STAT 5430

Lec 28, F, Apr 4

- Homework 7 posted, due, M, Apr 7
  (↑ testing)
- Exam 2 is coming up (2 weeks away)
  on W, April 16, 6:15-8:15 PM, 3rd floor seminar room
- No class on that W.
- I'll post: study guide (sufficiency/completeness/tests)
  - practice exams
  - bring new 1 page (front/back) formula sheet on exam 2 material
    (I'll post one to use if you'd like)
  - can bring calculator & previous formula sheet for exam 1
  - I'll provide table of distributions / STAT 542 facts on test as before

Hypothesis Testing II

Likelihood Ratio Tests: Large Sample Calibrations

$\leftarrow \lambda(\underset{\sim}{x})$

The following result describes the asymptotic distribution of the likelihood ratio statistic (under appropriate regularity conditions) & may be used to calibrate a LRT in a simple fashion when the sample size n is "sufficiently large."

Theorem: Let $\underset{\sim}{X}_1, \underset{\sim}{X}_2, \ldots$ be iid random vectors with common pdf/pmf $f(\underset{\sim}{x}|\theta)$, $\underset{\sim}{\theta} \in \Theta \subset \mathbb{R}^p$ (the parameter $\underset{\sim}{\theta}$ can be vector-valued). Let $\lambda_n(\underset{\sim}{X}_1, \underset{\sim}{X}_2, \ldots, \underset{\sim}{X}_n)$ denote the likelihood ratio statistic based on $\underset{\sim}{X}_1, \underset{\sim}{X}_2, \ldots, \underset{\sim}{X}_n$ for testing $H_0 : \underset{\sim}{\theta} \in \Theta_0 \subset \mathbb{R}^p$ vs $H_1 : \underset{\sim}{\theta} \notin \Theta_0$, where Θ_0 has the form

$r =$ # of parameters to be tested $(1 \le r \le p)$

$$\Theta_0 = \Big\{ \underset{\sim}{\theta} \equiv (\theta_1, \ldots, \theta_p) \in \Theta : \underbrace{\theta_1 = \theta_1^0, \ldots, \theta_r = \theta_r^0}_{\substack{\text{hypothesized values} \\ \text{for first } r \le p \text{ parameters}}} \Big\}$$

claimed values under H_0

for some $\theta_1^0, \ldots, \theta_r^0$, $r \le p$. That is, from the p parameters, we make a claim about exactly r of these parameters and the hypotheses are

$\leftarrow$ "not H_0"

$$\text{“}H_0 : \theta_1 = \theta_1^0, \ldots, \theta_r = \theta_r^0\text{” vs “}H_1 : \theta_i \neq \theta_i^0 \text{ for some } 1 \le i \le r\text{”}$$

$\leftarrow$ e.g. hold for exponential families

Then, under the Cramér-Rao type regularity conditions, it holds that:

$$\textit{if } H_0 \textit{ is true}, \quad -2\log\lambda_n(\underset{\sim}{X}_1, \underset{\sim}{X}_2, \ldots, \underset{\sim}{X}_n) \xrightarrow{d} \chi_r^2 \text{ as } n \to \infty.$$

$\leftarrow$ # of parameters under H_0

$\uparrow$ converges in distribution to χ_r^2

Remark: The above limiting distribution suggests the following testing procedure based on the $(1-\alpha)$-quantile of a χ_r^2 distribution, denoted as $\chi_{1-\alpha}^2(r)$ for which $P\big(\chi_r^2 \le \chi_{1-\alpha}^2(r)\big) = 1-\alpha$ and $P\big(\chi_r^2 > \chi_{1-\alpha}^2(r)\big) = \alpha$.

recall: We reject H_0 if $\lambda(\underset{\sim}{x}) \in [0,1]$ is too small $\Rightarrow$ reject H_0 if $-2\log\lambda(\underset{\sim}{x})$ is too big

