

# Bunker Hill Community College

Final Statistics Exam 2019-05-02

Exam ID 009

**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)  $P(\text{gray given tree}) = 0.369$
- (b)  $P(\text{pink}) = 0.406$
- (c)  $P(\text{bike}) = 0.234$
- (d)  $P(\text{wheel or yellow}) = 0.363$
- (e)  $P(\text{flower and gray}) = 0.0612$
- (f)  $P(\text{wheel given pink}) = 0.275$
2.  $P(\text{"not bike" given "red"}) = 0.754$
3.  $P(72.72 < X < 73.31) = 0.7656$
4. (a)  $P(X = 37) = 0.0449$
- (b)  $P(31 \leq X \leq 42) = 0.5809$
5. **(59.1, 76.7)**
6. (a)  $H_0 : \mu_2 - \mu_1 = 0$
- (b)  $H_0 : \mu_2 - \mu_1 \neq 0$
- (c)  $t^* = 2.09$
- (d)  $SE = 0.576$
- (e)  $|t_{\text{obs}}| = 2.34$
- (f)  $0.02 < p\text{-value} < 0.04$
- (g) **reject**
7. (a) **LB of p CI = 0.618 or 61.8%**
- (b) **UB of p CI = 0.656 or 65.6%**

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

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

(c)  $z^* = 2.33$

(d)  $SE = 0.048$

(e)  $|z_{\text{obs}}| = 2.59$

(f)  $p\text{-value} = 0.0096$

(g) **reject**

1. In a deck of strange cards, there are 849 cards. Each card has an image and a color. The amounts are shown in the table below.

	gray	pink	yellow
bike	89	98	12
flower	52	67	28
pig	85	32	45
tree	65	53	58
wheel	47	95	23

- (a) What is the probability a random card is gray given it is a tree?
- (b) What is the probability a random card is pink?
- (c) What is the probability a random card is a bike?
- (d) What is the probability a random card is either a wheel or yellow (or both)?
- (e) What is the probability a random card is both a flower and gray?
- (f) What is the probability a random card is a wheel given it is pink?

**Solution**

$$(a) P(\text{gray given tree}) = \frac{65}{65+53+58} = 0.369$$

$$(b) P(\text{pink}) = \frac{98+67+32+53+95}{849} = 0.406$$

$$(c) P(\text{bike}) = \frac{89+98+12}{849} = 0.234$$

$$(d) P(\text{wheel or yellow}) = \frac{47+95+23+12+28+45+58+23-23}{849} = 0.363$$

$$(e) P(\text{flower and gray}) = \frac{52}{849} = 0.0612$$

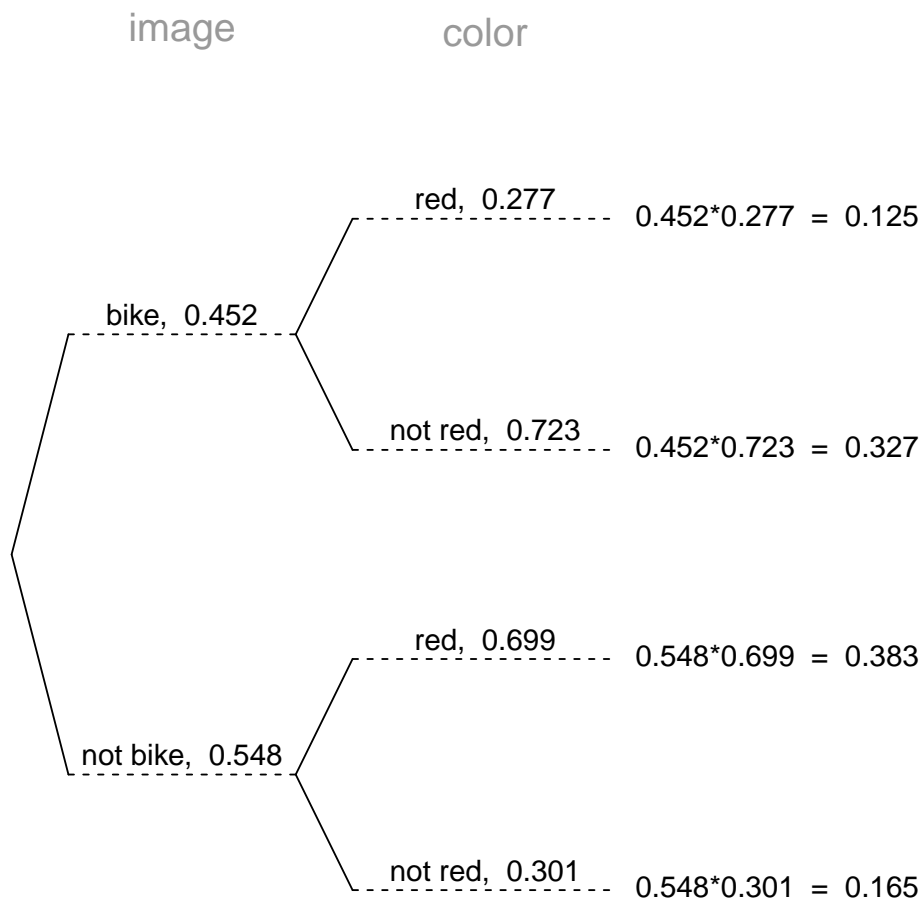
$$(f) P(\text{wheel given pink}) = \frac{95}{98+67+32+53+95} = 0.275$$

