A researcher designs an experiment in which a rat is attracted to the end of a ramp that divides, leading to doors of three different colors. The researcher sends the rat down the ramp $n = 90$ times and observes the choices listed in Table 14.1. Does the rat have (or acquire) a preference for one of the three doors?

Rat's Door Choices

	Door		
	Green	Red	Blue
Observed Count (O_i)	20	39	31

Solution If the rat has no preference in the choice of a door, you would expect in the long run that the rat would choose each door an equal number of times. That is, the null hypothesis is

$$H_0: p_1 = p_2 = p_3 = \frac{1}{3}$$

versus the alternative hypothesis

$$H_a: \text{At least one } p_i \text{ is different from } \frac{1}{3}$$

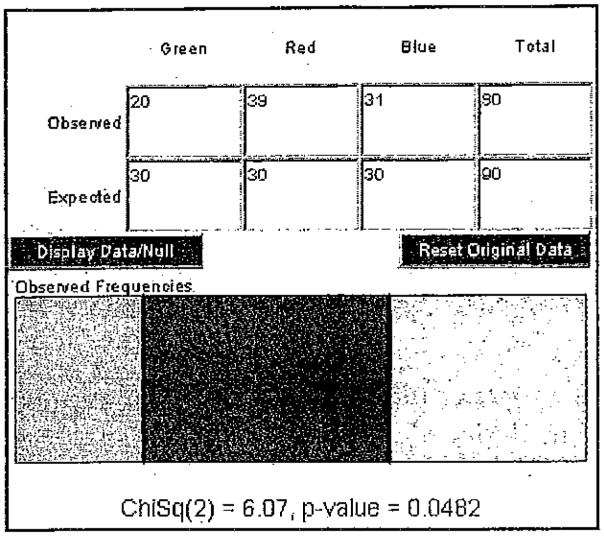
where p_i is the probability that the rat chooses door i, for $i = 1, 2,$ and 3. The expected cell counts are the same for each of the three categories—namely, $np_i = 90(1/3) = 30$. The chi-square test statistic can now be calculated as

$$X^2 = \Sigma\frac{(O_i - E_i)^2}{E_i}$$

$$= \frac{(20-30)^2}{30} + \frac{(39-30)^2}{30} + \frac{(31-30)^2}{30} = 6.067$$

For this example, the test statistic has $(k - 1) = 2$ degrees of freedom because the only linear restriction on the cell probabilities is that they must sum to 1. Hence, you can use Table 5 in Appendix I to find bounds for the right-tailed p-value. Since the observed value, $X^2 = 6.067$, lies between $\chi^2_{.050} = 5.99$ and $\chi^2_{.025} = 7.38$, the p-value is between .025 and .050. The researcher would report the results as significant at the 5% level ($P < .05$), meaning that the null hypothesis of no preference is rejected. There is sufficient evidence to indicate that the rat has a preference for one of the three doors.

What more can you say about the experiment once you have determined statistically that the rat has a preference? Look at the data to see where the differences lie. **The Goodness-of-Fit Test** applet, shown in Figure 14.1, will help.



The proportions of blood phenotypes A, B, AB, and O in the population of all Caucasians in the United States are .41, .10, .04, and .45, respectively. To determine whether or not the actual population proportions fit this set of reported probabilities, a random sample of 200 Americans were selected and their blood phenotypes were recorded. The observed and expected cell counts are shown in Table 14.2. The expected cell counts are calculated as $E_i = 200p_i$. Test the goodness of fit of these blood phenotype proportions.

Counts of Blood Phenotypes

	A	B	AB	O
Observed (O_i)	89	18	12	81
Expected (E_i)	82	20	8	90

Solution The hypothesis to be tested is determined by the model probabilities:

$$H_0 : p_1 = .41; p_2 = .10; p_3 = .04; p_4 = .45$$

versus

$$H_a : \text{At least one of the four probabilities is different from the specified value}$$

Then

$$X^2 = \Sigma \frac{(O_i - E_i)^2}{E_i}$$

$$= \frac{(89 - 82)^2}{82} + \cdots + \frac{(81 - 90)^2}{90} = 3.70$$

From Table 5 in Appendix I, indexing $df = (k - 1) = 3$, you can find that the observed value of the test statistic is less than $\chi^2_{.100} = 6.25$, so that the p-value is greater than .10. You do not have sufficient evidence to reject H_0; that is, you cannot declare that the blood phenotypes for American Caucasians are *different* from those reported earlier. The results are nonsignificant (NS).