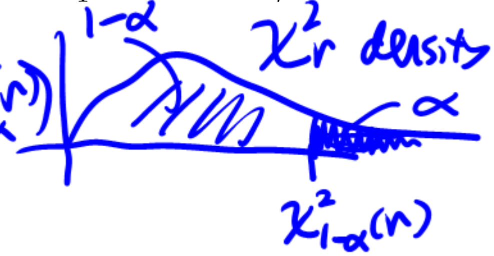
Namely, for n large (e.g., say $n \ge 30$ observations),

$$\varphi(\underset{\sim}{X}_1, \underset{\sim}{X}_2, \ldots, \underset{\sim}{X}_n) = \begin{cases} 1 & \text{if } \; -2\log\lambda_n(\underset{\sim}{X}_1, \underset{\sim}{X}_2, \ldots, \underset{\sim}{X}_n) > \chi_{1-\alpha}^2(r) \\ 0 & \text{otherwise} \end{cases}$$

or $\lambda_n(X_1, \ldots, X_n) < e^{-\chi_{1-\alpha}^2(r)/2}$

is an approximate size α LRT for testing "$H_0 : \theta_1 = \theta_1^0, \ldots, \theta_r = \theta_r^0$" vs "$H_1 : \theta_i \neq \theta_i^0$ for some $1 \le i \le r$."

i.e. $\max_{\underset{\sim}{\theta} \in \Theta_0} E\phi(\underset{\sim}{X}) = \max_{\underset{\sim}{\theta} \in \Theta_0} P_{\underset{\sim}{\theta}}\left(-2\log\lambda(\underset{\sim}{X}) > \chi_{1-\alpha}^2(r)\right) \approx \alpha$



Hypothesis Testing II

Likelihood Ratio Tests + Large Sample Calibration: Illustration

← ← data points as vectors in $\mathbb{R}^2$

Example: Let $\underset{\sim}{X}_1, \underset{\sim}{X}_2, \ldots$ be iid $N_2(\mu, A)$ random vectors, where $\mu = (\mu_1, \mu_2) \in \mathbb{R}^2$

unknown → ↖ known

and A is a known 2×2 positive definite matrix. Find a size α LRT for testing $H_0 : 2\mu_1 + 3\mu_2 = 0$ vs $H_1 : 2\mu_1 + 3\mu_2 \neq 0$, using χ^2-calibration.

<u>Solution</u>: Let $\underset{\sim}{\theta} = (\theta_1, \theta_2)$ where $\theta_1 = 2\mu_1 + 3\mu_2$
$\theta_2 = \mu_2$

Note: $$\underset{\sim}{\theta} = \underbrace{\begin{bmatrix} 2 & 3 \\ 0 & 1 \end{bmatrix}}_{B^{-1}} \underbrace{\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}}_{\underset{\sim}{\mu}}$$

$$\Rightarrow \underset{\sim}{\mu} = \underbrace{\begin{bmatrix} 2 & 3 \\ 0 & 1 \end{bmatrix}^{-1}}_{B} \underset{\sim}{\theta} \equiv B\underset{\sim}{\theta}$$

Hence, $\underset{\sim}{X}_1, \ldots \underset{\sim}{X}_n$ iid $N_2(B\underset{\sim}{\theta}, A)$ and, in terms of $\underset{\sim}{\theta}$, the testing problem is $\underbrace{H_0 : \theta_1 = 0 \text{ vs } H_1 : \theta_1 \neq 0}_{\text{this } H_0 \text{ form needed for } \chi^2\text{-calibration for } \lambda(\underset{\sim}{x})}$

