به نام زندگی

تخمین MVUE

همان طور که دیدیم، روند به دست آوردن تخمین MVUE بر این

پایه استوار است که مجموعه خانواده توزیع احتمال $\{P_\theta, \theta \in \Lambda\}$

یک مجموعه کامل باشد. در ادامه مطرح می کنیم که گروه زیادی از توابع احتمالی

که در عمل با آنها سروکار داریم، کامل هستند. برای این منظور خانواده توزیع های
نمایی را معرفی می کنیم.

* خانواده توزیع های نمایی Exponential family of densities

اگر توابع $c(\cdot)$، $h(\cdot)$، $Q_1(\cdot)$، ...، $Q_m(\cdot)$، $T_1(\cdot)$، ...، $T_m(\cdot)$
توابعی حقیقی باشند و داشته باشیم،

$$p_{\underline{\theta}}(\underline{y}) = P(\underline{y}|\underline{\theta}) = h(\underline{y}) \exp\left(\sum_{i=1}^{m} T_i(\underline{y})\, Q_i(\underline{\theta})\right) c(\underline{\theta})$$

$$\forall \underline{\theta} \in \Lambda, \quad \forall \underline{y} \in \Gamma$$

آنگاه $\{P_{\underline{\theta}}(\underline{y}), \underline{\theta} \in \Lambda\}$ را یک خانواده توزیع نمایی گوییم.

همان طور که دیده می‌شود، گروه زیادی از توابع احتمالی که با آنها سر و کار داریم، از خانواده توزیع‌های نمایی هستند.

در جمله ای هندسی برنولی

به عنوان مثال توزیع های گوسی ، رایلی ، نمایی ، لاپلاس ، پواسن و ... در حالت یک بعدی و بر داری از خانواده توزیع های نمایی هستند.

قضیه : خانواده توزیع های نمایی ، کامل هستند (اثبات به عنوان تمرین یا مطالعه)

اگر $c(\cdot)$ و $h(\cdot)$ و $T_1(\cdot), \dots, T_m(\cdot)$ توابع حقیقی باشند ،

$$\underline{\theta} \in \Lambda, \quad \underline{y} \in \Gamma, \quad \Lambda \subseteq \mathbb{R}^m, \quad \Gamma \subseteq \mathbb{R}^n$$

$$\underline{\theta} = [\theta_1, \dots, \theta_m]^T, \quad \underline{y} = (y_1, \dots, y_n]^T$$

و داشتیم

$$P_{\underline{\theta}}(\underline{y}) = c(\underline{\theta}) \exp\left(\sum_{i=1}^{m} \theta_i T_i(\underline{y})\right) h(\underline{y})$$

آنگاه که بردار $T(\underline{y}) = [T_1(\underline{y}), T_2(\underline{y}), \dots, T_m(\underline{y})]$ یک آمارگان کافی کامل است. (یعنی $\{P_{\underline{\theta}}(\underline{y}), \underline{\theta} \in \Lambda\}$ یک خانواده توزیع کامل است و $T(\underline{y})$ یک آمارگان کافی روی این خانواده توزیع است)

* نشان داده می‌شود که آماره‌های کامل ، مینیمال نیز هست.

* اگر چند آماره‌ی کامل پیدا کنیم حتماً یک رابطه‌ی یک به یک بین آنها وجود دارد.

در ادامه سعی خواهیم با یک مثال ، روند به دست آوردن تخمین MVUE را ببینیم.

مثال: تخمین دامنه‌ی یک سیگنال از روی مشاهدات نویزی آن

$$Y_k = \mu S_k + N_k \qquad k = 1, 2, \dots, n$$

دامنه سیگنال که می‌خواهیم تخمین بزنیم

سیگنال ارسالی معلوم

نویز

نمونه‌های نویز $N_1, \dots, N_n$ نمونه‌های نویز iid گوسی با توزیع $N(0, \sigma^2)$ و سیگنال $\underline{S} = [S_1, S_2, \dots, S_n]^T$ معلوم و μ پارامتر نامعلوم است که می‌خواهیم آن را تخمین بزنیم، همان دامنه سیگنال

