowling Green, Ohio 43403-0221 USA

Key Words and Phrases: Type II censoring; minimum variance unbiased estimation; one-truncation parameter density; pivot; shortest length confidence interval.

ABSTRACT

We consider uniform minimum variance unbiased estimation of a U-estimable function when the sample is (singly) Type II censored and comes from a one-truncation parameter density $f(x;\theta) = h(x)\,q(\theta)$. An explicit expression for the estimator is derived. Shortest length confidence interval for $q(\theta)$ is obtained.

1. INTRODUCTION

Let $X_1, X_2, \ldots, X_n$ be independent identically distributed random variables with common probability density function (pdf) $f(x;\theta)$

and let $X_{1:n} < X_{2:n} < \dots < X_{n:n}$ be the corresponding set of order statistics. Suppose the sample is (singly) Type II censored so that one observes only $X_{1:n}, X_{2:n}, \dots, X_{r:n}$, $1 \le r \le n$. In this paper we consider uniform minimum variance unbiased (UMVU) estimation of a U-estimable function $g(\theta)$ when the (censored) sample comes from a one-truncation parameter pdf

$$f_1(x;\theta) = q_1(\theta)\, h_1(x), \quad a < \theta < x < b \tag{1.1}$$

or

$$f_2(x;\theta) = q_2(\theta)\, h_2(x), \quad a < x < \theta < b \tag{1.2}$$

where $-\infty \le a < b \le \infty$ are known, h_1, h_2 are positive absolutely continuous functions and, q_1, q_2 are everywhere differentiable. The case $r = n$, that is, when a complete sample is available was treated by Tate (1959) who showed that the UMVU estimator for g for the family $\{f_1\}$ is given by

$$\varphi(X_{1:n}) = g(X_{1:n}) - \frac{g'(X_{1:n})}{nq_1(X_{1:n})\, h_1(X_{1:n})}, \tag{1.3}$$

and that for the family $\{f_2\}$ is given by

$$\varphi(X_{n:n}) = g(X_{n:n}) + \frac{g'(X_{n:n})}{nq_2(X_{n:n})\, h_2(X_{n:n})}, \tag{1.4}$$

where $g'(\theta) = \partial g(\theta)/\partial\theta$.

2. RESULTS

We first consider the case when $X_{1:n}, X_{2:n}, \dots, X_{r:n}$ is a type II censored sample from pdf f_1 given in (1.1). The likelihood function is given by

$$L_1(\underset{\sim}{x};\theta) = \frac{n!}{(n-r)!}\left\{\prod_{j=1}^{r} h_1(x_{j:n})\right\} q_1^n(\theta)\left\{\int_{x_{r:n}}^{b} h_1(x)dx\right\}^{n-r} I(x_{1:n} > \theta) \tag{2.1}$$

where $\underset{\sim}{x} = (x_{1:n}, x_{2:n}, \dots, x_{r:n})$ and $I(A)$ denotes the indicator function of set A. It follows from (2.1) that $X_{1:n}$ is a minimal sufficient statistic that is complete and hence the UMVU estimator of any U-estimable function g is given by $\varphi(X_{1:n})$ defined in (1.3).

Next we consider the pdf $f_2(x;\theta)$ defined in (1.2). For convenience we write $f_2(x;\theta) = f(x;\theta) = q(\theta)\, h(X)$. In this case the likelihood function is given by

$$L(\underset{\sim}{x};\theta) = \frac{n!}{(n-r)!}\left\{\prod_{j=1}^{r} h(x_{j:n})\right\} q^n(\theta)\left\{\int_{x_{r:n}}^{\theta} h(x)dx\right\}^{n-r} I(x_{r:n} < \theta) \tag{2.2}$$

and it follows that $X_{r:n}$ is minimal sufficient with pdf

$$f_{r:n}(x) = n\binom{n-1}{r-1} q^n(\theta)\left\{\int_a^x h(u)du\right\}^{r-1}\left\{\int_x^\theta h(u)du\right\}^{n-r} h(x). \tag{2.3}$$

If $\mathcal{E}_\theta \varphi(X_{r:n}) = 0$ for all $\theta \in (a,b)$ then one sees easily on differentiation that $\varphi(x) = 0$ a.e. and hence $X_{r:n}$ is a complete sufficient statistic.

