

فرایندهای آماری: جلسه 6 24 8 خرداد 1400

* مطالب جلسه:

• یک مثال از نحوه یافتن آمار کافی کین

• اصل likelihood

• شروع بحث تخمین نقطه‌ای (Point estimators)

• مقدمه

• روش‌های یافتن تخمین

• روش گشتاورها

• تخمین بیشینه likelihood (Maximum Likelihood Estimator)

• تخمین بیز (Bayes Estimators)

* در هر: آمار کافی کین $T(X)$ یک آمار کافی که تابعی از هر آمار کافی دلخواه $T'(X)$ به‌ای پارامتر θ باشد.

معنی: اگر $f(x|\theta)$ تابع چگالی یا پری احتمال مشاهدات x باشد و اگر تابع $T(x)$ وجود داشته باشد که برای هر x, y داشته باشیم

$$T(x) = T(y) \iff \frac{f(x|\theta)}{f(y|\theta)}$$

آنگاه $T(x)$ یک آمار کافی است.

مثال: فرض کنید $x_1, \dots, x_n$ n مشاهده iid از توزیع یکدانت در بازه $(\theta, \theta+1)$ باشد $(-\infty < \theta < +\infty)$

برای تابع چگالی احتمال مشترک مشاهدات داریم:

$$f(x|\theta) = \begin{cases} 1 & : \theta < x_i < \theta+1, i \in [1:n] \\ 0 & : \text{o.w.} \end{cases}$$

$$= \begin{cases} 1 & : \max_i x_i - 1 < \Theta < \min_i x_i \\ 0 & : \text{o.w.} \end{cases}$$

$T(X) = (\min_i X_i, \max_i X_i)$: با توجه به شکل مرتب آله انتخاب کنیم
 $= (X_{(1)}, X_{(n)})$

اگر x و y مستقل باشند $\Rightarrow T(x) = T(y) \sqrt{y}$ $\Leftrightarrow \frac{f(x|\theta)}{f(y|\theta)} = 1$

$$\begin{cases} \min_i x_i = \min_i y_i \\ \max_i x_i = \max_i y_i \end{cases}$$

$\Leftrightarrow$ طبق قضیه بالا $T(X) = (X_{(1)}, X_{(n)})$ یک آله کافی است.

تابعی کنیم $f(x|y)$ تابع چگالی یا همی احتمال مشترک مشاهدات X باشد.
یا تابعی از x مشاهد شده باشد، $X=y$ مشاهد شده، تعریف کنیم؛

محقق: X یک مدار گسسته باشد. اگر برای دو نقطه θ_1 و θ_2 داشته باشیم

امثل likelihood: α و β در محاسبه بازنه و دالته بازنه:

$$L(\theta(x)) = \underbrace{C(x, y)}_{\text{یک وجه مستقل}} \cdot L(\theta(y)) \quad \forall \theta$$

آنگاه نتیجه گیری ای که از این دو مشاهده انجام می دهیم باید یکسان باشد (در مورد θ).

توضیح: فرض کنیم داریم: $L(\theta_2|x) = 2L(\theta_1|x)$

حال فرض کنیم برای مشاهده دیگر y داشته باشیم: $L(\theta_1|y) = C(x,y)L(\theta_2|x)$

آنگاه:
$$\frac{L(\theta_2|y)}{L(\theta_1|y)} = \frac{C(x,y)L(\theta_2|x)}{C(x,y)L(\theta_1|x)} = 2$$

* داخل پرانتز: ابره یا لا ما در مورد تابع likelihood در آزمایشی مختلف می شود بیان کرد.

مثال: فرض کنیم X مقدار شیر آمدن صادر ۱۶ هر تاب یک که با احتمال θ آمدن θ باشد.

در آزمایشی دیگر Y مقدار هر تاب های لازم برای سه بار شیر آمدن باشد.

اگر مشاهده $X=3$ اتفاق افتاده باشد برای L_1 داریم:

$$L_1(\theta | X=3) = \binom{12}{3} \theta^3 (1-\theta)^9 = 220 \theta^3 (1-\theta)^9$$

و برای مشاهده $y=12$ تا ج L_2 به صورت زیر است:

$$L_2(\theta | Y=12) = \binom{11}{2} \theta^3 (1-\theta)^9 = 55 \theta^3 (1-\theta)^9$$

که برای این دو مشاهده یکسان طبق (مل likelihood) یا به نتیجه گیری یکسانی را کنه بکنیم.

* تغییراتی نقطه ای (Point Estimators) :

مدل : $f(\underline{x} | \underline{\theta})$

- اگر $\underline{\theta}$ را بدانیم $\rightarrow$ مدل مان را به صورت کاملی می کنیم.
- در بسیاری از موارد $\underline{\theta}$ معنی می دهد یا عملیاتی دارد.

• تعریف : یک تغییراتی نقطه ای هر تابع $w(x_1, \dots, x_n)$ از نتایج است. به بیان دیگر هر آمار یک تغییراتی نقطه ای است.

* روشی برای یافتن تخمینه ها :
روش گشتاورها (Method of Moments) :

فرض کنیم $X_1, \dots, X_n$ مشاهدات مستقل و یکسان که به صورت iid از توزیع

$$f(x | \theta_1, \dots, \theta_k)$$

متوینه دارای شده اند. روش گشتاورها به صورت زیر عمل می کند.

بر اساس مشاهدات و توانیم گشتاورهای نمونه (Sample moments) را بنویسیم:

$$m_1 = \frac{1}{n} \sum_{i=1}^n X_i$$

$$m_2 = \frac{1}{n} \sum_{i=1}^n X_i^2$$

$\vdots$

$$m_k = \frac{1}{n} \sum_{i=1}^n X_i^k$$

از طرقت دیگر با داده‌های نمونه $\rightarrow$ گسترده مرتبه اول تا مرتبه k ام را به دست

بیاوریم $\theta_1, \dots, \theta_k$ می‌نویسیم:

$$\mu_1 = \mathbb{E}[X], \mu_2 = \mathbb{E}[X^2], \dots, \mu_k = \mathbb{E}[X^k]$$

$\swarrow$ $\searrow$ $\swarrow$
تابعی از $\theta_1, \dots, \theta_k$

باید به دست آورد تعیین $\theta_1, \dots, \theta_k$ قمار و رهیم:

$$\begin{cases} \mu_1(\theta_1, \dots, \theta_k) = m_1 \\ \vdots \\ \mu_k(\theta_1, \dots, \theta_k) = m_k \end{cases} \xrightarrow[k \text{ معادله}]{\text{حل دستگاه}} (\hat{\theta}_1, \dots, \hat{\theta}_k)$$