Parametric estimation of exponential distribution in informative model of random censorship from both sides

A. A. Abdushukurov[1]***, D. R. Mansurov***[2]

[1]Moscow State University Tashkent branch, Tashkent, Uzbekistan, a_abdushukurov@rambler.ru
[2]Navoi State Pedagogical Institute, Navoi, Uzbekistan, mathematicianmd@gmail

Introduction. In biomedical studies of individuals for survival, in engineering tests of technical devices for reliability, there may be cases when the tested objects coming under observation after a certain random time after the start of testing. This phenomenon is called delayed entry or left random censoring. The random variable (r.v.) X that characterizes the lifetime of the tested object becomes to observation under the condition $X \geq L$. Here L is the moment when the object was placed under observation. In addition to this, r.v. X may also be censored from the right by some other r.v. Y. We are interested in r.v. X which will be subjected to random censorship from both sides by random vector $(L,\, Y)$. Let r.v.-s $L,\; X$ and Y are mutually independent with continuous distribution functions (d.f.) K , F and G respectively.

1. Informative model of random censorship from both sides. Let $\{(X_k,\, L_k,\, Y_k)\,,\; k \geq 1\}$ be a sequence of independent realizations of the triple $(X,\, L,\, Y)$ and

$$S^{(n)} = \{(\mathrm{Z}_i,\, \Delta_i)\,,\, i = 1,\, \dots,\, n\}$$

-the observed sample, where $\mathrm{Z}_i = \max\{L_i, \min\{X_i,\, Y_i\}\}$, $\Delta_i = \left(\delta_i^{(0)},\, \delta_i^{(1)},\, \delta_i^{(2)}\right)$, with $\delta_i^{(0)} = I\left(\min\left(X_i,\, Y_i\right) < L_i\right)$, $\delta_i^{(1)} = I\left(L_i \leq X_i < Y_i\right)$, $\delta_i^{(2)} = I\left(L_i \leq Y_i < X_i\right)$ and $I\left(A\right)$ standing for indicator of the event A. Note that in sample $S^{(n)}$ the number of observed r.v.-s X_i is equal to $\delta_1^{(1)} + \dots + \delta_n^{(1)}$. The statistical task is consist in estimating of d.f. F from a sample $S^{(n)}$. However, such a general statement of the estimating problem, d.f.-s K and G are considered as a nuisance. In this paper, we will investigate the evaluation of d.f. F in the case of informative censoring from two sources, when d.f.-s K and G functionally depend on F. To describe such a model by H and N we denote d.f.-s of r.v.-s Z_i and $V_i = \min\left(X_i,\, Y_i\right)$. Then it is easy to see, that

$$H\left(x\right) = K\left(x\right)N\left(x\right),\;\; N\left(x\right) = 1 - \left(1 - F\left(x\right)\right)\left(1 - G\left(x\right)\right),\;\; x \in \mathrm{R}^1. \qquad (1)$$

Assume that there are positive unknown parameters θ , β such that following representations are valid for all $x \in \mathrm{R}^1$:

$$\begin{cases} 1 - G\left(x\right) = \left(1 - F\left(x\right)\right)^{\theta}, \\ K\left(x\right) = \left(N\left(x\right)\right)^{\beta}, \end{cases} \qquad (2)$$

where the parameter β is responsible for the power of censoring from the left, and θ -from the right. In this model, the censoring parameters θ and β , which determine the deepness of censorship from both sides. From formulas (1) and (2), it is not difficult to derive the following representation for d.f. F :

$$1 - F(x) = \left[1 - (H(x))^{\lambda}\right]^{\gamma}, \quad x \in \mathrm{R}^1, \tag{3}$$

where $\lambda = \frac{1}{1+\beta}$, $\gamma = \frac{1}{1+\theta}$ and therefore, the closeness of the parameters λ and γ to 1, denotes the weakness of the censoring. Using representation (3), we can construct a semiparametric estimate for F over a sample $S^{(n)}$ by estimating a triple $(H(x), \lambda, \gamma)$:

$$F_n(x) = 1 - \left[1 - (H_n(x))^{\lambda_n}\right]^{\gamma_n}, \quad x \in \mathrm{R}^1. \tag{4}$$

2. Parametric estimation of exponential distribution. Now let's look at the case in model (2) $X_i \sim F(x, \alpha) = 1 - e^{-\frac{x}{\alpha}}$, $x \geq 0$, $\alpha > 0$, $Y_i \sim G(x, \alpha) = 1 - (1 - F(x, \alpha))^{\theta}$, $\theta > 0$ and $L_i \sim K(x, \alpha) = (N(x, \alpha))^{\beta}$, $\beta > 0$. Then in accordance with (3) we have $1 - e^{-\frac{x}{\alpha}} = 1 - \left[1 - (H(x))^{\lambda}\right]^{\gamma}$ and we get for $H(x)$ the expression $H(x) = \left(1 - e^{-\frac{x}{\alpha \cdot \gamma}}\right)^{\frac{1}{\lambda}}$. Differentiate this distribution over by x we have density of $H(x)$ as $h(x) = \frac{1}{\alpha\lambda\gamma} e^{-\frac{x}{\alpha\gamma}} \left(1 - e^{-\frac{x}{\alpha\gamma}}\right)^{\frac{1-\lambda}{\lambda}}$. Now we construct a maximum likelihood function for $h(x)$ using the sample from Z_i :

