Probability and Statistics

Fall 2000

LECT. 17: MOST POWERFUL TESTS, CB 8.3.2-8.3.3

Example

Let us consider some examples of applications of the Neyman-Pearson Lemma. Suppose X has an exponential distribution with arrival rate λ. We wish to test the hypothesis that $\lambda = 1$ against the alternative that $\lambda = 2$. The form of the critical region, determined by the ratio of densities, is

$$C_X = \{x | f_X(x; 2) \geq k \cdot f_X(x; 1)\}$$

$$= \{x | 2 \cdot \exp(-2x) \geq k \cdot \exp(-x)\}$$

$$= \{x | -2x \geq k' - x\}$$

$$= \{x | x \leq k''\}.$$

All that is left to determine is k''. Suppose we wish to test at the 0.05 level. Then we choose k'' to satisfy

$$0.05 = \int_0^{k''} \exp(-x) dx = 1 - \exp(-k''),$$

or

$$k'' = -\ln(0.95),$$

and the critical region is

$$C_X = [0, \ln(0.95)].$$

□

Example

Consider another example. Suppose $X_1, \ldots, X_N$ are iid normal with mean μ and unit variance. We wish to test the null hypothesis $\mu = \mu_0$ against the alternative hypothesis that $\mu = \mu_1$, for some μ_1 and μ_0 with $\mu_1 > \mu_0$. The ratio of density functions is

$$\frac{\mathcal{L}(\mu_1)}{\mathcal{L}(\mu_0)} = \frac{\exp\Big(-\sum_{i=1}^N (x_i - \mu_1)^2/2\Big)}{\exp\Big(-\sum_{i=1}^N (x_i - \mu_0)^2/2\Big)} = \frac{\exp\Big(\sum_{i=1}^N x_i\mu_1 - N\mu_1^2/2\Big)}{\exp\Big(\sum_{i=1}^N x_i\mu_0 - N\mu_0^2/2\Big)}.$$

This is larger than some cutoff point if

$$\bar{x} \geq k.$$

The critical region is therefore of the form

$$C_X = \{x_1, \ldots, x_N | \bar{x} \geq k\}.$$

Suppose we wish to test at the 0.05 level. Then

$$0.05 = Pr(\bar{x} \geq k | \mu = \mu_0).$$

Under the null the distribution of $\bar{x}$ is normal with mean μ_0 and variance $1/N$. The probability that such a random variable is larger than $\mu_0 + 1.645/\sqrt{N}$ is 0.05. Hence, the critical region is

$$C_X = \{x_1, \ldots, x_N | \bar{x} > \mu_0 + 1.645/\sqrt{N}\}.$$

□

Note that the second example the critical region does not depend on the value of the parameter under the alternative hypothesis, μ_1. Whether the alternative is $\mu_1 = \mu_0 + 1$ or $\mu_1 = \mu_0 + 4$ leads to exactly the same critical region. The test is therefore uniformly the most powerful test even for the <u>composite</u> alternative hypothesis $H_1 : \mu > \mu_0$.

Now consider testing, in the same normal model, the hypothesis

$$H_0 : \qquad \mu = \mu_0,$$

against the alternative

$$H_1 : \qquad \mu \neq \mu_0.$$

If the alternative is $\mu = \mu_1 > \mu_0$ the critical region for the most powerful test is of the form

$$C_X = \{x | x > k\}.$$

If the alternative is $\mu = \mu_1 < \mu_0$ the critical region of the most powerful test is of the form

$$C_X = \{x | x < k\}.$$

There is therefore no test that is most powerful for all values under the alternative. In other words, there is <u>no uniformly most powerful test</u>. We impose an additional restriction to get around this problem.

A test is <u>unbiased</u> if the power function $\beta(\theta_1) \geq \beta(\theta_0)$ for all $\theta_1 \in \Theta_0^c$ and all $\theta \in \Theta_0$. That is, the probability of rejecting the null hypothesis, or of an observation in the critical region, is at least as large for values of the parameters consistent with the alternative ($\theta \in \Theta_0^c$) as for values of the parameters consistent with the null hypothesis ($\theta \in \Theta_0$).

