0.1 Likelihood Ratio Tests

Likelihood ratio tests are a very general approach to testing.

Let $f(x;\theta)$ be either a probability density function or a probability distribution where θ is a real valued parameter taking values in an interval Θ that could be the whole real line. We call Θ the **parameter space**. An alternative hypothesis H_1 will restrict the parameter θ to some subset Θ_1 of the parameter space Θ. The null hypothesis H_0 is then the complement of Θ_0 with respect to Θ. For instance, if $f(x;\theta)$ is the negative exponential distribution with pdf

$$f(x;\theta) = \frac{1}{\theta}e^{-x/\theta} \quad \text{for} \quad x > 0$$

the alternative hypothesis that specifies that the mean is not equal to 3 has $\Theta = \{\theta : 0 < \theta < \infty\}$ and $\Theta_1 = \{\theta : \theta \neq 3\}$ so $\Theta_0 = \{\theta : \theta = 3\}$.

Initially, we will confine our discussion to cases where the parameter is completely specified under the null hypothesis so $H_0 : \theta = \theta_0$ for some value θ_0 in the parameter space. That is, the null hypothesis is simple so $\Theta_0 = \{\theta : \theta = \theta_0\}$ consists of a single point.

We will test H_0 versus H_1 on the basis of random sample $X_1, X_2, ..., X_n$ from $f(x;\theta)$. If the null hypothesis holds, we would expect the likelihood

$$L(\theta) = \prod_{i=1}^{n} f(x_i;\theta)$$

to be relatively large, when evaluated at the prevailing value θ_0. The reference for judging this comparison is taken to be the maximum of the likelihood over the whole parameter space

$$L(\hat{\theta}) = max_{\theta\in\Theta} \prod_{i=1}^{n} f(x_i;\theta)$$

This will always be at least as large as the likelihood $L(\theta_0)$, evaluated at the particular value θ_0. However, we cannot discredit the null hypothesis unless $L(\theta_0)$ is much smaller than $L(\hat{\theta})$. The **likelihood ratio test** is actually based on the **likelihood ratio**

$$\lambda = \frac{L(\theta_0)}{L(\hat{\theta})} = \frac{\prod_{i=1}^{n} f(x_i;\theta_0)}{max_{\theta\in\Theta} \prod_{i=1}^{n} f(x_i;\theta)}$$

and the null hypothesis $H_0 : \theta = \theta_0$ is rejected if λ is small.

Example Let $X_1, ..., X_n$ be a random sample of size n from a normal distribution with known variance. Obtain the likelihood ratio for testing $H_0 : \mu = \mu_0$. versus $H_1 : \mu \neq \mu_0$.

Solution The likelihood function is

$$L(\mu|x_1, ..., x_n) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(x_i-\mu)^2/2\sigma^2} = e^{-n(\bar{x}-\mu)^2/2\sigma^2} \times \frac{1}{(2\pi\sigma^2)^{n/2}} e^{-\sum_{i=1}^{n}(x_i-\bar{x})^2/2\sigma^2}$$

This likelihood is maximized over all values of μ when the exponent $n(\bar{x} - \mu)^2/2\sigma^2$ is minimized. Therefore, the maximum likelihood estimator $\hat{\mu} = \bar{x}$ and

$$L(\hat{\mu}|x_1, ..., x_n) = \frac{1}{(2\pi\sigma^2)^{n/2}} e^{-\sum_{i=1}^{n}(x_i-\bar{x})^2/2\sigma^2}$$

The likelihood ratio

$$\lambda = \frac{L(\mu_0)}{L(\hat{\mu})}$$

$$= \frac{e^{-n(\bar{x}-\mu_0)^2/2\sigma^2} \times \frac{1}{(2\pi\sigma^2)^{n/2}} e^{-\sum_{i=1}^{n}(x_i-\bar{x})^2/2\sigma^2}}{\frac{1}{(2\pi\sigma^2)^{n/2}} e^{-\sum_{i=1}^{n}(x_i-\bar{x})^2/2\sigma^2}}$$

$$= e^{-n(\bar{x}-\mu_0)^2/2\sigma^2}$$

The likelihood ratio is just a function of $n(\bar{x} - \mu_0)^2/\sigma^2$ and will be small when this quantity is large. Since

$$Z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}$$

has a standard normal distribution under H_0, the likelihood ratio test rejects H_0 if

$$Z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} < -c \quad \text{or} \quad > c$$

where c is some constant.

