Sign Test

ST551 Lecture 13

Charlotte Wickham 2017-10-18

Sign test

Data Setting

Population: $Y \sim \text{something with c.d.f } F_Y(y) = P(Y \leq y)$

Parameter: $M = F_Y^{-1}(0.5)$, the population median for the value M s.t $P(Y \le M) = 0.5$

Sample: n i.i.d from population: Y_1, \ldots, Y_n

Null hypothesis: $H_0: M = M_0$

Your Turn:

Consider the hypothesis $H_0: M = M_0$

Imagine transforming the Y_i , i = 1, ..., n to

$$X_i = \begin{cases} 1, & Y_i \leq M_0 \\ 0, & Y_i > M_0 \end{cases}$$

If the null hypothesis is true $M=M_0$, what is $P(X_i=1)$? = 0.5

(You can assume Y is a continuous distribution) $P(Y_i \leq M_o)$

4

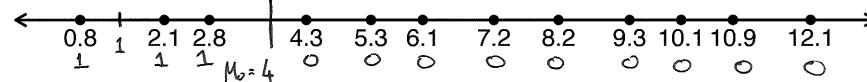
P(Vi & population)= 0.5

Sign test

To test $H_0: M=M_0$, perform a Binomial test on $X_i=\mathbf{1}\{Y_i\leq M_0\}$ with $H_0: p=0.5$.

Example

Consider a sample, n = 12, with the sample values:



Consider testing H_0 : M=4 versus a two-sided alternative

$$H_A: M \neq 4$$
 (at the $\alpha = 0.05$ level).

$$\hat{p}_{M_0} = 140/201 = \frac{1}{n} \hat{z} \times i = \frac{3}{12} = \frac{1}{4}$$

$$Z(p_0 = 0.5) = \frac{\hat{p}_{M_0} - p_0}{\sqrt{p_0(1 - p_0)/n}} = \frac{\frac{1}{4} - \frac{1}{2}}{\sqrt{\frac{1}{2}(\frac{1}{2})/12}} = -1.73$$

We <u>fail to reject</u> the null hypothesis.

Approximate Rimonial

Your turn: 95% Confidence Interval

We can *invert the test* by considering all M_0 for which we would fail to reject the null hypothesis $H_0: M = M_0$.

Would you reject for the value on your slip of paper?

Why do we only need to consider the actual sample values?

Your turn: 95% Confidence Interval

##	m_0	PM.	Reject?
## 1	0.8	/12	Reject Ho Reject Ho
## 2	2.1	3/12	Reject Ho
## 3	2.8	3/12	Fail to Reject Ho
## 4	4.3	4/12	Fail
## 5	5.3	5/12	Fail
## 6	6.1	6/12	Fail
## 7	7.2	7/12	Fail
## 8	8.2	8/12	their to.
## 9	9.3	9/12	Fail
## 10	10.1	19/12	Reject <
## 11	10.9	1/12	Fail Z=1.52? Reject Should be 5
## 12	12.1	12/12	Reject should be

95% confidence interval for M is (2.8, 0.1)

Confidence interval in general

Solve for M_0 that satisfy (i.e. not in rejection region)

$$\left| \frac{X/n - 0.5}{0.5\sqrt{n}} \right| < z_{1-\alpha/2}, \quad \text{where } X = \text{number of observations } \leq M_0$$

$$-z_{1-\alpha/2} < \frac{X/n - 0.5}{0.5/\sqrt{n}} < z_{1-\alpha/2}$$

$$n(0.5 - z_{1-\alpha/2} \frac{0.5}{\sqrt{n}}) < X < n(0.5 + z_{1-\alpha/2} \frac{0.5}{\sqrt{n}})$$

$$\frac{1}{2}(n - z_{1-\alpha/2} \sqrt{n}) < X < \frac{1}{2}(n + z_{1-\alpha/2} \sqrt{n})$$

Confidence interval in general

So, M_0 is in interval if the number of observations smaller than M_0 is between:

$$\frac{1}{2}(n-z_{1-\alpha/2}\sqrt{n})$$
 and $\frac{1}{2}(n+z_{1-\alpha/2}\sqrt{n})$

The smallest value that satisifies this is the

$$\left(\frac{1}{2}(n-z_{1-\alpha/2}\sqrt{n})\right)^{\text{th}}$$
 smallest observation

The largest value that satisifies this is the

$$\left(\frac{1}{2}(n+z_{1-\alpha/2}\sqrt{n})+1\right)^{\mathsf{th}}$$
 smallest observation

Confidence interval for median

Approximate (based on approximate Binomial test) confidence interval for the median:

$$\left(\left(\frac{n-z_{1-\alpha/2}\sqrt{n}}{2}\right)^{\text{th}}\text{ smallest observation,}\right.$$

$$\left(\frac{n+z_{1-\alpha/2}\sqrt{n}}{2}+1\right)^{\text{th}}\text{ smallest observation}\right)$$

May need to round (.)th to nearest integers

Example, continued

$$n = 12, \alpha = 0.05 \implies$$

$$\left(\left(\frac{12-1.96\sqrt{12}}{2} \right)^{th} \text{ smallest observation,} \right. \\ \left. \left(\frac{12+1.96\sqrt{12}}{2} + 1 \right)^{th} \text{ smallest observation} \right) \\ \left((2.61)^{th} \text{ smallest observation, } (10.40)^{th} \text{ smallest observation} \right) \\ \left(3^{rd} \text{ smallest observation, } 10^{th} \text{ smallest observation} \right) \\ \left(2.8, 10.1 \right)$$

Sign test for discrete distributions/data

- 1. Remove all values exactly equal to M_0
- 2. Proceed with test as usual (with a reduced sample size n)

Sign test: exactness

Finite sample exact? No

- Discrete nature of data means we can't achieve a lot of signficance levels
- Normal approximation is only an approximation...

Assymptotically exact? Yes

Sign test: consistency

The sign test test is consistent. Comes from Binomial test being consistent (which comes from Z-test being consistent).

Next time...

Signed Rank test