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** $W(\theta) = \mathbb{P}(\mathbf{X} \in C \,|\, \theta) = \mathbb{P}(\text{reject } H_0 \,|\, \theta)$.
- 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 $\theta \in \Theta_1$, $1 - W(\theta) = \mathbb{P}(\text{Type II error} \,|\, \theta)$.
- 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 α test of $H_0 : \mu = \mu_0$ versus $H_1 : \mu = \mu_1$ (where $\mu_1 > \mu_0$) is given by $C = \{\mathbf{x} : \sqrt{n}(\bar{x} - \mu_0)/\sigma_0 > z_\alpha\}$.
- This critical region depends on μ_0, n, σ_0, α, on the fact that $\mu_1 > \mu_0$, but not on the particular value of μ_1.

- Hence this C defines the most powerful size α test of $H_0 : \mu = \mu_0$ against any μ_1 that is larger than μ_0.
- This test is then uniformly most powerful size α for testing $H_0 : \mu = \mu_0$ against $H_1 : \mu > \mu_0$.

Definition 8.1

A test specified by a critical region C is **uniformly most powerful** (UMP) size α test for testing $H_0 : \theta \in \Theta_0$ against $H_1 : \theta \in \Theta_1$ if

(i) $\sup_{\theta \in \Theta_0} W(\theta) = \alpha$;

(ii) for any other test C^* with size $\leq \alpha$ and with power function W^* we have $W(\theta) \geq W^*(\theta)$ for all $\theta \in \Theta_1$.

- UMP tests may not exist.
- However likelihood ratio tests are often UMP.

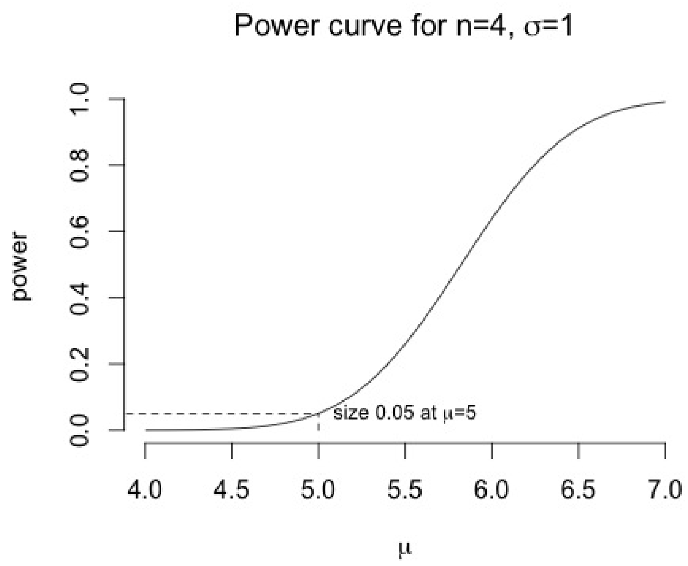
Example 8.2

Suppose $X_1, \ldots, X_n$ are iid $N(\mu_0, \sigma_0^2)$ where σ_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 α of H_0' against H_1' has $C = \{\mathbf{x} : \sqrt{n}(\bar{x} - \mu_0)/\sigma_0 > z_\alpha\}$.
- 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

$$\begin{aligned} W(\mu) &= \mathbb{P}_\mu(\text{reject } H_0) = \mathbb{P}_\mu\left(\frac{\sqrt{n}(\bar{X} - \mu_0)}{\sigma_0} > z_\alpha\right) \\ &= \mathbb{P}_\mu\left(\frac{\sqrt{n}(\bar{X} - \mu)}{\sigma_0} > z_\alpha + \frac{\sqrt{n}(\mu_0 - \mu)}{\sigma_0}\right) \\ &= 1 - \Phi\left(z_\alpha + \frac{\sqrt{n}(\mu_0 - \mu)}{\sigma_0}\right). \end{aligned}$$

```
power= 1 - pnorm( qnorm(0.95) + sqrt(n) * (mu0-x) / sigma0 )
```



- We know $W(\mu_0) = \alpha$. (just plug in)
- $W(\mu)$ is an increasing function of μ.
- So $\sup_{\mu \le \mu_0} W(\mu) = \alpha$, and (i) is satisfied.
- For (ii), observe that for any $\mu > \mu_0$, the Neyman Pearson size α 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 μ_1.)
- Let C^* and W^* belong to any other test of H_0 vs H_1 of size $\le \alpha$
- Then C^* can be regarded as a test of $H_0' : \mu = \mu_0$ vs H_1' of size $\le \alpha$, and NP-Lemma says that $W^*(\mu_1) \le W(\mu_1)$
- This holds for all $\mu_1 > \mu_0$ and so (ii) is satisfied.
- So C is UMP size α for H_0 vs H_1. $\square$

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 $\mathbf{x}$ to be
$$L_{\mathbf{x}}(H) = \sup_{\theta \in \Theta} f(\mathbf{x} \mid \theta).$$
- So far we have considered disjoint hypotheses Θ_0, Θ_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
$$\Lambda_{\mathbf{x}}(H_0; H_1) = \frac{L_{\mathbf{x}}(H_1)}{L_{\mathbf{x}}(H_0)} = \frac{\sup_{\theta \in \Theta_1} f(\mathbf{x} \mid \theta)}{\sup_{\theta \in \Theta_0} f(\mathbf{x} \mid \theta)} (\ge 1), \qquad (1)$$
with large values of $\Lambda_{\mathbf{x}}$ indicating departure from H_0.
- Notice that if $\Lambda^*_{\mathbf{x}} = \sup_{\theta \in \Theta \setminus \Theta_0} f(\mathbf{x} \mid \theta) / \sup_{\theta \in \Theta_0} f(\mathbf{x} \mid \theta)$, then $\Lambda_{\mathbf{x}} = \max\{1, \Lambda^*_{\mathbf{x}}\}$.