A total of $n = 309$ furniture defects were recorded and the defects were classified into four types: A, B, C, or D. At the same time, each piece of furniture was identified by the production shift in which it was manufactured. These counts are presented in a contingency table in Table 14.3.

Contingency Table

	Shift			
Type of Defects	1	2	3	Total
A	15	26	33	74
B	21	31	17	69
C	45	34	49	128
D	13	5	20	38
Total	94	96	119	309

The Chi-Square Test of Independence

The question of independence of the two methods of classification can be investigated using a test of hypothesis based on the chi-square statistic. These are the hypotheses:

H_0 : The two methods of classification are independent
H_a : The two methods of classification are dependent

To explain how to estimate these expected cell counts, we must revisit the concept of *independent events* from Chapter 4. Consider p_{ij}, the probability that an observation falls into row i and column j of the contingency table. If the rows and columns are independent, then

$$\begin{aligned} p_{ij} &= P(\text{observation falls in row } i \text{ and column } j) \\ &= P(\text{observation falls in row } i) \times P(\text{observation falls in column } j) \\ &= p_i p_j \end{aligned}$$

where p_i and p_j are the **unconditional** or **marginal probabilities** of falling into row i or column j, respectively. If you could obtain proper estimates of these marginal probabilities, you could use them in place of p_{ij} in the formula for the expected cell count.

Fortunately, these estimates do exist. In fact, they are exactly what you would intuitively choose:

- To estimate a row probability, use $\hat{p}_i = \dfrac{\text{Total observations in row } i}{\text{Total number of observations}} = \dfrac{r_i}{n}$,
- To estimate a column probability, use $\hat{p}_j = \dfrac{\text{Total observations in column } j}{\text{Total number of observations}}$

$$= \frac{c_j}{n}.$$

The estimate of the expected cell count for row i and column j follows from the independence assumption.

ESTIMATED EXPECTED CELL COUNT

$$\hat{E}_{ij} = n\left(\frac{r_i}{n}\right)\left(\frac{c_j}{n}\right) = \frac{r_i c_j}{n}$$

where r_i is the total for row i and c_j is the total for column j.

The chi-square test statistic for a contingency table with r rows and c columns is calculated as

$$X^2 = \Sigma\frac{(O_{ij} - \hat{E}_{ij})^2}{\hat{E}_{ij}}$$

and can be shown to have an approximate chi-square distribution with

$$df = (r - 1)(c - 1)$$

If the observed value of X^2 is too large, then the null hypothesis of independence is rejected.

Solution . The estimated expected cell counts are shown in parentheses in Table 14.4. For example, the estimated expected count for a type C defect produced during the second shift is

$$\hat{E}_{32} = \frac{r_3 c_2}{n} = \frac{(128)(96)}{309} = 39.77$$

Observed and Estimated Expected Cell Counts

	Shift			
Type of Defects	1	2	3	Total
A	15 (22.51)	26 (22.99)	33 (28.50)	74
B	21 (20.99)	31 (21.44)	17 (26.57)	69
C	45 (38.94)	34 (39.77)	49 (49.29)	128
D	13 (11.56)	5 (11.81)	20 (14.63)	38
Total	94	96	119	309

You can now use the values shown in Table 14.4 to calculate the test statistic as

$$X^2 = \Sigma\frac{(O_{ij} - \hat{E}_{ij})^2}{\hat{E}_{ij}}$$

$$= \frac{(15 - 22.51)^2}{22.51} + \frac{(26 - 22.99)^2}{22.99} + \cdots + \frac{(20 - 14.63)^2}{14.63}$$

$$= 19.18$$

When you index the chi-square distribution in Table 5 of Appendix I with

$$df = (r - 1)(c - 1) = (4 - 1)(3 - 1) = 6$$

the observed test statistic is greater than $\chi^2_{.005} = 18.5476$, which indicates that the p-value is less than .005. You can reject H_0 and declare the results to be highly significant ($P < .005$). There is sufficient evidence to indicate that the proportions of defect types vary from shift to shift.