

This tutorial explores hypothesis testing, the proportion test and the sign test.

### Proportion Test

Context	$n$ independent Binary trials with unknown success probability $\rho$
Hypothesis	$H_0 : \rho = \rho_0$
Test Statistic	$X = \# \text{ successes} \stackrel{H_0}{\sim} \text{Bin}(n, \rho_0)$

### Sign Test

Context	$n$ observations $\{x_i\}$ from a continuous distribution with unknown mean $\mu$ We want to test $H_0 : \mu = \mu_0$
Working	Work out signs $\{x_i - \mu_0\}$ (eliminating any zeroes)
Hypothesis	$H_0 : \rho_+ = 0.5$
Test Statistic	$X = \# + \text{signs} \stackrel{H_0}{\sim} \text{Bin}(n, 0.5)$

## 1. Hypothesis Testing

In your own words, explain the following concepts:

- (a) The purpose of hypothesis testing
- (b) The importance of assumptions
- (c) Test statistic
- (d) P-value

## 2. Proportion test

The proportion of families buying a certain brand of orange juice in a certain city is believed to be 0.6. A consumer group claims that this particular brand is now less popular than it was before. A random sample of 50 families from this city shows that only 22 bought that brand of juice.

Let  $X$  be the number of families in the sample buying the brand and let  $\rho$  be the probability that a randomly picked family prefers the brand.

- (a) Hypothesis: Explain why a one-sided test of  $H_0 : \rho = 0.6$  vs  $H_1 : \rho < 0.6$  is suitable to test the claim of this consumer group.
- (b) Assumptions: To model  $X$  as a Binomial random variable, what further assumptions are needed?
- (c) Test statistic: What is the test statistic and its distribution under  $H_0$ ?
- (d) P-value: Calculate the p-value using R, and by hand using a normal approximation. What is the relative error of this approximation?

```
> pbinom(22,50,0.6)
```

- (e) Conclusion: Draw your conclusion based on the p-value.

### 3. Proportion Test

A certain family has 7 children and they are all girls. Perform a 2 sided test of the hypothesis that in that family each child is either a boy or a girl independently with equal probability. Write all your steps. Calculate the p-value using both Binomial tables and R.

### 4. Sign Test

The following data are 8 measurements of moisture retention (%) using a new scaling system. This system is expected to be better (ie have greater retention) than the system previously in use, for which the mean retention was 96%.

97.5 96.2 97.3 96.0 99.8 93.0 94.2 95.5

- (a) Preparation: By comparing the data to 96, write down the signs.
- (b) Hypothesis: Explain why  $H_0 : \rho = 0.5$  vs  $H_1 : \rho > 0.5$  is the appropriate hypothesis.
- (c) Assumptions: To perform a sign test, what assumptions are needed?
- (d) Test statistic: What is the test statistic and its distribution under  $H_0$ ?
- (e) P-value: Calculate the p-value using Binomial Tables.
- (f) Conclusion: Draw your conclusion based on the p-value.

### 5. Sign Test

14 students taste-tested two different brands of drink (brand *A* and brand *B*), with the brands being hidden from them. The object of the exercise was to see if students preferred one brand over the other, but there was no indication of which this might be before the test. Overall, 8 subjects preferred brand *A*, 4 preferred brand *B* and 2 had no preference either way. Use a sign test to test whether there is a difference between the 2 brands. Write all your steps. Calculate the p-value using Binomial tables.

### 6. Sign Test for paired data

A new measuring technique is being considered to replace the standard technique. When 10 samples are measured by both techniques, the measurements are

Sample:	1	2	3	4	5	6	7	8	9	10
New	2.5	2.2	2.6	2.6	1.9	3.3	3.3	2.8	3.0	2.9
Standard	2.1	2.4	2.1	1.9	2.0	2.8	2.7	2.8	2.8	3.0

Test the hypothesis that there is no long-run systematic difference between the two techniques.

Hint: First, calculate Difference = New-Standard. Second, write down the signs of the Differences. Third, perform the sign test.