Likelihood Ratio Test (LRT)

Suppose $\boldsymbol{X} \sim P_\theta$, $\theta \in \Theta$, with joint pdf (or pmf) $f(\boldsymbol{x} \,|\, \theta)$.

For observed data $\boldsymbol{x}$, the likelihood function is $L(\theta \,|\, \boldsymbol{x}) \equiv f(\boldsymbol{x} \,|\, \theta)$.

The LRT statistic for testing

$$H_0 : \theta \in \Theta_0 \quad \text{versus} \quad H_1 : \theta \in \Theta_0^c$$

is given by

$$\lambda(\boldsymbol{x}) = \frac{\sup_{\theta \in \Theta_0} L(\theta \,|\, \boldsymbol{x})}{\sup_{\theta \in \Theta} L(\theta \,|\, \boldsymbol{x})} = \frac{L(\widehat{\theta}_0 \,|\, \boldsymbol{x})}{L(\widehat{\theta} \,|\, \boldsymbol{x})}$$

where $\widehat{\theta}_0 = \underset{\theta \in \Theta_0}{\operatorname{argmax}} \, L(\theta \,|\, \boldsymbol{x})$ and $\widehat{\theta} = \underset{\theta \in \Theta}{\operatorname{argmax}} \, L(\theta \,|\, \boldsymbol{x})$.

The LRT rejects for small values of $\lambda(\boldsymbol{x})$; the test has rejection region (critical region) given by

$$\mathcal{R} = \{\, \boldsymbol{x} : \lambda(\boldsymbol{x}) \le c \,\}$$

where c is chosen so that

$$\sup_{\theta \in \Theta_0} P_\theta(\lambda(\boldsymbol{X}) \le c) = \alpha \quad \text{(or failing that, } \le \alpha)$$

for some pre-specified value α (say, .05 or .01).

Sometimes the exact distribution of $\lambda(\boldsymbol{X})$ can be obtained and then used to find c giving an exact size α test.

But often this cannot be done, and we have to rely on the following asymptotic approximation.

Asymptotic Distribution of LRT Statistic

Consider a sequence of successively larger data sets

$$\boldsymbol{X_n} = (X_1, X_2, \ldots, X_n)$$

and let $\lambda_n(\boldsymbol{X_n})$ be the LRT statistic based on $\boldsymbol{X_n}$.

Theorem: If $\theta \in \Theta_0$, then (under regularity conditions)

$$-2 \log \lambda_n(\boldsymbol{X_n}) \xrightarrow{d} \chi^2_k \quad \text{as } n \to \infty$$

where $k \equiv (\dim \Theta) - (\dim \Theta_0)$.

Thus, if c^* satisfies $P(\chi^2_k \geq c^*) = \alpha$, then the rejection region $\mathcal{R} = \{\boldsymbol{x} : -2 \log \lambda(\boldsymbol{x}) \geq c^*\}$ gives an approximate size α test for large sample sizes.

Comment: The test statistics $\lambda(\boldsymbol{x})$ and $-2 \log(\boldsymbol{x})$ are equivalent since

$$\lambda(\boldsymbol{x}) \leq c \ \text{ iff } \ -2 \log \lambda(\boldsymbol{x}) \geq c^* \quad \text{where } c^* \equiv -2 \log c.$$

It is often convenient to replace the LRT statistic $\lambda(\boldsymbol{x})$ by an equivalent statistic obtained by applying a strictly monotone transformation.

Comment on the regularity conditions: Conditions are required on both the family of distributions $f(\boldsymbol{x} \,|\, \theta)$ and the set Θ_0. The family $f(\boldsymbol{x} \,|\, \theta)$ must satisfy conditions like those required for the consistency and asymptotic normality of the MLE (and the validity of the Fisher information). The set Θ_0 must be a lower dimensional subspace (or manifold) of Θ.

