

1. A human resources manager is interested in understanding the preferred training methods among employees. The manager surveyed a random sample of 115 employees across three departments. Each employee was asked to indicate their department (sales, marketing, or finance) and choose between two training methods (online course or in-person workshop). The results are summarized in the table provided.

	Online Courses	In-person Workshops	Total
Sales	15	32	47
Marketing	25	16	41
Finance	16	11	27
Total	56	59	115

The manager is wondering if there is an association between an employee's department and their preferred training method.

- a. Write the hypotheses for the appropriate hypothesis test.
- b. Assuming that the null hypothesis is true, fill in the table with the expected counts in each category. Round to the nearest hundredth.

	Online Courses	In-person Workshops
Sales		
Marketing		
Finance		

2. For each scenario, determine which type of test is appropriate. Give a reason for your answer.
- a. An athletic association is investigating the distribution of injuries to different body parts among athletes participating in three different sports: football, soccer, and volleyball. The association wants to determine if there is a difference in the distribution of injuries across these sports. To investigate, injury data from a random sample of 25 athletes in each sport is collected, and the proportions of injuries to various body parts are recorded.
 - b. A seed company claims that their packets of flower seeds contain a mix of different flower types where 20% are roses. To test this claim, a gardener randomly selects 80 seeds from a packet and plants them. After the seeds have grown into flowers, the gardener records the types of flowers that have bloomed.
 - c. A city planner wants to investigate whether there is an association between the type of neighborhood and the preferred mode of commuting to work. The planner collects data from a random sample of 523 residents. Each resident is asked whether they live in an urban, suburban, or rural neighborhood and if they prefer commuting by car, public transportation, bicycle, or walking.

3. A social scientist was interested if the type of home is related to the level of interaction with neighbors. A random sample of 200 residents of a particular city was asked to identify if they lived in a house or apartment and if they would consider their interactions with their neighbors as frequent or limited. The data is given in the table.

	Frequent Interaction	Limited Interaction	Total
House	79	53	132
Apartment	20	48	68
Total	99	101	200

Do the data provide convincing statistical evidence at the $\alpha = 0.05$ significance level of an association between home type and level of interaction with neighbors in this city?

4. A random sample of employed adults was surveyed regarding their work. The two-way table displays the results. Suppose an employee from the sample is selected at random.

	Works on weekends	Does not work on weekends	Total
Full-time	28	72	100
Part-time	8	14	22
Total	36	86	122

- What is the probability that the employee works on weekends?
- Given that the employee works full-time, what is the probability that the employee works on weekends?
- Given that the employee works part-time, what is the probability that the employee works on weekends?
- Are the events “works on weekends” and “full-time” independent? Explain.
- A researcher wonders if there is an association between whether an employee works full-time or part-time and whether the employee works on weekends or not. A chi-square test of independence from the sample data in the table gives a test statistic of $\chi^2 = 0.6064$. Calculate the associated p-value.
- Is there convincing statistical evidence that there is an association between whether an employee works full-time or part-time and whether the employee works on weekends or not? Use $\alpha = 0.05$.
- Explain the differences in the results from parts (d) and (f).

5. A social media company offers two main types of content formats: images and videos. The company wants to determine if there is an association between the type of content posted by users and their engagement levels, measured by the number of likes and shares. Data is collected from a random sample of user posts. Each post is categorized based on its content format (image or video) and its level of engagement (high or low). Which of the following hypothesis tests would be best used to investigate this question?
- A) one-sample z -test for p
 - B) two-sample z -test for $p_1 - p_2$
 - C) chi-square test for homogeneity
 - D) chi-square test of independence
6. A hotel manager is interested in determining if more customers take advantage of the complimentary breakfast in the winter season than in the summer season. The manager selects two independent random samples of guests: one from the summer season and another from the winter season. In the summer sample, 81 out of 200 guests choose to have the free breakfast, while in the winter sample, 93 out of 200 guests choose to have the free breakfast. Assuming the conditions for inference are met, which of the following hypothesis tests would be best used to analyze the data?
- A) A one-sample z -test for a population proportion
 - B) A chi-square test of independence
 - C) A two-sample z -test for a difference between population proportions
 - D) None of the above