برآوردگرهای UMVU در خانواده‌های لانه‌ای از توزیع‌ها

عبدالحمید رضائی رکن‌آبادی[1]

چکیده

اغلب وقتی برآوردگر $\delta = \delta(X_1, \ldots, X_n)$ در خانواده‌ی توزیع‌های $\mathcal{P}_1$، UMVU است، به غلط نتیجه‌گیری می‌شود که در کلاس $\mathcal{P}_2$ که $\mathcal{P}_2 \subset \mathcal{P}_1$ نیز برآوردگری UMVU خواهدبود. در این مقاله نادرستی این نتیجه‌گیری را مورد بحث قرار می‌دهیم.

فرض کنید $\mathcal{P}_1$ خانواده‌ی همه‌ی توزیع‌های پیوسته با امید ریاضی متناهی θ؛ $\mathcal{P}_2$ خانواده‌ی همه‌ی توزیع‌های متقارن حول θ و $\mathcal{P}_3$ خانواده‌ی همه‌ی توزیع‌های نرمال با میانگین θ و واریانس مجهول σ^2 باشند، بدیهی است :

$$\mathcal{P}_3 \subset \mathcal{P}_2 \subset \mathcal{P}_1 \cdot \qquad (۱)$$

اگر برآوردگر $\delta^* = \delta^*(X_1, \ldots, X_n)$ در کلاس $\mathcal{P}_1$ برآوردگری UMVU برای θ باشد، آیا این برآوردگر در کلاس‌های $\mathcal{P}_2$ و $\mathcal{P}_3$ نیز UMVU است؟ در برخورد اول چنین استنباط می‌شود که پاسخ این سؤال مثبت است. مثلاً اگر فردی کوتاه‌ترین فردِ بالغ در کشور بوده و ساکن مشهد باشد؛ قطعاً او کوتاه‌ترین فرد بالغ در استان خراسان و شهر مشهد نیز هست. در اینجا نیز $\mathcal{P}_1$ خانواده‌های $\mathcal{P}_2$ و $\mathcal{P}_3$ را در بردارد ؛ δ^* برآوردگر UMVU در $\mathcal{P}_1$، در $\mathcal{P}_2$ و $\mathcal{P}_3$ هم نا اریب است و ممکن است تصور شود در $\mathcal{P}_2$ و $\mathcal{P}_3$ نیز UMVU هست. اما واقعیت چنین نیست و پاسخ سؤال فوق همواره مثبت نیست.

اگر فرض کنیم Δ_1 و Δ_2 و Δ_3 به ترتیب کلاس همه‌ی برآوردگرهای نااریب θ (میانگین توزیع) ، در خانواده‌های $\mathcal{P}_1$ و $\mathcal{P}_2$ و $\mathcal{P}_3$ باشند، به ازای هر $\delta_1 \in \Delta_1$؛ چون δ_1 برآوردگری نااریب برای θ در $\mathcal{P}_1$ است، با فرض $\underset{\sim}{x} = (x_1, x_2, \ldots, x_n)$ داریم :

$$\forall f_1 \in \mathcal{P}_1 \quad E(\delta_1) = \int f_1(\underset{\sim}{x})\delta_1(\underset{\sim}{x})d\underset{\sim}{x} = \theta$$

و چون $\mathcal{P}_2 \subset \mathcal{P}_1$ پس ؛

$$\forall f_2 \in \mathcal{P}_2 \quad E(\delta_1) = \int f_2(\underset{\sim}{x})\delta_1(\underset{\sim}{x})d\underset{\sim}{x} = \theta$$

[1] عبدالحمید رضائی رکن‌آبادی، دانشکده علوم ریاضی دانشگاه فردوسی مشهد

یعنی $\delta_1 \in \Delta_2$. در نتیجه $\Delta_1 \subset \Delta_2$. بطور مشابه می توان نشان داد $\Delta_2 \subset \Delta_3$ و بنابراین :

$$\Delta_1 \subset \Delta_2 \subset \Delta_3 . \qquad (2)$$

در حقیقت بر اساس همین استدلال ثابت می‌شود که صرف نظر از نوع توزیع ، $\bar{X}$ برآوردگری نااریب برای θ (میانگین توزیع) است. در توزیع نرمال، نااریب بودن $\bar{X}$ را برای میانگین توزیع نرمال بدیهی می‌انگاریم.

