

# Bunker Hill Community College

Final Statistics Exam 2019-05-02

Exam ID 025

**Name:** \_\_\_\_\_

This take-home exam is due **Wednesday, May 8**, at the beginning of class.

You may use any notes, textbook, or online tools; however, you may not request help from any other human.

You will show your work on the pages with questions. When you are sure of your answers, you will **put those answers in the boxes** on the first few pages.

Unless you have an objection to doing so, please **copy the honor-code text below and sign**.

*I understand that outside help is NOT allowed on this exam. On my honor, the work herein is my own.*

---

---

---

**Signature:** \_\_\_\_\_

1. (a)  $P(\text{wheel and teal}) = 0.123$
- (b)  $P(\text{violet given pig}) = 0.366$
- (c)  $P(\text{pig given black}) = 0.149$
- (d)  $P(\text{wheel}) = 0.371$
- (e)  $P(\text{black}) = 0.179$
- (f)  $P(\text{wheel or teal}) = 0.576$
2.  $P(\text{"not ring" given "green"}) = 0.83$
3.  $P(65.97 < X < 66.22) = 0.7038$
4. (a)  $P(X = 33) = 0.1018$
- (b)  $P(29 \leq X \leq 40) = 0.8364$
5. **(15.2, 18.8)**
6. (a)  $H_0 : \mu_2 - \mu_1 = 0$
- (b)  $H_0 : \mu_2 - \mu_1 \neq 0$
- (c)  $t^* = 2.68$
- (d)  $SE = 22.027$
- (e)  $|t_{\text{obs}}| = 2.81$
- (f)  $0.01 < p\text{-value} < 0.02$
- (g) **reject**
7. (a) **LB of p CI = 0.294 or 29.4%**
- (b) **UB of p CI = 0.3 or 30%**

8. (a)  $H_0 : p_2 - p_1 = 0$

(b)  $H_A : p_2 - p_1 \neq 0$

(c)  $z^* = 1.96$

(d)  $SE = 0.066$

(e)  $|z_{\text{obs}}| = 2.02$

(f)  $p\text{-value} = 0.0434$

(g) **reject**

1. In a deck of strange cards, there are 375 cards. Each card has an image and a color. The amounts are shown in the table below.

	black	teal	violet
pig	10	35	26
shovel	32	42	91
wheel	25	46	68

- (a) What is the probability a random card is both a wheel and teal?
- (b) What is the probability a random card is violet given it is a pig?
- (c) What is the probability a random card is a pig given it is black?
- (d) What is the probability a random card is a wheel?
- (e) What is the probability a random card is black?
- (f) What is the probability a random card is either a wheel or teal (or both)?

