Topic 19

Extensions on the Likelihood Ratio

Two-Sided Tests

Outline

Overview

- The likelihood ratio test is a popular choice for composite hypothesis when Θ_0 is a subspace Θ the parameter space.
- The rationale for this approach is that the null hypothesis is unlikely to be true if the maximum likelihood on Θ_0 is sufficiently smaller that the likelihood maximized over Θ. Let
  - $\hat{\theta}_0$ be the parameter value that maximizes the likelihood for $\theta \in \Theta_0$ and
  - $\hat{\theta}$ be the parameter value that maximizes the likelihood for $\theta \in \Theta$.
- The likelihood ratio

$$\Lambda(\mathbf{x}) = \frac{L(\hat{\theta}_0|\mathbf{x})}{L(\hat{\theta}|\mathbf{x})}.$$

Overview

We have two optimization problems - maximize $L(\theta|\mathbf{x})$ on the parameter space Θ and on the null hypothesis space Θ_0.

The critical region for an α-level likelihood ratio test is

$$\{\Lambda(\mathbf{x}) \leq \lambda_\alpha\}.$$

As with any α level test, λ_α is chosen so that

$$P_\theta\{\Lambda(X) \leq \lambda_\alpha\} \leq \alpha \text{ for all } \theta \in \Theta_0.$$

NB. This ratio is the reciprocal from the version given by the Neyman-Pearson lemma. Thus, the critical region consists of those values that are *below* a critical value.

Normal Observations

Consider the two-sided hypothesis

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

Here the data are n independent $N(\mu, \sigma_0)$ random variables with known variance σ_0^2. The parameter space Θ is one dimensional giving the value μ for the mean. As we have seen before $\hat{\mu} = \bar{x}$. Θ_0 is the single point $\{\mu_0\}$ and so $\hat{\mu}_0 = \mu_0$.

$$L(\hat{\mu}_0|\mathbf{x}) = \left(\frac{1}{\sqrt{2\pi\sigma_0^2}}\right)^n \exp -\frac{1}{2\sigma_0^2}\sum_{i=1}^n (x_i - \mu_0)^2, \quad L(\hat{\mu}|\mathbf{x}) = \left(\frac{1}{\sqrt{2\pi\sigma_0^2}}\right)^n \exp -\frac{1}{2\sigma_0^2}\sum_{i=1}^n (x_i - \bar{x})^2$$

and

$$\Lambda(\mathbf{x}) = \exp -\frac{1}{2\sigma_0^2}\left(\sum_{i=1}^n ((x_i - \mu_0)^2 - (x_i - \bar{x})^2)\right) = \exp -\frac{n}{2\sigma_0^2}(\bar{x} - \mu_0)^2.$$

Notice that

$$-2\ln\Lambda(\mathbf{x}) = \frac{n}{\sigma_0^2}(\bar{x} - \mu_0)^2 = \left(\frac{\bar{x} - \mu_0}{\sigma_0/\sqrt{n}}\right)^2.$$

Normal Observations

Then, critical region,

$$\{\Lambda(\mathbf{x}) \le \lambda_\alpha\} = \left\{ \left(\frac{\bar{x} - \mu_0}{\sigma_0/\sqrt{n}} \right)^2 \ge -2 \ln \lambda_\alpha \right\}.$$

Under the null hypothesis, $(\bar{X} - \mu_0)/(\sigma_0/\sqrt{n})$ is a standard normal random variable, and thus $-2 \ln \Lambda(X)$ is the square of a single standard normal. This is the defining property of a χ-square random variable with 1 degree of freedom.

Naturally we can use both

$$\left(\frac{\bar{x} - \mu_0}{\sigma_0/\sqrt{n}} \right)^2 \quad \text{and} \quad \left| \frac{\bar{x} - \mu_0}{\sigma_0/\sqrt{n}} \right|.$$

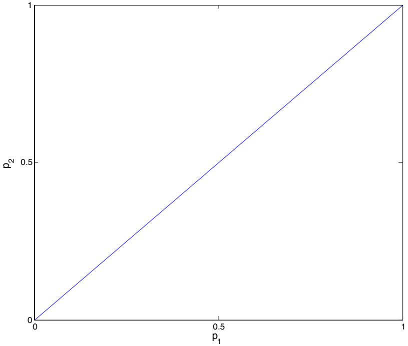
as a test statistic. We have seen the second choice in the example of a possible invasion of a model butterfly by a mimic.

