Chi-squared test example

Example: 2×3 contingency table

Patients suffering from a chronic condition are recruited to a randomized clinical trial with three treatment groups:

- Drug 1
- Drug 2
- Drug 3

The alleviation of the chronic condition was studied:

- Condition did not improve
- Condition did improve

Data:

Improved	Drug 1	2	3	Total
No	10	12	10	32
Yes	10	18	15	43
Total	20	30	25	75

Notation:

	Drug			
Improved	1	2	3	Total
No	n_{11}	n_{12}	n_{13}	$n_{1.}$
Yes	n_{21}	n_{22}	n_{23}	$n_{2.}$
Total	$n_{.1}$	$n_{.2}$	$n_{.3}$	n

Scientific question

Is there a difference between the improvement rates across drugs ?

Null Hypothesis:

$$H_0 \quad : \quad \text{Improvement rate is the same across drugs.}$$

Under the null hypothesis, there is no difference between the drugs, and we would consider the elements in the columns of the table as realizations of three independent Binomial sampling schemes with the same probability of success p.

The best estimate of the common improvement rate p is

$$\widehat{p} = \frac{n_{2\cdot}}{n} = \frac{43}{75} = 0.573$$

That is

$$\begin{aligned} n_{21} &\sim Binomial(n_{.1}, p) \\ n_{22} &\sim Binomial(n_{.2}, p) \\ n_{23} &\sim Binomial(n_{.3}, p) \end{aligned}$$

and thus, under H_0, we use the fitted values determined by the expectations of these distributions

$$\begin{aligned} \widehat{n}_{21} &= n_{.1}\widehat{p} = \frac{n_{.1} n_{2.}}{n} \\ \widehat{n}_{22} &= n_{.2}\widehat{p} = \frac{n_{.2} n_{2.}}{n} \\ \widehat{n}_{23} &= n_{.3}\widehat{p} = \frac{n_{.3} n_{2.}}{n} \end{aligned}$$

We can complete the "fitted" or "expected" table by noting that we must have

$$
\begin{aligned}
n_{.1} &= \widehat{n}_{11} + \widehat{n}_{21} \\
n_{.2} &= \widehat{n}_{12} + \widehat{n}_{22} \\
n_{.3} &= \widehat{n}_{13} + \widehat{n}_{23}
\end{aligned}
$$

This is sufficient to allow us to complete the "fitted" table

Fitted table under H_0

Improved	Drug 1	2	3	Total
No	8.533	12.800	10.667	32
Yes	11.467	17.200	14.333	43
Total	20.000	30.000	25.000	75

so that the test statistic is

$$X^2 = \sum_{i=1}^{2} \sum_{j=1}^{3} \frac{(n_{ij} - \widehat{n}_{ij})^2}{\widehat{n}_{ij}} = 0.599$$

Completing the test

Under H_0, we have that

$$X^2 \dot{\sim} \text{Chi-squared}((r-1)(c-1))$$

Here $r = 2$, $c = 3$, so

$$X^2 \dot{\sim} \text{Chi-squared}(2)$$

From the tables in McClave and Sincich Table VII, we find that

$$\text{Chisq}_{0.05}(2) = 5.991$$

so we **do not reject** the null hypothesis, and conclude that there is no evidence of a difference between the improvement rates for the three drugs.