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 σ_0^2 known, and we wish to test $H_0 : \mu = \mu_0$ against $H_1 : \mu \neq \mu_0$ (μ_0 is a given constant).

- Here $\Theta_0 = \{\mu_0\}$ and $\Theta = \mathbb{R}$.
- For the denominator in (1) we have $\sup_{\Theta_0} f(\mathbf{x} \mid \mu) = f(\mathbf{x} \mid \mu_0)$.
- For the numerator, we have $\sup_{\Theta} f(\mathbf{x} \mid \mu) = f(\mathbf{x} \mid \hat{\mu})$, where $\hat{\mu}$ is the mle, so $\hat{\mu} = \bar{x}$ (check).
- Hence
$$\Lambda_{\mathbf{x}}(H_0; H_1) = \frac{(2\pi\sigma_0^2)^{-n/2} \exp\left(-\frac{1}{2\sigma_0^2} \sum (x_i - \bar{x})^2\right)}{(2\pi\sigma_0^2)^{-n/2} \exp\left(-\frac{1}{2\sigma_0^2} \sum (x_i - \mu_0)^2\right)},$$
and we reject H_0 if $\Lambda_{\mathbf{x}}$ is 'large.'
- We find that
$$2 \log \Lambda_{\mathbf{x}}(H_0; H_1) = \frac{1}{\sigma_0^2} \left[\sum (x_i - \mu_0)^2 - \sum (x_i - \bar{x})^2 \right] = \frac{n}{\sigma_0^2} (\bar{x} - \mu_0)^2. \text{ (check)}$$
- Thus an equivalent test is to reject H_0 if $\left| \sqrt{n}(\bar{x} - \mu_0)/\sigma_0 \right|$ is large.

- Under H_0, $Z = \sqrt{n}(\bar{X} - \mu_0)/\sigma_0 \sim N(0,1)$ so the size α generalised likelihood test rejects H_0 if $\left|\sqrt{n}(\bar{x} - \mu_0)/\sigma_0\right| > z_{\alpha/2}$.
- Since $n(\bar{X} - \mu_0)^2/\sigma_0^2 \sim \chi_1^2$ if H_0 s true, this is equivalent to rejecting H_0 if $n(\bar{X} - \mu_0)^2/\sigma_0^2 > \chi_1^2(\alpha)$ (check that $z_{\alpha/2}^2 = \chi_1^2(\alpha)$). $\square$

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 $\left|\sqrt{n}(\bar{x} - \mu_0)/\sigma_0\right| > z_{\alpha/2}$. A symmetric $100(1-\alpha)\%$ confidence interval for μ is $\bar{x} \pm z_{\alpha/2}\,\sigma_0/\sqrt{n}$. Therefore we reject H_0 iff μ_0 is not in this confidence interval (check).
- In later lectures the close connection between confidence intervals and hypothesis tests is explored further.

The 'generalised likelihood ratio test'

- The next theorem allows us to use likelihood ratio tests even when we cannot find the exact relevant null distribution.
- First consider the 'size' or 'dimension' of our hypotheses: suppose that H_0 imposes p independent restrictions on Θ, so for example, if and we have
  - $H_0 : \theta_{i_1} = a_1, \ldots, \theta_{i_p} = a_p$ ($a_1, \ldots, a_p$ given numbers),
  - $H_0 : A\theta = \mathbf{b}$ (A $p \times k$, $\mathbf{b}$ $p \times 1$ given),
  - $H_0 : \theta_i = f_i(\phi)$, $i = 1, \ldots, k$, $\phi = (\phi_1, \ldots, \phi_{k-p})$.
- Then Θ has 'k free parameters' and Θ_0 has '$k-p$ free parameters.'
- We write $|\Theta_0| = k - p$ and $|\Theta| = k$.

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 $\mathbf{X} = (X_1, \ldots, X_n)$, X_i's iid, we have, if H_0 is true,

$$2 \log \Lambda_{\mathbf{X}}(H_0; H_1) \sim \chi_p^2.$$

If H_0 is not true, then $2 \log \Lambda$ tends to be larger. We reject H_0 if $2 \log \Lambda > c$ where $c = \chi_p^2(\alpha)$ for a test of approximately size α.

In Example 8.3, $|\Theta_1| - |\Theta_0| = 1$, and in this case we saw that under H_0, $2 \log \Lambda \sim \chi_1^2$ *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 $-2 \log \Lambda_{\mathbf{X}}(H_1; H_0)$.)