**Solution**

$$(a) P(\text{wheel and teal}) = \frac{46}{375} = 0.123$$

$$(b) P(\text{violet given pig}) = \frac{26}{10+35+26} = 0.366$$

$$(c) P(\text{pig given black}) = \frac{10}{10+32+25} = 0.149$$

$$(d) P(\text{wheel}) = \frac{25+46+68}{375} = 0.371$$

$$(e) P(\text{black}) = \frac{10+32+25}{375} = 0.179$$

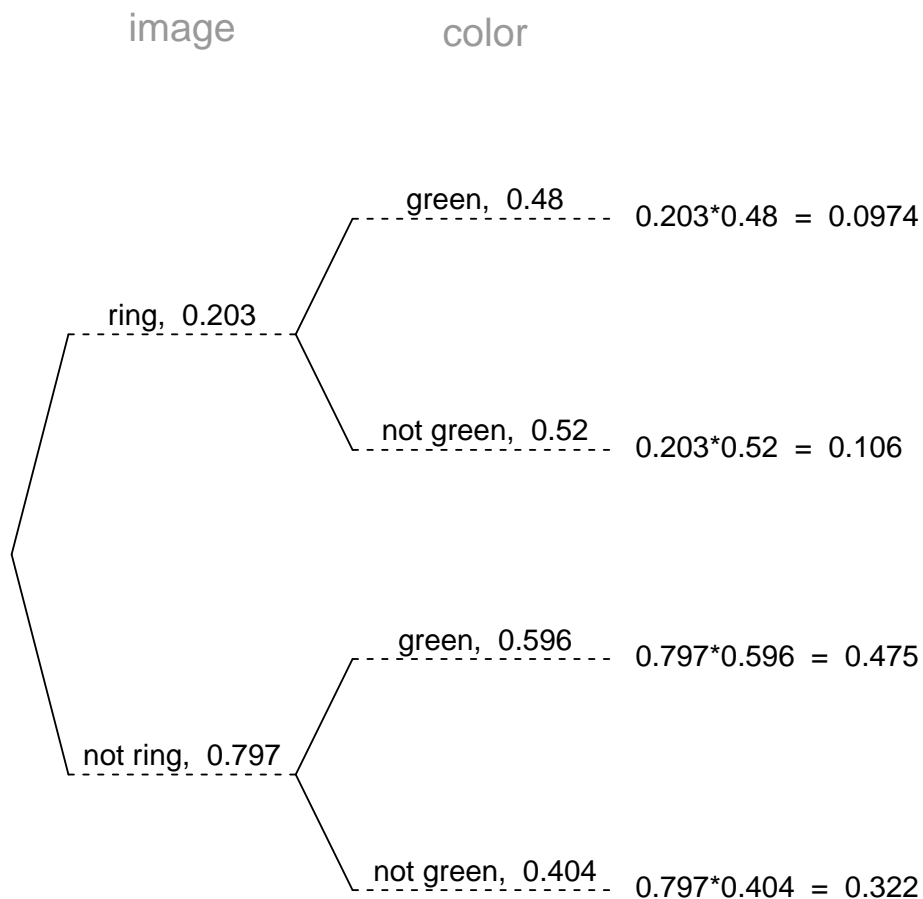
$$(f) P(\text{wheel or teal}) = \frac{25+46+68+35+42+46-46}{375} = 0.576$$

2. In a deck of strange cards, each card has an image and a color. The chance of drawing a ring is 20.3%. If a ring is drawn, there is a 48% chance that it is green. If a card that is not a ring is drawn, there is a 59.6% chance that it is green.

Now, someone draws a random card and reveals it is green. What is the chance the card is not a ring?

**Solution**

I'd recommend making a tree. Remember, on the first branch, we put simple probabilities. On the second branches we put conditional probabilities. The results (products) are joint probabilities.



Determine the appropriate conditional probability.

$$P(\text{"not ring" given "green"}) = \frac{0.475}{0.475 + 0.0974} = 0.83$$

3. In a very large pile of toothpicks, the mean length is 66.17 millimeters and the standard deviation is 1.17 millimeters. If you randomly sample 175 toothpicks, what is the chance the sample mean is between 65.97 and 66.22 millimeters?



**Solution**

Label the given information.

$$\mu = 66.17$$

$$\sigma = 1.17$$

$$n = 175$$

$$\bar{x}_{\text{lower}} = 65.97$$

$$\bar{x}_{\text{upper}} = 66.22$$

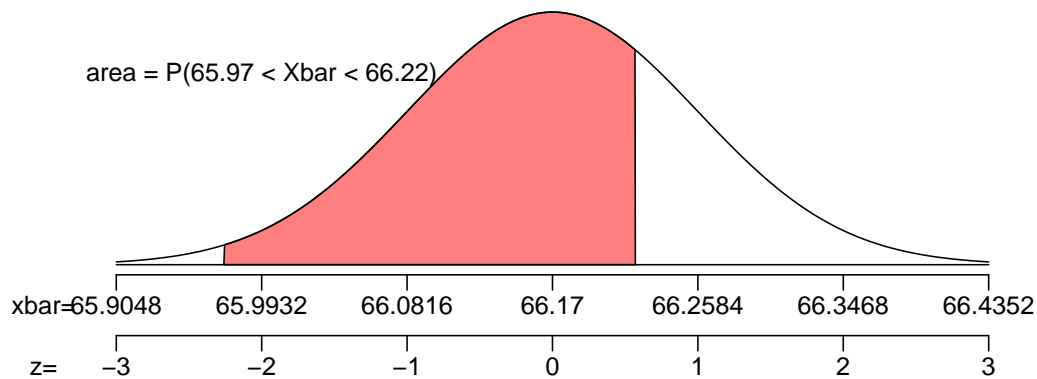
Find the standard error.

$$SE = \frac{\sigma}{\sqrt{n}} = \frac{1.17}{\sqrt{175}} = 0.0884$$

Describe the sampling distribution.

$$\bar{X} \sim \mathcal{N}(66.17, 0.0884)$$

Draw a sketch.



