به نام زندگی

تخمین MVUE

$$E_\theta\{\hat{\theta}(y)\} = \theta \longrightarrow \underset{\hat{\theta}}{\mathrm{Arg\,Min}}\ \mathrm{Var}_\theta(\hat{\theta}(y))$$

- آماره‌ی کافی sufficient statistic

- آماره‌ی کافی مینیمال Minimal suff. stat.

نکته‌ای که به آن اشاره کردیم، این است: به دست آوردن آماره‌های کافی مینیمال در حالت کلی، کار ساده‌ای نیست. علاوه بر این ممکن است برای برخی خانواده توزیع $\{P_\theta(y), \theta \in \Lambda\}$ آماره‌های کافی مینیمال وجود نداشته باشد. بنابراین به دنبال آماره‌های کافی مناسب (مفیدی) هستیم که بتوانیم برای تخمین θ از آن استفاده کنیم. قضیه‌ی زیر در این زمینه به ما کمک می‌کند.

قضیه فاکتورسازی Factorization

برای خانواده توزیع های $\{P_\theta(y);\ \theta \in \Lambda\}$ اگر $P_\theta(y)$ را به صورت زیر بازنویسی کنیم،

$$P_\theta(y) = g_\theta(T(y))\, h(y) \qquad \theta \in \Lambda,\quad y \in \Gamma$$

آنگاه $T(y)$ یک آماره کافی برای تخمین θ از روی مشاهدات y است (و بالعکس)

مثال: یک بردار تصادفی $\underline{X} = (X_1, X_2, \ldots, X_n)^T$ را در نظر بگیرید که در آن X_i ها متغیرهای تصادفی برنولی مستقل از هم با پارامتر θ هستند.

$$X_i \sim \text{Bernolli}(\theta) \quad , \quad X_1 \perp\!\!\!\perp X_2 \perp\!\!\!\perp \cdots \perp\!\!\!\perp X_n$$

یک آماره‌ی کافی برای تخمین θ از روی مشاهدات x، به دست بیاورید.

$$P_\theta(\underline{x}) = \theta^t (1-\theta)^{n-t}$$

که در آن t تعداد مکان‌های از بردار x است که مؤلفه آن برابر '1' است $:= \sum_{i=1}^{n} x_i$

$$P_\theta(\underline{x}) = \theta^{t}(1-\theta)^{n-t} \equiv g_\theta(T(\underline{x}))\,\underbrace{h(\underline{x})}_{1}$$

$$\Longrightarrow g_\theta(T(\underline{x})) \equiv \theta^{\overset{T(\underline{x})}{\sim}\; t}(1-\theta)^{n-\overset{T(x)}{\sim}\; t} \quad , \quad t \equiv T(\underline{x})$$

$\sum_{i=1}^{n} x_i$ که معادل تعداد '1' در بردار $\underline{x}$ است یا معادل

$$\Longrightarrow T(\underline{x}) = \sum_{i=1}^{n} x_i$$

ما این مقدمه به دنبال پیدا کردن جواب تخمین MVUE هستیم. برای این منظور مفاهیمی را بیان خواهیم کرد که از کنار هم گذاشتن آنها به پاسخ مورد نظر می‌رسیم

اولین مطلب با یک قضیه بیان می‌شود.

همان تابعی از θ که می‌خواهیم تخمین بزنیم

قضیه Rao - Blackwell

اگر $\hat{g}(\underline{y})$ یک تخمین‌گر بدون بایاس از $\underbrace{g(\theta)}$ باشد و $T(\underline{y})$ یک آماره کافی کامل روی کلیه توابع احتمال $\{P_\theta(y) ; \theta \in \Lambda\}$ باشد، آنگاه

$$\tilde{g}(T(\underline{y})) = E_\theta \left\{ \hat{g}(\underline{y}) \mid T(\underline{y}) = T(\underline{y}) \right\}$$

$$= E_{y|T,\theta} \left\{ \hat{g}(\underline{y}) \mid T(\underline{y}) = t(\underline{y}) \right\}$$

چون $T(\underline{y})$ یک آمارگان کافی است $p(y|T,\theta)$ تابعیت از θ ندارد.