ارسالی است . هدف تخمین MVUE است. $\mu \in \mathbb{R}$

۱- به دست آوردن یک آماره کافی کامل برروی خانواده توزیع‌های $\{P_\theta , \theta \in \Lambda\}$

$N_i \sim N(0, \sigma^2)$

$N_1 \perp\!\!\!\perp N_2 \ldots \perp\!\!\!\perp N_n$

$Y_k = \mu S_k + N_k \quad , \quad k = 1, \ldots, n$

$P_\theta(\underline{y}) = ?$

$P(\underline{y} | \mu) = ?$

به شرط معلوم بودن μ

$$\Longrightarrow \quad Y_k \sim N(\mu S_k, \sigma^2)$$

$$, \quad Y_1 \perp\!\!\!\perp Y_2 \perp\!\!\!\perp \dots \perp\!\!\!\perp Y_n \quad , P(y_k|\mu) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(y_k - \mu S_k)^2}{2\sigma^2}\right)$$

$$\Rightarrow P(\underline{y}|\mu) = \prod_{i=1}^{n} P(y_i|\mu)$$

$$= \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(y_i - \mu S_i)^2}{2\sigma^2}\right)$$

$$\Longrightarrow P(\underline{y} \mid \mu) = \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i - \mu s_i)^2\right)$$

برای به دست آوردن $T(\underline{y})$ از قضیه فاکتورسازی و کامل بودن خانواده توزیع‌های نمایی استفاده می‌کنیم.

$$P(\underline{y} \mid \mu) = \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}\left(y_i^2 + \mu^2 s_i^2 - 2\mu s_i y_i\right)\right)$$

$$\Rightarrow P(\underline{y} \mid \mu) = \underbrace{\left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n \exp\left(-\frac{\sum_{i=1}^{n} y_i^2}{2\sigma^2}\right)}_{h(y)} \underbrace{\exp\left(-\frac{\mu^2}{2\sigma^2}\sum_{i=1}^{n} s_i^2\right)}_{c(\theta)}$$

$$\exp\left(\underbrace{\frac{\mu}{\sigma^2}}_{\theta_1} \underbrace{\sum_{i=1}^{n} s_i y_i}_{T_1(\underline{y})}\right)$$

$$\theta_1 = \frac{\mu}{\sigma^2} \equiv \theta \quad , \quad T(\underline{y}) = T_1(\underline{y}) = \sum_{i=1}^{n} s_i y_i$$

۲- به دست آوردن یک تخمین بدون بایاس از $g(\theta)$ به نام $\hat{g}(\underline{y})$

$$\theta = \frac{\mu}{\sigma^2} \Longrightarrow \mu \equiv g(\theta) = \theta \sigma^2$$

با نگاهی به مشاهدات $y_k = \mu s_k + N_k$ متوجه می‌شویم که امید ریاضی تمام مشاهدات به شرط معلوم بودن μ برابر است با

$$E_\mu\{y_k\} = \mu s_k$$

بنابراین $\hat{g}(\underline{y})$ می‌تواند هر تابع خطی از مشاهدات باشد که با یک ضریب ثابت مقیاس شده باشد. بنابراین ساده‌ترین حالت را برای $\hat{g}(\underline{y})$ به صورت زیر در نظر می‌گیریم.

$$\hat{g}(\underline{y}) = \frac{y_1}{s_1}$$

$$E_\mu\{\hat{g}(\underline{y})\} = \frac{\mu s_1}{s_1} = \mu = g(\theta)$$

۳- به دست آوردن تخمین MVUE با استفاده از رابطه‌ی زیر

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

$$\hat{g}(T(\underline{y})) = E_{\mu}\left\{ \overbrace{\frac{Y_1}{S_1}}^{Z_1} \mid \overbrace{T(\underline{Y})}^{Z_2} = \overbrace{T(\underline{y})}^{z_2} \right\}$$

$$T(\underline{y}) = \sum_{i=1}^{n} S_i y_i \quad , \quad T(\underline{Y}) = \sum_{i=1}^{n} S_i Y_i$$

$$\hat{g}(\underline{y}) = \frac{y_1}{S_1} \quad , \quad \hat{g}(\underline{Y}) = \frac{Y_1}{S_1}$$

