Large-Sample Likelihood Ratio Tests

We will use the following hypothesis-testing framework. The data are $Y_1, \ldots, Y_n$. The distribution of these independent and identically distributed random variables depends on the parameter θ, and we are testing a null hypothesis H_0 using a large sample likelihood ratio test.

$$Y_1, \ldots, Y_n \overset{i.i.d.}{\sim} P_\theta,\ \theta \in \Theta,$$
$$H_0 : \theta \in \Theta_0 \text{ v.s. } H_A : \theta \in \Theta \cap \Theta_0^c,$$

The data have likelihood function

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

where $f(y_i; \theta)$ is the density or probability mass function evaluated at y_i.

Let $\widehat{\theta}$ denote the usual Maximum Likelihood Estimate (MLE). That is, it is the parameter value for which the likelihood function is greatest, over all $\theta \in \Theta$. And, let $\widehat{\theta}_0$ denote the *restricted* MLE. The restricted MLE is the parameter value for which the likelihood function is greatest, over all $\theta \in \Theta_0$. This MLE is *restricted* by the null hypothesis $H_0 : \theta \in \Theta_0$. It should be clear that $L(\widehat{\theta}_0) \leq L(\widehat{\theta})$, so that the *likelihood ratio.*

$$\lambda = \frac{L(\widehat{\theta}_0)}{L(\widehat{\theta})} \leq 1.$$

The likelihood ratio will equal one if and only if the overall MLE $\widehat{\theta}$ is located in Θ_0. In this case, there is no reason to reject the null hypothesis.

Usually, the likelihood ratio is strictly less than one. If it's a *lot* less than one, then the data are a lot less likely to have been observed under the null hypothesis than under the alternative hypothesis; if so, the the null hypothesis is questionable. This is the basis of the likelihood ratio tests.

If λ is small (close to zero), then $\ln \lambda$ is a large negative number, and $-2 \ln \lambda$ is a big positive number.

Tests will be based on

$$\begin{aligned} G &= -2 \ln \left(\frac{\max_{\theta \in \Theta_0} L(\theta)}{\max_{\theta \in \Theta} L(\theta)} \right) \\ &= -2 \ln \left(\frac{L(\widehat{\theta}_0)}{L(\widehat{\theta})} \right) \\ &= -2 \ln L(\widehat{\theta}_0) - [-2 \ln L(\widehat{\theta})]. \end{aligned} \tag{1}$$

Thus, the test statistic G is the *difference* between two -2 log likelihood functions. This means that to carry out a test, you can minimize $-2 \ln L(\theta)$ twice, first over all $\theta \in \Theta$, and then over all $\theta \in \Theta_0$. The test statistic is the difference between the two minimum values.

If the null hypothesis is true, then the test statistic G has, if the sample size is large, an approximate chisquare distribution, with degrees of freedom equal to the difference of the *dimension* of Θ and Θ_0. For example, if the null hypothesis is that 4 elements of θ equal zero, then the degrees of freedom are equal to 4. More generally, if the null hypothesis imposes k linear restrictions on θ, then the degrees of freedom equal k.

Think of the usual normal multiple regression model. Here, $\theta = (\boldsymbol{\beta}, \sigma^2)$. Consider the null hypothesis $H_0 : \mathbf{L}\boldsymbol{\beta} = \boldsymbol{\gamma}$, where $\mathbf{L}$ is a $k \times p$ matrix. This null hypothesis imposes k linear restrictions on the parameter, one for each row of $\mathbf{L}$. The dimension of Θ is $p+1$; the dimension of Θ_0 is $p+1-k$.

The p-value associated with the test statistic G is $Pr\{X > G\}$, where X is a chisquare random variable with k degrees of freedom. If $p < \alpha$, we reject H_0 and call the results "statistically significant."

Example Let $X_1, \ldots, X_{n_1}$ be a random sample from a Poisson distribution with parameter λ_1. Independently of the X values, let $Y_1, \ldots, Y_{n_2}$ be a random sample from a Poisson distribution with parameter λ_2. We will test $H_0 : \lambda_1 = \lambda_2$.

The parameter for this problem is $\theta = (\lambda_1, \lambda_2)$, and the null hypothesis imposes one linear restriction on the parameter. So, the degrees of freedom of the large-sample likelihood ratio chisquare test will equal one.

The likelihood function is

$$L(\theta) = \prod_{i=1}^{n_1} \frac{e^{-\lambda_1}\lambda_1^{x_i}}{x_i!} \prod_{i=1}^{n_2} \frac{e^{-\lambda_2}\lambda_2^{y_i}}{y_i!} = \frac{e^{-n_1\lambda_1}\lambda_1^{\sum_{i=1}^{n_1} x_i}}{\prod_{i=1}^{n_1} x_i!} \frac{e^{-n_2\lambda_2}\lambda_2^{\sum_{i=1}^{n_2} y_i}}{\prod_{i=1}^{n_2} y_i!}. \quad (2)$$

Partially differentiating the log (or -2 times the log) with respect to λ_1 and setting the result to zero, we get $\widehat{\lambda}_1 = \overline{x}$. Similarly, $\widehat{\lambda}_2 = \overline{y}$. Thus, we get $\widehat{\theta} = (\widehat{\lambda}_1, \widehat{\lambda}_2) = (\overline{x}, \overline{y})$.

Next we need to calculate the restricted MLE. There are two ways to do this, the easy way and the hard way. The hard way is to set $\lambda_1 = \lambda_2 = \lambda$ in (2), take the log and start differentiating with respect to λ. The smart way is to recognize that you've already done the problem once. With $\lambda_1 = \lambda_2 = \lambda$, this is just a single random sample from a Poisson distribution with parameter λ, and the MLE is the sample mean of all the data combined. That is,

$$\widehat{\lambda} = \frac{\sum_{i=1}^{n_1} x_i + \sum_{i=1}^{n_2} y_i}{n_1 + n_2} = \frac{n_1\overline{x} + n_2\overline{y}}{n_1 + n_2}.$$

The next step is to calculate the test statistic G as te difference between two -2 log likelihoods.