Chi-Square Goodness of Fit Test

When an analyst attempts to fit a statistical model to observed data, he or she may wonder how well the model actually reflects the data. How "close" are the observed values to those which would be expected under the fitted model? One statistical test that addresses this issue is the chi-square goodness of fit test. In general, the ***chi-square test statistic*** is of the form

$$X^2 = \sum \frac{(\text{observed - expected})^2}{\text{expected}}.$$

If the computed test statistic is large, then the observed and expected values are not close and the model is a poor fit to the data.

Example

A new casino game involves rolling 3 dice. The winnings are directly proportional to the total number of sixes rolled. Suppose a gambler plays the game 100 times, with the following observed counts:

Number of Sixes	Number of Rolls
0	48
1	35
2	15
3	3

The casino becomes suspicious of the gambler and wishes to determine whether the dice are fair. What do they conclude?

If a die is fair, we would expect the probability of rolling a 6 on any given toss to be 1/6. Assuming the 3 dice are independent (the roll of one die should not affect the roll of the others), we might assume that the number of sixes in three rolls is distributed Binomial (3,1/6). To determine whether the gambler's dice are fair, we may compare his results with the results expected under this distribution. The expected values for 0, 1, 2, and 3 sixes under the Binomial(3,1/6) distribution are the following:

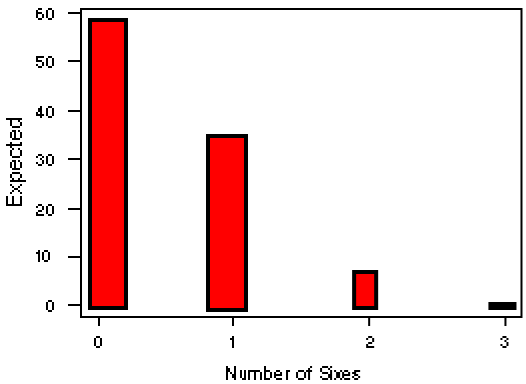
Null Hypothesis:
p_1 = P(roll 0 sixes) = P(X=0) = 0.58
p_2 = P(roll 1 six) = P(X=1) = 0.345
p_3 = P(roll 2 sixes) = P(X=2) = 0.07
p_4 = P(roll 3 sixes) = P(X=3) = 0.005.

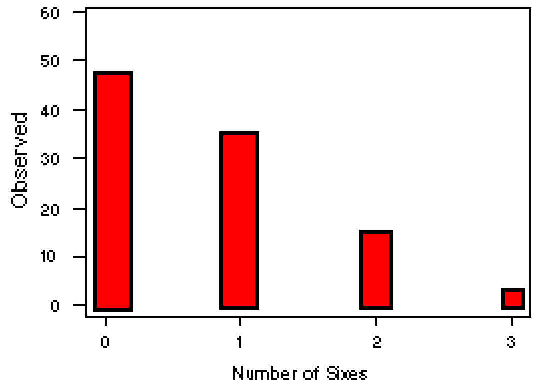
Since the gambler plays 100 times, the expected counts are the following:

Number of Sixes	Expected Counts	Observed Counts
0	58	48
1	34.5	35
2	7	15
3	0.5	3

Chi-Square Goodness of Fit Test

The two plots shown below provide a visual comparison of the expected and observed values:





The chi-square statistic is the sum of the squares of the plotted values,
$(48-58)^2/58 + (35-34.5)^2/58 + (15-7)^2/7 + (3-0.5)^2/0.5$
$= 1.72 + 0.007 + 9.14 + 12.5 = 23.367$.

Given this statistic, are the observed values likely under the assumed model?

The standardized counts (observed - expected)/sqrt(expected) for *k* possibilities are approximately normal, but they are not independent because one of the counts is entirely determined by the sum of the others (since the total of the observed and expected counts must sum to *n*). This results in a loss of one degree of freedom, so it turns out the the distribution of the chi-square test statistic based on *k* counts is approximately the chi-square distribution with $m = k-1$ degrees of freedom, denoted $\chi^2(k-1)$.

Example
In the gambling example above, the chi-square test statistic was calculated to be 23.367. Since $k = 4$ in this case (the possibilities are 0, 1, 2, or 3 sixes), the test statistic χ^2 is associated with the chi-square distribution with 3 degrees of freedom. If we are interested in a significance level of 0.05 we may reject the null hypothesis (that the dice are fair) if $\chi^2 \geq 7.815$, the value corresponding to the 0.05 significance level for the $\chi^2(3)$ distribution. Since 23.367 is clearly greater than 7.815, we may reject the null hypothesis that the dice are fair at the 0.05 significance level.

Given this information, the casino asked the gambler to take his dice (and his business) elsewhere.