UDC 519.24

Noninformative Statistical Models: An Alternative Proof of the Cramér–Rao Inequality under Random Censoring

Abdushukurov Abdurakhim Akhmedovich
Moscow State University Tashkent Branch, Tashkent, Uzbekistan
a_abdushukurov@rambler.ru
Muradov Rustamjon Sobitkhonovich
Namangan State Technical University. Namangan, Uzbekistan
rustamjonmuradov@gmail.com

RESUME

We consider a model of random right censoring generated by a pair of independent random variables and the corresponding observed sample of minima and censoring indicators. For the case of a noninformative censoring distribution, we derive the Fisher information for the parameter of interest and present an alternative proof of the Cramér–Rao lower bound. The proof is based on a direct application of the Cauchy–Bunyakovsky (Cauchy–Schwarz) inequality to the likelihood of the censored sample, under suitable regularity conditions.

Key words: Cramér–Rao lower bound, noninformative censoring and Fisher information.

1. Introduction.

The Cramer-Rao Lower Bound (CRLB) is one of the most fundamental results in mathematical statistics, establishing a lower limit on the variance of unbiased estimators. Classical results assume that the underlying model is regular, meaning that the Fisher information matrix exists, is finite, and is strictly positive. However, many modern statistical problems-especially in likelihood-free inference, weakly identifiable models, nonregular exponential families, and nonparametric structures-violate these assumptions. In such noninformative models, the Fisher information may be zero, infinite, or undefined. This motivates a modern reconsideration of the CRLB applicable under minimal assumptions.

Classical CRLB theory originates from the foundational works of Rao (1945) and Cramer (1946), where the Fisher information plays the central role in describing the efficiency of unbiased estimators [3,4]. Subsequent developments in Kendall & Stuart (1977) and Lehmann & Casella (1998) extended the theory to multivariate settings. However, starting from the 1980s, several researchers pointed out the limitations of CRLB in nonregular models. Ibragimov and Has'minskii (1981) studied estimators in nonregular families where likelihood derivatives fail to exist. Hajek (1970) introduced local asymptotic minimax results under minimal smoothness. Van der Vaart (1998, 2000) developed semi-parametric CRLB extensions using tangent spaces. Ghosh, Basu & Martin (2016) analyzed CRLB failures in misspecified and robust statistical models (see, [5-10]).

More recently, the problem of zero-information models has gained attention in the context of likelihood-free inference (Sisson et al., 2018), Bayesian noninformative priors (Bernardo & Smith, 1994), weakly identifiable models (Drton, 2009), information-geometry (Amari & Nagaoka, 2000). Several authors propose generalized densities of bounded variation (Pfanzagl, 1990) or consider alternative differential structures, leading to extensions of CRLB based on generalized score functions [10-14].

Let $(\Omega, \mathcal{A})$ be a measurable space on which two independent random variables ξ and η are defined. Assume that their distribution functions depend on an unknown scalar parameter θ $(s = 1)$ and are given by

$$F(x;\theta) = \mathbb{P}_\theta(\xi < x), \qquad G(x;\theta) = \mathbb{P}_\theta(\eta < x), \qquad \theta \in \Theta \subset \mathbb{R}. \tag{1}$$

We assume that both distributions are absolutely continuous with respect to Lebesgue measure, with corresponding densities $f(x;\theta)$ and $g(x;\theta)$.

The statistical model is such that instead of observing the pair (ξ, η) directly, we observe the pair (Z, δ) defined by

$$Z = \min(\xi, \eta), \qquad \delta = I\{\xi = Z\} = I\{\xi \le \eta\}.$$

Thus the random variable of interest ξ is observed only when $\xi \le \eta$, i.e. when $\delta = 1$.

If (X_i, Y_i) denotes the realization of (ξ, η) in the i-th experiment, then after n independent repetitions we observe the sample

$$\{(Z_i, \delta_i)\}_{i=1}^n, \qquad Z_i = \min(X_i, Y_i), \quad \delta_i = I\{Z_i = X_i\}, \quad i = 1, \dots, n.$$

Each pair (Z_i, δ_i) takes values in the sample space $U \times \{0,1\}$, where U is the set of possible values of Z. We equip $U \times \{0,1\}$ with the product σ-algebra generated by sets of the form $A \times \{0\}$ and $B \times \{1\}$, where A, B are Borel subsets of U. Let $\{Q_\theta, \theta \in \Theta\}$ denote the family of distributions of (Z, δ) on $U \times \{0,1\}$, dominated by a product measure

$$v(dx, dy) = \Delta_y \, dx,$$

where Δ_y is the counting measure on $\{0,1\}$: $\Delta_y(\{0\}) = \Delta_y(\{1\}) = 1$.