2. In a deck of strange cards, each card has an image and a color. The chance of drawing a bike is 45.2%. If a bike is drawn, there is a 27.7% chance that it is red. If a card that is not a bike is drawn, there is a 69.9% chance that it is red.

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

**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.

$$P(\text{"not bike" given "red"}) = \frac{0.383}{0.383 + 0.125} = 0.754$$

3. In a very large pile of toothpicks, the mean length is 73.12 millimeters and the standard deviation is 2.24 millimeters. If you randomly sample 100 toothpicks, what is the chance the sample mean is between 72.72 and 73.31 millimeters?



**Solution**

Label the given information.

$$\mu = 73.12$$

$$\sigma = 2.24$$

$$n = 100$$

$$\bar{x}_{\text{lower}} = 72.72$$

$$\bar{x}_{\text{upper}} = 73.31$$

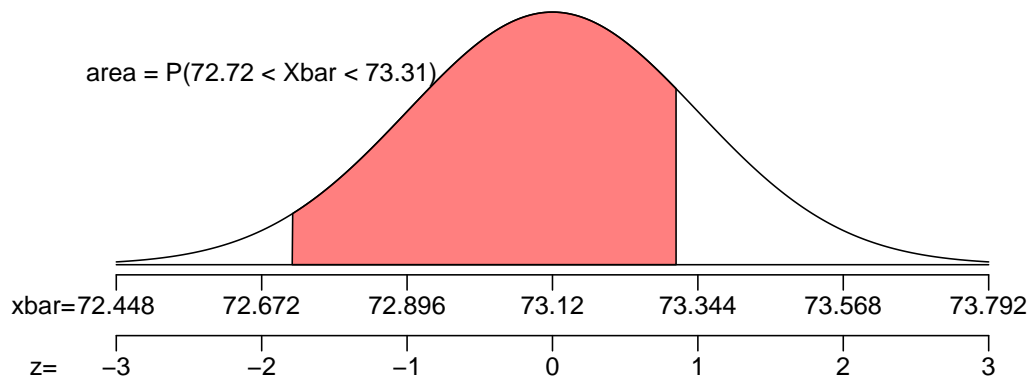
Find the standard error.

$$SE = \frac{\sigma}{\sqrt{n}} = \frac{2.24}{\sqrt{100}} = 0.224$$

Describe the sampling distribution.

$$\bar{X} \sim \mathcal{N}(73.12, 0.224)$$

Draw a sketch.



