Statistical Inference

Lecture 08a

ANU - RSFAS

Last Updated: Sun Apr 22 17:08:25 2018

Maximum Likelihood Ratio Tests

Section 4.6: The likelihood ratio test for testing

$$H_0 : \theta \in \omega \text{ versus } H_1 : \theta \in \Omega - \omega$$

$$\lambda(\boldsymbol{x}) = \frac{\max_{\Theta \in \omega} L(\theta; \boldsymbol{x})}{\max_{\Theta \in \Omega} L(\theta; \boldsymbol{x})}$$

if maximize θ
in null space

- Note:
  - $\max_{\Theta \in \omega} L(\theta; \boldsymbol{x})$ is a restricted maximization.
  - $\max_{\Theta \in \Omega} L(\theta; \boldsymbol{x})$ is a unrestricted maximization.

- We construct a test of the form:

$$C = \{\boldsymbol{x} : \lambda(\boldsymbol{x}) \leq k\}$$

- Note: $0 \leq \lambda \leq 1$, and λ will be close to 1 if H_0 is true.
- Where $0 \leq k \leq 1$.

Likelihood Ratio Tests

Example: Let $X_1, \ldots, X_n \overset{\text{iid}}{\sim} \text{normal}(\theta, 1)$.

- Test $H_0 : \theta = \theta_0$ versus $H_1 : \theta \neq \theta_0$.
- θ_0 is a number fixed by the experimenter prior to the experiment.

$$\max_{\Theta \in \omega} L(\theta; \boldsymbol{x}) = L(\theta_0; \boldsymbol{x})$$

$$\max_{\Theta \in \Omega} L(\theta; \boldsymbol{x}) = L(\hat{\theta}; \boldsymbol{x}) \Rightarrow \hat{\theta} = \bar{X}$$

Likelihood Ratio Tests

$$
\begin{aligned}
\lambda(\boldsymbol{x}) &= \frac{(2\pi)^{-n/2} exp[-\sum(x_i-\theta_0)^2/2]}{(2\pi)^{-n/2} exp[-\sum(x_i-\bar{x})^2/2]} \\
&= exp\left[\left(-\sum(x_i-\theta_0)^2+\sum(x_i-\bar{x})^2\right)/2\right] \\
&= exp\left[\left(-\left[\sum(x_i-\bar{x})^2+n(\bar{x}-\theta_0)^2\right]+\sum(x_i-\bar{x})^2\right)/2\right] \\
&= exp\left[-n(\bar{x}-\theta_0)^2/2\right]
\end{aligned}
$$

Likelihood Ratio Tests

$$
\begin{aligned}
C &= \{\lambda(\boldsymbol{x}) \leq k\} \\
&= \left\{exp\left[-n(\bar{x}-\theta_0)^2/2\right] \leq k\right\} \\
&= \{-n(\bar{x}-\theta_0)^2/2 \leq log(k)\} \\
&= \{(\bar{x}-\theta_0)^2 > [-2log(k)]/n\} \\
&\Rightarrow \{|\bar{x}-\theta_0| > \sqrt{[-2log(k)]/n}\} \\
&\Rightarrow \left\{\frac{|\bar{x}-\theta_0|}{1/\sqrt{n}} > \frac{\sqrt{[-2log(k)]/n}}{1/\sqrt{n}}\right\} \\
&= \left\{|Z| > \frac{\sqrt{[-2log(k)]/n}}{1/\sqrt{n}}\right\}
\end{aligned}
$$

Likelihood Ratio Tests

- Now we have:

$$C = \left\{ |Z| > \sqrt{n}\sqrt{[-2log(k)]/n} \right\} = \{|Z| > k^*\}$$

- Under the null hypothesis $\theta = \theta_0$. So $Z \sim \mathrm{normal}(0, 1)$.

$$\begin{aligned} P(|Z| > k^*) &= P(Z > k^*) + P(Z < -k^*) = \alpha \\ &= 2P(Z < -k^*) = \alpha \\ &= P(Z < -k^*) = \alpha/2 \\ &= P(Z < k^{**}) = \alpha/2 \end{aligned}$$

Likelihood Ratio Tests

- Suppose $\alpha = 0.05$, then $k^{**} = -1.96$

```r
qnorm(0.05/2)
```

```
## [1] -1.959964
```

- So we will reject H_0 if:

$$\left\{ \left| \frac{(\bar{x} - \theta_0)}{1/\sqrt{n}} \right| > 1.96 \right\}$$

Likelihood Ratio Tests

Eg. Let $X_1, \ldots, X_n \overset{\text{iid}}{\sim} \mathrm{Poisson}(\theta)$.