In this example we see that minus twice the log likelihood ratio

$$-2log(\lambda) = Z^2$$

has a chi square distribution with 1 degree of freedom.

Example Let $X_1, ..., X_n$ be a random sample of size n from the Poisson distribution

$$p(x|\theta) = \frac{\theta^x e^{-\theta}}{x!}$$

Obtain the likelihood statistic for testing $H_0 : \theta = \theta_0$ versus $H_1 : \theta \neq \theta_0$

Solution The likelihood function is

$$L(\theta|x_1, ..., x_n) = \prod_{i=1}^{n} \frac{\theta^{x_i} e^{-\theta}}{x_i!} = \theta^{\sum_{i=1}^n x_i} e^{-n\theta} \times \frac{1}{\prod_{i=1}^n x_i!}$$

If $\sum_{i=1}^n x_i = 0$, then $L(\theta) = e^{-n\theta}$ has its maximum of 1 at $\hat{\theta} = 0$. Otherwise, the maximum likelihood estimator is $\hat{\theta} = \sum_{i=1}^n x_i/n = \bar{x}$. This same formula works for the special case.

$$Ł(\hat{\theta}) = \bar{x}^{\sum_{i=1}^n x_i} e^{-n\bar{x}} \times \frac{1}{\prod_{i=1}^n x_i!}$$

and the likelihood ratio is

$$\lambda = \frac{L(\theta_0)}{L(\hat{\theta})} = \left(\frac{\theta_0}{\bar{x}}\right)^{\sum x_i} e^{n(\bar{x}-\theta_0)} \quad \text{for} \quad \sum x_i > 0$$

and $\lambda = e^{-n\theta_0}$ otherwise.

Composite Null Hypothesis The likelihood ratio approach has to be modified slightly when the null hypothesis is composite. We use the same notation as above for θ, Θ_0, and Θ_1. When testing the null hypothesis $H_0 : \mu = \mu_0$ concerning a normal mean when σ^2 is unknown, the parameter space

$$\Theta = \{(\mu, \sigma^2) : -\infty < \mu < \infty, 0 < \sigma^2 < \infty\}$$

is a subset of 2-dimensional Euclidean space R^2. The null hypothesis is composite and

$$\Theta_0 = \{(\mu, \sigma^2) : \mu = \mu_0 \ , \quad 0 < \sigma^2 < \infty\}$$

which is a horizontal half line.

Since the null hypothesis is composite, it is not certain which value of the parameter(s) prevails even under H_0. Consequently, we take the maximum of the likelihood over Θ_0.

$$\max_{\theta \in \Theta_0} L(\theta) = \max_{\theta \in \Theta_0} \prod_{i=1}^{n} f(x_i; \theta)$$

This is again compared with with maximum of the likelihood over the whole parameter space. The generalized likelihood ratio statistic is defined as

$$\lambda = \frac{\max_{\theta \in \Theta_0} L(\theta)}{\max_{\theta \in \Theta} L(\theta)}$$

$$= \frac{\max_{\theta \in \Theta_0} \prod_{i=1}^{n} f(x_i; \theta)}{\max_{\theta \in \Theta} \prod_{i=1}^{n} f(x_i; \theta)}$$

and $H_0 : \theta \in \Theta_0$ is rejected for small values of λ. The exact null distribution is often difficult to obtain.

