Lecture 8. Composite hypotheses

Composite hypotheses, types of error and power

- For composite hypotheses like \( H : \theta \geq 0 \), the error probabilities do not have a single value.
- Define the power function \( W(\theta) = P(X \in C | \theta) = P(\text{reject } H_0 | \theta) \).
- We want \( W(\theta) \) to be small on \( H_0 \) and large on \( H_1 \).
- Define the size of the test to be \( \alpha = \sup_{\theta \in \Theta_0} W(\theta) \).
- For \( \theta \in \Theta_1 \), \( 1 - W(\theta) = P(\text{Type II error} | \theta) \).
- Sometimes the Neyman–Pearson theory can be extended to one-sided alternatives.
- For example, in Example 7.3 we have shown that the most powerful size \( \alpha \) test of \( H_0 : \mu = \mu_0 \) versus \( H_1 : \mu = \mu_1 \) (where \( \mu_1 > \mu_0 \)) is given by \( C = \{ x : \sqrt{n}(\bar{x} - \mu_0)/\sigma_0 > z_\alpha \} \).
- This critical region depends on \( \mu_0, n, \sigma_0, \alpha \), on the fact that \( \mu_1 > \mu_0 \), but not on the particular value of \( \mu_1 \).

Example 8.2

Suppose \( X_1, \ldots, X_n \) are iid \( N(\mu_0, \sigma_0^2) \) where \( \sigma_0 \) is known, and we wish to test \( H_0 : \mu \leq \mu_0 \) against \( H_1 : \mu > \mu_0 \).

- First consider testing \( H_0' : \mu = \mu_0 \) against \( H_1' : \mu = \mu_1 \) where \( \mu_1 > \mu_0 \) (as in Example 7.3).
- As in Example 7.3, the Neyman-Pearson test of size \( \alpha \) of \( H_0' \) against \( H_1' \) has \( C = \{ x : \sqrt{n}(\bar{x} - \mu_0)/\sigma_0 > z_\alpha \} \).
- We will show that \( C \) is in fact UMP for the composite hypotheses \( H_0 \) against \( H_1 \).
- For \( \mu \in \mathbb{R} \), the power function is

\[
W(\mu) = P(\text{reject } H_0) = P_{\mu} \left( \frac{\sqrt{n}(\bar{x} - \mu_0)}{\sigma_0} > z_\alpha \right) \\
= P_{\mu} \left( \frac{\sqrt{n}(\bar{x} - \mu)}{\sigma_0} > z_\alpha + \frac{\sqrt{n}(\mu_0 - \mu)}{\sigma_0} \right) \\
= 1 - \Phi \left( z_\alpha + \frac{\sqrt{n}(\mu_0 - \mu)}{\sigma_0} \right).
\]
Generalised likelihood ratio tests

We now consider likelihood ratio tests for more general situations.

Define the likelihood of a composite hypothesis $H : \theta \in \Theta$ given data $x$ to be

$$L_x(H) = \sup_{\theta \in \Theta} f(x|\theta).$$

So far we have considered disjoint hypotheses $\Theta_0, \Theta_1$, but often we are not interested in any specific alternative, and it is easier to take $\Theta_1 = \Theta$ rather than $\Theta_1 = \Theta \setminus \Theta_0$.

Then

$$\Lambda_x(H_0; H_1) = \frac{L_x(H_1)}{L_x(H_0)} = \sup_{\theta \in \Theta_1} \frac{f(x|\theta)}{\sup_{\theta \in \Theta_0} f(x|\theta)} \geq 1,$$

(1)

with large values of $\Lambda_x$ indicating departure from $H_0$.

Notice that if $\Lambda_x^* = \sup_{\theta \in \Theta_1 \setminus \Theta_0} f(x|\theta) / \sup_{\theta \in \Theta_0} f(x|\theta)$, then $\Lambda_x = \max\{1, \Lambda_x^*\}$.

We know $W(\mu_0) = \alpha$ (just plug in)

$W(\mu)$ is an increasing function of $\mu$.

So $\sup_{\mu \leq \mu_0} W(\mu) = \alpha$, and (i) is satisfied.

For (ii), observe that for any $\mu > \mu_0$, the Neyman Pearson size $\alpha$ test of $H_0'$ vs $H_1'$ has critical region $C$ (the calculation in Example 7.3 depended only on the fact that $\mu > \mu_0$ and not on the particular value of $\mu_1$).

