Topic 17: Extensions on the Likelihood Ratio

November 17, 2009

1 One-Sided Tests

We begin with a composite hypothesis test

$$H_0 : \theta \in \Theta_0 \quad \text{versus} \quad H_1 : \theta \in \Theta_1$$

with $\Theta_0 \cap \Theta_1 = \emptyset$ and $\Theta_0 \cup \Theta_1 = \Theta$. Let C be the critical region for an α level test.

One good property of a test is that its power function π is greater for any value of θ than the power function of any other test for any value of θ. Such a test is called **uniformly most powerful**.

In general, a hypothesis will not have a uniformly most powerful test. However, in several procedures involving simple hypothesis, the test statistic did not depend on the specific value of the alternative. For example, in the case of independent normal data with unknown mean μ and known variance σ^2, we have simple hypothesis

$$H_0 : \mu = \mu_0 \quad \text{versus} \quad H_1 : \mu = \mu_1$$

with $\mu_1 > \mu_0$. The critical region is determined by the likelihood ratio test is

$$C = \{\mathbf{x}; \bar{x} \geq k_\alpha\}.$$

irrespective of the value of μ_1.

In this case, the power function $\pi(\mu_1) = P_{\mu_1}\{\bar{X} \geq k_\alpha\}$. increases as μ_1 increases and so the test becomes more powerful with increasing μ_1.

In general, we look for a test statistic $T(\mathbf{x})$ (like $\bar{x}$ in the example above. Next, we check that the likelihood ratio,

$$\frac{L(\theta_2|\mathbf{x})}{L(\theta_1|\mathbf{x})}, \quad \theta_1 < \theta_2.$$

depends on value of the data $\mathbf{x}$ only through the value of statistic $T(\mathbf{x})$ and, in addition, this ratio is a monotone function of $T(\mathbf{x})$. If these conditions hold, then $C = \{\mathbf{x}; T(\mathbf{x}) \geq k_\alpha\}$ is the critical region for a uniformly most powerful α level test for the one-sided alternative hypothesis

$$H_0 : \theta \leq \theta_0 \quad \text{versus} \quad H_1 : \theta > \theta_0.$$

These conditions are satisfied for the case above as well as the tests for p, the probability of success in Bernoulli trials.

Example 1 (One sample one proportion z-test)**.** *For the hypotheses*

$$H_0 : p \leq 0.7 \quad \text{versus} \quad H_1 : p > 0.7.$$

for a test of the probability that a feral bee hive survives a winter. The test statistic is still

$$Z = \frac{\bar{X} - p}{\sqrt{p(1-p)/n}}$$

and, again, for an α level test, the critical value is z_α where α is the probability that a standard normal is at least z_α

```
> prop.test(88,112,0.7,alternative=c("greater"))

1-sample proportions test with continuity correction

data:  88 out of 112, null probability 0.7
X-squared = 3.5208, df = 1, p-value = 0.0303
alternative hypothesis: true p is greater than 0.7
95 percent confidence interval:
 0.7107807 1.0000000
sample estimates:
        p
0.7857143
```

2 Likelihood Ratio Tests

The likelihood ratio test is a popular choice for composite hypothesis. Θ_0 is a subspace of the whole parameter space.

$$\Lambda(\mathbf{x}) = \frac{\sup\{L(\theta|\mathbf{x}); \theta \in \Theta_0\}}{\sup\{L(\theta|\mathbf{x}); \theta \in \Theta\}}$$

The rejection region for an α-level test is $\{\Lambda(\mathbf{x}) \leq \lambda_0\}$ where λ_0 is chosen so that

$$P_\theta\{\Lambda(X) \leq \lambda_0\} \leq \alpha \text{ for all } \theta \in \Theta_0.$$

Let $\hat{\theta}_0$ be the parameter value that maximizes the likelihood for $\theta_0 \in \Theta_0$ and $\hat{\theta}$ be the parameter value that maximizes the likelihood for $\theta_0 \in \Theta$. Then,

$$\Lambda(\mathbf{x}) = \frac{L(\hat{\theta}_0|\mathbf{x})}{L(\hat{\theta}|\mathbf{x})}.$$

Example 2. *Let $\Theta = \mathbb{R}$ and consider the* **two-sided hypothesis**

$$H_0 : \mu = \mu_0 \quad \text{versus} \quad H_1 : \mu \neq \mu_0.$$

Here the data are n independent $N(\mu, \sigma^2)$ random variables $X_1 \ldots, X_n$ with known variance σ^2. Then, $\hat{\mu}_0 = \mu_0$ and $\hat{\mu} = \bar{x}$. Consequently,

$$L(\hat{\mu}_0|\mathbf{x}) = \left(\frac{1}{2\pi\sigma^2}\right)^n \exp -\frac{1}{2\sigma^2}\sum_{i=1}^n (x_i - \mu_0)^2, \quad L(\hat{\mu}|\mathbf{x}) = \left(\frac{1}{2\pi\sigma^2}\right)^n \exp -\frac{1}{2\sigma^2}\sum_{i=1}^n (x_i - \bar{x})^2$$

and

$$\Lambda(\mathbf{x}) = \exp -\frac{1}{2\sigma^2}\left(\sum_{i=1}^n ((x_i - \mu_0)^2 - (x_i - \bar{x})^2)\right) = \exp -\frac{n}{2\sigma^2}(\bar{x} - \mu_0)^2.$$

Now notice that

$$-2\ln\Lambda(\mathbf{x}) = \frac{n}{\sigma^2}(\bar{x} - \mu_0)^2 = \left(\frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}\right)^2.$$

Because $(\bar{X} - \mu_0)/(\sigma/\sqrt{n})$ is a standard normal random variable, $-2\ln\Lambda(X)$ is the square of a standard normal, hence, a χ-square random variable with 1 degree of freedom.

Naturally we can use both

$$\left(\frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}\right)^2 \quad \text{and} \quad \left|\frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}\right|.$$

as a test statistic. For the first, the critical value is just the square of the critical value for the second choice.

3 Chi-square test

This exact computation for normal data yields, owing to the central limit theorem, an asymptotic result that is contained in the following theorem.

Theorem 3. *Whenever the maximum likelihood estimate has an asymptotically normal distribution, let $\Lambda_n(\mathbf{x})$ be the likelihood ratio criterion for an n dimensional parameter space:*

$$H_0 : \theta_1 = c_1 \text{ for all } i = 1, \dots, k \quad \text{versus} \quad H_1 : \theta_1 \neq c_1 \text{ for some } i = 1, \dots, k$$

Then under H_0,

$$-2 \ln \Lambda_n(X)$$

converges in distribution to a χ^2_{n-k} random variable.