Lemma: Let $T, n > 0$. Define

$$H(\sigma^2) = (2\pi\sigma^2)^{-n/2} e^{-T/(2\sigma^2)} \quad \text{for } \sigma^2 > 0\,.$$

Then

$$\begin{aligned} \underset{\sigma^2>0}{\operatorname{argmax}}\, H(\sigma^2) &= T/n \equiv \hat{\sigma}^2 \quad \text{and} \\ \sup_{\sigma^2>0} H(\sigma^2) &= H\left(\hat{\sigma}^2\right) = \left(2\pi\hat{\sigma}^2\right)^{-n/2} e^{-n/2}\,. \end{aligned}$$

Example: Observe $X_1, \ldots, X_n$ iid $N(0, \sigma^2)$. Find LRT of

$$H_0 : \sigma^2 = \sigma_0^2 \quad \text{versus} \quad H_1 : \sigma^2 \neq \sigma_0^2$$

Here:

$\Theta = (0, \infty)$ and $\Theta_0 = \{\sigma_0^2\}$.

$L(\sigma^2) = (2\pi\sigma^2)^{-n/2} \exp\left(-(2\sigma^2)^{-1} \sum_i x_i^2\right)$.

$\underset{\Theta}{\operatorname{argmax}}\, L(\sigma^2) = n^{-1} \sum_i x_i^2 \equiv \hat{\sigma}^2 \quad \left(\text{so that} \quad \sum_i x_i^2 = n\hat{\sigma}^2\right)$.

$$\lambda(\boldsymbol{x}) = \frac{L(\sigma_0^2)}{L(\hat{\sigma}^2)} = \frac{(2\pi\sigma_0^2)^{-n/2} \exp\left(-(2\sigma_0^2)^{-1} n\hat{\sigma}^2\right)}{(2\pi\hat{\sigma}^2)^{-n/2} \exp\left(-n/2\right)}.$$

$$= e^{n/2} \left(\frac{\hat{\sigma}^2}{\sigma_0^2}\right)^{n/2} \exp\left[-\frac{n}{2}\left(\frac{\hat{\sigma}^2}{\sigma_0^2}\right)\right]$$

$$= \psi\left(\frac{\hat{\sigma}^2}{\sigma_0^2}\right) \quad \text{where } \psi(u) \equiv e^{n/2} u^{n/2} e^{-(n/2)u}.$$

Example (continued)

With $X_1, X_2, \ldots, X_n$ iid $N(0, \sigma^2)$, the LRT of

$$H_0: \sigma^2 = \sigma_0^2 \quad \text{vs.} \quad H_1: \sigma^2 \neq \sigma_0^2$$

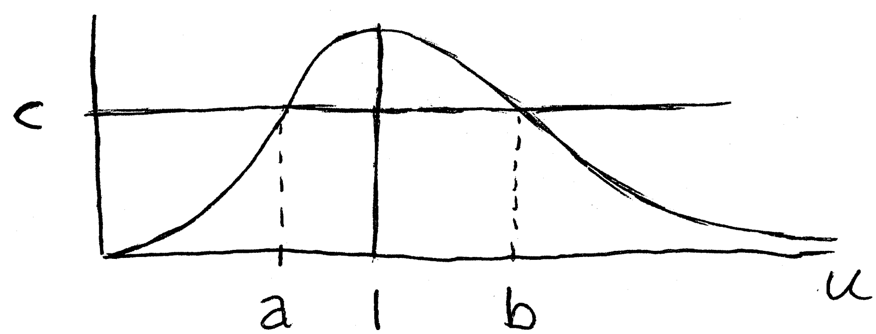
rejects in the region

$$\mathcal{R} = \{x : \lambda(x) \leq c\} \quad \text{where}$$

$$\lambda(x) = \psi\left(\frac{\hat{\sigma}^2}{\sigma_0^2}\right), \quad \hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n} x_i^2, \text{ and}$$

$$\psi(u) = e^{n/2} u^{n/2} e^{-(n/2)u}.$$

The function ψ is maximized at $u = 1$ and looks like this:



Find a, b with $a < b$ and $\psi(a) = \psi(b) = c$.