Calculate a z scores.

$$z_{\text{lower}} = \frac{x_{\text{lower}} - \mu}{SE} = \frac{65.97 - 66.17}{0.0884} = -2.26$$

$$z_{\text{upper}} = \frac{x_{\text{upper}} - \mu}{SE} = \frac{66.22 - 66.17}{0.0884} = 0.57$$

Determine the probability.

$$\begin{aligned} P(65.97 < X < 66.22) &= \Phi(z_{\text{upper}}) - \Phi(z_{\text{lower}}) \\ &= \Phi(0.57) - \Phi(-2.26) \\ &= 0.7038 \end{aligned}$$

4. In a game, there is a 56% chance to win a round. You will play 60 rounds.
- (a) What is the probability of winning exactly 33 rounds?
  - (b) What is the probability of winning at least 29 but at most 40 rounds?

**Solution**

We use the formula for binomial probabilities.

$$P(X = k) = \binom{n}{k} (p)^k (1 - p)^{n-k}$$

$$P(X = 33) = \binom{60}{33} (0.56)^{33} (1 - 0.56)^{60-33}$$

$$P(X = 33) = \binom{60}{33} (0.56)^{33} (0.44)^{27}$$

$$P(X = 33) = 0.1018$$

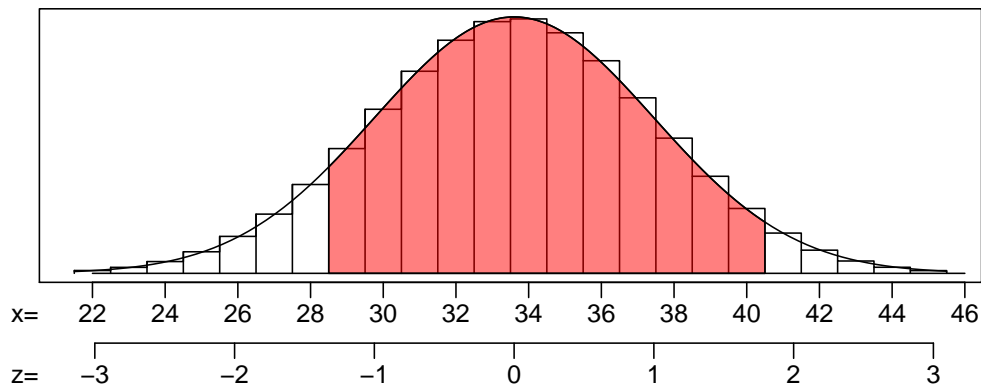
Find the mean.

$$\mu = np = (60)(0.56) = 33.6$$

Find the standard deviation.

$$\sigma = \sqrt{np(1 - p)} = \sqrt{(60)(0.56)(1 - 0.56)} = 3.845$$

Make a sketch, specifically try to picture whether you need to add or subtract 0.5 for the continuity correction.



Find the z scores.

$$z_1 = \frac{28.5 - 33.6}{3.845} = -1.2$$

$$z_2 = \frac{40.5 - 33.6}{3.845} = 1.66$$

Calculate the probability.

$$P(29 \leq X \leq 40) = \Phi(1.66) - \Phi(-1.2) = 0.8364$$

(a)  $P(X = 33) = 0.1018$

(b)  $P(29 \leq X \leq 40) = 0.8364$

5. As an ornithologist, you wish to determine the average body mass of *Oporornis formosus*. You randomly sample 13 adults of *Oporornis formosus*, resulting in a sample mean of 17 grams and a sample standard deviation of 1.92 grams. Determine a 99.5% confidence interval of the true population mean.

