R Handout: Likelihood Ratio Test

This handout will take the Iris data as an example to show the computation and application of Likelihood ratio test in practice. Suppose we are interested in the following hypothesis testing for the mean of sepal width of Virginica (from the Iris data set) $H_0 : \mu = \mu_0 = 3$ vs. $H_1 : \mu = \mu_1 \neq 3$ under the assumption that the observations X_i's are i.i.d. the normal distribution with $N(\mu, \sigma^2 = 0.1040)$. Also, let the significance level be $\alpha = 0.05$.

1. THE CONCEPT OF THE LIKELIHOOD RATIO TEST

1. Exact Method

   The likelihood

   $$L(\mu; \underline{x}) = \prod_{i=1}^{n} (2\pi\sigma^2)^{-1/2} e^{-(x_i - \mu)^2/(2\sigma^2)}$$

   The log-likelihood

   $$lnL(\mu; \underline{x}) = -\frac{n}{2}ln(2\pi\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - \mu)^2 \tag{1}$$

   It is easy to show the m.l.e. be $\hat{\mu} = \bar{x}$. The log-likelihood ratio

   $$\delta = ln(\frac{L(\mu_0; \underline{x})}{L(\hat{\mu}; \underline{x})}) = lnL(\mu_0; \underline{x}) - lnL(\hat{\mu}; \underline{x})$$

   $$= -\frac{n}{2}ln(2\pi\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - \mu_0)^2 + (\frac{n}{2}ln(2\pi\sigma^2) + \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - \hat{\mu})^2)$$

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

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

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

1

So

$$\delta = -\frac{n}{2\sigma^2}(\bar{x} - \mu_0)^2$$

Because $\bar{X} \sim N(\mu, \frac{\sigma^2}{n})$, $\frac{\bar{X}-\mu}{\sqrt{\sigma^2/n}} \sim N(0,1) \Rightarrow (\frac{\bar{X}-\mu}{\sqrt{\sigma^2/n}})^2 \sim \chi_1^2$, according to the equal-tail assumption, the size

$$P(-2\delta \leq C_1 \cup -2\delta \geq -C_2|H_0) = P(-2\delta \leq C_1 \cup -2\delta \geq C_2|\mu = \mu_0) = \alpha$$

Under the null hypothesis, let $\Delta = -2\delta \sim \chi_1^2$. The critical region

$$\Rightarrow \{\Delta \leq 0.0010 \cup \Delta \geq 5.0239\}$$

Wilks (1938) [2] also proved that, in general, Δ is asymptotically following the chi-squared distribution with the degrees of freedom equal to the difference between the degrees of freedom under the alternative hypothesis and the degrees of freedom under the null hypothesis. This result is called **Wilks' Theorem**.

2. Parametric Bootstrap

Suppose the analytical form of Δ is too complicated, we have no clue w

3. Non-parametric Bootstrap

4. Comparison

2. APPLICATION: NAIVE BAYES CLASSIFIER

■References References

[1] Wilks, Samuel S. (1938) *The Large-Sample Distribution of the Likelihood Ratio for Testing Composite Hypotheses*, The Annals of Mathematical Statistics. **9**: 60 - 62.

[2] @miscDua:2017 , author = "Dheeru, Dua and Karra Taniskidou, Efi", year = "2017", title = "UCI Machine Learning Repository", url = "http://archive.ics.uci.edu/ml", institution = "University of California, Irvine, School of Information and Computer Sciences" https://archive.ics.uci.edu/ml/datasets/iris