STAT 135, Concepts of Statistics

Helmut Pitters

Hypothesis testing 4

Department of Statistics
University of California, Berkeley

March 16, 2017

Hypothesis testing.

Likelihood ratio test for multinomial distribution. Last example: Pearson's chi squared test for goodness of fit of a multinomial distribution.

Now: LRT for multinomial distribution.

Let

$$M := \{(p_1, \ldots, p_m) \colon p_i \geq 0, \sum_i p_i = 1\}$$

denote the set of all possible probabilities for a multinomial distribution with m cells. Consider hypotheses[1]

$$H_0 \colon p \in M_0 := \{(p_1(\theta), \ldots, p_m(\theta)) \in M \colon \theta \in \Theta_0\},$$

vs.

$$H_A \colon (p_1, \ldots, p_m) \in M$$

such that $(p_1, \ldots, p_m) \neq (p_1(\theta), \ldots, p_m(\theta))$ for all $\theta \in \Theta_0$.

[1] Notice that the cell probabilities p_i depend on some parameter θ.

Hypothesis testing.

Likelihood ratio test for multinomial distribution. For random sample of size n LR is found to be

$$\Lambda = \frac{\sup_{\theta \in \Theta_0} \operatorname{lik}(p(\theta))}{\sup_{p \in M} \operatorname{lik}(p)} = \prod_{i=1}^{m} \left(\frac{p_i(\hat{\theta})}{\hat{p}_i}\right)^{x_i}, \quad \operatorname{lik}(p) = \binom{n}{x_1, \ldots, x_m} \prod_i p_i^{x_i},$$

where $\hat{p}_i = x_i/n$ and $\hat{\theta}$ is MLE for θ under restriction $\theta \in \Theta_0$.[2]
Therefore obtain

$$-2 \log \Lambda = -2n \sum_{i=1}^{m} \hat{p}_i \log\left(\frac{p_i(\hat{\theta})}{\hat{p}_i}\right) = 2 \sum_{i=1}^{m} O_i \log \frac{O_i}{E_i}$$

where

$$O_i = n\hat{p}_i = \text{\# observed counts in category } i,$$

$$E_i = np_i(\hat{\theta}) = \text{\# expected counts in category } i.$$

One can show that as sample size n grows without bounds

$$-2 \log \Lambda \qquad \text{and Pearson's chi squared statistics}$$

are asymptotically equivalent under H_0.

[2]LRT suggests a connection between testing a hypothesis and estimating a

Hypothesis testing.

Example: Mendel's pea experiment. Gregor Mendel supported his theory of inheritance with his famous pea experiment, in which he crossed round yellow seeds with wrinkled green seeds. According to his theory, the relative frequencies of peas of different types are

round yellow	round green	wrinkled yellow	wrinkled green
$\frac{9}{16}$	$\frac{3}{16}$	$\frac{3}{16}$	$\frac{1}{16}$

Table: Relative frequencies of types of beans.

In $n = 556$ trials Mendel observed

$$O = (315, 101, 108, 32) \qquad \text{peas of each type.}$$

Under

$$H_0\colon p = (\frac{9}{16}, \frac{3}{16}, \frac{3}{16}, \frac{1}{16}),$$

the hypothesis claimed by Mendel's theory, we'd expect

$$E = 556 \times (\frac{9}{16}, \frac{3}{16}, \frac{3}{16}, \frac{1}{16}) = (312.75, 104.25, 104.25, 34.75)$$

peas of each type.

Hypothesis testing.

Example: Mendel's pea experiment. Evaluating Pearson's chi squared statistic yields

$$X^2 = \sum_{i \in \{\text{type of peas}\}} \frac{(O_i - E_i)^2}{E_i} = \frac{(315 - 312.75)^2}{312.75} + \frac{(101 - 104.25)^2}{104.25} + \frac{(108 - 104.25)^2}{104.25} + \frac{(32 - 34.75)^2}{34.75} \approx 0.604.$$

Since large values of X^2 contradict H_0, the p-value is

$$\mathbb{P}\{Y \geq 0.604\} = 0.896$$

where $Y \sim \chi^2_3$ has chi squared distribution with 3 df ($= 4$ categories of peas -1).
That is in about 90% of cases we see for X^2 a value of 0.604 or larger. This result is in very good agreement with H_0 (and hence Mendel's theory) — too good to be true?

Hypothesis testing.

Example: Mendel's pea experiment. Notice that with Λ denoting the likelihood ratio, we obtain

$$-2 \log \Lambda = 2 \sum_{i \in \{\text{type of peas}\}} O_i \log \frac{O_i}{E_i} \approx 0.618$$

which is very close to the value of Pearson's chi squared statistic X^2.

Hypothesis testing.

"Recipe" for testing hypotheses.

1. Specify null hypothesis H_0 and alternative hypothesis H_A.
2. Specify test statistic T that discriminates between H_0 and H_A.
3. Specify "extreme" values of T under H_0 in direction of H_A, suggesting that H_A better explains data than H_0.
   E.g. reject H_0 for
   - large values of T, or
   - small values of T, or
   - large values of $|T - t|$ for some t,
   - etc.
4. If significance level α given, can work out rejection region if null distribution (of T) is known or can be approximated.
5. Evaluate T from data and
   - decide whether to reject H_0, or
   - work out p-value.

Hypothesis testing.

Duality of confidence intervals and hypothesis tests. Next theorem shows that procedures of testing a hypothesis and estimating a parameter are in fact intimately related. To see this, have to slightly generalize the notion of confidence interval to a confidence set.
Recall that $X = (X_1, \ldots, X_n) \sim \mathbb{P}_\theta$ denotes vector from which the sample is drawn whose joint distribution depends on some (unknown) parameter $\theta \in \Theta$.[3] Let's assume Θ is a subset of $\mathbb{R}$.

Definition (Confidence set)

A (random) subset $C(X)$ of $\mathbb{R}$ is called an α *confidence set for* θ if

$$\mathbb{P}\left\{\theta \in C(X)\right\} = \alpha.$$

[3] Think of specific examples as the ones we saw in the introduction of the terminology in the Neyman-Pearson paradigm.

Duality of confidence intervals and hypothesis tests.

Theorem

(I) Suppose that for each $\theta_0 \in \Theta$ there exists a level α test of

$$H_0\colon \theta = \theta_0$$

with acceptance region $A(\theta_0)$, i.e.

$$\text{test rejects } H_0 \text{ if } X \notin A(\theta_0).$$

Then

$$C(X) := \{\theta\colon X \in A(\theta)\}$$

is a $1-\alpha$ confidence set for θ.

Proof. By definition of $C(X)$

$$\mathbb{P}\left\{\theta_0 \in C(X) | \theta = \theta_0\right\} = \mathbb{P}\left\{X \in A(\theta_0) | \theta = \theta_0\right\} = 1 - \alpha.$$

Duality of confidence intervals and hypothesis tests.

Theorem

(II) In the other direction, start with $1-\alpha$ confidence set $C(X)$ for θ. A level α test for

$$H_0\colon \theta = \theta_0$$

is given by decision rule

$$\textit{reject } H_0 \textit{ if } X \notin A(\theta_0) \coloneqq \{x\colon \theta_0 \in C(x)\}$$

Proof. By definition of $A(\theta_0)$

$$\mathbb{P}\{X \notin A(\theta_0) | \theta = \theta_0\} = \mathbb{P}\{\theta_0 \notin C(X) | \theta = \theta_0\} = \alpha.$$