

# Bunker Hill Community College

Final Statistics Exam 2019-05-02

Exam ID 018

**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)  $P(\text{bike given indigo}) = 0.326$

(b)  $P(\text{flower and white}) = 0.0464$

(c)  $P(\text{indigo given bike}) = 0.247$

(d)  $P(\text{yellow}) = 0.159$

(e)  $P(\text{flower}) = 0.416$

(f)  $P(\text{flower or yellow}) = 0.514$

2.  $P(\text{"pig" given "not blue"}) = 0.324$

3.  $P(64.77 < X < 65.17) = 0.7298$

4. (a)  $P(X = 91) = 0.051$

(b)  $P(81 \leq X \leq 96) = 0.7745$

5.  $(10.1, 11)$

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

(b)  $H_0 : \mu_2 - \mu_1 \neq 0$

(c)  $t^* = 2.41$

(d)  $SE = 17.347$

(e)  $|t_{\text{obs}}| = 2.31$

(f)  $0.02 < p\text{-value} < 0.04$

(g) **retain**

7. (a) **LB of p CI = 0.553 or 55.3%**

(b) **UB of p CI = 0.571 or 57.1%**

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

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

(c)  $z^* = 1.96$

(d)  $SE = 0.041$

(e)  $|z_{\text{obs}}| = 2.16$

(f)  $p\text{-value} = 0.0308$

(g) **reject**

1. In a deck of strange cards, there are 841 cards. Each card has an image and a color. The amounts are shown in the table below.

	indigo	pink	teal	white	yellow
bike	59	90	18	12	60
cat	38	74	25	93	22
flower	84	96	79	39	52

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

**Solution**

$$(a) P(\text{bike given indigo}) = \frac{59}{59+38+84} = 0.326$$

$$(b) P(\text{flower and white}) = \frac{39}{841} = 0.0464$$

$$(c) P(\text{indigo given bike}) = \frac{59}{59+90+18+12+60} = 0.247$$

$$(d) P(\text{yellow}) = \frac{60+22+52}{841} = 0.159$$

$$(e) P(\text{flower}) = \frac{84+96+79+39+52}{841} = 0.416$$

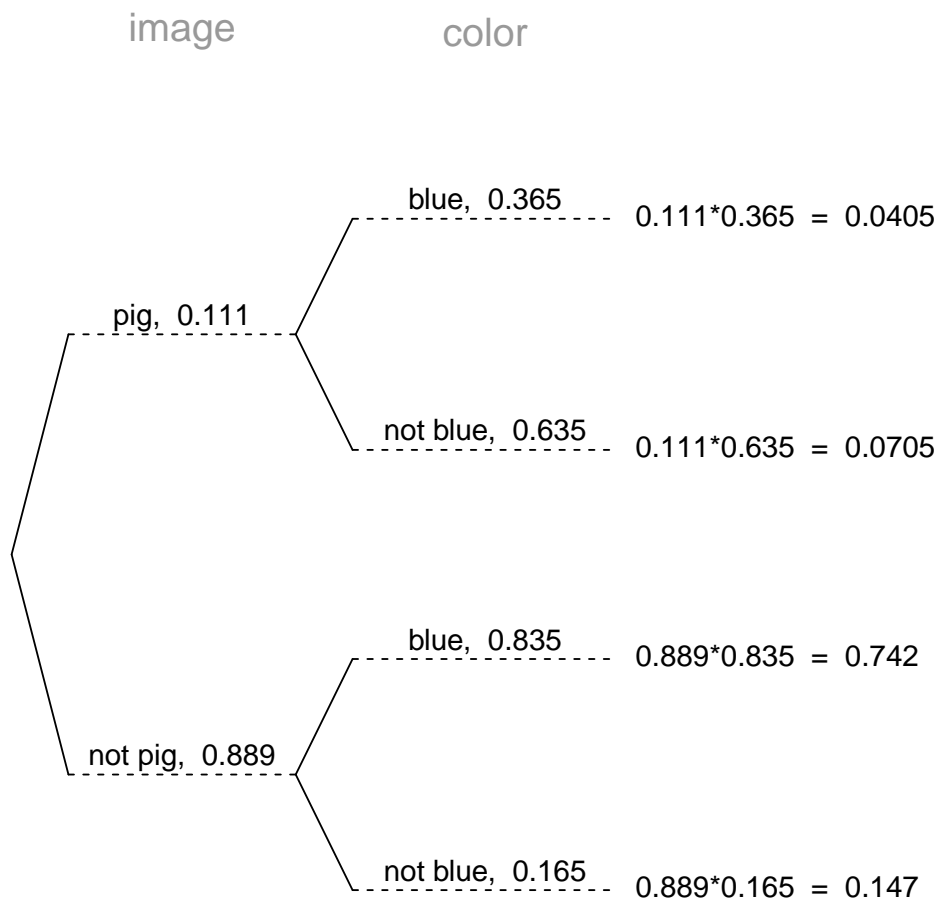
$$(f) P(\text{flower or yellow}) = \frac{84+96+79+39+52+60+22+52-52}{841} = 0.514$$

