

# Bunker Hill Community College

Final Statistics Exam 2019-05-02

Exam ID 001

**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)  $P(\text{wheel given red}) = 0.61$

(b)  $P(\text{wheel or gray}) = 0.499$

(c)  $P(\text{violet given wheel}) = 0.103$

(d)  $P(\text{orange}) = 0.217$

(e)  $P(\text{cat}) = 0.296$

(f)  $P(\text{cat and orange}) = 0.0628$

2.  $P(\text{"not kite" given "gray"}) = 0.881$

3.  $P(68.39 < X < 68.91) = 0.7856$

4. (a)  $P(X = 19) = 0.1032$

(b)  $P(13 \leq X \leq 25) = 0.8758$

5. **(17.1, 20.4)**

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

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

(c)  $t^* = 2.72$

(d)  $SE = 8.005$

(e)  $|t_{\text{obs}}| = 2.5$

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

(g) **retain**

7. (a) **LB of p CI = 0.709 or 70.9%**

(b) **UB of p CI = 0.763 or 76.3%**

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

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

(c)  $Z^* = 2.58$

(d)  $SE = 0.029$

(e)  $|Z_{\text{obs}}| = 2.76$

(f)  $p\text{-value} = 0.0058$

(g) **reject**

1. In a deck of strange cards, there are 764 cards. Each card has an image and a color. The amounts are shown in the table below.

|       | black | gray | orange | red | violet |
|-------|-------|------|--------|-----|--------|
| cat   | 45    | 97   | 48     | 19  | 17     |
| dog   | 82    | 32   | 54     | 27  | 91     |
| wheel | 43    | 47   | 64     | 72  | 26     |

- (a) What is the probability a random card is a wheel given it is red?
- (b) What is the probability a random card is either a wheel or gray (or both)?
- (c) What is the probability a random card is violet given it is a wheel?
- (d) What is the probability a random card is orange?
- (e) What is the probability a random card is a cat?
- (f) What is the probability a random card is both a cat and orange?

**Solution**

(a)  $P(\text{wheel given red}) = \frac{72}{19+27+72} = 0.61$

(b)  $P(\text{wheel or gray}) = \frac{43+47+64+72+26+97+32+47-47}{764} = 0.499$

(c)  $P(\text{violet given wheel}) = \frac{26}{43+47+64+72+26} = 0.103$

(d)  $P(\text{orange}) = \frac{48+54+64}{764} = 0.217$

(e)  $P(\text{cat}) = \frac{45+97+48+19+17}{764} = 0.296$

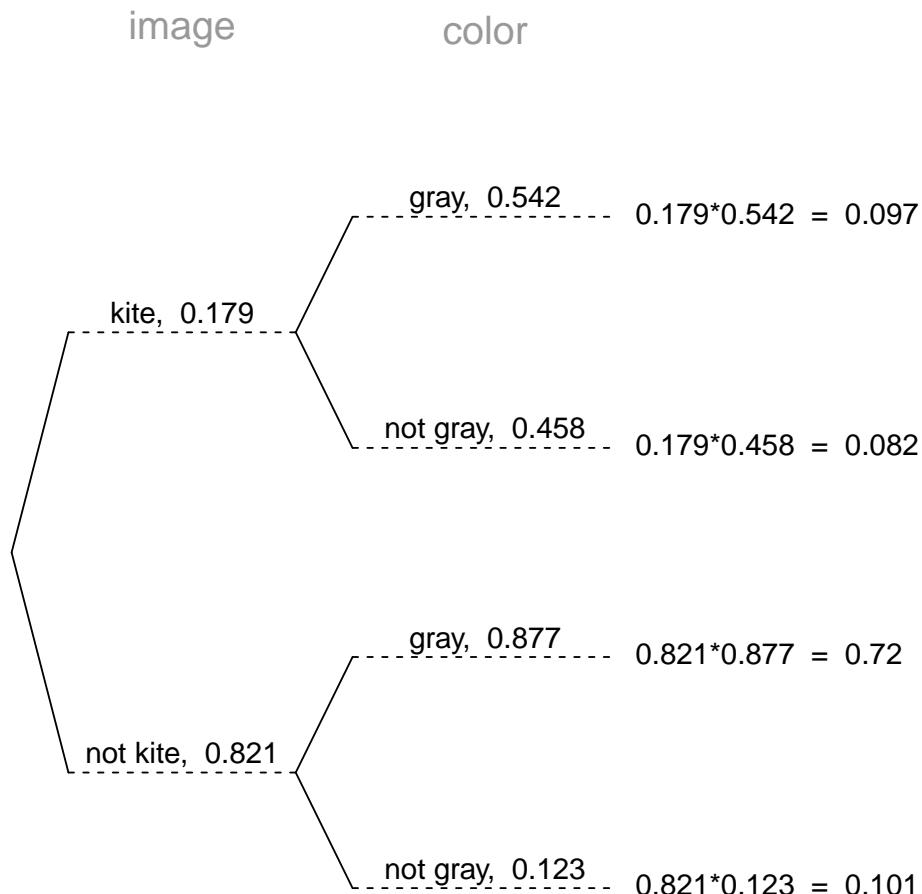
(f)  $P(\text{cat and orange}) = \frac{48}{764} = 0.0628$

2. In a deck of strange cards, each card has an image and a color. The chance of drawing a kite is 17.9%. If a kite is drawn, there is a 54.2% chance that it is gray. If a card that is not a kite is drawn, there is a 87.7% chance that it is gray.

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

**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.

$$P(\text{"not kite" given "gray"}) = \frac{0.72}{0.72 + 0.097} = 0.881$$

3. In a very large pile of toothpicks, the mean length is 68.61 millimeters and the standard deviation is 2.92 millimeters. If you randomly sample 200 toothpicks, what is the chance the sample mean is between 68.39 and 68.91 millimeters?

**Solution**

Label the given information.

$$\mu = 68.61$$

$$\sigma = 2.92$$

$$n = 200$$

$$\bar{x}_{\text{lower}} = 68.39$$

$$\bar{x}_{\text{upper}} = 68.91$$

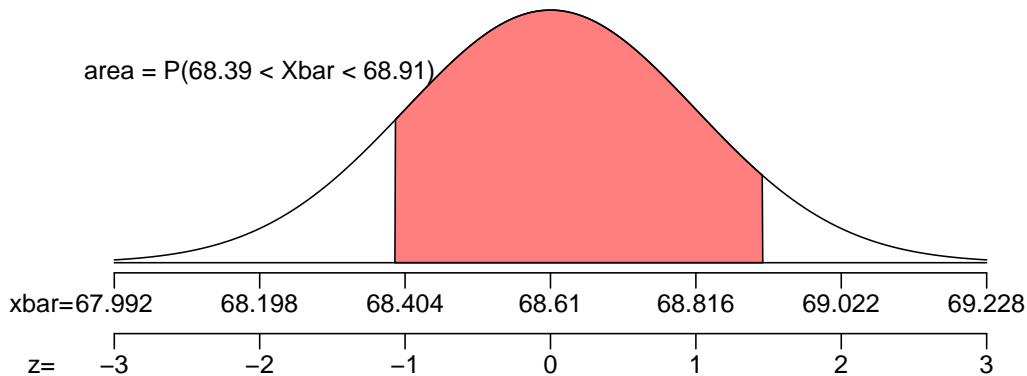
Find the standard error.

$$SE = \frac{\sigma}{\sqrt{n}} = \frac{2.92}{\sqrt{200}} = 0.206$$

Describe the sampling distribution.

$$\bar{X} \sim \mathcal{N}(68.61, 0.206)$$

Draw a sketch.



Calculate a  $z$  scores.

$$z_{\text{lower}} = \frac{x_{\text{lower}} - \mu}{SE} = \frac{68.39 - 68.61}{0.206} = -1.07$$

$$z_{\text{upper}} = \frac{x_{\text{upper}} - \mu}{SE} = \frac{68.91 - 68.61}{0.206} = 1.46$$

Determine the probability.

$$\begin{aligned} P(68.39 < X < 68.91) &= \Phi(z_{\text{upper}}) - \Phi(z_{\text{lower}}) \\ &= \Phi(1.46) - \Phi(-1.07) \\ &= 0.7856 \end{aligned}$$

4. In a game, there is a 31% chance to win a round. You will play 65 rounds.
- (a) What is the probability of winning exactly 19 rounds?
  - (b) What is the probability of winning at least 13 but at most 25 rounds?

**Solution**

We use the formula for binomial probabilities.

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

$$P(X = 19) = \binom{65}{19} (0.31)^{19} (1 - 0.31)^{65-19}$$

$$P(X = 19) = \binom{65}{19} (0.31)^{19} (0.69)^{46}$$

$$P(X = 19) = 0.1032$$

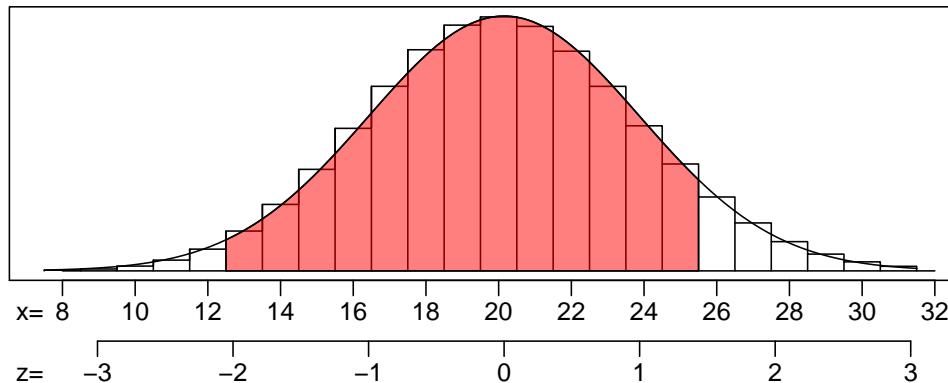
Find the mean.

$$\mu = np = (65)(0.31) = 20.15$$

Find the standard deviation.

$$\sigma = \sqrt{np(1 - p)} = \sqrt{(65)(0.31)(1 - 0.31)} = 3.7287$$

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



Find the z scores.

$$z_1 = \frac{12.5 - 20.15}{3.7287} = -1.92$$

$$z_2 = \frac{25.5 - 20.15}{3.7287} = 1.3$$

Calculate the probability.

$$P(13 \leq X \leq 25) = \Phi(1.3) - \Phi(-1.92) = 0.8758$$

(a)  $P(X = 19) = 0.1032$

(b)  $P(13 \leq X \leq 25) = 0.8758$

5. As an ornithologist, you wish to determine the average body mass of *Vireo olivaceus*. You randomly sample 20 adults of *Vireo olivaceus*, resulting in a sample mean of 18.75 grams and a sample standard deviation of 2.6 grams. Determine a 99% confidence interval of the true population mean.

**Solution**

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

$$n = 20$$

$$\bar{x} = 18.75$$

$$s = 2.6$$

$$CL = 0.99$$

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

$$df = n - 1 = 19$$

Determine the critical  $t$  value,  $t^*$ , such that  $P(|T| < t^*) = 0.99$ .

$$t^* = 2.86$$

Calculate the standard error.

$$SE = \frac{s}{\sqrt{n}} = \frac{2.6}{\sqrt{20}} = 0.581$$

We want to make an inference about the population mean.

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

Determine the bounds.

$$\begin{aligned} CI &= (\bar{x} - t^* SE, \bar{x} + t^* SE) \\ &= (18.75 - 2.86 \times 0.581, 18.75 + 2.86 \times 0.581) \\ &= (17.1, 20.4) \end{aligned}$$

We are 99% confident that the population mean is between 17.1 and 20.4.

6. A treatment group of size 15 has a mean of 123 and standard deviation of 22.6. A control group of size 32 has a mean of 103 and standard deviation of 31. If you decided to use a significance level of 0.01, 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 36.

- (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.

$$n_1 = 15$$

$$\bar{x}_1 = 123$$

$$s_1 = 22.6$$

$$n_2 = 32$$

$$\bar{x}_2 = 103$$

$$s_2 = 31$$

$$\alpha = 0.01$$

$$df = 36$$

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  $P(|T| > t^*) = 0.01$  by using a  $t$  table.

$$t^* = 2.72$$

Calculate the standard error.

$$\begin{aligned} SE &= \sqrt{\frac{(s_1)^2}{n_1} + \frac{(s_2)^2}{n_2}} \\ &= \sqrt{\frac{(22.6)^2}{15} + \frac{(31)^2}{32}} \\ &= 8.005 \end{aligned}$$

Determine the test statistic.

$$\begin{aligned} t_{\text{obs}} &= \frac{(\bar{x}_2 - \bar{x}_1) - (\mu_2 - \mu_1)_0}{SE} \\ &= \frac{(103 - 123) - (0)}{8.005} \\ &= -2.5 \end{aligned}$$

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 retain the null hypothesis.

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

- (b)  $H_A : \mu_2 - \mu_1 \neq 0$
- (c)  $t^* = 2.72$
- (d)  $SE = 8.005$
- (e)  $|t_{\text{obs}}| = 2.5$
- (f)  $0.01 < p\text{-value} < 0.02$
- (g) retain the null

7. From a very large population, a random sample of 1100 individuals was taken. In that sample, 73.6% were special. 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 condifence interval.

**Solution**

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

$$z^* = 2.05$$

Calculate the standard error.

$$SE = \sqrt{\frac{\hat{p}(1 - \hat{p})}{n}} = \sqrt{\frac{(0.736)(1 - 0.736)}{1100}} = 0.0133$$

Calculate the margin of error.

$$ME = z^* SE = (2.05)(0.0133) = 0.0273$$

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.

$$(0.709, 0.763)$$

We are 96% confident that the true population proportion is between 70.9% and 76.3%.

- (a) The lower bound = 0.709, which can also be expressed as 70.9%.
- (b) The upper bound = 0.763, which can also be expressed as 76.3%.

8. An experiment is run with a treatment group of size 147 and a control group of size 170. The results are summarized in the table below.

|           | treatment | control |
|-----------|-----------|---------|
| green     | 130       | 164     |
| not green | 17        | 6       |

Using a significance level of 0.01, determine whether the treatment causes an effect on the proportion of cases that are green.

- (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  $P(|Z| > z^*) = 0.01$ .

$$z^* = \Phi^{-1} \left( 1 - \frac{\alpha}{2} \right) = 2.58$$

Determine the sample proportions.

$$\hat{p}_1 = \frac{130}{147} = 0.884$$

$$\hat{p}_2 = \frac{164}{170} = 0.965$$

Determine the difference of sample proportions.

$$\hat{p}_2 - \hat{p}_1 = 0.965 - 0.884 = 0.081$$

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

$$\hat{p} = \frac{130 + 164}{147 + 170} = 0.927$$

Determine the standard error.

$$\begin{aligned} SE &= \sqrt{\frac{\hat{p}(1 - \hat{p})}{n_1} + \frac{\hat{p}(1 - \hat{p})}{n_2}} \\ &= \sqrt{\frac{(0.927)(0.073)}{147} + \frac{(0.927)(0.073)}{170}} \\ &= 0.0293 \end{aligned}$$

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

$$H_0 : \hat{P}_2 - \hat{P}_1 \sim \mathcal{N}(0, 0.0293)$$

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.

$$\begin{aligned} z &= \frac{(\hat{p}_2 - \hat{p}_1) - (p_2 - p_1)_0}{SE} \\ &= \frac{(0.965 - 0.884) - 0}{0.0293} \\ &= 2.76 \end{aligned}$$

Determine the  $p$ -value.

$$\begin{aligned} p\text{-value} &= 2 \cdot \Phi(-|z|) \\ &= 2 \cdot \Phi(-2.76) \\ &= 0.0058 \end{aligned}$$

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)  $z^* = 2.58$
- (d)  $SE = 0.0293$
- (e)  $|z_{\text{obs}}| = 2.76$
- (f)  $p\text{-value} = 0.0058$
- (g) reject the null

# Bunker Hill Community College

Final Statistics Exam 2019-05-02

Exam ID 002

**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)  $P(\text{pink given bike}) = 0.321$

(b)  $P(\text{tree given pink}) = 0.21$

(c)  $P(\text{red}) = 0.283$

(d)  $P(\text{bike and red}) = 0.112$

(e)  $P(\text{bike or red}) = 0.502$

(f)  $P(\text{pig}) = 0.208$

2.  $P(\text{"not tree" given "not yellow"}) = 0.511$

3.  $P(61.55 < X < 62.3) = 0.8251$

4. (a)  $P(X = 30) = 0.0884$

(b)  $P(34 \leq X \leq 39) = 0.2244$

5.  $(54.2, 57)$

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

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

(c)  $t^* = 1.99$

(d)  $SE = 20.518$

(e)  $|t_{\text{obs}}| = 1.95$

(f)  $0.05 < p\text{-value} < 0.1$

(g) **retain**

7. (a) **LB of p CI = 0.0182 or 1.82%**

(b) **UB of p CI = 0.0198 or 1.98%**

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

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

(c)  $Z^* = 2.05$

(d)  $SE = 0.166$

(e)  $|Z_{\text{obs}}| = 2.14$

(f)  $p\text{-value} = 0.0324$

(g) **reject**

1. In a deck of strange cards, there are 827 cards. Each card has an image and a color. The amounts are shown in the table below.

|       | gray | pink | red | white |
|-------|------|------|-----|-------|
| bike  | 24   | 88   | 93  | 69    |
| pig   | 87   | 27   | 42  | 16    |
| tree  | 21   | 38   | 47  | 58    |
| wheel | 51   | 28   | 52  | 86    |

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

**Solution**

$$(a) P(\text{pink given bike}) = \frac{88}{24+88+93+69} = 0.321$$

$$(b) P(\text{tree given pink}) = \frac{38}{88+27+38+28} = 0.21$$

$$(c) P(\text{red}) = \frac{93+42+47+52}{827} = 0.283$$

$$(d) P(\text{bike and red}) = \frac{93}{827} = 0.112$$

$$(e) P(\text{bike or red}) = \frac{24+88+93+69+93+42+47+52-93}{827} = 0.502$$

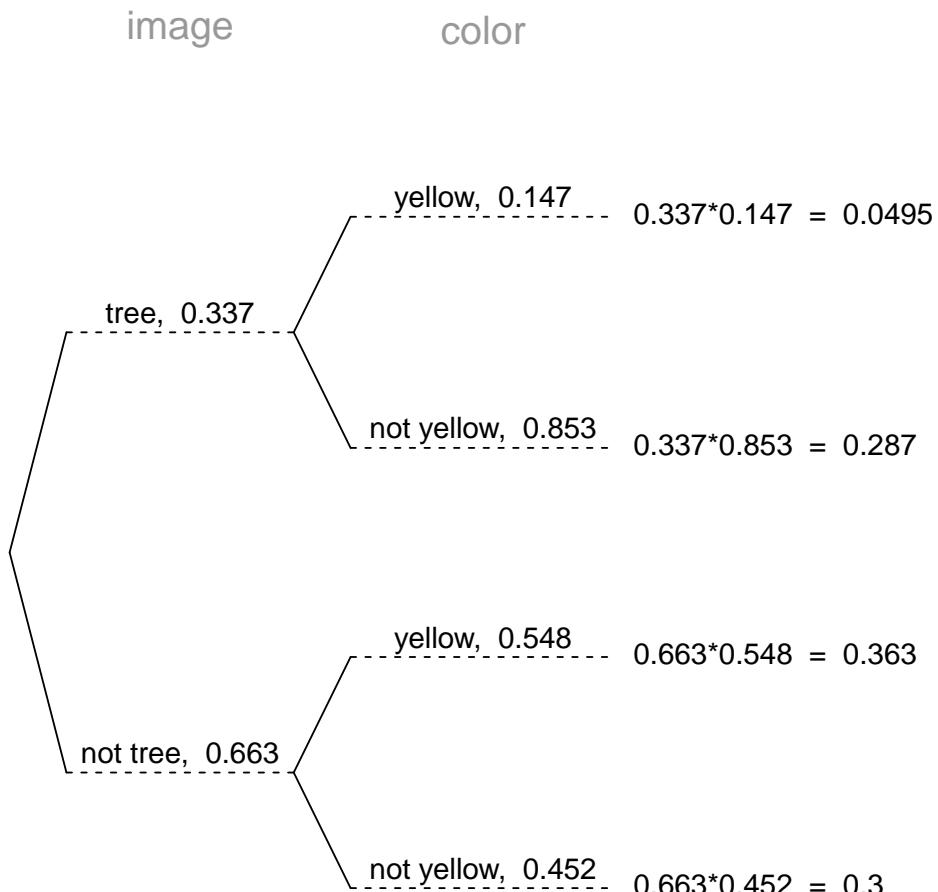
$$(f) P(\text{pig}) = \frac{87+27+42+16}{827} = 0.208$$

2. In a deck of strange cards, each card has an image and a color. The chance of drawing a tree is 33.7%. If a tree is drawn, there is a 14.7% chance that it is yellow. If a card that is not a tree is drawn, there is a 54.8% chance that it is yellow.

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

**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.

$$P(\text{"not tree" given "not yellow"}) = \frac{0.3}{0.3 + 0.287} = 0.511$$

3. In a very large pile of toothpicks, the mean length is 61.98 millimeters and the standard deviation is 3.52 millimeters. If you randomly sample 169 toothpicks, what is the chance the sample mean is between 61.55 and 62.3 millimeters?

**Solution**

Label the given information.

$$\mu = 61.98$$

$$\sigma = 3.52$$

$$n = 169$$

$$\bar{x}_{\text{lower}} = 61.55$$

$$\bar{x}_{\text{upper}} = 62.3$$

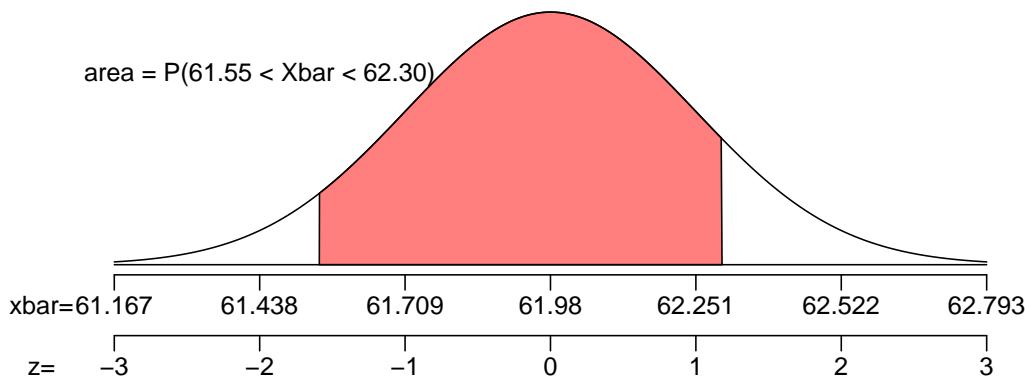
Find the standard error.

$$SE = \frac{\sigma}{\sqrt{n}} = \frac{3.52}{\sqrt{169}} = 0.271$$

Describe the sampling distribution.

$$\bar{X} \sim \mathcal{N}(61.98, 0.271)$$

Draw a sketch.



Calculate a  $z$  scores.

$$z_{\text{lower}} = \frac{x_{\text{lower}} - \mu}{SE} = \frac{61.55 - 61.98}{0.271} = -1.59$$

$$z_{\text{upper}} = \frac{x_{\text{upper}} - \mu}{SE} = \frac{62.3 - 61.98}{0.271} = 1.18$$

Determine the probability.

$$\begin{aligned} P(61.55 < X < 62.3) &= \Phi(z_{\text{upper}}) - \Phi(z_{\text{lower}}) \\ &= \Phi(1.18) - \Phi(-1.59) \\ &= 0.8251 \end{aligned}$$

4. In a game, there is a 39% chance to win a round. You will play 80 rounds.
- (a) What is the probability of winning exactly 30 rounds?
  - (b) What is the probability of winning at least 34 but at most 39 rounds?

**Solution**

We use the formula for binomial probabilities.

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

$$P(X = 30) = \binom{80}{30} (0.39)^{30} (1 - 0.39)^{80-30}$$

$$P(X = 30) = \binom{80}{30} (0.39)^{30} (0.61)^{50}$$

$$P(X = 30) = 0.0884$$

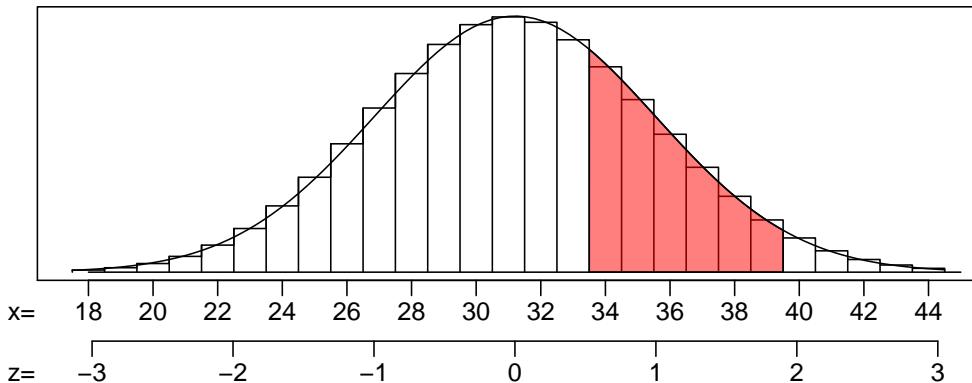
Find the mean.

$$\mu = np = (80)(0.39) = 31.2$$

Find the standard deviation.

$$\sigma = \sqrt{np(1 - p)} = \sqrt{(80)(0.39)(1 - 0.39)} = 4.3626$$

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



Find the z scores.

$$z_1 = \frac{33.5 - 31.2}{4.3626} = 0.64$$

$$z_2 = \frac{39.5 - 31.2}{4.3626} = 1.79$$

Calculate the probability.

$$P(34 \leq X \leq 39) = \Phi(1.79) - \Phi(0.64) = 0.2244$$

(a)  $P(X = 30) = 0.0884$

(b)  $P(34 \leq X \leq 39) = 0.2244$

5. As an ornithologist, you wish to determine the average body mass of *Hylocichla mustelina*. You randomly sample 29 adults of *Hylocichla mustelina*, resulting in a sample mean of 55.61 grams and a sample standard deviation of 5.63 grams. Determine a 80% confidence interval of the true population mean.

**Solution**

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

$$n = 29$$

$$\bar{x} = 55.61$$

$$s = 5.63$$

$$CL = 0.8$$

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  $P(|T| < t^*) = 0.8$ .

$$t^* = 1.31$$

Calculate the standard error.

$$SE = \frac{s}{\sqrt{n}} = \frac{5.63}{\sqrt{29}} = 1.05$$

We want to make an inference about the population mean.

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

Determine the bounds.

$$\begin{aligned} CI &= (\bar{x} - t^* SE, \bar{x} + t^* SE) \\ &= (55.61 - 1.31 \times 1.05, 55.61 + 1.31 \times 1.05) \\ &= (54.2, 57) \end{aligned}$$

We are 80% confident that the population mean is between 54.2 and 57.

6. A treatment group of size 32 has a mean of 1060 and standard deviation of 77.5. A control group of size 40 has a mean of 1020 and standard deviation of 96.6. If you decided to use a significance level of 0.05, 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 69.

- (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.

$$n_1 = 32$$

$$\bar{x}_1 = 1060$$

$$s_1 = 77.5$$

$$n_2 = 40$$

$$\bar{x}_2 = 1020$$

$$s_2 = 96.6$$

$$\alpha = 0.05$$

$$df = 69$$

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  $P(|T| > t^*) = 0.05$  by using a  $t$  table.

$$t^* = 1.99$$

Calculate the standard error.

$$\begin{aligned} SE &= \sqrt{\frac{(s_1)^2}{n_1} + \frac{(s_2)^2}{n_2}} \\ &= \sqrt{\frac{(77.5)^2}{32} + \frac{(96.6)^2}{40}} \\ &= 20.518 \end{aligned}$$

Determine the test statistic.

$$\begin{aligned} t_{\text{obs}} &= \frac{(\bar{x}_2 - \bar{x}_1) - (\mu_2 - \mu_1)_0}{SE} \\ &= \frac{(1020 - 1060) - (0)}{20.518} \\ &= -1.95 \end{aligned}$$

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

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

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

$$0.05 < p\text{-value} < 0.1$$

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)  $t^* = 1.99$
- (d)  $SE = 20.518$
- (e)  $|t_{\text{obs}}| = 1.95$
- (f)  $0.05 < p\text{-value} < 0.1$
- (g) retain the null

7. From a very large population, a random sample of 52000 individuals was taken. In that sample, 1.9% were angry. 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 condifence interval.

**Solution**

Determine  $z^*$  such that  $P(|Z| < z^*) = 0.8$ .

$$z^* = 1.28$$

Calculate the standard error.

$$SE = \sqrt{\frac{\hat{p}(1 - \hat{p})}{n}} = \sqrt{\frac{(0.019)(1 - 0.019)}{52000}} = 0.000599$$

Calculate the margin of error.

$$ME = z^* SE = (1.28)(0.000599) = 0.000767$$

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.

$$(0.0182, 0.0198)$$

We are 80% confident that the true population proportion is between 1.82% and 1.98%.

- (a) The lower bound = 0.0182, which can also be expressed as 1.82%.
- (b) The upper bound = 0.0198, which can also be expressed as 1.98%.

8. An experiment is run with a treatment group of size 16 and a control group of size 21. The results are summarized in the table below.

|             | treatment | control |
|-------------|-----------|---------|
| abysmal     | 5         | 14      |
| not abysmal | 11        | 7       |

Using a significance level of 0.04, determine whether the treatment causes an effect on the proportion of cases that are abysmal.

- (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  $P(|Z| > z^*) = 0.04$ .

$$z^* = \Phi^{-1} \left( 1 - \frac{\alpha}{2} \right) = 2.05$$

Determine the sample proportions.

$$\hat{p}_1 = \frac{5}{16} = 0.312$$

$$\hat{p}_2 = \frac{14}{21} = 0.667$$

Determine the difference of sample proportions.

$$\hat{p}_2 - \hat{p}_1 = 0.667 - 0.312 = 0.355$$

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

$$\hat{p} = \frac{5 + 14}{16 + 21} = 0.514$$

Determine the standard error.

$$\begin{aligned} SE &= \sqrt{\frac{\hat{p}(1 - \hat{p})}{n_1} + \frac{\hat{p}(1 - \hat{p})}{n_2}} \\ &= \sqrt{\frac{(0.514)(0.486)}{16} + \frac{(0.514)(0.486)}{21}} \\ &= 0.166 \end{aligned}$$

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

$$H_0 : \hat{P}_2 - \hat{P}_1 \sim \mathcal{N}(0, 0.166)$$

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.

$$\begin{aligned} z &= \frac{(\hat{p}_2 - \hat{p}_1) - (p_2 - p_1)_0}{SE} \\ &= \frac{(0.667 - 0.312) - 0}{0.166} \\ &= 2.14 \end{aligned}$$

Determine the  $p$ -value.

$$\begin{aligned} p\text{-value} &= 2 \cdot \Phi(-|z|) \\ &= 2 \cdot \Phi(-2.14) \\ &= 0.0324 \end{aligned}$$

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)  $z^* = 2.05$
- (d)  $SE = 0.166$
- (e)  $|Z_{\text{obs}}| = 2.14$
- (f)  $p\text{-value} = 0.0324$
- (g) reject the null

# Bunker Hill Community College

Final Statistics Exam 2019-05-02

Exam ID 003

**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)  $P(\text{wheel and orange}) = 0.0495$

(b)  $P(\text{blue}) = 0.18$

(c)  $P(\text{tree}) = 0.229$

(d)  $P(\text{pig given red}) = 0.208$

(e)  $P(\text{indigo given tree}) = 0.307$

(f)  $P(\text{wheel or orange}) = 0.408$

2.  $P(\text{"cat" given "not green"}) = 0.22$

3.  $P(66.76 < X < 66.99) = 0.6043$

4. (a)  $P(X = 37) = 0.0722$

(b)  $P(38 \leq X \leq 48) = 0.3432$

5.  $(63.1, 70.9)$

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

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

(c)  $t^* = 2.13$

(d)  $SE = 6.063$

(e)  $|t_{\text{obs}}| = 1.85$

(f)  $0.05 < p\text{-value} < 0.1$

(g) **retain**

7. (a) **LB of p CI = 0.0981 or 9.81%**

(b) **UB of p CI = 0.104 or 10.4%**

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

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

(c)  $Z^* = 1.64$

(d)  $SE = 0.021$

(e)  $|Z_{\text{obs}}| = 1.88$

(f)  $p\text{-value} = 0.0602$

(g) **reject**

1. In a deck of strange cards, there are 1253 cards. Each card has an image and a color. The amounts are shown in the table below.

|        | blue | indigo | orange | red | violet |
|--------|------|--------|--------|-----|--------|
| flower | 14   | 67     | 47     | 40  | 10     |
| pig    | 52   | 98     | 48     | 51  | 20     |
| shovel | 53   | 41     | 90     | 55  | 27     |
| tree   | 70   | 88     | 73     | 21  | 35     |
| wheel  | 36   | 61     | 62     | 78  | 16     |

- (a) What is the probability a random card is both a wheel and orange?
- (b) What is the probability a random card is blue?
- (c) What is the probability a random card is a tree?
- (d) What is the probability a random card is a pig given it is red?
- (e) What is the probability a random card is indigo given it is a tree?
- (f) What is the probability a random card is either a wheel or orange (or both)?

**Solution**

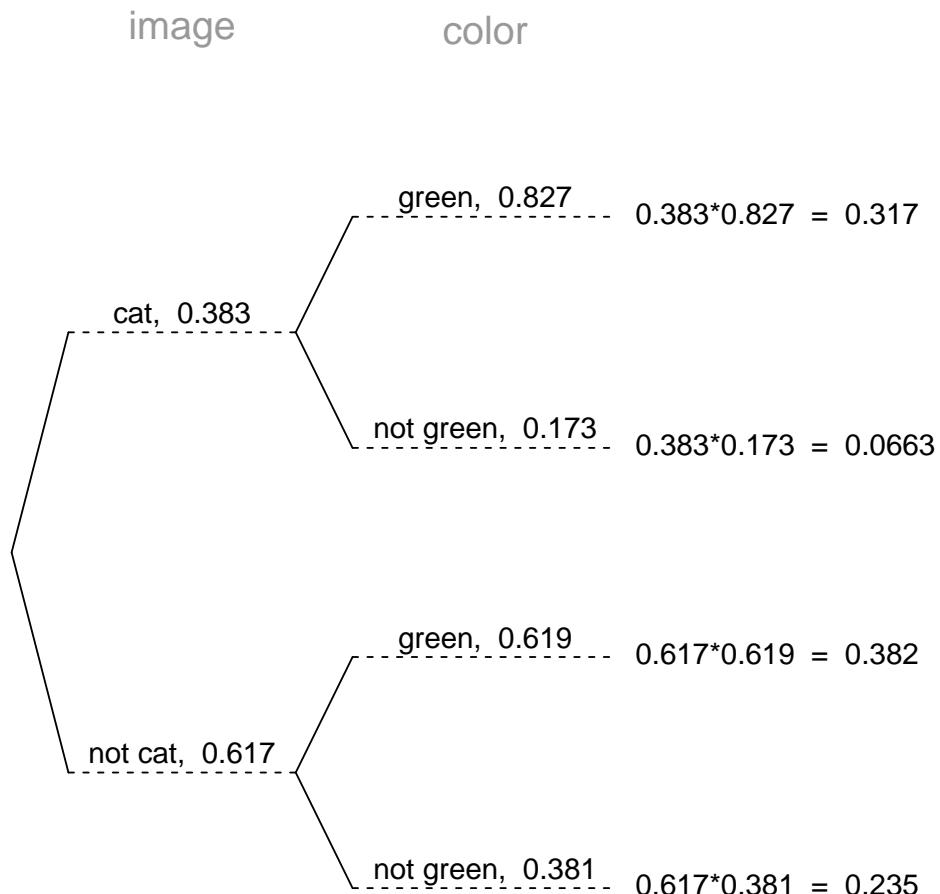
- (a)  $P(\text{wheel and orange}) = \frac{62}{1253} = 0.0495$
- (b)  $P(\text{blue}) = \frac{14+52+53+70+36}{1253} = 0.18$
- (c)  $P(\text{tree}) = \frac{70+88+73+21+35}{1253} = 0.229$
- (d)  $P(\text{pig given red}) = \frac{51}{40+51+55+21+78} = 0.208$
- (e)  $P(\text{indigo given tree}) = \frac{88}{70+88+73+21+35} = 0.307$
- (f)  $P(\text{wheel or orange}) = \frac{36+61+62+78+16+47+48+90+73+62-62}{1253} = 0.408$

2. In a deck of strange cards, each card has an image and a color. The chance of drawing a cat is 38.3%. If a cat is drawn, there is a 82.7% chance that it is green. If a card that is not a cat is drawn, there is a 61.9% chance that it is green.

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

**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.

$$P(\text{"cat" given "not green"}) = \frac{0.0663}{0.0663 + 0.235} = 0.22$$

3. In a very large pile of toothpicks, the mean length is 66.87 millimeters and the standard deviation is 1.79 millimeters. If you randomly sample 175 toothpicks, what is the chance the sample mean is between 66.76 and 66.99 millimeters?

**Solution**

Label the given information.

$$\mu = 66.87$$

$$\sigma = 1.79$$

$$n = 175$$

$$\bar{x}_{\text{lower}} = 66.76$$

$$\bar{x}_{\text{upper}} = 66.99$$

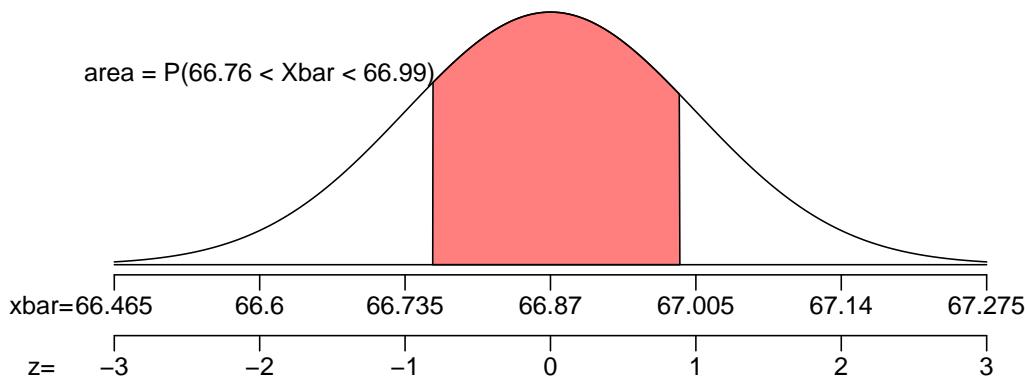
Find the standard error.

$$SE = \frac{\sigma}{\sqrt{n}} = \frac{1.79}{\sqrt{175}} = 0.135$$

Describe the sampling distribution.

$$\bar{X} \sim \mathcal{N}(66.87, 0.135)$$

Draw a sketch.



Calculate a  $z$  scores.

$$z_{\text{lower}} = \frac{x_{\text{lower}} - \mu}{SE} = \frac{66.76 - 66.87}{0.135} = -0.81$$

$$z_{\text{upper}} = \frac{x_{\text{upper}} - \mu}{SE} = \frac{66.99 - 66.87}{0.135} = 0.89$$

Determine the probability.

$$\begin{aligned} P(66.76 < X < 66.99) &= \Phi(z_{\text{upper}}) - \Phi(z_{\text{lower}}) \\ &= \Phi(0.89) - \Phi(-0.81) \\ &= 0.6043 \end{aligned}$$

4. In a game, there is a 20% chance to win a round. You will play 180 rounds.
- (a) What is the probability of winning exactly 37 rounds?
  - (b) What is the probability of winning at least 38 but at most 48 rounds?

**Solution**

We use the formula for binomial probabilities.

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

$$P(X = 37) = \binom{180}{37} (0.2)^{37} (1 - 0.2)^{180-37}$$

$$P(X = 37) = \binom{180}{37} (0.2)^{37} (0.8)^{143}$$

$$P(X = 37) = 0.0722$$

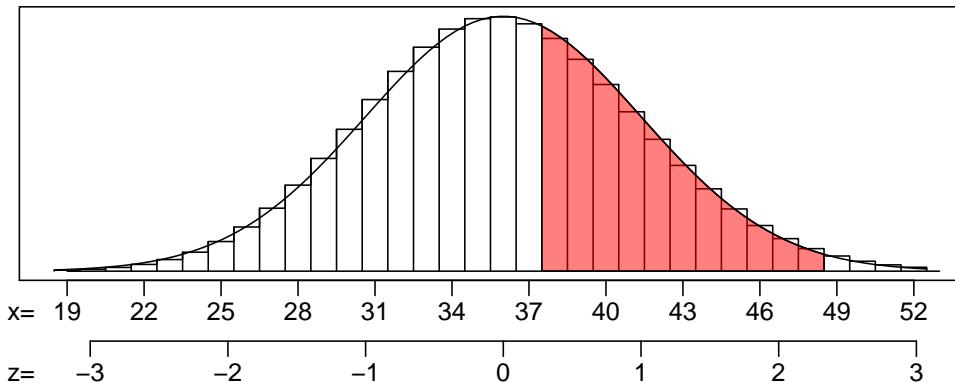
Find the mean.

$$\mu = np = (180)(0.2) = 36$$

Find the standard deviation.

$$\sigma = \sqrt{np(1 - p)} = \sqrt{(180)(0.2)(1 - 0.2)} = 5.3666$$

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



Find the z scores.

$$z_1 = \frac{37.5 - 36}{5.3666} = 0.37$$

$$z_2 = \frac{48.5 - 36}{5.3666} = 2.24$$

Calculate the probability.

$$P(38 \leq X \leq 48) = \Phi(2.24) - \Phi(0.37) = 0.3432$$

(a)  $P(X = 37) = 0.0722$

(b)  $P(38 \leq X \leq 48) = 0.3432$

5. As an ornithologist, you wish to determine the average body mass of *Porzana carolina*. You randomly sample 27 adults of *Porzana carolina*, resulting in a sample mean of 66.97 grams and a sample standard deviation of 9.8 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.

$$n = 27$$

$$\bar{x} = 66.97$$

$$s = 9.8$$

$$CL = 0.95$$

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

$$df = n - 1 = 26$$

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

$$t^* = 2.06$$

Calculate the standard error.

$$SE = \frac{s}{\sqrt{n}} = \frac{9.8}{\sqrt{27}} = 1.89$$

We want to make an inference about the population mean.

