Section 11, 11/11/19

Generalized Likelihood Ratio Test (GLRT) (page 339)

In cases where we are testing a simple vs composite hypothesis, we can resort the Generalized Likelihood Ratio Test. Consider

$H_0 : \theta = \theta_0$
$H_1 : \theta \neq \theta_0$

Given a set of observations $X = (X_1, \ldots, X_n)$ We evaluate the ratio of likelihoods

$\Lambda^* = \frac{f(X|\theta_0)}{\max_{\theta \in \Theta} f(X|\theta)}$

The numerator is simply the likelihood at the null hypothesis parameter. The denominator the likelihood at the MLE, NOT the argmax of the parameters. The parameter space for the MLE is Θ, specified by the alternative hypothesis.

Example A

Testing a Normal Mean
Let $X_1, \ldots, X_n$ be i.i.d. and normally distributed with mean μ and variance σ^2, where σ is known. We wish to test H_0: $\mu = \mu_0$ against H_1: $\mu \neq \mu_0$, where μ_0 is a prescribed number. The role of θ is played by μ, and $\omega_0 = \{\mu_0\}$, $\omega_1 = \{\mu | \mu \neq \mu_0\}$, and $\Omega = \{-\infty < \mu < \infty\}$.

Since ω_0 consists of only one point, the numerator of the likelihood ratio statistic is

$$\frac{1}{(\sigma\sqrt{2\pi})^n}e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(X_i-\mu_0)^2}$$

For the denominator, we have to maximize the likelihood for $\mu \in \Omega$, which is achieved when μ is the mle $\overline{X}$. The denominator is the likelihood of $X_1, X_2, \ldots, X_n$ evaluated with $\mu = \overline{X}$:

$$\frac{1}{(\sigma\sqrt{2\pi})^n}e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(X_i-\overline{X})^2}$$

The likelihood ratio statistic is, therefore,

$$\Lambda = \exp\left(-\frac{1}{2\sigma^2}\left[\sum_{i=1}^{n}(X_i-\mu_0)^2 - \sum_{i=1}^{n}(X_i-\overline{X})^2\right]\right)$$

Rejecting for small values of Λ is equivalent to rejecting for large values of

$$-2\log\Lambda = \frac{1}{\sigma^2}\left(\sum_{i=1}^{n}(X_i-\mu_0)^2 - \sum_{i=1}^{n}(X_i-\overline{X})^2\right)$$

Using the identity

$$\sum_{i=1}^{n}(X_i-\mu_0)^2 = \sum_{i=1}^{n}(X_i-\overline{X})^2 + n(\overline{X}-\mu_0)^2$$

we see that the likelihood ratio test rejects for large values of $-2\log\Lambda = n(\overline{X} - \mu_0)^2/\sigma^2$. The distribution of this statistic under H_0 is chi-square with 1 degree of freedom. This follows, since under H_0, $\overline{X} \sim N(\mu_0, \sigma^2/n)$, which implies that $\sqrt{n}(\overline{X}-\mu_0)/\sigma \sim N(0, 1)$ and hence its square, $-2\log\Lambda \sim \chi_1^2$. Knowing the null distribution of the test statistic makes possible the construction of a rejection region for any significance level α: The test rejects when

$$\frac{n}{\sigma^2}(\overline{X}-\mu_0)^2 > \chi_1^2(\alpha)$$

Again using the fact that a chi-square random variable with 1 degree of freedom is the square of a standard normal random variable, we can rewrite this relation to show that the rejection region for the test is

$$|\overline{X}-\mu_0| \geq \frac{\sigma}{\sqrt{n}}z(\alpha/2) \qquad \blacksquare$$

Wilk's Theorem

Under appropriate smoothness conditions of the probability density, the null distribution of $-2\log\Lambda$ approaches a chi-squared distribution with $dim\Theta_1 - dim\Theta_0$ degrees of freedom as $n \to \infty$.

Notice that for the prior example, $dim\Theta_1 - dim\Theta_0 = 1 - 0 = 1$ degree of freedom as was shown.

