The word *contingent* has a few different meanings, one of which refers to a *dependence* on some other factor. We use the term *contingency table* because we test for *independence* between the row and column variables. We first define a *test of independence* and we provide key elements of the test in the box that follows.

DEFINITION In a **test of independence,** we test the null hypothesis that in a contingency table, the row and column variables are independent. (That is, there is no dependency between the row variable and the column variable.)

Objective

Conduct a hypothesis test for independence between the row variable and column variable in a contingency table.

Notation

O represents the *observed frequency* in a cell of a contingency table.

E represents the *expected frequency* in a cell, found by assuming that the row and column variables are independent.

r represents the number of rows in a contingency table (not including labels).

c represents the number of columns in a contingency table (not including labels).

Requirements

1. The sample data are randomly selected.

2. The sample data are represented as frequency counts in a two-way table.

3. For every cell in the contingency table, the expected frequency E is at least 5. (There is no requirement that every observed frequency must be at least 5. Also, there is no requirement that the population must have a normal distribution or any other specific distribution.)

Null and Alternative Hypotheses

The null and alternative hypotheses are as follows:

H_0: The row and column variables are independent.

H_1: The row and column variables are dependent.

Test Statistic for a Test of Independence

$$\chi^2 = \sum \frac{(O - E)^2}{E}$$

where O is the observed frequency in a cell and E is the expected frequency found by evaluating

$$E = \frac{(\text{row total})\,(\text{column total})}{(\text{grand total})}$$

P-Values

P-values are typically provided by technology, or a range of *P*-values can be found from Table A-4.

Critical Values

1. The critical values are found in Table A-4 using

$$\textbf{degrees of freedom} = (r - 1)(c - 1)$$

where r is the number of rows and c is the number of columns.

2. Tests of independence with a contingency table are always *right-tailed.*