

Probability and Statistics: Practice Set 7

Paul Beggs

December 2, 2025

1. (2 points) Suppose a researcher wishes to see if students who sleep less than 8 hours before an exam have a mean score of 80 or better on the exam. Their null hypothesis is that the mean score is less than or equal to 80. They use $\alpha = 0.05$ and sample $n = 36$ students and determine $p = 0.03$ and so reject H_0 . If a later survey of all student exam scores shows that in fact, students who sleep less than 8 hours have a mean exam score of 78.2, what kind of error has the initial researcher made?

Solution. The researcher made a Type I Error.

2. (2 points) A researcher wants to test an alternative hypothesis that eating a fun sized bag of M&Ms before an exam improve test scores. They make a control group of size $n = 23$ and treatment group of size $n = 26$ and $\alpha = 0.05$. They run their t -test, but find $p = 0.0861$. They then run another version of the same experiment, with sample sizes $n = 28$ for the control and $n = 25$ for the treatment, but find $p = 0.0912$. They repeat this again, using various similar sample sizes and find $p = 0.5023$ and then $p = 0.0781$ and finally $p = 0.0234$. They reason that they have found a true result, since it only took 5 attempts, and at $\alpha = 0.05$, they should only encounter spurious results about 1 in 20 times. Briefly, critique their process.

Solution. This researcher has engaged in p -hacking. Wherein, instead of accepting the first valid p -value, they kept trying to reach the $p < \alpha$ threshold for their data. This leads to an increase in Type I/II errors. Also, their reasoning for the spurious results is also flawed: they have not found a “true result” because it only took 5 attempts. Instead, they kept rolling the dice until they encountered a Type I Error.

3. (2 points) You flip a coin 300 times and get 128 heads. Use the calculator to build a 99% confidence interval for the proportion of heads this coin produces. What can you say about the fairness of the coin?

Solution. Using 1-PropZInt with $x = 128$ and $n = 300$, we get $(0.353, 0.500)$ as our 99% confidence interval. Since the value 0.5 is included within the 99% CI, we cannot conclude that the coin is unfair at $\alpha = 0.1$.

4. (3 points) Week-old ducklings have a mean weight of 423 g. You are concerned about a new industrial plant upstream from your favorite pond, and so weigh a random sample of $n = 43$ week-old ducklings. You find $\bar{x} = 413.7$ g and $s = 33.9$. At $\alpha = 0.05$, is there evidence that the plant is causing the ducks to be underweight? Use the NHST steps.

Solution. Our population consists of week-old ducklings. Our test is one-tailed because we think the ducklings are less in weight. We hypothesize that:

- $H_0: \mu \geq 423$ – The average duckling weighs at least 423 grams.
- $H_1: \mu < 423$ – The average duckling weighs less than 423 grams.

For our sample, we will use a 1-sample t -test. This yields $t_{\text{samp}} = -1.799$, $p = 0.04$. Since $p < \alpha$, we **reject** the null hypothesis. We find the 90% confidence interval to be $(405, 422.4)$. Our effect size is calculated like so:

$$d = \frac{\bar{x} - \mu_0}{s} = \frac{413.7 - 423}{33.9} = -0.274.$$

Therefore, our effect is small. These data lead us to conclude that the industrial plant upstream has a negative small effect upon week-old ducklings.

5. (3 points) Prof. Seme gives the same 5 NHST questions to his MATH 215 and MATH 310 students, and assigns each student a grade. We will treat each as a random sample of all Intro Stats and Math-Major-Adjacent Stats students. Using the data in the table below, is there evidence, at $\alpha = 0.01$, that there is a difference in performance between the two groups? Use the NHST steps.

	MATH 215	MATH 310
Mean Score	74.7	81.2
Sample St. Dev.	12.3	9.1
Count	35	23

Solution. We have two populations: one consisting of all students who belong to Intro Stats (MATH 215), and the other being all students in Math-Major-Adjacent Stats (MATH 310). Our test is two-tailed because we are investigating if there is any difference at all. Thus, we posit the following hypotheses:

- $H_0: \mu_{215} = \mu_{310}$ – There is no difference in-between MATH 215 and MATH 310.
- $H_1: \mu_{215} \neq \mu_{310}$ – There is a difference in-between MATH 215 and MATH 310.