- Test $H_0 : \theta = \theta_0$ versus $H_1 : \theta \neq \theta_0$.
- θ_0 is a number fixed by the experimenter prior to the experiment.

$$\max_{\Theta_0} L(\theta; \boldsymbol{x}) = L(\theta_0; \boldsymbol{x})$$

$$\max_{\Theta} L(\theta; \boldsymbol{x}) = L(\hat{\theta}; \boldsymbol{x}) \Rightarrow \hat{\theta} = \bar{X}$$

Likelihood Ratio Tests

$$
\begin{aligned}
\lambda(\boldsymbol{x}) &= \frac{\frac{exp(-n\theta_0)\theta_0^{\sum x_i}}{\prod x_i!}}{\frac{exp(-n\hat{\theta})\hat{\theta}^{\sum x_i}}{\prod x_i!}} \\
&= \frac{exp(-n\theta_0)\theta_0^{\sum x_i}}{exp(-n\hat{\theta})\hat{\theta}^{\sum x_i}} \\
&= exp(-n(\theta_0 - \hat{\theta}))\left(\frac{\theta_0}{\hat{\theta}}\right)^{\sum x_i} \\
&= exp(-n(\theta_0 - \hat{\theta}))\left(\frac{\theta_0}{\hat{\theta}}\right)^{n\bar{x}} \\
&= exp(-n(\theta_0 - \bar{x}))\left(\frac{\theta_0}{\bar{x}}\right)^{n\bar{x}}
\end{aligned}
$$

Likelihood Ratio Tests

- The rejection region is of the form:

$$C = \{\boldsymbol{x} : \lambda(\boldsymbol{x}) \leq k\} \quad = \quad \left\{ exp(n(\bar{x} - \theta_0)) \left(\frac{\theta_0}{\bar{x}}\right)^{n\bar{x}} \leq k \right\}$$

- Notice again that this is based on a sufficient statistic.
- If we could determine the distribution of $\lambda(\boldsymbol{X})$ we could then determine k for a given α!
- Looks a bit tricky here!!

Likelihood Ratio Tests - Asymptotics

Theorem (Section 4.6.1): For testing $H_0 : \theta \in \omega$ versus $H_1 : \theta \in \Omega$,

- suppose $X_1, \ldots, X_n \overset{\text{iid}}{\sim} f(x;\theta)$ and $\hat{\theta}$ is the MLE of θ and $f(x;\theta)$ satisfies the regularity conditions (smoothness).
- Then under H_0, as $n \to \infty$,

$$-2log[\lambda(\boldsymbol{x})] \overset{D}{\to} \chi_1^2$$

Likelihood Ratio Tests - Asymptotics

Proof:

- Do a two-step Taylor series expansion of $\ell(\theta; \boldsymbol{x})$ around $\hat{\theta}$:

$$\ell(\theta; \boldsymbol{x}) = \ell(\hat{\theta}; \boldsymbol{x}) + \ell'(\hat{\theta}; \boldsymbol{x})(\theta - \hat{\theta}) + \ell''(\hat{\theta}; \boldsymbol{x})\frac{(\theta - \hat{\theta})^2}{2} + \cdots$$

- $\ell'(\hat{\theta}; \boldsymbol{x}) = 0$ and dropping $(\cdots)$, we have:

$$\ell(\theta; \boldsymbol{x}) = \ell(\hat{\theta}; \boldsymbol{x}) + \ell''(\hat{\theta}; \boldsymbol{x})\frac{(\theta - \hat{\theta})^2}{2}$$

Likelihood Ratio Tests - Asymptotics

- Now consider:

$\lambda = \dfrac{L(\theta_0)}{L(\hat{\theta})} = -2\log(\lambda)$

$$\underline{-2log(\lambda)} = -2[\ell(\theta_0; \boldsymbol{x}) - \ell(\hat{\theta}; \boldsymbol{x})]$$

- Substitute Taylor's approximation for $\ell(\theta_0; \boldsymbol{x})$:

$$\begin{aligned}
-2log(\lambda) &= -2\ell(\theta_0; \boldsymbol{x}) + 2\ell(\hat{\theta}; \boldsymbol{x}) \\
&= -2\left[\ell(\hat{\theta}; \boldsymbol{x}) + \ell''(\hat{\theta}; \boldsymbol{x})\frac{(\theta - \hat{\theta})^2}{2}\right] + 2\ell(\hat{\theta}; \boldsymbol{x}) \\
&= -\ell''(\hat{\theta}; \boldsymbol{x})(\theta - \hat{\theta})^2
\end{aligned}$$

Likelihood Ratio Tests - Asymptotics

just like.