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

Determine the bounds.

$$\begin{aligned} CI &= (\bar{x} - t^* SE, \bar{x} + t^* SE) \\ &= (66.97 - 2.06 \times 1.89, 66.97 + 2.06 \times 1.89) \\ &= (63.1, 70.9) \end{aligned}$$

We are 95% confident that the population mean is between 63.1 and 70.9.

6. A treatment group of size 21 has a mean of 108 and standard deviation of 18.3. A control group of size 20 has a mean of 96.8 and standard deviation of 20.4. 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 38.

- (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.

$$n_1 = 21$$

$$\bar{x}_1 = 108$$

$$s_1 = 18.3$$

$$n_2 = 20$$

$$\bar{x}_2 = 96.8$$

$$s_2 = 20.4$$

$$\alpha = 0.04$$

$$df = 38$$

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.

$$t^* = 2.13$$

Calculate the standard error.

$$\begin{aligned} SE &= \sqrt{\frac{(s_1)^2}{n_1} + \frac{(s_2)^2}{n_2}} \\ &= \sqrt{\frac{(18.3)^2}{21} + \frac{(20.4)^2}{20}} \\ &= 6.063 \end{aligned}$$

Determine the test statistic.

$$\begin{aligned} t_{\text{obs}} &= \frac{(\bar{x}_2 - \bar{x}_1) - (\mu_2 - \mu_1)_0}{SE} \\ &= \frac{(96.8 - 108) - (0)}{6.063} \\ &= -1.85 \end{aligned}$$

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

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

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

$$0.05 < p\text{-value} < 0.1$$

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)  $t^* = 2.13$
- (d)  $SE = 6.063$
- (e)  $|t_{\text{obs}}| = 1.85$
- (f)  $0.05 < p\text{-value} < 0.1$
- (g) retain the null

7. From a very large population, a random sample of 59000 individuals was taken. In that sample, 10.1% were asleep. Determine a 98% confidence interval of the population proportion.

- (a) Find the lower bound of the confidence interval.
- (b) Find the upper bound of the condifence interval.

**Solution**

Determine  $z^*$  such that  $P(|Z| < z^*) = 0.98$ .

$$z^* = 2.33$$

Calculate the standard error.

$$SE = \sqrt{\frac{\hat{p}(1 - \hat{p})}{n}} = \sqrt{\frac{(0.101)(1 - 0.101)}{59000}} = 0.00124$$

Calculate the margin of error.

$$ME = z^* SE = (2.33)(0.00124) = 0.00289$$

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.

$$(0.0981, 0.104)$$

We are 98% confident that the true population proportion is between 9.81% and 10.4%.

- (a) The lower bound = 0.0981, which can also be expressed as 9.81%.
- (b) The upper bound = 0.104, which can also be expressed as 10.4%.

8. An experiment is run with a treatment group of size 235 and a control group of size 268. The results are summarized in the table below.

|          | treatment | control |
|----------|-----------|---------|
| cold     | 19        | 11      |
| not cold | 216       | 257     |

Using a significance level of 0.1, determine whether the treatment causes an effect on the proportion of cases that are cold.

- (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  $P(|Z| > z^*) = 0.1$ .

$$z^* = \Phi^{-1} \left( 1 - \frac{\alpha}{2} \right) = 1.64$$

Determine the sample proportions.

$$\hat{p}_1 = \frac{19}{235} = 0.0809$$

$$\hat{p}_2 = \frac{11}{268} = 0.041$$

Determine the difference of sample proportions.

$$\hat{p}_2 - \hat{p}_1 = 0.041 - 0.0809 = -0.0399$$

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

$$\hat{p} = \frac{19 + 11}{235 + 268} = 0.0596$$

Determine the standard error.

$$\begin{aligned} SE &= \sqrt{\frac{\hat{p}(1 - \hat{p})}{n_1} + \frac{\hat{p}(1 - \hat{p})}{n_2}} \\ &= \sqrt{\frac{(0.0596)(0.9404)}{235} + \frac{(0.0596)(0.9404)}{268}} \\ &= 0.0212 \end{aligned}$$

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

$$H_0 : \hat{P}_2 - \hat{P}_1 \sim \mathcal{N}(0, 0.0212)$$

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.

$$\begin{aligned} z &= \frac{(\hat{p}_2 - \hat{p}_1) - (p_2 - p_1)_0}{SE} \\ &= \frac{(0.041 - 0.0809) - 0}{0.0212} \\ &= -1.88 \end{aligned}$$

Determine the  $p$ -value.

$$\begin{aligned} p\text{-value} &= 2 \cdot \Phi(-|z|) \\ &= 2 \cdot \Phi(-1.88) \\ &= 0.0602 \end{aligned}$$

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)  $z^* = 1.64$
- (d)  $SE = 0.0212$
- (e)  $|Z_{\text{obs}}| = 1.88$
- (f)  $p\text{-value} = 0.0602$
- (g) reject the null

# Bunker Hill Community College

Final Statistics Exam 2019-05-02

Exam ID 004

**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)  $P(\text{horn given black}) = 0.4$

(b)  $P(\text{pig or teal}) = 0.614$

(c)  $P(\text{teal given pig}) = 0.358$

(d)  $P(\text{horn and indigo}) = 0.0707$

(e)  $P(\text{black}) = 0.269$

(f)  $P(\text{pig}) = 0.367$

2.  $P(\text{"not horn" given "black"}) = 0.821$

3.  $P(71.36 < X < 71.65) = 0.493$

4. (a)  $P(X = 41) = 0.0718$

(b)  $P(33 \leq X \leq 51) = 0.8808$

5.  $(41.3, 43.5)$

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

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

(c)  $t^* = 2.45$

(d)  $SE = 4.179$

(e)  $|t_{\text{obs}}| = 2.39$

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

(g) **retain**

7. (a) **LB of p CI = 0.0569 or 5.69%**

(b) **UB of p CI = 0.0731 or 7.31%**

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

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

(c)  $Z^* = 2.05$

(d)  $SE = 0.056$

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

(f)  $p\text{-value} = 0.0208$

(g) **reject**

1. In a deck of strange cards, there are 594 cards. Each card has an image and a color. The amounts are shown in the table below.

|        | black | gray | indigo | teal |
|--------|-------|------|--------|------|
| flower | 51    | 44   | 17     | 91   |
| horn   | 64    | 11   | 42     | 56   |
| pig    | 45    | 32   | 63     | 78   |

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

**Solution**

(a)  $P(\text{horn given black}) = \frac{64}{51+64+45} = 0.4$

(b)  $P(\text{pig or teal}) = \frac{45+32+63+78+91+56+78-78}{594} = 0.614$

(c)  $P(\text{teal given pig}) = \frac{78}{45+32+63+78} = 0.358$

(d)  $P(\text{horn and indigo}) = \frac{42}{594} = 0.0707$

(e)  $P(\text{black}) = \frac{51+64+45}{594} = 0.269$

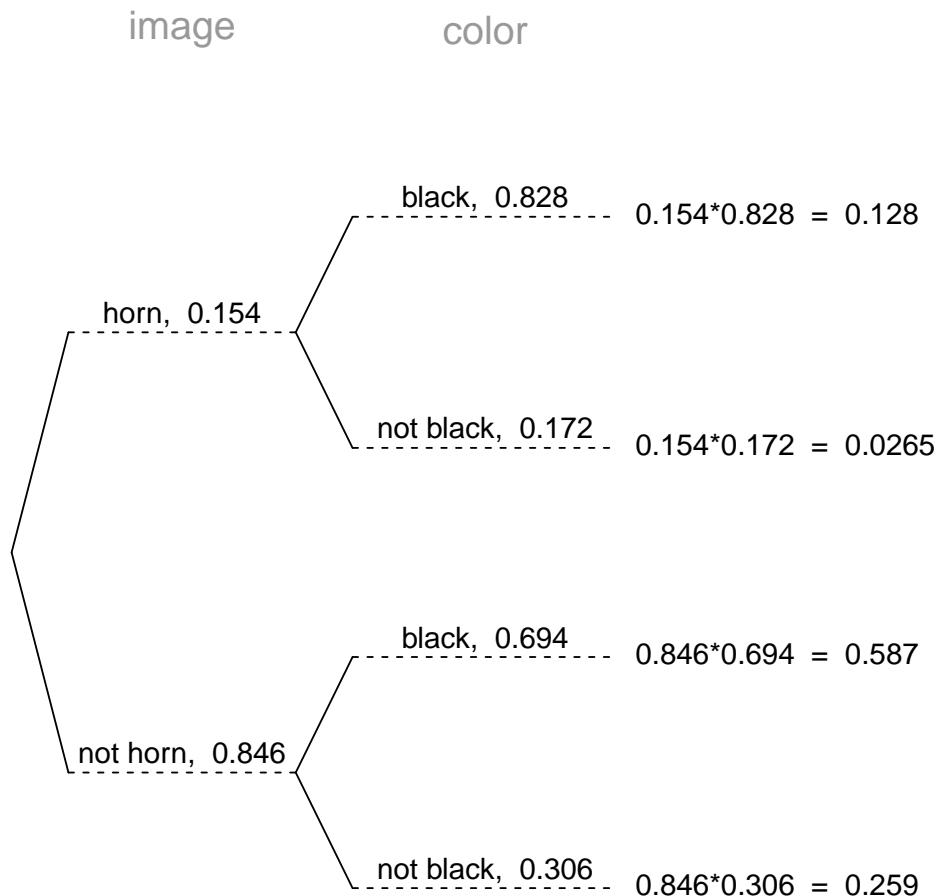
(f)  $P(\text{pig}) = \frac{45+32+63+78}{594} = 0.367$

2. In a deck of strange cards, each card has an image and a color. The chance of drawing a horn is 15.4%. If a horn is drawn, there is a 82.8% chance that it is black. If a card that is not a horn is drawn, there is a 69.4% chance that it is black.

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

**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.

$$P(\text{"not horn" given "black"}) = \frac{0.587}{0.587 + 0.128} = 0.821$$

3. In a very large pile of toothpicks, the mean length is 71.49 millimeters and the standard deviation is 2.84 millimeters. If you randomly sample 169 toothpicks, what is the chance the sample mean is between 71.36 and 71.65 millimeters?

**Solution**

Label the given information.

$$\mu = 71.49$$

$$\sigma = 2.84$$

$$n = 169$$

$$\bar{x}_{\text{lower}} = 71.36$$

$$\bar{x}_{\text{upper}} = 71.65$$

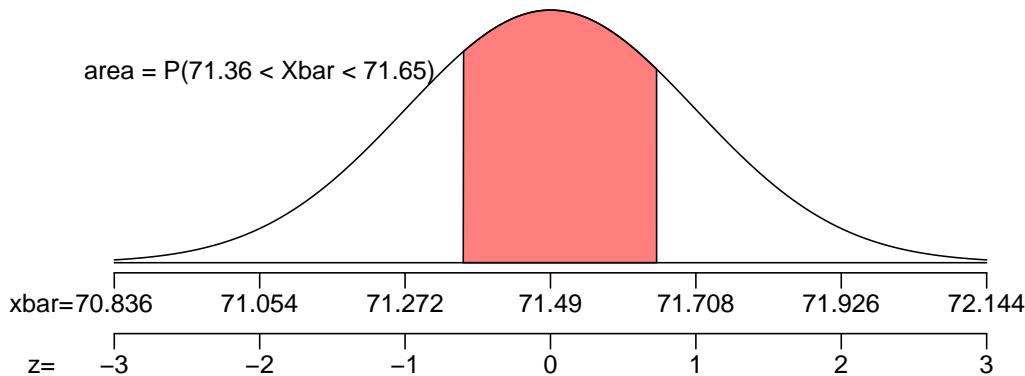
Find the standard error.

$$SE = \frac{\sigma}{\sqrt{n}} = \frac{2.84}{\sqrt{169}} = 0.218$$

Describe the sampling distribution.

$$\bar{X} \sim \mathcal{N}(71.49, 0.218)$$

Draw a sketch.



Calculate a  $z$  scores.

$$z_{\text{lower}} = \frac{x_{\text{lower}} - \mu}{SE} = \frac{71.36 - 71.49}{0.218} = -0.6$$

$$z_{\text{upper}} = \frac{x_{\text{upper}} - \mu}{SE} = \frac{71.65 - 71.49}{0.218} = 0.73$$

Determine the probability.

$$\begin{aligned} P(71.36 < \bar{X} < 71.65) &= \Phi(z_{\text{upper}}) - \Phi(z_{\text{lower}}) \\ &= \Phi(0.73) - \Phi(-0.6) \\ &= 0.493 \end{aligned}$$

4. In a game, there is a 27% chance to win a round. You will play 148 rounds.
- (a) What is the probability of winning exactly 41 rounds?
  - (b) What is the probability of winning at least 33 but at most 51 rounds?

**Solution**

We use the formula for binomial probabilities.

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

$$P(X = 41) = \binom{148}{41} (0.27)^{41} (1 - 0.27)^{148-41}$$

$$P(X = 41) = \binom{148}{41} (0.27)^{41} (0.73)^{107}$$

$$P(X = 41) = 0.0718$$

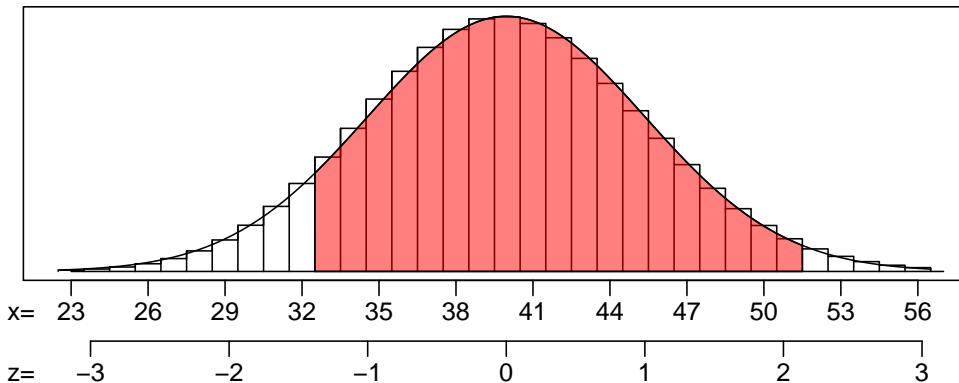
Find the mean.

$$\mu = np = (148)(0.27) = 39.96$$

Find the standard deviation.

$$\sigma = \sqrt{np(1 - p)} = \sqrt{(148)(0.27)(1 - 0.27)} = 5.401$$

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



Find the z scores.

$$z_1 = \frac{32.5 - 39.96}{5.401} = -1.29$$

$$z_2 = \frac{51.5 - 39.96}{5.401} = 2.04$$

Calculate the probability.

$$P(33 \leq X \leq 51) = \Phi(2.04) - \Phi(-1.29) = 0.8808$$

(a)  $P(X = 41) = 0.0718$

(b)  $P(33 \leq X \leq 51) = 0.8808$

5. As an ornithologist, you wish to determine the average body mass of *Catharus fuscescens*. You randomly sample 26 adults of *Catharus fuscescens*, resulting in a sample mean of 42.41 grams and a sample standard deviation of 4.23 grams. Determine a 80% confidence interval of the true population mean.

**Solution**

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

$$n = 26$$

$$\bar{x} = 42.41$$

$$s = 4.23$$

$$CL = 0.8$$

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

$$df = n - 1 = 25$$

Determine the critical  $t$  value,  $t^*$ , such that  $P(|T| < t^*) = 0.8$ .

$$t^* = 1.32$$

Calculate the standard error.

$$SE = \frac{s}{\sqrt{n}} = \frac{4.23}{\sqrt{26}} = 0.83$$

We want to make an inference about the population mean.

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

Determine the bounds.

$$\begin{aligned} CI &= (\bar{x} - t^* SE, \bar{x} + t^* SE) \\ &= (42.41 - 1.32 \times 0.83, 42.41 + 1.32 \times 0.83) \\ &= (41.3, 43.5) \end{aligned}$$

We are 80% confident that the population mean is between 41.3 and 43.5.

6. A treatment group of size 18 has a mean of 110 and standard deviation of 14.7. A control group of size 30 has a mean of 100 and standard deviation of 12.8. 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 32.

- (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.

$$n_1 = 18$$

$$\bar{x}_1 = 110$$

$$s_1 = 14.7$$

$$n_2 = 30$$

$$\bar{x}_2 = 100$$

$$s_2 = 12.8$$

$$\alpha = 0.02$$

$$df = 32$$

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  $P(|T| > t^*) = 0.02$  by using a  $t$  table.

$$t^* = 2.45$$

Calculate the standard error.

$$\begin{aligned} SE &= \sqrt{\frac{(s_1)^2}{n_1} + \frac{(s_2)^2}{n_2}} \\ &= \sqrt{\frac{(14.7)^2}{18} + \frac{(12.8)^2}{30}} \\ &= 4.179 \end{aligned}$$

Determine the test statistic.

$$\begin{aligned} t_{\text{obs}} &= \frac{(\bar{x}_2 - \bar{x}_1) - (\mu_2 - \mu_1)_0}{SE} \\ &= \frac{(100 - 110) - (0)}{4.179} \\ &= -2.39 \end{aligned}$$

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)  $t^* = 2.45$
- (d)  $SE = 4.179$
- (e)  $|t_{\text{obs}}| = 2.39$
- (f)  $0.02 < p\text{-value} < 0.04$
- (g) retain the null

7. From a very large population, a random sample of 6200 individuals was taken. In that sample, 6.5% were purple. Determine a 99% confidence interval of the population proportion.

- (a) Find the lower bound of the confidence interval.
- (b) Find the upper bound of the condifence interval.

**Solution**

Determine  $z^*$  such that  $P(|Z| < z^*) = 0.99$ .

$$z^* = 2.58$$

Calculate the standard error.

$$SE = \sqrt{\frac{\hat{p}(1 - \hat{p})}{n}} = \sqrt{\frac{(0.065)(1 - 0.065)}{6200}} = 0.00313$$

Calculate the margin of error.

$$ME = z^* SE = (2.58)(0.00313) = 0.00808$$

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.

$$(0.0569, 0.0731)$$

We are 99% confident that the true population proportion is between 5.69% and 7.31%.

- (a) The lower bound = 0.0569, which can also be expressed as 5.69%.
- (b) The upper bound = 0.0731, which can also be expressed as 7.31%.

8. An experiment is run with a treatment group of size 144 and a control group of size 186. The results are summarized in the table below.

|             | treatment | control |
|-------------|-----------|---------|
| abysmal     | 59        | 100     |
| not abysmal | 85        | 86      |

Using a significance level of 0.04, determine whether the treatment causes an effect on the proportion of cases that are abysmal.

- (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  $P(|Z| > z^*) = 0.04$ .

$$z^* = \Phi^{-1} \left( 1 - \frac{\alpha}{2} \right) = 2.05$$

Determine the sample proportions.

$$\hat{p}_1 = \frac{59}{144} = 0.41$$

$$\hat{p}_2 = \frac{100}{186} = 0.538$$

Determine the difference of sample proportions.

$$\hat{p}_2 - \hat{p}_1 = 0.538 - 0.41 = 0.128$$

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

$$\hat{p} = \frac{59 + 100}{144 + 186} = 0.482$$

Determine the standard error.

$$\begin{aligned} SE &= \sqrt{\frac{\hat{p}(1 - \hat{p})}{n_1} + \frac{\hat{p}(1 - \hat{p})}{n_2}} \\ &= \sqrt{\frac{(0.482)(0.518)}{144} + \frac{(0.482)(0.518)}{186}} \\ &= 0.0555 \end{aligned}$$

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

$$H_0 : \hat{P}_2 - \hat{P}_1 \sim \mathcal{N}(0, 0.0555)$$

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.

$$\begin{aligned} z &= \frac{(\hat{p}_2 - \hat{p}_1) - (p_2 - p_1)_0}{SE} \\ &= \frac{(0.538 - 0.41) - 0}{0.0555} \\ &= 2.31 \end{aligned}$$

Determine the  $p$ -value.

$$\begin{aligned} p\text{-value} &= 2 \cdot \Phi(-|z|) \\ &= 2 \cdot \Phi(-2.31) \\ &= 0.0208 \end{aligned}$$

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)  $z^* = 2.05$
- (d)  $SE = 0.0555$
- (e)  $|z_{\text{obs}}| = 2.31$
- (f)  $p\text{-value} = 0.0208$
- (g) reject the null

# Bunker Hill Community College

Final Statistics Exam 2019-05-02

Exam ID 005

**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)  $P(\text{flower}) = 0.434$

(b)  $P(\text{wheel or yellow}) = 0.596$

(c)  $P(\text{red}) = 0.281$

(d)  $P(\text{pig given white}) = 0.0855$

(e)  $P(\text{white given flower}) = 0.0901$

(f)  $P(\text{flower and white}) = 0.0391$

2.  $P(\text{"not shovel" given "not white"}) = 0.533$

3.  $P(65.28 < X < 65.85) = 0.839$

4. (a)  $P(X = 14) = 0.1225$

(b)  $P(8 \leq X \leq 21) = 0.9508$

5.  $(24.5, 26.5)$

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

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

(c)  $t^* = 2.02$

(d)  $SE = 0.01$

(e)  $|t_{\text{obs}}| = 2.3$

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

(g) **reject**

7. (a) **LB of p CI = 0.852 or 85.2%**

(b) **UB of p CI = 0.874 or 87.4%**

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

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

(c)  $Z^* = 2.33$

(d)  $SE = 0.043$

(e)  $|Z_{\text{obs}}| = 2.5$

(f)  $p\text{-value} = 0.0124$

(g) **reject**

1. In a deck of strange cards, there are 512 cards. Each card has an image and a color. The amounts are shown in the table below.

|        | indigo | red | white | yellow |
|--------|--------|-----|-------|--------|
| flower | 30     | 91  | 20    | 81     |
| pig    | 27     | 29  | 10    | 48     |
| wheel  | 44     | 24  | 87    | 21     |

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

**Solution**

(a)  $P(\text{flower}) = \frac{30+91+20+81}{512} = 0.434$

(b)  $P(\text{wheel or yellow}) = \frac{44+24+87+21+81+48+21-21}{512} = 0.596$

(c)  $P(\text{red}) = \frac{91+29+24}{512} = 0.281$

(d)  $P(\text{pig given white}) = \frac{10}{20+10+87} = 0.0855$

(e)  $P(\text{white given flower}) = \frac{20}{30+91+20+81} = 0.0901$

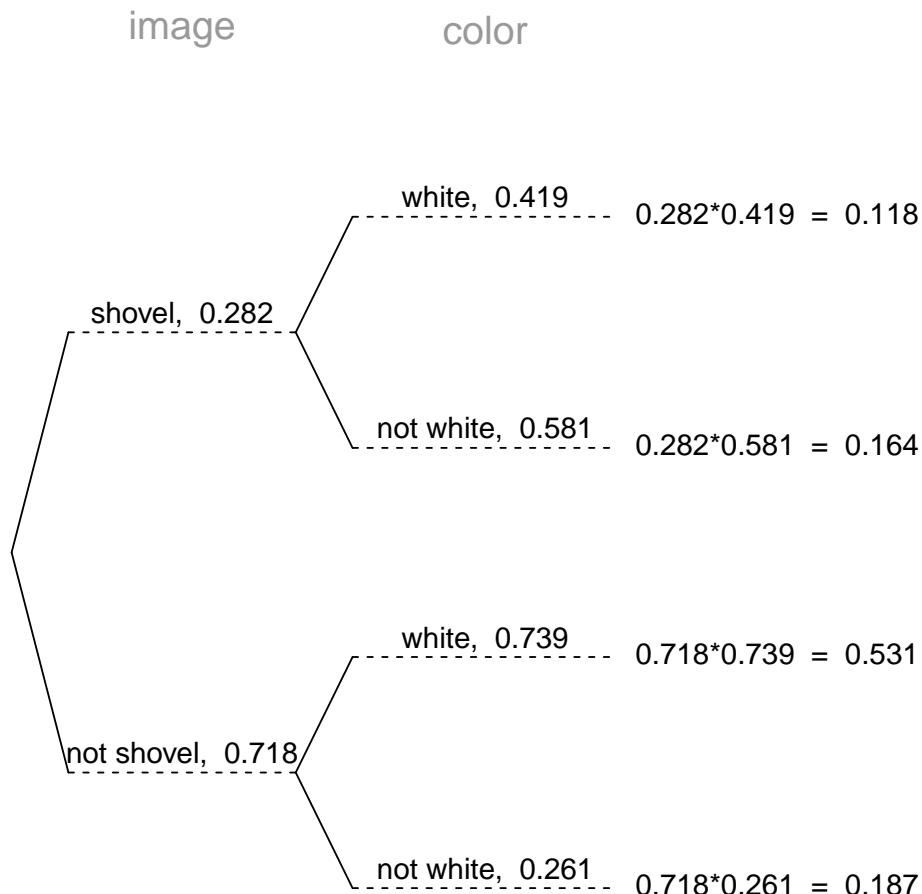
(f)  $P(\text{flower and white}) = \frac{20}{512} = 0.0391$

2. In a deck of strange cards, each card has an image and a color. The chance of drawing a shovel is 28.2%. If a shovel is drawn, there is a 41.9% chance that it is white. If a card that is not a shovel is drawn, there is a 73.9% chance that it is white.

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

**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.

$$P(\text{"not shovel" given "not white"}) = \frac{0.187}{0.187 + 0.164} = 0.533$$

3. In a very large pile of toothpicks, the mean length is 65.69 millimeters and the standard deviation is 1.94 millimeters. If you randomly sample 150 toothpicks, what is the chance the sample mean is between 65.28 and 65.85 millimeters?

**Solution**

Label the given information.

$$\mu = 65.69$$

$$\sigma = 1.94$$

$$n = 150$$

$$\bar{x}_{\text{lower}} = 65.28$$

$$\bar{x}_{\text{upper}} = 65.85$$

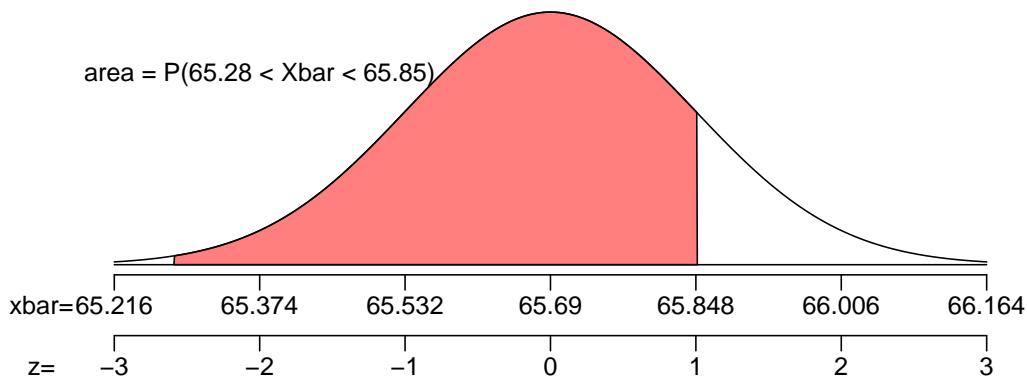
Find the standard error.

$$SE = \frac{\sigma}{\sqrt{n}} = \frac{1.94}{\sqrt{150}} = 0.158$$

Describe the sampling distribution.

$$\bar{X} \sim \mathcal{N}(65.69, 0.158)$$

Draw a sketch.



Calculate a  $z$  scores.

$$z_{\text{lower}} = \frac{\bar{x}_{\text{lower}} - \mu}{SE} = \frac{65.28 - 65.69}{0.158} = -2.59$$

$$z_{\text{upper}} = \frac{\bar{x}_{\text{upper}} - \mu}{SE} = \frac{65.85 - 65.69}{0.158} = 1.01$$

Determine the probability.

$$\begin{aligned} P(65.28 < \bar{X} < 65.85) &= \Phi(z_{\text{upper}}) - \Phi(z_{\text{lower}}) \\ &= \Phi(1.01) - \Phi(-2.59) \\ &= 0.839 \end{aligned}$$

4. In a game, there is a 26% chance to win a round. You will play 52 rounds.
- (a) What is the probability of winning exactly 14 rounds?
  - (b) What is the probability of winning at least 8 but at most 21 rounds?

**Solution**

We use the formula for binomial probabilities.

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

$$P(X = 14) = \binom{52}{14} (0.26)^{14} (1 - 0.26)^{52-14}$$

$$P(X = 14) = \binom{52}{14} (0.26)^{14} (0.74)^{38}$$

$$P(X = 14) = 0.1225$$

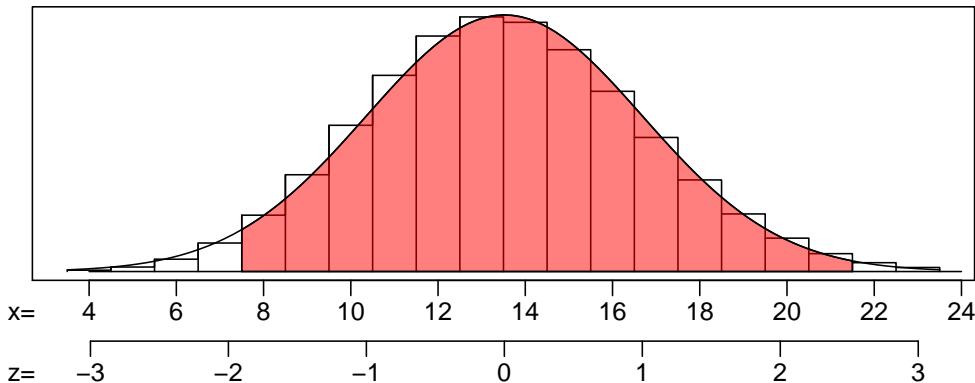
Find the mean.

$$\mu = np = (52)(0.26) = 13.52$$

Find the standard deviation.

$$\sigma = \sqrt{np(1 - p)} = \sqrt{(52)(0.26)(1 - 0.26)} = 3.163$$

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



Find the z scores.

$$z_1 = \frac{7.5 - 13.52}{3.163} = -1.75$$

$$z_2 = \frac{21.5 - 13.52}{3.163} = 2.36$$

Calculate the probability.

$$P(8 \leq X \leq 21) = \Phi(2.36) - \Phi(-1.75) = 0.9508$$

(a)  $P(X = 14) = 0.1225$

(b)  $P(8 \leq X \leq 21) = 0.9508$

5. As an ornithologist, you wish to determine the average body mass of *Passer domesticus*. You randomly sample 25 adults of *Passer domesticus*, resulting in a sample mean of 25.5 grams and a sample standard deviation of 2.47 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.

$$n = 25$$

$$\bar{x} = 25.5$$

$$s = 2.47$$

$$CL = 0.95$$

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

$$df = n - 1 = 24$$

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

$$t^* = 2.06$$

Calculate the standard error.

$$SE = \frac{s}{\sqrt{n}} = \frac{2.47}{\sqrt{25}} = 0.494$$

We want to make an inference about the population mean.

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

Determine the bounds.

$$\begin{aligned} CI &= (\bar{x} - t^* SE, \bar{x} + t^* SE) \\ &= (25.5 - 2.06 \times 0.494, 25.5 + 2.06 \times 0.494) \\ &= (24.5, 26.5) \end{aligned}$$

We are 95% confident that the population mean is between 24.5 and 26.5.

6. A treatment group of size 20 has a mean of 1.02 and standard deviation of 0.0297. A control group of size 25 has a mean of 0.998 and standard deviation of 0.0345. If you decided to use a significance level of 0.05, 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 42.

- (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.

$$n_1 = 20$$

$$\bar{x}_1 = 1.02$$

$$s_1 = 0.0297$$

$$n_2 = 25$$

$$\bar{x}_2 = 0.998$$

$$s_2 = 0.0345$$

$$\alpha = 0.05$$

$$df = 42$$

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  $P(|T| > t^*) = 0.05$  by using a  $t$  table.

$$t^* = 2.02$$

Calculate the standard error.

$$\begin{aligned} SE &= \sqrt{\frac{(s_1)^2}{n_1} + \frac{(s_2)^2}{n_2}} \\ &= \sqrt{\frac{(0.0297)^2}{20} + \frac{(0.0345)^2}{25}} \\ &= 0.01 \end{aligned}$$

Determine the test statistic.

$$\begin{aligned} t_{\text{obs}} &= \frac{(\bar{x}_2 - \bar{x}_1) - (\mu_2 - \mu_1)_0}{SE} \\ &= \frac{(0.998 - 1.02) - (0)}{0.01} \\ &= -2.3 \end{aligned}$$

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)  $t^* = 2.02$
- (d)  $SE = 0.01$
- (e)  $|t_{\text{obs}}| = 2.3$
- (f)  $0.02 < p\text{-value} < 0.04$
- (g) reject the null

7. From a very large population, a random sample of 5200 individuals was taken. In that sample, 86.3% were glowing. Determine a 98% confidence interval of the population proportion.

- (a) Find the lower bound of the confidence interval.
- (b) Find the upper bound of the condifence interval.

**Solution**

Determine  $z^*$  such that  $P(|Z| < z^*) = 0.98$ .

$$z^* = 2.33$$

Calculate the standard error.

$$SE = \sqrt{\frac{\hat{p}(1 - \hat{p})}{n}} = \sqrt{\frac{(0.863)(1 - 0.863)}{5200}} = 0.00477$$

Calculate the margin of error.

$$ME = z^* SE = (2.33)(0.00477) = 0.0111$$

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.

$$(0.852, 0.874)$$

We are 98% confident that the true population proportion is between 85.2% and 87.4%.

- (a) The lower bound = 0.852, which can also be expressed as 85.2%.
- (b) The upper bound = 0.874, which can also be expressed as 87.4%.

8. An experiment is run with a treatment group of size 259 and a control group of size 227. The results are summarized in the table below.

|             | treatment | control |
|-------------|-----------|---------|
| special     | 183       | 136     |
| not special | 76        | 91      |

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.

$$\hat{p}_1 = \frac{183}{259} = 0.707$$

$$\hat{p}_2 = \frac{136}{227} = 0.599$$

Determine the difference of sample proportions.

$$\hat{p}_2 - \hat{p}_1 = 0.599 - 0.707 = -0.108$$

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

$$\hat{p} = \frac{183 + 136}{259 + 227} = 0.656$$

Determine the standard error.

$$\begin{aligned} SE &= \sqrt{\frac{\hat{p}(1 - \hat{p})}{n_1} + \frac{\hat{p}(1 - \hat{p})}{n_2}} \\ &= \sqrt{\frac{(0.656)(0.344)}{259} + \frac{(0.656)(0.344)}{227}} \\ &= 0.0432 \end{aligned}$$

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

$$H_0 : \hat{P}_2 - \hat{P}_1 \sim \mathcal{N}(0, 0.0432)$$

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.

$$\begin{aligned} z &= \frac{(\hat{p}_2 - \hat{p}_1) - (p_2 - p_1)_0}{SE} \\ &= \frac{(0.599 - 0.707) - 0}{0.0432} \\ &= -2.5 \end{aligned}$$

Determine the  $p$ -value.

$$\begin{aligned} p\text{-value} &= 2 \cdot \Phi(-|z|) \\ &= 2 \cdot \Phi(-2.5) \\ &= 0.0124 \end{aligned}$$

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)  $SE = 0.0432$
- (e)  $|z_{\text{obs}}| = 2.5$
- (f)  $p\text{-value} = 0.0124$
- (g) reject the null

# Bunker Hill Community College

Final Statistics Exam 2019-05-02

Exam ID 006

**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)  $P(\text{shovel}) = 0.247$

(b)  $P(\text{gray given bike}) = 0.311$

(c)  $P(\text{gem and orange}) = 0.0291$

(d)  $P(\text{bike given yellow}) = 0.239$

(e)  $P(\text{gem or violet}) = 0.444$

(f)  $P(\text{violet}) = 0.258$

2.  $P(\text{"not shovel" given "not blue"}) = 0.777$

3.  $P(69.04 < X < 69.24) = 0.6538$

4. (a)  $P(X = 59) = 0.0411$

(b)  $P(70 \leq X \leq 76) = 0.1751$

5. **(12.5, 13.2)**

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

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

(c)  $t^* = 1.72$

(d)  $SE = 0.489$

(e)  $|t_{\text{obs}}| = 1.43$

(f)  $0.1 < p\text{-value} < 0.2$

(g) **retain**

7. (a) **LB of p CI = 0.715 or 71.5%**

(b) **UB of p CI = 0.743 or 74.3%**

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

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

(c)  $Z^* = 1.28$

(d)  $SE = 0.038$

(e)  $|Z_{\text{obs}}| = 1.39$

(f)  $p\text{-value} = 0.1646$

(g) **reject**

1. In a deck of strange cards, there are 928 cards. Each card has an image and a color. The amounts are shown in the table below.

|        | gray | orange | violet | yellow |
|--------|------|--------|--------|--------|
| bike   | 79   | 86     | 37     | 52     |
| cat    | 50   | 53     | 69     | 54     |
| gem    | 56   | 27     | 46     | 90     |
| shovel | 61   | 59     | 87     | 22     |

- (a) What is the probability a random card is a shovel?
- (b) What is the probability a random card is gray given it is a bike?
- (c) What is the probability a random card is both a gem and orange?
- (d) What is the probability a random card is a bike given it is yellow?
- (e) What is the probability a random card is either a gem or violet (or both)?
- (f) What is the probability a random card is violet?

**Solution**

(a)  $P(\text{shovel}) = \frac{61+59+87+22}{928} = 0.247$

(b)  $P(\text{gray given bike}) = \frac{79}{79+86+37+52} = 0.311$

(c)  $P(\text{gem and orange}) = \frac{27}{928} = 0.0291$

(d)  $P(\text{bike given yellow}) = \frac{52}{52+54+90+22} = 0.239$

(e)  $P(\text{gem or violet}) = \frac{56+27+46+90+37+69+46+87-46}{928} = 0.444$

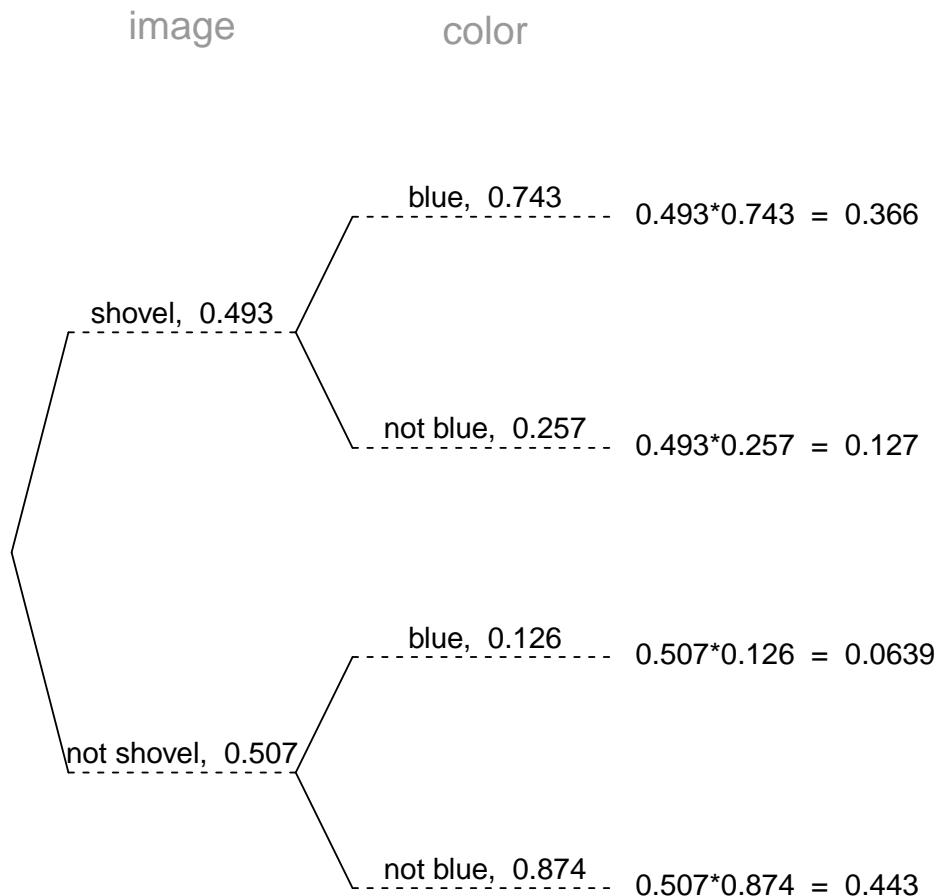
(f)  $P(\text{violet}) = \frac{37+69+46+87}{928} = 0.258$

2. In a deck of strange cards, each card has an image and a color. The chance of drawing a shovel is 49.3%. If a shovel is drawn, there is a 74.3% chance that it is blue. If a card that is not a shovel is drawn, there is a 12.6% chance that it is blue.

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

**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.

$$P(\text{"not shovel" given "not blue"}) = \frac{0.443}{0.443 + 0.127} = 0.777$$

3. In a very large pile of toothpicks, the mean length is 69.19 millimeters and the standard deviation is 1.01 millimeters. If you randomly sample 120 toothpicks, what is the chance the sample mean is between 69.04 and 69.24 millimeters?

**Solution**

Label the given information.

$$\mu = 69.19$$

$$\sigma = 1.01$$

$$n = 120$$

$$\bar{x}_{\text{lower}} = 69.04$$

$$\bar{x}_{\text{upper}} = 69.24$$

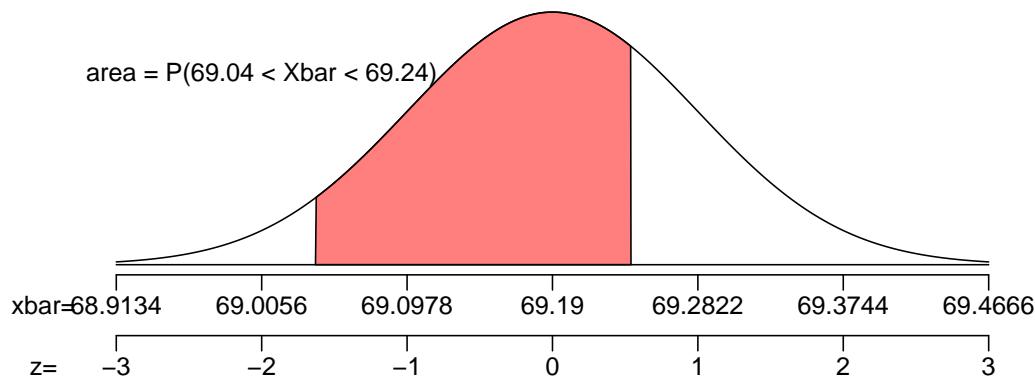
Find the standard error.

$$SE = \frac{\sigma}{\sqrt{n}} = \frac{1.01}{\sqrt{120}} = 0.0922$$

Describe the sampling distribution.

$$\bar{X} \sim \mathcal{N}(69.19, 0.0922)$$

Draw a sketch.



