LECTURE 17

Chapter 9.3: The duality of confidence intervals and hypothesis tests

There is a duality between confidence intervals and hypothesis tests.

Example A Let $X_1, \ldots, X_n$ be a random sample from a normal distribution having unknown mean μ and known variance σ^2.

We are interested in testing

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

At significance level α, consider the following test (or decision rule):
Reject H_0 if $|\bar{X} - \mu_0| > \sigma z(\alpha/2)/\sqrt{n}$, and do not reject (or accept as in the textbook) H_0 otherwise.

Thus the test accepts H_0 when

$$|\bar{X} - \mu_0| < \sigma z(\alpha/2)/\sqrt{n},$$

or equivalently,

$$-\sigma z(\alpha/2)/\sqrt{n} < \bar{X} - \mu_0 < \sigma z(\alpha/2)/\sqrt{n}.$$

The latter statement is also equivalent to a $100(1-\alpha)\%$ CI for μ_0 given by

$$\bar{X} - \sigma z(\alpha/2)/\sqrt{n} < \mu_0 < \bar{X} + \sigma z(\alpha/2)/\sqrt{n}.$$

In other words, the CI consists precisely of all those values of μ_0 for which the null hypothesis $H_0 : \mu = \mu_0$ is accepted.

The theorem below shows that the duality between CIs and hypothesis tests holds more generally.

Theorem A Suppose that for every value θ_0 in Θ there is a test at level α of the hypothesis $H_0 : \theta = \theta_0$.

Denote the acceptance region of the test by $A(\theta_0)$. Then the set

$$C(X) = \{\theta : X \in A(\theta)\}$$

is a $100(1-\alpha)\%$ confidence region for θ.

Proof Because A is the acceptance region of a test at level α,

$$P[X \in A(\theta)|\theta = \theta_0] = 1 - \alpha.$$

Now $\theta \in C(X)$ if and only if $X \in A(\theta)$. Hence

$$\begin{aligned} P[\theta_0 \in C(X)|\theta = \theta_0] &= P[X \in A(\theta_0)|\theta = \theta_0] \\ &= 1 - \alpha, \end{aligned}$$

by the definition of $C(x)$. $\square$

To state Theorem A in words: A $100(1-\alpha)\%$ confidence region (or interval if θ is a scalar) for θ consists of all those values for which the hypothesis that $\theta = \theta_0$ will not be rejected at level α.

Theorem B Suppose that $C(X)$ is a $100(1-\alpha)\%$ confidence region for θ; i.e. for every θ_0,

$$P[\theta_0 \in C(X)|\theta = \theta_0] = 1 - \alpha.$$

Then an acceptance region for a test at level α of the hypothesis $H_0 : \theta = \theta_0$ is

$$A(\theta_0) = \{X|\theta_0 \in C(X)\}.$$

Proof The test has level α because

$$\begin{aligned} & P[X \in A(\theta_0)|\theta = \theta_0] \\ = \; & P[\theta_0 \in C(X)|\theta = \theta_0] \\ = \; & 1 - \alpha. \end{aligned}$$

This proves Theorem B. $\square$

Chapter 9.4: Generalized likelihood ratio tests

Recall that the likelihood ratio test is optimal for testing a simple null hypothesis versus a simple alternative hypothesis.

The **generalized likelihood ratio test** extends the idea of a likelihood ratio test for testing a composite null hypothesis H_0 versus a composite alternative hypothesis H_1.

More specifically, let $X_1, \ldots, X_n$ be an i.i.d. sample with joint probability density function $f(x|\theta)$, $x = (x_1, \ldots, x_n)$.

The null hypothesis is

$$H_0 : \theta \in \omega_0$$

where ω_0 is a subset of the set of all possible values of θ.

The alternative hypothesis is

$$H_1 : \theta \in \omega_1$$

where ω_1 is another subset of the set of all possible values of θ and $\omega_0 \cap \omega_1 = \emptyset$. Let

$$\Omega = \omega_0 \cup \omega_1.$$

The generalized likelihood ratio test statistic is defined as

$$\Lambda = \frac{\max_{\theta \in \omega_0} \text{lik}(\theta)}{\max_{\theta \in \Omega} \text{lik}(\theta)}.$$

From the form of Λ, we deduce that small values of Λ tend to discredit H_0.

Then the generalized likelihood ratio test would reject H_0 if

$$\Lambda \leq \lambda_0,$$

where λ_0 is a constant determined by

$$P(\Lambda \leq \lambda_0 | H_0) = \alpha,$$

the desired significance level of the test.

Note

- A generalized likelihood ratio test is a likelihood ratio test when ω_0, ω_1 are singletons (i.e. each set consisting of only one value of θ).
- Generalized likelihood ratio tests are not typically optimal but they perform reasonally well and have much broader utility (in terms of applications).

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

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

where μ_0 is a prescribed number. Here

$$\begin{aligned} \omega_0 &= \{\mu_0\}, \\ \omega_1 &= \{\mu | \mu \neq \mu_0\}, \\ \Omega &= \mathbb{R}. \end{aligned}$$

We observe that

$$\max_{\mu \in \omega_0} \text{lik}(\mu)$$
$$= \frac{1}{(2\pi\sigma^2)^{n/2}} e^{-\sum_{i=1}^n (X_i - \mu_0)^2/(2\sigma^2)}.$$

Likewise we have

$$\max_{\mu \in \Omega} \text{lik}(\mu)$$
$$= \frac{1}{(2\pi\sigma^2)^{n/2}} e^{-\sum_{i=1}^n (X_i - \bar{X})^2/(2\sigma^2)},$$

since the mle of μ is $\bar{X}$. Thus the generalized likelihood ratio statistic is

$$\Lambda = \frac{\max_{\theta \in \omega_0} \text{lik}(\theta)}{\max_{\theta \in \Omega} \text{lik}(\theta)}$$
$$= e^{-\frac{1}{2\sigma^2}[\sum_{i=1}^n (X_i - \mu_0)^2 - \sum_{i=1}^n (X_i - \bar{X})^2]}.$$

Also

$$
\begin{aligned}
& -2\log\Lambda \\
= & \frac{1}{\sigma^2}[\sum_{i=1}^{n}(X_i-\mu_0)^2-\sum_{i=1}^{n}(X_i-\bar{X})^2] \\
= & \frac{n(\bar{X}-\mu_0)^2}{\sigma^2}.
\end{aligned}
$$

Note that small values of Λ corresponds to large values of $-2\log\Lambda$.

Under H_0, $\bar{X} \sim N(\mu_0, \sigma^2/n)$. Hence

$$-2\log\Lambda \sim \chi_1^2.$$

This implies that, at significance level α, the generalized likelihood ratio test rejects H_0 if

$$-2\log\Lambda = \frac{n(\bar{X}-\mu_0)^2}{\sigma^2} > \chi_1^2(\alpha),$$

where $\chi_1^2(\alpha)$ is the $100(1-\alpha)$th percentile of the χ_1^2 distribution.

Taking square roots, we get the rejection region of the generalized likelihood ratio test to be

$$|\bar{X} - \mu_0| > \frac{\sigma}{\sqrt{n}} z(\alpha/2).$$