Introduction to Probability and Statistics

Session 3 Exercises

Israel Leyva-Mayorga

Exercise 1: Adapted from Chapter 8, problem 25

It is claimed that a certain type of bipolar transistor has a mean value of current gain that is at least 210. A sample of these transistors is tested. If the sample mean value of the current gain is 200 and it is known that the standard deviation is 35, would the claim be rejected at the 5 percent level of significance if

a) The sample size is 25?

b) The sample size is 64?

Tips: First, formulate the null hypothesis H_0. Then, identify the type of test:

- Is it one-side or two-side test?
- What is the test statistic T that you should use?
- Will you compare T to the threshold c or the p-value to α?

Try to interpret the results.

Exercise 2: Chapter 8, problem 56

According to the U.S. Bureau of Census, 25.5 percent of the population of those age 18 or over smoked in 1990. A scientist has recently claimed that this percentage has since increased, and to prove her claim she randomly sampled 500 individuals from this population. If 138 of them were smokers, is her claim proved? Use the 5 percent level of significance.

Tips: Which type of distribution do the RVs have when there are only two options for the outcomes? What is the type of distribution that results from the sum of several of these RVs?

Several of the exercises were taken from the book "Introduction to Probability and Statistics for Engineers and Scientists" by Sheldon M. Ross, 3rd Ed.

Exercise 3: The students were asked to choose between the following two options for the rest of the course using an online poll.

1) Israel gives the rest of the theory this Friday morning. Students do the hypothesis testing exercises on their own. We do the workshop exercises on April 12.
2) Israel gives half of the theory this Friday and the rest on April 12. We do the exercises for hypothesis testing and the workshop mixed with the lectures.

Naturally, not all the students responded to the poll. As Fig. 1 with a total of $n = 5$ responses, 3 of them selecting option 2).

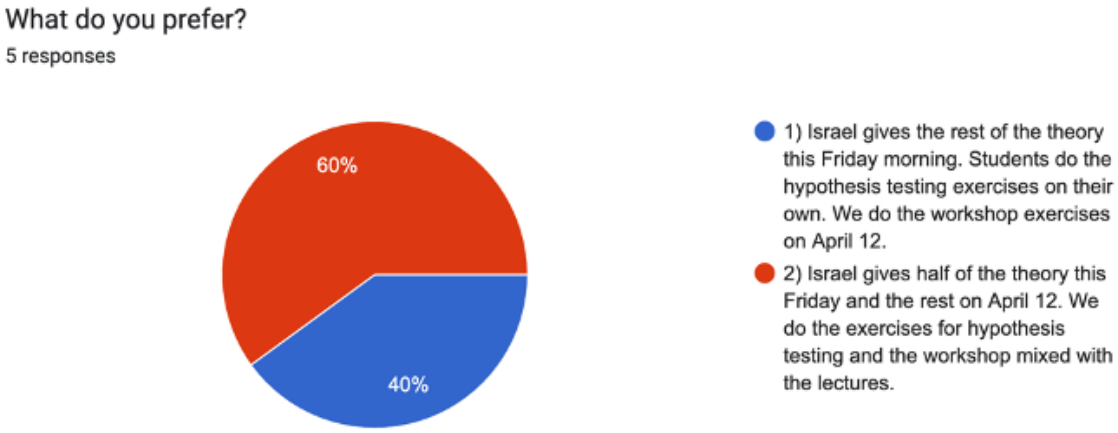

Fig. 1. Results of the voting for the next sessions of statistics.

To make a decision that is fair, Israel decided to conduct a test. First, he defined $X_i \sim \text{Bernoulli}(p)$, where $X_i = 0$ if a student selects 1) and $X_i = 1$ the students selects 2). With this, he set $p_0 = 0.5$, defined the null hypothesis H_0: "the students prefer option 2)", which can be formulated as $p \geq p_0$. Then, he estimated

$$\hat{p}_5 = \frac{1}{5}\sum_{i=1}^{n} X_i = 0.6$$

a) Perform a one-sided test with $\alpha = 0.05$ using the null hypothesis H_0 and calculating the p-value for the Binomial distribution $v = P(X \leq k; n, p_0)$, where $k = 3$ is the number of students that chose option 2). Can you reject Israel's hypothesis?

b) Change the null hypothesis to be H_0': "the students prefer option 1)". What is the mathematical formulation of this? Can you reject this hypothesis with the collected evidence?

c) Assume that the result of the poll is now $k = 1$, meaning that only one student preferred option 2). Can the original null hypothesis H_0 be rejected?

Exercise 4: Repeat **Exercise 1** but now assuming that the real standard deviation is not known and the estimated one from the sample is 35.

Exercise 5: Chapter 8, problem 43

A question of medical importance is whether jogging leads to a reduction in one's heart rate. To test this hypothesis, 8 nonjogging volunteers agreed to begin a 1-month jogging program. After the month, their pulse rates were determined and compared with their earlier values. If the data are as follows, can we conclude that jogging has an effect on the pulse rates with a 5 percent level of significance?

Subject	1	2	3	4	5	6	7	8
Pulse rate before	74	86	98	102	78	84	79	70
Pulse rate after	70	85	90	110	71	80	69	74

Tips: Which type of test should you use? The paired t-test. How should you formulate the null-hypothesis so that it represents that jogging has no effect on pulse rates?

Exercise 6:

Two different types of cable insulation have been tested. We need to determine the voltage level at which failures occur. The tests revealed that the individual cables failed at the following voltage levels.

We know that the voltage level that cables of type A can withstand is normally distributed with unknown mean μ_A and known variance $\sigma_A^2 = 40$. We also know that the voltage level that cables of type B can withstand is normally distributed with unknown mean μ_B and known variance $\sigma_B^2 = 100$.

With these data, can you reject the hypothesis $\mu_A = \mu_B$ with

TABLE I

VOLTAGE LEVEL AT WHICH FAILURES OCCURRED

Type A cables		Type B cables	
36	54	52	60
44	52	64	44
41	37	38	48
53	51	68	46
38	44	66	70
36	35	52	62
34	44		

a) 5 percent level of significance?

b) 10 percent level of significance?

c) 1 percent level of significance?

Tips: Which type of test should you use? What happens to the result of a test when we increase or decrease the level of significance?

Exercise 7: Chapter 9, problem 12

The following data set presents the heights of 12 male law school classmates whose law school examination scores were roughly equal. It also gives their annual salaries 5 years after graduation. Each of them went into corporate law. The height is in inches and the salary in units of $1,000.

Height	Salary
64	91
65	94
66	88
67	103
69	77
70	96
72	105
72	88
74	122
74	102
75	90
76	114

a) Do the above data establish the hypothesis that a lawyer's salary is related to his height? Use the 5 percent level of significance.

b) What was the null hypothesis in part a)