Calculate a  $z$  scores.

$$z_{\text{lower}} = \frac{\bar{x}_{\text{lower}} - \mu}{SE} = \frac{69.04 - 69.19}{0.0922} = -1.63$$

$$z_{\text{upper}} = \frac{\bar{x}_{\text{upper}} - \mu}{SE} = \frac{69.24 - 69.19}{0.0922} = 0.54$$

Determine the probability.

$$\begin{aligned} P(69.04 < \bar{X} < 69.24) &= \Phi(z_{\text{upper}}) - \Phi(z_{\text{lower}}) \\ &= \Phi(0.54) - \Phi(-1.63) \\ &= 0.6538 \end{aligned}$$

4. In a game, there is a 34% chance to win a round. You will play 191 rounds.
- (a) What is the probability of winning exactly 59 rounds?
  - (b) What is the probability of winning at least 70 but at most 76 rounds?

**Solution**

We use the formula for binomial probabilities.

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

$$P(X = 59) = \binom{191}{59} (0.34)^{59} (1 - 0.34)^{191-59}$$

$$P(X = 59) = \binom{191}{59} (0.34)^{59} (0.66)^{132}$$

$$P(X = 59) = 0.0411$$

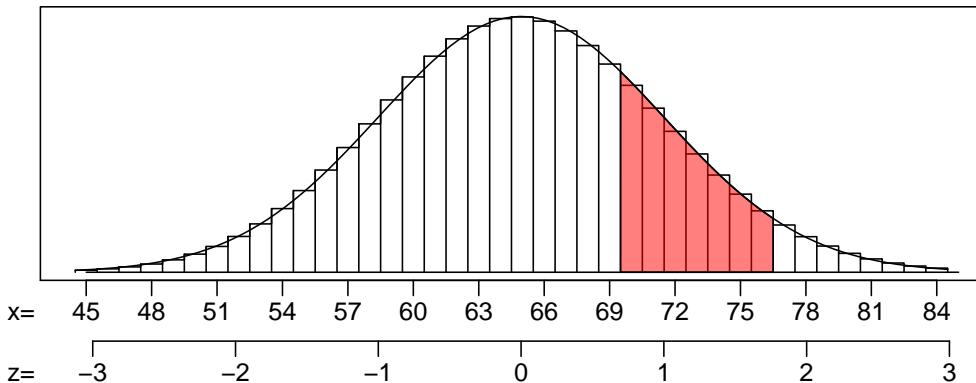
Find the mean.

$$\mu = np = (191)(0.34) = 64.94$$

Find the standard deviation.

$$\sigma = \sqrt{np(1 - p)} = \sqrt{(191)(0.34)(1 - 0.34)} = 6.5468$$

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



Find the z scores.

$$z_1 = \frac{69.5 - 64.94}{6.5468} = 0.77$$

$$z_2 = \frac{76.5 - 64.94}{6.5468} = 1.69$$

Calculate the probability.

$$P(70 \leq X \leq 76) = \Phi(1.69) - \Phi(0.77) = 0.1751$$

(a)  $P(X = 59) = 0.0411$

(b)  $P(70 \leq X \leq 76) = 0.1751$

5. As an ornithologist, you wish to determine the average body mass of *Vermivora peregrina*. You randomly sample 35 adults of *Vermivora peregrina*, resulting in a sample mean of 12.82 grams and a sample standard deviation of 1.26 grams. Determine a 90% confidence interval of the true population mean.

**Solution**

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

$$n = 35$$

$$\bar{x} = 12.82$$

$$s = 1.26$$

$$CL = 0.9$$

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

$$df = n - 1 = 34$$

Determine the critical  $t$  value,  $t^*$ , such that  $P(|T| < t^*) = 0.9$ .

$$t^* = 1.69$$

Calculate the standard error.

$$SE = \frac{s}{\sqrt{n}} = \frac{1.26}{\sqrt{35}} = 0.213$$

We want to make an inference about the population mean.

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

Determine the bounds.

$$\begin{aligned} CI &= (\bar{x} - t^* SE, \bar{x} + t^* SE) \\ &= (12.82 - 1.69 \times 0.213, 12.82 + 1.69 \times 0.213) \\ &= (12.5, 13.2) \end{aligned}$$

We are 90% confident that the population mean is between 12.5 and 13.2.

6. A treatment group of size 17 has a mean of 10.5 and standard deviation of 1.04. A control group of size 13 has a mean of 11.2 and standard deviation of 1.51. If you decided to use a significance level of 0.1, 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 20.

- (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.

$$n_1 = 17$$

$$\bar{x}_1 = 10.5$$

$$s_1 = 1.04$$

$$n_2 = 13$$

$$\bar{x}_2 = 11.2$$

$$s_2 = 1.51$$

$$\alpha = 0.1$$

$$df = 20$$

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  $P(|T| > t^*) = 0.1$  by using a  $t$  table.

$$t^* = 1.72$$

Calculate the standard error.

$$\begin{aligned} SE &= \sqrt{\frac{(s_1)^2}{n_1} + \frac{(s_2)^2}{n_2}} \\ &= \sqrt{\frac{(1.04)^2}{17} + \frac{(1.51)^2}{13}} \\ &= 0.489 \end{aligned}$$

Determine the test statistic.

$$\begin{aligned} t_{\text{obs}} &= \frac{(\bar{x}_2 - \bar{x}_1) - (\mu_2 - \mu_1)_0}{SE} \\ &= \frac{(11.2 - 10.5) - (0)}{0.489} \\ &= 1.43 \end{aligned}$$

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

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

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

$$0.1 < p\text{-value} < 0.2$$

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)  $t^* = 1.72$
- (d)  $SE = 0.489$
- (e)  $|t_{\text{obs}}| = 1.43$
- (f)  $0.1 < p\text{-value} < 0.2$
- (g) retain the null

7. From a very large population, a random sample of 3800 individuals was taken. In that sample, 72.9% were super. 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 condifence interval.

**Solution**

Determine  $z^*$  such that  $P(|Z| < z^*) = 0.95$ .

$$z^* = 1.96$$

Calculate the standard error.

$$SE = \sqrt{\frac{\hat{p}(1 - \hat{p})}{n}} = \sqrt{\frac{(0.729)(1 - 0.729)}{3800}} = 0.00721$$

Calculate the margin of error.

$$ME = z^* SE = (1.96)(0.00721) = 0.0141$$

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.

$$(0.715, 0.743)$$

We are 95% confident that the true population proportion is between 71.5% and 74.3%.

- (a) The lower bound = 0.715, which can also be expressed as 71.5%.
- (b) The upper bound = 0.743, which can also be expressed as 74.3%.

8. An experiment is run with a treatment group of size 167 and a control group of size 135. The results are summarized in the table below.

|                | treatment | control |
|----------------|-----------|---------|
| omnivorous     | 16        | 20      |
| not omnivorous | 151       | 115     |

Using a significance level of 0.2, determine whether the treatment causes an effect on the proportion of cases that are omnivorous.

- (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  $P(|Z| > z^*) = 0.2$ .

$$z^* = \Phi^{-1} \left( 1 - \frac{\alpha}{2} \right) = 1.28$$

Determine the sample proportions.

$$\hat{p}_1 = \frac{16}{167} = 0.0958$$

$$\hat{p}_2 = \frac{20}{135} = 0.148$$

Determine the difference of sample proportions.

$$\hat{p}_2 - \hat{p}_1 = 0.148 - 0.0958 = 0.0522$$

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

$$\hat{p} = \frac{16 + 20}{167 + 135} = 0.119$$

Determine the standard error.

$$\begin{aligned} SE &= \sqrt{\frac{\hat{p}(1 - \hat{p})}{n_1} + \frac{\hat{p}(1 - \hat{p})}{n_2}} \\ &= \sqrt{\frac{(0.119)(0.881)}{167} + \frac{(0.119)(0.881)}{135}} \\ &= 0.0375 \end{aligned}$$

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

$$H_0 : \hat{P}_2 - \hat{P}_1 \sim \mathcal{N}(0, 0.0375)$$

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.

$$\begin{aligned} z &= \frac{(\hat{p}_2 - \hat{p}_1) - (p_2 - p_1)_0}{SE} \\ &= \frac{(0.148 - 0.0958) - 0}{0.0375} \\ &= 1.39 \end{aligned}$$

Determine the  $p$ -value.

$$\begin{aligned} p\text{-value} &= 2 \cdot \Phi(-|z|) \\ &= 2 \cdot \Phi(-1.39) \\ &= 0.1646 \end{aligned}$$

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)  $z^* = 1.28$
- (d)  $SE = 0.0375$
- (e)  $|Z_{\text{obs}}| = 1.39$
- (f)  $p\text{-value} = 0.1646$
- (g) reject the null

# Bunker Hill Community College

Final Statistics Exam 2019-05-02

Exam ID 007

**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)  $P(\text{white}) = 0.0828$

(b)  $P(\text{gem given blue}) = 0.18$

(c)  $P(\text{yellow given horn}) = 0.269$

(d)  $P(\text{gem or yellow}) = 0.505$

(e)  $P(\text{gem and blue}) = 0.0366$

(f)  $P(\text{flower}) = 0.389$

2.  $P(\text{"not horn" given "blue"}) = 0.691$

3.  $P(65.96 < X < 66.74) = 0.7792$

4. (a)  $P(X = 51) = 0.1022$

(b)  $P(49 \leq X \leq 59) = 0.749$

5. **(42.1, 49.1)**

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

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

(c)  $t^* = 3.05$

(d)  $SE = 0.389$

(e)  $|t_{\text{obs}}| = 2.88$

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

(g) **retain**

7. (a) **LB of p CI = 0.0893 or 8.93%**

(b) **UB of p CI = 0.101 or 10.1%**

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

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

(c)  $z^* = 2.81$

(d)  $SE = 0$

(e)  $|Z_{\text{obs}}| = \text{NaN}$

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

(g) **reject**

1. In a deck of strange cards, there are 628 cards. Each card has an image and a color. The amounts are shown in the table below.

|        | blue | indigo | white | yellow |
|--------|------|--------|-------|--------|
| flower | 61   | 85     | 10    | 88     |
| gem    | 23   | 40     | 28    | 81     |
| horn   | 44   | 97     | 14    | 57     |

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

**Solution**

(a)  $P(\text{white}) = \frac{10+28+14}{628} = 0.0828$

(b)  $P(\text{gem given blue}) = \frac{23}{61+23+44} = 0.18$

(c)  $P(\text{yellow given horn}) = \frac{57}{44+97+14+57} = 0.269$

(d)  $P(\text{gem or yellow}) = \frac{23+40+28+81+88+81+57-81}{628} = 0.505$

(e)  $P(\text{gem and blue}) = \frac{23}{628} = 0.0366$

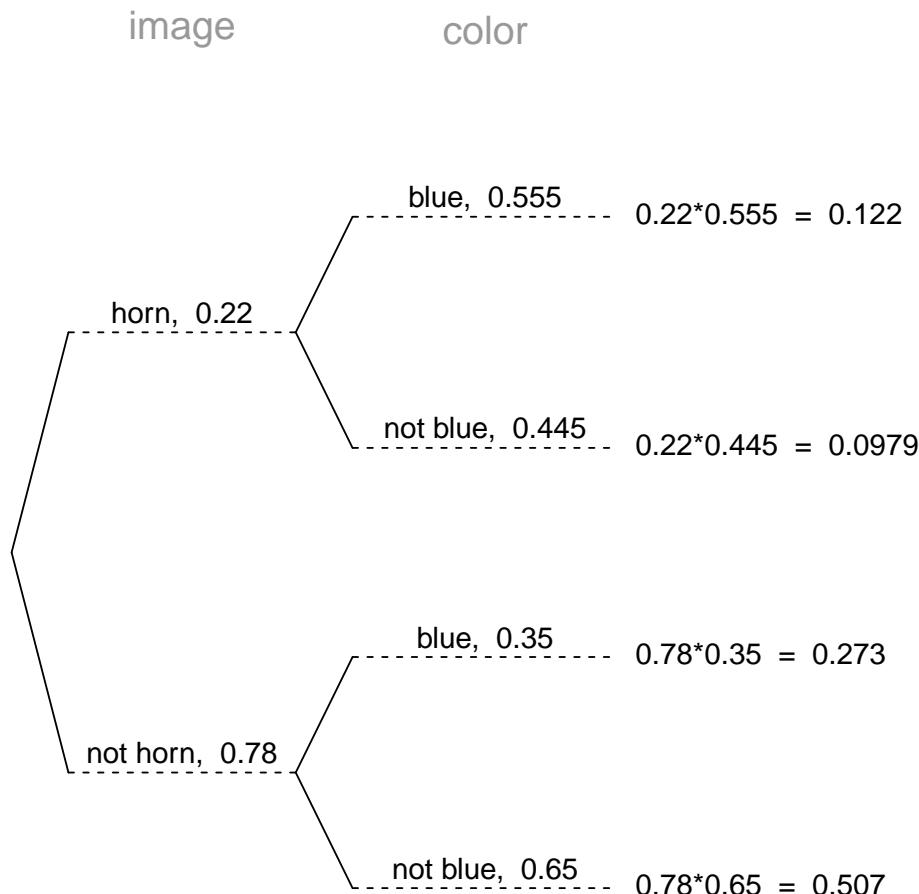
(f)  $P(\text{flower}) = \frac{61+85+10+88}{628} = 0.389$

2. In a deck of strange cards, each card has an image and a color. The chance of drawing a horn is 22%. If a horn is drawn, there is a 55.5% chance that it is blue. If a card that is not a horn is drawn, there is a 35% chance that it is blue.

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

**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.

$$P(\text{"not horn" given "blue"}) = \frac{0.273}{0.273 + 0.122} = 0.691$$

3. In a very large pile of toothpicks, the mean length is 66.22 millimeters and the standard deviation is 3.19 millimeters. If you randomly sample 121 toothpicks, what is the chance the sample mean is between 65.96 and 66.74 millimeters?

**Solution**

Label the given information.

$$\mu = 66.22$$

$$\sigma = 3.19$$

$$n = 121$$

$$\bar{x}_{\text{lower}} = 65.96$$

$$\bar{x}_{\text{upper}} = 66.74$$

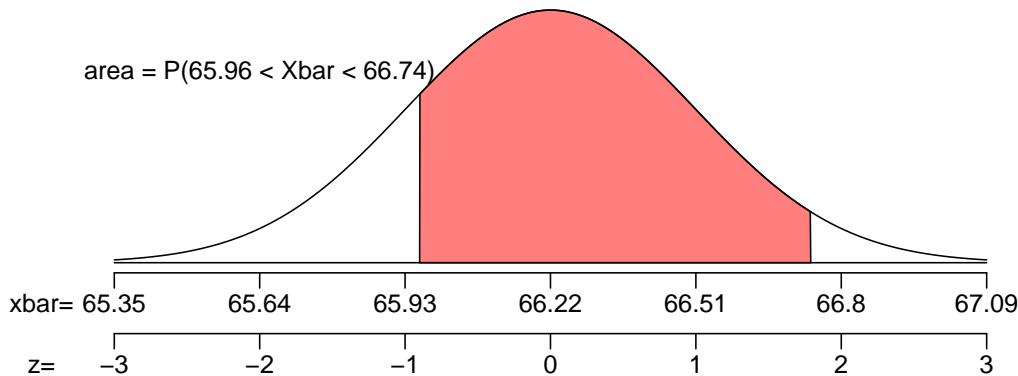
Find the standard error.

$$SE = \frac{\sigma}{\sqrt{n}} = \frac{3.19}{\sqrt{121}} = 0.29$$

Describe the sampling distribution.

$$\bar{X} \sim \mathcal{N}(66.22, 0.29)$$

Draw a sketch.



Calculate a  $z$  scores.

$$z_{\text{lower}} = \frac{\bar{x}_{\text{lower}} - \mu}{SE} = \frac{65.96 - 66.22}{0.29} = -0.9$$

$$z_{\text{upper}} = \frac{\bar{x}_{\text{upper}} - \mu}{SE} = \frac{66.74 - 66.22}{0.29} = 1.79$$

Determine the probability.

$$\begin{aligned} P(65.96 < X < 66.74) &= \Phi(z_{\text{upper}}) - \Phi(z_{\text{lower}}) \\ &= \Phi(1.79) - \Phi(-0.9) \\ &= 0.7792 \end{aligned}$$

4. In a game, there is a 73% chance to win a round. You will play 71 rounds.
- (a) What is the probability of winning exactly 51 rounds?
  - (b) What is the probability of winning at least 49 but at most 59 rounds?

**Solution**

We use the formula for binomial probabilities.

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

$$P(X = 51) = \binom{71}{51} (0.73)^{51} (1 - 0.73)^{71-51}$$

$$P(X = 51) = \binom{71}{51} (0.73)^{51} (0.27)^{20}$$

$$P(X = 51) = 0.1022$$

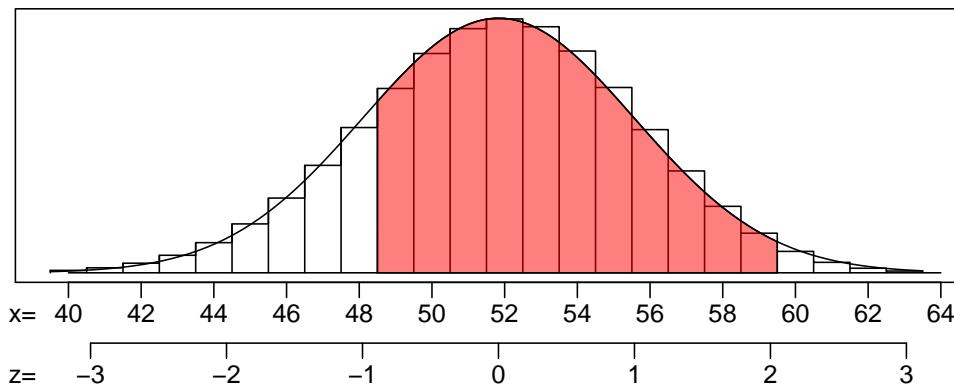
Find the mean.

$$\mu = np = (71)(0.73) = 51.83$$

Find the standard deviation.

$$\sigma = \sqrt{np(1 - p)} = \sqrt{(71)(0.73)(1 - 0.73)} = 3.7409$$

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



Find the z scores.

$$z_1 = \frac{48.5 - 51.83}{3.7409} = -0.76$$

$$z_2 = \frac{59.5 - 51.83}{3.7409} = 1.92$$

Calculate the probability.

$$P(49 \leq X \leq 59) = \Phi(1.92) - \Phi(-0.76) = 0.749$$

(a)  $P(X = 51) = 0.1022$

(b)  $P(49 \leq X \leq 59) = 0.749$

5. As an ornithologist, you wish to determine the average body mass of *Agelaius Phoeniceus*. You randomly sample 25 adults of *Agelaius Phoeniceus*, resulting in a sample mean of 45.6 grams and a sample standard deviation of 5.6 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.

$$n = 25$$

$$\bar{x} = 45.6$$

$$s = 5.6$$

$$CL = 0.995$$

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

$$df = n - 1 = 24$$

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

$$t^* = 3.09$$

Calculate the standard error.

$$SE = \frac{s}{\sqrt{n}} = \frac{5.6}{\sqrt{25}} = 1.12$$

We want to make an inference about the population mean.

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

Determine the bounds.

$$\begin{aligned} CI &= (\bar{x} - t^* SE, \bar{x} + t^* SE) \\ &= (45.6 - 3.09 \times 1.12, 45.6 + 3.09 \times 1.12) \\ &= (42.1, 49.1) \end{aligned}$$

We are 99.5% confident that the population mean is between 42.1 and 49.1.

6. A treatment group of size 31 has a mean of 11 and standard deviation of 0.98. A control group of size 9 has a mean of 9.88 and standard deviation of 1.04. If you decided to use a significance level of 0.01, 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.

$$n_1 = 31$$

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

$$s_1 = 0.98$$

$$n_2 = 9$$

$$\bar{x}_2 = 9.88$$

$$s_2 = 1.04$$

$$\alpha = 0.01$$

$$df = 12$$

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  $P(|T| > t^*) = 0.01$  by using a  $t$  table.

$$t^* = 3.05$$

Calculate the standard error.

$$\begin{aligned} SE &= \sqrt{\frac{(s_1)^2}{n_1} + \frac{(s_2)^2}{n_2}} \\ &= \sqrt{\frac{(0.98)^2}{31} + \frac{(1.04)^2}{9}} \\ &= 0.389 \end{aligned}$$

Determine the test statistic.

$$\begin{aligned} t_{\text{obs}} &= \frac{(\bar{x}_2 - \bar{x}_1) - (\mu_2 - \mu_1)_0}{SE} \\ &= \frac{(9.88 - 11) - (0)}{0.389} \\ &= -2.88 \end{aligned}$$

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 retain the null hypothesis.

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

- (b)  $H_A : \mu_2 - \mu_1 \neq 0$
- (c)  $t^* = 3.05$
- (d)  $SE = 0.389$
- (e)  $|t_{\text{obs}}| = 2.88$
- (f)  $0.01 < p\text{-value} < 0.02$
- (g) retain the null

7. From a very large population, a random sample of 4400 individuals was taken. In that sample, 9.5% were salty. 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 condifence interval.

**Solution**

Determine  $z^*$  such that  $P(|Z| < z^*) = 0.8$ .

$$z^* = 1.28$$

Calculate the standard error.

$$SE = \sqrt{\frac{\hat{p}(1 - \hat{p})}{n}} = \sqrt{\frac{(0.095)(1 - 0.095)}{4400}} = 0.00442$$

Calculate the margin of error.

$$ME = z^* SE = (1.28)(0.00442) = 0.00566$$

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.

$$(0.0893, 0.101)$$

We are 80% confident that the true population proportion is between 8.93% and 10.1%.

- (a) The lower bound = 0.0893, which can also be expressed as 8.93%.
- (b) The upper bound = 0.101, which can also be expressed as 10.1%.

8. An experiment is run with a treatment group of size 16 and a control group of size 27. The results are summarized in the table below.

|          | treatment | control |
|----------|-----------|---------|
| sick     | 16        | 27      |
| not sick | 0         | 0       |

Using a significance level of 0.005, determine whether the treatment causes an effect on the proportion of cases that are sick.

- (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  $P(|Z| > z^*) = 0.005$ .

$$z^* = \Phi^{-1} \left( 1 - \frac{\alpha}{2} \right) = 2.81$$

Determine the sample proportions.

$$\hat{p}_1 = \frac{16}{16} = 1$$

$$\hat{p}_2 = \frac{27}{27} = 1$$

Determine the difference of sample proportions.

$$\hat{p}_2 - \hat{p}_1 = 1 - 1 = 0$$

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

$$\hat{p} = \frac{16 + 27}{16 + 27} = 1$$

Determine the standard error.

$$\begin{aligned} SE &= \sqrt{\frac{\hat{p}(1 - \hat{p})}{n_1} + \frac{\hat{p}(1 - \hat{p})}{n_2}} \\ &= \sqrt{\frac{(1)(0)}{16} + \frac{(1)(0)}{27}} \\ &= 0 \end{aligned}$$

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

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

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.

$$\begin{aligned} z &= \frac{(\hat{p}_2 - \hat{p}_1) - (p_2 - p_1)_0}{SE} \\ &= \frac{(1 - 1) - 0}{0} \\ &= NaN \end{aligned}$$

Determine the  $p$ -value.

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

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)  $z^* = 2.81$
- (d)  $SE = 0$
- (e)  $|Z_{\text{obs}}| = \text{NaN}$
- (f)  $p\text{-value} = \text{NaN}$
- (g) reject the null

# Bunker Hill Community College

Final Statistics Exam 2019-05-02

Exam ID 008

**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)  $P(\text{pig given blue}) = 0.228$

(b)  $P(\text{teal}) = 0.269$

(c)  $P(\text{pig or gray}) = 0.407$

(d)  $P(\text{horn}) = 0.185$

(e)  $P(\text{red given wheel}) = 0.212$

(f)  $P(\text{flower and red}) = 0.0663$

2.  $P(\text{"not pig" given "yellow"}) = 0.316$

3.  $P(69.67 < X < 70.23) = 0.785$

4. (a)  $P(X = 37) = 0.0742$

(b)  $P(35 \leq X \leq 48) = 0.5848$

5. **(19, 22.1)**

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

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

(c)  $t^* = 1.99$

(d)  $SE = 3.921$

(e)  $|t_{\text{obs}}| = 1.79$

(f)  $0.05 < p\text{-value} < 0.1$

(g) **retain**

7. (a) **LB of p CI = 0.934 or 93.4%**

(b) **UB of p CI = 0.938 or 93.8%**

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

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

(c)  $Z^* = 1.96$

(d)  $SE = 0.024$

(e)  $|Z_{\text{obs}}| = 2.03$

(f)  $p\text{-value} = 0.0424$

(g) **reject**

1. In a deck of strange cards, there are 966 cards. Each card has an image and a color. The amounts are shown in the table below.

|        | blue | gray | red | teal |
|--------|------|------|-----|------|
| flower | 98   | 87   | 64  | 89   |
| horn   | 27   | 48   | 16  | 88   |
| pig    | 55   | 44   | 71  | 10   |
| wheel  | 61   | 78   | 57  | 73   |

- (a) What is the probability a random card is a pig given it is blue?
- (b) What is the probability a random card is teal?
- (c) What is the probability a random card is either a pig or gray (or both)?
- (d) What is the probability a random card is a horn?
- (e) What is the probability a random card is red given it is a wheel?
- (f) What is the probability a random card is both a flower and red?

**Solution**

(a)  $P(\text{pig given blue}) = \frac{55}{98+27+55+61} = 0.228$

(b)  $P(\text{teal}) = \frac{89+88+10+73}{966} = 0.269$

(c)  $P(\text{pig or gray}) = \frac{55+44+71+10+87+48+44+78-44}{966} = 0.407$

(d)  $P(\text{horn}) = \frac{27+48+16+88}{966} = 0.185$

(e)  $P(\text{red given wheel}) = \frac{57}{61+78+57+73} = 0.212$

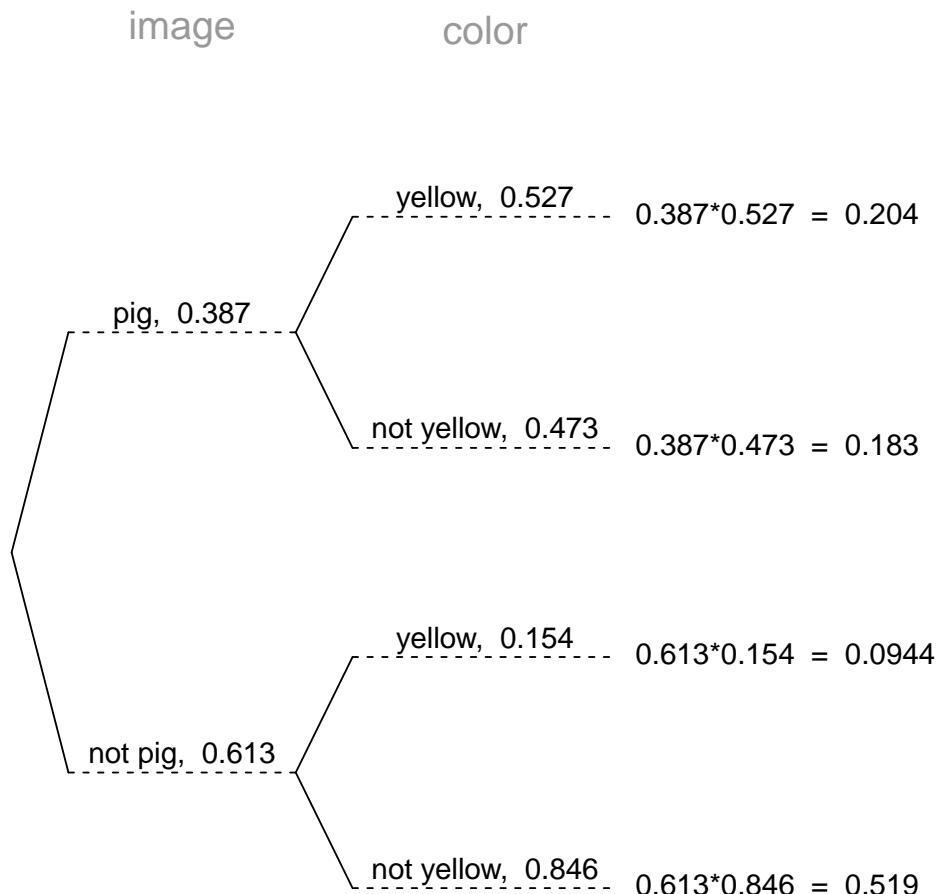
(f)  $P(\text{flower and red}) = \frac{64}{966} = 0.0663$

2. In a deck of strange cards, each card has an image and a color. The chance of drawing a pig is 38.7%. If a pig is drawn, there is a 52.7% chance that it is yellow. If a card that is not a pig is drawn, there is a 15.4% chance that it is yellow.

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

$$P(\text{"not pig" given "yellow"}) = \frac{0.0944}{0.0944 + 0.204} = 0.316$$

3. In a very large pile of toothpicks, the mean length is 69.91 millimeters and the standard deviation is 2.89 millimeters. If you randomly sample 169 toothpicks, what is the chance the sample mean is between 69.67 and 70.23 millimeters?

**Solution**

Label the given information.

$$\mu = 69.91$$

$$\sigma = 2.89$$

$$n = 169$$

$$\bar{x}_{\text{lower}} = 69.67$$

$$\bar{x}_{\text{upper}} = 70.23$$

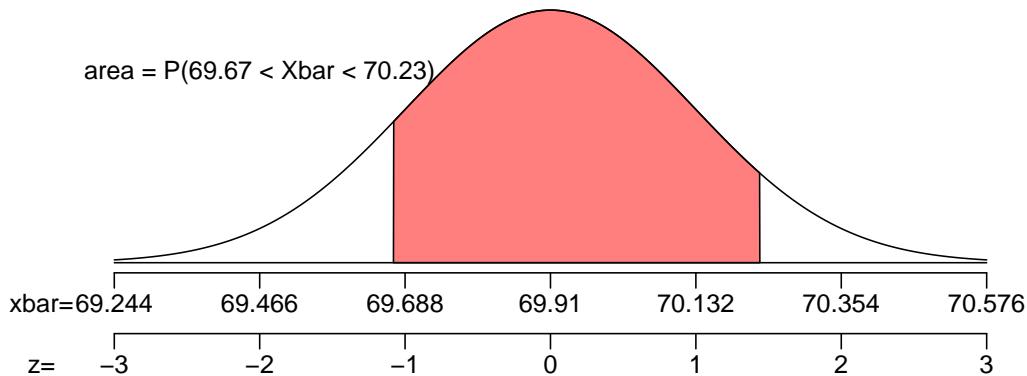
Find the standard error.

$$SE = \frac{\sigma}{\sqrt{n}} = \frac{2.89}{\sqrt{169}} = 0.222$$

Describe the sampling distribution.

$$\bar{X} \sim \mathcal{N}(69.91, 0.222)$$

Draw a sketch.



Calculate a  $z$  scores.

$$z_{\text{lower}} = \frac{\bar{x}_{\text{lower}} - \mu}{SE} = \frac{69.67 - 69.91}{0.222} = -1.08$$

$$z_{\text{upper}} = \frac{\bar{x}_{\text{upper}} - \mu}{SE} = \frac{70.23 - 69.91}{0.222} = 1.44$$

Determine the probability.

$$\begin{aligned} P(69.67 < \bar{X} < 70.23) &= \Phi(z_{\text{upper}}) - \Phi(z_{\text{lower}}) \\ &= \Phi(1.44) - \Phi(-1.08) \\ &= 0.785 \end{aligned}$$

4. In a game, there is a 23% chance to win a round. You will play 158 rounds.
- (a) What is the probability of winning exactly 37 rounds?
  - (b) What is the probability of winning at least 35 but at most 48 rounds?

**Solution**

We use the formula for binomial probabilities.

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

$$P(X = 37) = \binom{158}{37} (0.23)^{37} (1 - 0.23)^{158-37}$$

$$P(X = 37) = \binom{158}{37} (0.23)^{37} (0.77)^{121}$$

$$P(X = 37) = 0.0742$$

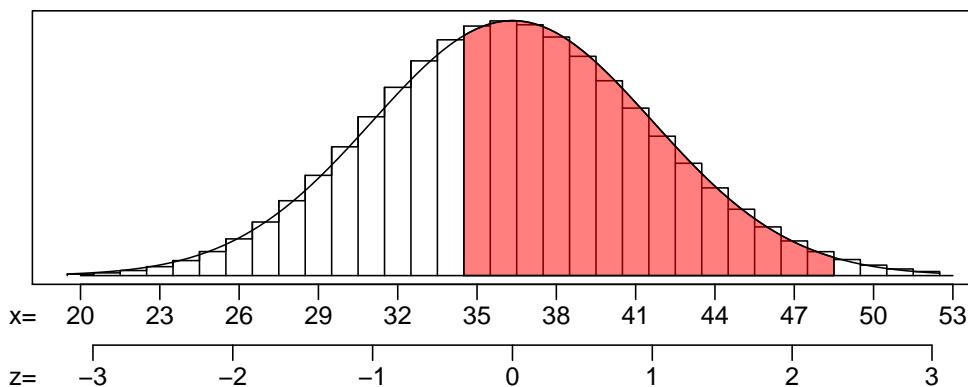
Find the mean.

$$\mu = np = (158)(0.23) = 36.34$$

Find the standard deviation.

$$\sigma = \sqrt{np(1 - p)} = \sqrt{(158)(0.23)(1 - 0.23)} = 5.2898$$

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



Find the z scores.

$$z_1 = \frac{34.5 - 36.34}{5.2898} = -0.25$$

$$z_2 = \frac{48.5 - 36.34}{5.2898} = 2.2$$

Calculate the probability.

$$P(35 \leq X \leq 48) = \Phi(2.2) - \Phi(-0.25) = 0.5848$$

(a)  $P(X = 37) = 0.0742$

(b)  $P(35 \leq X \leq 48) = 0.5848$

5. As an ornithologist, you wish to determine the average body mass of *Seiurus noveboracensis*. You randomly sample 33 adults of *Seiurus noveboracensis*, resulting in a sample mean of 20.51 grams and a sample standard deviation of 3.64 grams. Determine a 98% confidence interval of the true population mean.

**Solution**

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

$$n = 33$$

$$\bar{x} = 20.51$$

$$s = 3.64$$

$$CL = 0.98$$

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

$$df = n - 1 = 32$$

Determine the critical  $t$  value,  $t^*$ , such that  $P(|T| < t^*) = 0.98$ .

$$t^* = 2.45$$

Calculate the standard error.

$$SE = \frac{s}{\sqrt{n}} = \frac{3.64}{\sqrt{33}} = 0.634$$

We want to make an inference about the population mean.

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

Determine the bounds.

$$\begin{aligned} CI &= (\bar{x} - t^* SE, \bar{x} + t^* SE) \\ &= (20.51 - 2.45 \times 0.634, 20.51 + 2.45 \times 0.634) \\ &= (19, 22.1) \end{aligned}$$

We are 98% confident that the population mean is between 19 and 22.1.

6. A treatment group of size 33 has a mean of 102 and standard deviation of 15.4. A control group of size 40 has a mean of 109 and standard deviation of 18.1. If you decided to use a significance level of 0.05, 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 70.

- (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.

$$n_1 = 33$$

$$\bar{x}_1 = 102$$

$$s_1 = 15.4$$

$$n_2 = 40$$

$$\bar{x}_2 = 109$$

$$s_2 = 18.1$$

$$\alpha = 0.05$$

$$df = 70$$

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  $P(|T| > t^*) = 0.05$  by using a  $t$  table.

$$t^* = 1.99$$

Calculate the standard error.

$$\begin{aligned} SE &= \sqrt{\frac{(s_1)^2}{n_1} + \frac{(s_2)^2}{n_2}} \\ &= \sqrt{\frac{(15.4)^2}{33} + \frac{(18.1)^2}{40}} \\ &= 3.921 \end{aligned}$$

Determine the test statistic.

$$\begin{aligned} t_{\text{obs}} &= \frac{(\bar{x}_2 - \bar{x}_1) - (\mu_2 - \mu_1)_0}{SE} \\ &= \frac{(109 - 102) - (0)}{3.921} \\ &= 1.79 \end{aligned}$$

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

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

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

$$0.05 < p\text{-value} < 0.1$$

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)  $t^* = 1.99$
- (d)  $SE = 3.921$
- (e)  $|t_{\text{obs}}| = 1.79$
- (f)  $0.05 < p\text{-value} < 0.1$
- (g) retain the null

7. From a very large population, a random sample of 49000 individuals was taken. In that sample, 93.6% were asleep. 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 condifence interval.

**Solution**

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

$$z^* = 2.05$$

Calculate the standard error.

$$SE = \sqrt{\frac{\hat{p}(1 - \hat{p})}{n}} = \sqrt{\frac{(0.936)(1 - 0.936)}{49000}} = 0.00111$$

Calculate the margin of error.

$$ME = z^* SE = (2.05)(0.00111) = 0.00228$$

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.

$$(0.934, 0.938)$$

We are 96% confident that the true population proportion is between 93.4% and 93.8%.

- (a) The lower bound = 0.934, which can also be expressed as 93.4%.
- (b) The upper bound = 0.938, which can also be expressed as 93.8%.

8. An experiment is run with a treatment group of size 204 and a control group of size 243. The results are summarized in the table below.

|            | treatment | control |
|------------|-----------|---------|
| glossy     | 185       | 232     |
| not glossy | 19        | 11      |

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.

$$\hat{p}_1 = \frac{185}{204} = 0.907$$

$$\hat{p}_2 = \frac{232}{243} = 0.955$$

Determine the difference of sample proportions.

$$\hat{p}_2 - \hat{p}_1 = 0.955 - 0.907 = 0.048$$

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

$$\hat{p} = \frac{185 + 232}{204 + 243} = 0.933$$

Determine the standard error.

$$\begin{aligned} SE &= \sqrt{\frac{\hat{p}(1 - \hat{p})}{n_1} + \frac{\hat{p}(1 - \hat{p})}{n_2}} \\ &= \sqrt{\frac{(0.933)(0.067)}{204} + \frac{(0.933)(0.067)}{243}} \\ &= 0.0237 \end{aligned}$$

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

$$H_0 : \hat{P}_2 - \hat{P}_1 \sim \mathcal{N}(0, 0.0237)$$

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.

$$\begin{aligned} z &= \frac{(\hat{p}_2 - \hat{p}_1) - (p_2 - p_1)_0}{SE} \\ &= \frac{(0.955 - 0.907) - 0}{0.0237} \\ &= 2.03 \end{aligned}$$

Determine the  $p$ -value.

$$\begin{aligned} p\text{-value} &= 2 \cdot \Phi(-|z|) \\ &= 2 \cdot \Phi(-2.03) \\ &= 0.0424 \end{aligned}$$

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)  $SE = 0.0237$
- (e)  $|Z_{\text{obs}}| = 2.03$
- (f)  $p\text{-value} = 0.0424$
- (g) reject the null

# 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)  $P(\text{tree given orange}) = 0.163$

(b)  $P(\text{wheel and gray}) = 0.0172$

(c)  $P(\text{wheel or orange}) = 0.439$

(d)  $P(\text{flower}) = 0.202$

(e)  $P(\text{orange}) = 0.362$

(f)  $P(\text{gray given gem}) = 0.18$

2.  $P(\text{"gem" given "pink"}) = 0.0891$

3.  $P(70.39 < X < 70.68) = 0.8111$

4. (a)  $P(X = 97) = 0.0554$

(b)  $P(78 \leq X \leq 96) = 0.5916$

5. **(12, 13.1)**

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

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

(c)  $t^* = 2.51$

(d)  $SE = 5.215$

(e)  $|t_{\text{obs}}| = 2.21$

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

(g) **retain**

7. (a) **LB of p CI = 0.663 or 66.3%**

(b) **UB of p CI = 0.701 or 70.1%**

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

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

(c)  $Z^* = 2.33$

(d)  $SE = 0.025$

(e)  $|Z_{\text{obs}}| = 2.49$

(f)  $p\text{-value} = 0.0128$

(g) **reject**

1. In a deck of strange cards, there are 816 cards. Each card has an image and a color. The amounts are shown in the table below.

|        | gray | orange | red |
|--------|------|--------|-----|
| bike   | 56   | 99     | 55  |
| flower | 77   | 64     | 24  |
| gem    | 27   | 40     | 83  |
| tree   | 62   | 48     | 74  |
| wheel  | 14   | 44     | 49  |

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

**Solution**

(a)  $P(\text{tree given orange}) = \frac{48}{99+64+40+48+44} = 0.163$

(b)  $P(\text{wheel and gray}) = \frac{14}{816} = 0.0172$

(c)  $P(\text{wheel or orange}) = \frac{14+44+49+99+64+40+48+44-44}{816} = 0.439$

(d)  $P(\text{flower}) = \frac{77+64+24}{816} = 0.202$

(e)  $P(\text{orange}) = \frac{99+64+40+48+44}{816} = 0.362$

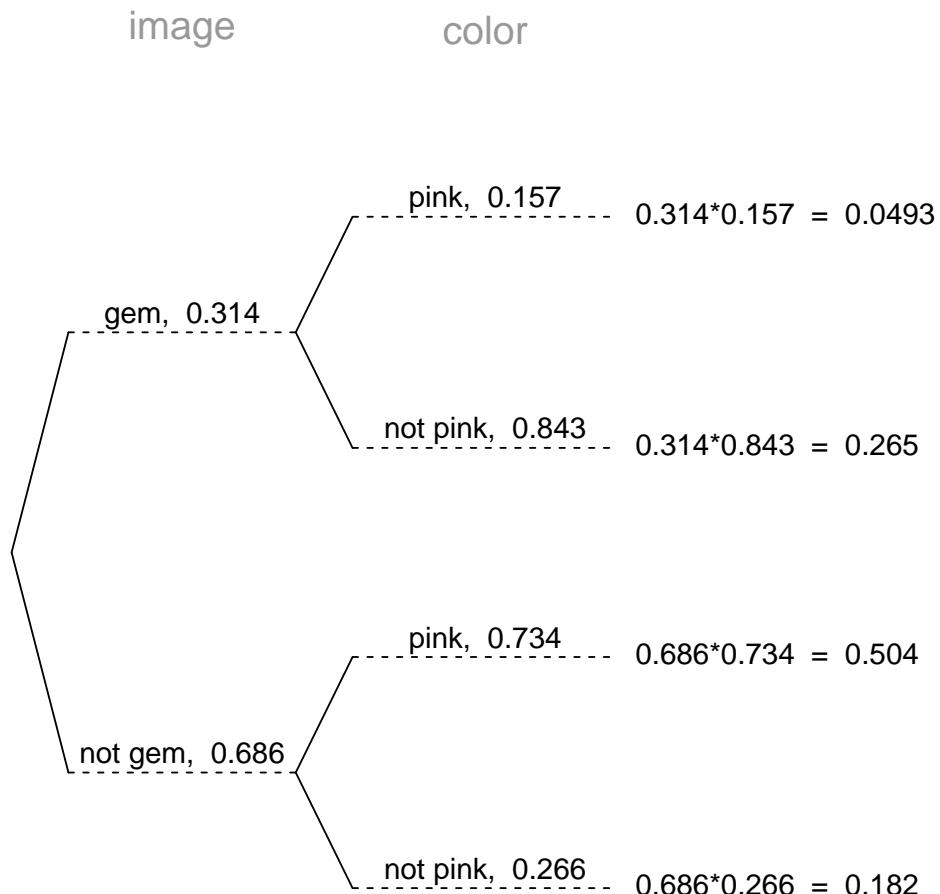
(f)  $P(\text{gray given gem}) = \frac{27}{27+40+83} = 0.18$

2. In a deck of strange cards, each card has an image and a color. The chance of drawing a gem is 31.4%. If a gem is drawn, there is a 15.7% chance that it is pink. If a card that is not a gem is drawn, there is a 73.4% chance that it is pink.