$$L(Z, \alpha) = \ln\left((\alpha\lambda\gamma)^{-n} \cdot e^{-\frac{\sum_{i=1}^n Z_i}{\alpha\gamma}} \left(\left(1 - e^{-\frac{Z_1}{\alpha\gamma}}\right)\left(1 - e^{-\frac{Z_1}{\alpha\gamma}}\right) \times \ldots \times \right.\right.$$
$$\left.\left. \times \left(1 - e^{-\frac{Z_n}{\alpha\gamma}}\right)\right)^{\frac{1-\lambda}{\lambda}}\right) = -n\ln(\alpha\lambda\gamma) - \frac{\sum_{i=1}^n Z_i}{\alpha\gamma} + \frac{1-\lambda}{\lambda}\sum_{i=1}^n \ln\left(1 - e^{-\frac{Z_i}{\alpha\gamma}}\right).$$

Now calculate $\frac{\partial L(Z, \alpha)}{\partial \alpha}$:

$$\frac{\partial L(Z, \alpha)}{\partial \alpha} = -\frac{n}{\alpha} + \frac{\sum_{i=1}^n Z_i}{\alpha^2\gamma} + \frac{1-\lambda}{\lambda}\sum_{i=1}^n \frac{-\frac{Z_i}{\alpha^2\gamma} e^{-\frac{Z_i}{\alpha\gamma}}}{1 - e^{-\frac{Z_i}{\alpha\gamma}}}.$$

We solve maximum likelihood equation $\frac{\partial L(Z, \alpha)}{\partial \alpha} = 0$ by α and get

$$\alpha = \frac{1}{\lambda\gamma n} \cdot \sum_{i=1}^n Z_i \cdot \left(1 - \frac{1-\lambda}{(H(Z_i))^{\lambda}}\right). \tag{5}$$

From (5), instead of (H, λ, γ) of supplying appropriate estimates $(H_n, \lambda_n, \gamma_n)$, we obtain an semiparametric estimate for the parameter α :

$$\alpha_n = \frac{1}{\lambda_n \gamma_n n} \cdot \sum_{i=1}^n Z_i \cdot \left(1 - \frac{1-\lambda_n}{(H_n(Z_i))^{\lambda_n}}\right). \tag{6}$$

From (6), we obtain an estimate $F(x, \alpha_n)$ for $F(x, \alpha)$. Now we draw the following graphs using numerical methods, in order to compare the above estimate (4) and estimate constructed using (5). In conclusion, we can say that the estimate $F(x, \alpha_n)$ would be very good (Figures 1-4).

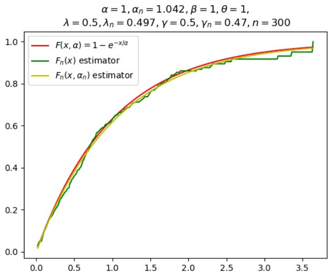

Figure 1: $\beta = 1, \quad \theta = 1, \quad n = 300$

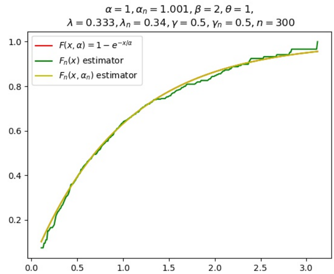

Figure 2: $\beta = 2, \quad \theta = 1, \quad n = 300$

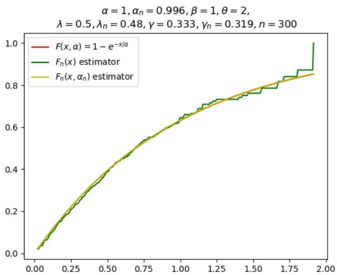

Figure 3: $\beta = 1, \quad \theta = 2, \quad n = 300$

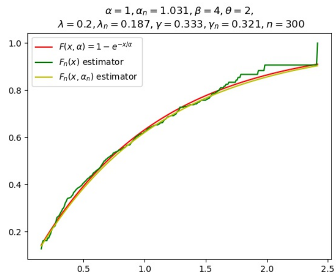

Figure 4: $\beta = 4, \quad \theta = 2, \quad n = 300$

References

1. A. A. Abdushukurov, *Estimation of probability density and intensity function of the Koziol-Green model of random cencoring*, Sankhya: Indian J. Statistics, Ser.A. v. 48, (1987), p. 150-168

2. A. A. Abdushukurov, *Random censorship model from both sides and independence test for it*, Report of Acad. Sci. Rep. Uz. Issue 11, (1994), p. 150-168(In Russian).