**Solution**

We are given the sample size, sample mean, sample standard deviation, and confidence level.

$$n = 13$$

$$\bar{x} = 17$$

$$s = 1.92$$

$$CL = 0.995$$

Determine the degrees of freedom (because we don't know  $\sigma$  and we are doing inference so we need to use the  $t$  distribution).

$$df = n - 1 = 12$$

Determine the critical  $t$  value,  $t^*$ , such that  $P(|T| < t^*) = 0.995$ .

$$t^* = 3.43$$

Calculate the standard error.

$$SE = \frac{s}{\sqrt{n}} = \frac{1.92}{\sqrt{13}} = 0.533$$

We want to make an inference about the population mean.

$$\mu \approx \bar{x} \pm t^* SE$$

Determine the bounds.

$$\begin{aligned} CI &= (\bar{x} - t^* SE, \bar{x} + t^* SE) \\ &= (17 - 3.43 \times 0.533, 17 + 3.43 \times 0.533) \\ &= (15.2, 18.8) \end{aligned}$$

We are 99.5% confident that the population mean is between 15.2 and 18.8.

6. A treatment group of size 9 has a mean of 1040 and standard deviation of 59.5. A control group of size 40 has a mean of 978 and standard deviation of 60.6. If you decided to use a significance level of 0.02, is there sufficient evidence to conclude the treatment causes an effect?

By using the Welch-Satterthwaite equation, I've calculated the degrees of freedom should be 12.

- (a) State the null hypothesis.
- (b) State the alternative hypothesis.
- (c) Evaluate the critical value. (The critical value is either  $z^*$  or  $t^*$ . Determine its value.)
- (d) Determine the standard error of the relevant sampling distribution.
- (e) Evaluate the absolute value of the test statistic. (The test statistic is either  $z_{\text{obs}}$  or  $t_{\text{obs}}$ . Determine its absolute value.)
- (f) If possible, evaluate the  $p$ -value. Otherwise, describe an interval containing the  $p$ -value.
- (g) Do we reject or retain the null?

**Solution**

We are given unpaired data. We are considering a difference of means. Label the given information.

$$\begin{aligned}n_1 &= 9 \\ \bar{x}_1 &= 1040 \\ s_1 &= 59.5 \\ n_2 &= 40 \\ \bar{x}_2 &= 978 \\ s_2 &= 60.6 \\ \alpha &= 0.02 \\ df &= 12\end{aligned}$$

State the hypotheses.

$$\begin{aligned}H_0 : \mu_2 - \mu_1 &= 0 \\ H_A : \mu_2 - \mu_1 &\neq 0\end{aligned}$$

We are using a two-tail test. Find  $t^*$  such that  $P(|T| > t^*) = 0.02$  by using a  $t$  table.

$$t^* = 2.68$$

Calculate the standard error.

$$\begin{aligned}SE &= \sqrt{\frac{(s_1)^2}{n_1} + \frac{(s_2)^2}{n_2}} \\ &= \sqrt{\frac{(59.5)^2}{9} + \frac{(60.6)^2}{40}} \\ &= 22.027\end{aligned}$$

Determine the test statistic.

$$\begin{aligned}t_{\text{obs}} &= \frac{(\bar{x}_2 - \bar{x}_1) - (\mu_2 - \mu_1)_0}{SE} \\ &= \frac{(978 - 1040) - (0)}{22.027} \\ &= -2.81\end{aligned}$$

Compare  $|t_{\text{obs}}|$  and  $t^*$ .

$$|t_{\text{obs}}| > t^*$$

We can determine an interval for the  $p$ -value using the  $t$  table.

$$0.01 < p\text{-value} < 0.02$$

Compare  $p$ -value and  $\alpha$ .

$$p\text{-value} < \alpha$$

We conclude that we should reject the null hypothesis.

$$(a) H_0 : \mu_2 - \mu_1 = 0$$

(b)  $H_A : \mu_2 - \mu_1 \neq 0$

(c)  $t^* = 2.68$

(d)  $SE = 22.027$

(e)  $|t_{\text{obs}}| = 2.81$

(f)  $0.01 < p\text{-value} < 0.02$

(g) reject the null



7. From a very large population, a random sample of 76000 individuals was taken. In that sample, 29.7% were cold. Determine a 96% confidence interval of the population proportion.
- (a) Find the lower bound of the confidence interval.
  - (b) Find the upper bound of the confidence interval.