2. In a deck of strange cards, each card has an image and a color. The chance of drawing a pig is 11.1%. If a pig is drawn, there is a 36.5% chance that it is blue. If a card that is not a pig is drawn, there is a 83.5% chance that it is blue.

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

**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.

$$P(\text{"pig" given "not blue"}) = \frac{0.0705}{0.0705 + 0.147} = 0.324$$

3. In a very large pile of toothpicks, the mean length is 64.85 millimeters and the standard deviation is 1.42 millimeters. If you randomly sample 125 toothpicks, what is the chance the sample mean is between 64.77 and 65.17 millimeters?



**Solution**

Label the given information.

$$\mu = 64.85$$

$$\sigma = 1.42$$

$$n = 125$$

$$\bar{x}_{\text{lower}} = 64.77$$

$$\bar{x}_{\text{upper}} = 65.17$$

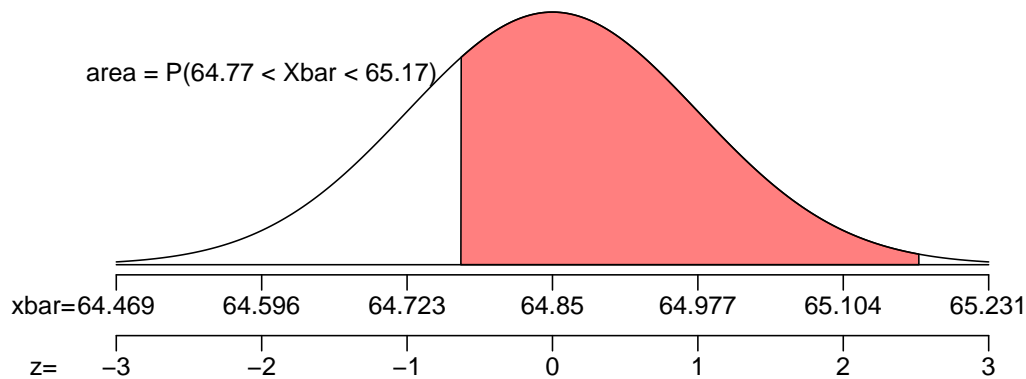
Find the standard error.

$$SE = \frac{\sigma}{\sqrt{n}} = \frac{1.42}{\sqrt{125}} = 0.127$$

Describe the sampling distribution.

$$\bar{X} \sim \mathcal{N}(64.85, 0.127)$$

Draw a sketch.



