

# When To Use Fisher's Exact Test

by Keith M. Bower, M.S.

*Reprinted with permission from the American Society for Quality*

Six Sigma practitioners occasionally conduct studies to assess differences between items such as operators or machines. When the experimental data are measured on a continuous scale (measuring nozzle diameter in microns, for example), a procedure such as "Student's" two-sample t-test may be appropriate.<sup>1</sup> When the response variable is recorded using counts, however, Karl Pearson's  $\chi^2$  test may be employed.<sup>2</sup>

But when the number of observations obtained for analysis is small, the  $\chi^2$  test may produce misleading results. A more appropriate form of analysis (when presented with a 2 \* 2 contingency table) is to use R.A. Fisher's exact test.

## Example

On the *Late Show With David Letterman*, the host (David) and the show's musical director (Paul Shaffer) frequently assess whether particular items will or will not float when placed in a tank of water. Let's assume Letterman guessed correctly for eight of nine items, and Shaffer guessed correctly for only four items. Let's also assume all the items have the same probability of being guessed.

**Figure 1**

	<b>Guessed correctly</b>	<b>Guessed incorrectly</b>	<b>Total</b>
<b>Letterman</b>	8	1	<b>9</b>
<b>Shaffer</b>	4	5	<b>9</b>
<b>Total</b>	<b>12</b>	<b>6</b>	<b>18</b>

You would typically use the  $\chi^2$  test when presented with the contingency table results in Figure 1. In this case, the  $\chi^2$  test assesses what the expected frequencies would be if the null hypothesis (equal proportions) was true. For example, if there were no difference between Letterman and Shaffer's guesses, you would expect Letterman to have been correct six times (see Figure 2). This is calculated as  $(9 * 12) / 18 = 108 / 18 = 6$ .

**Figure 2**

Rows: Player	Columns: Result		
	Correct	Incorrect	All
David	8 6.00	1 3.00	9 9.00
Paul	4 6.00	5 3.00	9 9.00
All	12 12.00	6 6.00	18 18.00

Chi-Square = 4.000, DF = 1, P-Value = 0.046  
 2 cells with expected counts less than 5.0

Cell Contents --  
 Count  
 Exp Freq

The resulting p-value, 0.046, from the  $\chi^2$  test indicates there is a statistically significant difference (at the  $\alpha = 0.05$  level) in the success rates between Letterman and Shaffer.

As Fisher discusses, however, “The treatment of frequencies by means of  $\chi^2$  is an approximation, which is useful for the comparative simplicity of the calculations. The exact treatment is somewhat more laborious, though necessary in cases of doubt.”<sup>3</sup>

Some practitioners will experience a problem when an expected value is less than five (this is what Fisher alludes to in his statement of doubt). Sometimes it's appropriate to group certain categories to avoid the problem, but this is clearly not possible when there are only two categories. As shown in Figure 2, there are two cells in which the expected counts are less than five.

Fisher's exact test considers all the possible cell combinations that would still result in the marginal frequencies as highlighted (namely 9, 9 and 12, 6). The test is exact because it uses the exact hypergeometric distribution rather than the approximate chi-square distribution to compute the p-value.

The resulting p-value using Fisher's exact test is 0.1312. Therefore, you would fail to reject the null hypothesis of equal proportions at the  $\alpha = 0.05$  level. This contradicts the results from the  $\chi^2$  test and indicates the  $\chi^2$  test provided a poor approximation to the exact results.

The computations involved in Fisher's exact test may be extremely time consuming to calculate by hand, but are in the sidebar "Calculations for Fisher's Exact Test" for illustration.<sup>4</sup> Clearly, it's much easier to use a statistical software package to obtain these results.

## **Implications**

It's appropriate to use Fisher's exact test, in particular when dealing with small counts. The  $\chi^2$  test is basically an approximation of the results from the exact test, so erroneous results could potentially be obtained from the few observations. This could lead to incorrect conclusions in Six Sigma projects.

## **References**

1. For more information on the two-sample t-test, see "The Two-Sample t-Test and Randomization Test" by Keith M. Bower, *Six Sigma Forum*, June 13, 2003.
2. Karl Pearson, "On the Criterion That a Given System of Deviations From the Probable in the Case of Correlated System of Variables Is Such That It Can Be Reasonably Supposed To Have Arisen From Random Sampling," *Philosophical Magazine*, Series V, No. 1, (1900), pp. 157-175.

3. Ronald A. Fisher, *Statistical Methods for Research Workers*, 14th edition, Hafner Publishing, 1970, p. 96.
4. To compute Fisher's exact test, refer to answer 210 at <http://www.minitab.com/support/answers>

## Bibliography

Fleiss, Joseph L., *Statistical Methods for Rates and Proportions*, John Wiley and Sons, 1981.

Montgomery, Douglas C., and George C. Runger, *Applied Statistics and Probability for Engineers*, second edition, John Wiley and Sons, 1999.

## SIDEBAR

### Calculations for Fisher's Exact Test

The hypergeometric probability distribution is used to compute the probability of the observed results (see Table 1).

**Table 1**

	Correct	Incorrect	<b>Total</b>
Letterman	8	1	<b>9</b>
Shaffer	4	5	<b>9</b>
<b>Total</b>	<b>12</b>	<b>6</b>	<b>18</b>

The remaining tables that will be consistent with the marginal frequencies of 9, 9 and 12, 6, along with their associated probabilities, are shown in Table 2.

To compute Fisher's exact test results, look at the tables with probabilities less than or equal to the probability of the observed results (0.061085972). They are highlighted with an \*. Add these probabilities together, along with the probability of the observed results, to obtain the p-value for the test.

**Table 2**

<b>Table</b>	<b>Associated Probability</b>
<b>9 0</b> <b>3 6</b>	$\frac{9! \cdot 9! \cdot 12! \cdot 6!}{18! \cdot 9! \cdot 0! \cdot 3! \cdot 6!} = 0.004524887 *$
<b>7 2</b> <b>5 4</b>	$\frac{9! \cdot 9! \cdot 12! \cdot 6!}{18! \cdot 7! \cdot 2! \cdot 5! \cdot 4!} = 0.244343891$
<b>6 3</b> <b>6 3</b>	$\frac{9! \cdot 9! \cdot 12! \cdot 6!}{18! \cdot 6! \cdot 3! \cdot 6! \cdot 3!} = 0.380090498$
<b>5 4</b> <b>7 2</b>	$\frac{9! \cdot 9! \cdot 12! \cdot 6!}{18! \cdot 5! \cdot 4! \cdot 7! \cdot 2!} = 0.244343891$
<b>4 5</b> <b>8 1</b>	$\frac{9! \cdot 9! \cdot 12! \cdot 6!}{18! \cdot 4! \cdot 5! \cdot 8! \cdot 1!} = 0.061085973 *$
<b>3 6</b> <b>9 0</b>	$\frac{9! \cdot 9! \cdot 12! \cdot 6!}{18! \cdot 3! \cdot 6! \cdot 9! \cdot 0!} = 0.004524887 *$

This particular p-value is 0.13122.