Z_1 و Z_2 دو متغیر تصادفی توأماً گوسی هستند، زیرا هر ترکیب خطی دلخواه از Z_1 و Z_2، در واقع یک ترکیب خطی از $y_1, \dots, y_n$ است و می‌دانیم $\{y_i\}_{i=1}^{n}$ متغیرهای تصادفی گوسی مستقل از هم هستند، بنابراین ترکیب یک متغیر تصادفی گوسی می‌کند. (به شرط معلوم بودن μ)

$$g^*(T(\underline{y})) = E_\mu\{Z_1 | Z_2\} = m_{Z_1} + C_{Z_1 Z_2} \, Var^{-1}_{Z_2} (Z_2 - m_{Z_2})$$

$$Z_1 = \frac{y_1}{s_1} \quad , \quad Z_2 = T(\underline{y}) = \sum_{i=1}^{n} s_i y_i$$

$$m_{z_1} = E_{\mu}\left\{ \frac{y_1}{s_1} \right\} = \mu$$

$$m_{z_2} = E_{\mu}\left\{ \sum_{i=1}^{n} s_i \cdot y_i \right\} = \sum_{i=1}^{n} s_i \, \overbrace{E_{\mu}\{y_i\}}^{\mu s_i} = \mu \sum_{i=1}^{n} s_i^2$$

$$m_{z_2} = \mu n \overline{s^2}$$

$$\overline{s^2} = \frac{\sum_{i=1}^{n} s_i^2}{n} \equiv$$ توان سیگنال ارسال

$$C_{z_1 z_2} = E_{\mu}\left\{ (z_1 - m_{z_1})(z_2 - m_{z_2}) \right\} = E_{\mu}\left\{ \left(\frac{y_1}{s_1} - \mu\right)\left(\sum_{i=1}^{n} s_i y_i - \mu n \overline{s^2} \right) \right\}$$

$$C_{z_1 z_2} = E_\mu \left\{ \frac{y_1}{s_1} \sum_{i=1}^{n} s_i y_i - \dots \right\} = \sigma^2$$

$$E_\mu \{ y_i \cdot y_j \} = \begin{cases} E y_i^2 = \mu^2 s_i^2 + \sigma^2 & i = j \\ E y_i E y_j = \mu^2 s_i s_j & i \neq j \end{cases}$$

$$y_i = \mu s_i + n_i \quad , \quad y_j = \mu s_j + n_j$$

$$Var_{z_2} = Var\left(\sum_{i=1}^{n} s_i y_i \right) = E\left(z_2 - m_{z_2} \right)^2$$

$$Var_{z_2} = E\left(\sum_{i=1}^{n} s_i y_i - \mu n \overline{s^2}\right)^2$$

$$= E\left(\left(\sum_{i=1}^{n} s_i y_i\right)\left(\sum_{j=1}^{n} s_j y_j\right) + \mu^2 n^2 (\overline{s^2})^2 - 2\mu n \overline{s^2} \sum_{i=1}^{n} s_i y_i\right)$$

$$\Rightarrow Var(z_2) = n\sigma^2 \overline{s^2}$$

$$\Rightarrow g^*(T(\underline{y})) = \frac{1}{n\overline{s^2}} \sum_{i=1}^{n} s_i y_i = \frac{T(\underline{y})}{n\overline{s^2}}$$

(تابعی از $T(\underline{y})$ است)

$$Var(g^*(T(\underline{y}))) = \frac{1}{(n\overline{s^2})^2} \underbrace{Var(T(\underline{y}))}_{Var_{Z_2}} = \frac{\sigma^2}{n\overline{s^2}}$$

$$= \text{Min Var}$$

همان طور که دیدیم در روش MVUE بر روی مجموعه توابع احتمال کامل $p_\theta(y)$ کار می‌کنیم. علاقمند هستیم که روشهای تخمین دیگری نیز معرفی کنیم که برای تمام خانواده‌ی توزیع‌ها، قابل استفاده باشد.

بنابراین در ادامه روش تخمین ML را برای تخمین پارامترهای نامعلوم، معرفی خواهیم کرد.