**Solution**

Determine  $z^*$  such that  $P(|Z| < z^*) = 0.96$ .

$$z^* = 2.05$$

Calculate the standard error.

$$SE = \sqrt{\frac{\hat{p}(1 - \hat{p})}{n}} = \sqrt{\frac{(0.297)(1 - 0.297)}{76000}} = 0.00166$$

Calculate the margin of error.

$$ME = z^* SE = (2.05)(0.00166) = 0.0034$$

To find the confidence interval's bounds, find the sample proportion plus or minus the margin of error.

$$p \approx \hat{p} \pm ME$$

Determine the interval.

$$(0.294, 0.3)$$

We are 96% confident that the true population proportion is between 29.4% and 30%.

- (a) The lower bound = 0.294, which can also be expressed as 29.4%.
- (b) The upper bound = 0.3, which can also be expressed as 30%.

8. An experiment is run with a treatment group of size 78 and a control group of size 96. The results are summarized in the table below.

	treatment	control
glossy	25	18
not glossy	53	78

Using a significance level of 0.05, determine whether the treatment causes an effect on the proportion of cases that are glossy.

- (a) State the null hypothesis.
- (b) State the alternative hypothesis.
- (c) Evaluate the critical value. (The critical value is either  $z^*$  or  $t^*$ . Determine its value.)
- (d) Determine the standard error of the relevant sampling distribution.
- (e) Evaluate the absolute value of the test statistic. (The test statistic is either  $z_{\text{obs}}$  or  $t_{\text{obs}}$ . Determine its absolute value.)
- (f) If possible, evaluate the  $p$ -value. Otherwise, describe an interval containing the  $p$ -value.
- (g) Do we reject or retain the null?

**Solution**

State the hypotheses.

$$H_0 : p_2 - p_1 = 0$$

$$H_A : p_2 - p_1 \neq 0$$

Find  $z^*$  such that  $P(|Z| > z^*) = 0.05$ .

$$z^* = \Phi^{-1} \left( 1 - \frac{\alpha}{2} \right) = 1.96$$

Determine the sample proportions.

$$\hat{p}_1 = \frac{25}{78} = 0.321$$

$$\hat{p}_2 = \frac{18}{96} = 0.188$$

Determine the difference of sample proportions.

$$\hat{p}_2 - \hat{p}_1 = 0.188 - 0.321 = -0.133$$

Determine the pooled proportion (because the null assumes the population proportions are equal).

$$\hat{p} = \frac{25 + 18}{78 + 96} = 0.247$$

Determine the standard error.

$$\begin{aligned} SE &= \sqrt{\frac{\hat{p}(1 - \hat{p})}{n_1} + \frac{\hat{p}(1 - \hat{p})}{n_2}} \\ &= \sqrt{\frac{(0.247)(0.753)}{78} + \frac{(0.247)(0.753)}{96}} \\ &= 0.0657 \end{aligned}$$

We can be more specific about what the null hypothesis claims.

$$H_0 : \hat{P}_2 - \hat{P}_1 \sim \mathcal{N}(0, 0.0657)$$

We want to describe how unusual our observation is under the null by finding the  $p$ -value. To do so, first find the  $z$  score.

$$\begin{aligned} z &= \frac{(\hat{p}_2 - \hat{p}_1) - (p_2 - p_1)_0}{SE} \\ &= \frac{(0.188 - 0.321) - 0}{0.0657} \\ &= -2.02 \end{aligned}$$

Determine the  $p$ -value.

$$\begin{aligned} p\text{-value} &= 2 \cdot \Phi(-|z|) \\ &= 2 \cdot \Phi(-2.02) \\ &= 0.0434 \end{aligned}$$

Compare the  $p$ -value to the significance level.

$$p\text{-value} < \alpha$$

So, we reject the null hypothesis. Thus the difference in proportions is significant.

(a)  $H_0 : p_2 - p_1 = 0$

(b)  $H_A : p_2 - p_1 \neq 0$

(c)  $z^* = 1.96$

(d)  $SE = 0.0657$

(e)  $|z_{\text{obs}}| = 2.02$

(f)  $p\text{-value} = 0.0434$

(g) reject the null