We will use a 2-sample t -test where we do not assume that the two populations have the same variance (not pooled), as we are working with two separate populations. We find $t_{\text{samp}} = -2.309$, $p = 0.02$. Since $p > \alpha$, we **fail to reject** H_0 , we do not have sufficient evidence that there is a difference in scores between MATH 215 and MATH 310. We find the 99% confidence interval to be $(-14.01, 1.01)$, which we note includes 0. Now, we find s_p and the effect size:

$$s_p = \sqrt{\frac{12.3^2 + 9.1^2}{2}} = 10.82, \quad \implies \quad d = \frac{\bar{x}_1 - \bar{x}_2}{s_p} = -0.601.$$

While the sample data shows a medium effect size, we do not have enough evidence to support that there is a difference in the populations.

6. (3 points) We are also interested in whether the variances of the two classes (from the previous problem) are different. So, to repeat, Prof. Seme gives the same 5 NHST questions to his MATH 215 and MATH 310 students, and assigns each student a grade. We will treat each as a random sample of all Intro Stats and Math-Major-Adjacent Stats students. Using the data in the table below, is there evidence, at $\alpha = 0.05$, that there is a difference the variance in performance between the two groups? Use the NHST steps.

	MATH 215	MATH 310
Mean Score	74.7	81.2
Sample St. Dev.	12.3	9.1
Count	35	23

Solution. The population consists of all students who belong to Intro Stats (MATH 215) and all Math-Major-Adjacent Stats (MATH 310). Our test is two-sided as we are only looking for a *difference* in variances. Thus, we have hypotheses:

- $H_0: \sigma_{215}^2 = \sigma_{310}^2$ – The two variances are the same.
- $H_1: \sigma_{215}^2 \neq \sigma_{310}^2$ – The two variances are different.

We will use a 2-sample F -test. We find that $F_{\text{samp}} = 1.827$, $p = 0.14$. Since $p > \alpha$, we **fail to reject** H_0 .

We do not have evidence to suggest that the variances are different.

7. (3 points) A professor wonders if students are more likely to be absent on Friday in the 3-day-a-week class than Monday or Wednesday (i.e. not exactly equal across all days). The professor uses a single course one semester as though it is a random sample, and counts that of 100 total absences, 41 were on a Friday. Is there evidence, at $\alpha = 0.05$ to support the professor's suspicion about Fridays? Use the NHST steps.

Solution. The population consists of all absences in the semester. We are examining the proportion of absences that occur on Fridays (p). If absences are distributed equally across the 3 class days, we expect $p = 1/3$.

Because the professor suspects students are *more* likely to be absent on Fridays, this is a one-tailed test. Thus, our hypotheses are:

- $H_0: p = 1/3$ – Friday absences occur at the expected rate.
- $H_a: p > 1/3$ – Friday absences occur more frequently than random chance.

We use a 1-proportion Z -test. We find that $z_{\text{samp}} = 1.63$, $p = 0.052$. Since $p > \alpha$ we **fail to reject** H_0 . We find the 90% confidence interval to be $(0.329, 0.491)$, which we note contains $\frac{1}{3}$. Our effect size is calculated below:

$$d = 2 \arcsin \left(\sqrt{\hat{p}} \right) - 2 \arcsin \left(\sqrt{p_0} \right) = 2 \arcsin \left(\sqrt{41/100} \right) - 2 \arcsin \left(\sqrt{.33} \right) = 0.159.$$

Our effect is small.

We conclude that there is not sufficient evidence at $\alpha = 0.05$ to support the claim that absences are more likely on Fridays, though the result is borderline. Additionally, even if they were significant, our result would have a small effect.

8. (3 points) A random sample of Hendrix and Ozarks students shows that 25 of 48 Hendrix students have taken a statistics class in college and 17 of 50 Ozarks students have. Is there evidence, at $\alpha = 0.1$ that there is a difference in the proportion of students who take stats at the two schools? Use the NHST steps.

Solution. We have two populations here: all Hendrix students that have taken a statistics class in college, and all Ozarks Students who have also taken a statistics class. Since we are only interested in determining if there is a difference in the population, this is a two-tailed test. Our hypotheses are:

- $H_0: p_O = p_H$ – There is no difference between the proportion of Ozarks students that take statistics and that of Hendrix students.
- $H_1: p_O \neq p_H$ – There is a difference between the proportion of Ozarks students that take statistics and that of Hendrix students.