Calculate a z scores.

$$z_{\text{lower}} = \frac{\bar{x}_{\text{lower}} - \mu}{SE} = \frac{64.77 - 64.85}{0.127} = -0.63$$

$$z_{\text{upper}} = \frac{\bar{x}_{\text{upper}} - \mu}{SE} = \frac{65.17 - 64.85}{0.127} = 2.52$$

Determine the probability.

$$\begin{aligned} P(64.77 < \bar{X} < 65.17) &= \Phi(z_{\text{upper}}) - \Phi(z_{\text{lower}}) \\ &= \Phi(2.52) - \Phi(-0.63) \\ &= 0.7298 \end{aligned}$$

4. In a game, there is a 53% chance to win a round. You will play 164 rounds.
- (a) What is the probability of winning exactly 91 rounds?
  - (b) What is the probability of winning at least 81 but at most 96 rounds?

**Solution**

We use the formula for binomial probabilities.

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

$$P(X = 91) = \binom{164}{91} (0.53)^{91} (1 - 0.53)^{164-91}$$

$$P(X = 91) = \binom{164}{91} (0.53)^{91} (0.47)^{73}$$

$$P(X = 91) = 0.051$$

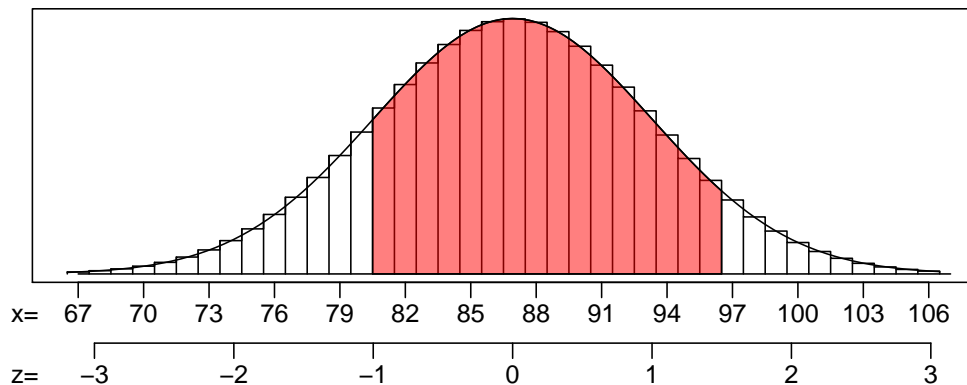
Find the mean.

$$\mu = np = (164)(0.53) = 86.92$$

Find the standard deviation.

$$\sigma = \sqrt{np(1 - p)} = \sqrt{(164)(0.53)(1 - 0.53)} = 6.3916$$

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



Find the z scores.

$$z_1 = \frac{80.5 - 86.92}{6.3916} = -1$$

$$z_2 = \frac{96.5 - 86.92}{6.3916} = 1.5$$

Calculate the probability.

$$P(81 \leq X \leq 96) = \Phi(1.5) - \Phi(-1) = 0.7745$$

(a)  $P(X = 91) = 0.051$

(b)  $P(81 \leq X \leq 96) = 0.7745$

5. As an ornithologist, you wish to determine the average body mass of *Dendroica palmarum*. You randomly sample 29 adults of *Dendroica palmarum*, resulting in a sample mean of 10.54 grams and a sample standard deviation of 1.2 grams. Determine a 95% confidence interval of the true population mean.

**Solution**

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

$$n = 29$$

$$\bar{x} = 10.54$$

$$s = 1.2$$

$$CL = 0.95$$

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

$$df = n - 1 = 28$$

Determine the critical  $t$  value,  $t^*$ , such that  $P(|T| < t^*) = 0.95$ .

$$t^* = 2.05$$

Calculate the standard error.

$$SE = \frac{s}{\sqrt{n}} = \frac{1.2}{\sqrt{29}} = 0.223$$

We want to make an inference about the population mean.

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

Determine the bounds.

$$\begin{aligned} CI &= (\bar{x} - t^* SE, \bar{x} + t^* SE) \\ &= (10.54 - 2.05 \times 0.223, 10.54 + 2.05 \times 0.223) \\ &= (10.1, 11) \end{aligned}$$

We are 95% confident that the population mean is between 10.1 and 11.

