College Prep Stats
Extra Practice and Review

1. The author drilled a hole in a die and filled it with a lead weight, then proceeded to roll it 200 times. Here are the observed frequencies for the outcomes of 1, 2, 3, 4, 5, and 6, respectively: 27, 31, 42, 40, 28, 32. Use a 0.05 significance level to test the claim that the outcomes are not equally likely.

a) State the null hypothesis and the alternative hypothesis.

b) Calculate the test statistic.

c) Determine the *p*-value.

d) What is the conclusion?

2. Among the four northwestern states, Washington has 51% of the total population, Oregon has 30%, Idaho has 11%, and Montana has 8%. A market researcher selects a sample of 1000 subjects, with 450 in Washington, 340 in Oregon, 150 in Idaho, and 60 in Montana. At the 0.05 significance level, test the claim that the sample of 1000 subjects has a distribution that agrees with the distribution of state populations.

a) State the null hypothesis and the alternative hypothesis.

b) Calculate the test statistic.

c) Determine the *p*-value.

d) What is the conclusion?

3. Researchers investigated the issue of race and equality of access to clinical trials. The table below shows the population distribution and the numbers of participants in clinical trials involving lung cancer (based on data from "Participation in Cancer Clinical Trials," by Murthy, Krumholz, and Gross, Journal of the American Medical Association, Vol. 291, No. 22). Use a 0.01 significance level to test the claim that the distribution of clinical trial participants fits well with the population distribution.

Race/Ethnicity	White Non-Hispanic	Hispanic	Black	Asian/Pacific Islander	American Indian/Alaskan Native
Distribution of Population	75.6%	9.1%	10.8%	3.8%	0.7%
Number in Lung Cancer Clinical Trials	3855	60	316	54	12

a) State the null hypothesis and the alternative hypothesis.

b) Calculate the test statistic.

c) Determine the *p*-value.

d) What is the conclusion?

4. Randomly selected nonfatal occupational injuries and illnesses are categorized according to the day of the week that they first occurred, and the results are listed below (based on data from the Bureau of Labor Statistics). Use a 0.05 significance level to test the claim that such injuries and illnesses occur with equal frequency on the different days of the week.

Day	Mon	Tues	Weds	Thurs	Fri
Number	23	23	21	21	19

a) State the null hypothesis and the alternative hypothesis.

b) Calculate the test statistic.

c) Determine the *p*-value.

d) What is the conclusion?

5. Records of randomly selected births were obtained and categorized according to the day of the week that they occurred (based on data from the National Center for Health Statistics). Because babies are unfamiliar with our schedule of weekdays, a reasonable claim is that births occur on the different days with equal frequency. Use a 0.01 significance level to test that claim.

Day	Sun	Mon	Tues	Weds	Thurs	Fri	Sat
Number of births	77	110	124	122	120	123	97

a) State the null hypothesis and the alternative hypothesis.

b) Calculate the test statistic.

c) Determine the *p*-value.

d) What is the conclusion?

6. Mars, Inc. claims that its M&M plain candies are distributed with the following color percentages: 16% green, 20% orange, 14% yellow, 24% blue, 13% red, and 13% brown. Refer to the table below and use the sample data to test the claim that the color distribution is as claimed by Mars, Inc. Use a 0.05 significance level.

a) State the null hypothesis and the alternative hypothesis.

b) Calculate the test statistic.

Color	Observed Frequency *O*
Green	19
Orange	25
Yellow	8
Blue	27
Red	13
Brown	8

c) Determine the *p*-value.

d) What is the conclusion?

According to Benford's law, a variety of different data sets include numbers with leading (first) digits that follow the distribution shown in the table below. In Numbers 2 & 3, test for goodness-of-fit with Benford's law.

Leading digit	1	2	3	4	5	6	7	8	9
Benford's Law: distribution of leading digits	30.1%	17.6%	12.5%	9.7%	7.9%	6.7%	5.8%	5.1%	4.6%

7. When working for the Brooklyn District Attorney, investigator Robert Burton analyzed the leading digits of the amounts from 784 checks issued by seven suspect companies. The frequencies were found to be 0, 15, 0, 76, 479, 183, 8, 23, and 0, and those digits correspond to the leading digits of 1, 2, 3, 4, 5, 6, 7, 8, and 9, respectively. If the observed frequencies are substantially different from the frequencies expected with Benford's law, the check amounts appear to result from fraud. Use a 0.01 significance level to test for goodness-of-fit with Benford's law.

a) State the null hypothesis and the alternative hypothesis.

b) Calculate the test statistic.

c) Determine the *p*-value.

d) What is the conclusion?

8. Amounts of recent political contributions are randomly selected, and the leading digits are found to have frequencies of 52, 40, 23, 20, 21, 9, 8, 9, and 30. (Those observed frequencies correspond to the leading digits of 1, 2, 3, 4, 5, 6, 7, 8, and 9, respectively, and they are based on data from "Breaking the (Benford) Law: Statistical Fraud Detection in Campaign Finance," by Cho and Gaines, American Statistician, Vol. 61, No. 3.) Using a 0.01 significance level, test the observed frequencies for goodness-of-fit with Benford's law.

a) State the null hypothesis and the alternative hypothesis.

b) Calculate the test statistic.

c) Determine the *p*-value.

d) What is the conclusion?