Two-Sample Proportions

For the two-sided two-sample α-level likelihood ratio test for population proportions p_1 and p_2, based on the hypothesis

$$H_0 : p_1 = p_2 \quad \text{versus} \quad H_1 : p_1 \neq p_2,$$

- we maximize the likelihood over the subspace $\Theta_0 = \{(p_1, p_2); p_1 = p_2\}$ (the blue line) and
- over the entire parameter space, $\Theta = [0, 1] \times [0, 1]$, shown as the square, and
- then take the ratio, simplify and make appropriate approximations.



The data are observations on n_1 Bernoulli trials, $x_{1,1}, x_{1,2}, \ldots, x_{1,n_1}$ from the first population and, independently, n_2 Bernoulli trials, $x_{2,1}, x_{2,2}, \ldots, x_{2,n_2}$ from the second.

Two-Sample Proportions

The likelihood ratio test is approximately equivalent to the critical region

$$|z| \geq z_{\alpha/2}$$

where

$$z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}_0(1-\hat{p}_0)\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}}$$

with $\hat{p}_i$, the sample proportion of successes from the observations from population i and $\hat{p}_0$, the pooled proportion

$$\hat{p}_0 = \frac{1}{n_1+n_2}\left((x_{1,1}+\cdots+x_{1,n_1})+(x_{2,1}+\cdots+x_{2,n_2})\right) = \frac{n_1\hat{p}_1+n_2\hat{p}_2}{n_1+n_2}.$$

Two-Sample Proportions

The subsequent winter had 167 out of 250 hives surviving. To test if the two survival probabilities are significantly different:

```
> prop.test(c(250,167),c(332,250))

2-sample test for equality of proportions with continuity correction

data:  c(250, 167) out of c(332, 250)
X-squared = 4.664, df = 1, p-value = 0.0308
alternative hypothesis: two.sided
95 percent confidence interval:
 0.006942351 0.163081746
sample estimates:
  prop 1   prop 2
0.753012 0.668000
```

Two-Sample Proportions

Exercise.

1. Give the 95% confidence interval for the difference in proportions. Does it contain the value 0?
2. Modify the R command to find a 98% confidence interval for the difference in proportions. Does it contain the value 0?
3. Explain your answers to the first two parts - why does one contain 0 and the other not? Could you have guessed this in advance by looking at the p-value.
4. Assume that the second winter was more severe. State a hypothesis for appropriate for this situation and modify the R commands and give the p-value. Could you have given this p-value based on the information in the output on the previous slide?

Two-Sample Proportions

Power analyses can be executed in R using the `power.prop.test` command. If we want to be able to detect a difference between two proportions $p_1 = 0.7$ and $p_2 = 0.6$ in a one-sided test with a significance level of $\alpha = 0.05$ and power $1 - \beta = 0.8$.

```
> power.prop.test(p1=0.70,p2=0.6,sig.level=0.05,power=0.8,
  alternative = c("one.sided"))
     Two-sample comparison of proportions power calculation
              n = 280.2581
             p1 = 0.7
             p2 = 0.6
      sig.level = 0.05
          power = 0.8
    alternative = one.sided
 NOTE: n is number in *each* group
```

We will need a sample of $n = 281$ from each group.

Two-Sample Proportions

If we vary p_2 and determine the power.

```
> power.prop.test(n=250,p1=0.70,p2=c(0.6,0.65),sig.level=0.05,
   alternative = c("one.sided"))
              p2 = 0.60, 0.65
           power = 0.7589896, 0.3256442
```

Now, let's vary sample size.

```
> power.prop.test(n=c(250,350,450,550),p1=0.70,p2=0.60,sig.level=0.05,
   alternative = c("one.sided"))
               n = 250, 350, 450, 550
           power = 0.7589896, 0.8717915, 0.9342626, 0.9672670
```

Exercise. Determine the reduction in power when the significance level $\alpha = 0.02$ for the sample sizes above. Why is the power reduced?