6. A treatment group of size 31 has a mean of 1020 and standard deviation of 69.1. A control group of size 20 has a mean of 1060 and standard deviation of 54.2. 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 47.

- (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.

$$\begin{aligned}n_1 &= 31 \\ \bar{x}_1 &= 1020 \\ s_1 &= 69.1 \\ n_2 &= 20 \\ \bar{x}_2 &= 1060 \\ s_2 &= 54.2 \\ \alpha &= 0.02 \\ df &= 47\end{aligned}$$

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.

$$t^* = 2.41$$

Calculate the standard error.

$$\begin{aligned}SE &= \sqrt{\frac{(s_1)^2}{n_1} + \frac{(s_2)^2}{n_2}} \\ &= \sqrt{\frac{(69.1)^2}{31} + \frac{(54.2)^2}{20}} \\ &= 17.347\end{aligned}$$

Determine the test statistic.

$$\begin{aligned}t_{\text{obs}} &= \frac{(\bar{x}_2 - \bar{x}_1) - (\mu_2 - \mu_1)_0}{SE} \\ &= \frac{(1060 - 1020) - (0)}{17.347} \\ &= 2.31\end{aligned}$$

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 retain the null hypothesis.

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

- (b)  $H_A : \mu_2 - \mu_1 \neq 0$
- (c)  $t^* = 2.41$
- (d)  $SE = 17.347$
- (e)  $|t_{\text{obs}}| = 2.31$
- (f)  $0.02 < p\text{-value} < 0.04$
- (g) retain the null



7. From a very large population, a random sample of 5300 individuals was taken. In that sample, 56.2% were special. Determine a 80% 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  $P(|Z| < z^*) = 0.8$ .

$$z^* = 1.28$$

Calculate the standard error.

$$SE = \sqrt{\frac{\hat{p}(1 - \hat{p})}{n}} = \sqrt{\frac{(0.562)(1 - 0.562)}{5300}} = 0.00682$$

Calculate the margin of error.

$$ME = z^* SE = (1.28)(0.00682) = 0.00873$$

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.

$$(0.553, 0.571)$$

We are 80% confident that the true population proportion is between 55.3% and 57.1%.

- (a) The lower bound = 0.553, which can also be expressed as 55.3%.
- (b) The upper bound = 0.571, which can also be expressed as 57.1%.

8. An experiment is run with a treatment group of size 285 and a control group of size 314. The results are summarized in the table below.

	treatment	control
sorry	121	161
not sorry	164	153

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

- (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.

$$\hat{p}_1 = \frac{121}{285} = 0.425$$

$$\hat{p}_2 = \frac{161}{314} = 0.513$$

Determine the difference of sample proportions.

$$\hat{p}_2 - \hat{p}_1 = 0.513 - 0.425 = 0.088$$

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

$$\hat{p} = \frac{121 + 161}{285 + 314} = 0.471$$

Determine the standard error.

$$\begin{aligned} SE &= \sqrt{\frac{\hat{p}(1 - \hat{p})}{n_1} + \frac{\hat{p}(1 - \hat{p})}{n_2}} \\ &= \sqrt{\frac{(0.471)(0.529)}{285} + \frac{(0.471)(0.529)}{314}} \\ &= 0.0408 \end{aligned}$$

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

$$H_0 : \hat{P}_2 - \hat{P}_1 \sim \mathcal{N}(0, 0.0408)$$

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.

$$\begin{aligned} z &= \frac{(\hat{p}_2 - \hat{p}_1) - (p_2 - p_1)_0}{SE} \\ &= \frac{(0.513 - 0.425) - 0}{0.0408} \\ &= 2.16 \end{aligned}$$

Determine the  $p$ -value.

$$\begin{aligned} p\text{-value} &= 2 \cdot \Phi(-|z|) \\ &= 2 \cdot \Phi(-2.16) \\ &= 0.0308 \end{aligned}$$

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)  $SE = 0.0408$

(e)  $|z_{\text{obs}}| = 2.16$

(f)  $p\text{-value} = 0.0308$

(g) reject the null