مثال زیر نشان می‌دهد زیرمجموعه‌ها در رابطه‌ی (۲) ممکن است، به‌صورت محض باشند.

- **مثال ۱** — با فرض $\delta_0(X_1,\ldots,X_n) = \frac{X_{(1)}+X_{(n)}}{2}$ که در آن $X_{(1)}$ و $X_{(n)}$ به‌ترتیب آماره‌های ترتیبی اول و n ام دریک نمونه‌ی تصادفی n تایی از توزیع X است، ثابت می‌شود δ_0 برآوردگری نااریب برای مرکز تقارنِ هر توزیع پیوسته‌ی متقارن است ([۱]، تمرین ۵.۳.۱)؛ یعنی $\delta_0 \in \Delta_2$. اما اگر f_1 در $\mathcal{P}_1$ به‌صورت نمایی

$$f_1(x) = \begin{cases} \frac{1}{\theta}e^{-\frac{x}{\theta}} & x > 0 \\ 0 & x \leq 0 \end{cases}$$

فرض شود، به ازای $n = 3$، δ_0 برآوردگری اریب برای θ، میانگین توزیع فوق، است و اریبی آن $\frac{\theta}{12}$ است؛ یعنی $\delta_0 \notin \Delta_1$. پس لااقل عضوی در Δ_2 وجود دارد که در Δ_1 نیست.

براساس رابطه‌ی (۲) می‌توان گفت، اصولاً یافتن بهترین برآوردگر در کلاس Δ_3 (مربوط به خانواده توزیع‌های $\mathcal{P}_3$) مشکلتر از یافتن بهترین برآوردگر در زیرکلاسهای Δ_1 و Δ_2 (مربوط به توزیع‌های به ترتیب $\mathcal{P}_2$ و $\mathcal{P}_3$) است.

مثالهای زیر منفی بودن پاسخ سؤال مطرح شده در شروع بحث را ثابت می‌کنند:

- **مثال ۲** — $\bar{X}$ برآوردگر UMVU برای θ در خانواده‌ی $\mathcal{P}_1$ است ([۱]، صفحهٔ ۱۰۲) ولی $\bar{X}$ برآوردگر UMVU برای θ در خانواده‌ی $\mathcal{P}_2$ نیست زیرا به‌طور کلی ثابت می‌شود در کلاس $\mathcal{P}_2$ برآوردگر UMVU برای θ وجود ندارد ([۱]، صفحه‌ی ۱۳۷، تمرین ۲.۴.۴).

- **مثال ۳** — اگر $\mathcal{P}_4$ نیز خانواده‌ی توزیع‌های نرمال با میانگین معلوم θ_0 و واریانس مجهول σ^2 باشد؛ بدیهی است، $\mathcal{P}_4 \subset \mathcal{P}_3$. در خانواده‌ی $\mathcal{P}_3$ با استفاده از خواص خانواده‌ی نمایی ثابت می‌شود آماره‌ی $(\sum X_i, \sum X_i^2)$ بسنده‌ی کامل است. پس بنا به قضیه‌ی لهمن شفه برآوردگر نااریب $S^2 = \frac{1}{n-1}\sum(X_i - \bar{X})^2$ برای σ^2 که تابعی از این آماره‌ی بسنده‌ی کامل است، برآوردگری UMVU برای σ^2 است. اما بطور مشابه می‌توان نشان داد برآوردگر UMVU برای σ^2 در خانواده‌ی محدودتر $\mathcal{P}_4$ عبارت‌است از $S_0^2 = \frac{1}{n}\sum(X_i - \theta_0)^2$ که با S^2 متفاوت است.

هر چند در بحث بالا نتیجه گرفتیم UMVU بودن در خانواده‌ی بزرگتر، UMVU بودن در زیرخانواده‌ی کوچکتر از آن را در حالت کلی نتیجه نمی دهد، ولی مسلماً این امکان وجود دارد که یک برآوردگر UMVU در خانواده‌ی وسیع‌تر، در زیرخانواده‌ی کوچکتر از آن نیز UMVU باشد.
مثلاً $\bar{X}$ برآوردگری UMVU از میانگین توزیع، هم در خانواده‌ی $\mathcal{P}_1$ و هم در زیرخانواده‌ی کوچکتر $\mathcal{P}_3$ است.

مراجع

[1] Lehmann. E. L. *Theory of Point Estimation,* (1983).