Week 10: t distribution and t-tests

Stat 201: Statistics I

April 14, 2019

Tests for proportions

Proportion tests are conducted with the `prop.test` function.

```
prop.test(x, n, p, alternative, correct=FALSE)
```

- `x` is the number of successes. Can be a single number for one sample tests, or a vector for two sample tests.
- `n` is the number of trials. Must be the same length as the `x` parameter.
- `p` is the null proportion to test against.
- `alternative` is the form of the alternative hypothesis ("two.sided", "less", "greater").
- `correct` is whether to apply a continuity correction. Defaults to `TRUE`, but should be set to `FALSE` to match results from the formulas.

One sample proportion test

- A sample has 62 successes out of 212 trials. Test whether the population proportion is less than 40%.

```r
prop.test(62, 212, 0.4, alternative="less", correct=FALSE)
```

```
##
##  1-sample proportions test without continuity correction
##
## data:  62 out of 212, null probability 0.4
## X-squared = 10.217, df = 1, p-value = 0.0006958
## alternative hypothesis: true p is less than 0.4
## 95 percent confidence interval:
##  0.0000000 0.3461987
## sample estimates:
##         p
## 0.2924528
```

Two sample proportion test

- A sample has 62 successes out of 212 trials and another has 59 successes out of 173 trials. Test whether the population proportions are the same.

```r
prop.test(c(62, 59), c(212, 173), correct=FALSE)
```

```
##
##  2-sample test for equality of proportions without continuity
##  correction
##
## data:  c(62, 59) out of c(212, 173)
## X-squared = 1.0435, df = 1, p-value = 0.307
## alternative hypothesis: two.sided
## 95 percent confidence interval:
##  -0.14207371  0.04489845
## sample estimates:
##    prop 1    prop 2
## 0.2924528 0.3410405
```

Goodness-of-fit tests

Goodness-of-fit tests are conducted with the `chisq.test` function.

```
chisq.test(x, p)
```

- `x` is a vector of category counts.
- `p` is a vector of probabilities (same length as `x`). If not included, defaults to equal probabilities.

Goodness-of-fit test example

- A 6-sided die is rolled 100 times. The results are 24, 16, 13, 16, 17, 14 for each number respectively. Perform a goodness-of-fit test of equal probabilities.

```r
chisq.test(c(24, 16, 13, 16, 17, 14))
```

```
##
##  Chi-squared test for given probabilities
##
## data:  c(24, 16, 13, 16, 17, 14)
## X-squared = 4.52, df = 5, p-value = 0.4772
```

Goodness-of-fit test example

- Test whether a sample with category counts of 13, 48 and 31 matches an expected distribution of 25%, 50% and 25%.

```r
chisq.test(c(13, 48, 31), p=c(.25, .5, .25))
```

```
##
##  Chi-squared test for given probabilities
##
## data:  c(13, 48, 31)
## X-squared = 7.2174, df = 2, p-value = 0.02709
```

Tests for independence

Tests for independence are alse conducted with the `chisq.test` function.

```
chisq.test(x, correct)
```

- `x` is a matrix representing a contingency table.
- `correct` is whether to apply continuity correction. Should be set to `FALSE`.

Matrices

A matrix is created with the `matrix` function.

```
matrix(data, n.row)
```

- `data` is a vector containing the values to populate the matrix. The order of the values should be 1st column - top to bottom, 2nd column - top to bottom, etc.
- `n.row` is the number of rows of the matrix. The number of columns will then be deterined by the length of `data`

Matrix example

- Save the contingency table as a matrix.

	1A	1B
2A	10	30
2B	20	40

```
x <- matrix(c(10, 20, 30, 40), nrow=2)
x
```

```
##      [,1] [,2]
## [1,]   10   30
## [2,]   20   40
```

Test of independence example

- Perform a test of independence for the sample in the table below.

	Positive	Negative
Cancer	74	13
No cancer	26	887

```r
x <- matrix(c(74, 26, 13, 887), nrow=2)
chisq.test(x, correct=FALSE)
```

```
##
##  Pearson's Chi-squared test
##
## data:  x
## X-squared = 596.48, df = 1, p-value < 2.2e-16
```