Let us consider this approach in detail for the case with a normal distribution with unknown mean and known variance. Let $X_1, \ldots, X_N$ be independent and normally distributed with unknown mean μ and known variance σ^2. We are interested in testing the null hypothesis

$$H_0 : \; \mu = \mu_0,$$

against the alternative

$$H_1 : \; \mu = \mu_1 \neq \mu_0.$$

Let us consider the ratio of density functions to determine the critical region:

$$\frac{f(x_1, \ldots, x_N | \mu_1)}{f(x_1, \ldots, x_N | \mu_0)} = \frac{(2\pi\sigma^2)^{N/2} \exp\left(-\frac{1}{2\sigma^2}\left(\sum_{i=1}^N x_i^2 - 2\mu_1 \sum_{i=1}^N x_i + N\mu_1^2\right)\right)}{(2\pi\sigma^2)^{N/2} \exp\left(-\frac{1}{2\sigma^2}\left(\sum_{i=1}^N x_i^2 - 2\mu_0 \sum_{i=1}^N x_i + N\mu_0^2\right)\right)}$$

$$= \exp\Big(\frac{1}{2\sigma^2} \cdot (\mu_1 - \mu_0) \sum x_i\Big) \cdot \exp(-(\mu_1^2 - \mu_0^2)N/(2\sigma^2)).$$

Hence if we are looking for a uniformly most powerful test against the alternative hypothesis $H_1: \ \mu = \mu_1 > \mu_0$, the critical region ought to be of the form

$$C_X = \Big\{x_1, \ldots, x_N \Big| \sum x_i \geq k\Big\}.$$

If we were to test against the alternative hypothesis $H_1: \ \mu = \mu_1 < \mu_0$, the critical region ought to be of the form

$$C_X = \Big\{x_1, \ldots, x_N \Big| \sum x_i \leq k\Big\}.$$

It therefore appears sensible to base a test on the value of $\bar{x}$, the sample average, which is a sufficient statistic for μ. It seems obvious that the critical region should be of the form

$$C_X\Big\{x_1, \ldots, x_N \Big| \bar{x} < a, \bar{x} > b\Big\}.$$

Unbiasedness of the test implies that $\mu_0 - a = b - \mu_0$. Hence the critical region is

$$C_X = \Big\{x_1, \ldots, x_N \Big| \bar{x} < \mu_0 - c, \bar{x} > \mu_0 + c\Big\},$$

with the value of c determined by the size of the test. Under the null hypothesis the distribution of $\bar{x}$ is normal with mean μ_0 and variance σ^2/N. Hence, if we wish to test at the 10% level, recalling that for a standard normal random variable Z

$$Pr(-1.645 < Z < 1.645) = 0.90,$$

the critical region is

$$C_X = \Big\{x_1, \ldots, x_N \Big| \bar{x} < \mu_0 - 1.645 \cdot \sigma/\sqrt{N}, \bar{x} > \mu_0 + 1.645 \cdot \sigma/\sqrt{N}\Big\}.$$

This is the <u>uniformly most powerful unbiased test</u>.

If we wish to test at the 5% level, the critical region is

$$C_X = \left\{ x_1, \ldots, x_N \middle| \bar{x} < \mu_0 - 1.96 \cdot \sigma/\sqrt{N}, \bar{x} > \mu_0 + 1.96 \cdot \sigma/\sqrt{N} \right\}.$$

Equivalently we can use the critical region

$$C_X = \left\{ x_1, \ldots, x_N \middle| N \cdot (\bar{x} - \mu_0)^2/\sigma^2 > 3.84 \right\},$$

which uses the Chi–squared distribution for the square of a standard normal random variable. In fact a common way of doing the test is to calculate the test statistic, here $N \cdot (\bar{x} - \mu_0)^2/\sigma^2$ which under the null hypothesis has a known distribution, in this case a chi–square one distribution. We reject the null hypothesis if the test statistic exceeds the critical value, in this case 3.84 at the 5% level or 2.728 at the 10% level.