Calculate a z scores.

$$z_{\text{lower}} = \frac{\bar{x}_{\text{lower}} - \mu}{SE} = \frac{72.72 - 73.12}{0.224} = -1.79$$

$$z_{\text{upper}} = \frac{\bar{x}_{\text{upper}} - \mu}{SE} = \frac{73.31 - 73.12}{0.224} = 0.85$$

Determine the probability.

$$\begin{aligned} P(72.72 < \bar{X} < 73.31) &= \Phi(z_{\text{upper}}) - \Phi(z_{\text{lower}}) \\ &= \Phi(0.85) - \Phi(-1.79) \\ &= 0.7656 \end{aligned}$$

4. In a game, there is a 13% chance to win a round. You will play 245 rounds.
- (a) What is the probability of winning exactly 37 rounds?
  - (b) What is the probability of winning at least 31 but at most 42 rounds?

**Solution**

We use the formula for binomial probabilities.

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

$$P(X = 37) = \binom{245}{37} (0.13)^{37} (1 - 0.13)^{245-37}$$

$$P(X = 37) = \binom{245}{37} (0.13)^{37} (0.87)^{208}$$

$$P(X = 37) = 0.0449$$

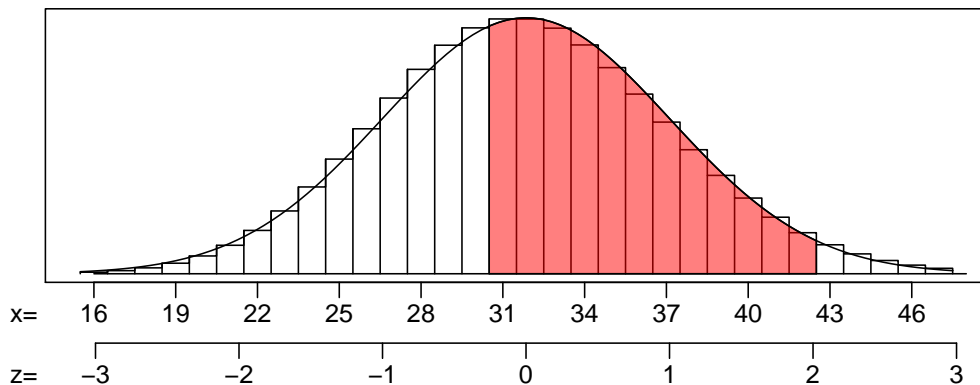
Find the mean.

$$\mu = np = (245)(0.13) = 31.85$$

Find the standard deviation.

$$\sigma = \sqrt{np(1 - p)} = \sqrt{(245)(0.13)(1 - 0.13)} = 5.264$$

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



Find the z scores.

$$z_1 = \frac{30.5 - 31.85}{5.264} = -0.26$$

$$z_2 = \frac{42.5 - 31.85}{5.264} = 2.02$$

Calculate the probability.

$$P(31 \leq X \leq 42) = \Phi(2.02) - \Phi(-0.26) = 0.5809$$

(a)  $P(X = 37) = 0.0449$

(b)  $P(31 \leq X \leq 42) = 0.5809$

5. As an ornithologist, you wish to determine the average body mass of *Cuculus americanus*. You randomly sample 19 adults of *Cuculus americanus*, resulting in a sample mean of 67.9 grams and a sample standard deviation of 12 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.

$$n = 19$$

$$\bar{x} = 67.9$$

$$s = 12$$

$$CL = 0.995$$

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

$$df = n - 1 = 18$$

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

$$t^* = 3.2$$

Calculate the standard error.

$$SE = \frac{s}{\sqrt{n}} = \frac{12}{\sqrt{19}} = 2.75$$

We want to make an inference about the population mean.

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

Determine the bounds.

$$\begin{aligned} CI &= (\bar{x} - t^* SE, \bar{x} + t^* SE) \\ &= (67.9 - 3.2 \times 2.75, 67.9 + 3.2 \times 2.75) \\ &= (59.1, 76.7) \end{aligned}$$

We are 99.5% confident that the population mean is between 59.1 and 76.7.

6. A treatment group of size 37 has a mean of 11 and standard deviation of 2.54. A control group of size 38 has a mean of 9.65 and standard deviation of 2.45. If you decided to use a significance level of 0.04, 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 72.

