Topic 18

Composite Hypotheses

Partitioning the Parameter Space

Outline

Partitioning the Parameter Space

The Power Function

Partitioning the Parameter Space

Simple hypotheses limit us to a decision between one of two possible states of nature. This limitation does not allow us, under the procedures of hypothesis testing to address the basic question:

> *Does the parameter value θ_0 increase, decrease or change at all under under a different experimental condition?*

This leads us to consider composite hypotheses. In this case, the parameter space Θ is divided into two disjoint regions, Θ_0 and Θ_1. The hypothesis test is now written

$$H_0 : \theta \in \Theta_0 \quad \text{versus} \quad H_1 : \theta \in \Theta_1.$$

Again, H_0 is called the null hypothesis and H_1 the alternative hypothesis.

Partitioning the Parameter Space

For the three alternatives to the question posed above, we have, for a one dimensional parameter space

- increase would lead to the choices $\Theta_0 = \{\theta; \theta \leq \theta_0\}$ and $\Theta_1 = \{\theta; \theta > \theta_0\}$,
- decrease would lead to the choices $\Theta_0 = \{\theta; \theta \geq \theta_0\}$ and $\Theta_1 = \{\theta; \theta < \theta_0\}$, and
- change would lead to the choices $\Theta_0 = \{\theta_0\}$ and $\Theta_1 = \{\theta; \theta \neq \theta_0\}$

for some choice of parameter value θ_0.The effect that we are meant to show, here the nature of the change, is contained in Θ_1. The first two options given above are called one-sided tests. The third is called a two-sided test.

Rejecting the null hypothesis, critical regions, and type I and type II errors have the same meaning for a composite hypotheses. Significance level and power will necessitate an extension of the ideas for simple hypotheses.

The Power Function

Power is now a function of the parameter value θ. If our test is to reject H_0 whenever the data fall in a critical region C, then the power function is defined as

$$\pi(\theta) = P_\theta\{X \in C\}.$$

that gives the probability of rejecting the null hypothesis for a given parameter value.

- For $\theta \in \Theta_0$, $\pi(\theta)$ is the probability of making a type I error, i.e., rejecting the null hypothesis when it is indeed true.
- For $\theta \in \Theta_1$, $1 - \pi(\theta)$ is the probability of making a type II error, i.e., failing to reject the null hypothesis when it is false.

The ideal power function has

$$\pi(\theta) \approx 0 \text{ for all } \theta \in \Theta_0 \text{ and } \pi(\theta) \approx 1 \text{ for all } \theta \in \Theta_1.$$

The Power Function

- The goal is to make the chance for error small.
- One strategy is to consider a method analogous to that employed in the Neyman-Pearson lemma. Thus, we must *simultaneously*,
  - fix a (significance) level α, now defined to be the largest value of $\pi(\theta)$ in the region Θ_0 defined by the null hypothesis, and.

    By focusing on the value of the parameter in Θ_0 that is most likely to result in an error, we insure that the probability of a type I error is no more that α *irrespective* of the value for $\theta \in \Theta_0$.
  - Look for a critical region that makes the power function as large as possible for values of the parameter $\theta \in \Theta_1$.

The Power Function

Example. Let $X_1, X_2, \ldots, X_n$ be independent $N(\mu, \sigma_0)$ random variables with σ_0 known and μ unknown. For the composite hypothesis for the one-sided test

$$H_0 : \mu \leq \mu_0 \quad \text{versus} \quad H_1 : \mu > \mu_0,$$

we use the test statistic from the likelihood ratio test and reject H_0 if the statistic $\bar{x}$ is too large. Thus, the critical region

$$C = \{\mathbf{x}; \bar{x} \geq k(\mu_0)\}.$$

If μ is the true mean, then the power function

$$\pi(\mu) = P_\mu\{X \in C\} = P_\mu\{\bar{X} \geq k(\mu_0)\}.$$

The value of $k(\mu_0)$ depends on the level of the test.

The Power Function

- As the actual mean μ increases, then the probability that the sample mean $\bar{X}$ exceeds a particular value $k(\mu_0)$ also increases.
- In other words, π is an increasing function.
- Thus, the maximum value of π on $\Theta_0 = \{\mu; \mu \leq \mu_0\}$ takes place for $\mu = \mu_0$.
- Consequently, to obtain level α for the hypothesis test, set

$$\alpha = \pi(\mu_0) = P_{\mu_0}\{\bar{X} \geq k(\mu_0)\}.$$

We now use this to find the value $k(\mu_0)$. When μ_0 is the value of the mean, we standardize to give a standard normal random variable

$$Z = \frac{\bar{X} - \mu_0}{\sigma_0/\sqrt{n}}.$$

Choose z_α so that $P\{Z \geq z_\alpha\} = \alpha$. Thus, $P_{\mu_0}\{Z \geq z_\alpha\} = P_{\mu_0}\{\bar{X} \geq \mu_0 + \frac{\sigma_0}{\sqrt{n}} z_\alpha\}$ and $k(\mu_0) = \mu_0 + (\sigma_0/\sqrt{n}) z_\alpha$.

The Power Function

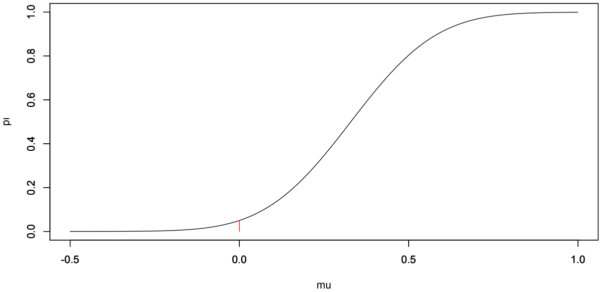
Example, If μ is the true state of nature, then

$$Z = \frac{\bar{X} - \mu}{\sigma_0/\sqrt{n}}$$

is a standard normal random variable. Use this to show that the power function

$$\pi(\mu) = 1 - \Phi\left(z_\alpha - \frac{\mu - \mu_0}{\sigma_0/\sqrt{n}}\right)$$

where Φ is the distribution function for the standard normal.



Power function for the one-sided test with alternative greater. The size of the test α is given by the height of the red segment. Notice that $\pi(\mu) < \alpha$ for all $\mu < \mu_0$ and $\pi(\mu) > \alpha$ for all $\mu > \mu_0$.

The Power Function

We have seen the expression

$$\pi(\mu) = 1 - \Phi\left(z_\alpha - \frac{\mu - \mu_0}{\sigma_0/\sqrt{n}}\right)$$

in several contexts.

- If we fix n, the number of observations and the alternative value $\mu = \mu_1 > \mu_0$ and determine the power $1 - \beta$ as a function of the significance level α, then we have the receiver operator characteristic.
- If we fix μ_1 the alternative value and the significance level α, then we can determine the power as a function of n the number of observations.
- If we fix n and the significance level α, then we can determine the power function $\pi(\mu)$, the power as a function of the alternative value μ.

Exercise. Give the appropriate expression for π for a less than alternative and use this to plot the power function for the example with a model species and its mimic. Take $\alpha = 0.05$, $\mu_0 = 10$, $\sigma_0 = 3$, and $n = 16$ observations,