Now, someone draws a random card and reveals it is pink. What is the chance the card is a gem?

**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.

$$P(\text{"gem" given "pink"}) = \frac{0.0493}{0.0493 + 0.504} = 0.0891$$

3. In a very large pile of toothpicks, the mean length is 70.54 millimeters and the standard deviation is 1.35 millimeters. If you randomly sample 150 toothpicks, what is the chance the sample mean is between 70.39 and 70.68 millimeters?

**Solution**

Label the given information.

$$\mu = 70.54$$

$$\sigma = 1.35$$

$$n = 150$$

$$\bar{x}_{\text{lower}} = 70.39$$

$$\bar{x}_{\text{upper}} = 70.68$$

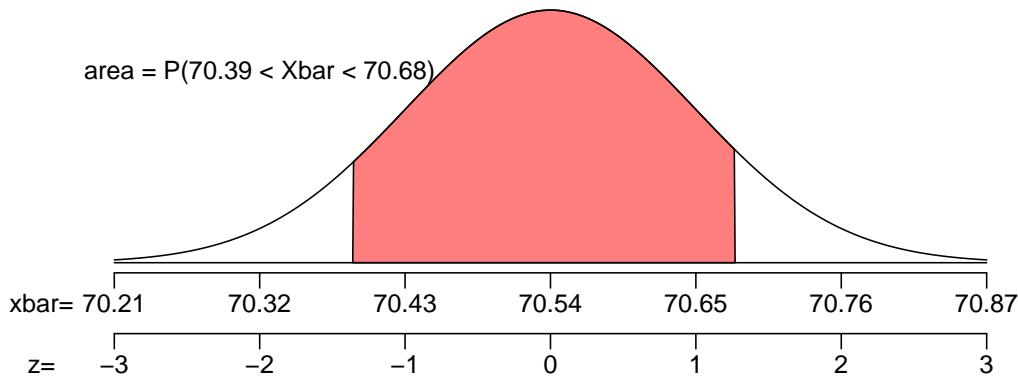
Find the standard error.

$$SE = \frac{\sigma}{\sqrt{n}} = \frac{1.35}{\sqrt{150}} = 0.11$$

Describe the sampling distribution.

$$\bar{X} \sim \mathcal{N}(70.54, 0.11)$$

Draw a sketch.



Calculate a  $z$  scores.

$$z_{\text{lower}} = \frac{x_{\text{lower}} - \mu}{SE} = \frac{70.39 - 70.54}{0.11} = -1.36$$

$$z_{\text{upper}} = \frac{x_{\text{upper}} - \mu}{SE} = \frac{70.68 - 70.54}{0.11} = 1.27$$

Determine the probability.

$$\begin{aligned} P(70.39 < \bar{X} < 70.68) &= \Phi(z_{\text{upper}}) - \Phi(z_{\text{lower}}) \\ &= \Phi(1.27) - \Phi(-1.36) \\ &= 0.8111 \end{aligned}$$

4. In a game, there is a 53% chance to win a round. You will play 178 rounds.
- (a) What is the probability of winning exactly 97 rounds?
  - (b) What is the probability of winning at least 78 but at most 96 rounds?

**Solution**

We use the formula for binomial probabilities.

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

$$P(X = 97) = \binom{178}{97} (0.53)^{97} (1 - 0.53)^{178-97}$$

$$P(X = 97) = \binom{178}{97} (0.53)^{97} (0.47)^{81}$$

$$P(X = 97) = 0.0554$$

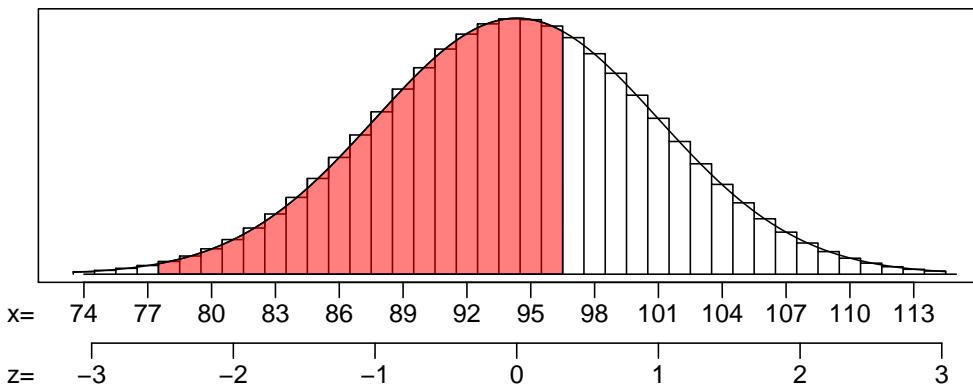
Find the mean.

$$\mu = np = (178)(0.53) = 94.34$$

Find the standard deviation.

$$\sigma = \sqrt{np(1 - p)} = \sqrt{(178)(0.53)(1 - 0.53)} = 6.6588$$

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



Find the z scores.

$$z_1 = \frac{77.5 - 94.34}{6.6588} = -2.45$$

$$z_2 = \frac{96.5 - 94.34}{6.6588} = 0.25$$

Calculate the probability.

$$P(78 \leq X \leq 96) = \Phi(0.25) - \Phi(-2.45) = 0.5916$$

(a)  $P(X = 97) = 0.0554$

(b)  $P(78 \leq X \leq 96) = 0.5916$

5. As an ornithologist, you wish to determine the average body mass of *Vermivora peregrina*. You randomly sample 23 adults of *Vermivora peregrina*, resulting in a sample mean of 12.52 grams and a sample standard deviation of 1.24 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.

$$n = 23$$

$$\bar{x} = 12.52$$

$$s = 1.24$$

$$CL = 0.95$$

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

$$df = n - 1 = 22$$

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

$$t^* = 2.07$$

Calculate the standard error.

$$SE = \frac{s}{\sqrt{n}} = \frac{1.24}{\sqrt{23}} = 0.259$$

We want to make an inference about the population mean.

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

Determine the bounds.

$$\begin{aligned} CI &= (\bar{x} - t^* SE, \bar{x} + t^* SE) \\ &= (12.52 - 2.07 \times 0.259, 12.52 + 2.07 \times 0.259) \\ &= (12, 13.1) \end{aligned}$$

We are 95% confident that the population mean is between 12 and 13.1.

6. A treatment group of size 14 has a mean of 98.5 and standard deviation of 16.4. A control group of size 16 has a mean of 110 and standard deviation of 11.3. 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 22.

- (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.

$$n_1 = 14$$

$$\bar{x}_1 = 98.5$$

$$s_1 = 16.4$$

$$n_2 = 16$$

$$\bar{x}_2 = 110$$

$$s_2 = 11.3$$

$$\alpha = 0.02$$

$$df = 22$$

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  $P(|T| > t^*) = 0.02$  by using a  $t$  table.

$$t^* = 2.51$$

Calculate the standard error.

$$\begin{aligned} SE &= \sqrt{\frac{(s_1)^2}{n_1} + \frac{(s_2)^2}{n_2}} \\ &= \sqrt{\frac{(16.4)^2}{14} + \frac{(11.3)^2}{16}} \\ &= 5.215 \end{aligned}$$

Determine the test statistic.

$$\begin{aligned} t_{\text{obs}} &= \frac{(\bar{x}_2 - \bar{x}_1) - (\mu_2 - \mu_1)_0}{SE} \\ &= \frac{(110 - 98.5) - (0)}{5.215} \\ &= 2.21 \end{aligned}$$

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)  $t^* = 2.51$
- (d)  $SE = 5.215$
- (e)  $|t_{\text{obs}}| = 2.21$
- (f)  $0.02 < p\text{-value} < 0.04$
- (g) retain the null

7. From a very large population, a random sample of 5000 individuals was taken. In that sample, 68.2% were blue. Determine a 99.5% confidence interval of the population proportion.

- (a) Find the lower bound of the confidence interval.
- (b) Find the upper bound of the condifence interval.

**Solution**

Determine  $z^*$  such that  $P(|Z| < z^*) = 0.995$ .

$$z^* = 2.81$$

Calculate the standard error.

$$SE = \sqrt{\frac{\hat{p}(1 - \hat{p})}{n}} = \sqrt{\frac{(0.682)(1 - 0.682)}{5000}} = 0.00659$$

Calculate the margin of error.

$$ME = z^* SE = (2.81)(0.00659) = 0.0185$$

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.

$$(0.663, 0.701)$$

We are 99.5% confident that the true population proportion is between 66.3% and 70.1%.

- (a) The lower bound = 0.663, which can also be expressed as 66.3%.
- (b) The upper bound = 0.701, which can also be expressed as 70.1%.

8. An experiment is run with a treatment group of size 273 and a control group of size 275. The results are summarized in the table below.

|          | treatment | control |
|----------|-----------|---------|
| pink     | 238       | 257     |
| not pink | 35        | 18      |

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

- (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.

$$\hat{p}_1 = \frac{238}{273} = 0.872$$

$$\hat{p}_2 = \frac{257}{275} = 0.935$$

Determine the difference of sample proportions.

$$\hat{p}_2 - \hat{p}_1 = 0.935 - 0.872 = 0.063$$

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

$$\hat{p} = \frac{238 + 257}{273 + 275} = 0.903$$

Determine the standard error.

$$\begin{aligned} SE &= \sqrt{\frac{\hat{p}(1 - \hat{p})}{n_1} + \frac{\hat{p}(1 - \hat{p})}{n_2}} \\ &= \sqrt{\frac{(0.903)(0.097)}{273} + \frac{(0.903)(0.097)}{275}} \\ &= 0.0253 \end{aligned}$$

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

$$H_0 : \hat{P}_2 - \hat{P}_1 \sim \mathcal{N}(0, 0.0253)$$

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.

$$\begin{aligned} z &= \frac{(\hat{p}_2 - \hat{p}_1) - (p_2 - p_1)_0}{SE} \\ &= \frac{(0.935 - 0.872) - 0}{0.0253} \\ &= 2.49 \end{aligned}$$

Determine the  $p$ -value.

$$\begin{aligned} p\text{-value} &= 2 \cdot \Phi(-|z|) \\ &= 2 \cdot \Phi(-2.49) \\ &= 0.0128 \end{aligned}$$

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)  $SE = 0.0253$
- (e)  $|z_{\text{obs}}| = 2.49$
- (f)  $p\text{-value} = 0.0128$
- (g) reject the null

# Bunker Hill Community College

Final Statistics Exam 2019-05-02

Exam ID 010

**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)  $P(\text{green given wheel}) = 0.341$

(b)  $P(\text{wheel or pink}) = 0.431$

(c)  $P(\text{bike}) = 0.304$

(d)  $P(\text{wheel and gray}) = 0.0169$

(e)  $P(\text{indigo}) = 0.18$

(f)  $P(\text{shovel given pink}) = 0.27$

2.  $P(\text{"kite" given "not indigo"}) = 0.45$

3.  $P(71.51 < X < 71.95) = 0.5183$

4. (a)  $P(X = 83) = 0.0583$

(b)  $P(70 \leq X \leq 77) = 0.0986$

5.  $(64.3, 75.1)$

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

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

(c)  $t^* = 2.2$

(d)  $SE = 28.699$

(e)  $|t_{\text{obs}}| = 2.09$

(f)  $0.05 < p\text{-value} < 0.1$

(g) **retain**

7. (a) **LB of p CI = 0.671 or 67.1%**

(b) **UB of p CI = 0.705 or 70.5%**

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

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

(c)  $Z^* = 2.33$

(d)  $SE = 0.028$

(e)  $|Z_{\text{obs}}| = 2.61$

(f)  $p\text{-value} = 0.009$

(g) **reject**

1. In a deck of strange cards, there are 947 cards. Each card has an image and a color. The amounts are shown in the table below.

|        | gray | green | indigo | pink |
|--------|------|-------|--------|------|
| bike   | 81   | 87    | 32     | 88   |
| horn   | 53   | 80    | 44     | 25   |
| shovel | 39   | 58    | 65     | 78   |
| wheel  | 16   | 74    | 29     | 98   |

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

**Solution**

(a)  $P(\text{green given wheel}) = \frac{74}{16+74+29+98} = 0.341$

(b)  $P(\text{wheel or pink}) = \frac{16+74+29+98+88+25+78+98-98}{947} = 0.431$

(c)  $P(\text{bike}) = \frac{81+87+32+88}{947} = 0.304$

(d)  $P(\text{wheel and gray}) = \frac{16}{947} = 0.0169$

(e)  $P(\text{indigo}) = \frac{32+44+65+29}{947} = 0.18$

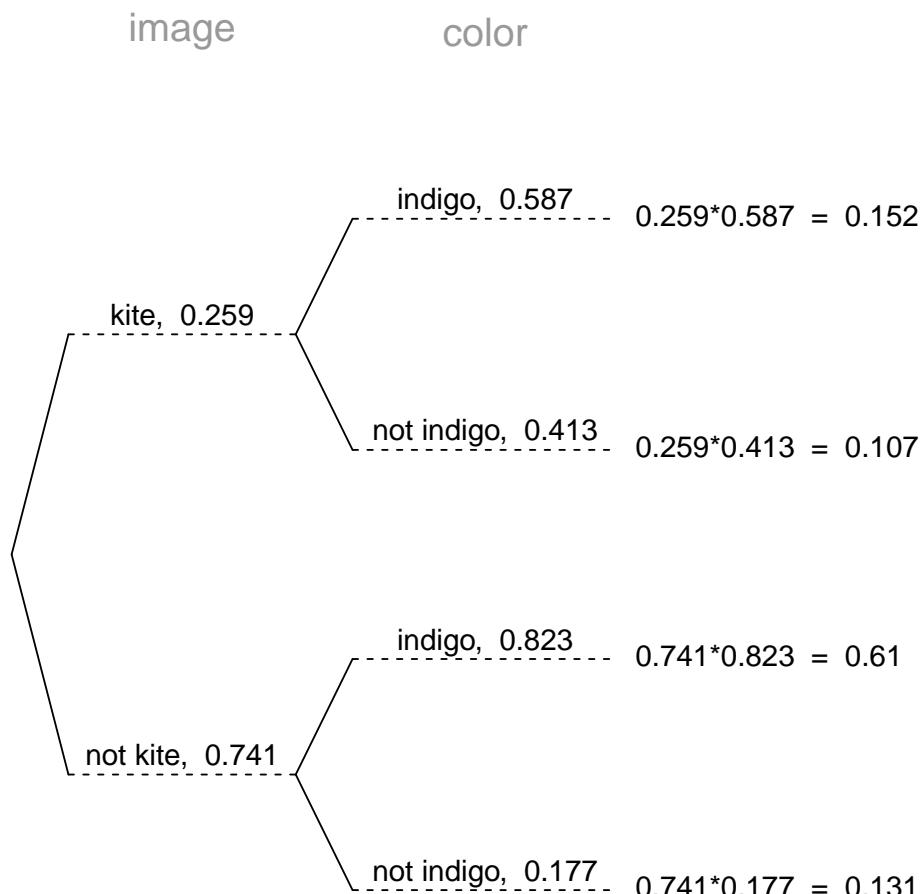
(f)  $P(\text{shovel given pink}) = \frac{78}{88+25+78+98} = 0.27$

2. In a deck of strange cards, each card has an image and a color. The chance of drawing a kite is 25.9%. If a kite is drawn, there is a 58.7% chance that it is indigo. If a card that is not a kite is drawn, there is a 82.3% chance that it is indigo.

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

**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.

$$P(\text{"kite" given "not indigo"}) = \frac{0.107}{0.107 + 0.131} = 0.45$$

3. In a very large pile of toothpicks, the mean length is 71.75 millimeters and the standard deviation is 3.42 millimeters. If you randomly sample 120 toothpicks, what is the chance the sample mean is between 71.51 and 71.95 millimeters?

**Solution**

Label the given information.

$$\mu = 71.75$$

$$\sigma = 3.42$$

$$n = 120$$

$$\bar{x}_{\text{lower}} = 71.51$$

$$\bar{x}_{\text{upper}} = 71.95$$

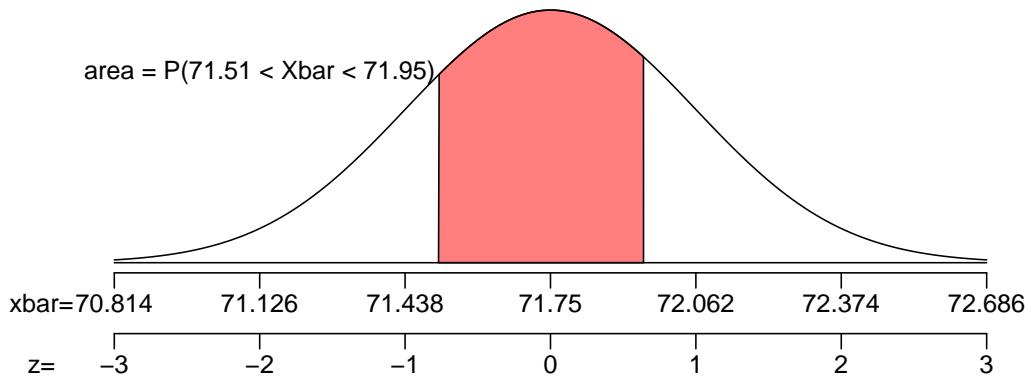
Find the standard error.

$$SE = \frac{\sigma}{\sqrt{n}} = \frac{3.42}{\sqrt{120}} = 0.312$$

Describe the sampling distribution.

$$\bar{X} \sim \mathcal{N}(71.75, 0.312)$$

Draw a sketch.



Calculate a  $z$  scores.

$$z_{\text{lower}} = \frac{x_{\text{lower}} - \mu}{SE} = \frac{71.51 - 71.75}{0.312} = -0.77$$

$$z_{\text{upper}} = \frac{x_{\text{upper}} - \mu}{SE} = \frac{71.95 - 71.75}{0.312} = 0.64$$

Determine the probability.

$$\begin{aligned} P(71.51 < X < 71.95) &= \Phi(z_{\text{upper}}) - \Phi(z_{\text{lower}}) \\ &= \Phi(0.64) - \Phi(-0.77) \\ &= 0.5183 \end{aligned}$$

4. In a game, there is a 50% chance to win a round. You will play 170 rounds.
- (a) What is the probability of winning exactly 83 rounds?
  - (b) What is the probability of winning at least 70 but at most 77 rounds?

**Solution**

We use the formula for binomial probabilities.

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

$$P(X = 83) = \binom{170}{83} (0.5)^{83} (1 - 0.5)^{170-83}$$

$$P(X = 83) = \binom{170}{83} (0.5)^{83} (0.5)^{87}$$

$$P(X = 83) = 0.0583$$

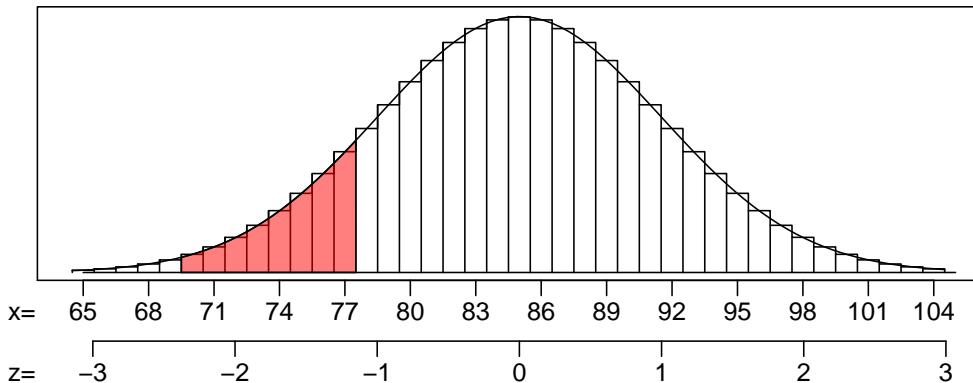
Find the mean.

$$\mu = np = (170)(0.5) = 85$$

Find the standard deviation.

$$\sigma = \sqrt{np(1 - p)} = \sqrt{(170)(0.5)(1 - 0.5)} = 6.5192$$

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



Find the z scores.

$$z_1 = \frac{69.5 - 85}{6.5192} = -2.3$$

$$z_2 = \frac{77.5 - 85}{6.5192} = -1.23$$

Calculate the probability.

$$P(70 \leq X \leq 77) = \Phi(-1.23) - \Phi(-2.3) = 0.0986$$

(a)  $P(X = 83) = 0.0583$

(b)  $P(70 \leq X \leq 77) = 0.0986$

5. As an ornithologist, you wish to determine the average body mass of *Porzana carolina*. You randomly sample 27 adults of *Porzana carolina*, resulting in a sample mean of 69.7 grams and a sample standard deviation of 11.3 grams. Determine a 98% confidence interval of the true population mean.

**Solution**

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

$$n = 27$$

$$\bar{x} = 69.7$$

$$s = 11.3$$

$$CL = 0.98$$

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

$$df = n - 1 = 26$$

Determine the critical  $t$  value,  $t^*$ , such that  $P(|T| < t^*) = 0.98$ .

$$t^* = 2.48$$

Calculate the standard error.

$$SE = \frac{s}{\sqrt{n}} = \frac{11.3}{\sqrt{27}} = 2.17$$

We want to make an inference about the population mean.

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

Determine the bounds.

$$\begin{aligned} CI &= (\bar{x} - t^* SE, \bar{x} + t^* SE) \\ &= (69.7 - 2.48 \times 2.17, 69.7 + 2.48 \times 2.17) \\ &= (64.3, 75.1) \end{aligned}$$

We are 98% confident that the population mean is between 64.3 and 75.1.

6. A treatment group of size 12 has a mean of 1090 and standard deviation of 63.8. A control group of size 11 has a mean of 1030 and standard deviation of 73. 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 19.

- (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.

$$n_1 = 12$$

$$\bar{x}_1 = 1090$$

$$s_1 = 63.8$$

$$n_2 = 11$$

$$\bar{x}_2 = 1030$$

$$s_2 = 73$$

$$\alpha = 0.04$$

$$df = 19$$

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.

$$t^* = 2.2$$

Calculate the standard error.

$$\begin{aligned} SE &= \sqrt{\frac{(s_1)^2}{n_1} + \frac{(s_2)^2}{n_2}} \\ &= \sqrt{\frac{(63.8)^2}{12} + \frac{(73)^2}{11}} \\ &= 28.699 \end{aligned}$$

Determine the test statistic.

$$\begin{aligned} t_{\text{obs}} &= \frac{(\bar{x}_2 - \bar{x}_1) - (\mu_2 - \mu_1)_0}{SE} \\ &= \frac{(1030 - 1090) - (0)}{28.699} \\ &= -2.09 \end{aligned}$$

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

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

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

$$0.05 < p\text{-value} < 0.1$$

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)  $t^* = 2.2$
- (d)  $SE = 28.699$
- (e)  $|t_{\text{obs}}| = 2.09$
- (f)  $0.05 < p\text{-value} < 0.1$
- (g) retain the null

7. From a very large population, a random sample of 1900 individuals was taken. In that sample, 68.8% were cold. Determine a 90% confidence interval of the population proportion.

- (a) Find the lower bound of the confidence interval.
- (b) Find the upper bound of the condifence interval.

**Solution**

Determine  $z^*$  such that  $P(|Z| < z^*) = 0.9$ .

$$z^* = 1.64$$

Calculate the standard error.

$$SE = \sqrt{\frac{\hat{p}(1 - \hat{p})}{n}} = \sqrt{\frac{(0.688)(1 - 0.688)}{1900}} = 0.0106$$

Calculate the margin of error.

$$ME = z^* SE = (1.64)(0.0106) = 0.0174$$

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.

$$(0.671, 0.705)$$

We are 90% confident that the true population proportion is between 67.1% and 70.5%.

- (a) The lower bound = 0.671, which can also be expressed as 67.1%.
- (b) The upper bound = 0.705, which can also be expressed as 70.5%.

8. An experiment is run with a treatment group of size 275 and a control group of size 226. The results are summarized in the table below.

|          | treatment | control |
|----------|-----------|---------|
| cold     | 237       | 211     |
| not cold | 38        | 15      |

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

- (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.

$$\hat{p}_1 = \frac{237}{275} = 0.862$$

$$\hat{p}_2 = \frac{211}{226} = 0.934$$

Determine the difference of sample proportions.

$$\hat{p}_2 - \hat{p}_1 = 0.934 - 0.862 = 0.072$$

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

$$\hat{p} = \frac{237 + 211}{275 + 226} = 0.894$$

Determine the standard error.

$$\begin{aligned} SE &= \sqrt{\frac{\hat{p}(1 - \hat{p})}{n_1} + \frac{\hat{p}(1 - \hat{p})}{n_2}} \\ &= \sqrt{\frac{(0.894)(0.106)}{275} + \frac{(0.894)(0.106)}{226}} \\ &= 0.0276 \end{aligned}$$

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

$$H_0 : \hat{P}_2 - \hat{P}_1 \sim \mathcal{N}(0, 0.0276)$$

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.

$$\begin{aligned} z &= \frac{(\hat{p}_2 - \hat{p}_1) - (p_2 - p_1)_0}{SE} \\ &= \frac{(0.934 - 0.862) - 0}{0.0276} \\ &= 2.61 \end{aligned}$$

Determine the  $p$ -value.

$$\begin{aligned} p\text{-value} &= 2 \cdot \Phi(-|z|) \\ &= 2 \cdot \Phi(-2.61) \\ &= 0.009 \end{aligned}$$

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)  $SE = 0.0276$
- (e)  $|Z_{\text{obs}}| = 2.61$
- (f)  $p\text{-value} = 0.009$
- (g) reject the null

# Bunker Hill Community College

Final Statistics Exam 2019-05-02

Exam ID 011

**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)  $P(\text{tree}) = 0.26$

(b)  $P(\text{shovel or indigo}) = 0.327$

(c)  $P(\text{shovel and red}) = 0.0816$

(d)  $P(\text{yellow given dog}) = 0.0676$

(e)  $P(\text{pig given pink}) = 0.2$

(f)  $P(\text{pink}) = 0.201$

2.  $P(\text{"not wheel" given "green"}) = 0.311$

3.  $P(68.15 < X < 68.51) = 0.7666$

4. (a)  $P(X = 25) = 0.0815$

(b)  $P(23 \leq X \leq 34) = 0.8863$

5.  $(9.89, 10.2)$

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

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

(c)  $t^* = 2.66$

(d)  $SE = 3.192$

(e)  $|t_{\text{obs}}| = 2.51$

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

(g) **retain**

7. (a) **LB of p CI = 0.872 or 87.2%**

(b) **UB of p CI = 0.884 or 88.4%**

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

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

(c)  $Z^* = 2.05$

(d)  $SE = 0.053$

(e)  $|Z_{\text{obs}}| = 2.33$

(f)  $p\text{-value} = 0.0198$

(g) **reject**

1. In a deck of strange cards, there are 895 cards. Each card has an image and a color. The amounts are shown in the table below.

|        | green | indigo | pink | red | yellow |
|--------|-------|--------|------|-----|--------|
| dog    | 16    | 26     | 86   | 65  | 14     |
| pig    | 34    | 18     | 36   | 68  | 90     |
| shovel | 31    | 19     | 45   | 73  | 41     |
| tree   | 84    | 40     | 13   | 20  | 76     |

- (a) What is the probability a random card is a tree?
- (b) What is the probability a random card is either a shovel or indigo (or both)?
- (c) What is the probability a random card is both a shovel and red?
- (d) What is the probability a random card is yellow given it is a dog?
- (e) What is the probability a random card is a pig given it is pink?
- (f) What is the probability a random card is pink?

**Solution**

(a)  $P(\text{tree}) = \frac{84+40+13+20+76}{895} = 0.26$

(b)  $P(\text{shovel or indigo}) = \frac{31+19+45+73+41+26+18+19+40-19}{895} = 0.327$

(c)  $P(\text{shovel and red}) = \frac{73}{895} = 0.0816$

(d)  $P(\text{yellow given dog}) = \frac{14}{16+26+86+65+14} = 0.0676$

(e)  $P(\text{pig given pink}) = \frac{36}{86+36+45+13} = 0.2$

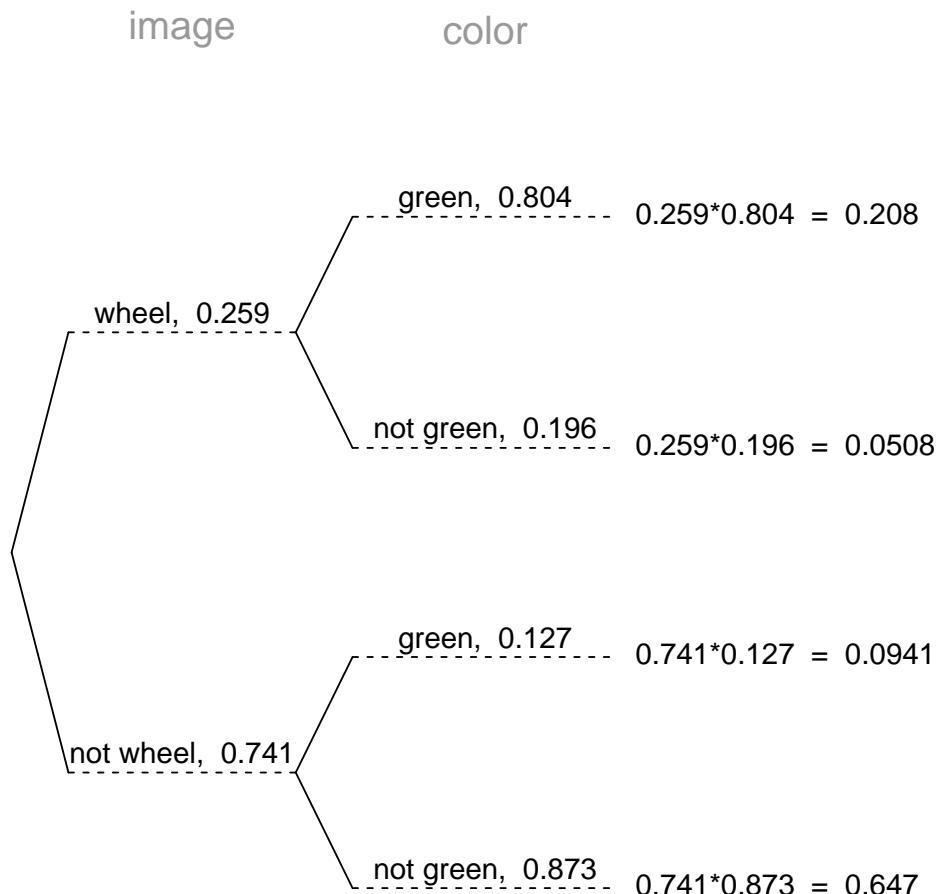
(f)  $P(\text{pink}) = \frac{86+36+45+13}{895} = 0.201$

2. In a deck of strange cards, each card has an image and a color. The chance of drawing a wheel is 25.9%. If a wheel is drawn, there is a 80.4% chance that it is green. If a card that is not a wheel is drawn, there is a 12.7% 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 wheel?

**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.

$$P(\text{"not wheel" given "green"}) = \frac{0.0941}{0.0941 + 0.208} = 0.311$$

3. In a very large pile of toothpicks, the mean length is 68.4 millimeters and the standard deviation is 1.99 millimeters. If you randomly sample 225 toothpicks, what is the chance the sample mean is between 68.15 and 68.51 millimeters?

**Solution**

Label the given information.

$$\mu = 68.4$$

$$\sigma = 1.99$$

$$n = 225$$

$$\bar{x}_{\text{lower}} = 68.15$$

$$\bar{x}_{\text{upper}} = 68.51$$

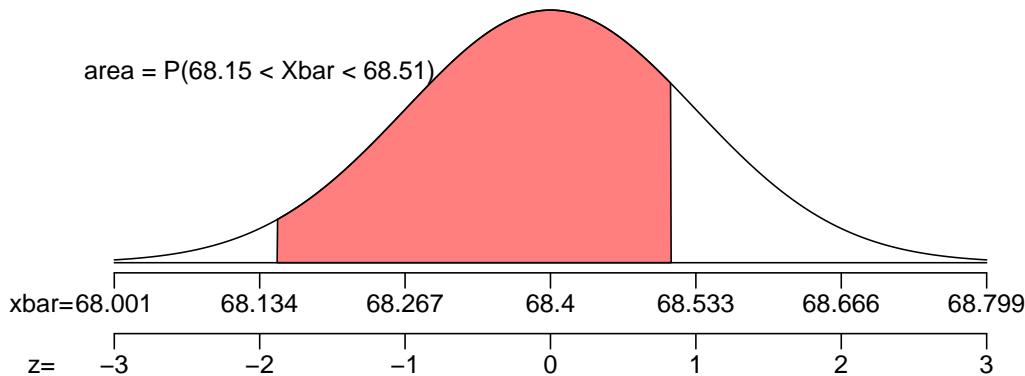
Find the standard error.

$$SE = \frac{\sigma}{\sqrt{n}} = \frac{1.99}{\sqrt{225}} = 0.133$$

Describe the sampling distribution.

$$\bar{X} \sim \mathcal{N}(68.4, 0.133)$$

Draw a sketch.



Calculate a  $z$  scores.

$$z_{\text{lower}} = \frac{\bar{x}_{\text{lower}} - \mu}{SE} = \frac{68.15 - 68.4}{0.133} = -1.88$$

$$z_{\text{upper}} = \frac{\bar{x}_{\text{upper}} - \mu}{SE} = \frac{68.51 - 68.4}{0.133} = 0.83$$

Determine the probability.

$$\begin{aligned} P(68.15 < X < 68.51) &= \Phi(z_{\text{upper}}) - \Phi(z_{\text{lower}}) \\ &= \Phi(0.83) - \Phi(-1.88) \\ &= 0.7666 \end{aligned}$$

4. In a game, there is a 58% chance to win a round. You will play 48 rounds.
- (a) What is the probability of winning exactly 25 rounds?
  - (b) What is the probability of winning at least 23 but at most 34 rounds?

**Solution**

We use the formula for binomial probabilities.

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

$$P(X = 25) = \binom{48}{25} (0.58)^{25} (1 - 0.58)^{48-25}$$

$$P(X = 25) = \binom{48}{25} (0.58)^{25} (0.42)^{23}$$

$$P(X = 25) = 0.0815$$

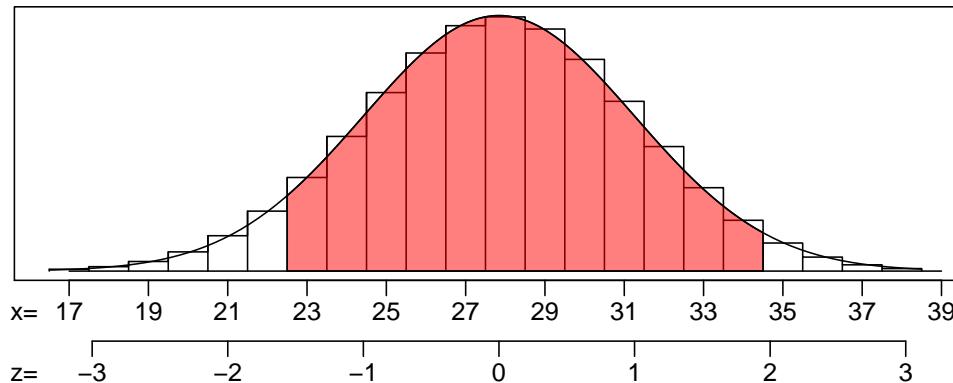
Find the mean.

$$\mu = np = (48)(0.58) = 27.84$$

Find the standard deviation.

$$\sigma = \sqrt{np(1 - p)} = \sqrt{(48)(0.58)(1 - 0.58)} = 3.4195$$

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



Find the z scores.

$$z_1 = \frac{22.5 - 27.84}{3.4195} = -1.42$$

$$z_2 = \frac{34.5 - 27.84}{3.4195} = 1.8$$

Calculate the probability.

$$P(23 \leq X \leq 34) = \Phi(1.8) - \Phi(-1.42) = 0.8863$$

(a)  $P(X = 25) = 0.0815$

(b)  $P(23 \leq X \leq 34) = 0.8863$

5. As an ornithologist, you wish to determine the average body mass of *Vireo griseus*. You randomly sample 33 adults of *Vireo griseus*, resulting in a sample mean of 10.04 grams and a sample standard deviation of 0.653 grams. Determine a 80% confidence interval of the true population mean.

**Solution**

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

$$n = 33$$

$$\bar{x} = 10.04$$

$$s = 0.653$$

$$CL = 0.8$$

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

$$df = n - 1 = 32$$

Determine the critical  $t$  value,  $t^*$ , such that  $P(|T| < t^*) = 0.8$ .

$$t^* = 1.31$$

Calculate the standard error.

$$SE = \frac{s}{\sqrt{n}} = \frac{0.653}{\sqrt{33}} = 0.114$$

We want to make an inference about the population mean.

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

Determine the bounds.

$$\begin{aligned} CI &= (\bar{x} - t^* SE, \bar{x} + t^* SE) \\ &= (10.04 - 1.31 \times 0.114, 10.04 + 1.31 \times 0.114) \\ &= (9.89, 10.2) \end{aligned}$$

We are 80% confident that the population mean is between 9.89 and 10.2.

6. A treatment group of size 33 has a mean of 109 and standard deviation of 14. A control group of size 29 has a mean of 101 and standard deviation of 11.1. If you decided to use a significance level of 0.01, 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 59.

- (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.

$$n_1 = 33$$

$$\bar{x}_1 = 109$$

$$s_1 = 14$$

$$n_2 = 29$$

$$\bar{x}_2 = 101$$

$$s_2 = 11.1$$

$$\alpha = 0.01$$

$$df = 59$$

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  $P(|T| > t^*) = 0.01$  by using a  $t$  table.

$$t^* = 2.66$$

Calculate the standard error.

$$\begin{aligned} SE &= \sqrt{\frac{(s_1)^2}{n_1} + \frac{(s_2)^2}{n_2}} \\ &= \sqrt{\frac{(14)^2}{33} + \frac{(11.1)^2}{29}} \\ &= 3.192 \end{aligned}$$

Determine the test statistic.

$$\begin{aligned} t_{\text{obs}} &= \frac{(\bar{x}_2 - \bar{x}_1) - (\mu_2 - \mu_1)_0}{SE} \\ &= \frac{(101 - 109) - (0)}{3.192} \\ &= -2.51 \end{aligned}$$

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 retain the null hypothesis.

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

- (b)  $H_A : \mu_2 - \mu_1 \neq 0$
- (c)  $t^* = 2.66$
- (d)  $SE = 3.192$
- (e)  $|t_{\text{obs}}| = 2.51$
- (f)  $0.01 < p\text{-value} < 0.02$
- (g) retain the null

7. From a very large population, a random sample of 5400 individuals was taken. In that sample, 87.8% were tasty. 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 condifence interval.

**Solution**

Determine  $z^*$  such that  $P(|Z| < z^*) = 0.8$ .

$$z^* = 1.28$$

Calculate the standard error.

$$SE = \sqrt{\frac{\hat{p}(1 - \hat{p})}{n}} = \sqrt{\frac{(0.878)(1 - 0.878)}{5400}} = 0.00445$$

Calculate the margin of error.

$$ME = z^* SE = (1.28)(0.00445) = 0.0057$$

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.

$$(0.872, 0.884)$$

We are 80% confident that the true population proportion is between 87.2% and 88.4%.

- (a) The lower bound = 0.872, which can also be expressed as 87.2%.
- (b) The upper bound = 0.884, which can also be expressed as 88.4%.

8. An experiment is run with a treatment group of size 165 and a control group of size 173. The results are summarized in the table below.

|            | treatment | control |
|------------|-----------|---------|
| folksy     | 110       | 94      |
| not folksy | 55        | 79      |

Using a significance level of 0.04, determine whether the treatment causes an effect on the proportion of cases that are folksy.