- (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.

$$n_1 = 37$$

$$\bar{x}_1 = 11$$

$$s_1 = 2.54$$

$$n_2 = 38$$

$$\bar{x}_2 = 9.65$$

$$s_2 = 2.45$$

$$\alpha = 0.04$$

$$df = 72$$

State the hypotheses.

$$H_0 : \mu_2 - \mu_1 = 0$$

$$H_A : \mu_2 - \mu_1 \neq 0$$

We are using a two-tail test. Find  $t^*$  such that  $P(|T| > t^*) = 0.04$  by using a  $t$  table.

$$t^* = 2.09$$

Calculate the standard error.

$$\begin{aligned} SE &= \sqrt{\frac{(s_1)^2}{n_1} + \frac{(s_2)^2}{n_2}} \\ &= \sqrt{\frac{(2.54)^2}{37} + \frac{(2.45)^2}{38}} \\ &= 0.576 \end{aligned}$$

Determine the test statistic.

$$\begin{aligned} t_{\text{obs}} &= \frac{(\bar{x}_2 - \bar{x}_1) - (\mu_2 - \mu_1)_0}{SE} \\ &= \frac{(9.65 - 11) - (0)}{0.576} \\ &= -2.34 \end{aligned}$$

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

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

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

$$0.02 < p\text{-value} < 0.04$$

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)  $t^* = 2.09$
- (d)  $SE = 0.576$
- (e)  $|t_{\text{obs}}| = 2.34$
- (f)  $0.02 < p\text{-value} < 0.04$
- (g) reject the null



7. From a very large population, a random sample of 2400 individuals was taken. In that sample, 63.7% were asleep. Determine a 95% 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  $P(|Z| < z^*) = 0.95$ .

$$z^* = 1.96$$

Calculate the standard error.

$$SE = \sqrt{\frac{\hat{p}(1 - \hat{p})}{n}} = \sqrt{\frac{(0.637)(1 - 0.637)}{2400}} = 0.00982$$

Calculate the margin of error.

$$ME = z^* SE = (1.96)(0.00982) = 0.0192$$

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.

$$(0.618, 0.656)$$

We are 95% confident that the true population proportion is between 61.8% and 65.6%.

- (a) The lower bound = 0.618, which can also be expressed as 61.8%.
- (b) The upper bound = 0.656, which can also be expressed as 65.6%.

8. An experiment is run with a treatment group of size 204 and a control group of size 217. The results are summarized in the table below.

	treatment	control
special	66	97
not special	138	120

Using a significance level of 0.02, determine whether the treatment causes an effect on the proportion of cases that are special.

- (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  $P(|Z| > z^*) = 0.02$ .

$$z^* = \Phi^{-1} \left( 1 - \frac{\alpha}{2} \right) = 2.33$$

Determine the sample proportions.

$$\hat{p}_1 = \frac{66}{204} = 0.324$$

$$\hat{p}_2 = \frac{97}{217} = 0.447$$

Determine the difference of sample proportions.

$$\hat{p}_2 - \hat{p}_1 = 0.447 - 0.324 = 0.123$$

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

$$\hat{p} = \frac{66 + 97}{204 + 217} = 0.387$$

Determine the standard error.

$$\begin{aligned} SE &= \sqrt{\frac{\hat{p}(1 - \hat{p})}{n_1} + \frac{\hat{p}(1 - \hat{p})}{n_2}} \\ &= \sqrt{\frac{(0.387)(0.613)}{204} + \frac{(0.387)(0.613)}{217}} \\ &= 0.0475 \end{aligned}$$

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

$$H_0 : \hat{P}_2 - \hat{P}_1 \sim \mathcal{N}(0, 0.0475)$$

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.

$$\begin{aligned} z &= \frac{(\hat{p}_2 - \hat{p}_1) - (p_2 - p_1)_0}{SE} \\ &= \frac{(0.447 - 0.324) - 0}{0.0475} \\ &= 2.59 \end{aligned}$$

Determine the  $p$ -value.

$$\begin{aligned} p\text{-value} &= 2 \cdot \Phi(-|z|) \\ &= 2 \cdot \Phi(-2.59) \\ &= 0.0096 \end{aligned}$$

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)  $z^* = 2.33$

(d)  $SE = 0.0475$

(e)  $|z_{\text{obs}}| = 2.59$

(f)  $p\text{-value} = 0.0096$

(g) reject the null