Let $g(\theta)$ be a U-estimable function. Then there exists a function φ such that

$$\begin{aligned} g(\theta) &= \mathcal{E}_\theta \varphi(X_{r:n}) \\ &= n\binom{n-1}{r-1} q^n(\theta)\int_a^\theta \varphi(x)\left\{\int_a^x h(u)du\right\}^{r-1}\left\{\int_x^\theta h(u)du\right\}^{n-r} h(x)dx \end{aligned} \tag{2.4}$$

and this φ is the (essentially) unique UMVU estimator by Lehmann-Scheffe theorem. We need a few simple facts. First

$$1 = \int_a^\theta q(\theta)\, h(x)dx$$

so that $q(\theta) = \left\{\int_a^\theta h(x)dx\right\}^{-1}$ and, moreover, on differentiation with respect to θ

$$q'(\theta) = \frac{\partial q(\theta)}{\partial \theta} = -q^2(\theta)h(\theta). \tag{2.5}$$

Next, we need a result on differentiation of a function defined by integrals. Let $\psi(x,t) \in C'$ for $a \leq t \leq b$, $A \leq x \leq B$ and suppose that $\Psi(x) = \int_{g(x)}^{h(x)} \psi(x,t)dt$. Then (see Widder (1961), p. 353)

$$\Psi'(x) = \int_{g(x)}^{h(x)} \frac{\partial \psi(x,t)}{\partial x} dt - \psi(x,g(x))\, g'(x) + \psi(x,h(x))\, h'(x). \tag{2.6}$$

We are ready to solve the integral equation (2.4). For $k = 0,1,\ldots,n - r$ set

$$I_k(\theta) = \int_a^\theta \varphi(x) \left[\int_a^x h(u)du\right]^{r-1} \left[\int_x^\theta h(u)du\right]^{n-r-k} h(x)dx.$$

Then (2.4) can be rewritten as

$$I_0(\theta) = \frac{(r-1)!\,(n-r)!}{n!}\, g(\theta)\, q^{-n}(\theta) = \frac{(r-1)!\,(n-r)!}{n!}\, s(\theta), \tag{2.7}$$

where $s(\theta) = g(\theta)\, q^{-n}(\theta)$. For each differentiable function $\omega(\theta)$ on (a,b), define an operator $\mathcal{D}$ by

$$(\mathcal{D}\omega)(\theta) = h^{-1}(\theta) \frac{d}{d\theta}\, \omega(\theta), \tag{2.8}$$

and for $k \geq 2$ define

$$(\mathcal{D}^k\omega)(\theta) = (\mathcal{D}(\mathcal{D}^{k-1}\omega))(\theta). \tag{2.9}$$

Then $\mathcal{D}$ defines a linear operator on (a,b).

Clearly

$$I_{n-r}(\theta) = \int_a^\theta \varphi(x) \left[\int_a^x h(u)du\right]^{r-1} h(x)dx,$$

and it follows that

$$\frac{d}{d\theta} I_{n-r}(\theta) = \varphi(\theta) \left[\int_a^\theta h(u)du\right]^{r-1} h(\theta).$$

In view of (2.8) we have

$$\varphi(\theta) = q^{r-1}(\theta)(\mathcal{D}I_{n-r})(\theta). \qquad (2.10)$$

Using (2.6) we see easily that

$$\frac{d}{d\theta} I_k(\theta) = (n - r - k)\, h(\theta)\, I_{k+1}(\theta),$$

and hence for $k = 0,1,\ldots,n - r - 1$

$$I_{k+1}(\theta) = \frac{1}{n - r - k} (\mathcal{D}I_k)(\theta). \qquad (2.11)$$

By the linearity of $\mathcal{D}$ we see from (2.11) that

$$I_{n-r}(\theta) = \frac{1}{(n - r)!} (\mathcal{D}^{n-r} I_0)(\theta)$$

and hence from (2.10) and (2.7), that

$$\varphi(\theta) = \frac{(r - 1)!}{n!} q^{r-1}(\theta)(\mathcal{D}^{n-r+1} s)(\theta).$$

We have thus proved the following result.

<u>THEOREM 1</u>: *The UMVU estimator of any U-estimable function* $g(\theta)$ *based on a type II censored sample from pdf* f *given in (1.2) is of the form*

$$\varphi(X_{r:n}) = \frac{(r - 1)!}{n!} q^{r-1}(X_{r:n})(\mathcal{D}^{n-r+1} s)(\theta), \qquad (2.12)$$

where $s(\theta) = g(\theta)q^{-n}(\theta)$ *and* $\mathcal{D}$ *is a linear operator defined by (2.8) and (2.9).*

Substituting $r = n$ in (2.12) we get (1.4).

<u>EXAMPLE 2.1</u>: Let $f(x;\theta) = 1/\theta$, $0 < x < \theta$ and $g(\theta) = e^{-\theta}$. In this case $s(\theta) = \theta^n e^{-\theta}$ and hence

$$(\mathcal{D}^{n-r+1} s)(\theta) = \sum_{k=0}^{n-r+1} (-1)^{n-r+1-k} \binom{n - r + 1}{k} \frac{n!}{(n - k)!} \theta^{n-k} e^{-\theta}.$$

It follows that the UMVU estimator of $g(\theta)$ is given by

$$\varphi(X_{r:n}) = (r-1)!(X_{r:n})^{r-1} \sum_{k=0}^{n-r+1} (-1)^{n-r+1-k} \binom{n-r+1}{k} \frac{1}{(n-k)!} \cdot (X_{r:n})^{n-k} e^{-X_{r:n}}.$$

