



## **Practice Exercises: Lesson 4.2**

Diez, D. M., Çetinkaya-Rundel, M., Barr, C. D. (2019). OpenIntro Statistics (4th ed.). OpenIntro.  
<https://www.openintro.org/book/os/> CC BY-SA 3.0

STAT 1201  
Introduction to Probability and Statistics

ONLINE AND DISTANCE EDUCATION

## Exercises

**6.31 True or false, Part I.** Determine if the statements below are true or false. For each false statement, suggest an alternative wording to make it a true statement.

- The chi-square distribution, just like the normal distribution, has two parameters, mean and standard deviation.
- The chi-square distribution is always right skewed, regardless of the value of the degrees of freedom parameter.
- The chi-square statistic is always positive.
- As the degrees of freedom increases, the shape of the chi-square distribution becomes more skewed.

**6.32 True or false, Part II.** Determine if the statements below are true or false. For each false statement, suggest an alternative wording to make it a true statement.

- As the degrees of freedom increases, the mean of the chi-square distribution increases.
- If you found  $\chi^2 = 10$  with  $df = 5$  you would fail to reject  $H_0$  at the 5% significance level.
- When finding the p-value of a chi-square test, we always shade the tail areas in both tails.
- As the degrees of freedom increases, the variability of the chi-square distribution decreases.

**6.33 Open source textbook.** A professor using an open source introductory statistics book predicts that 60% of the students will purchase a hard copy of the book, 25% will print it out from the web, and 15% will read it online. At the end of the semester he asks his students to complete a survey where they indicate what format of the book they used. Of the 126 students, 71 said they bought a hard copy of the book, 30 said they printed it out from the web, and 25 said they read it online.

- State the hypotheses for testing if the professor's predictions were inaccurate.
- How many students did the professor expect to buy the book, print the book, and read the book exclusively online?
- This is an appropriate setting for a chi-square test. List the conditions required for a test and verify they are satisfied.
- Calculate the chi-squared statistic, the degrees of freedom associated with it, and the p-value.
- Based on the p-value calculated in part (d), what is the conclusion of the hypothesis test? Interpret your conclusion in this context.

**6.34 Barking deer.** Microhabitat factors associated with forage and bed sites of barking deer in Hainan Island, China were examined. In this region woods make up 4.8% of the land, cultivated grass plot makes up 14.7%, and deciduous forests make up 39.6%. Of the 426 sites where the deer forage, 4 were categorized as woods, 16 as cultivated grassplot, and 61 as deciduous forests. The table below summarizes these data.<sup>39</sup>

Woods	Cultivated grassplot	Deciduous forests	Other	Total
4	16	61	345	426

- Write the hypotheses for testing if barking deer prefer to forage in certain habitats over others.
- What type of test can we use to answer this research question?
- Check if the assumptions and conditions required for this test are satisfied.
- Do these data provide convincing evidence that barking deer prefer to forage in certain habitats over others? Conduct an appropriate hypothesis test to answer this research question.

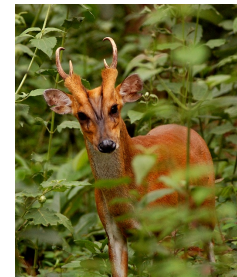


Photo by Shrikant Rao  
(<http://flic.kr/p/4Xjdkk>)  
CC BY 2.0 license

<sup>39</sup>Liwei Teng et al. "Forage and bed sites characteristics of Indian muntjac (*Muntiacus muntjak*) in Hainan Island, China". In: *Ecological Research* 19.6 (2004), pp. 675–681.

**6.43 College smokers.** We are interested in estimating the proportion of students at a university who smoke. Out of a random sample of 200 students from this university, 40 students smoke.

- Calculate a 95% confidence interval for the proportion of students at this university who smoke, and interpret this interval in context. (Reminder: Check conditions.)
- If we wanted the margin of error to be no larger than 2% at a 95% confidence level for the proportion of students who smoke, how big of a sample would we need?

**6.44 Acetaminophen and liver damage.** It is believed that large doses of acetaminophen (the active ingredient in over the counter pain relievers like Tylenol) may cause damage to the liver. A researcher wants to conduct a study to estimate the proportion of acetaminophen users who have liver damage. For participating in this study, he will pay each subject \$20 and provide a free medical consultation if the patient has liver damage.

- If he wants to limit the margin of error of his 98% confidence interval to 2%, what is the minimum amount of money he needs to set aside to pay his subjects?
- The amount you calculated in part (a) is substantially over his budget so he decides to use fewer subjects. How will this affect the width of his confidence interval?

**6.45 Life after college.** We are interested in estimating the proportion of graduates at a mid-sized university who found a job within one year of completing their undergraduate degree. Suppose we conduct a survey and find out that 348 of the 400 randomly sampled graduates found jobs. The graduating class under consideration included over 4500 students.

- Describe the population parameter of interest. What is the value of the point estimate of this parameter?
- Check if the conditions for constructing a confidence interval based on these data are met.
- Calculate a 95% confidence interval for the proportion of graduates who found a job within one year of completing their undergraduate degree at this university, and interpret it in the context of the data.
- What does “95% confidence” mean?
- Now calculate a 99% confidence interval for the same parameter and interpret it in the context of the data.
- Compare the widths of the 95% and 99% confidence intervals. Which one is wider? Explain.

**6.46 Diabetes and unemployment.** A Gallup poll surveyed Americans about their employment status and whether or not they have diabetes. The survey results indicate that 1.5% of the 47,774 employed (full or part time) and 2.5% of the 5,855 unemployed 18-29 year olds have diabetes.<sup>48</sup>

- Create a two-way table presenting the results of this study.
- State appropriate hypotheses to test for difference in proportions of diabetes between employed and unemployed Americans.
- The sample difference is about 1%. If we completed the hypothesis test, we would find that the p-value is very small (about 0), meaning the difference is statistically significant. Use this result to explain the difference between statistically significant and practically significant findings.

**6.47 Rock-paper-scissors.** Rock-paper-scissors is a hand game played by two or more people where players choose to sign either rock, paper, or scissors with their hands. For your statistics class project, you want to evaluate whether players choose between these three options randomly, or if certain options are favored above others. You ask two friends to play rock-paper-scissors and count the times each option is played. The following table summarizes the data:

Rock	Paper	Scissors
43	21	35

Use these data to evaluate whether players choose between these three options randomly, or if certain options are favored above others. Make sure to clearly outline each step of your analysis, and interpret your results in context of the data and the research question.

---

<sup>48</sup>Gallup Wellbeing, Employed Americans in Better Health Than the Unemployed, data collected Jan. 2, 2011 - May 21, 2012.