We will use a two sample Z -test for proportion. We find that $z_{\text{samp}} = 1.81$, $p = 0.07$. Since $p > \alpha$, we **reject** H_0 . Since our sample was two-sided and $\alpha = 0.1$, we find the 90% confidence interval to be $(0.0189, 0.342)$, which we note does not include 0. We find the effect size to be

$$d = 2 \arcsin \left(\sqrt{\hat{p}_1} \right) - 2 \arcsin \left(\sqrt{\hat{p}_2} \right) = 2 \arcsin \left(\sqrt{25/48} \right) - 2 \arcsin \left(\sqrt{17/50} \right) = 0.367.$$

Thus, our effect is small-medium. Overall, we support a claim that there is a difference between the proportion of students who take stats at Ozarks and Hendrix. In addition, our effect size tells us that our result is small to medium.

9. (3 points) It is known that, over the past 10 years, 50% of Hendrix students come from AR, 29% from TX, 10% from OK, 5% from MO, 3% from TN, and 3% elsewhere. A survey of 120 students who have recently taken Calculus I shows that 54 are from AR, 23 from TX, 17 from OK, 8 from MO, 8 from TN, and 10 elsewhere. Is there evidence, at $\alpha = 0.05$, that the distribution of students who take Calculus I is different from the overall Hendrix population?

Solution. We will use the χ^2 -GOF-Test. From the given data, we have the following table:

	AR	TX	OK	MO	TN	Else.
Observed	54	23	17	8	8	10
Expected	60	34.8	12	6	3.6	3.6

By typing these data into my TI-84 for L_1 and L_2 , and then using **X2GOF-Test**, I got $\chi^2_{\text{samp}} = 24.107$, $p = 0$. Since $p < \alpha$, we **reject** H_0 and conclude that there is a difference between the students who take Calculus I and the overall Hendrix population.

10. (3 points) You are interested in whether there is a relationship between a student's major area (Humanities, Nat Sci, Soc Sci, or Interdisciplinary) and the amount of sleep they get. You survey students and find the data in the table below. What can you say, at $\alpha = 0.05$? Use the NHST steps.

Sleep	Human.	Nat. Sci.	Soc. Sci.	Inter.
8+ Hours	12	34	36	9
6.5 - 7.9 Hours	17	31	42	11
6.4- Hours	13	29	43	12

Solution. We have two categories: Sleep amount and major area. We want to see if there is some dependence between the two categories. Thus, we have the following hypotheses:

- H_0 : Sleep amount and major are independent.
- H_1 : Sleep amount and majors are dependent.

We will use a χ^2 -test of independence. We find that $\chi^2 = 2.055$, and $p = 0.915$. Since $p > \alpha$, we **fail to reject** H_0 . We lack evidence to support the claim that there is a relationship between the amount of sleep a student gets and their major.

11. (3 points) You want to know if Hendrix students have differences in their average number of M&Ms eaten per week, depending on their home state. You run an ANOVA, at $\alpha = 0.05$ and find the following table:

Groups	Count	Sum	Average	Variance
AR	34	6791.78	199.76	194.84
LA	29	6238.93	215.14	226.15
MO	25	4582.62	183.30	298.45
OK	25	4909.81	196.39	144.80
TX	25	4517.76	180.71	271.74

ANOVA

Source of Variation	SS	df	MS	F	P-value	F crit
Between Groups	20975.92	4	5243.98	23.31	1.2879E-14	2.44
Within Groups	29921.96	133	224.98			
Total	50897.89	137				

Solution. We are investigating the means of samples from 5 different states. Our hypotheses are

- $H_0: \mu_{AR} = \mu_{LA} = \mu_{MO} = \mu_{OK} = \mu_{TX}$ - The mean number of M&Ms eaten are the same among all groups.
- H_1 : There is at least one group that has a different weekly mean M&Ms eaten.

Using the given table, we see that $F_{\text{samp}} = 23.31$, $p = 1.29\text{E} -14$. Since $p < \alpha$, we **reject** H_0 . We find the effect size to be

$$\eta^2 = \frac{SS_{\text{between}}}{SS_{\text{total}}} = \frac{20975.92}{50897.89} = 0.412.$$

We have a very large effect size. Thus, we have found a strong relationship between home state and M&M consumption. It appears that the home state and candy eating habits are strongly related.