Thus $\mathcal{R} = \left\{x : \frac{\hat{\sigma}^2}{\sigma_0^2} \leq a(c) \text{ or } \frac{\hat{\sigma}^2}{\sigma_0^2} \geq b(c)\right\}$.

We reject when $\hat{\sigma}^2/\sigma_0^2$ departs far enough from 1.

Obtaining an exact level α test

$$P_{\sigma_0^2}(X \in R) = 1 - P_{\sigma_0^2}\left(a(c) < \frac{\hat{\sigma}^2}{\sigma_0^2} < b(c)\right)$$

$$= 1 - P_{\sigma_0^2}\left(n\,a(c) < \frac{n\hat{\sigma}^2}{\sigma_0^2} < n\,b(c)\right).$$

$$\frac{n\hat{\sigma}^2}{\sigma_0^2} = \sum_{i=1}^{n}\left(\frac{X_i}{\sigma_0}\right)^2 \sim \chi_n^2 \text{ under } H_0.$$

An exact level α is thus obtained by choosing c so that

$$P(n\,a(c) < \chi_n^2 < n\,b(c)) = 1 - \alpha.$$

Finding c requires computation.
An easier approach is to reject H_0 when

$$\frac{n\hat{\sigma}^2}{\sigma_0^2} \leq \chi_n^2(\alpha/2) \text{ or } \frac{n\hat{\sigma}^2}{\sigma_0^2} \geq \chi_n^2(1 - \tfrac{\alpha}{2})$$

↓ ↙

These are the values which cut off probability $\alpha/2$ in the left and right tails of the χ_n^2 distn. (Given in tables.)

Example continued: A variation.

Find the LRT with size α of

$$H_0 : \sigma^2 \le \sigma_0^2 \quad \text{versus} \quad H_1 : \sigma^2 > \sigma_0^2$$

Now we have:

$\Theta_0 = (0, \sigma_0^2]$.

$$\underset{\sigma^2 \in \Theta_0}{\operatorname{argmax}} L(\sigma^2) \equiv \widehat{\sigma}_0^2 = \begin{cases} \widehat{\sigma}^2 & \text{if } \widehat{\sigma}^2 \le \sigma_0^2 \\ \sigma_0^2 & \text{if } \widehat{\sigma}^2 > \sigma_0^2 \end{cases}$$

> since the likelihood function falls away monotonically on each side of $\widehat{\sigma}^2$.

$$\lambda(\boldsymbol{x}) = \frac{L(\widehat{\sigma}_0^2)}{L(\widehat{\sigma}^2)} = \begin{cases} 1 & \text{if } \widehat{\sigma}^2 \le \sigma_0^2 \\ L(\sigma_0^2)/L(\widehat{\sigma}^2) & \text{if } \widehat{\sigma}^2 > \sigma_0^2 \end{cases}$$

$$= \begin{cases} 1 & \text{if } \widehat{\sigma}^2 \le \sigma_0^2 \\ \psi\left(\dfrac{\widehat{\sigma}^2}{\sigma_0^2}\right) & \text{if } \widehat{\sigma}^2 > \sigma_0^2 \end{cases}$$

Since $\psi(u)$ decreases for $u \ge 1$, we have $\lambda(\boldsymbol{x}) \le c$ iff $\widehat{\sigma}^2/\sigma_0^2 \ge c^*$ iff $S \equiv \sum_i X_i^2/\sigma_0^2 \ge c'$ where c' is chosen to give size α.

$$\sup_{\sigma^2 \in \Theta_0} P_{\sigma^2}(S \ge c') = P_{\sigma_0^2}(S \ge c') = \alpha$$

if we choose c' such that $P(\chi_n^2 \ge c') = \alpha$.

Example continued: Another variation

Observe $X_1, \ldots, X_n$ iid $N(\mu, \sigma^2)$. Find LRT of

$$H_0 : \sigma^2 = \sigma_0^2 , \mu \in \mathbb{R} \quad \text{versus} \quad H_1 : \sigma^2 \neq \sigma_0^2 , \mu \in \mathbb{R}$$

Now we have:

$\Theta = \{(\mu, \sigma^2) : \mu \in \mathbb{R}, \sigma^2 > 0\}$ and
$\Theta_0 = \{(\mu, \sigma^2) : \mu \in \mathbb{R}, \sigma^2 = \sigma_0^2\}$.

