Test for the probability of a binomial distribution

Chao Cheng

August 29, 2022

For an i.i.d sample from a bernoulli distribution

$$x_1, \cdots, x_n \overset{\text{i.i.d.}}{\sim} Bernoulli(p),$$

The likelihood of the data is

$$f\left(x_1, \cdots, x_n\right) = \prod_{i=1}^{n} p^{x_i} \left(1-p\right)^{1-x_i} = p^{\sum x_i} \left(1-p\right)^{n-\sum x_i}.$$

MLE for p is $\bar{x} = \frac{1}{n}\sum x_i$ and

$$\sum_{i=1}^{n} x_i \sim Binom(n, p).$$

So here are mainly two situations: One is to test the probability p against some given value p_0. The other is to compare the probability between two independent random samples $x_1, \cdots, x_n$ and $y_1, \cdots, y_m$.

Case1: One sample $x_1, \cdots, x_n$ from $Bernoulli(p)$, and test p against a given p_0.

Case2: Two samples: $x_1, \cdots, x_2$ from $Bernoulli(p_1)$ and $y_1, \cdots, y_m$ from $Bernoulli(p_2)$. And test whether $p_1 = p_2$.

1 Normal approximation

1.1 Case 1

Note that

$$\mathrm{E}X = p, \quad \mathrm{Var}X = p\left(1-p\right).$$

Then by CLT we have

$$\bar{x} \overset{\text{asymp}}{\sim} N\left(p,\ \frac{p\left(1-p\right)}{n}\right).$$

For $H_0:\ p = p_0$, we propose a test statistic

$$Z = \frac{\bar{x} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}.$$

Then Z is asymptotically standard normal under H_0.

Also we know that under H_1:

$$
\begin{aligned}
Z =& \frac{\bar{x} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} \\
=& \frac{\bar{x} - p}{\sqrt{\frac{p(1-p)}{n}}} \cdot \sqrt{\frac{p\,(1-p)}{p_0\,(1-p_0)}} + \frac{p - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} \\
\sim& N\left(\frac{p - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}},\quad \frac{p\,(1-p)}{p_0\,(1-p_0)}\right).
\end{aligned}
$$

So the power of the test can be easily computed.

1.2 Case 2

So we have

$$
\bar{x} \overset{\text{asymp}}{\sim} N\left(p_1,\ \frac{p_1\,(1-p_1)}{n}\right), \quad \text{and} \quad \bar{y} \overset{\text{asymp}}{\sim} N\left(p_2,\ \frac{p_2\,(1-p_2)}{m}\right).
$$

A test statistic can be

$$
Z = \frac{\bar{x} - \bar{y}}{\sqrt{\hat{p}\,(1-\hat{p})\left(\frac{1}{n} + \frac{1}{m}\right)}},
$$

where $\hat{p} = \frac{n\bar{x}+m\bar{y}}{n+m}$. This test statistic can be found at

https://stats.stackexchange.com/questions/361015/proof-of-the-standard-error-of-the-distribution-between-two-normal-distributions/361048#361048

https://stats.stackexchange.com/questions/113602/test-if-two-binomial-distributions-are-statistically-different-from-each-other

Here this $\hat{p}\,(1-\hat{p})$ can be seen as an estimate for the variance $p\,(1-p)$ when H_0 is true by directly plugging in $\hat{p}$. This is **NOT** a pooled variance for these two samples, which should always be no greater than $\hat{p}\,(1-\hat{p})$.

The power of this test statistic is hard to compute under H_1.

Note: One can also use the same idea in the "t-test.pdf" notes and propose the test statistic

$$
T = \frac{\bar{x} - p_0}{\sqrt{S_x/n}},
$$

where $S_x = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})$ for Case 1.

And for Case 2

$$
T = \frac{\bar{x} - \bar{y} - \Delta}{\sqrt{\left(\frac{1}{n} + \frac{1}{m}\right) S_p}},
$$

where $S_p^2 = \frac{(n-1)S_x^2+(m-1)S_y^2}{n+m-2}$ and $\Delta = p_1 - p_2$. But again it is hard to evaluate the testing power of these statistics.

2 Chi-square approximation

See the notes of "chisq_test.pdf" for details.

3 Exact test

3.1 Case 1: Clopper-Pearson test

The Clopper-Pearson method is an early method. It's called exact method because it's directly based on p.m.f of binomial distribution. Let $X = \sum_{i=1}^{n} x_i$. Then $X \sim Binom(n, p)$ and the p.m.f is

$$f(x; p) = P(X = x|p) = C_n^x p^x (1-p)^{n-x}, \tag{1}$$

for $x = 0, 1, \cdots, n$. One thing to point out is that though the $|p$ notation, (1) is frequenist's point of view, not bayesian's. Now let's recall that p-value is the probability under H_0 that something **as or more extreme than** what we have observed happens. Then after observing $X = x_0$, for one-sided test:

- $H_0 : p \le p_0$ against $H_1 : p > p_0$ for some given p_0. The p-value at this observed x_0 is

$$p_{val}(x_0) = \sum_{x=x_0}^{n} f(x; p_0). \tag{2}$$

- $H_0 : p \ge p_0$ against $H_1 : p < p_0$ for some given p_0. The p-value at this observed x_0 is

$$p_{val}(x_0) = \sum_{x=0}^{x_0} f(x; p_0). \tag{3}$$

For the two-sided test. This is a little complicated. Let index set

$$\mathcal{I} = \{x|P(X = x|p_0) \le P(X = x_0|p_0), \quad 0 \le x \le n\}.$$

Then $\mathcal{I}$ contains all possible realizations of X with its probability no greater than the probability of our observation. Then the p-value of $H_0 : p = p_0$ at this observed x_0 is given by

$$p_{val}(x_0) = \sum_{x \in \mathcal{I}} f(x; p_0). \tag{4}$$

3.1.1 Power analysis

The probability to reject H_0 of Clopper-Pearson test at given underlying p can be computed by

$$P(\text{Reject } H_0|p) = \sum_{x=0}^{n} P(X = x|p) \cdot I_{\{p_{val}(x) \le \alpha\}} = \sum_{x=0}^{n} f(x; p) \cdot I_{\{p_{val}(x) \le \alpha\}}, \tag{5}$$

where α is the significant level of the test and $p_{val}(x)$ is computed for different types of H_0 based on (2), (3) and (4).

3.1.2 Confidence interval

3.2 Case 2: Fisher's exact test