# STAT 517: Statistical Inference

## Lecture 8: Maximum Likelihood Tests

Prof. Michael Levine

March 5, 2015

# Basic Setup

- The model: observations $X_1, \ldots, X_n \sim f(x, \theta)$ with $\theta \in \Omega$
- Hypotheses: $H_0 : \theta = \theta_0$ vs. $H_a : \theta \neq \theta_0$
- The likelihood function is $L(\theta) = \prod_{i=1}^{n} f(X_i; \theta)$ and the log-likelihood is $l(\theta) = \sum_{i=1}^{n} \log f(X_i; \theta)$
- Keep in mind that asymptotically $L(\theta_0)$ is the maximum value of $L(\theta)$

# Likelihood Ratio Test

- Observe that the ratio

$$\Lambda = \frac{L(\theta_0)}{L(\hat{\theta})}$$

- $\Lambda$ is always less than 1 but close to it if $H_0$ is true...
- Decision rule: reject $H_0$ at the level of significance $\alpha$ if $\Lambda \leq c$ where

$$\alpha = P_{\theta_0}[\Lambda \leq c]$$

- The resulting test is called the **likelihood ratio test** or LRT

# Example

- Take $f(x;\theta) = \frac{1}{\theta} e^{-x/\theta}$ - an exponential pdf, $0 < \theta < \infty$
- We know that $\hat{\theta} = \bar{X}$ is the MLE and $L(\theta) = \theta^{-n} e^{-\frac{n}{\theta}\bar{X}}$
- The likelihood ratio has the form (up to $e^n$) of $g(t) = t^n e^{-nt}$ where $t = \frac{\bar{X}}{\theta_0} > 0$
- Easy to verify that $g^{'}(1) = 0$ and $t = 1$ is an actual maximum; thus, $g(t) \le c$ iff $t < c_1$ or $t \ge c_2$
- Under $H_0$, $\frac{2}{\theta_0} \sum_{i=1}^{n} X_i \sim \chi^2_{2n}$ and so the decision rule is to reject $H_0$ if

$$\frac{2}{\theta_0} \sum_{i=1}^{n} X_i \le \chi^2_{2n,1-\alpha/2}$$

or

$$\frac{2}{\theta_0} \sum_{i=1}^{n} X_i \ge \chi^2_{2n,\alpha/2}$$

# Example

- Now, take $f(x;\theta) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-\theta)^2/2\sigma^2}$ with known $\sigma^2 > 0$
- Hypotheses: $H_0 : \theta = \theta_0$ vs. $H_a : \theta \neq \theta_0$
- Observe that the MLE $\hat{\theta} = \bar{X}$ and the likelihood ratio is

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

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

# Example

- The likelihood ratio is

$$\Lambda = e^{-\frac{1}{2\sigma^2} n(\bar{x}-\theta_0)^2}$$

- $\Lambda \leq c$ is equivalent to $-2 \log \Lambda \geq -2 \log c$ where

$$-2 \log \Lambda = \left( \frac{\bar{X} - \theta_0}{\sigma \sqrt{n}} \right)^2 \sim \chi^2_1$$

  under $H_0$
- The decision rule: reject $H_0$ if $-2 \log \Lambda = \left( \frac{\bar{X}-\theta_0}{\sigma\sqrt{n}} \right)^2 \geq \chi^2_{1,\alpha}$

# Asymptotic result

- Under the regularity conditions needed for asymptotic efficiency of MLE, under $H_0 : \theta = \theta_0$ we have

$$-2 \log \Lambda \xrightarrow{D} \chi^2_1$$

- This, of course, suggests the decision rule "Reject $H_0$ if $\chi^2_L \geq \chi^2_{1,\alpha}$
- The resulting test has the asymptotic level of significance $\alpha$

# Wald type test

- A natural test statistic based on the asymptotic distribution of $\hat{\theta}$

$$\chi_W^2 = \left\{ \sqrt{nI(\hat{\theta})}(\hat{\theta} - \theta_0) \right\}^2$$

- $I(\hat{\theta}) \xrightarrow{p} I(\theta_0)$ under $H_0$ because $I(\theta)$ is a continuous function
- Thus, under $H_0$, $\chi_W^2 \sim \chi_1^2$ asymptotically; the decision rule is "reject $H_0$ if $\chi_W^2 \geq \chi_{1,\alpha}^2$
- This test has asymptotic level $\alpha$; moreover, one can show that

$$\chi_W^2 - \chi_L^2 \xrightarrow{p} 0$$

# Scores-type test

- The score vector is

$$S(\theta) = \left\{ \frac{\partial \log f(X_i;\theta)}{\partial \theta} \right\}'$$

- Note that $\frac{1}{\sqrt{n}} l'(\theta_0) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \frac{\partial \log f(X_i;\theta_0)}{\partial \theta}$ and define the statistic

$$\chi_R^2 = \left( \frac{l'(\theta_0)}{\sqrt{nI(\theta_0)}} \right)^2$$

- Can show that

$$\chi_R^2 = \chi_W^2 + R_{0n}$$

where $R_{0n} \xrightarrow{p} 0$

- The decision rule is "reject $H_0$ if $\chi_R^2 \geq \chi_{1,\alpha}^2$"

# Example

- Take a beta $X \sim (\theta, 1)$ with the pdf $f(x) = \theta x^{\theta-1}$, $0 < x < 1$
- Test $H_0 : \theta = 1$ (meaning $X \sim Unif[0,1]$) vs. $H_a : \theta \neq 1$
- Recall that the MLE of $\theta$ is $\hat{\theta} = -\frac{n}{\sum_{i=1}^n \log X_i}$
- Observe that

$$L(\hat{\theta}) = \left(-\sum_{i=1}^n \log X_i\right)^{-n} \exp\left\{-\sum_{i=1}^n \log X_i\right\} \exp\left\{n(\log n - 1)\right\}$$

and $L(1) = 1$

- Therefore, the likelihood ratio is $\Lambda = \frac{1}{L(\hat{\theta})}$ and the test statistic is

$$\begin{aligned}\chi_L^2 &= -2\log\Lambda \\ &= 2\left\{-\sum_{i=1}^n \log X_i - n\log\left(-\sum_{i=1}^n \log X_i\right) - n + n\log n\right\}\end{aligned}$$

# Example

- Recall that $I(\theta) = \theta^{-2}$ that can be estimated consistently by $\hat{\theta}^{-2}$
- The Wald-type test statistic is

$$\chi^2_W = \left( \sqrt{\frac{n}{\hat{\theta}^2}} \left( \hat{\theta} - 1 \right) \right)^2 = n \left\{ 1 - \frac{1}{\hat{\theta}} \right\}^2$$

- To obtain the scores type test note that

$$l^{'}(1) = \sum_{i=1}^{n} \log X_i + n$$

- The test statistic is

$$\chi^2_R = \left\{ \frac{\sum_{i=1}^{n} \log X_i + n}{\sqrt{n}} \right\}^2$$

- Note that in this particular case $\chi^2_L = \chi^2_R$