For each observed pair (Z_i, δ_i) the likelihood contribution admits a density $k(x, y; \theta)$ with respect to v, which can be written as

$$k(x, y; \theta) = \{f(x;\theta)G(x;\theta)\}^y \{g(x;\theta)F(x;\theta)\}^{1-y}, \qquad x \in U, \ y \in \{0,1\}. \tag{2}$$

Introduce the following notation for derivatives with respect to θ:

$$\dot f(x;\theta) = \frac{\partial}{\partial\theta} f(x;\theta), \quad \dot g(x;\theta) = \frac{\partial}{\partial\theta} g(x;\theta),$$

$$\dot F(x;\theta) = \frac{\partial}{\partial\theta} F(x;\theta), \quad \dot G(x;\theta) = \frac{\partial}{\partial\theta} G(x;\theta),$$

and the corresponding (one-dimensional) score-type functions

$$\lambda(x;\theta) = \frac{\dot f(x;\theta)}{f(x;\theta)}, \qquad \mu(x;\theta) = \frac{\dot g(x;\theta)}{g(x;\theta)}.$$

For the joint density we write

$$\varphi(x, y; \theta) = \frac{\partial}{\partial\theta} \log k(x, y; \theta) = \frac{1}{k(x, y; \theta)} \frac{\partial}{\partial\theta} k(x, y; \theta).$$

We now recall an important special case of random right censoring, known as the *proportional intensities model* (PIM), also referred to as the Koziol–Green model (see, e.g., [1-4]).

Definition 1. The pair of distribution functions (F, G) (or equivalently, the pair (ξ, η)) is said to satisfy the proportional intensities model if there exists a positive constant $\beta > 0$ such that for all x in the support of ξ,

$$G(x;\theta) = \big(F(x;\theta)\big)^\beta. \tag{3}$$

In the literature the PIM is often called the Koziol–Green model. One of its key properties is that the observed variables Z and δ become independent.

Theorem 1. [3] *The pair (F, G) satisfies the proportional intensities model () if and only if the random variables*

$$Z = \min(\xi, \eta), \qquad \delta = I\{\xi = Z\}$$

are independent.

2. Noninformative Censoring Model and Fisher Information.

In the discussion above, the censoring distribution G was allowed to depend on the parameter θ, which leads to an *informative* censoring model. We now consider the case where G does *not* depend on θ, i.e. we assume that

$$F(x;\theta) \quad \text{depends on } \theta, \qquad G(x) \quad \text{does not depend on } \theta.$$

Such a model is called a *noninformative* (or noninformatively censored) model.

In this setting we present an alternative proof of the Cramér–Rao inequality. Assume that the distributions of ξ and η are given by $F(x;\theta)$ and $G(x)$, with a scalar parameter $\theta \in \Theta \subset \mathbb{R}$. The observed sample is

$$C^{(n)} = \{(Z_i, \delta_i)\}_{i=1}^{n}.$$

The joint density of $C^{(n)}$ with respect to the product measure

$$\nu_\theta(dx, dy) = \prod_{i=1}^{n} v(dx_i, dy_i),$$

on the sample space

$$\mathcal{Y}^{(n)} = \big(\mathbb{R} \times \{0,1\}\big)^n$$

is given by

$$k^{(n)}(x, y; \theta) = \prod_{i=1}^{n} k(x_i, y_i; \theta), \qquad (x, y) = (x_1, y_1, \ldots, x_n, y_n) \in \mathcal{Y}^{(n)},$$

where now

$$k(x, y; \theta) = \{f(x;\theta)G(x)\}^y \{g(x)F(x;\theta)\}^{1-y}, \qquad x \in \mathbb{R},\ y \in \{0,1\}. \tag{4}$$

Here G and g do not depend on θ, while F and f do, and

$$F(x;\theta) = \int_{(-\infty,x]} f(t;\theta)\, dt, \qquad g(x) = \frac{d}{dx} G(x).$$

The Fisher information of the sample $C^{(n)}$ is denoted $I_n(\theta)$ and has the form

$$I_n(\theta) = n\, I(\theta),$$

where

$$I(\theta) = \int_{\mathbb{R}} \left(\frac{\partial}{\partial\theta} \ln f(x;\theta) \right)^2 G(x)\, dF(x;\theta) + \int_{\mathbb{R}} \left(\frac{\partial}{\partial\theta} \ln F(x;\theta) \right)^2 F(x;\theta)\, dG(x). \tag{5}$$

We now impose the following regularity conditions on the family $\{Q_\theta, \theta \in \Theta\}$.