EXAMPLE 2.2: In many applications $g(\theta) = q(\theta)$. In this case $s(\theta) = q^{-n+1}(\theta)$ and $(\mathcal{D}^{n-r+1}s)(\theta) = (n-1)!\, q^{2-r}(\theta)/(r-2)!$. It follows that for $r > 1$, $\varphi(X_{r:n}) = (r-1)\, q(X_{r:n})/n$ is the UMVU estimator of $q(\theta)$. It is easy to check that for $r > 2$

$$\mathcal{E}q^2(X_{r:n}) = \frac{n(n-1)}{(r-1)(r-2)} q^2(\theta)$$

so that

$$\mathrm{var}\{q(X_{r:n})\} = \frac{n(n-r+1)}{(r-1)^2\,(r-2)} q^2(\theta), \quad r > 2,$$

and

$$\mathrm{var}\{\varphi(X_{r:n})\} = \frac{n-r+1}{n(r-2)} q^2(\theta). \quad r > 2.$$

3. AN APPLICATION

As an application we find the shortest length confidence interval for $q(\theta)$ based on $q(X_{r:n})$. We show that the distribution of $Y = q(\theta)/q(X_{r:n})$ is independent of θ and hence Y is a pivot for $q(\theta)$ (or θ). Indeed from (2.3) and (2.5)

$$f_{r:n}(x) = n\binom{n-1}{r-1}\left[\frac{q(\theta)}{q(x)}\right]^{r-1}\left[1 - \frac{q(\theta)}{q(x)}\right]^{n-r} q(\theta)\ h(x)$$

for $a < x < \theta$. From (2.5) it is clear that q is a decreasing function of θ and hence $q(\theta) < q(x)$, that is, $0 < y < 1$. Also $\frac{dx}{dy} = [h(x)\, q(\theta)]^{-1}$. It follows that

$$f_Y(y) = n\binom{n-1}{r-1} y^{r-1}\,(1-y)^{n-r}, \quad 0 < y < 1$$

and $f_Y(y) = 0$ elsewhere.

TABLE 1

MULTIPLIERS FOR THE SHORTEST CONFIDENCE INTERVAL
$(\alpha_1 q(X_{r:n}), \alpha_2 q(X_{r:n}))$.

		1 - α = .95		1 - α = .99	
		α_1	α_2	α_1	α_2
n=5	r				
	2	.026032	.670141	.008306	.781987
	3	.146614	.853369	.082827	.917114
	4	.329843	.973972	.217962	.991699
n=10	r				
	2	.007293	.397927	.001848	.505378
	3	.046423	.522402	.022650	.618977
	4	.107179	.633594	.065843	.719031
	5	.181959	.732031	.124598	.804883
	6	.268429	.818484	.194886	.875125
	7	.367239	.893652	.281190	.934375
	8	.477620	.953594	.38098	.977344
	9	.602194	.992700	.494692	.998160
n=20	r				
	2	.002771	.217568	.000621	.289070
	3	.018725	.289070	.008399	.360986
	4	.041161	.353125	.022574	.425000
	5	.086972	.437500	.066134	.550000
	6	.125230	.505000	.094006	.600000
	7	.158064	.550000	.125331	.650000
	8	.196275	.600000	.162970	.752500
	9	.238025	.650000	.197412	.752500
	10	.298126	.750000	.223358	.750025
	11	.336510	.762000	.247290	.775000
	12	.328514	.752500	.169531	.800000
	13	.402162	.805000	.336403	.850000
	14	.468980	.855500	.406294	.900000
	15	.530490	.899500	.466496	.950000
	16	.584832	.928500	.514464	.955250
	17	.644450	.956250	.572727	.974766
	18	.710895	.981250	.638774	.991387
	19	.782510	.997214	.710895	.999371

It is now easy to construct a $1 - \alpha$ level shortest length confidence interval for $q(\theta)$ based on $q(X_{r:n})$. The confidence interval is given by $(\alpha_1 q(X_{r:n}), \alpha_2 q(X_{r:n}))$ where α_1, α_2 are determined simultaneously from

$$\int_{\alpha_1}^{\alpha_2} f_Y(y)\, dy = 1 - \alpha, \text{ and } f_Y(\alpha_1) = f_Y(\alpha_2).$$

In Table 1 we have numerically computed values of (α_1, α_2) for selected values of $1 - \alpha$, n and r. It should be noted that α_1 and α_2 satisfy

$$|f_Y(\alpha_1) - f_Y(\alpha_2)| < 10^{-6}$$

and

$$\left| \int_{\alpha_1}^{\alpha_2} f_Y(y)dy - (1 - \alpha) \right| < 10^{-6}.$$

BIBLIOGRAPHY

Tate, R. F. (1959). Unbiased estimation: functions of location and scale parameters, Ann. Math. statist. 30, 341-366.

Widder, D. (1961). Advanced Calculus. Prentice-Hall, Englewood Cliffs, NJ.

Received May 1988; Revised November 1988.

Recommended by S. Kocherlakota, University of Manitoba, Winnipeg, CANADA.

Refereed Anonymously.