- (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  $P(|Z| > z^*) = 0.04$ .

$$z^* = \Phi^{-1} \left( 1 - \frac{\alpha}{2} \right) = 2.05$$

Determine the sample proportions.

$$\hat{p}_1 = \frac{110}{165} = 0.667$$

$$\hat{p}_2 = \frac{94}{173} = 0.543$$

Determine the difference of sample proportions.

$$\hat{p}_2 - \hat{p}_1 = 0.543 - 0.667 = -0.124$$

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

$$\hat{p} = \frac{110 + 94}{165 + 173} = 0.604$$

Determine the standard error.

$$\begin{aligned} SE &= \sqrt{\frac{\hat{p}(1 - \hat{p})}{n_1} + \frac{\hat{p}(1 - \hat{p})}{n_2}} \\ &= \sqrt{\frac{(0.604)(0.396)}{165} + \frac{(0.604)(0.396)}{173}} \\ &= 0.0532 \end{aligned}$$

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

$$H_0 : \hat{P}_2 - \hat{P}_1 \sim \mathcal{N}(0, 0.0532)$$

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.

$$\begin{aligned} z &= \frac{(\hat{p}_2 - \hat{p}_1) - (p_2 - p_1)_0}{SE} \\ &= \frac{(0.543 - 0.667) - 0}{0.0532} \\ &= -2.33 \end{aligned}$$

Determine the  $p$ -value.

$$\begin{aligned} p\text{-value} &= 2 \cdot \Phi(-|z|) \\ &= 2 \cdot \Phi(-2.33) \\ &= 0.0198 \end{aligned}$$

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)  $z^* = 2.05$
- (d)  $SE = 0.0532$
- (e)  $|Z_{\text{obs}}| = 2.33$
- (f)  $p\text{-value} = 0.0198$
- (g) reject the null

# Bunker Hill Community College

Final Statistics Exam 2019-05-02

Exam ID 012

**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)  $P(\text{pig given black}) = 0.117$

(b)  $P(\text{pig and green}) = 0.0868$

(c)  $P(\text{green}) = 0.299$

(d)  $P(\text{horn or green}) = 0.483$

(e)  $P(\text{green given bike}) = 0.156$

(f)  $P(\text{wheel}) = 0.166$

2.  $P(\text{"not shovel" given "not yellow"}) = 0.577$

3.  $P(60.8 < X < 61.26) = 0.7409$

4. (a)  $P(X = 18) = 0.0924$

(b)  $P(22 \leq X \leq 29) = 0.2799$

5. **(13.8, 16)**

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

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

(c)  $t^* = 2.46$

(d)  $SE = 0.756$

(e)  $|t_{\text{obs}}| = 2.51$

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

(g) **reject**

7. (a) **LB of p CI = 0.274 or 27.4%**

(b) **UB of p CI = 0.31 or 31%**

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)  $|Z_{\text{obs}}| = 2.08$

(f)  $p\text{-value} = 0.0376$

(g) **reject**

1. In a deck of strange cards, there are 1002 cards. Each card has an image and a color. The amounts are shown in the table below.

|        | black | green | orange |
|--------|-------|-------|--------|
| bike   | 41    | 20    | 67     |
| flower | 88    | 85    | 79     |
| horn   | 91    | 71    | 93     |
| pig    | 38    | 87    | 76     |
| wheel  | 66    | 37    | 63     |

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

**Solution**

(a)  $P(\text{pig given black}) = \frac{38}{41+88+91+38+66} = 0.117$

(b)  $P(\text{pig and green}) = \frac{87}{1002} = 0.0868$

(c)  $P(\text{green}) = \frac{20+85+71+87+37}{1002} = 0.299$

(d)  $P(\text{horn or green}) = \frac{91+71+93+20+85+71+87+37-71}{1002} = 0.483$

(e)  $P(\text{green given bike}) = \frac{20}{41+20+67} = 0.156$

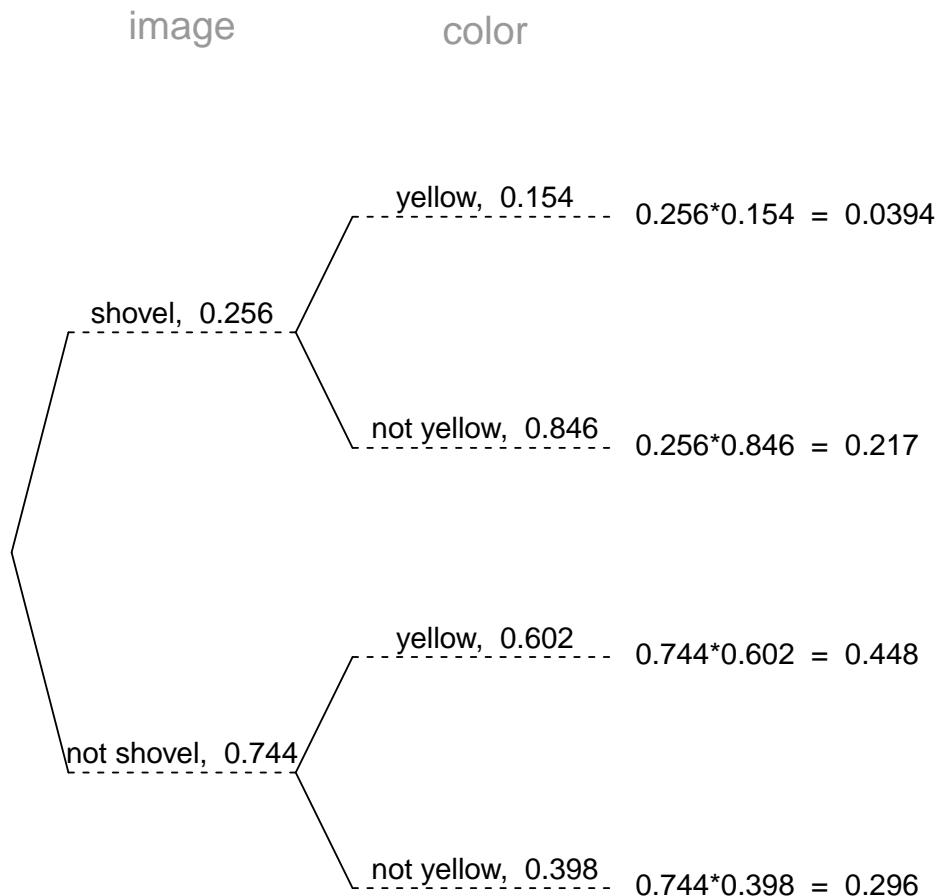
(f)  $P(\text{wheel}) = \frac{66+37+63}{1002} = 0.166$

2. In a deck of strange cards, each card has an image and a color. The chance of drawing a shovel is 25.6%. If a shovel is drawn, there is a 15.4% chance that it is yellow. If a card that is not a shovel is drawn, there is a 60.2% chance that it is yellow.

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

**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.

$$P(\text{"not shovel" given "not yellow"}) = \frac{0.296}{0.296 + 0.217} = 0.577$$

3. In a very large pile of toothpicks, the mean length is 61.04 millimeters and the standard deviation is 2.45 millimeters. If you randomly sample 144 toothpicks, what is the chance the sample mean is between 60.8 and 61.26 millimeters?

**Solution**

Label the given information.

$$\mu = 61.04$$

$$\sigma = 2.45$$

$$n = 144$$

$$\bar{x}_{\text{lower}} = 60.8$$

$$\bar{x}_{\text{upper}} = 61.26$$

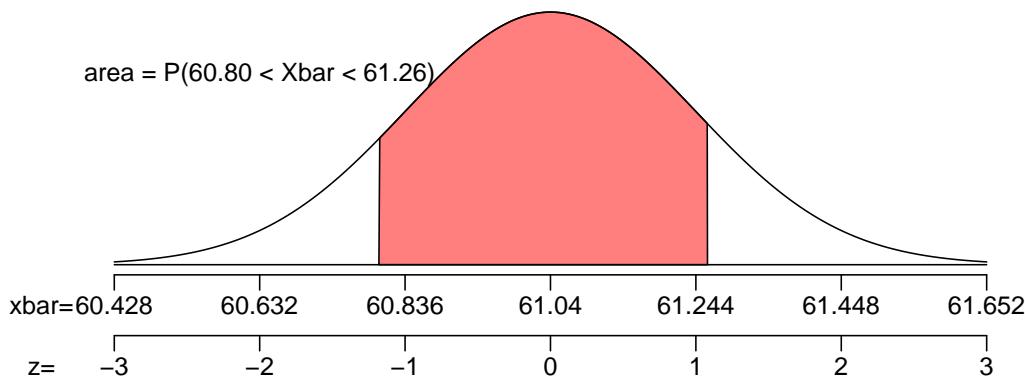
Find the standard error.

$$SE = \frac{\sigma}{\sqrt{n}} = \frac{2.45}{\sqrt{144}} = 0.204$$

Describe the sampling distribution.

$$\bar{X} \sim \mathcal{N}(61.04, 0.204)$$

Draw a sketch.



Calculate a  $z$  scores.

$$z_{\text{lower}} = \frac{\bar{x}_{\text{lower}} - \mu}{SE} = \frac{60.8 - 61.04}{0.204} = -1.18$$

$$z_{\text{upper}} = \frac{\bar{x}_{\text{upper}} - \mu}{SE} = \frac{61.26 - 61.04}{0.204} = 1.08$$

Determine the probability.

$$\begin{aligned} P(60.8 < \bar{X} < 61.26) &= \Phi(z_{\text{upper}}) - \Phi(z_{\text{lower}}) \\ &= \Phi(1.08) - \Phi(-1.18) \\ &= 0.7409 \end{aligned}$$

4. In a game, there is a 18% chance to win a round. You will play 110 rounds.
- (a) What is the probability of winning exactly 18 rounds?
  - (b) What is the probability of winning at least 22 but at most 29 rounds?

**Solution**

We use the formula for binomial probabilities.

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

$$P(X = 18) = \binom{110}{18} (0.18)^{18} (1 - 0.18)^{110-18}$$

$$P(X = 18) = \binom{110}{18} (0.18)^{18} (0.82)^{92}$$

$$P(X = 18) = 0.0924$$

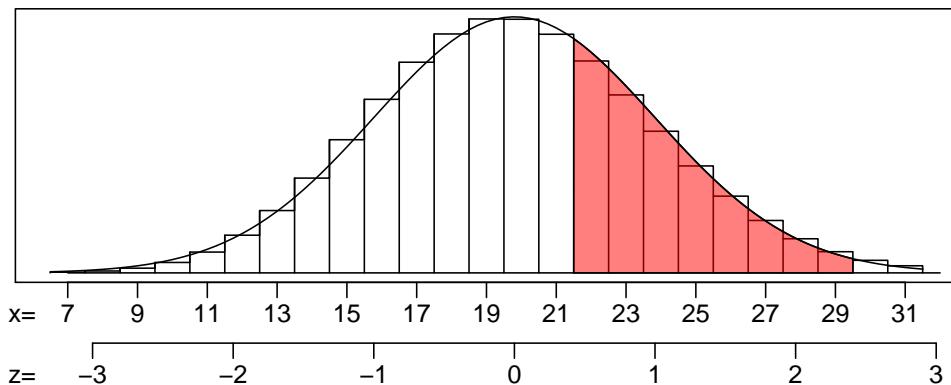
Find the mean.

$$\mu = np = (110)(0.18) = 19.8$$

Find the standard deviation.

$$\sigma = \sqrt{np(1 - p)} = \sqrt{(110)(0.18)(1 - 0.18)} = 4.0294$$

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



Find the z scores.

$$z_1 = \frac{21.5 - 19.8}{4.0294} = 0.55$$

$$z_2 = \frac{29.5 - 19.8}{4.0294} = 2.28$$

Calculate the probability.

$$P(22 \leq X \leq 29) = \Phi(2.28) - \Phi(0.55) = 0.2799$$

(a)  $P(X = 18) = 0.0924$

(b)  $P(22 \leq X \leq 29) = 0.2799$

5. As an ornithologist, you wish to determine the average body mass of *Passerina cyanea*. You randomly sample 24 adults of *Passerina cyanea*, resulting in a sample mean of 14.92 grams and a sample standard deviation of 2.16 grams. Determine a 98% confidence interval of the true population mean.

**Solution**

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

$$n = 24$$

$$\bar{x} = 14.92$$

$$s = 2.16$$

$$CL = 0.98$$

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

$$df = n - 1 = 23$$

Determine the critical  $t$  value,  $t^*$ , such that  $P(|T| < t^*) = 0.98$ .

$$t^* = 2.5$$

Calculate the standard error.

$$SE = \frac{s}{\sqrt{n}} = \frac{2.16}{\sqrt{24}} = 0.441$$

We want to make an inference about the population mean.

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

Determine the bounds.

$$\begin{aligned} CI &= (\bar{x} - t^* SE, \bar{x} + t^* SE) \\ &= (14.92 - 2.5 \times 0.441, 14.92 + 2.5 \times 0.441) \\ &= (13.8, 16) \end{aligned}$$

We are 98% confident that the population mean is between 13.8 and 16.

6. A treatment group of size 17 has a mean of 12.1 and standard deviation of 1.9. A control group of size 17 has a mean of 10.2 and standard deviation of 2.47. 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 30.

- (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.

$$n_1 = 17$$

$$\bar{x}_1 = 12.1$$

$$s_1 = 1.9$$

$$n_2 = 17$$

$$\bar{x}_2 = 10.2$$

$$s_2 = 2.47$$

$$\alpha = 0.02$$

$$df = 30$$

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  $P(|T| > t^*) = 0.02$  by using a  $t$  table.

$$t^* = 2.46$$

Calculate the standard error.

$$\begin{aligned} SE &= \sqrt{\frac{(s_1)^2}{n_1} + \frac{(s_2)^2}{n_2}} \\ &= \sqrt{\frac{(1.9)^2}{17} + \frac{(2.47)^2}{17}} \\ &= 0.756 \end{aligned}$$

Determine the test statistic.

$$\begin{aligned} t_{\text{obs}} &= \frac{(\bar{x}_2 - \bar{x}_1) - (\mu_2 - \mu_1)_0}{SE} \\ &= \frac{(10.2 - 12.1) - (0)}{0.756} \\ &= -2.51 \end{aligned}$$

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)  $t^* = 2.46$
- (d)  $SE = 0.756$
- (e)  $|t_{\text{obs}}| = 2.51$
- (f)  $0.01 < p\text{-value} < 0.02$
- (g) reject the null

7. From a very large population, a random sample of 5300 individuals was taken. In that sample, 29.2% were super. Determine a 99.5% confidence interval of the population proportion.

- (a) Find the lower bound of the confidence interval.
- (b) Find the upper bound of the condifence interval.

**Solution**

Determine  $z^*$  such that  $P(|Z| < z^*) = 0.995$ .

$$z^* = 2.81$$

Calculate the standard error.

$$SE = \sqrt{\frac{\hat{p}(1 - \hat{p})}{n}} = \sqrt{\frac{(0.292)(1 - 0.292)}{5300}} = 0.00625$$

Calculate the margin of error.

$$ME = z^* SE = (2.81)(0.00625) = 0.0176$$

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.

$$(0.274, 0.31)$$

We are 99.5% confident that the true population proportion is between 27.4% and 31%.

- (a) The lower bound = 0.274, which can also be expressed as 27.4%.
- (b) The upper bound = 0.31, which can also be expressed as 31%.

8. An experiment is run with a treatment group of size 117 and a control group of size 71. The results are summarized in the table below.

|          | treatment | control |
|----------|-----------|---------|
| sick     | 92        | 46      |
| not sick | 25        | 25      |

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

- (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.

$$\hat{p}_1 = \frac{92}{117} = 0.786$$

$$\hat{p}_2 = \frac{46}{71} = 0.648$$

Determine the difference of sample proportions.

$$\hat{p}_2 - \hat{p}_1 = 0.648 - 0.786 = -0.138$$

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

$$\hat{p} = \frac{92 + 46}{117 + 71} = 0.734$$

Determine the standard error.

$$\begin{aligned} SE &= \sqrt{\frac{\hat{p}(1 - \hat{p})}{n_1} + \frac{\hat{p}(1 - \hat{p})}{n_2}} \\ &= \sqrt{\frac{(0.734)(0.266)}{117} + \frac{(0.734)(0.266)}{71}} \\ &= 0.0665 \end{aligned}$$

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

$$H_0 : \hat{P}_2 - \hat{P}_1 \sim \mathcal{N}(0, 0.0665)$$

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.

$$\begin{aligned} z &= \frac{(\hat{p}_2 - \hat{p}_1) - (p_2 - p_1)_0}{SE} \\ &= \frac{(0.648 - 0.786) - 0}{0.0665} \\ &= -2.08 \end{aligned}$$

Determine the  $p$ -value.

$$\begin{aligned} p\text{-value} &= 2 \cdot \Phi(-|z|) \\ &= 2 \cdot \Phi(-2.08) \\ &= 0.0376 \end{aligned}$$

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)  $SE = 0.0665$
- (e)  $|Z_{\text{obs}}| = 2.08$
- (f)  $p\text{-value} = 0.0376$
- (g) reject the null

# Bunker Hill Community College

Final Statistics Exam 2019-05-02

Exam ID 013

**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)  $P(\text{orange given dog}) = 0.335$

(b)  $P(\text{horn and gray}) = 0.0253$

(c)  $P(\text{horn given gray}) = 0.128$

(d)  $P(\text{dog}) = 0.322$

(e)  $P(\text{gray}) = 0.197$

(f)  $P(\text{horn or red}) = 0.402$

2.  $P(\text{"cat" given "orange"}) = 0.432$

3.  $P(70.05 < X < 70.54) = 0.6763$

4. (a)  $P(X = 28) = 0.0814$

(b)  $P(33 \leq X \leq 40) = 0.1555$

5.  $(9.3, 9.94)$

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

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

(c)  $t^* = 1.68$

(d)  $SE = 0.021$

(e)  $|t_{\text{obs}}| = 1.46$

(f)  $0.1 < p\text{-value} < 0.2$

(g) **retain**

7. (a) **LB of p CI = 0.412 or 41.2%**

(b) **UB of p CI = 0.506 or 50.6%**

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

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

(c)  $Z^* = 2.81$

(d)  $SE = 0.118$

(e)  $|Z_{\text{obs}}| = 2.68$

(f)  $p\text{-value} = 0.0074$

(g) **retain**

1. In a deck of strange cards, there are 752 cards. Each card has an image and a color. The amounts are shown in the table below.

|       | gray | green | orange | red |
|-------|------|-------|--------|-----|
| dog   | 61   | 76    | 81     | 24  |
| horn  | 19   | 63    | 34     | 85  |
| wheel | 68   | 90    | 74     | 77  |

- (a) What is the probability a random card is orange given it is a dog?
- (b) What is the probability a random card is both a horn and gray?
- (c) What is the probability a random card is a horn given it is gray?
- (d) What is the probability a random card is a dog?
- (e) What is the probability a random card is gray?
- (f) What is the probability a random card is either a horn or red (or both)?

**Solution**

$$(a) P(\text{orange given dog}) = \frac{81}{61+76+81+24} = 0.335$$

$$(b) P(\text{horn and gray}) = \frac{19}{752} = 0.0253$$

$$(c) P(\text{horn given gray}) = \frac{19}{61+19+68} = 0.128$$

$$(d) P(\text{dog}) = \frac{61+76+81+24}{752} = 0.322$$

$$(e) P(\text{gray}) = \frac{61+19+68}{752} = 0.197$$

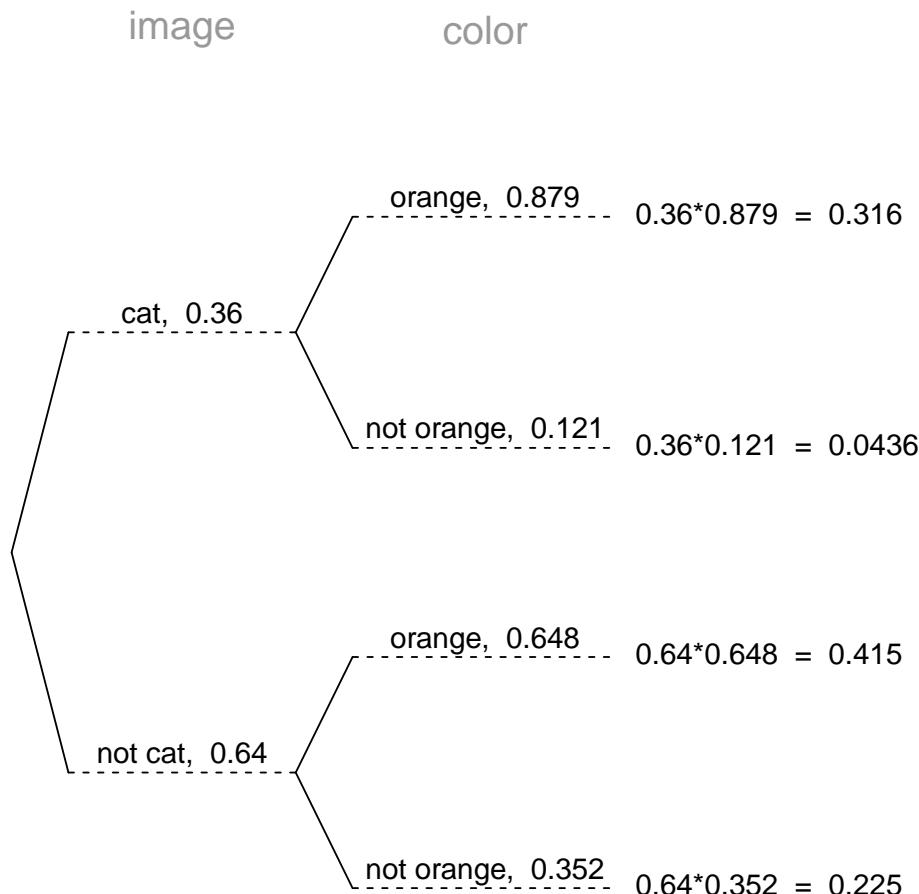
$$(f) P(\text{horn or red}) = \frac{19+63+34+85+24+85+77-85}{752} = 0.402$$

2. In a deck of strange cards, each card has an image and a color. The chance of drawing a cat is 36%. If a cat is drawn, there is a 87.9% chance that it is orange. If a card that is not a cat is drawn, there is a 64.8% chance that it is orange.

Now, someone draws a random card and reveals it is orange. What is the chance the card is a cat?

**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.

$$P(\text{"cat" given "orange"}) = \frac{0.316}{0.316 + 0.415} = 0.432$$

3. In a very large pile of toothpicks, the mean length is 70.4 millimeters and the standard deviation is 3.36 millimeters. If you randomly sample 225 toothpicks, what is the chance the sample mean is between 70.05 and 70.54 millimeters?

**Solution**

Label the given information.

$$\mu = 70.4$$

$$\sigma = 3.36$$

$$n = 225$$

$$\bar{x}_{\text{lower}} = 70.05$$

$$\bar{x}_{\text{upper}} = 70.54$$

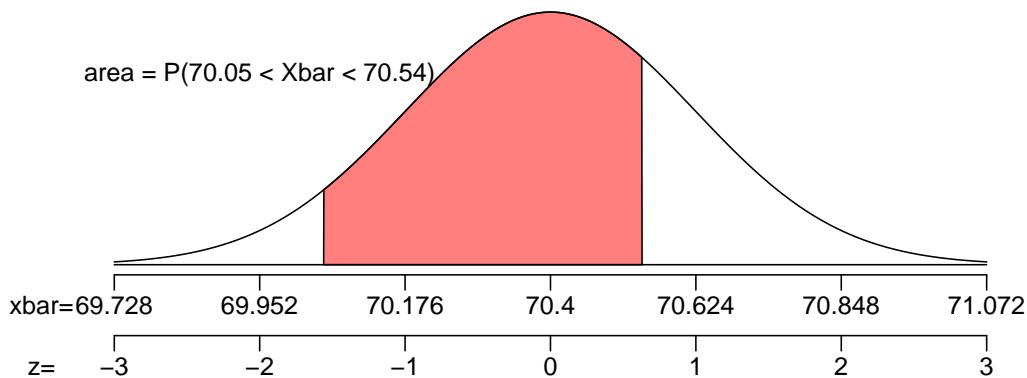
Find the standard error.

$$SE = \frac{\sigma}{\sqrt{n}} = \frac{3.36}{\sqrt{225}} = 0.224$$

Describe the sampling distribution.

$$\bar{X} \sim \mathcal{N}(70.4, 0.224)$$

Draw a sketch.



Calculate a  $z$  scores.

$$z_{\text{lower}} = \frac{\bar{x}_{\text{lower}} - \mu}{SE} = \frac{70.05 - 70.4}{0.224} = -1.56$$

$$z_{\text{upper}} = \frac{\bar{x}_{\text{upper}} - \mu}{SE} = \frac{70.54 - 70.4}{0.224} = 0.63$$

Determine the probability.

$$\begin{aligned} P(70.05 < X < 70.54) &= \Phi(z_{\text{upper}}) - \Phi(z_{\text{lower}}) \\ &= \Phi(0.63) - \Phi(-1.56) \\ &= 0.6763 \end{aligned}$$

4. In a game, there is a 15% chance to win a round. You will play 188 rounds.
- (a) What is the probability of winning exactly 28 rounds?
  - (b) What is the probability of winning at least 33 but at most 40 rounds?

**Solution**

We use the formula for binomial probabilities.

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

$$P(X = 28) = \binom{188}{28} (0.15)^{28} (1 - 0.15)^{188-28}$$

$$P(X = 28) = \binom{188}{28} (0.15)^{28} (0.85)^{160}$$

$$P(X = 28) = 0.0814$$

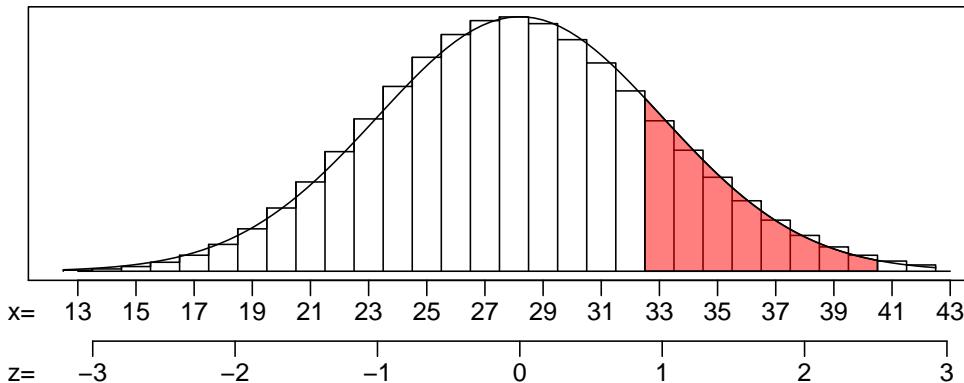
Find the mean.

$$\mu = np = (188)(0.15) = 28.2$$

Find the standard deviation.

$$\sigma = \sqrt{np(1 - p)} = \sqrt{(188)(0.15)(1 - 0.15)} = 4.8959$$

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



Find the z scores.

$$z_1 = \frac{32.5 - 28.2}{4.8959} = 0.98$$

$$z_2 = \frac{40.5 - 28.2}{4.8959} = 2.41$$

Calculate the probability.

$$P(33 \leq X \leq 40) = \Phi(2.41) - \Phi(0.98) = 0.1555$$

(a)  $P(X = 28) = 0.0814$

(b)  $P(33 \leq X \leq 40) = 0.1555$

5. As an ornithologist, you wish to determine the average body mass of *Dendroica magnolia*. You randomly sample 24 adults of *Dendroica magnolia*, resulting in a sample mean of 9.62 grams and a sample standard deviation of 1.17 grams. Determine a 80% confidence interval of the true population mean.

**Solution**

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

$$n = 24$$

$$\bar{x} = 9.62$$

$$s = 1.17$$

$$CL = 0.8$$

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

$$df = n - 1 = 23$$

Determine the critical  $t$  value,  $t^*$ , such that  $P(|T| < t^*) = 0.8$ .

$$t^* = 1.32$$

Calculate the standard error.

$$SE = \frac{s}{\sqrt{n}} = \frac{1.17}{\sqrt{24}} = 0.239$$

We want to make an inference about the population mean.

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

Determine the bounds.

$$\begin{aligned} CI &= (\bar{x} - t^* SE, \bar{x} + t^* SE) \\ &= (9.62 - 1.32 \times 0.239, 9.62 + 1.32 \times 0.239) \\ &= (9.3, 9.94) \end{aligned}$$

We are 80% confident that the population mean is between 9.3 and 9.94.

6. A treatment group of size 27 has a mean of 1.02 and standard deviation of 0.066. A control group of size 23 has a mean of 1.05 and standard deviation of 0.0778. If you decided to use a significance level of 0.1, 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 43.

- (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.

$$n_1 = 27$$

$$\bar{x}_1 = 1.02$$

$$s_1 = 0.066$$

$$n_2 = 23$$

$$\bar{x}_2 = 1.05$$

$$s_2 = 0.0778$$

$$\alpha = 0.1$$

$$df = 43$$

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  $P(|T| > t^*) = 0.1$  by using a  $t$  table.

$$t^* = 1.68$$

Calculate the standard error.

$$\begin{aligned} SE &= \sqrt{\frac{(s_1)^2}{n_1} + \frac{(s_2)^2}{n_2}} \\ &= \sqrt{\frac{(0.066)^2}{27} + \frac{(0.0778)^2}{23}} \\ &= 0.021 \end{aligned}$$

Determine the test statistic.

$$\begin{aligned} t_{\text{obs}} &= \frac{(\bar{x}_2 - \bar{x}_1) - (\mu_2 - \mu_1)_0}{SE} \\ &= \frac{(1.05 - 1.02) - (0)}{0.021} \\ &= 1.46 \end{aligned}$$

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

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

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

$$0.1 < p\text{-value} < 0.2$$

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)  $t^* = 1.68$
- (d)  $SE = 0.021$
- (e)  $|t_{\text{obs}}| = 1.46$
- (f)  $0.1 < p\text{-value} < 0.2$
- (g) retain the null

7. From a very large population, a random sample of 440 individuals was taken. In that sample, 45.9% were tasty. 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 condifence interval.

**Solution**

Determine  $z^*$  such that  $P(|Z| < z^*) = 0.95$ .

$$z^* = 1.96$$

Calculate the standard error.

$$SE = \sqrt{\frac{\hat{p}(1 - \hat{p})}{n}} = \sqrt{\frac{(0.459)(1 - 0.459)}{440}} = 0.0238$$

Calculate the margin of error.

$$ME = z^* SE = (1.96)(0.0238) = 0.0466$$

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.

$$(0.412, 0.506)$$

We are 95% confident that the true population proportion is between 41.2% and 50.6%.

- (a) The lower bound = 0.412, which can also be expressed as 41.2%.
- (b) The upper bound = 0.506, which can also be expressed as 50.6%.

8. An experiment is run with a treatment group of size 49 and a control group of size 20. The results are summarized in the table below.

|           | treatment | control |
|-----------|-----------|---------|
| green     | 9         | 10      |
| not green | 40        | 10      |

Using a significance level of 0.005, determine whether the treatment causes an effect on the proportion of cases that are green.

- (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  $P(|Z| > z^*) = 0.005$ .

$$z^* = \Phi^{-1} \left( 1 - \frac{\alpha}{2} \right) = 2.81$$

Determine the sample proportions.

$$\hat{p}_1 = \frac{9}{49} = 0.184$$

$$\hat{p}_2 = \frac{10}{20} = 0.5$$

Determine the difference of sample proportions.

$$\hat{p}_2 - \hat{p}_1 = 0.5 - 0.184 = 0.316$$

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

$$\hat{p} = \frac{9 + 10}{49 + 20} = 0.275$$

Determine the standard error.

$$\begin{aligned} SE &= \sqrt{\frac{\hat{p}(1 - \hat{p})}{n_1} + \frac{\hat{p}(1 - \hat{p})}{n_2}} \\ &= \sqrt{\frac{(0.275)(0.725)}{49} + \frac{(0.275)(0.725)}{20}} \\ &= 0.118 \end{aligned}$$

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

$$H_0 : \hat{P}_2 - \hat{P}_1 \sim \mathcal{N}(0, 0.118)$$

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.

$$\begin{aligned} z &= \frac{(\hat{p}_2 - \hat{p}_1) - (p_2 - p_1)_0}{SE} \\ &= \frac{(0.5 - 0.184) - 0}{0.118} \\ &= 2.68 \end{aligned}$$

Determine the  $p$ -value.

$$\begin{aligned} p\text{-value} &= 2 \cdot \Phi(-|z|) \\ &= 2 \cdot \Phi(-2.68) \\ &= 0.0074 \end{aligned}$$

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

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

So, we retain the null hypothesis. Thus the difference in proportions is not significant.

- (a)  $H_0 : p_2 - p_1 = 0$
- (b)  $H_A : p_2 - p_1 \neq 0$
- (c)  $z^* = 2.81$
- (d)  $SE = 0.118$
- (e)  $|z_{\text{obs}}| = 2.68$
- (f)  $p\text{-value} = 0.0074$
- (g) retain the null

# Bunker Hill Community College

Final Statistics Exam 2019-05-02

Exam ID 014

**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)  $P(\text{horn or red}) = 0.449$

(b)  $P(\text{horn}) = 0.242$

(c)  $P(\text{wheel and red}) = 0.0746$

(d)  $P(\text{orange}) = 0.357$

(e)  $P(\text{shovel given orange}) = 0.0514$

(f)  $P(\text{teal given dog}) = 0.453$

2.  $P(\text{"tree" given "green"}) = 0.0532$

3.  $P(61.96 < X < 62.41) = 0.6884$

4. (a)  $P(X = 32) = 0.1093$

(b)  $P(29 \leq X \leq 35) = 0.6159$

5. **(11.5, 12.6)**

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

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

(c)  $t^* = 2.13$

(d)  $SE = 0.049$

(e)  $|t_{\text{obs}}| = 1.83$

(f)  $0.05 < p\text{-value} < 0.1$

(g) **retain**

7. (a) **LB of p CI = 0.496 or 49.6%**

(b) **UB of p CI = 0.516 or 51.6%**

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

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

(c)  $Z^* = 2.81$

(d)  $SE = 0.141$

(e)  $|Z_{\text{obs}}| = 2.75$

(f)  $p\text{-value} = 0.006$

(g) **retain**

1. In a deck of strange cards, there are 818 cards. Each card has an image and a color. The amounts are shown in the table below.

|        | orange | red | teal |
|--------|--------|-----|------|
| dog    | 75     | 41  | 96   |
| gem    | 63     | 25  | 52   |
| horn   | 94     | 21  | 83   |
| shovel | 15     | 42  | 72   |
| wheel  | 45     | 61  | 33   |

- (a) What is the probability a random card is either a horn or red (or both)?
- (b) What is the probability a random card is a horn?
- (c) What is the probability a random card is both a wheel and red?
- (d) What is the probability a random card is orange?
- (e) What is the probability a random card is a shovel given it is orange?
- (f) What is the probability a random card is teal given it is a dog?

**Solution**

$$(a) P(\text{horn or red}) = \frac{94+21+83+41+25+21+42+61-21}{818} = 0.449$$

$$(b) P(\text{horn}) = \frac{94+21+83}{818} = 0.242$$

$$(c) P(\text{wheel and red}) = \frac{61}{818} = 0.0746$$

$$(d) P(\text{orange}) = \frac{75+63+94+15+45}{818} = 0.357$$

$$(e) P(\text{shovel given orange}) = \frac{15}{75+63+94+15+45} = 0.0514$$

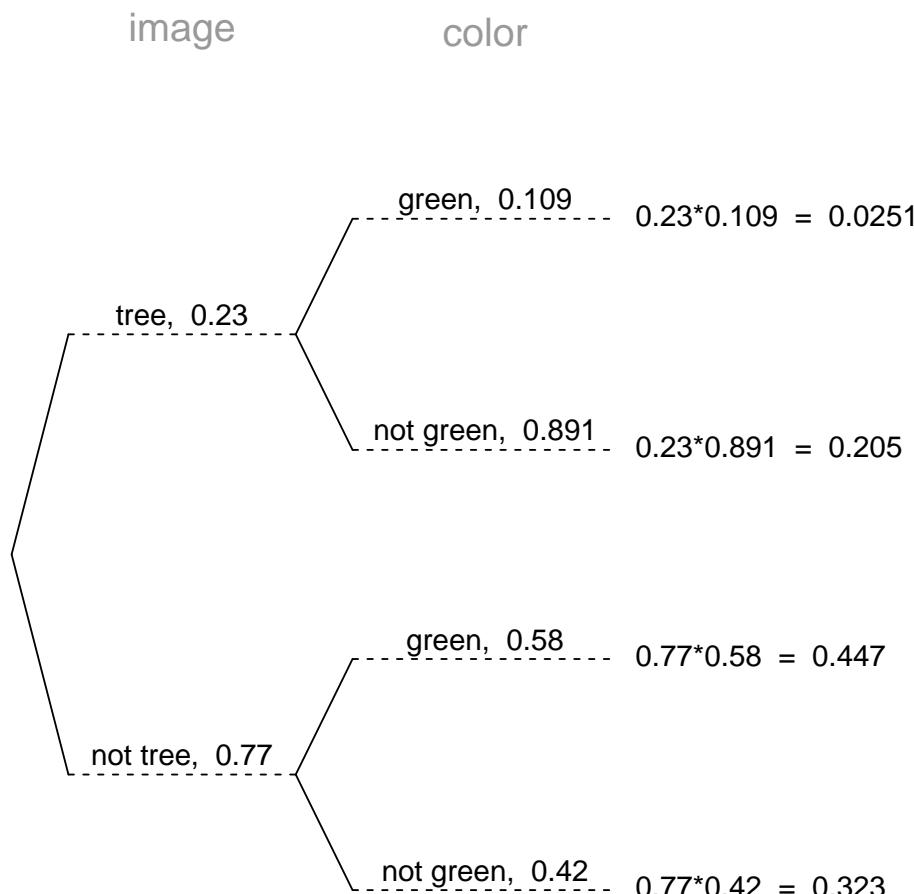
$$(f) P(\text{teal given dog}) = \frac{96}{75+41+96} = 0.453$$

2. In a deck of strange cards, each card has an image and a color. The chance of drawing a tree is 23%. If a tree is drawn, there is a 10.9% chance that it is green. If a card that is not a tree is drawn, there is a 58% chance that it is green.

Now, someone draws a random card and reveals it is green. What is the chance the card is a tree?

**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.

$$P(\text{"tree" given "green"}) = \frac{0.0251}{0.0251 + 0.447} = 0.0532$$

3. In a very large pile of toothpicks, the mean length is 62.13 millimeters and the standard deviation is 3.02 millimeters. If you randomly sample 196 toothpicks, what is the chance the sample mean is between 61.96 and 62.41 millimeters?

**Solution**

Label the given information.

$$\mu = 62.13$$

$$\sigma = 3.02$$

$$n = 196$$

$$\bar{x}_{\text{lower}} = 61.96$$

$$\bar{x}_{\text{upper}} = 62.41$$

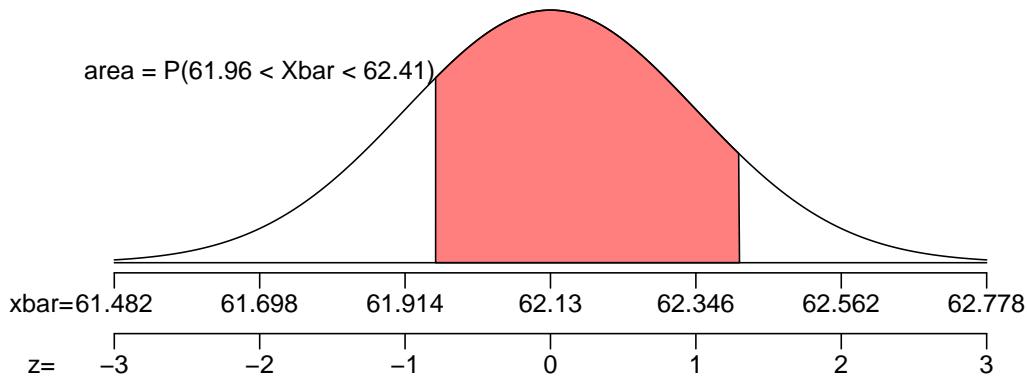
Find the standard error.

$$SE = \frac{\sigma}{\sqrt{n}} = \frac{3.02}{\sqrt{196}} = 0.216$$

Describe the sampling distribution.

$$\bar{X} \sim \mathcal{N}(62.13, 0.216)$$

Draw a sketch.



Calculate a  $z$  scores.

$$z_{\text{lower}} = \frac{\bar{x}_{\text{lower}} - \mu}{SE} = \frac{61.96 - 62.13}{0.216} = -0.79$$

$$z_{\text{upper}} = \frac{\bar{x}_{\text{upper}} - \mu}{SE} = \frac{62.41 - 62.13}{0.216} = 1.3$$

Determine the probability.

$$\begin{aligned} P(61.96 < \bar{X} < 62.41) &= \Phi(z_{\text{upper}}) - \Phi(z_{\text{lower}}) \\ &= \Phi(1.3) - \Phi(-0.79) \\ &= 0.6884 \end{aligned}$$

4. In a game, there is a 75% chance to win a round. You will play 45 rounds.
- (a) What is the probability of winning exactly 32 rounds?
  - (b) What is the probability of winning at least 29 but at most 35 rounds?

**Solution**

We use the formula for binomial probabilities.

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

$$P(X = 32) = \binom{45}{32} (0.75)^{32} (1 - 0.75)^{45-32}$$

$$P(X = 32) = \binom{45}{32} (0.75)^{32} (0.25)^{13}$$

$$P(X = 32) = 0.1093$$

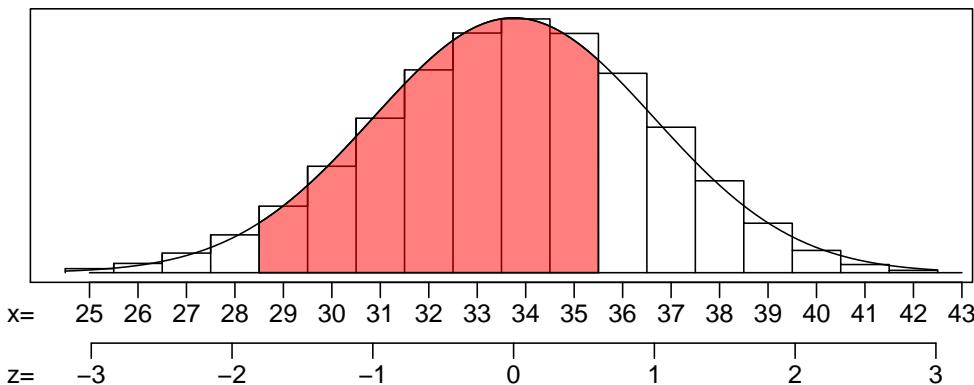
Find the mean.

$$\mu = np = (45)(0.75) = 33.75$$

Find the standard deviation.

$$\sigma = \sqrt{np(1 - p)} = \sqrt{(45)(0.75)(1 - 0.75)} = 2.9047$$

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



Find the z scores.

$$z_1 = \frac{28.5 - 33.75}{2.9047} = -1.64$$

$$z_2 = \frac{35.5 - 33.75}{2.9047} = 0.43$$

Calculate the probability.

$$P(29 \leq X \leq 35) = \Phi(0.43) - \Phi(-1.64) = 0.6159$$

(a)  $P(X = 32) = 0.1093$

(b)  $P(29 \leq X \leq 35) = 0.6159$

5. As an ornithologist, you wish to determine the average body mass of *Dendroica coronata*. You randomly sample 14 adults of *Dendroica coronata*, resulting in a sample mean of 12.09 grams and a sample standard deviation of 1.54 grams. Determine a 80% confidence interval of the true population mean.