(C1) The support

$$\{x : 0 < F(x;\theta) < 1\}$$

does not depend on θ.

(C2) The derivatives $\partial f(x;\theta)/\partial\theta$ and $\partial F(x;\theta)/\partial\theta$ exist for all x and θ, and

$$\int_{\mathbb{R}} \frac{\partial f(x;\theta)}{\partial\theta} G(x)\, dx = \frac{\partial}{\partial\theta} \int_{\mathbb{R}} f(x;\theta) G(x)\, dx,$$
$$\int_{\mathbb{R}} \frac{\partial F(x;\theta)}{\partial\theta} dG(x) = \frac{\partial}{\partial\theta} \int_{\mathbb{R}} F(x;\theta)\, dG(x),$$

and, moreover, the following limit and differentiation under the integral are valid:

$$\lim_{\Delta\downarrow 0} \frac{1}{\Delta} \int_{\mathcal{Y}^{(n)}} \left[k^{(n)}(x, y; \theta + \Delta) - k^{(n)}(x, y; \theta) \right] \nu_\theta(dx, dy)$$
$$= \int_{\mathcal{Y}^{(n)}} \frac{\partial}{\partial\theta} k^{(n)}(x, y; \theta)\, \nu_\theta(dx, dy),$$
$$\int_{\mathcal{Y}^{(n)}} \left(\frac{\partial}{\partial\theta} \ln k^{(n)}(x, y; \theta) \right)^2 k^{(n)}(x, y; \theta)\, \nu_\theta(dx, dy) < \infty.$$

(C3) For all $\theta \in \Theta$ we have

$$0 < I(\theta) < \infty,$$

and for any unbiased estimator θ_n of θ we have

$$0 < d_n(\theta) = \mathbb{E}_\theta\big[(\theta_n - \theta)^2\big] < \infty.$$

We now formulate and prove the Cramér–Rao inequality for the noninformative censoring model.

Theorem 2. *Assume that conditions* (C1)–(C3) *hold. Then for any unbiased estimator* θ_n *of* θ *based on the sample* $C^{(n)}$*, and for all* $\theta \in \Theta$*, we have*

$$d_n(\theta) = \mathbb{E}_\theta\big[(\theta_n - \theta)^2\big] \;\geq\; \frac{1}{n\, I(\theta)}, \tag{6}$$

where $I(\theta)$ *is given by (5).*

Proof. Denote by ν_θ the probability measure of the sample $C^{(n)}$ on $\mathcal{Y}^{(n)}$, with density $k^{(n)}(x, y; \theta)$ with respect to the dominating measure. By unbiasedness we have

$$\int_{\mathcal{Y}^{(n)}} \theta_n(x, y)\, k^{(n)}(x, y; \theta)\, \nu_\theta(dx, dy) = \theta \qquad \text{for all } \theta \in \Theta.$$

Hence,

$$\int_{\mathcal{Y}^{(n)}} \big(\theta_n(x, y) - \theta\big)\, k^{(n)}(x, y; \theta)\, \nu_\theta(dx, dy) = 0.$$

Consider the function

$$\Delta(\theta, \Delta) = \int_{\mathcal{Y}^{(n)}} \big(\theta_n(x, y) - \theta\big) \big[k^{(n)}(x, y; \theta + \Delta) - k^{(n)}(x, y; \theta)\big]\, \nu_\theta(dx, dy).$$

By the above identity, $\Delta(\theta, \Delta) = 0$ for all sufficiently small Δ. Using condition (C2) and differentiating under the integral sign, we obtain

$$0 = \lim_{\Delta \downarrow 0} \frac{\Delta(\theta, \Delta)}{\Delta} = \int_{\mathcal{Y}^{(n)}} \big(\theta_n(x, y) - \theta\big)\, \frac{\partial}{\partial \theta} k^{(n)}(x, y; \theta)\, \nu_\theta(dx, dy).$$

Equivalently,

$$\int_{\mathcal{Y}^{(n)}} (\theta_n - \theta)\, \frac{\partial}{\partial \theta} \ln k^{(n)}(x, y; \theta)\, k^{(n)}(x, y; \theta)\, \nu_\theta(dx, dy) = 0.$$

Applying the Cauchy–Bunyakovsky (Cauchy–Schwarz) inequality to the last integral, we obtain

$$\left| \int_{\mathcal{Y}^{(n)}} (\theta_n - \theta)\, \frac{\partial}{\partial \theta} \ln k^{(n)}\, k^{(n)}\, d\nu_\theta \right|^2 \leq \left(\int_{\mathcal{Y}^{(n)}} (\theta_n - \theta)^2 k^{(n)}\, d\nu_\theta \right) \times \left(\int_{\mathcal{Y}^{(n)}} \left(\frac{\partial}{\partial \theta} \ln k^{(n)} \right)^2 k^{(n)}\, d\nu_\theta \right).$$