So, we reject H_0 if $-2\log\lambda(\underset{\sim}{x}) > \chi^2_{1-\alpha}(1)$

or $\lambda(\underset{\sim}{x}) < e^{-\chi^2_{1-\alpha}(1)/2}$

Need to find LRS $\lambda(\underset{\sim}{x})$ as follows:

$$L(\underset{\sim}{\theta}) \equiv \text{joint pdf of } \underset{\sim}{X}_1, \ldots, \underset{\sim}{X}_n = f(\underset{\sim}{x}_1, \ldots, \underset{\sim}{x}_n \mid \underset{\sim}{\theta})$$
$$= \prod_{i=1}^{n} \left(\left(\frac{1}{2\pi\sqrt{|A|}} \right) e^{-\frac{1}{2}(\underset{\sim}{x}_i - B\underset{\sim}{\theta})^T A^{-1} (\underset{\sim}{x}_i - B\underset{\sim}{\theta})} \right)$$

↖ $|A| \equiv \det(A)$

$$= \left(\frac{1}{2\pi}\frac{1}{\sqrt{|A|}}\right)^n e^{-\frac{1}{2}\sum_{i=1}^{n}(\underset{\sim}{X}_i - B\underset{\sim}{\varphi})^T A^{-1}(\underset{\sim}{X}_i - B\underset{\sim}{\varphi})}$$

$$= \left(\frac{1}{2\pi}\frac{1}{\sqrt{|A|}}\right)^n e^{-\frac{1}{2}\sum_{i=1}^{n}(\underset{\sim}{Y}_i - \underset{\sim}{\varphi})^T B^T A^{-1} B(\underset{\sim}{Y}_i - \underset{\sim}{\varphi})}$$

where $\underset{\sim}{Y}_i = B^{-1}\underset{\sim}{X}_i = \begin{bmatrix} 2 & 3 \\ 0 & 1 \end{bmatrix}\begin{pmatrix} X_{i1} \\ X_{i2} \end{pmatrix} = \begin{pmatrix} Y_{i1} \\ Y_{i2} \end{pmatrix}$

So, the MLE over Θ is $\underset{\sim}{\hat{\varphi}} = \underset{\sim}{\bar{Y}}_n = \frac{1}{n}\sum_{i=1}^{n} \underset{\sim}{Y}_i$

(just sample means)

Under $H_0: \varphi_1 = 0$,

$f(\underset{\sim}{X}_1, \ldots, \underset{\sim}{X}_n \mid \underset{\sim}{\varphi} = (\varphi_1, \varphi_2), \varphi_1 = 0)$

$$= \left(\frac{1}{2\pi\sqrt{|A|}}\right)^n e^{-\frac{1}{2}\sum_{i=1}^{n}\{\sigma_{22}(Y_{i2}-\varphi_2)^2 + 2\sigma_{12}Y_{i1}(Y_{i2}-\varphi_2) + \sigma_{11}Y_{i1}^2\}}$$

where $B^T A^{-1} B = \begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{12} & \sigma_{22} \end{bmatrix}$

& Maximizer for $\underset{\sim}{\varphi}$ over Θ_0 is

$\underset{\sim}{\tilde{\Phi}} = (0, \tilde{\Phi}_2)$, $\tilde{\Phi}_2 = \bar{Y}_{2n} + \frac{\sigma_{12}}{\sigma_{22}}\bar{Y}_{1n}$ where $\bar{Y}_{jn} = \frac{1}{n}\sum_{i=1}^{n} Y_{ij}$, $j = 1, 2$

$$-2\log\lambda(\underset{\sim}{X}) = -2\log\frac{L(\tilde{\Phi})}{L(\hat{\Phi})} = \sum_{i=1}^{n}(\underset{\sim}{Y}_i - \underset{\sim}{\hat{\Phi}})^T B^T A^{-1} B(\underset{\sim}{Y}_i - \underset{\sim}{\hat{\Phi}}) - \sum_{i=1}^{n}(\underset{\sim}{Y}_i - \underset{\sim}{\tilde{\varphi}}) B^T A^{-1} B(\underset{\sim}{Y}_i - \underset{\sim}{\tilde{\varphi}})$$