بنابراین $E_\theta\{\hat{g}(\underline{y}) \mid T(\underline{y}) = T(\underline{y})\}$ نیز تابعیت از θ ندارد.

نیز یک تخمین‌گر بدون بایاس از $g(\theta)$ است و واریانس آن کمتر یا مساوی با واریانس

$\hat{g}(\underline{y})$ است یعنی داریم

$$Var_\theta\left(\tilde{g}(T(\underline{y}))\right) \leq Var_\theta\left(\hat{g}(\underline{y})\right)$$

تساوی زمانی برقرار است که داشته باشیم

$$P_\theta\left\{\tilde{g}(T(\underline{y})) = \hat{g}(\underline{y})\right\} = 1$$

یا $\tilde{g}(T(\underline{y})) \overset{w.p.1}{=} \hat{g}(\underline{y})$

(تساوی با احتمال 1)

$$P_r\left\{\tilde{g}(T(\underline{y})) = \hat{g}(\underline{y}) \mid \theta = \theta\right\} = 1$$

Almost every where

(اثبات قضیه به عنوان تمرین برای مطالعه)

$$E\, g(X,Y) = E_{X|\theta}\Big\{ \overbrace{E_{Y|X,\theta}\{ g(X,Y) \mid X=x \}}^{\text{تابعی از } X=x} \Big\}$$

راهنمایی: توجه به رابطه زنجیره‌ای

$$= E_{Y|\theta}\Big\{ \overbrace{E_{X|Y,\theta}\{ g(X,Y) \mid Y=y \}}^{\text{تابعی از } Y=y} \Big\}$$

$$E X^2 \geq E^2 X \quad \longrightarrow \quad E X^2|Y \geq E^2 X|Y$$

توجه به نامساوی شوارتز

بنابراین از قضیه Rao-Blackwell می‌توان نتیجه گیری کرد که با داشتن یک تخمین گر بدون بایاس $\hat{g}(\underline{y})$ از $g(\theta)$ و یک آماره‌ی کافی مناسب (

می‌توانیم واریانس تخمین را با به دست آوردن تخمین گر ($\bar{g}(T(\underline{y}))$ کاهش دهیم

$$\bar{g}(T(\underline{y})) = E_{\theta}\left\{ \hat{g}(\underline{y}) \,\middle|\, T(\underline{y}) = T(\underline{y}) \right\}$$

$T(\underline{y})$ را با قضیه فاکتورسازی می‌توان به دست آورد.

همچنین می‌توان گفت که: اگر $T(y)$ یک آماره کافی (مناسب) روی مجموعه خانواده توزیع‌های $\{P_\theta(y);\ \theta \in \Lambda\}$ باشد و فقط یک تابع از $T(y)$ وجود داشته باشد که تخمین‌گر بدون بایاس از $g(\theta)$ باشد، در این صورت این تخمین‌گر همان MVUE است.

(این تابع منحصر به فرد از $T(y)$ را با $g^*(T(y))$ نمایش می‌دهیم)

اثبات: اگر $\hat{g}(\underline{y})$ یک تخمین‌گر بدون اریبی دلخواه از $g(\theta)$ باشد.

آنگاه طبق قضیه Rao-Blackwell،

$$\bar{g}(T(\underline{y})) = E_{\theta}\{\hat{g}(\underline{y}) \mid T(\underline{y}) = T(\underline{y})\}$$

نیز یک تخمین‌گر بدون اریبی از $g(\theta)$ است که واریانس آن از $\hat{g}(\underline{y})$ کمتر است.

از طرف دیگر $\bar{g}(T(\underline{y}))$ تابعی از $T(\underline{y})$ است.

از طرف دیگر می‌دانیم که $g^*(T(\underline{y}))$ تنها تابع از $T(\underline{y})$ است که تخمین بدون اریبی $g(\theta)$ را به دست می‌دهد.