Let $C^*$ and $W^*$ belong to any other test of $H_0$ vs $H_1$ of size $\leq \alpha$.

Then $C^*$ can be regarded as a test of $H_0$ vs $H_1$ of size $\leq \alpha$, and NP-Lemma says that $W^*(\mu_1) \leq W(\mu_1)$.

This holds for all $\mu_1 > \mu_0$ and so (ii) is satisfied.

So $C$ is UMP size $\alpha$ for $H_0$ vs $H_1$. □

Example 8.3

Single sample: testing a given mean, known variance ($z$-test). Suppose that $X_1, \ldots, X_n$ are iid $N(\mu, \sigma_0^2)$, with $\sigma_0^2$ known, and we wish to test $H_0 : \mu = \mu_0$ against $H_1 : \mu \neq \mu_0$ ($\mu_0$ is a given constant).

Here $\Theta_0 = \{\mu_0\}$ and $\Theta = \mathbb{R}$.

For the denominator in (1) we have $\sup_{\Theta_0} f(x|\mu) = f(x|\mu_0)$.

For the numerator, we have $\sup_{\Theta_0} f(x|\mu) = f(x|\hat{\mu})$, where $\hat{\mu}$ is the mle, so $\hat{\mu} = \bar{x}$ (check).

Hence

$$\Lambda_x = \frac{(2\pi\sigma_0^2)^{-n/2} \exp \left(-\frac{1}{2\sigma_0^2} \sum (x_i - \bar{x})^2 \right)}{(2\pi\sigma_0^2)^{-n/2} \exp \left(-\frac{1}{2\sigma_0^2} \sum (x_i - \mu_0)^2 \right)},$$

and we reject $H_0$ if $\Lambda_x$ is 'large.'

We find that

$$2 \log \Lambda_x = \frac{1}{\sigma_0^2} \left[ \sum (x_i - \mu_0)^2 - \sum (x_i - \bar{x})^2 \right] = \frac{n}{\sigma_0^2} (\bar{x} - \mu_0)^2.$$ (check)

Thus an equivalent test is to reject $H_0$ if $|\sqrt{n}(\bar{x} - \mu_0)/\sigma_0|$ is large.
8. Generalised likelihood ratio tests

8.3. The 'generalised likelihood ratio test'

The 'generalised likelihood ratio test'

If \( H_0 \) is true, then the test statistic is

\[
-2 \log \Lambda \sim \chi^2_p
\]

for large \( n \), as the Theorem shows. □

Notes:

- This is a ‘two-tailed’ test - i.e. reject \( H_0 \) both for high and low values of \( \bar{x} \).
- We reject \( H_0 \) if \( |\sqrt{n}(\bar{x} - \mu_0)/\sigma_0| > z_{\alpha/2} \). A symmetric \( 100(1 - \alpha)\% \) confidence interval for \( \mu \) is \( \bar{x} \pm z_{\alpha/2} \sigma_0/\sqrt{n} \). Therefore we reject \( H_0 \) iff \( \mu_0 \) is not in this confidence interval (check).
- In later lectures the close connection between confidence intervals and hypothesis tests is explored further.

Theorem 8.4

(not proved)

Suppose \( \Theta_0 \subseteq \Theta_1, |\Theta_1| - |\Theta_0| = p \). Then under regularity conditions, as \( n \to \infty \), with \( X = (X_1, \ldots, X_n) \), \( X_i \)'s iid, we have, if \( H_0 \) is true,

\[
2 \log \Lambda_X(H_0; H_1) \sim \chi^2_p.
\]

If \( H_0 \) is not true, then \( 2 \log \Lambda \) tends to be larger. We reject \( H_0 \) if \( 2 \log \Lambda > c \) where

\[
c = \chi^2_p(\alpha)
\]

for a test of approximately size \( \alpha \).

In Example 8.3, \( |\Theta_1| - |\Theta_0| = 1 \), and in this case we saw that under \( H_0 \),

\[
2 \log \Lambda \sim \chi^2_1
\]

exactly for all \( n \) in that particular case, rather than just approximately for large \( n \) as the Theorem shows.

(Often the likelihood ratio is calculated with the null hypothesis in the numerator, and so the test statistic is \( -2 \log \Lambda_X(H_1; H_0) \).)