Step 6: State the conclusion in terms of the original problem.

At the 0.10 level, there is insufficient evidence to conclude that the distribution of the card types in the set differs from the company claim that the proportion of each card type is approximately equal.

16.2 Exercises

Basic Concepts

1. Describe what the test statistic for the chi-square test for goodness of fit measures.
2. What is a multinomial probability distribution? What more familiar probability distribution discussed previously in the text is a multinomial probability distribution related to?
3. List the four requirements for a multinomial experiment.
4. What are the null and alternative hypotheses for a chi-square test for goodness of fit?
5. What is the test statistic for a chi-square test for goodness of fit?
6. How many degrees of freedom does the test statistic for the chi-square test for goodness of fit have?
7. What assumptions are necessary for a chi-square test for goodness of fit?
8. How are the expected values determined in a chi-square test for goodness of fit?

Exercises

9. An internet provider claims that the service calls they receive are equally distributed among the five working days of the week. A survey of 85 randomly selected service calls produced the following results.

Service Calls					
	Monday	**Tuesday**	**Wednesday**	**Thursday**	**Friday**
Number of Calls	15	20	25	15	10

   a. Is the company's claim refuted by the data at $\alpha = 0.05$?

   b. What assumptions were made in the test for part **a.**?

10. Suppose a consumer affairs representative for Mars Incorporated claims that M&M's plain chocolate candies are mixed such that each large production batch has "precisely" the following ratios of colored candies: 30% brown, 20% yellow, 20% red, 10% orange, 10% green, and 10% blue. To test this claim, a professor distributed small sample bags of M&M's to students and had them count the number of candies of each color. The counts of the students were then pooled with the following results.

Candy Colors							
	Brown	Yellow	Red	Orange	Green	Blue	Total
Number of Candies	84	79	75	49	36	47	370

a. If the representative's claim is true, what would be the expected number of candies in each of the color categories for 370 candies?

b. Is the representative's claim refuted by the data at $\alpha = 0.01$?

c. What assumptions were made in performing the test for part **b.**?

11. A highway department executive claims that the number of fatal accidents which occur in her state does not vary from month to month. A survey of 170 fatal accidents produced the following results.

Accidents												
	Jan.	Feb.	Mar.	Apr.	May	Jun.	July	Aug.	Sept.	Oct.	Nov.	Dec.
Accidents	18	16	7	5	8	12	15	18	15	11	20	25

a. Is the executive's claim refuted by the data at $\alpha = 0.01$?

b. What assumptions were made in the test for part **a.**?

12. A psychologist conducted an attitude survey of 200 randomly selected individuals several years ago. The individuals were asked to pick the one category which most accurately described their attitudes. The results of the survey were as follows.

1st Attitude Survey	
Attitude	Percent of Respondents
Optimistic	15%
Slightly Optimistic	30%
Slightly Pessimistic	30%
Pessimistic	25%

The psychologist believes that these attitudes have changed over time. To test this theory, he randomly selects 200 individuals and asks them the same questions. The results of the second survey are as follows.

2nd Attitude Survey	
Attitude	Percent of Respondents
Optimistic	20%
Slightly Optimistic	40%
Slightly Pessimistic	30%
Pessimistic	10%

a. Can the psychologist conclude that the attitudes have changed over time at $\alpha = 0.01$?

b. What assumptions were made in the test for part **a.**?

13. A manager for an insurance company believes that customers have the following preferences for life insurance products: 20% prefer Whole Life, 30% prefer Universal Life, and 50% prefer Life Annuities. The results of a survey of 240 customers were tabulated.

Life Insurance Preferences	
Product	**Number**
Whole Life	60
Universal Life	72
Life Annuities	108

a. State the null and alternative hypothesis for testing the insurance manager's claim.

b. Calculate the expected number of customers who prefer Whole Life policies.

c. Calculate the test statistic. (**Note:** keep intermediate calculations to six decimal places.)

d. Can we refute the insurance manager's claim for customers preferring each insurance product at a significance level of 0.05?

16.3 The Chi-Square Test for Association

Our interest sometimes extends beyond one variable to summarizing the relationship between two qualitative variables. For example, a radio executive might be interested in knowing if the marital status of an individual affects that person's preference for music. Here the qualitative variables of interest are marital status, which could take on values such as single, married, divorced, or widowed, and preference for music, which could take on the values Classical, Jazz, Easy Listening, Rock, Rap, and Country. Other examples of relationships between two qualitative variables which might be of interest are age and political preference, education level and job performance, income level and occupation, etc. When we are interested in this type of relationship, we often make use of a contingency table.

Contingency Table (or Two-Way Frequency Table)

A **contingency table** organizes data on two characteristics simultaneously. Each cell in a contingency table contains either a count or a proportion which represents the number of observations falling into that cell. It is important to note that contingency tables are composed of two variables, each satisfying the properties of the multinomial distribution.

DEFINITION

Table 16.3.1 shows the general form of a contingency table.

Table 16.3.1 - General Form of a Contingency Table

		Factor A		
		Level A1	**Level A2**	**Total**
Factor B	**Level B1**	n_1	n_2	$n_1 + n_2$
	Level B2	n_3	n_4	$n_3 + n_4$
Total		$n_1 + n_3$	$n_2 + n_4$	$n = n_1 + n_2 + n_3 + n_4$