بنابراین باید داشته باشیم $g^*(T(y)) \equiv \bar{g}(T(y))$ که واریانس آن نیز از

هر تخمین گر دلخواه بدون اریب $\hat{g}(y)$ کمتر است $\leftarrow$ MVUE

* برای اینکه مراحل به دست آوردن تخمین MVUE را کامل کنیم به دنبال تابع منحصر به فرد از $T(y)$ هستیم. برای رسیدن به این تابع لازم است مفهومی را مطرح کنیم.

اولین مفهوم، مفهوم کامل بودن یک خانواده توزیع است.

* مفهوم کامل بودن Completeness

یک مجموعه از خانواده ی توزیع ها به صورت $\{P_\theta(\underline{y});\theta\in\Lambda\}$ را کامل
می گوییم اگر عبارت

$$\forall\theta\in\Lambda;\quad E_\theta\{f(y)\}=0$$

تنها نتیجه ای که به دست می دهد، این باشد که

$$f(y)=0\qquad\forall\theta\in\Lambda$$

برای برقراری تساوی، تساوی با احتمال 1 نیز کافی است یعنی

$$P_\theta\{f(\underline{y})=0\}=1 \quad \text{یا} \quad f(\underline{y}) \overset{w.p.1}{=\!=} 0$$

$$\forall\theta\in\Lambda$$

مفهوم کامل بودن یک خانواده از توزیع‌ها، مشابه به مفهوم کامل بودن یک مجموعه از بردارهای $V=\{\underline{v}_1,\underline{v}_2,\dots,\underline{v}_n\}$ در یک فضای برداری N بعدی است.

یک مجموعه از بردارهای $V=\{\underline{v}_1,\dots,\underline{v}_n\}$ را یک مجموعه متعامد کامل می‌گوییم اگر

هیچ برداری در این فضای n بعدی وجود نداشته باشد که بر اعضای V عمود باشد اما عضو V نباشد.

$$\exists\ \underline{y} \in \mathbb{R}^n \ ; \quad \langle \underline{y}, \underline{v}_i \rangle = 0 \implies \underline{y} = \underline{0}$$

تنها جواب $\underline{y} = \underline{0}$ باشد

مثال: نشان دهید که خانواده توزیع های دو جمله ای با پارامتر θ، کامل هست.

$$P_\theta(y) = \binom{n}{y} \theta^y (1-\theta)^{n-y} , \quad 0 \leq \theta \leq 1$$

$$\Lambda = [0,1] , \quad \theta \in \Lambda$$

$y \in \{0, 1, \ldots, n\} = \Gamma$

برای اینکه نشان دهیم خانواده توزیع های $\{P_\theta(y), \theta \in \Lambda\}$ کامل است

نشان می دهیم اگر برای یک تابع دلخواه $f(y)$ داشته باشیم $E_\theta\{f(y)\} = 0$

تنها جوابی که برای $f(y)$ وجود دارد، $f(y) = 0$ است

$$E_\theta\{f(y)\} = \sum_y f(y)\, P_\theta(y) = 0 \quad \forall \theta \in \Lambda$$

$$= \sum_{y=0}^{n} f(y) \binom{n}{y} \theta^{y} (1-\theta)^{n-y} = 0$$

$$= (1-\theta)^{n} \sum_{y=0}^{n} f(y) \binom{n}{y} \Bigg(\underbrace{\frac{\theta}{1-\theta}}_{x}\Bigg)^{y} = 0$$

$$= (1-\theta)^{n} \sum_{y=0}^{n} \underbrace{f(y) \binom{n}{y}}_{a(y) = x^{y} \text{ ضریب}} x^{y} = 0$$

$$a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n = 0$$

$x = f(a_0, \ldots, a_n)$ جواب‌های معادله = ریشه‌های معادله

یا $a_1 = a_2 = \cdots = a_n = 0$

$$E_\theta\{f(Y)\} = (1-\theta)^n \sum_{y=0}^{n} a_y\, x^y = 0 \quad \forall \theta \in \Omega$$

