M 358K

Supplement to Chapter 26: CONNECTION BETWEEN THE ONE-SAMPLE TEST FOR PROPORTIONS AND THE CHI-SQUARED GOODNESS-OF-FIT TEST

In 2007, 31.8% of births in the U.S. were by Caesarean section. Researchers are interested in whether or not that rate has changed. Suppose they have data on a random sample of n births in the U.S. in 2011. Of these, n_1 were births by Caesarean section, n_2 were not. (Note that $n_1 + n_2 = n$.)

There are two hypothesis tests the researchers could use with these data to answer their question. Both tests have the same null and alternate hypotheses. If we let p denote the proportion of births by Caesarean section in the U.S. in 2011, these null and alternate hypotheses are

H_0: p = 0.318.
H_A: p ≠ 0.318

One hypothesis test the researchers could use is the one-proportion z-test from Chapter 21. They would calculate the sample proportion $\hat{p} = n_1/n$ of births by C-section in the sample, and the standard deviation of $\hat{p}$ assuming the null hypothesis is true:

$$SD(\hat{p}) = \sqrt{\frac{(0.318)(1-0.318)}{n}}.$$

The test statistic would be

$$z = \frac{\hat{p}-0.318}{SD(\hat{p})},$$

which has approximately a standard normal distribution.

The second hypothesis test the researchers could use is the chi-square goodness-of-fit test, in Chapter 26. For this, the information could be arranged in a table as follows:

Type of Birth	**Observed Count**	**Expected Count**
C-Section	n_1	0.318n
Not C-Section	n_2	0.682n

Then calculate the chi-square statistic

$$\chi^2 = \sum \frac{(Obs-Exp)^2}{Exp}$$
$$= \frac{(n_1-0.318n)^2}{0.318n} + \frac{(n_2-0.682n)^2}{0.682n},$$

which has approximately a $\chi^2(1)$ distribution.

Will these two hypothesis tests give the same result? We can get a hint for what might be happening by remembering that the square of a standard normal random variable has a $\chi^2(1)$ distribution. So let's look at the square of the z-statistic for the test for a proportion. To make the reasoning more general, we'll use p_0 instead of 0.318 for the proportion in the null hypothesis. So

$$z = \frac{\hat{p}-p_0}{SD(\hat{p})} = \frac{\frac{n_1}{n}-p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}.$$

Squaring z and using some algebra gives

$$z^2 = \frac{\left(\frac{n_1}{n}-p_0\right)^2}{\frac{p_0(1-p_0)}{n}}$$

$$= \frac{(n_1-np_0)^2}{n^2}\,\frac{n}{p_0(1-p_0)}$$

$$= \frac{(n_1-np_0)^2}{np_0(1-p_0)}$$

Using p_0 instead of 0.318 in the formula for the χ^2 statistic above and using some algebra gives

$$\chi^2 = \frac{(n_1-p_0n)^2}{p_0n} + \frac{[n_2-(1-p_0)n]^2}{(1-p_0)n}$$

$$= \frac{(n_1-p_0n)^2}{p_0n} + \frac{[(n-n_1)-(1-p_0)n]^2}{(1-p_0)n}$$

$$= \frac{(n_1-p_0n)^2}{p_0n} + \frac{(p_0n-n_1)^2}{(1-p_0)n}$$

$$= (n_1 - np_0)^2\left[\frac{1}{np_0} + \frac{1}{n(1-p_0)}\right]$$

Using common denominator $np_0(1-p_0)$ to add the fractions $\frac{1}{np_0}$ and $\frac{1}{n(1-p_0)}$ gives $\frac{(1-p_0)+p_0}{np_0(1-p_0)} = \frac{1}{np_0(1-p_0)}$, which shows that indeed the square of the z-statistic for the 1-proportion z-test is just the χ^2 statistic for the chi-square goodness of fit test.

Now consider p-values. Let z_{obs} denote the value of the z-statistic obtained for the one proportion test. The p-value for that test is then

$$P_{prop} = \text{Prob}(Z < |z_{obs}|),$$

where Z denotes a standard random variable.

From what we have shown above, the value of the χ^2statistic calculated from the same data for the chi-square goodness-of-fit test is $(z_{obs})^2$. So the p-value for the chi-square goodness-of-fit test is

$$P_{chi} = \text{Prob}(Y < (z_{obs})^2),$$

where Y denotes a random variable with a $\chi^2(1)$ distribution. But since Z is standard normal, $Z^2 \sim \chi^2(1)$ (by the definition of χ^2 random variable).

$$P_{chi} = \text{Prob}\,(Z^2 < (z_{obs})^2)$$

Since $Z^2 < (z_{obs})^2$ if and only if $Z < |z_{obs}|$, we have

$$P_{chi} = \text{Prob}\,(Z < |z_{obs}|) = P_{prop}.$$

In other words, *the one-proportion z-test and the chi-squared goodness-of-fit test will give the same p-value when applied to the same data.* Consequently, one test will reject the null hypothesis if and only if the other test rejects it.

Thus we can consider the chi-squared goodness-of-fit test (when we allow more than two categories) to be a generalization of the z-test for a single proportion.