Likelihood Ratio Tests

Likelihood ratio tests (LRTs) have been used to compare two **nested** models. The form of the test is suggested by its name,

$$\mathbf{LRT = -2\,log}_e\left(\frac{\mathcal{L}_s(\hat{\theta})}{\mathcal{L}_g(\hat{\theta})}\right),$$

the ratio of two likelihood functions; the simpler model (s) has fewer parameters than the general (g) model. Asymptotically, the test statistic is distributed as a chi-squared random variable, with degrees of freedom equal to the difference in the number of parameters between the two models.

Likelihood ratio tests compare two models provided the simpler model is a special case of the more complex model (i.e., "nested"). LRTs can be presented as a difference in the log-likelihoods (recall that log(A/B) = logA – logB) and this is often handy as they can be expressed in terms of deviance. Then,

$$\begin{aligned}\text{LRT} &= -2\left(\log_e(\mathcal{L}_\text{s}) - \log_e(\mathcal{L}_\text{g})\right)\\ &= -2\log_e(\mathcal{L}_\text{s}) + 2\log_e(\mathcal{L}_\text{g})\\ &= \text{deviance}_\text{s} - \text{deviance}_\text{g}.\end{aligned}$$

Thus, the LRT can be computed as a difference in the deviance for the two models (ignoring the term for the saturated model). This is convenient as the deviance is a value of interest in other respects.

Say we flipped two coins, we could have

$$\text{H}_\text{o}: p_1 = p_2 \equiv p \qquad \text{(the simpler model, 1 parameter)},$$

vs. the alternative of different probabilities of heads, hence

$$\text{H}_\text{a}: p_1 \neq p_2. \qquad \text{(the more general model, 2 parameters)}.$$

Compute MLEs under both models and compute the deviances (D),

$$D_s = -2\log_e(\mathcal{L}(\hat{p})), \qquad K_s = 1, \text{ the number of parameters}$$

$$D_g = -2\log_e(\mathcal{L}(\hat{p}_1, \hat{p}_2)), \qquad K_g = 2, \text{ the number of parameters.}$$

The the LRT $= D_s - D_g$, $\text{df} = K_g - K_s = 1$.

Thus, this test statistic in approximately χ^2 with 1 df under the null hypothesis. The approximation improves as sample size increases. Note, too that the log-likelihood for the saturated model is a constant and the same for both of the above models; thus it was deleted in this example.

Testing of null hypotheses has seen decreasing use in many areas of applied science over the past 2 decades. We will made some reference to LRTs so that students can better understand existing literature that makes use of these methods. Problems with statistical hypothesis testing will be outlined at a latter point.