The left-hand side is zero, which yields

$$0 \leq d_n(\theta)\, I_n(\theta),$$

where

$$d_n(\theta) = \int_{\mathcal{Y}^{(n)}} (\theta_n - \theta)^2 k^{(n)}\, d\nu_\theta = \mathbb{E}_\theta\big[(\theta_n - \theta)^2\big],$$

and

$$I_n(\theta) = \int_{\mathcal{Y}^{(n)}} \left(\frac{\partial}{\partial \theta} \ln k^{(n)}(x, y; \theta) \right)^2 k^{(n)}(x, y; \theta)\, \nu_\theta(dx, dy) = n\, I(\theta).$$

Using condition (C3) we have $0 < I(\theta) < \infty$, so we can divide by $I_n(\theta)$ and obtain the inequality (6). The theorem is proved.

REFERENCES

1. A. A. Abdushukurov and L. V. Kim, Lower Cramér–Rao and Bhattacharyya bounds for randomly censored observations. *Journal of Soviet Mathematics*, **5**, (1987), 2171-2185.
2. A. A. Abdushukurov, Statistics of Incomplete Observations: Asymptotic Theory of Estimation for Nonclassical Models, *University*, *Tashkent*, (2009), 269p. (In russian)
3. H. Cramer, *Mathematical Methods of Statistics.* Princeton University Press, 1946.
4. C. R. Rao, Information and accuracy attainable in the estimation of statistical parameters, *Bulletin of the Calcutta Mathematical Society*, **37** (1945), 81-91.
5. E. L. Lehmann and G. Casella, *Theory of Point Estimation.* Springer, 1998.
6. A. W. van der Vaart, *Asymptotic Statistics.* Cambridge University Press, 1998.
7. I. A. Ibragimov and R. Z. Has'minskii, *Statistical Estimation: Asymptotic Theory.* Springer, 1981.
8. S. Amari and H. Nagaoka, *Methods of Information Geometry.* American Mathematical Society, 2000.
9. J. M. Bernardo and A. F. M. Smith, *Bayesian Theory.* John Wiley & Sons, 1994.
10. A. Ghosh, A. Basu, and R. Martin, Robust alternative to the Fisher information, *Statistical Papers*, **57** (2016), 239-252.
11. M. Drton, Likelihood ratio tests and weak identifiability, *Biometrika*, **96** (2009), 101-114.
12. J. Hajek, A characterization of limiting distributions in regular estimation problems, *Annals of Mathematical Statistics*, **41** (1970), 154-161.
13. J. Pfanzagl, *Estimation in Semiparametric Models.* Springer, 1990.
14. S. Sisson, Y. Fan, and M. Beaumont, *Handbook of Approximate Bayesian Computation.* CRC Press, 2018.

REZYUME

Bog‘liqsiz tasodifiy miqdorlar juftligi tomonidan hosil qilingan o‘ng tomondan tasodifiy senzuralangan model va unga mos ravishda kuzatiladigan minimumlar hamda senzuralanish indikatorlaridan iborat tanlanma ko‘rib chiqiladi. Senzuralanish taqsimoti informativ bo‘lmagan hol uchun baholanuvchi parametrga doir Fisher informatsiyasi chiqariladi va Cramer-Rao quyi chegarasining boshqacha isboti keltiriladi. Isbot senzuralangan tanlanmaning haqiqatga o‘xshashlik funksiyasiga Koshi–Bunyakovskiy (Koshi–Shvarts) tengsizliklarini bevosita qo‘llashga asoslanadi, bunda ba’zi regulyarlik shartlari bajarilishi talab qilinadi.

Kalit so‘zlar: Cramer-Rao quyi chegarasi, noinformativ senzuralanish, Fisher informatsiyasi.

РЕЗЮМЕ

Мы рассматриваем модель случайного цензурирования справа, порождённую парой независимых случайных величин, и соответствующую наблюдаемую выборку, состоящую из минимумов и индикаторов цензурирования. В случае неинформативного распределения цензурирования мы выводим функцию информации Фишера для интересующего параметра и представляем альтернативное доказательство нижней границы Крамера-Рао. Данное доказательство основано на прямом применении неравенства Коши-Буняковского (Коши-Шварца) к функции правдоподобия цензурированной выборки при выполнении соответствующих регулярностных условий.

Ключевые слова: нижняя граница Крамера-Рао, неинформативное цензурирование, информация Фишера.