Example Let $X_1, ..., X_n$ be a random sample of size n from a normal distribution with unknown mean and variance. Obtain the likelihood ratio statistic for testing $H_0 : \mu = \mu_0$. versus $H_1 : \mu \neq \mu_0$.

Solution As above

$$\Theta = \{(\mu, \sigma^2) : -\infty < \mu < \infty, 0 < \sigma^2 < \infty\}$$

and

$$\Theta_0 = \{(\mu, \sigma^2) : \mu = \mu_0, 0 < \sigma^2 < \infty\}$$

The likelihood function is

$$L(\mu, \sigma^2 | x_1, ..., x_n) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(x_i - \mu)^2/2\sigma^2} = e^{-n(\bar{x}-\mu)^2/2\sigma^2} \times \frac{1}{(2\pi\sigma^2)^{n/2}} e^{-\sum_{i=1}^{n}(x_i - \bar{x})^2/2\sigma^2}$$

This likelihood is maximized, over Θ, by taking $\hat{\mu} = \bar{x}$ so the first term vanishes. Then, we find $\hat{\sigma}^2 = \sum_{i=1}^{n}(x_i - \bar{x})^2/n$ and the maximum of the likelihood is

$$L(\hat{\mu}, \hat{\sigma}^2) | x_1, ..., x_n) = \frac{1}{(2\pi\hat{\sigma}^2)^{n/2}} e^{-n/2}$$

Under $H_0 : \theta \in \Theta_0$, the likelihood

$$L(\mu_0) = \frac{1}{(2\pi\sigma^2)^{n/2}} e^{-\sum_{i=1}^{n}(x_i - \mu_0)^2/2\sigma^2}$$

where only σ^2 can be varied. We obtain the maximum likelihood estimator

$$\hat{\sigma_0}^2 = \frac{\sum_{i=1}^n (x_i - \mu_0)^2}{n}$$

where the subscript 0 reminds us that the maximum was taken over H_0. The maximum of the likelihood is

$$L(\mu_0, \hat{\sigma}_0^2) = \frac{1}{(2\pi\hat{\sigma}_0^2)^{n/2}} e^{-n/2}$$

Consequently, the likelihood ratio statistic is

$$\lambda = \frac{L(\mu_0, \hat{\sigma}_0^2)}{L(\hat{\mu}, \hat{\sigma}^2)}$$

$$= \left(\frac{\hat{\sigma}^2}{\hat{\sigma}_0^2} \right)^{n/2}$$

and the test rejects H_0 if this is small. Equivalently, the test rejects H_0 for large values of

$$\frac{\hat{\sigma}_0^2}{\hat{\sigma}^2} = \frac{\sum_{i=1}^n (x_i - \mu_0)^2}{\sum_{i=1}^n (x_i - \bar{x})^2} = \frac{\sum_{i=1}^n (x_i - \bar{x})^2 + n(\bar{x} - \mu_0)^2}{\sum_{i=1}^n (x_i - \bar{x})^2}$$

$$= 1 + \frac{n(\bar{x} - \mu_0)^2}{\sum_{i=1}^n (x_i - \bar{x})^2}$$

The fraction is just $T^2/(n-1)$ where T is the student's t variable

$$T = \frac{(\bar{X} - \mu_0)}{S/\sqrt{n}}$$

with $n-1$ degrees of freedom. That is, we have shown that the two-sided t test which rejects $H_0 : \mu = \mu_0$ if $T < t_{\alpha/2}$ or $T > t_{\alpha/2}$ is equivalent to the likelihood ratio test when the population is normal.

Most of the standard statistical tests that apply to normal distributions are likelihood ratio tests.

Large sample Distribution

In cases where the maximum likelihood estimators have a limiting normal distribution, the distribution of minus twice the log likelihood ratio

$$-2\ln(\lambda)$$

converges in distribution to a chi square distribution where the degrees of freedom are determined as

dimension(Θ) - dimension (Θ_0)

In the last example above, the degrees of freedom are 2 - 1 = 1.