$L(\mu, \sigma^2) = (2\pi\sigma^2)^{-n/2} \exp\left(-(2\sigma^2)^{-1} \sum_i (x_i - \mu)^2\right)$.

$\underset{(\mu,\sigma^2)\in\Theta}{\text{argmax}}\, L(\mu, \sigma^2) = (\bar{x}, \widehat{\sigma}^2)$ where $\widehat{\sigma}^2 \equiv \frac{1}{n} \sum_i (x_i - \bar{x})^2$.

$\underset{(\mu,\sigma^2)\in\Theta_0}{\text{argmax}}\, L(\mu, \sigma^2) = (\bar{x}, \sigma_0^2)$

$$\lambda(\boldsymbol{x}) = \frac{L(\bar{x}, \sigma_0^2)}{L(\bar{x}, \widehat{\sigma}^2)} = \frac{(2\pi\sigma_0^2)^{-n/2} \exp\left(-(2\sigma_0^2)^{-1} n\widehat{\sigma}^2\right)}{(2\pi\widehat{\sigma}^2)^{-n/2} \exp\left(-n/2\right)}.$$

$$= e^{n/2} \left(\frac{\widehat{\sigma}^2}{\sigma_0^2}\right)^{n/2} \exp\left[-\frac{n}{2}\left(\frac{\widehat{\sigma}^2}{\sigma_0^2}\right)\right]$$

$$= \psi\left(\frac{\widehat{\sigma}^2}{\sigma_0^2}\right) \quad \text{where } \psi(u) \equiv e^{n/2} u^{n/2} e^{-(n/2)u}.$$

Just like before but with a different definition of $\widehat{\sigma}^2$. Now determine critical values using $n\widehat{\sigma}^2/\sigma_0^2 \sim \chi^2_{n-1}$ under H_0.

Example:

Observe independent samples:

$X_1, \ldots, X_m$ iid $N(\mu_x, \sigma^2)\,, \quad Y_1, \ldots, Y_n$ iid $N(\mu_y, \sigma^2)\,.$

Find the LRT of

$$H_0 : \mu_x = \mu_y \quad \text{versus} \quad H_1 : \mu_x \neq \mu_y\,.$$

Work:

$$L(\mu_x, \mu_y, \sigma^2) =$$
$$(2\pi\sigma^2)^{\frac{-(m+n)}{2}} e^{\frac{-1}{2\sigma^2}\left\{\sum_{i=1}^m (X_i - \mu_x)^2 + \sum_{j=1}^n (Y_i - \mu_y)^2\right\}}$$

$$\Theta : \underset{(\mu_x,\mu_y,\sigma^2)}{\text{argmax}} \; L(\mu_x, \mu_y, \sigma^2) = (\bar{x}, \bar{y}, \hat{\sigma}^2)$$

where

$$\hat{\sigma}^2 = \frac{1}{m+n}\left(\sum_{i=1}^m (X_i - \bar{x})^2 + \sum_{j=1}^n (Y_i - \bar{y})^2\right)$$

$$\Theta_0 : \underset{(\mu,\sigma^2)}{\text{argmax}} \, L(\mu, \mu, \sigma^2) = \left(\hat{\mu}_0\,,\, \hat{\sigma}_0^2\right)$$

where

$$
\begin{aligned}
\widehat{\mu}_0 &= \frac{1}{m+n}\left(\sum_i X_i + \sum_j Y_j\right) \\
\widehat{\sigma}^2 &= \frac{1}{m+n}\left(\sum_{i=1}^{m}(X_i - \widehat{\mu}_0)^2 + \sum_{j=1}^{n}(Y_i - \widehat{\mu}_0)^2\right)
\end{aligned}
$$