$E_\theta\{f(Y)\}$ یک چندجمله ای درجه ی nام برحسب x با ضرایب a_y

است. بنابراین n جواب مختلف برای x برحسب $\{a_y\}_{y=0}^{n}$ برای آن

وجود دارد. اما این جوابها قابل قبول نیستند چون $x = \frac{\theta}{1-\theta}$

یک پارامتر آزاد در این معادله نیست باید $E_\theta\{f(Y)\} = 0 \quad \forall \theta \in \Omega$

$$\xRightarrow[\text{بدیهی}]{\text{فقط جواب}} \quad a_y = 0 \qquad \forall y \in \{0, 1, \ldots, n\} = \Gamma$$

$$a_0 = a_1 = \cdots = a_n = 0$$

$$a_y = \binom{n}{y} f(y) = 0 \qquad \forall y \in \{0, 1, \ldots, n\}$$

$$\Rightarrow \quad f(y) = 0 \quad , \forall y \in \{0, 1, \ldots, n\} \, , \forall \theta \in \Lambda$$

$$\Rightarrow \quad P_\theta\{f(y) = 0\} = 1 \quad ! \quad f(y) \overset{w.p.1}{=\!=} 0$$

* آماره‌های کافی کامل (تعریف)

اگر خانواده توزیع‌های $\{P_\theta(\underline{y}) ; \underline{\theta} \in \Lambda\}$ یک خانواده توزیع‌های کامل باشند و $T(\underline{y})$ یک آماره کافی روی مجموعه $\{P_\theta(\underline{y}) ; \underline{\theta} \in \Lambda\}$ باشد، آنگاه $T(\underline{y})$ را یک آماره کافی کامل می‌گوییم.

به کمک این مفاهیم می‌خواهیم به تابعی منحصر به فرد از $T(\underline{y})$ که کمترین پراش را بین برآوردهای نااریب $g(\theta)$ به دست می‌دهد، برسیم.

فرض کنیم $T(\underline{y})$ یک آماره کافی روی مجموعه ی کامل $\{P_\theta(\underline{y});\ \underline{\theta} \in \Lambda\}$

باشد (یعنی $T(\underline{y})$ یک آماره کافی کامل) و فرض می کنیم

$$E_\theta\{|Y|\} < \infty \quad , \quad \forall \theta \in \Lambda$$

و تابع $f(y)$ را به صورت زیر تعریف می کنیم

در حالت کلی تابع $T(y)$

$$f(y) = y - E_\theta\{Y \mid T(Y) = T(y)\}$$

چون $T(y)$ آماره ای کافی است تابعیت از θ ندارد

$$f(y) = y - E_{y|T,\theta}\{Y / T(Y) = T(y)\}$$

می‌خواهیم نشان دهیم که

$$E_\theta\{f(y)\} = 0$$

است.

یعنی $\boxed{E_{y|\theta}\{f(y)\} = 0}$

$$E_\theta\{f(y)\} = E_\theta\{y - E_\theta\{y \mid T(y) = T(y)\}\}$$

$$= E_\theta\{y\} - E_\theta\{E_\theta\{y \mid T(y) = T(y)\}\}$$

$$= E_{y|\theta}\{y\} - E_{y|\theta}\{\underbrace{E_{y|T,\theta}\{y \mid T(y) = T(y)\}}_{\text{تابعی از } T(y)=T(y)}\}$$

یادآوری: قضیه اساسی امید ریاضی

$$Z = g(X)$$

$$EZ = \int z\, f_Z(z)\, dz = \int g(x)\, f_X(x)\, dx$$

$$\Rightarrow \; E_\theta\{f(y)\} = E_\theta\{y\} - E_\theta\{E_\theta\{y \mid T(y) = T(y)\}\}$$

$$= E_\theta\{y\} - E_{T|\theta}\{E_{y|T,\theta}\{y \mid T(y) = T(y)\}\}$$

رابطه زنجیره‌ای

$$= E_\theta\{y\} - E_\theta\{y\} = 0$$

$$\Rightarrow \quad E_\theta\{f(y)\}=0 \qquad \forall \theta \in \Lambda$$

با توجه به اینکه مجموعه $\{P_\theta(y), \theta \in \Lambda\}$ کامل است، نتیجه‌ای که می‌توان گرفت، این است که

$$P_\theta\{f(y)=0\}=1$$

یا:

$$f(y) \overset{w.p.1}{=\!=} 0$$

$$\Rightarrow \quad y - E_\theta\{y \mid T(y)=T(y)\}=0$$

$$\Rightarrow \quad y = \underbrace{E_\theta \{ y \mid T(y) = T(y) \}}_{\text{تابعی از } T(y)=T(y) \text{ است}}$$

به عبارت دیگر، به این نتیجه رسیدیم که y تابعی از $T(y)$ است.