Distributions Based on the Normal (page 193)

t-distribution

If $Z \sim N(0,1)$ and $U \sim \chi^2_n$ are independent, then $Z/\sqrt{U/n}$ follows a t-dist with n dof. By math,

$f(t) = \frac{\Gamma[(n+1)/2]}{\sqrt{n\pi}\Gamma(n/2)}\left(1 + \frac{t^2}{n}\right)^{-(n+1)/2}$

This distribution is symmetric about zero and as dof $\to \infty$, t-distribution $\to N(0,1)$

f-distribution

If U, V are independent chi-square random variables with m and n dof respectively, then $W = \frac{U/m}{V/n}$ follows the F-dist with m and n dof $F_{m,n}$.

chi-squared distribution

Let $Z \sim N(0,1)$, the standard normal distribution, then $U = Z^2$ is the chi-squared distribution with 1 degree of freedom (dof), i.e. χ^2_1. This distribution is equivalent to $\Gamma(\frac{1}{2}, \frac{1}{2})$.

If $U_1, U_2, \ldots, U_n$ are i.i.d. chi-squared random variables with 1 dof, then $V = U_1 + \ldots + U_n \sim \chi^2_n$, the chi-square dist with n dof, equivalent to $\Gamma(\frac{n}{2}, \frac{1}{2})$ as the sum of independent Gammas with the same λ parameter is also a Gamma with λ. Also,

- $E(V) = n$
- Var$(V) = n$

and if $U \sim \chi^2_n$, $V \sim \chi^2_m$ are independent, then $U + V \sim \chi^2_{m+n}$.

Theorem A:

The random variable $\bar{X}$ and the vector $(X_1 - \bar{X}, X_2 - \bar{X}, \ldots, X_n - \bar{X})$ of random variables are independent.

Proof

At the level of this course, it is difficult to give a proof that provides sufficient insight into why this result is true; a rigorous proof essentially depends on geometric properties of the multivariate normal distribution, which this book does not cover. We present a proof based on moment-generating functions; in particular, we will show that the joint moment-generating function

$$M(s, t_1, \dots, t_n) = E\{\exp[s\overline{X} + t_1(X_1 - \overline{X}) + \cdots + t_n(X_n - \overline{X})]\}$$

factors into the product of two moment-generating functions—one of $\overline{X}$ and the other of $(X_1 - \overline{X}), \dots, (X_n - \overline{X})$. The factoring implies (Section 4.5) that the random variables are independent of each other and is accomplished through some algebraic trickery. First we observe that since

$$\sum_{i=1}^{n} t_i(X_i - \overline{X}) = \sum_{i=1}^{n} t_i X_i - n\overline{X}\,\bar{t}$$

then

$$s\overline{X} + \sum_{i=1}^{n} t_i(X_i - \overline{X}) = \sum_{i=1}^{n} \left[\frac{s}{n} + (t_i - \bar{t})\right] X_i$$

$$= \sum_{i=1}^{n} a_i X_i$$

where

$$a_i = \frac{s}{n} + (t_i - \bar{t})$$

Furthermore, we observe that

$$\sum_{i=1}^{n} a_i = s$$

$$\sum_{i=1}^{n} a_i^2 = \frac{s^2}{n} + \sum_{i=1}^{n} (t_i - \bar{t})^2$$

Now we have

$$M(s, t_1, \dots, t_n) = M_{X_1 \cdots X_n}(a_1, \dots, a_n)$$

and since the X_i are independent normal random variables, we have

$$\begin{aligned}
M(s, t_1, \ldots, t_n) &= \prod_{i=1}^{n} M_{X_i}(a_i) \\
&= \prod_{i=1}^{n} \exp\left(\mu a_i + \frac{\sigma^2}{2} a_i^2\right) \\
&= \exp\left(\mu \sum_{i=1}^{n} a_i + \frac{\sigma^2}{2} \sum_{i=1}^{n} a_i^2\right) \\
&= \exp\left[\mu s + \frac{\sigma^2}{2}\left(\frac{s^2}{n}\right) + \frac{\sigma^2}{2} \sum_{i=1}^{n} (t_i - \bar{t})^2\right] \\
&= \exp\left(\mu s + \frac{\sigma^2}{2n} s^2\right) \exp\left[\frac{\sigma^2}{2} \sum_{i=1}^{n} (t_i - \bar{t})^2\right]
\end{aligned}$$

The first factor is the mgf of $\overline{X}$. Since the mgf of the vector $(X_1 - \overline{X}, \ldots, X_n - \overline{X})$ can be obtained by setting $s = 0$ in M, the second factor is this mgf. ■