**Solution**

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

$$n = 14$$

$$\bar{x} = 12.09$$

$$s = 1.54$$

$$CL = 0.8$$

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

$$df = n - 1 = 13$$

Determine the critical  $t$  value,  $t^*$ , such that  $P(|T| < t^*) = 0.8$ .

$$t^* = 1.35$$

Calculate the standard error.

$$SE = \frac{s}{\sqrt{n}} = \frac{1.54}{\sqrt{14}} = 0.412$$

We want to make an inference about the population mean.

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

Determine the bounds.

$$\begin{aligned} CI &= (\bar{x} - t^* SE, \bar{x} + t^* SE) \\ &= (12.09 - 1.35 \times 0.412, 12.09 + 1.35 \times 0.412) \\ &= (11.5, 12.6) \end{aligned}$$

We are 80% confident that the population mean is between 11.5 and 12.6.

6. A treatment group of size 19 has a mean of 1.12 and standard deviation of 0.166. A control group of size 25 has a mean of 1.03 and standard deviation of 0.155. 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 37.

- (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.

$$\begin{aligned}n_1 &= 19 \\ \bar{x}_1 &= 1.12 \\ s_1 &= 0.166 \\ n_2 &= 25 \\ \bar{x}_2 &= 1.03 \\ s_2 &= 0.155 \\ \alpha &= 0.04 \\ df &= 37\end{aligned}$$

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  $P(|T| > t^*) = 0.04$  by using a  $t$  table.

$$t^* = 2.13$$

Calculate the standard error.

$$\begin{aligned}SE &= \sqrt{\frac{(s_1)^2}{n_1} + \frac{(s_2)^2}{n_2}} \\ &= \sqrt{\frac{(0.166)^2}{19} + \frac{(0.155)^2}{25}} \\ &= 0.049\end{aligned}$$

Determine the test statistic.

$$\begin{aligned}t_{\text{obs}} &= \frac{(\bar{x}_2 - \bar{x}_1) - (\mu_2 - \mu_1)_0}{SE} \\ &= \frac{(1.03 - 1.12) - (0)}{0.049} \\ &= -1.83\end{aligned}$$

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

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

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

$$0.05 < p\text{-value} < 0.1$$

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)  $t^* = 2.13$
- (d)  $SE = 0.049$
- (e)  $|t_{\text{obs}}| = 1.83$
- (f)  $0.05 < p\text{-value} < 0.1$
- (g) retain the null

7. From a very large population, a random sample of 6300 individuals was taken. In that sample, 50.6% were angry. Determine a 90% confidence interval of the population proportion.

- (a) Find the lower bound of the confidence interval.
- (b) Find the upper bound of the condifence interval.

**Solution**

Determine  $z^*$  such that  $P(|Z| < z^*) = 0.9$ .

$$z^* = 1.64$$

Calculate the standard error.

$$SE = \sqrt{\frac{\hat{p}(1 - \hat{p})}{n}} = \sqrt{\frac{(0.506)(1 - 0.506)}{6300}} = 0.0063$$

Calculate the margin of error.

$$ME = z^* SE = (1.64)(0.0063) = 0.0103$$

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.

$$(0.496, 0.516)$$

We are 90% confident that the true population proportion is between 49.6% and 51.6%.

- (a) The lower bound = 0.496, which can also be expressed as 49.6%.
- (b) The upper bound = 0.516, which can also be expressed as 51.6%.

8. An experiment is run with a treatment group of size 55 and a control group of size 14. The results are summarized in the table below.

|           | treatment | control |
|-----------|-----------|---------|
| happy     | 14        | 9       |
| not happy | 41        | 5       |

Using a significance level of 0.005, determine whether the treatment causes an effect on the proportion of cases that are happy.

- (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  $P(|Z| > z^*) = 0.005$ .

$$z^* = \Phi^{-1} \left( 1 - \frac{\alpha}{2} \right) = 2.81$$

Determine the sample proportions.

$$\hat{p}_1 = \frac{14}{55} = 0.255$$

$$\hat{p}_2 = \frac{9}{14} = 0.643$$

Determine the difference of sample proportions.

$$\hat{p}_2 - \hat{p}_1 = 0.643 - 0.255 = 0.388$$

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

$$\hat{p} = \frac{14 + 9}{55 + 14} = 0.333$$

Determine the standard error.

$$\begin{aligned} SE &= \sqrt{\frac{\hat{p}(1 - \hat{p})}{n_1} + \frac{\hat{p}(1 - \hat{p})}{n_2}} \\ &= \sqrt{\frac{(0.333)(0.667)}{55} + \frac{(0.333)(0.667)}{14}} \\ &= 0.141 \end{aligned}$$

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

$$H_0 : \hat{P}_2 - \hat{P}_1 \sim \mathcal{N}(0, 0.141)$$

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.

$$\begin{aligned} z &= \frac{(\hat{p}_2 - \hat{p}_1) - (p_2 - p_1)_0}{SE} \\ &= \frac{(0.643 - 0.255) - 0}{0.141} \\ &= 2.75 \end{aligned}$$

Determine the  $p$ -value.

$$\begin{aligned} p\text{-value} &= 2 \cdot \Phi(-|z|) \\ &= 2 \cdot \Phi(-2.75) \\ &= 0.006 \end{aligned}$$

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

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

So, we retain the null hypothesis. Thus the difference in proportions is not significant.

- (a)  $H_0 : p_2 - p_1 = 0$
- (b)  $H_A : p_2 - p_1 \neq 0$
- (c)  $z^* = 2.81$
- (d)  $SE = 0.141$
- (e)  $|z_{\text{obs}}| = 2.75$
- (f)  $p\text{-value} = 0.006$
- (g) retain the null

# Bunker Hill Community College

Final Statistics Exam 2019-05-02

Exam ID 015

**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)  $P(\text{flower or white}) = 0.426$

(b)  $P(\text{pink given bike}) = 0.109$

(c)  $P(\text{orange}) = 0.218$

(d)  $P(\text{wheel and blue}) = 0.0372$

(e)  $P(\text{bike given blue}) = 0.128$

(f)  $P(\text{cat}) = 0.32$

2.  $P(\text{"not llama" given "yellow"}) = 0.886$

3.  $P(60.5 < X < 60.79) = 0.8051$

4. (a)  $P(X = 140) = 0.0519$

(b)  $P(144 \leq X \leq 150) = 0.307$

5. **(9.36, 11.3)**

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

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

(c)  $t^* = 1.69$

(d)  $SE = 8.867$

(e)  $|t_{\text{obs}}| = 1.76$

(f)  $0.05 < p\text{-value} < 0.1$

(g) **reject**

7. (a) **LB of p CI = 0.0974 or 9.74%**

(b) **UB of p CI = 0.113 or 11.3%**

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

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

(c)  $Z^* = 2.05$

(d)  $SE = 0.068$

(e)  $|Z_{\text{obs}}| = 2.21$

(f)  $p\text{-value} = 0.0272$

(g) **reject**

1. In a deck of strange cards, there are 1128 cards. Each card has an image and a color. The amounts are shown in the table below.

|        | blue | orange | pink | teal | white |
|--------|------|--------|------|------|-------|
| bike   | 32   | 28     | 18   | 68   | 19    |
| cat    | 85   | 95     | 52   | 74   | 55    |
| flower | 91   | 79     | 50   | 71   | 90    |
| wheel  | 42   | 44     | 94   | 16   | 25    |

- (a) What is the probability a random card is either a flower or white (or both)?
- (b) What is the probability a random card is pink given it is a bike?
- (c) What is the probability a random card is orange?
- (d) What is the probability a random card is both a wheel and blue?
- (e) What is the probability a random card is a bike given it is blue?
- (f) What is the probability a random card is a cat?

**Solution**

$$(a) P(\text{flower or white}) = \frac{91+79+50+71+90+19+55+90+25-90}{1128} = 0.426$$

$$(b) P(\text{pink given bike}) = \frac{18}{32+28+18+68+19} = 0.109$$

$$(c) P(\text{orange}) = \frac{28+95+79+44}{1128} = 0.218$$

$$(d) P(\text{wheel and blue}) = \frac{42}{1128} = 0.0372$$

$$(e) P(\text{bike given blue}) = \frac{32}{32+85+91+42} = 0.128$$

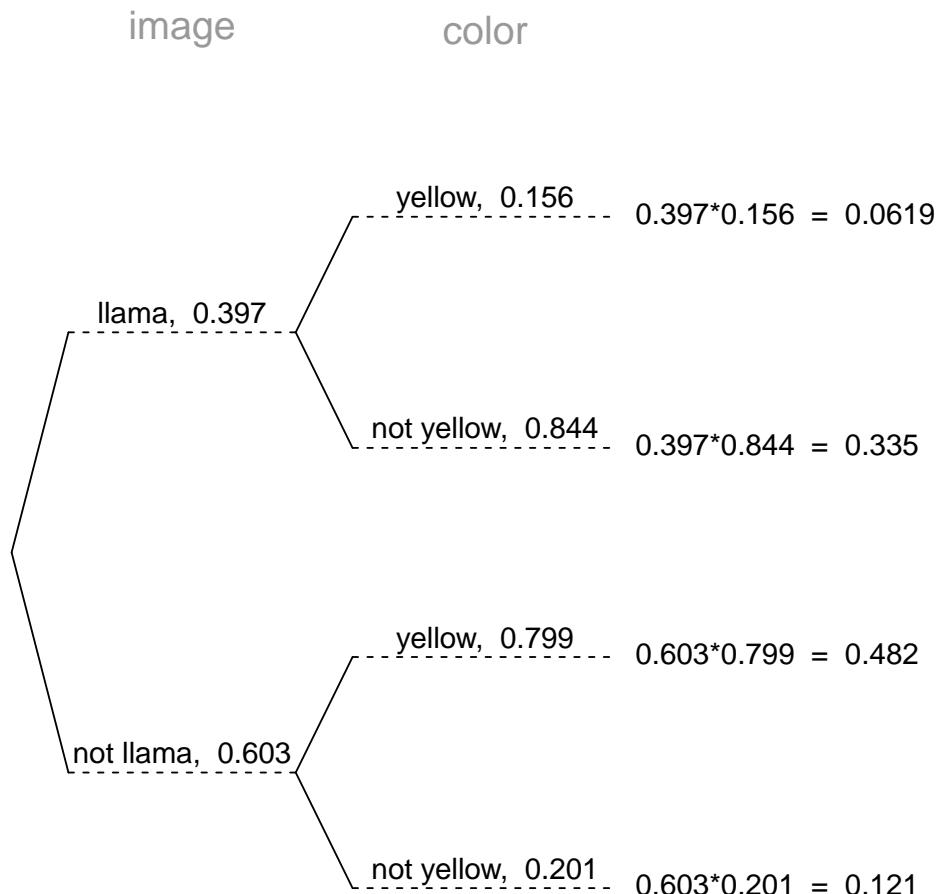
$$(f) P(\text{cat}) = \frac{85+95+52+74+55}{1128} = 0.32$$

2. In a deck of strange cards, each card has an image and a color. The chance of drawing a llama is 39.7%. If a llama is drawn, there is a 15.6% chance that it is yellow. If a card that is not a llama is drawn, there is a 79.9% chance that it is yellow.

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

**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.

$$P(\text{"not llama" given "yellow"}) = \frac{0.482}{0.482 + 0.0619} = 0.886$$

3. In a very large pile of toothpicks, the mean length is 60.6 millimeters and the standard deviation is 1.35 millimeters. If you randomly sample 175 toothpicks, what is the chance the sample mean is between 60.5 and 60.79 millimeters?

**Solution**

Label the given information.

$$\mu = 60.6$$

$$\sigma = 1.35$$

$$n = 175$$

$$\bar{x}_{\text{lower}} = 60.5$$

$$\bar{x}_{\text{upper}} = 60.79$$

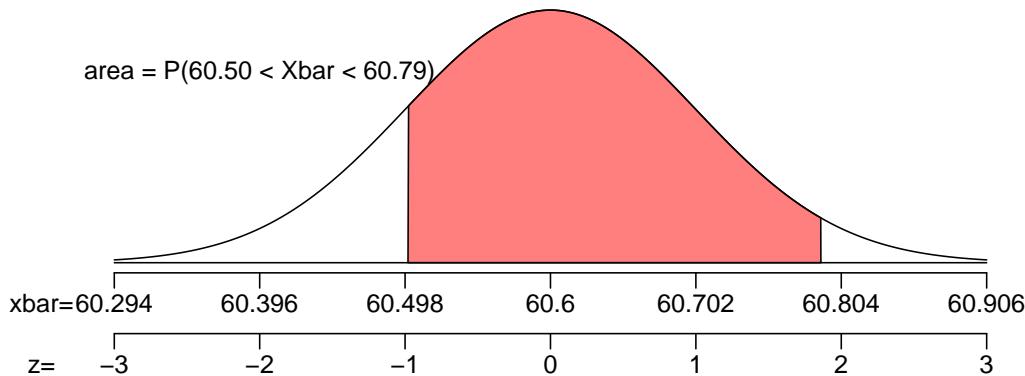
Find the standard error.

$$SE = \frac{\sigma}{\sqrt{n}} = \frac{1.35}{\sqrt{175}} = 0.102$$

Describe the sampling distribution.

$$\bar{X} \sim \mathcal{N}(60.6, 0.102)$$

Draw a sketch.



Calculate a  $z$  scores.

$$z_{\text{lower}} = \frac{\bar{x}_{\text{lower}} - \mu}{SE} = \frac{60.5 - 60.6}{0.102} = -0.98$$

$$z_{\text{upper}} = \frac{\bar{x}_{\text{upper}} - \mu}{SE} = \frac{60.79 - 60.6}{0.102} = 1.86$$

Determine the probability.

$$\begin{aligned} P(60.5 < \bar{X} < 60.79) &= \Phi(z_{\text{upper}}) - \Phi(z_{\text{lower}}) \\ &= \Phi(1.86) - \Phi(-0.98) \\ &= 0.8051 \end{aligned}$$

4. In a game, there is a 70% chance to win a round. You will play 205 rounds.
- (a) What is the probability of winning exactly 140 rounds?
  - (b) What is the probability of winning at least 144 but at most 150 rounds?

**Solution**

We use the formula for binomial probabilities.

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

$$P(X = 140) = \binom{205}{140} (0.7)^{140} (1 - 0.7)^{205-140}$$

$$P(X = 140) = \binom{205}{140} (0.7)^{140} (0.3)^{65}$$

$$P(X = 140) = 0.0519$$

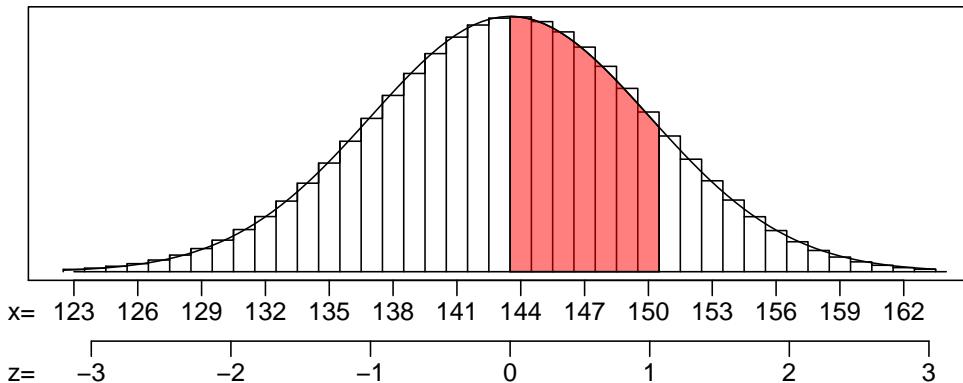
Find the mean.

$$\mu = np = (205)(0.7) = 143.5$$

Find the standard deviation.

$$\sigma = \sqrt{np(1 - p)} = \sqrt{(205)(0.7)(1 - 0.7)} = 6.5612$$

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



Find the z scores.

$$z_1 = \frac{143.5 - 143.5}{6.5612} = 0.08$$

$$z_2 = \frac{150.5 - 143.5}{6.5612} = 0.99$$

Calculate the probability.

$$P(144 \leq X \leq 150) = \Phi(0.99) - \Phi(0.08) = 0.307$$

(a)  $P(X = 140) = 0.0519$

(b)  $P(144 \leq X \leq 150) = 0.307$

5. As an ornithologist, you wish to determine the average body mass of *Dendroica dominica*. You randomly sample 25 adults of *Dendroica dominica*, resulting in a sample mean of 10.35 grams and a sample standard deviation of 1.99 grams. Determine a 98% confidence interval of the true population mean.

**Solution**

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

$$n = 25$$

$$\bar{x} = 10.35$$

$$s = 1.99$$

$$CL = 0.98$$

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

$$df = n - 1 = 24$$

Determine the critical  $t$  value,  $t^*$ , such that  $P(|T| < t^*) = 0.98$ .

$$t^* = 2.49$$

Calculate the standard error.

$$SE = \frac{s}{\sqrt{n}} = \frac{1.99}{\sqrt{25}} = 0.398$$

We want to make an inference about the population mean.

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

Determine the bounds.

$$\begin{aligned} CI &= (\bar{x} - t^* SE, \bar{x} + t^* SE) \\ &= (10.35 - 2.49 \times 0.398, 10.35 + 2.49 \times 0.398) \\ &= (9.36, 11.3) \end{aligned}$$

We are 98% confident that the population mean is between 9.36 and 11.3.

6. A treatment group of size 20 has a mean of 94.4 and standard deviation of 25. A control group of size 19 has a mean of 110 and standard deviation of 30. If you decided to use a significance level of 0.1, 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 35.

- (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.

$$n_1 = 20$$

$$\bar{x}_1 = 94.4$$

$$s_1 = 25$$

$$n_2 = 19$$

$$\bar{x}_2 = 110$$

$$s_2 = 30$$

$$\alpha = 0.1$$

$$df = 35$$

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  $P(|T| > t^*) = 0.1$  by using a  $t$  table.

$$t^* = 1.69$$

Calculate the standard error.

$$\begin{aligned} SE &= \sqrt{\frac{(s_1)^2}{n_1} + \frac{(s_2)^2}{n_2}} \\ &= \sqrt{\frac{(25)^2}{20} + \frac{(30)^2}{19}} \\ &= 8.867 \end{aligned}$$

Determine the test statistic.

$$\begin{aligned} t_{\text{obs}} &= \frac{(\bar{x}_2 - \bar{x}_1) - (\mu_2 - \mu_1)_0}{SE} \\ &= \frac{(110 - 94.4) - (0)}{8.867} \\ &= 1.76 \end{aligned}$$

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

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

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

$$0.05 < p\text{-value} < 0.1$$

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)  $t^* = 1.69$
- (d)  $SE = 8.867$
- (e)  $|t_{\text{obs}}| = 1.76$
- (f)  $0.05 < p\text{-value} < 0.1$
- (g) reject the null

7. From a very large population, a random sample of 8900 individuals was taken. In that sample, 10.5% were purple. Determine a 98% confidence interval of the population proportion.

- (a) Find the lower bound of the confidence interval.
- (b) Find the upper bound of the condifence interval.

**Solution**

Determine  $z^*$  such that  $P(|Z| < z^*) = 0.98$ .

$$z^* = 2.33$$

Calculate the standard error.

$$SE = \sqrt{\frac{\hat{p}(1 - \hat{p})}{n}} = \sqrt{\frac{(0.105)(1 - 0.105)}{8900}} = 0.00325$$

Calculate the margin of error.

$$ME = z^* SE = (2.33)(0.00325) = 0.00757$$

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.

$$(0.0974, 0.113)$$

We are 98% confident that the true population proportion is between 9.74% and 11.3%.

- (a) The lower bound = 0.0974, which can also be expressed as 9.74%.
- (b) The upper bound = 0.113, which can also be expressed as 11.3%.

8. An experiment is run with a treatment group of size 99 and a control group of size 118. The results are summarized in the table below.

|           | treatment | control |
|-----------|-----------|---------|
| green     | 61        | 55      |
| not green | 38        | 63      |

Using a significance level of 0.04, determine whether the treatment causes an effect on the proportion of cases that are green.

- (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  $P(|Z| > z^*) = 0.04$ .

$$z^* = \Phi^{-1} \left( 1 - \frac{\alpha}{2} \right) = 2.05$$

Determine the sample proportions.

$$\hat{p}_1 = \frac{61}{99} = 0.616$$

$$\hat{p}_2 = \frac{55}{118} = 0.466$$

Determine the difference of sample proportions.

$$\hat{p}_2 - \hat{p}_1 = 0.466 - 0.616 = -0.15$$

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

$$\hat{p} = \frac{61 + 55}{99 + 118} = 0.535$$

Determine the standard error.

$$\begin{aligned} SE &= \sqrt{\frac{\hat{p}(1 - \hat{p})}{n_1} + \frac{\hat{p}(1 - \hat{p})}{n_2}} \\ &= \sqrt{\frac{(0.535)(0.465)}{99} + \frac{(0.535)(0.465)}{118}} \\ &= 0.068 \end{aligned}$$

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

$$H_0 : \hat{P}_2 - \hat{P}_1 \sim \mathcal{N}(0, 0.068)$$

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.

$$\begin{aligned} z &= \frac{(\hat{p}_2 - \hat{p}_1) - (p_2 - p_1)_0}{SE} \\ &= \frac{(0.466 - 0.616) - 0}{0.068} \\ &= -2.21 \end{aligned}$$

Determine the  $p$ -value.

$$\begin{aligned} p\text{-value} &= 2 \cdot \Phi(-|z|) \\ &= 2 \cdot \Phi(-2.21) \\ &= 0.0272 \end{aligned}$$

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)  $z^* = 2.05$
- (d)  $SE = 0.068$
- (e)  $|z_{\text{obs}}| = 2.21$
- (f)  $p\text{-value} = 0.0272$
- (g) reject the null

# Bunker Hill Community College

Final Statistics Exam 2019-05-02

Exam ID 016

**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)  $P(\text{blue}) = 0.389$

(b)  $P(\text{red given gem}) = 0.294$

(c)  $P(\text{wheel and red}) = 0.0574$

(d)  $P(\text{gem or blue}) = 0.531$

(e)  $P(\text{gem given red}) = 0.234$

(f)  $P(\text{cat}) = 0.193$

2.  $P(\text{"not cat" given "violet"}) = 0.433$

3.  $P(60.84 < X < 61.37) = 0.8953$

4. (a)  $P(X = 49) = 0.0444$

(b)  $P(46 \leq X \leq 54) = 0.3119$

5.  $(8.95, 11.1)$

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

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

(c)  $t^* = 2.7$

(d)  $SE = 48.57$

(e)  $|t_{\text{obs}}| = 2.49$

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

(g) **retain**

7. (a) **LB of p CI = 0.429 or 42.9%**

(b) **UB of p CI = 0.451 or 45.1%**

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

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

(c)  $Z^* = 1.96$

(d)  $SE = 0.042$

(e)  $|Z_{\text{obs}}| = 2.14$

(f)  $p\text{-value} = 0.0324$

(g) **reject**

1. In a deck of strange cards, there are 1027 cards. Each card has an image and a color. The amounts are shown in the table below.

|       | blue | green | red |
|-------|------|-------|-----|
| bike  | 61   | 28    | 32  |
| cat   | 95   | 57    | 46  |
| dog   | 78   | 84    | 85  |
| gem   | 86   | 77    | 68  |
| wheel | 80   | 91    | 59  |

- (a) What is the probability a random card is blue?
- (b) What is the probability a random card is red given it is a gem?
- (c) What is the probability a random card is both a wheel and red?
- (d) What is the probability a random card is either a gem or blue (or both)?
- (e) What is the probability a random card is a gem given it is red?
- (f) What is the probability a random card is a cat?

**Solution**

(a)  $P(\text{blue}) = \frac{61+95+78+86+80}{1027} = 0.389$

(b)  $P(\text{red given gem}) = \frac{68}{86+77+68} = 0.294$

(c)  $P(\text{wheel and red}) = \frac{59}{1027} = 0.0574$

(d)  $P(\text{gem or blue}) = \frac{86+77+68+61+95+78+86+80-86}{1027} = 0.531$

(e)  $P(\text{gem given red}) = \frac{68}{32+46+85+68+59} = 0.234$

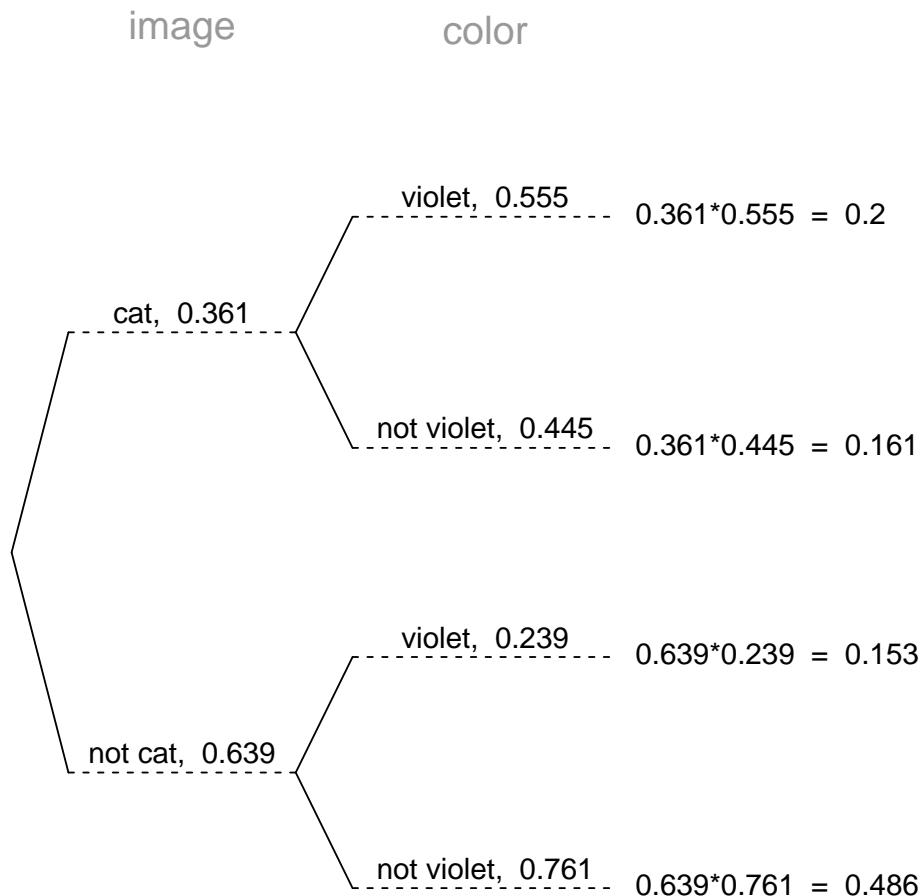
(f)  $P(\text{cat}) = \frac{95+57+46}{1027} = 0.193$

2. In a deck of strange cards, each card has an image and a color. The chance of drawing a cat is 36.1%. If a cat is drawn, there is a 55.5% chance that it is violet. If a card that is not a cat is drawn, there is a 23.9% chance that it is violet.

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

**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.

$$P(\text{"not cat" given "violet"}) = \frac{0.153}{0.153 + 0.2} = 0.433$$

3. In a very large pile of toothpicks, the mean length is 61.02 millimeters and the standard deviation is 1.96 millimeters. If you randomly sample 196 toothpicks, what is the chance the sample mean is between 60.84 and 61.37 millimeters?

**Solution**

Label the given information.

$$\mu = 61.02$$

$$\sigma = 1.96$$

$$n = 196$$

$$\bar{x}_{\text{lower}} = 60.84$$

$$\bar{x}_{\text{upper}} = 61.37$$

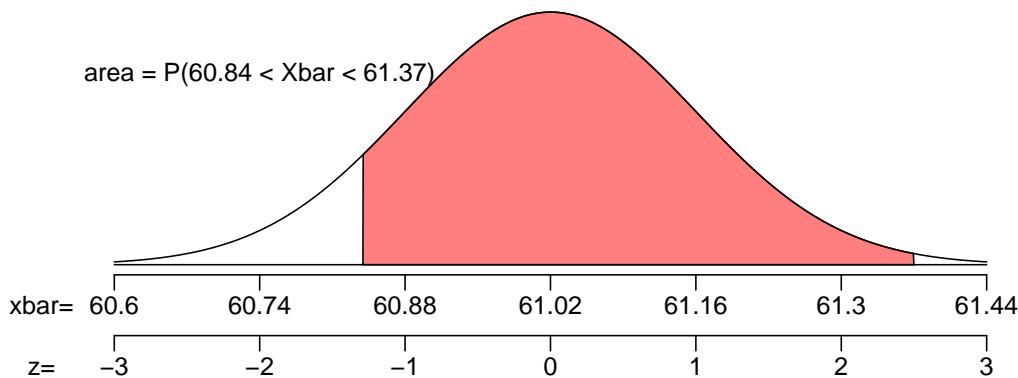
Find the standard error.

$$SE = \frac{\sigma}{\sqrt{n}} = \frac{1.96}{\sqrt{196}} = 0.14$$

Describe the sampling distribution.

$$\bar{X} \sim \mathcal{N}(61.02, 0.14)$$

Draw a sketch.



Calculate a  $z$  scores.

$$z_{\text{lower}} = \frac{x_{\text{lower}} - \mu}{SE} = \frac{60.84 - 61.02}{0.14} = -1.29$$

$$z_{\text{upper}} = \frac{x_{\text{upper}} - \mu}{SE} = \frac{61.37 - 61.02}{0.14} = 2.5$$

Determine the probability.

$$\begin{aligned} P(60.84 < X < 61.37) &= \Phi(z_{\text{upper}}) - \Phi(z_{\text{lower}}) \\ &= \Phi(2.5) - \Phi(-1.29) \\ &= 0.8953 \end{aligned}$$

4. In a game, there is a 22% chance to win a round. You will play 199 rounds.
- (a) What is the probability of winning exactly 49 rounds?
  - (b) What is the probability of winning at least 46 but at most 54 rounds?

**Solution**

We use the formula for binomial probabilities.

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

$$P(X = 49) = \binom{199}{49} (0.22)^{49} (1 - 0.22)^{199-49}$$

$$P(X = 49) = \binom{199}{49} (0.22)^{49} (0.78)^{150}$$

$$P(X = 49) = 0.0444$$

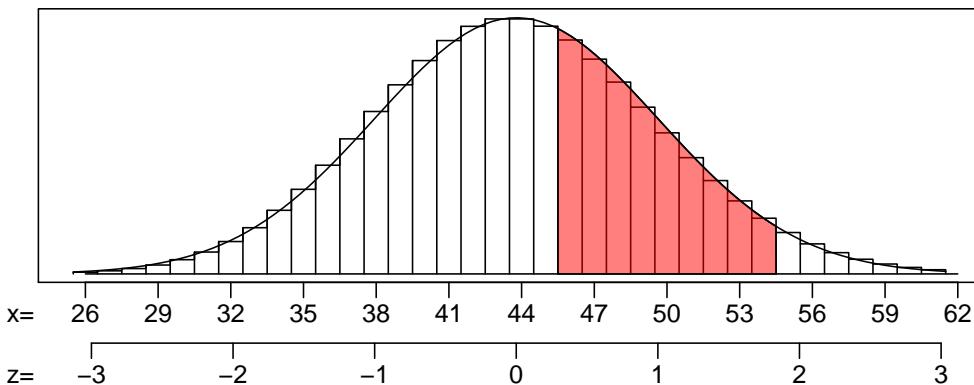
Find the mean.

$$\mu = np = (199)(0.22) = 43.78$$

Find the standard deviation.

$$\sigma = \sqrt{np(1 - p)} = \sqrt{(199)(0.22)(1 - 0.22)} = 5.8437$$

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



Find the z scores.

$$z_1 = \frac{45.5 - 43.78}{5.8437} = 0.38$$

$$z_2 = \frac{54.5 - 43.78}{5.8437} = 1.75$$

Calculate the probability.

$$P(46 \leq X \leq 54) = \Phi(1.75) - \Phi(0.38) = 0.3119$$

(a)  $P(X = 49) = 0.0444$

(b)  $P(46 \leq X \leq 54) = 0.3119$

5. As an ornithologist, you wish to determine the average body mass of *Dendroica palmarum*. You randomly sample 13 adults of *Dendroica palmarum*, resulting in a sample mean of 10.01 grams and a sample standard deviation of 1.26 grams. Determine a 99% confidence interval of the true population mean.

**Solution**

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

$$n = 13$$

$$\bar{x} = 10.01$$

$$s = 1.26$$

$$CL = 0.99$$

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  $P(|T| < t^*) = 0.99$ .

$$t^* = 3.05$$

Calculate the standard error.

$$SE = \frac{s}{\sqrt{n}} = \frac{1.26}{\sqrt{13}} = 0.349$$

We want to make an inference about the population mean.

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

Determine the bounds.

$$\begin{aligned} CI &= (\bar{x} - t^* SE, \bar{x} + t^* SE) \\ &= (10.01 - 3.05 \times 0.349, 10.01 + 3.05 \times 0.349) \\ &= (8.95, 11.1) \end{aligned}$$

We are 99% confident that the population mean is between 8.95 and 11.1.

6. A treatment group of size 19 has a mean of 1120 and standard deviation of 158. A control group of size 31 has a mean of 999 and standard deviation of 180. If you decided to use a significance level of 0.01, 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 42.

- (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.

$$n_1 = 19$$

$$\bar{x}_1 = 1120$$

$$s_1 = 158$$

$$n_2 = 31$$

$$\bar{x}_2 = 999$$

$$s_2 = 180$$

$$\alpha = 0.01$$

$$df = 42$$

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  $P(|T| > t^*) = 0.01$  by using a  $t$  table.

$$t^* = 2.7$$

Calculate the standard error.

$$\begin{aligned} SE &= \sqrt{\frac{(s_1)^2}{n_1} + \frac{(s_2)^2}{n_2}} \\ &= \sqrt{\frac{(158)^2}{19} + \frac{(180)^2}{31}} \\ &= 48.57 \end{aligned}$$

Determine the test statistic.

$$\begin{aligned} t_{\text{obs}} &= \frac{(\bar{x}_2 - \bar{x}_1) - (\mu_2 - \mu_1)_0}{SE} \\ &= \frac{(999 - 1120) - (0)}{48.57} \\ &= -2.49 \end{aligned}$$

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 retain the null hypothesis.

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

- (b)  $H_A : \mu_2 - \mu_1 \neq 0$
- (c)  $t^* = 2.7$
- (d)  $SE = 48.57$
- (e)  $|t_{\text{obs}}| = 2.49$
- (f)  $0.01 < p\text{-value} < 0.02$
- (g) retain the null

7. From a very large population, a random sample of 12000 individuals was taken. In that sample, 44% were glowing. Determine a 98% confidence interval of the population proportion.

- (a) Find the lower bound of the confidence interval.
- (b) Find the upper bound of the condifence interval.

**Solution**

Determine  $z^*$  such that  $P(|Z| < z^*) = 0.98$ .

$$z^* = 2.33$$

Calculate the standard error.

$$SE = \sqrt{\frac{\hat{p}(1 - \hat{p})}{n}} = \sqrt{\frac{(0.44)(1 - 0.44)}{12000}} = 0.00453$$

Calculate the margin of error.

$$ME = z^* SE = (2.33)(0.00453) = 0.0106$$

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.

$$(0.429, 0.451)$$

We are 98% confident that the true population proportion is between 42.9% and 45.1%.

- (a) The lower bound = 0.429, which can also be expressed as 42.9%.
- (b) The upper bound = 0.451, which can also be expressed as 45.1%.

8. An experiment is run with a treatment group of size 185 and a control group of size 220. The results are summarized in the table below.

|          | treatment | control |
|----------|-----------|---------|
| pink     | 52        | 42      |
| not pink | 133       | 178     |

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

- (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.

$$\hat{p}_1 = \frac{52}{185} = 0.281$$

$$\hat{p}_2 = \frac{42}{220} = 0.191$$

Determine the difference of sample proportions.

$$\hat{p}_2 - \hat{p}_1 = 0.191 - 0.281 = -0.09$$

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

$$\hat{p} = \frac{52 + 42}{185 + 220} = 0.232$$

Determine the standard error.

$$\begin{aligned} SE &= \sqrt{\frac{\hat{p}(1 - \hat{p})}{n_1} + \frac{\hat{p}(1 - \hat{p})}{n_2}} \\ &= \sqrt{\frac{(0.232)(0.768)}{185} + \frac{(0.232)(0.768)}{220}} \\ &= 0.0421 \end{aligned}$$

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

$$H_0 : \hat{P}_2 - \hat{P}_1 \sim \mathcal{N}(0, 0.0421)$$

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.

$$\begin{aligned} z &= \frac{(\hat{p}_2 - \hat{p}_1) - (p_2 - p_1)_0}{SE} \\ &= \frac{(0.191 - 0.281) - 0}{0.0421} \\ &= -2.14 \end{aligned}$$

Determine the  $p$ -value.

$$\begin{aligned} p\text{-value} &= 2 \cdot \Phi(-|z|) \\ &= 2 \cdot \Phi(-2.14) \\ &= 0.0324 \end{aligned}$$

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)  $SE = 0.0421$
- (e)  $|Z_{\text{obs}}| = 2.14$
- (f)  $p\text{-value} = 0.0324$
- (g) reject the null

# Bunker Hill Community College

Final Statistics Exam 2019-05-02

Exam ID 017

**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)  $P(\text{dog given white}) = 0.211$

(b)  $P(\text{pig and violet}) = 0.0423$

(c)  $P(\text{violet given wheel}) = 0.208$

(d)  $P(\text{white}) = 0.341$

(e)  $P(\text{dog or white}) = 0.528$

(f)  $P(\text{wheel}) = 0.289$

2.  $P(\text{"llama" given "not white"}) = 0.0695$

3.  $P(62.52 < X < 63.24) = 0.8035$

4. (a)  $P(X = 18) = 0.094$

(b)  $P(14 \leq X \leq 27) = 0.897$

5.  $(9.58, 10.8)$

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

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

(c)  $t^* = 1.69$

(d)  $SE = 2.667$

(e)  $|t_{\text{obs}}| = 1.95$

(f)  $0.05 < p\text{-value} < 0.1$

(g) **reject**

7. (a) **LB of p CI = 0.449 or 44.9%**

(b) **UB of p CI = 0.535 or 53.5%**

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

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

(c)  $Z^* = 2.58$

(d)  $SE = 0.063$

(e)  $|Z_{\text{obs}}| = 2.71$

(f)  $p\text{-value} = 0.0068$

(g) **reject**

1. In a deck of strange cards, there are 781 cards. Each card has an image and a color. The amounts are shown in the table below.

|       | indigo | red | violet | white |
|-------|--------|-----|--------|-------|
| bike  | 21     | 54  | 17     | 38    |
| dog   | 55     | 34  | 57     | 56    |
| pig   | 72     | 31  | 33     | 87    |
| wheel | 65     | 29  | 47     | 85    |

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

**Solution**

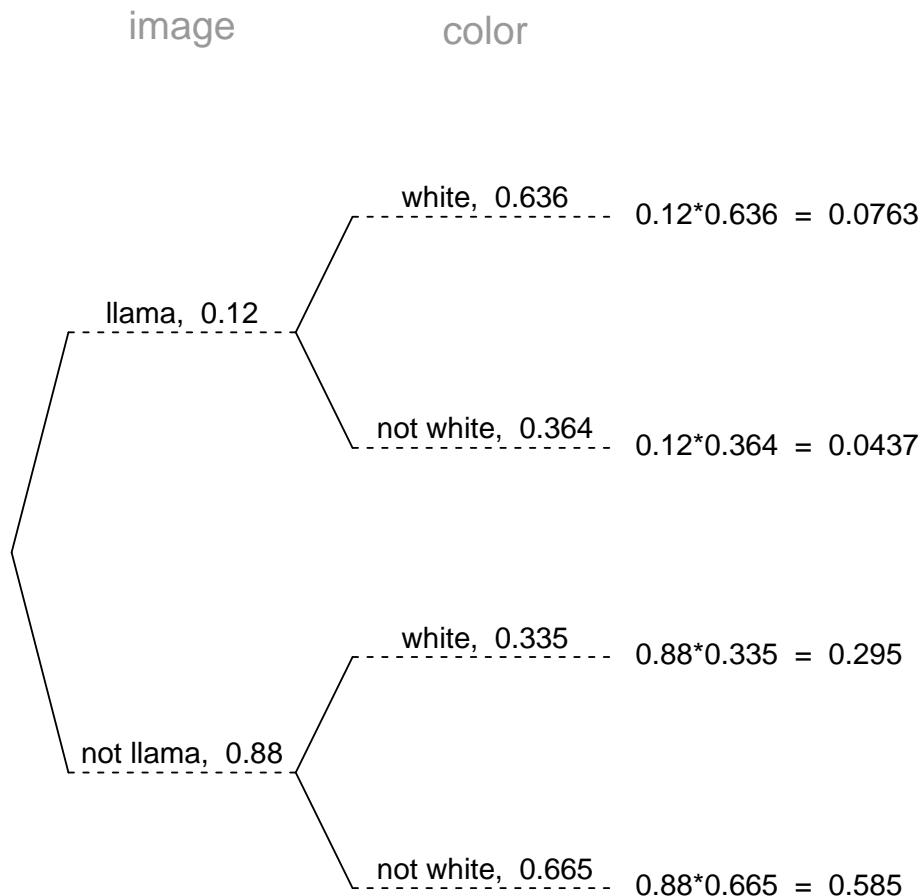
- (a)  $P(\text{dog given white}) = \frac{56}{38+56+87+85} = 0.211$
- (b)  $P(\text{pig and violet}) = \frac{33}{781} = 0.0423$
- (c)  $P(\text{violet given wheel}) = \frac{47}{65+29+47+85} = 0.208$
- (d)  $P(\text{white}) = \frac{38+56+87+85}{781} = 0.341$
- (e)  $P(\text{dog or white}) = \frac{55+34+57+56+38+56+87+85-56}{781} = 0.528$
- (f)  $P(\text{wheel}) = \frac{65+29+47+85}{781} = 0.289$

2. In a deck of strange cards, each card has an image and a color. The chance of drawing a llama is 12%. If a llama is drawn, there is a 63.6% chance that it is white. If a card that is not a llama is drawn, there is a 33.5% chance that it is white.