- Now, $-\frac{1}{n}\ell''(\hat{\theta};\boldsymbol{x}) \overset{LLN}{\rightarrow} i(\theta_0)$.

$\bar{X} \xrightarrow{LLN} \mu$

$\ell''(\hat{\theta}) \rightarrow i(\theta_0)$

$$
\begin{aligned}
-2log(\lambda) &= -\ell''(\hat{\theta};\boldsymbol{x})(\theta-\hat{\theta})^2 \\
&= ni(\theta)(\hat{\theta}-\theta)^2 \\
&= \left[\sqrt{ni(\theta)}(\hat{\theta}-\theta)\right]^2 \\
&= \left[\frac{\sqrt{n}(\hat{\theta}-\theta)}{1/\sqrt{i(\theta)}}\right]^2
\end{aligned}
$$

- We showed:

$$\sqrt{n}(\hat{\theta}-\theta) \overset{D}{\rightarrow} \mathrm{normal}(0, i(\theta)^{-1})$$

Likelihood Ratio Tests - Asymptotics

- So:

$$\frac{\sqrt{n}(\hat{\theta}-\theta)}{1/\sqrt{i(\theta)}} = Z \xrightarrow{D} \mathrm{normal}(0,1)$$

- Thus:

$$-2log(\lambda) = Z^2 \xrightarrow{D} \chi^2_1$$

Likelihood Ratio Tests - Asymptotics

- Back to our Poisson example:

$$C = \{\boldsymbol{x} : \lambda(\boldsymbol{x}) \leq k\} \quad = \quad \left\{ exp(n(\bar{x} - \theta_0)) \left(\frac{\theta_0}{\bar{x}}\right)^{n\bar{x}} \leq k \right\}$$

- Consider the asymptotic distribution:

$$\begin{aligned} -2log(\lambda) &= -2log\left[exp(n(\bar{x} - \theta_0)) \left(\frac{\theta_0}{\bar{x}}\right)^{n\bar{x}} \right] \\ &= 2n\left[(\bar{x} - \theta_0) + \bar{x}\,log\left(\frac{\theta_0}{\bar{x}}\right) \right] \sim \chi^2_1 \end{aligned}$$

Likelihood Ratio Tests - Asymptotics

- If we reject when $\{\lambda \leq k\}$, then we reject when

$$\{-2log(\lambda) > -2log(k)\} = \{-2log(\lambda) > k^*\}$$

- What value of k^* should we pick so that $\alpha = 0.05$?

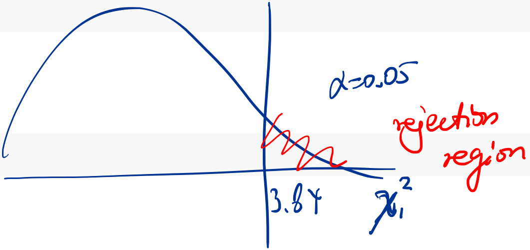
$$P(-2log(\lambda) > k^*) = 0.05$$

```r
qchisq(0.95, 1)
```

```
## [1] 3.841459
```

```r
1-pchisq(3.841,1)
```

```
## [1] 0.05001368
```



$x_1, \dots, x_n \overset{iid}{\sim} N(\mu, \sigma^2)$

$\mu \leq -2$

$H_0: \mu = \mu_0$, $\cancel{\sigma^2 = 1}$ $\quad H_1: \mu \neq \mu_0$, $\cancel{\sigma^2 \neq 1}$ $\Big\}$ ~~2 constraints~~

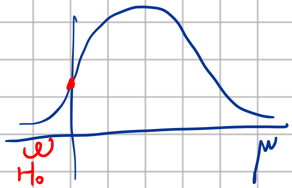
MLRT:

$$\lambda = \frac{\max_{\cancel{\theta \in \omega}} L(\mu_0, \sigma_0^2)}{\max_{\cancel{\theta \in \Omega}} L(\mu, \sigma^2)}$$

$$\lambda = \frac{L(\mu_0, \hat{\sigma}^2)}{L(\hat{\mu}, \hat{\sigma}^2)}$$

- small Sample (Example 4.8)
- $-2\log(\lambda) \sim \chi^2_2$ # of constraints under H_0

χ_1^2



Likelihood Ratio Tests - Asymptotics

Theorem A: This theorem extends the previous one to allow for more parameters. It can be shown:

$$-2log(\lambda) \xrightarrow{D} \chi^2_\nu$$

where $\nu = \#\text{number of constraints set in } H_0$.

- Another way to think about it is: Let p be the number of parameters estimated (are free) under H_1. And let p_0 be the the number of parameters estimated (are free) under H_0.
- Then $\nu = p - p_0$.