~~3/30/2020~~

Note 1.

Page 1 STA 200C: Hypothesis testing

Purpose and outline:

① Testing methods:

Likelihood ratio tests

Union-interaction / Interaction-union tests

Bayesian tests

② Basic concepts:

Power function

Size and level

P-values

Confidence intervals

③ Large-sample theory

④ Other topics:

Robustness

Nonparametric methods

Mutiple testing and FDR.

Statistical model

$X \sim f(x|\theta)$ $\qquad\qquad \theta \in \Theta$

$\qquad$ pdf / pmf $\qquad\qquad\qquad$ parametric space

$\underset{\sim}{X} = (X_1, \cdots, X_n)$ is called a random sample from $f(x|\theta)$. If all X_i have distribution given by $f(x|\theta)$, and the X_i are independent.

Statistical procedures aim to make inference about θ based on $\underset{\sim}{X}$.

1) Point estimation: find $\hat{\theta}$ based on $\underset{\sim}{X}$;
2) Confidence interval: find $[\hat{\theta}_L, \hat{\theta}_U]$ based on $\underset{\sim}{X}$:
3) Hypothesis testing: Make decision on $\theta \in \Theta_0$ or $\theta \in \Theta_1$,

where $\left| \begin{array}{l} \Theta_0 \cup \Theta_1 = \Theta \\ \Theta_0 \cap \Theta_1 = \phi. \end{array} \right.$

In this course, we aim to establish generic methods, concepts and theory for hypothesis testing.

The following three examples will be used repeatedly throughout the course:

Example 1. $X_1, \cdots, X_n \overset{iid}{\sim} N(\theta, \sigma^2)$, σ^2 is known

$$\text{pdf}: f(x|\theta) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\theta}{\sigma}\right)^2}$$

Shift family of normal distribution

$\theta \in \Theta = (-\infty, \infty)$

Then

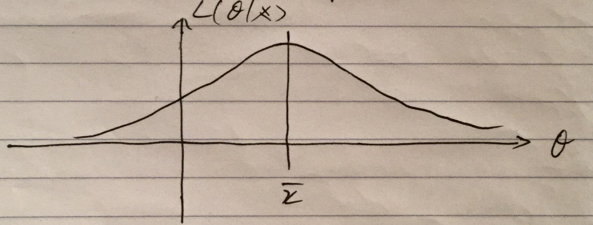
$$f(\mathbf{x}\mid\theta) = \prod_{i=1}^{n} f(x_i\mid\theta)$$

$$= \left(\frac{1}{\sigma\sqrt{2\pi}}\right)^n \exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\theta)^2\right\}$$

$$= \left(\frac{1}{\sigma\sqrt{2\pi}}\right)^n \exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}x_i^2\right\}\exp\left\{\frac{1}{2\sigma^2}\left(2n\theta\bar{x} - n\theta^2\right)\right\}$$

The likelihood function $L(\theta\mid\mathbf{x}) = f(\mathbf{x}\mid\theta)$

Sufficient statistic: $T(\mathbf{x}) = \bar{x}$



MLE: $\hat{\theta} = \bar{x}$.

Example 2: Bernoulli family

$$X_1, \dots, X_n \overset{iid}{\sim} Ber(\theta),$$

$$\theta \in \Theta = (0, 1)$$

$$\text{pmf: } f(x\mid\theta) = \begin{cases} \theta & \text{if } x=1 \\ 1-\theta & \text{if } x=0 \end{cases}$$

$$= \theta^x(1-\theta)^{1-x}$$

Then $f(\mathbf{x}|\theta) = \prod_{i=1}^{n} f(x_i|\theta)$

$$= \prod_{i=1}^{n} \theta^{x_i} (1-\theta)^{1-x_i}$$

$$= \theta^{\sum_{i=1}^{n} x_i} (1-\theta)^{n-\sum_{i=1}^{n} x_i}$$

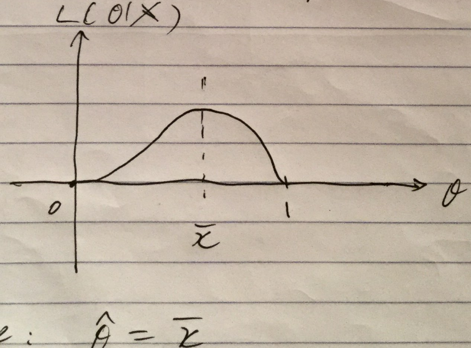
$$= \theta^{n\bar{x}} (1-\theta)^{n(1-\bar{x})}$$

$$= \exp\{n\bar{x}\log\theta + n(1-\bar{x})\log(1-\theta)\}$$

$$= \exp\{n\bar{x}(\log\theta - \log(1-\theta)) + n\log(1-\theta)\}$$

The likelihood function: $L(\theta|\mathbf{x}) = f(\mathbf{x}|\theta)$.

Sufficient statistic: $T(\mathbf{x}) = \bar{x}$.



mle: $\hat{\theta} = \bar{x}$

Example 3 : Shift family of exponential distribution

$X_1, \dots, X_n$ iid exponential with pdf

$$f(x \mid \theta) = \begin{cases} e^{-(x-\theta)} & x \geq \theta \\ 0 & x < \theta \end{cases}$$

$$\theta \in \Theta = (-\infty, \infty)$$

Then

$$f(\underset{\sim}{x} \mid \theta) = \prod_{i=1}^{n} f(x_i \mid \theta)$$

$$= \begin{cases} e^{-\sum_{i=1}^{n}(x_i - \theta)} & x_{(1)} \geq \theta \\ 0 & x_{(1)} < \theta . \end{cases}$$

where $x_{(1)} = \min_i x_i$.

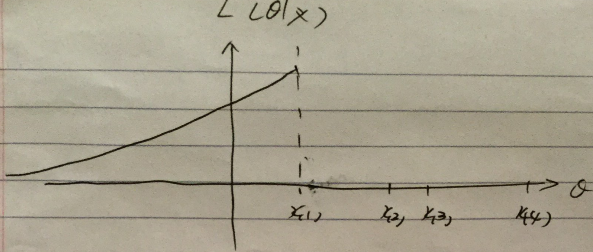
Then $L(\theta \mid \underset{\sim}{x}) = f(\underset{\sim}{x} \mid \theta)$

$$= \begin{cases} e^{-\sum_{i=1}^{n}(x_i - \theta)} & \theta \leq x_{(1)} \\ 0 & \theta > x_{(1)} \end{cases}$$

$$= e^{-n\bar{x}} e^{n\theta} \mathbb{1}\{\theta \leq x_{(1)}\}$$

The sufficient statistic is $T(\underset{\sim}{x}) = x_{(1)}$

Page 6



mle : $\hat{\theta} = x_{(1)}$