

A scientist wants to study the null hypothesis “boy and girl births are equally likely *and* independent within a family”. To this end he samples 1000 random two-child families and obtains the following statistics:

$$\frac{\begin{array}{ccc} 2 \text{ boys} & 1 \text{ boy and 1 girl} & 2 \text{ girls} \\ 222 & 492 & 286 \end{array}}{}$$

Based on this data, can he reject the null hypothesis with 95% confidence?

Solution: Under the null hypothesis, the number of girls in a family should be a Binomial($2, \frac{1}{2}$) random variable, so among 1000 samples the scientist would expect to see 250 two-boy families, 500 one-boy one-girl families, and 250 two-girl families. The chisquare statistic is

$$X^2 = \frac{(222 - 250)^2}{250} + \frac{(492 - 500)^2}{500} + \frac{(286 - 250)^2}{250} = 8.448.$$

The p-value is the probability that a $\chi^2(3 - 1)$ random variable takes value at least 8.448, which is about 0.015. As this is under 5% the scientist can reject the null hypothesis with 95% confidence.