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

**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.

$$P(\text{"llama" given "not white"}) = \frac{0.0437}{0.0437 + 0.585} = 0.0695$$

3. In a very large pile of toothpicks, the mean length is 62.84 millimeters and the standard deviation is 3.3 millimeters. If you randomly sample 144 toothpicks, what is the chance the sample mean is between 62.52 and 63.24 millimeters?

**Solution**

Label the given information.

$$\mu = 62.84$$

$$\sigma = 3.3$$

$$n = 144$$

$$\bar{x}_{\text{lower}} = 62.52$$

$$\bar{x}_{\text{upper}} = 63.24$$

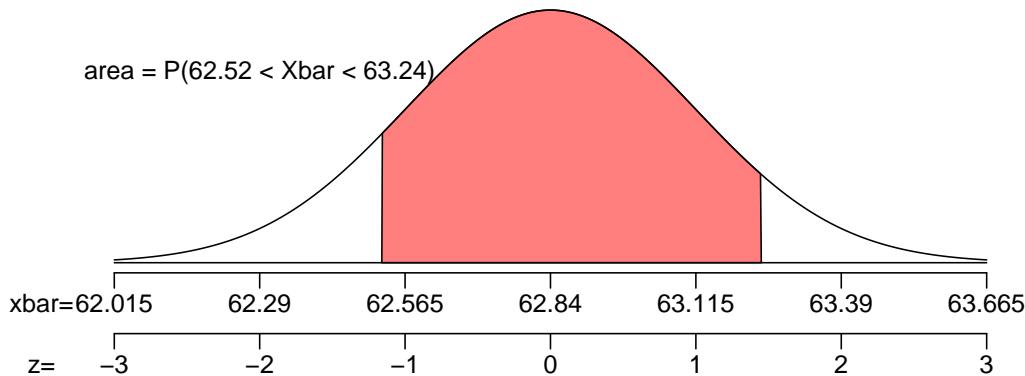
Find the standard error.

$$SE = \frac{\sigma}{\sqrt{n}} = \frac{3.3}{\sqrt{144}} = 0.275$$

Describe the sampling distribution.

$$\bar{X} \sim \mathcal{N}(62.84, 0.275)$$

Draw a sketch.



Calculate a  $z$  scores.

$$z_{\text{lower}} = \frac{x_{\text{lower}} - \mu}{SE} = \frac{62.52 - 62.84}{0.275} = -1.16$$

$$z_{\text{upper}} = \frac{x_{\text{upper}} - \mu}{SE} = \frac{63.24 - 62.84}{0.275} = 1.45$$

Determine the probability.

$$\begin{aligned} P(62.52 < X < 63.24) &= \Phi(z_{\text{upper}}) - \Phi(z_{\text{lower}}) \\ &= \Phi(1.45) - \Phi(-1.16) \\ &= 0.8035 \end{aligned}$$

4. In a game, there is a 22% chance to win a round. You will play 90 rounds.
- (a) What is the probability of winning exactly 18 rounds?
  - (b) What is the probability of winning at least 14 but at most 27 rounds?

**Solution**

We use the formula for binomial probabilities.

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

$$P(X = 18) = \binom{90}{18} (0.22)^{18} (1 - 0.22)^{90-18}$$

$$P(X = 18) = \binom{90}{18} (0.22)^{18} (0.78)^{72}$$

$$P(X = 18) = 0.094$$

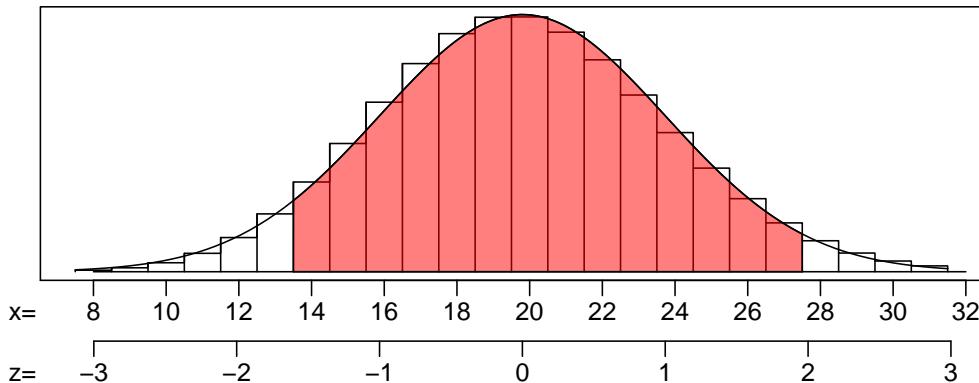
Find the mean.

$$\mu = np = (90)(0.22) = 19.8$$

Find the standard deviation.

$$\sigma = \sqrt{np(1 - p)} = \sqrt{(90)(0.22)(1 - 0.22)} = 3.9299$$

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



Find the z scores.

$$z_1 = \frac{13.5 - 19.8}{3.9299} = -1.48$$

$$z_2 = \frac{27.5 - 19.8}{3.9299} = 1.83$$

Calculate the probability.

$$P(14 \leq X \leq 27) = \Phi(1.83) - \Phi(-1.48) = 0.897$$

(a)  $P(X = 18) = 0.094$

(b)  $P(14 \leq X \leq 27) = 0.897$

5. As an ornithologist, you wish to determine the average body mass of *Vireo griseus*. You randomly sample 16 adults of *Vireo griseus*, resulting in a sample mean of 10.18 grams and a sample standard deviation of 0.729 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.

$$n = 16$$

$$\bar{x} = 10.18$$

$$s = 0.729$$

$$CL = 0.995$$

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

$$df = n - 1 = 15$$

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

$$t^* = 3.29$$

Calculate the standard error.

$$SE = \frac{s}{\sqrt{n}} = \frac{0.729}{\sqrt{16}} = 0.182$$

We want to make an inference about the population mean.

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

Determine the bounds.

$$\begin{aligned} CI &= (\bar{x} - t^* SE, \bar{x} + t^* SE) \\ &= (10.18 - 3.29 \times 0.182, 10.18 + 3.29 \times 0.182) \\ &= (9.58, 10.8) \end{aligned}$$

We are 99.5% confident that the population mean is between 9.58 and 10.8.

6. A treatment group of size 16 has a mean of 99.8 and standard deviation of 7.01. A control group of size 21 has a mean of 105 and standard deviation of 9.21. If you decided to use a significance level of 0.1, 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 34.

- (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.

$$n_1 = 16$$

$$\bar{x}_1 = 99.8$$

$$s_1 = 7.01$$

$$n_2 = 21$$

$$\bar{x}_2 = 105$$

$$s_2 = 9.21$$

$$\alpha = 0.1$$

$$df = 34$$

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  $P(|T| > t^*) = 0.1$  by using a  $t$  table.

$$t^* = 1.69$$

Calculate the standard error.

$$\begin{aligned} SE &= \sqrt{\frac{(s_1)^2}{n_1} + \frac{(s_2)^2}{n_2}} \\ &= \sqrt{\frac{(7.01)^2}{16} + \frac{(9.21)^2}{21}} \\ &= 2.667 \end{aligned}$$

Determine the test statistic.

$$\begin{aligned} t_{\text{obs}} &= \frac{(\bar{x}_2 - \bar{x}_1) - (\mu_2 - \mu_1)_0}{SE} \\ &= \frac{(105 - 99.8) - (0)}{2.667} \\ &= 1.95 \end{aligned}$$

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

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

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

$$0.05 < p\text{-value} < 0.1$$

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)  $t^* = 1.69$
- (d)  $SE = 2.667$
- (e)  $|t_{\text{obs}}| = 1.95$
- (f)  $0.05 < p\text{-value} < 0.1$
- (g) reject the null

7. From a very large population, a random sample of 360 individuals was taken. In that sample, 49.2% were sweet. Determine a 90% confidence interval of the population proportion.
- (a) Find the lower bound of the confidence interval.
  - (b) Find the upper bound of the condifence interval.

**Solution**

Determine  $z^*$  such that  $P(|Z| < z^*) = 0.9$ .

$$z^* = 1.64$$

Calculate the standard error.

$$SE = \sqrt{\frac{\hat{p}(1 - \hat{p})}{n}} = \sqrt{\frac{(0.492)(1 - 0.492)}{360}} = 0.0263$$

Calculate the margin of error.

$$ME = z^* SE = (1.64)(0.0263) = 0.0431$$

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.

$$(0.449, 0.535)$$

We are 90% confident that the true population proportion is between 44.9% and 53.5%.

- (a) The lower bound = 0.449, which can also be expressed as 44.9%.
- (b) The upper bound = 0.535, which can also be expressed as 53.5%.

8. An experiment is run with a treatment group of size 115 and a control group of size 92. The results are summarized in the table below.

|          | treatment | control |
|----------|-----------|---------|
| pink     | 91        | 57      |
| not pink | 24        | 35      |

Using a significance level of 0.01, determine whether the treatment causes an effect on the proportion of cases that are pink.

- (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  $P(|Z| > z^*) = 0.01$ .

$$z^* = \Phi^{-1} \left( 1 - \frac{\alpha}{2} \right) = 2.58$$

Determine the sample proportions.

$$\hat{p}_1 = \frac{91}{115} = 0.791$$

$$\hat{p}_2 = \frac{57}{92} = 0.62$$

Determine the difference of sample proportions.

$$\hat{p}_2 - \hat{p}_1 = 0.62 - 0.791 = -0.171$$

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

$$\hat{p} = \frac{91 + 57}{115 + 92} = 0.715$$

Determine the standard error.

$$\begin{aligned} SE &= \sqrt{\frac{\hat{p}(1 - \hat{p})}{n_1} + \frac{\hat{p}(1 - \hat{p})}{n_2}} \\ &= \sqrt{\frac{(0.715)(0.285)}{115} + \frac{(0.715)(0.285)}{92}} \\ &= 0.0631 \end{aligned}$$

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

$$H_0 : \hat{P}_2 - \hat{P}_1 \sim \mathcal{N}(0, 0.0631)$$

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.

$$\begin{aligned} z &= \frac{(\hat{p}_2 - \hat{p}_1) - (p_2 - p_1)_0}{SE} \\ &= \frac{(0.62 - 0.791) - 0}{0.0631} \\ &= -2.71 \end{aligned}$$

Determine the  $p$ -value.

$$\begin{aligned} p\text{-value} &= 2 \cdot \Phi(-|z|) \\ &= 2 \cdot \Phi(-2.71) \\ &= 0.0068 \end{aligned}$$

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)  $z^* = 2.58$
- (d)  $SE = 0.0631$
- (e)  $|z_{\text{obs}}| = 2.71$
- (f)  $p\text{-value} = 0.0068$
- (g) reject the null

# 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)  $P(\text{tree given white}) = 0.129$

(b)  $P(\text{white}) = 0.26$

(c)  $P(\text{white given tree}) = 0.157$

(d)  $P(\text{shovel or teal}) = 0.487$

(e)  $P(\text{tree}) = 0.213$

(f)  $P(\text{tree and white}) = 0.0334$

2.  $P(\text{"not cat" given "teal"}) = 0.73$

3.  $P(74 < X < 74.7) = 0.7567$

4. (a)  $P(X = 28) = 0.0905$

(b)  $P(21 \leq X \leq 32) = 0.7656$

5. **(11.7, 12.3)**

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

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

(c)  $t^* = 2.62$

(d)  $SE = 8.061$

(e)  $|t_{\text{obs}}| = 2.79$

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

(g) **reject**

7. (a) **LB of p CI = 0.917 or 91.7%**

(b) **UB of p CI = 0.921 or 92.1%**

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

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

(c)  $Z^* = 2.33$

(d)  $SE = 0.097$

(e)  $|Z_{\text{obs}}| = 2.46$

(f)  $p\text{-value} = 0.0138$

(g) **reject**

1. In a deck of strange cards, there are 809 cards. Each card has an image and a color. The amounts are shown in the table below.

|        | green | teal | violet | white |
|--------|-------|------|--------|-------|
| bike   | 32    | 48   | 62     | 29    |
| cat    | 76    | 59   | 58     | 83    |
| shovel | 15    | 18   | 86     | 71    |
| tree   | 28    | 97   | 20     | 27    |

- (a) What is the probability a random card is a tree given it is white?
- (b) What is the probability a random card is white?
- (c) What is the probability a random card is white given it is a tree?
- (d) What is the probability a random card is either a shovel or teal (or both)?
- (e) What is the probability a random card is a tree?
- (f) What is the probability a random card is both a tree and white?

**Solution**

(a)  $P(\text{tree given white}) = \frac{27}{29+83+71+27} = 0.129$

(b)  $P(\text{white}) = \frac{29+83+71+27}{809} = 0.26$

(c)  $P(\text{white given tree}) = \frac{27}{28+97+20+27} = 0.157$

(d)  $P(\text{shovel or teal}) = \frac{15+18+86+71+48+59+18+97-18}{809} = 0.487$

(e)  $P(\text{tree}) = \frac{28+97+20+27}{809} = 0.213$

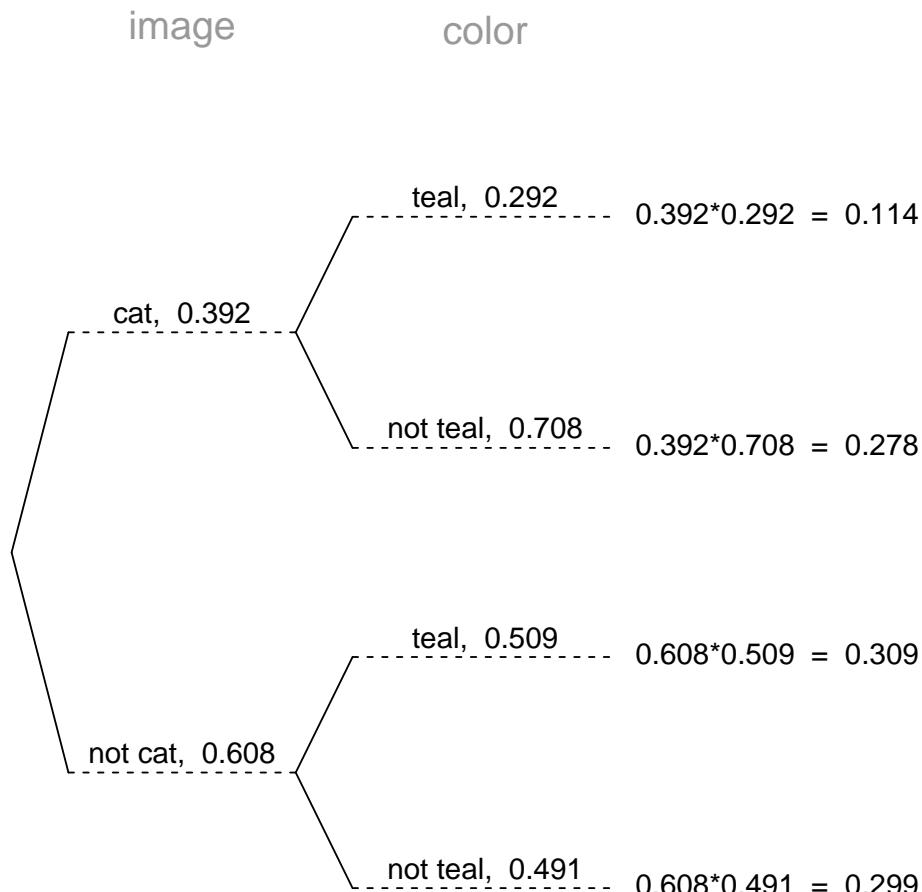
(f)  $P(\text{tree and white}) = \frac{27}{809} = 0.0334$

2. In a deck of strange cards, each card has an image and a color. The chance of drawing a cat is 39.2%. If a cat is drawn, there is a 29.2% chance that it is teal. If a card that is not a cat is drawn, there is a 50.9% chance that it is teal.

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

**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.

$$P(\text{"not cat" given "teal"}) = \frac{0.309}{0.309 + 0.114} = 0.73$$

3. In a very large pile of toothpicks, the mean length is 74.2 millimeters and the standard deviation is 3.33 millimeters. If you randomly sample 169 toothpicks, what is the chance the sample mean is between 74 and 74.7 millimeters?

**Solution**

Label the given information.

$$\mu = 74.2$$

$$\sigma = 3.33$$

$$n = 169$$

$$\bar{x}_{\text{lower}} = 74$$

$$\bar{x}_{\text{upper}} = 74.7$$

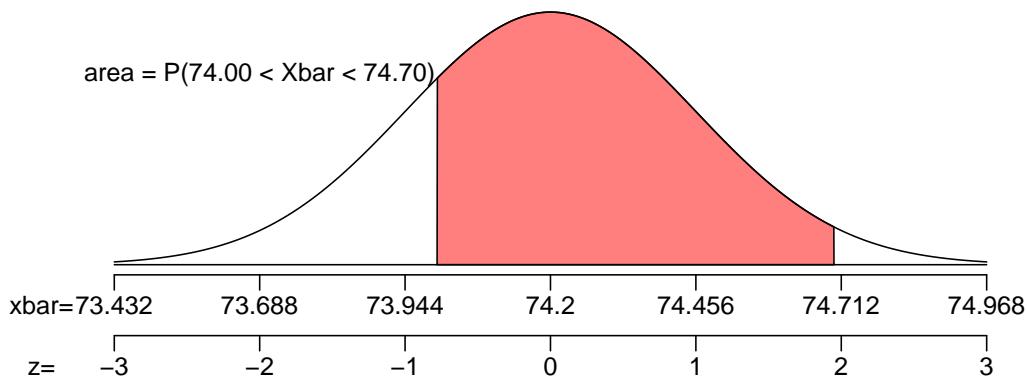
Find the standard error.

$$SE = \frac{\sigma}{\sqrt{n}} = \frac{3.33}{\sqrt{169}} = 0.256$$

Describe the sampling distribution.

$$\bar{X} \sim \mathcal{N}(74.2, 0.256)$$

Draw a sketch.



Calculate a  $z$  scores.

$$z_{\text{lower}} = \frac{\bar{x}_{\text{lower}} - \mu}{SE} = \frac{74 - 74.2}{0.256} = -0.78$$

$$z_{\text{upper}} = \frac{\bar{x}_{\text{upper}} - \mu}{SE} = \frac{74.7 - 74.2}{0.256} = 1.95$$

Determine the probability.

$$\begin{aligned} P(74 < X < 74.7) &= \Phi(z_{\text{upper}}) - \Phi(z_{\text{lower}}) \\ &= \Phi(1.95) - \Phi(-0.78) \\ &= 0.7567 \end{aligned}$$

4. In a game, there is a 31% chance to win a round. You will play 90 rounds.
- (a) What is the probability of winning exactly 28 rounds?
  - (b) What is the probability of winning at least 21 but at most 32 rounds?

**Solution**

We use the formula for binomial probabilities.

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

$$P(X = 28) = \binom{90}{28} (0.31)^{28} (1 - 0.31)^{90-28}$$

$$P(X = 28) = \binom{90}{28} (0.31)^{28} (0.69)^{62}$$

$$P(X = 28) = 0.0905$$

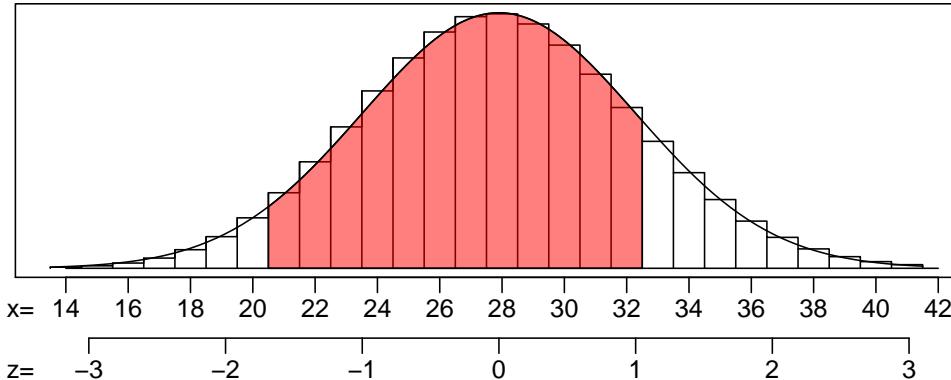
Find the mean.

$$\mu = np = (90)(0.31) = 27.9$$

Find the standard deviation.

$$\sigma = \sqrt{np(1 - p)} = \sqrt{(90)(0.31)(1 - 0.31)} = 4.3876$$

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



Find the z scores.

$$z_1 = \frac{20.5 - 27.9}{4.3876} = -1.57$$

$$z_2 = \frac{32.5 - 27.9}{4.3876} = 0.93$$

Calculate the probability.

$$P(21 \leq X \leq 32) = \Phi(0.93) - \Phi(-1.57) = 0.7656$$

(a)  $P(X = 28) = 0.0905$

(b)  $P(21 \leq X \leq 32) = 0.7656$

5. As an ornithologist, you wish to determine the average body mass of *Dendroica coronata*. You randomly sample 34 adults of *Dendroica coronata*, resulting in a sample mean of 11.97 grams and a sample standard deviation of 1.28 grams. Determine a 80% confidence interval of the true population mean.

**Solution**

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

$$n = 34$$

$$\bar{x} = 11.97$$

$$s = 1.28$$

$$CL = 0.8$$

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

$$df = n - 1 = 33$$

Determine the critical  $t$  value,  $t^*$ , such that  $P(|T| < t^*) = 0.8$ .

$$t^* = 1.31$$

Calculate the standard error.

$$SE = \frac{s}{\sqrt{n}} = \frac{1.28}{\sqrt{34}} = 0.22$$

We want to make an inference about the population mean.

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

Determine the bounds.

$$\begin{aligned} CI &= (\bar{x} - t^* SE, \bar{x} + t^* SE) \\ &= (11.97 - 1.31 \times 0.22, 11.97 + 1.31 \times 0.22) \\ &= (11.7, 12.3) \end{aligned}$$

We are 80% confident that the population mean is between 11.7 and 12.3.

6. A treatment group of size 10 has a mean of 93.5 and standard deviation of 22.3. A control group of size 21 has a mean of 116 and standard deviation of 17.9. 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 14.

- (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.

$$n_1 = 10$$

$$\bar{x}_1 = 93.5$$

$$s_1 = 22.3$$

$$n_2 = 21$$

$$\bar{x}_2 = 116$$

$$s_2 = 17.9$$

$$\alpha = 0.02$$

$$df = 14$$

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  $P(|T| > t^*) = 0.02$  by using a  $t$  table.

$$t^* = 2.62$$

Calculate the standard error.

$$\begin{aligned} SE &= \sqrt{\frac{(s_1)^2}{n_1} + \frac{(s_2)^2}{n_2}} \\ &= \sqrt{\frac{(22.3)^2}{10} + \frac{(17.9)^2}{21}} \\ &= 8.061 \end{aligned}$$

Determine the test statistic.

$$\begin{aligned} t_{\text{obs}} &= \frac{(\bar{x}_2 - \bar{x}_1) - (\mu_2 - \mu_1)_0}{SE} \\ &= \frac{(116 - 93.5) - (0)}{8.061} \\ &= 2.79 \end{aligned}$$

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)  $t^* = 2.62$
- (d)  $SE = 8.061$
- (e)  $|t_{\text{obs}}| = 2.79$
- (f)  $0.01 < p\text{-value} < 0.02$
- (g) reject the null

7. From a very large population, a random sample of 90000 individuals was taken. In that sample, 91.9% were messy. 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 condifence interval.

**Solution**

Determine  $z^*$  such that  $P(|Z| < z^*) = 0.95$ .

$$z^* = 1.96$$

Calculate the standard error.

$$SE = \sqrt{\frac{\hat{p}(1 - \hat{p})}{n}} = \sqrt{\frac{(0.919)(1 - 0.919)}{90000}} = 0.000909$$

Calculate the margin of error.

$$ME = z^* SE = (1.96)(0.000909) = 0.00178$$

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.

$$(0.917, 0.921)$$

We are 95% confident that the true population proportion is between 91.7% and 92.1%.

- (a) The lower bound = 0.917, which can also be expressed as 91.7%.
- (b) The upper bound = 0.921, which can also be expressed as 92.1%.

8. An experiment is run with a treatment group of size 62 and a control group of size 46. The results are summarized in the table below.

|           | treatment | control |
|-----------|-----------|---------|
| happy     | 27        | 31      |
| not happy | 35        | 15      |

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

- (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.

$$\hat{p}_1 = \frac{27}{62} = 0.435$$

$$\hat{p}_2 = \frac{31}{46} = 0.674$$

Determine the difference of sample proportions.

$$\hat{p}_2 - \hat{p}_1 = 0.674 - 0.435 = 0.239$$

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

$$\hat{p} = \frac{27 + 31}{62 + 46} = 0.537$$

Determine the standard error.

$$\begin{aligned} SE &= \sqrt{\frac{\hat{p}(1 - \hat{p})}{n_1} + \frac{\hat{p}(1 - \hat{p})}{n_2}} \\ &= \sqrt{\frac{(0.537)(0.463)}{62} + \frac{(0.537)(0.463)}{46}} \\ &= 0.097 \end{aligned}$$

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

$$H_0 : \hat{P}_2 - \hat{P}_1 \sim \mathcal{N}(0, 0.097)$$

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.

$$\begin{aligned} z &= \frac{(\hat{p}_2 - \hat{p}_1) - (p_2 - p_1)_0}{SE} \\ &= \frac{(0.674 - 0.435) - 0}{0.097} \\ &= 2.46 \end{aligned}$$

Determine the  $p$ -value.

$$\begin{aligned} p\text{-value} &= 2 \cdot \Phi(-|z|) \\ &= 2 \cdot \Phi(-2.46) \\ &= 0.0138 \end{aligned}$$

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)  $SE = 0.097$
- (e)  $|Z_{\text{obs}}| = 2.46$
- (f)  $p\text{-value} = 0.0138$
- (g) reject the null

# Bunker Hill Community College

Final Statistics Exam 2019-05-02

Exam ID 019

**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)  $P(\text{cat or white}) = 0.456$

(b)  $P(\text{white}) = 0.418$

(c)  $P(\text{tree given white}) = 0.168$

(d)  $P(\text{red given dog}) = 0.216$

(e)  $P(\text{dog and white}) = 0.105$

(f)  $P(\text{bike}) = 0.272$

2.  $P(\text{"ring" given "violet"}) = 0.706$

3.  $P(67.06 < X < 67.88) = 0.8396$

4. (a)  $P(X = 37) = 0.0718$

(b)  $P(37 \leq X \leq 50) = 0.4721$

5.  $(11.2, 12.8)$

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

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

(c)  $t^* = 1.67$

(d)  $SE = 0.526$

(e)  $|t_{\text{obs}}| = 1.79$

(f)  $0.05 < p\text{-value} < 0.1$

(g) **reject**

7. (a) **LB of p CI = 0.0795 or 7.95%**

(b) **UB of p CI = 0.0927 or 9.27%**

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

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

(c)  $Z^* = 2.81$

(d)  $SE = 0.115$

(e)  $|Z_{\text{obs}}| = 2.74$

(f)  $p\text{-value} = 0.0062$

(g) **retain**

1. In a deck of strange cards, there are 895 cards. Each card has an image and a color. The amounts are shown in the table below.

|        | gray | red | white |
|--------|------|-----|-------|
| bike   | 87   | 71  | 85    |
| cat    | 22   | 12  | 56    |
| dog    | 58   | 42  | 94    |
| flower | 68   | 70  | 76    |
| tree   | 81   | 10  | 63    |

- (a) What is the probability a random card is either a cat or white (or both)?
- (b) What is the probability a random card is white?
- (c) What is the probability a random card is a tree given it is white?
- (d) What is the probability a random card is red given it is a dog?
- (e) What is the probability a random card is both a dog and white?
- (f) What is the probability a random card is a bike?

**Solution**

$$(a) P(\text{cat or white}) = \frac{22+12+56+85+56+94+76+63-56}{895} = 0.456$$

$$(b) P(\text{white}) = \frac{85+56+94+76+63}{895} = 0.418$$

$$(c) P(\text{tree given white}) = \frac{63}{85+56+94+76+63} = 0.168$$

$$(d) P(\text{red given dog}) = \frac{42}{58+42+94} = 0.216$$

$$(e) P(\text{dog and white}) = \frac{94}{895} = 0.105$$

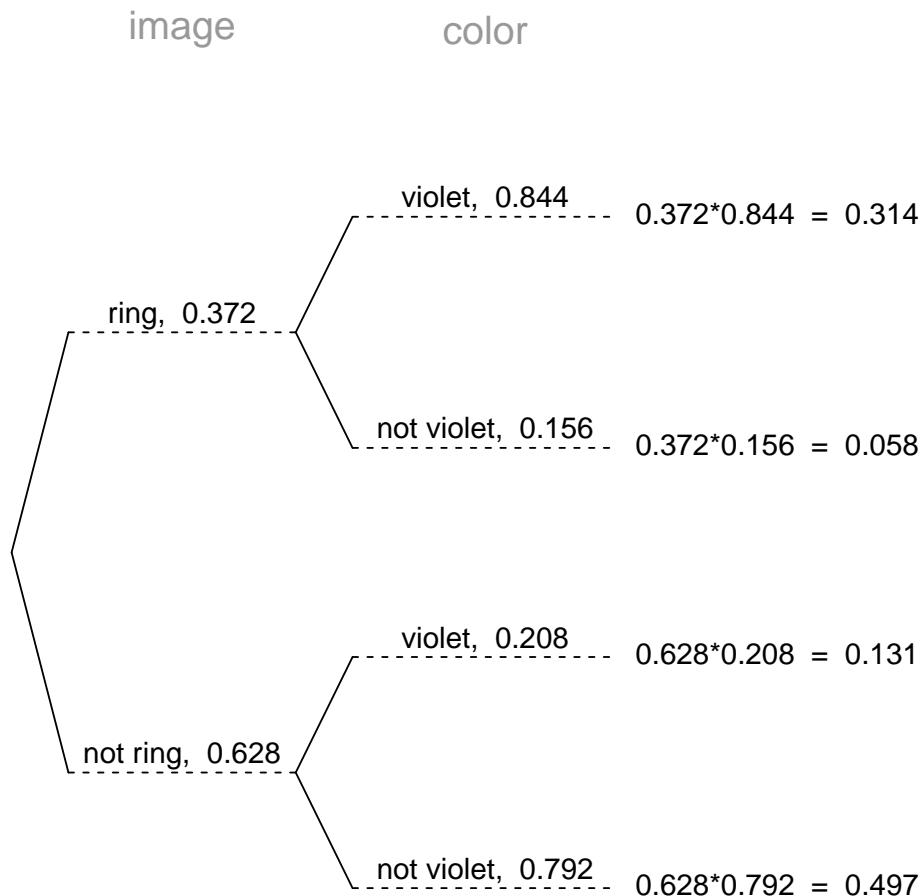
$$(f) P(\text{bike}) = \frac{87+71+85}{895} = 0.272$$

2. In a deck of strange cards, each card has an image and a color. The chance of drawing a ring is 37.2%. If a ring is drawn, there is a 84.4% chance that it is violet. If a card that is not a ring is drawn, there is a 20.8% chance that it is violet.

Now, someone draws a random card and reveals it is violet. What is the chance the card is 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.

$$P(\text{"ring" given "violet"}) = \frac{0.314}{0.314 + 0.131} = 0.706$$

3. In a very large pile of toothpicks, the mean length is 67.52 millimeters and the standard deviation is 3.17 millimeters. If you randomly sample 121 toothpicks, what is the chance the sample mean is between 67.06 and 67.88 millimeters?

**Solution**

Label the given information.

$$\mu = 67.52$$

$$\sigma = 3.17$$

$$n = 121$$

$$\bar{x}_{\text{lower}} = 67.06$$

$$\bar{x}_{\text{upper}} = 67.88$$

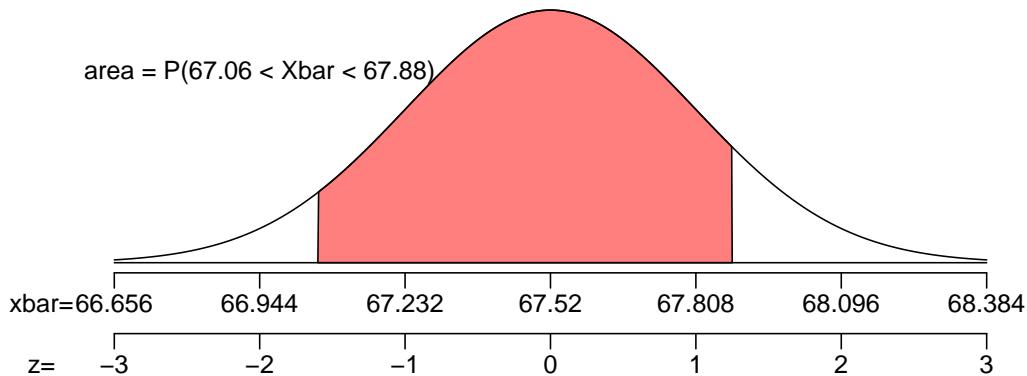
Find the standard error.

$$SE = \frac{\sigma}{\sqrt{n}} = \frac{3.17}{\sqrt{121}} = 0.288$$

Describe the sampling distribution.

$$\bar{X} \sim \mathcal{N}(67.52, 0.288)$$

Draw a sketch.



Calculate a  $z$  scores.

$$z_{\text{lower}} = \frac{\bar{x}_{\text{lower}} - \mu}{SE} = \frac{67.06 - 67.52}{0.288} = -1.6$$

$$z_{\text{upper}} = \frac{\bar{x}_{\text{upper}} - \mu}{SE} = \frac{67.88 - 67.52}{0.288} = 1.25$$

Determine the probability.

$$\begin{aligned} P(67.06 < \bar{X} < 67.88) &= \Phi(z_{\text{upper}}) - \Phi(z_{\text{lower}}) \\ &= \Phi(1.25) - \Phi(-1.6) \\ &= 0.8396 \end{aligned}$$

4. In a game, there is a 17% chance to win a round. You will play 216 rounds.
- (a) What is the probability of winning exactly 37 rounds?
  - (b) What is the probability of winning at least 37 but at most 50 rounds?

**Solution**

We use the formula for binomial probabilities.

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

$$P(X = 37) = \binom{216}{37} (0.17)^{37} (1 - 0.17)^{216-37}$$

$$P(X = 37) = \binom{216}{37} (0.17)^{37} (0.83)^{179}$$

$$P(X = 37) = 0.0718$$

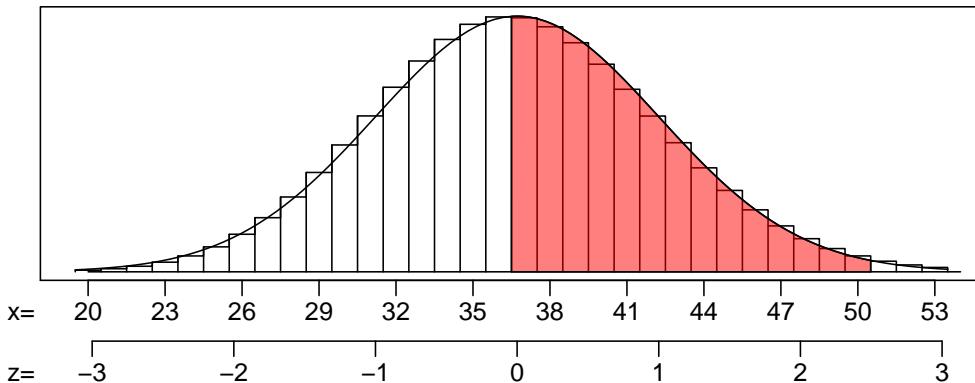
Find the mean.

$$\mu = np = (216)(0.17) = 36.72$$

Find the standard deviation.

$$\sigma = \sqrt{np(1 - p)} = \sqrt{(216)(0.17)(1 - 0.17)} = 5.5207$$

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



Find the z scores.

$$z_1 = \frac{36.5 - 36.72}{5.5207} = 0.05$$

$$z_2 = \frac{50.5 - 36.72}{5.5207} = 2.41$$

Calculate the probability.

$$P(37 \leq X \leq 50) = \Phi(2.41) - \Phi(0.05) = 0.4721$$

(a)  $P(X = 37) = 0.0718$

(b)  $P(37 \leq X \leq 50) = 0.4721$

5. As an ornithologist, you wish to determine the average body mass of *Dendroica coronata*. You randomly sample 24 adults of *Dendroica coronata*, resulting in a sample mean of 11.99 grams and a sample standard deviation of 1.25 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.

$$n = 24$$

$$\bar{x} = 11.99$$

$$s = 1.25$$

$$CL = 0.995$$

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

$$df = n - 1 = 23$$

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

$$t^* = 3.1$$

Calculate the standard error.

$$SE = \frac{s}{\sqrt{n}} = \frac{1.25}{\sqrt{24}} = 0.255$$

We want to make an inference about the population mean.

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

Determine the bounds.

$$\begin{aligned} CI &= (\bar{x} - t^* SE, \bar{x} + t^* SE) \\ &= (11.99 - 3.1 \times 0.255, 11.99 + 3.1 \times 0.255) \\ &= (11.2, 12.8) \end{aligned}$$

We are 99.5% confident that the population mean is between 11.2 and 12.8.

6. A treatment group of size 37 has a mean of 9.96 and standard deviation of 2.22. A control group of size 26 has a mean of 10.9 and standard deviation of 1.93. If you decided to use a significance level of 0.1, 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 58.

- (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.

$$n_1 = 37$$

$$\bar{x}_1 = 9.96$$

$$s_1 = 2.22$$

$$n_2 = 26$$

$$\bar{x}_2 = 10.9$$

$$s_2 = 1.93$$

$$\alpha = 0.1$$

$$df = 58$$

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  $P(|T| > t^*) = 0.1$  by using a  $t$  table.

$$t^* = 1.67$$

Calculate the standard error.

$$\begin{aligned} SE &= \sqrt{\frac{(s_1)^2}{n_1} + \frac{(s_2)^2}{n_2}} \\ &= \sqrt{\frac{(2.22)^2}{37} + \frac{(1.93)^2}{26}} \\ &= 0.526 \end{aligned}$$

Determine the test statistic.

$$\begin{aligned} t_{\text{obs}} &= \frac{(\bar{x}_2 - \bar{x}_1) - (\mu_2 - \mu_1)_0}{SE} \\ &= \frac{(10.9 - 9.96) - (0)}{0.526} \\ &= 1.79 \end{aligned}$$

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

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

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

$$0.05 < p\text{-value} < 0.1$$

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)  $t^* = 1.67$
- (d)  $SE = 0.526$
- (e)  $|t_{\text{obs}}| = 1.79$
- (f)  $0.05 < p\text{-value} < 0.1$
- (g) reject the null

7. From a very large population, a random sample of 7600 individuals was taken. In that sample, 8.61% were glowing. 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 condifence interval.

**Solution**

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

$$z^* = 2.05$$

Calculate the standard error.

$$SE = \sqrt{\frac{\hat{p}(1 - \hat{p})}{n}} = \sqrt{\frac{(0.0861)(1 - 0.0861)}{7600}} = 0.00322$$

Calculate the margin of error.

$$ME = z^* SE = (2.05)(0.00322) = 0.0066$$

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.

$$(0.0795, 0.0927)$$

We are 96% confident that the true population proportion is between 7.95% and 9.27%.

- (a) The lower bound = 0.0795, which can also be expressed as 7.95%.
- (b) The upper bound = 0.0927, which can also be expressed as 9.27%.

8. An experiment is run with a treatment group of size 23 and a control group of size 43. The results are summarized in the table below.

|             | treatment | control |
|-------------|-----------|---------|
| special     | 11        | 7       |
| not special | 12        | 36      |

Using a significance level of 0.005, 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  $P(|Z| > z^*) = 0.005$ .

$$z^* = \Phi^{-1} \left( 1 - \frac{\alpha}{2} \right) = 2.81$$

Determine the sample proportions.

$$\hat{p}_1 = \frac{11}{23} = 0.478$$

$$\hat{p}_2 = \frac{7}{43} = 0.163$$

Determine the difference of sample proportions.

$$\hat{p}_2 - \hat{p}_1 = 0.163 - 0.478 = -0.315$$

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

$$\hat{p} = \frac{11 + 7}{23 + 43} = 0.273$$

Determine the standard error.

$$\begin{aligned} SE &= \sqrt{\frac{\hat{p}(1 - \hat{p})}{n_1} + \frac{\hat{p}(1 - \hat{p})}{n_2}} \\ &= \sqrt{\frac{(0.273)(0.727)}{23} + \frac{(0.273)(0.727)}{43}} \\ &= 0.115 \end{aligned}$$

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

$$H_0 : \hat{P}_2 - \hat{P}_1 \sim \mathcal{N}(0, 0.115)$$

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.

$$\begin{aligned} z &= \frac{(\hat{p}_2 - \hat{p}_1) - (p_2 - p_1)_0}{SE} \\ &= \frac{(0.163 - 0.478) - 0}{0.115} \\ &= -2.74 \end{aligned}$$

Determine the  $p$ -value.

$$\begin{aligned} p\text{-value} &= 2 \cdot \Phi(-|z|) \\ &= 2 \cdot \Phi(-2.74) \\ &= 0.0062 \end{aligned}$$

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

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

So, we retain the null hypothesis. Thus the difference in proportions is not significant.

