STAT 231

November 14, 2016.

Roadmap

- 5 min recap
- Hypothesis testing using the likelihood function
- Applications
  - Binomial
  - Exponential
- Summary
- Clicker Questions

$$H_0 : \theta = \theta_0$$

$\{y_1, \ldots, y_n\}$.

- Construct the test statistic D
- Compute the observed value d.
- Compute the p-value :

$$P(D \geq d; H_0 \text{ is true})$$

$D \geq 0$

$D = 0$ best news for H_0

$D >> 0$ evidence against H_0.

We know the distn. of D.

It is not always easy to find D.

- There is an alternative method that we can use for any distribution, if n is large.

That involves the likelihood function

Example :

$n = 200$ $\qquad$ $y = 110$

$y = \#$ of successes.

Given this data set, can we check whether my friend has special powers

θ = probability of success.

$$H_0 : \theta = 0.5$$

Method 1.

$$D = \left| \frac{\tilde{\theta} - \theta_0}{\sqrt{\frac{\theta_0(1-\theta_0)}{n}}} \right| \qquad \theta_0 = 0.5$$

$$d = \left| \frac{\hat{\theta} - \theta_0}{\sqrt{\frac{\theta_0(1-\theta_0)}{n}}} \right|$$

$H_0 : \theta = 0.5$

Calculate the p-value

$$\Delta(\theta) = -2 \log \frac{L(\theta)}{L(\tilde{\theta})}$$

$\tilde{\theta}$ = MLE.

We can use Λ as our test statistic.

$$\begin{cases} \cdot \Lambda \geqslant 0 \,;\ \Lambda = 0 \text{ best evidence.} \\ \cdot \Lambda >> 0 \implies \text{evidence against } H_0 \\ \therefore \Lambda \sim \chi^2_1 \text{ for large } n. \end{cases}$$

<u>Step 1</u>: Compute $\lambda(\theta_0) = -2 \log \dfrac{L(\theta_0)}{L(\hat{\theta})}$

<u>Step 2</u>: Compute the p-value

$$P(\Lambda \geqslant \lambda) = \underbrace{P(Z^2 \geqslant \lambda)}_{\text{because } \Lambda \text{ follows } \chi^2(1)}$$

Example: $n = 200 \quad y = 110$. (BINOMIAL)

$H_0: \theta = 0.5$

$$L(\theta) = {}^nC_y \, \theta^y (1-\theta)^{n-y}$$

$n = 200$
$y = 110$

$$\hat{\theta} = y/n = 110/200 = 0.55$$

Step 1: Compute $\lambda(\theta_0)$ $\quad \theta_0 = 0.5$

$$\lambda(\theta) = -2 \log \frac{L(\theta)}{L(\hat{\theta})}$$

$$\lambda(0.5) = -2 \log \frac{L(0.5)}{L(0.55)} = 2.003$$

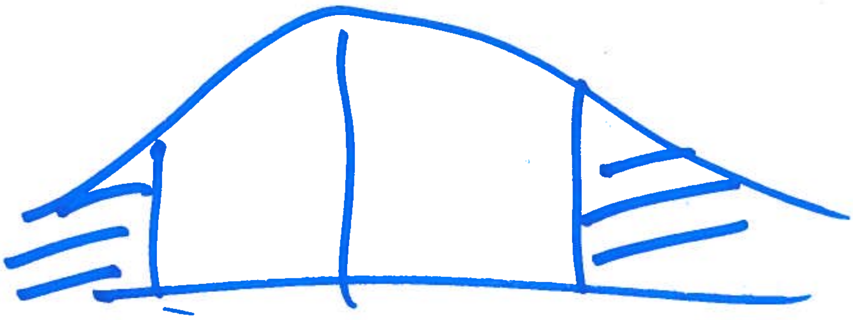
Step 2: Calculate the p-value.

$$P(\Lambda \geqslant \lambda)$$

$$= P(W \geqslant 2.003) \qquad W \sim \chi^2_1.$$

$$= P(Z^2 \geqslant 2.003)$$

$$= P\left(|Z| \geqslant \sqrt{2.003}\right) = 0. \approx 16\%.$$



Given the p-value, there is no evidence against the null hypothesis

Example: $Y_1, \ldots Y_n \sim Exp(\theta)$

A random sample is drawn from the data set $\{y_1, \ldots y_n\}$

$n = 50$; $\sum_{i=1}^{50} y_i = 93840$

$\bar{y} = 1876.8$.

Test whether

$$\left.\begin{array}{ll} & \theta = 2000 \\ \text{vs} & \theta \neq 2000 \end{array}\right\}$$

Method 1: Covered in class where

$D \sim \chi^2_{2n}$ (test-statistic).

$\mathscr{D}$ $$f(y) = \frac{1}{\theta} e^{-y/\theta}$$

Likelihood function:

$$= \frac{1}{\theta^n} e^{-\frac{1}{\theta}\sum y_i}$$

$$\uparrow$$

$$\frac{1}{\theta} e^{-y_1/\theta} \cdot \frac{1}{\theta} e^{-y_2/\theta} \cdots \frac{1}{\theta} e^{-y_n/\theta}$$

$$\boxed{L(\theta) = \theta^{-n} e^{-n\bar{y}/\theta}}$$

Step 1: Calculate $\lambda(\theta_0)$

$\theta_0 = 2000$

$\hat{\theta} = 1876.8$

$$\lambda(\theta_0) = -2 \log \frac{L(\theta_0)}{L(\hat{\theta})} = 0.1979$$

Step 2: Compute the p-value:

$$P(\Delta \geq \lambda(\theta_0))$$

$$= P(W \geq 0.1979)$$

$$= P(Z^2 \geq 0.1979)$$

$$= P(|Z| \geq \sqrt{0.1979})$$

$$\approx 70\%.$$

There is no evidence against the null hypothesis

We can use this method as long as

- n is large
- $\hat{\theta}$ = MLE can be calculated from $L(\theta)$.

$n > 30 \Rightarrow$ LARGE

- How to find the range of the p-value ~~fo~~ from the T- and the χ^2 table.

- A lot of tests are not two-sided.. but ~~xxx~~ one-sided.

300 cups of 30 prizes

$H_0 : \theta = 1/6$

$H_1 : \theta < \frac{1}{6}$

$D = \max\{50 - y, 0\}$

$D \geq 0$ ✓

$D = 0$, ✓

$D >> 0 \Rightarrow$ evidence against.

Clicker Questions:

Let $Y_1, \ldots, Y_{20}$ be $G(\mu, \sigma)$ independent.

$\bar{Y} = \frac{1}{n} \sum Y_i$

$S^2 = \frac{1}{\underset{n-1}{19}} \underline{\underline{\sum (Y_i - \bar{Y})^2}}$.

$W \sim \chi^2_{25}$.

All r.v.s are independent.

Q 1

- What is the distribution of $\frac{\bar{Y}-\mu}{S/\sqrt{20}}$.

$$\frac{\bar{Y}-\mu}{S/\sqrt{n}} \sim T_{n-1}$$

a. χ^2_{19}

b. $T_{19} \longrightarrow 79\%$

c. Z

d. Can't say

Q2

What distribution does $\dfrac{\sum_{i=1}^{20} (Y_i - \bar{Y})^2}{\sigma^2}$ follow?

$= \dfrac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1}$

(a) T_{19}

(b) χ^2_{19} ✓

(c) Z

(d) CAN'T SAY.