Hypothesis test for population proportion

Date: 2017-10-05

What it is

A hypothesis test for p determines whether observed data is consistent with a hypothesized population proportion.

When to use it

The data conditions for this hypothesis test are similar to those for the confidence interval for p (a fixed sample of independent binary outcomes). In addition, there should be a specific numeric value of interest (e.g., $p = 0.5$, or a CDC vaccination coverage goal) to use as the null hypothesis. If there is no specific value of interest, a confidence interval is probably more appropriate.

How to use it

As with all hypothesis tests, begin with explicitly stating hypotheses. For this test, there are three options for the null hypothesis: $H_0 : p = p_0$, $H_0 : p \leq p_0$, or $H_0 : p \geq p_0$, where p_0 is the proportion of interest. A significance level, α, should also be chosen, usually 0.05.

In R, use the function `binom.test()`. Like with the confidence interval for p, we need to get the sample size and the number of successes. Suppose we want test whether the data is approximately 50% female[1]. The null hypothesis would be $H_0 : p = 0.5$, so we set the inputs `p=0.5` and `alternative='two.sided'` for the test:

[1] This is a common way to test representativeness: if the data is not consistent with a population with 50% females, the sample is potentially biased.

```
table(d$sex)
```

```
##
## female   male
##    267    233
```

```
binom.test(x = 267, n = 500, p = 0.5, alternative = "two.sided",
    conf.level = 0.95)
```

```
##
##  Exact binomial test
##
## data:  267 and 500
## number of successes = 267, number of
## trials = 500, p-value = 0.1399
## alternative hypothesis: true probability of success is not equal to 0.5
```

```
## 95 percent confidence interval:
##  0.4891844 0.5784114
## sample estimates:
## probability of success
##                  0.534
```

We see that about 53% of the sample is female, and this is within reason for a sample of this size if the true proportion is 0.5 ($p - value = 0.14$)[2]. Had the p-value been less than 0.05 (our chosen α), we would have had evidence that the data was not consistent with a 50/50 gender split.

[2] See also that the confidence interval covers the 0.5.

Why it works

The hypothesis test beging by assuming $p = p_0$. If this is actually true, then the Central Limit Theorem tells us that $\hat{p}$ is normaly distributed[3] with mean p_0 and variance $\frac{p_0(1-p_0)}{n}$. Then we know what values of $\hat{p}$ to expect from a sample.

[3] Using mathematical shorthand, $\hat{p} \sim \mathcal{N}\left(p_0, \frac{p_0(1-p_0)}{n}\right)$

Since about 95% of values are within 2 standard errors of p_0, anything outside the range $p_0 - 2\sqrt{\frac{p_0(1-p_0)}{n}}$ to $p_0 + 2\sqrt{\frac{p_0(1-p_0)}{n}}$ is "unexpected" and can be taken as evidence that $p \neq p_0$. These values of $\hat{p}$ far from p_0 correspond exactly to values of $\hat{p}$ that will give small p-values.