از طرف دیگر می‌دانیم که $T(y)$ نیز تابعی از y است.

$\Leftarrow$ $T(y)$ یک تابع یک به یک است

$\Longleftarrow$ برای هر تجربه، توابع احتمال کامل، آماره‌های کافی مینیمال وجود ندارد.

به عبارت دیگر $T(y)$ که از قضیه فاکتورسازی به دست می‌آید، برای هر تجربه توابع احتمال کامل، یک آماره‌ی مناسب برای تخمین $g(\theta)$ است.

آخرین مرحله در به دست آوردن $MVUE$، این است که نشان دهیم $g^*(T(y))$ تنها تابعی از $T(y)$ است که یک تخمین بدون اریبی از $g(\theta)$ به دست می‌دهد، در نتیجه $MVUE$ است.

با بیان دیگر، اگر نشان دهیم که $g^*(T(y))$ و $\tilde{g}(T(y))$

یکسان هستند، منحصر به فرد بودن $g^*(T(y))$ مشخص می‌شود.

برای این منظور، از خاصیت کامل بودن مجموعه توابع احتمال $\{P_\theta ; \theta \in \Lambda\}$

به صورت زیر کمک می‌گیریم:

$$f(y) = g^{*}(T(y)) - \tilde{g}(T(y)) \quad \forall \theta \in \Lambda$$

$$E_{\theta}\{f(y)\} = E_{\theta}\{g^{*}(T(y)) - \tilde{g}(T(y))\} \quad \forall \theta \in \Lambda$$

$$= E_{y|T,\theta}\{g^{*}(T(y)) - \tilde{g}(T(y)) \mid T(y),\theta\} \quad \forall \theta \in \Lambda$$

$$= \underbrace{E_{\theta}\{g^{*}(T(y))\}}_{g(\theta)} - \underbrace{E_{\theta}\{\tilde{g}(T(y))\}}_{g(\theta)} = 0 \quad \forall \theta \in \Lambda$$

خانوادهٔ توزیع‌های $\Longrightarrow$ کامل

$$g^*(T(y)) \overset{w.p.1}{=} \tilde{g}(T(y))$$

بنابراین $g^*(T(y))$ منحصر به فرد است.

به این ترتیب روند به دست آوردن تخمین MVUE برای تخمین $g(\theta)$

از روی مشاهدات y به صورت زیر خواهد بود

۱- به دست آوردن یک آماره‌ی کافی کامل $T(y)$ برای خانواده‌ی $\{P_\theta ; \theta \in \Lambda\}$

(قضیه‌ی فاکتورسازی در به دست آوردن $T(y)$ کمک کننده است - $\{P_\theta , \theta \in \Lambda\}$ یک کمیه‌ی کامل است)

۲- پیدا کردن یک تخمین بدون بایاس $\hat{g}(y)$ از $g(\theta)$

(خلاقیت حل مسأله + نرم آماره‌ی کافی)

۳ـ به دست آوردن تخمین MVUE با کمک رابطه‌ی زیر

$$g^{*}(T(\underline{y})) \equiv \tilde{g}(T(\underline{y})) = E_{\theta}\left\{ \hat{g}(y) \,\middle|\, T(\underline{y}) = T(\underline{y}) \right\}$$

$$= E_{y|T,\theta}\left\{ \hat{g}(y) \,\middle|\, T(\underline{y}) = T(\underline{y}) ,\theta = \theta \right\}$$

(عملیات E ریاضی کاسبی است امید ریاضی شرطی)