- (a)  $H_0 : p_2 - p_1 = 0$
- (b)  $H_A : p_2 - p_1 \neq 0$
- (c)  $z^* = 2.81$
- (d)  $SE = 0.115$
- (e)  $|z_{\text{obs}}| = 2.74$
- (f)  $p\text{-value} = 0.0062$
- (g) retain the null

# Bunker Hill Community College

Final Statistics Exam 2019-05-02

Exam ID 020

**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)  $P(\text{teal}) = 0.205$

(b)  $P(\text{horn given red}) = 0.241$

(c)  $P(\text{bike or black}) = 0.363$

(d)  $P(\text{teal given wheel}) = 0.165$

(e)  $P(\text{bike}) = 0.24$

(f)  $P(\text{bike and teal}) = 0.0685$

2.  $P(\text{"gem" given "black"}) = 0.265$

3.  $P(72.47 < X < 72.81) = 0.7557$

4. (a)  $P(X = 20) = 0.1113$

(b)  $P(24 \leq X \leq 29) = 0.2023$

5.  $(9.35, 10.4)$

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

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

(c)  $t^* = 2.11$

(d)  $SE = 0.009$

(e)  $|t_{\text{obs}}| = 2.21$

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

(g) **reject**

7. (a) **LB of p CI = 0.522 or 52.2%**

(b) **UB of p CI = 0.548 or 54.8%**

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

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

(c)  $Z^* = 1.64$

(d)  $SE = 0.042$

(e)  $|Z_{\text{obs}}| = 1.76$

(f)  $p\text{-value} = 0.0784$

(g) **reject**

1. In a deck of strange cards, there are 1271 cards. Each card has an image and a color. The amounts are shown in the table below.

|        | black | green | pink | red | teal |
|--------|-------|-------|------|-----|------|
| bike   | 34    | 85    | 68   | 31  | 87   |
| dog    | 71    | 24    | 47   | 19  | 32   |
| flower | 20    | 69    | 55   | 88  | 35   |
| horn   | 14    | 26    | 65   | 72  | 63   |
| wheel  | 51    | 40    | 42   | 89  | 44   |

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

**Solution**

(a)  $P(\text{teal}) = \frac{87+32+35+63+44}{1271} = 0.205$

(b)  $P(\text{horn given red}) = \frac{72}{31+19+88+72+89} = 0.241$

(c)  $P(\text{bike or black}) = \frac{34+85+68+31+87+34+71+20+14+51-34}{1271} = 0.363$

(d)  $P(\text{teal given wheel}) = \frac{44}{51+40+42+89+44} = 0.165$

(e)  $P(\text{bike}) = \frac{34+85+68+31+87}{1271} = 0.24$

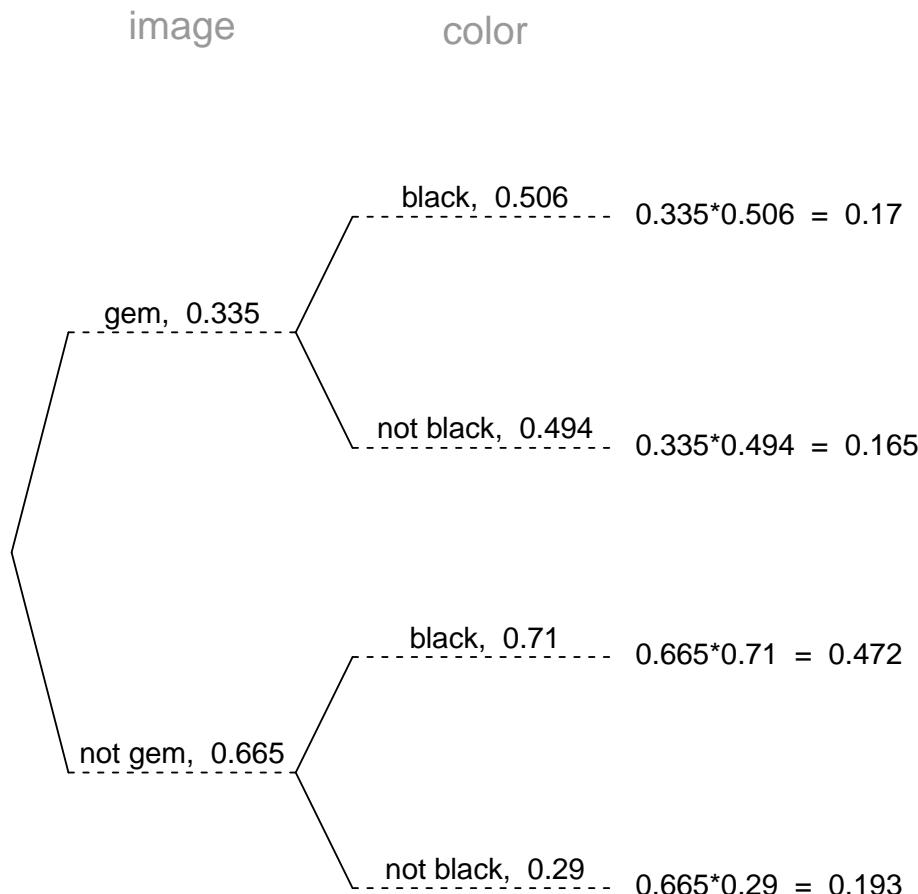
(f)  $P(\text{bike and teal}) = \frac{87}{1271} = 0.0685$

2. In a deck of strange cards, each card has an image and a color. The chance of drawing a gem is 33.5%. If a gem is drawn, there is a 50.6% chance that it is black. If a card that is not a gem is drawn, there is a 71% chance that it is black.

Now, someone draws a random card and reveals it is black. What is the chance the card is a gem?

**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.

$$P(\text{"gem" given "black"}) = \frac{0.17}{0.17 + 0.472} = 0.265$$

3. In a very large pile of toothpicks, the mean length is 72.61 millimeters and the standard deviation is 2.14 millimeters. If you randomly sample 225 toothpicks, what is the chance the sample mean is between 72.47 and 72.81 millimeters?

**Solution**

Label the given information.

$$\mu = 72.61$$

$$\sigma = 2.14$$

$$n = 225$$

$$\bar{x}_{\text{lower}} = 72.47$$

$$\bar{x}_{\text{upper}} = 72.81$$

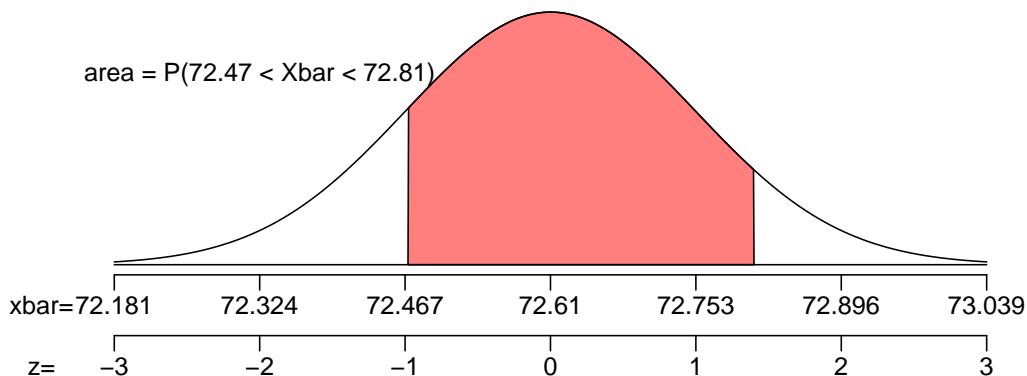
Find the standard error.

$$SE = \frac{\sigma}{\sqrt{n}} = \frac{2.14}{\sqrt{225}} = 0.143$$

Describe the sampling distribution.

$$\bar{X} \sim \mathcal{N}(72.61, 0.143)$$

Draw a sketch.



Calculate a  $z$  scores.

$$z_{\text{lower}} = \frac{x_{\text{lower}} - \mu}{SE} = \frac{72.47 - 72.61}{0.143} = -0.98$$

$$z_{\text{upper}} = \frac{x_{\text{upper}} - \mu}{SE} = \frac{72.81 - 72.61}{0.143} = 1.4$$

Determine the probability.

$$\begin{aligned} P(72.47 < X < 72.81) &= \Phi(z_{\text{upper}}) - \Phi(z_{\text{lower}}) \\ &= \Phi(1.4) - \Phi(-0.98) \\ &= 0.7557 \end{aligned}$$

4. In a game, there is a 51% chance to win a round. You will play 42 rounds.
- (a) What is the probability of winning exactly 20 rounds?
  - (b) What is the probability of winning at least 24 but at most 29 rounds?

**Solution**

We use the formula for binomial probabilities.

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

$$P(X = 20) = \binom{42}{20} (0.51)^{20} (1 - 0.51)^{42-20}$$

$$P(X = 20) = \binom{42}{20} (0.51)^{20} (0.49)^{22}$$

$$P(X = 20) = 0.1113$$

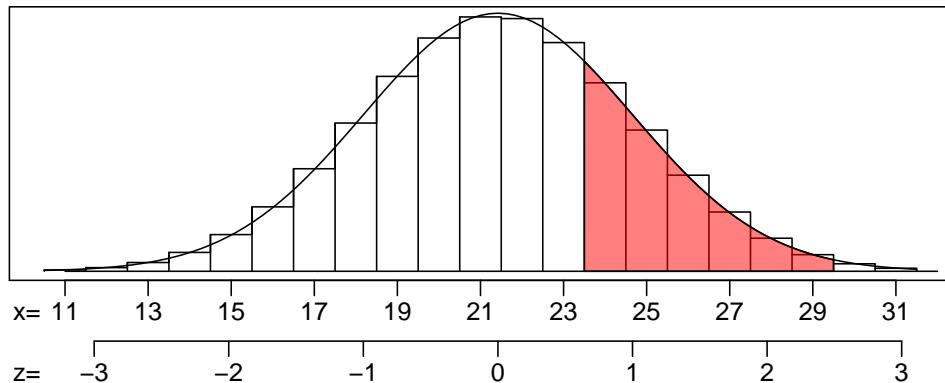
Find the mean.

$$\mu = np = (42)(0.51) = 21.42$$

Find the standard deviation.

$$\sigma = \sqrt{np(1 - p)} = \sqrt{(42)(0.51)(1 - 0.51)} = 3.2397$$

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



Find the z scores.

$$z_1 = \frac{23.5 - 21.42}{3.2397} = 0.8$$

$$z_2 = \frac{29.5 - 21.42}{3.2397} = 2.34$$

Calculate the probability.

$$P(24 \leq X \leq 29) = \Phi(2.34) - \Phi(0.8) = 0.2023$$

(a)  $P(X = 20) = 0.1113$

(b)  $P(24 \leq X \leq 29) = 0.2023$

5. As an ornithologist, you wish to determine the average body mass of *Dendroica dominica*. You randomly sample 27 adults of *Dendroica dominica*, resulting in a sample mean of 9.87 grams and a sample standard deviation of 1.59 grams. Determine a 90% confidence interval of the true population mean.

**Solution**

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

$$n = 27$$

$$\bar{x} = 9.87$$

$$s = 1.59$$

$$CL = 0.9$$

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

$$df = n - 1 = 26$$

Determine the critical  $t$  value,  $t^*$ , such that  $P(|T| < t^*) = 0.9$ .

$$t^* = 1.71$$

Calculate the standard error.

$$SE = \frac{s}{\sqrt{n}} = \frac{1.59}{\sqrt{27}} = 0.306$$

We want to make an inference about the population mean.

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

Determine the bounds.

$$\begin{aligned} CI &= (\bar{x} - t^* SE, \bar{x} + t^* SE) \\ &= (9.87 - 1.71 \times 0.306, 9.87 + 1.71 \times 0.306) \\ &= (9.35, 10.4) \end{aligned}$$

We are 90% confident that the population mean is between 9.35 and 10.4.

6. A treatment group of size 26 has a mean of 1 and standard deviation of 0.0315. A control group of size 29 has a mean of 1.02 and standard deviation of 0.0356. 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 52.

- (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.

$$n_1 = 26$$

$$\bar{x}_1 = 1$$

$$s_1 = 0.0315$$

$$n_2 = 29$$

$$\bar{x}_2 = 1.02$$

$$s_2 = 0.0356$$

$$\alpha = 0.04$$

$$df = 52$$

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.

$$t^* = 2.11$$

Calculate the standard error.

$$\begin{aligned} SE &= \sqrt{\frac{(s_1)^2}{n_1} + \frac{(s_2)^2}{n_2}} \\ &= \sqrt{\frac{(0.0315)^2}{26} + \frac{(0.0356)^2}{29}} \\ &= 0.009 \end{aligned}$$

Determine the test statistic.

$$\begin{aligned} t_{\text{obs}} &= \frac{(\bar{x}_2 - \bar{x}_1) - (\mu_2 - \mu_1)_0}{SE} \\ &= \frac{(1.02 - 1) - (0)}{0.009} \\ &= 2.21 \end{aligned}$$

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)  $t^* = 2.11$
- (d)  $SE = 0.009$
- (e)  $|t_{\text{obs}}| = 2.21$
- (f)  $0.02 < p\text{-value} < 0.04$
- (g) reject the null

7. From a very large population, a random sample of 4000 individuals was taken. In that sample, 53.5% were broken. Determine a 90% confidence interval of the population proportion.

- (a) Find the lower bound of the confidence interval.
- (b) Find the upper bound of the condifence interval.

**Solution**

Determine  $z^*$  such that  $P(|Z| < z^*) = 0.9$ .

$$z^* = 1.64$$

Calculate the standard error.

$$SE = \sqrt{\frac{\hat{p}(1 - \hat{p})}{n}} = \sqrt{\frac{(0.535)(1 - 0.535)}{4000}} = 0.00789$$

Calculate the margin of error.

$$ME = z^* SE = (1.64)(0.00789) = 0.0129$$

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.

$$(0.522, 0.548)$$

We are 90% confident that the true population proportion is between 52.2% and 54.8%.

- (a) The lower bound = 0.522, which can also be expressed as 52.2%.
- (b) The upper bound = 0.548, which can also be expressed as 54.8%.

8. An experiment is run with a treatment group of size 296 and a control group of size 284. The results are summarized in the table below.

|                 | treatment | control |
|-----------------|-----------|---------|
| fluorescent     | 151       | 124     |
| not fluorescent | 145       | 160     |

Using a significance level of 0.1, determine whether the treatment causes an effect on the proportion of cases that are fluorescent.

- (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  $P(|Z| > z^*) = 0.1$ .

$$z^* = \Phi^{-1} \left( 1 - \frac{\alpha}{2} \right) = 1.64$$

Determine the sample proportions.

$$\hat{p}_1 = \frac{151}{296} = 0.51$$

$$\hat{p}_2 = \frac{124}{284} = 0.437$$

Determine the difference of sample proportions.

$$\hat{p}_2 - \hat{p}_1 = 0.437 - 0.51 = -0.073$$

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

$$\hat{p} = \frac{151 + 124}{296 + 284} = 0.474$$

Determine the standard error.

$$\begin{aligned} SE &= \sqrt{\frac{\hat{p}(1 - \hat{p})}{n_1} + \frac{\hat{p}(1 - \hat{p})}{n_2}} \\ &= \sqrt{\frac{(0.474)(0.526)}{296} + \frac{(0.474)(0.526)}{284}} \\ &= 0.0415 \end{aligned}$$

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

$$H_0 : \hat{P}_2 - \hat{P}_1 \sim \mathcal{N}(0, 0.0415)$$

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.

$$\begin{aligned} z &= \frac{(\hat{p}_2 - \hat{p}_1) - (p_2 - p_1)_0}{SE} \\ &= \frac{(0.437 - 0.51) - 0}{0.0415} \\ &= -1.76 \end{aligned}$$

Determine the  $p$ -value.

$$\begin{aligned} p\text{-value} &= 2 \cdot \Phi(-|z|) \\ &= 2 \cdot \Phi(-1.76) \\ &= 0.0784 \end{aligned}$$

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)  $z^* = 1.64$
- (d)  $SE = 0.0415$
- (e)  $|z_{\text{obs}}| = 1.76$
- (f)  $p\text{-value} = 0.0784$
- (g) reject the null

# Bunker Hill Community College

Final Statistics Exam 2019-05-02

Exam ID 021

**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)  $P(\text{yellow given cat}) = 0.439$

(b)  $P(\text{wheel}) = 0.281$

(c)  $P(\text{cat given blue}) = 0.408$

(d)  $P(\text{wheel or yellow}) = 0.603$

(e)  $P(\text{cat and teal}) = 0.209$

(f)  $P(\text{teal}) = 0.495$

2.  $P(\text{"not tree" given "green"}) = 0.802$

3.  $P(63.31 < X < 63.96) = 0.878$

4. (a)  $P(X = 17) = 0.1012$

(b)  $P(9 \leq X \leq 16) = 0.6524$

5. **(52.1, 59.2)**

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

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

(c)  $t^* = 2.01$

(d)  $SE = 0.023$

(e)  $|t_{\text{obs}}| = 1.93$

(f)  $0.05 < p\text{-value} < 0.1$

(g) **retain**

7. (a) **LB of p CI = 0.852 or 85.2%**

(b) **UB of p CI = 0.86 or 86%**

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

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

(c)  $Z^* = 1.96$

(d)  $SE = 0.07$

(e)  $|Z_{\text{obs}}| = 2$

(f)  $p\text{-value} = 0.0456$

(g) **reject**

1. In a deck of strange cards, there are 459 cards. Each card has an image and a color. The amounts are shown in the table below.

|       | blue | teal | yellow |
|-------|------|------|--------|
| cat   | 29   | 96   | 98     |
| dog   | 16   | 41   | 50     |
| wheel | 26   | 90   | 13     |

- (a) What is the probability a random card is yellow given it is a cat?
- (b) What is the probability a random card is a wheel?
- (c) What is the probability a random card is a cat given it is blue?
- (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 cat and teal?
- (f) What is the probability a random card is teal?

**Solution**

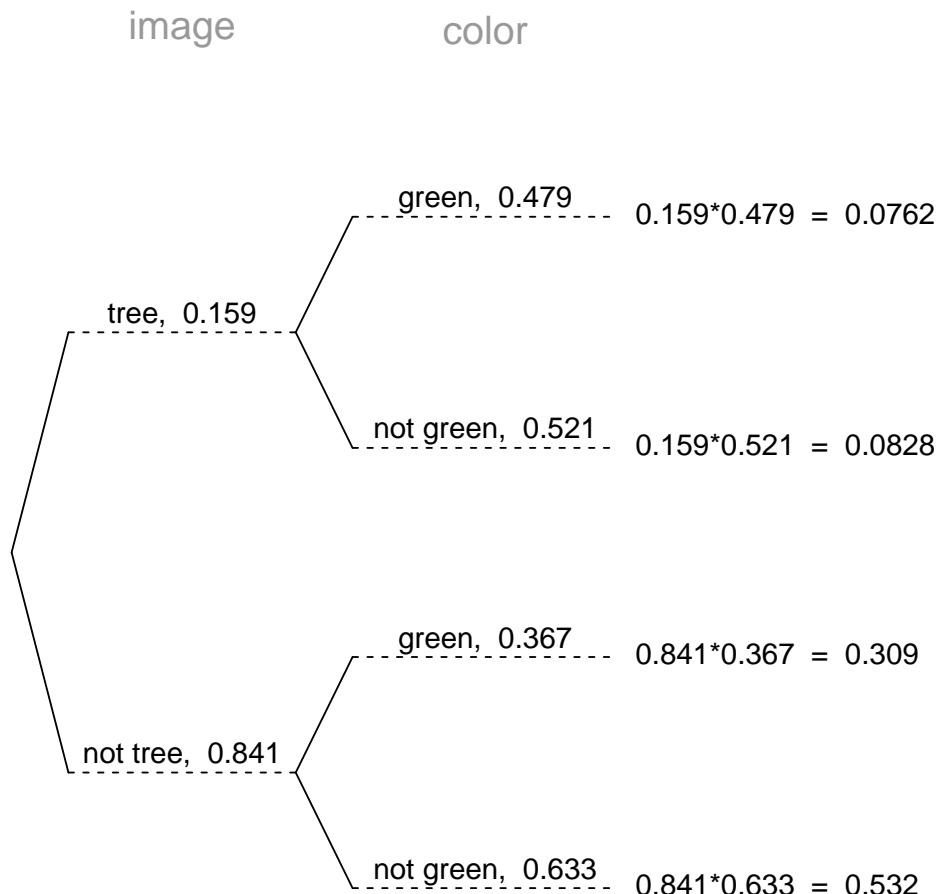
- (a)  $P(\text{yellow given cat}) = \frac{98}{29+96+98} = 0.439$
- (b)  $P(\text{wheel}) = \frac{26+90+13}{459} = 0.281$
- (c)  $P(\text{cat given blue}) = \frac{29}{29+16+26} = 0.408$
- (d)  $P(\text{wheel or yellow}) = \frac{26+90+13+98+50+13-13}{459} = 0.603$
- (e)  $P(\text{cat and teal}) = \frac{96}{459} = 0.209$
- (f)  $P(\text{teal}) = \frac{96+41+90}{459} = 0.495$

2. In a deck of strange cards, each card has an image and a color. The chance of drawing a tree is 15.9%. If a tree is drawn, there is a 47.9% chance that it is green. If a card that is not a tree is drawn, there is a 36.7% 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 tree?

**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.

$$P(\text{"not tree" given "green"}) = \frac{0.309}{0.309 + 0.0762} = 0.802$$

3. In a very large pile of toothpicks, the mean length is 63.65 millimeters and the standard deviation is 2.3 millimeters. If you randomly sample 120 toothpicks, what is the chance the sample mean is between 63.31 and 63.96 millimeters?

**Solution**

Label the given information.

$$\mu = 63.65$$

$$\sigma = 2.3$$

$$n = 120$$

$$\bar{x}_{\text{lower}} = 63.31$$

$$\bar{x}_{\text{upper}} = 63.96$$

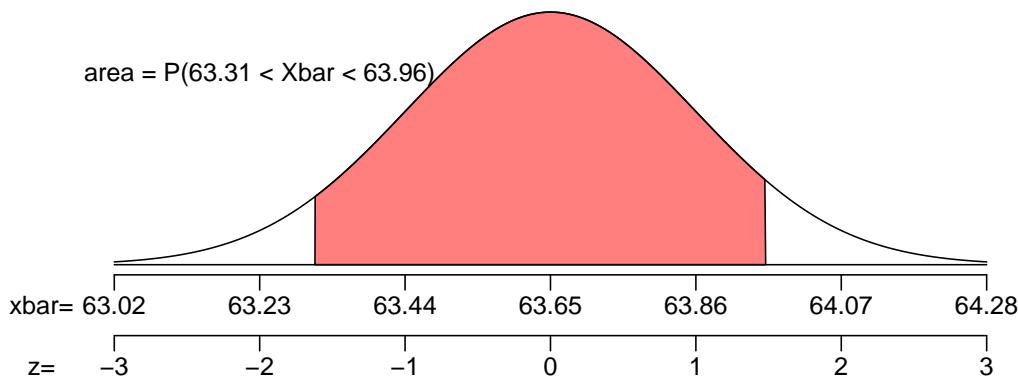
Find the standard error.

$$SE = \frac{\sigma}{\sqrt{n}} = \frac{2.3}{\sqrt{120}} = 0.21$$

Describe the sampling distribution.

$$\bar{X} \sim \mathcal{N}(63.65, 0.21)$$

Draw a sketch.



Calculate a  $z$  scores.

$$z_{\text{lower}} = \frac{\bar{x}_{\text{lower}} - \mu}{SE} = \frac{63.31 - 63.65}{0.21} = -1.62$$

$$z_{\text{upper}} = \frac{\bar{x}_{\text{upper}} - \mu}{SE} = \frac{63.96 - 63.65}{0.21} = 1.48$$

Determine the probability.

$$\begin{aligned} P(63.31 < \bar{X} < 63.96) &= \Phi(z_{\text{upper}}) - \Phi(z_{\text{lower}}) \\ &= \Phi(1.48) - \Phi(-1.62) \\ &= 0.878 \end{aligned}$$

4. In a game, there is a 46% chance to win a round. You will play 32 rounds.
- (a) What is the probability of winning exactly 17 rounds?
  - (b) What is the probability of winning at least 9 but at most 16 rounds?

**Solution**

We use the formula for binomial probabilities.

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

$$P(X = 17) = \binom{32}{17} (0.46)^{17} (1 - 0.46)^{32-17}$$

$$P(X = 17) = \binom{32}{17} (0.46)^{17} (0.54)^{15}$$

$$P(X = 17) = 0.1012$$

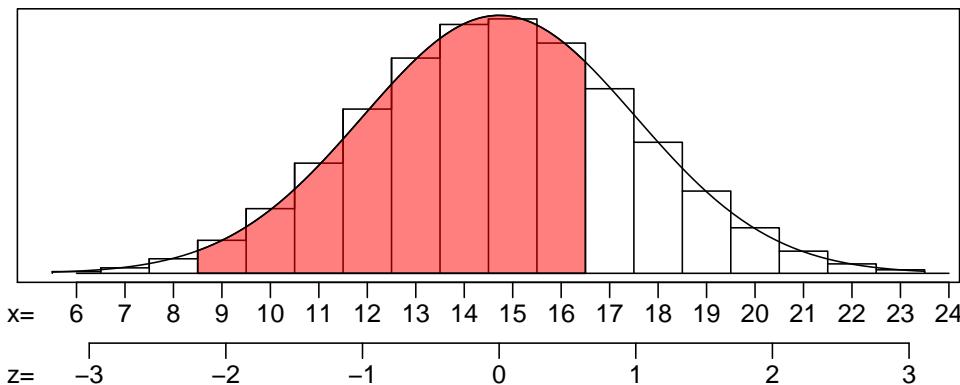
Find the mean.

$$\mu = np = (32)(0.46) = 14.72$$

Find the standard deviation.

$$\sigma = \sqrt{np(1 - p)} = \sqrt{(32)(0.46)(1 - 0.46)} = 2.8194$$

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



Find the z scores.

$$z_1 = \frac{8.5 - 14.72}{2.8194} = -2.03$$

$$z_2 = \frac{16.5 - 14.72}{2.8194} = 0.45$$

Calculate the probability.

$$P(9 \leq X \leq 16) = \Phi(0.45) - \Phi(-2.03) = 0.6524$$

(a)  $P(X = 17) = 0.1012$

(b)  $P(9 \leq X \leq 16) = 0.6524$

5. As an ornithologist, you wish to determine the average body mass of *Hylocichla mustelina*. You randomly sample 35 adults of *Hylocichla mustelina*, resulting in a sample mean of 55.64 grams and a sample standard deviation of 7.77 grams. Determine a 99% confidence interval of the true population mean.

**Solution**

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

$$n = 35$$

$$\bar{x} = 55.64$$

$$s = 7.77$$

$$CL = 0.99$$

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

$$df = n - 1 = 34$$

Determine the critical  $t$  value,  $t^*$ , such that  $P(|T| < t^*) = 0.99$ .

$$t^* = 2.73$$

Calculate the standard error.

$$SE = \frac{s}{\sqrt{n}} = \frac{7.77}{\sqrt{35}} = 1.31$$

We want to make an inference about the population mean.

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

Determine the bounds.

$$\begin{aligned} CI &= (\bar{x} - t^* SE, \bar{x} + t^* SE) \\ &= (55.64 - 2.73 \times 1.31, 55.64 + 2.73 \times 1.31) \\ &= (52.1, 59.2) \end{aligned}$$

We are 99% confident that the population mean is between 52.1 and 59.2.

6. A treatment group of size 35 has a mean of 0.996 and standard deviation of 0.0738. A control group of size 28 has a mean of 1.04 and standard deviation of 0.101. If you decided to use a significance level of 0.05, 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 48.

- (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.

$$\begin{aligned} n_1 &= 35 \\ \bar{x}_1 &= 0.996 \\ s_1 &= 0.0738 \\ n_2 &= 28 \\ \bar{x}_2 &= 1.04 \\ s_2 &= 0.101 \\ \alpha &= 0.05 \\ df &= 48 \end{aligned}$$

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  $P(|T| > t^*) = 0.05$  by using a  $t$  table.

$$t^* = 2.01$$

Calculate the standard error.

$$\begin{aligned} SE &= \sqrt{\frac{(s_1)^2}{n_1} + \frac{(s_2)^2}{n_2}} \\ &= \sqrt{\frac{(0.0738)^2}{35} + \frac{(0.101)^2}{28}} \\ &= 0.023 \end{aligned}$$

Determine the test statistic.

$$\begin{aligned} t_{\text{obs}} &= \frac{(\bar{x}_2 - \bar{x}_1) - (\mu_2 - \mu_1)_0}{SE} \\ &= \frac{(1.04 - 0.996) - (0)}{0.023} \\ &= 1.93 \end{aligned}$$

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

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

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

$$0.05 < p\text{-value} < 0.1$$

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)  $t^* = 2.01$
- (d)  $SE = 0.023$
- (e)  $|t_{\text{obs}}| = 1.93$
- (f)  $0.05 < p\text{-value} < 0.1$
- (g) retain the null

7. From a very large population, a random sample of 49000 individuals was taken. In that sample, 85.6% were glowing. Determine a 99.5% confidence interval of the population proportion.

- (a) Find the lower bound of the confidence interval.
- (b) Find the upper bound of the condifence interval.

**Solution**

Determine  $z^*$  such that  $P(|Z| < z^*) = 0.995$ .

$$z^* = 2.81$$

Calculate the standard error.

$$SE = \sqrt{\frac{\hat{p}(1 - \hat{p})}{n}} = \sqrt{\frac{(0.856)(1 - 0.856)}{49000}} = 0.00159$$

Calculate the margin of error.

$$ME = z^* SE = (2.81)(0.00159) = 0.00447$$

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.

$$(0.852, 0.86)$$

We are 99.5% confident that the true population proportion is between 85.2% and 86%.

- (a) The lower bound = 0.852, which can also be expressed as 85.2%.
- (b) The upper bound = 0.86, which can also be expressed as 86%.

8. An experiment is run with a treatment group of size 73 and a control group of size 109. The results are summarized in the table below.

|           | treatment | control |
|-----------|-----------|---------|
| angry     | 29        | 28      |
| not angry | 44        | 81      |

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

- (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.

$$\hat{p}_1 = \frac{29}{73} = 0.397$$

$$\hat{p}_2 = \frac{28}{109} = 0.257$$

Determine the difference of sample proportions.

$$\hat{p}_2 - \hat{p}_1 = 0.257 - 0.397 = -0.14$$

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

$$\hat{p} = \frac{29 + 28}{73 + 109} = 0.313$$

Determine the standard error.

$$\begin{aligned} SE &= \sqrt{\frac{\hat{p}(1 - \hat{p})}{n_1} + \frac{\hat{p}(1 - \hat{p})}{n_2}} \\ &= \sqrt{\frac{(0.313)(0.687)}{73} + \frac{(0.313)(0.687)}{109}} \\ &= 0.0701 \end{aligned}$$

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

$$H_0 : \hat{P}_2 - \hat{P}_1 \sim \mathcal{N}(0, 0.0701)$$

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.

$$\begin{aligned} z &= \frac{(\hat{p}_2 - \hat{p}_1) - (p_2 - p_1)_0}{SE} \\ &= \frac{(0.257 - 0.397) - 0}{0.0701} \\ &= -2 \end{aligned}$$

Determine the  $p$ -value.

$$\begin{aligned} p\text{-value} &= 2 \cdot \Phi(-|z|) \\ &= 2 \cdot \Phi(-2) \\ &= 0.0456 \end{aligned}$$

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)  $SE = 0.0701$
- (e)  $|z_{\text{obs}}| = 2$
- (f)  $p\text{-value} = 0.0456$
- (g) reject the null

# Bunker Hill Community College

Final Statistics Exam 2019-05-02

Exam ID 022

**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)  $P(\text{gem or orange}) = 0.405$

(b)  $P(\text{gem}) = 0.262$

(c)  $P(\text{teal}) = 0.378$

(d)  $P(\text{wheel and gray}) = 0.219$

(e)  $P(\text{gem given gray}) = 0.133$

(f)  $P(\text{teal given wheel}) = 0.184$

2.  $P(\text{"not dog" given "not white"}) = 0.31$

3.  $P(60.24 < X < 60.69) = 0.895$

4. (a)  $P(X = 73) = 0.0575$

(b)  $P(66 \leq X \leq 81) = 0.7044$

5. **(20.6, 23.1)**

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

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

(c)  $t^* = 2.38$

(d)  $SE = 0.05$

(e)  $|t_{\text{obs}}| = 2.42$

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

(g) **reject**

7. (a) **LB of p CI = 0.44 or 44%**

(b) **UB of p CI = 0.514 or 51.4%**

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

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

(c)  $Z^* = 1.96$

(d)  $SE = 0.042$

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

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

(g) **reject**

1. In a deck of strange cards, there are 447 cards. Each card has an image and a color. The amounts are shown in the table below.

|       | gray | orange | teal |
|-------|------|--------|------|
| gem   | 22   | 49     | 46   |
| horn  | 45   | 16     | 90   |
| wheel | 98   | 48     | 33   |

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

**Solution**

(a)  $P(\text{gem or orange}) = \frac{22+49+46+49+16+48-49}{447} = 0.405$

(b)  $P(\text{gem}) = \frac{22+49+46}{447} = 0.262$

(c)  $P(\text{teal}) = \frac{46+90+33}{447} = 0.378$

(d)  $P(\text{wheel and gray}) = \frac{98}{447} = 0.219$

(e)  $P(\text{gem given gray}) = \frac{22}{22+45+98} = 0.133$

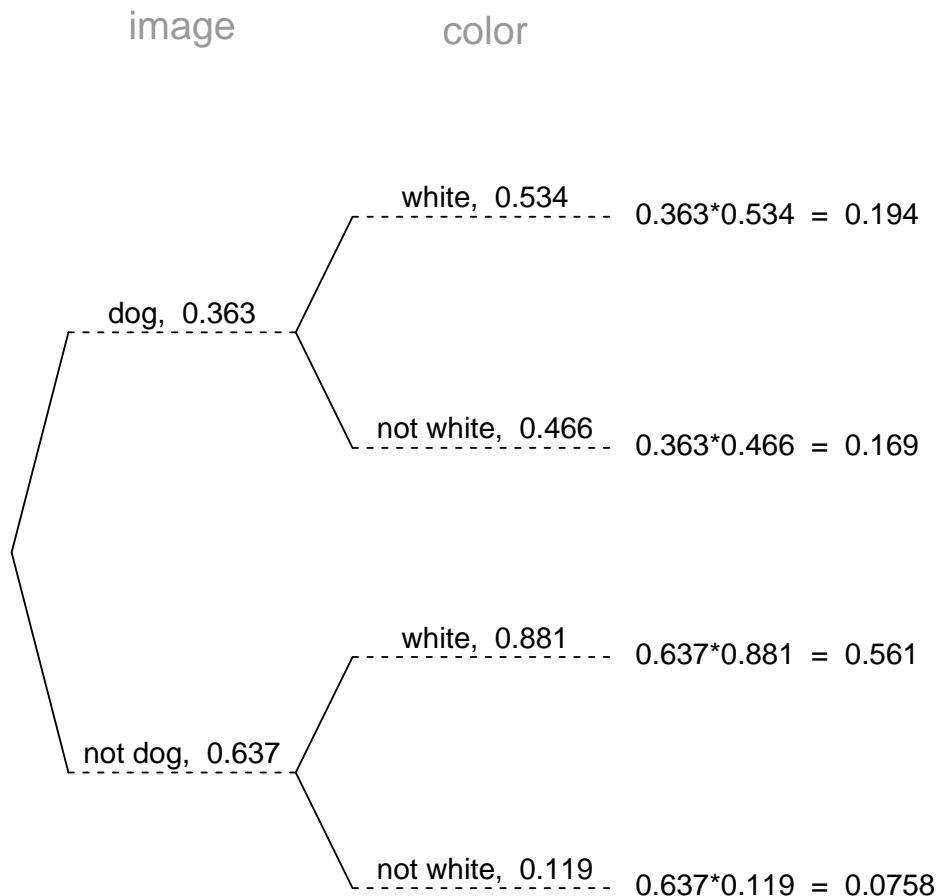
(f)  $P(\text{teal given wheel}) = \frac{33}{98+48+33} = 0.184$

2. In a deck of strange cards, each card has an image and a color. The chance of drawing a dog is 36.3%. If a dog is drawn, there is a 53.4% chance that it is white. If a card that is not a dog is drawn, there is a 88.1% chance that it is white.

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

**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.

$$P(\text{"not dog" given "not white"}) = \frac{0.0758}{0.0758 + 0.169} = 0.31$$

3. In a very large pile of toothpicks, the mean length is 60.51 millimeters and the standard deviation is 1.44 millimeters. If you randomly sample 121 toothpicks, what is the chance the sample mean is between 60.24 and 60.69 millimeters?

**Solution**

Label the given information.

$$\mu = 60.51$$

$$\sigma = 1.44$$

$$n = 121$$

$$\bar{x}_{\text{lower}} = 60.24$$

$$\bar{x}_{\text{upper}} = 60.69$$

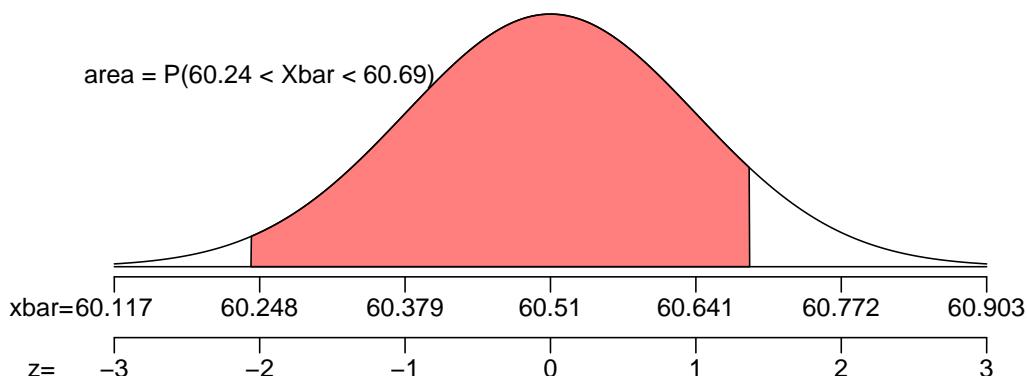
Find the standard error.

$$SE = \frac{\sigma}{\sqrt{n}} = \frac{1.44}{\sqrt{121}} = 0.131$$

Describe the sampling distribution.

$$\bar{X} \sim \mathcal{N}(60.51, 0.131)$$

Draw a sketch.



Calculate a  $z$  scores.

$$z_{\text{lower}} = \frac{\bar{x}_{\text{lower}} - \mu}{SE} = \frac{60.24 - 60.51}{0.131} = -2.06$$

$$z_{\text{upper}} = \frac{\bar{x}_{\text{upper}} - \mu}{SE} = \frac{60.69 - 60.51}{0.131} = 1.37$$

Determine the probability.

$$\begin{aligned} P(60.24 < \bar{X} < 60.69) &= \Phi(z_{\text{upper}}) - \Phi(z_{\text{lower}}) \\ &= \Phi(1.37) - \Phi(-2.06) \\ &= 0.895 \end{aligned}$$

4. In a game, there is a 60% chance to win a round. You will play 115 rounds.
- (a) What is the probability of winning exactly 73 rounds?
  - (b) What is the probability of winning at least 66 but at most 81 rounds?

**Solution**

We use the formula for binomial probabilities.

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

$$P(X = 73) = \binom{115}{73} (0.6)^{73} (1 - 0.6)^{115-73}$$

$$P(X = 73) = \binom{115}{73} (0.6)^{73} (0.4)^{42}$$

$$P(X = 73) = 0.0575$$

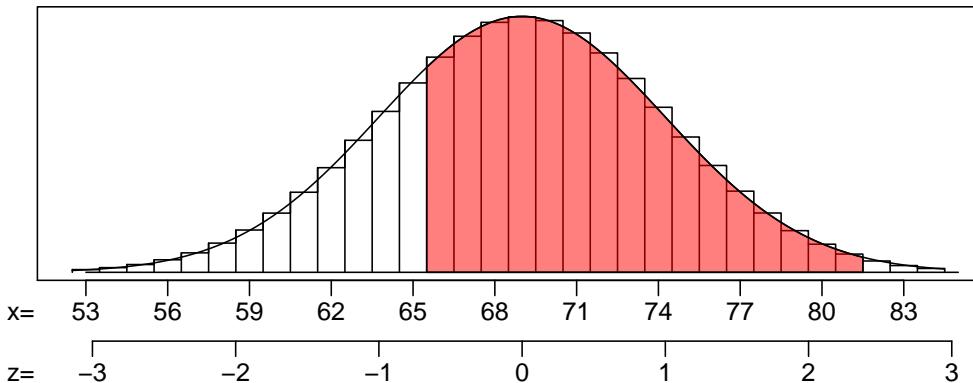
Find the mean.

$$\mu = np = (115)(0.6) = 69$$

Find the standard deviation.

$$\sigma = \sqrt{np(1 - p)} = \sqrt{(115)(0.6)(1 - 0.6)} = 5.2536$$

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



Find the z scores.

$$z_1 = \frac{65.5 - 69}{5.2536} = -0.57$$

$$z_2 = \frac{81.5 - 69}{5.2536} = 2.28$$

Calculate the probability.

$$P(66 \leq X \leq 81) = \Phi(2.28) - \Phi(-0.57) = 0.7044$$

(a)  $P(X = 73) = 0.0575$

(b)  $P(66 \leq X \leq 81) = 0.7044$

5. As an ornithologist, you wish to determine the average body mass of *Icterus spurius*. You randomly sample 13 adults of *Icterus spurius*, resulting in a sample mean of 21.81 grams and a sample standard deviation of 2.07 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.

$$n = 13$$

$$\bar{x} = 21.81$$

$$s = 2.07$$

$$CL = 0.95$$

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

$$df = n - 1 = 12$$

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

$$t^* = 2.18$$

Calculate the standard error.

$$SE = \frac{s}{\sqrt{n}} = \frac{2.07}{\sqrt{13}} = 0.574$$

We want to make an inference about the population mean.

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

Determine the bounds.

$$\begin{aligned} CI &= (\bar{x} - t^* SE, \bar{x} + t^* SE) \\ &= (21.81 - 2.18 \times 0.574, 21.81 + 2.18 \times 0.574) \\ &= (20.6, 23.1) \end{aligned}$$

We are 95% confident that the population mean is between 20.6 and 23.1.

6. A treatment group of size 35 has a mean of 1.03 and standard deviation of 0.188. A control group of size 38 has a mean of 1.15 and standard deviation of 0.235. 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 69.

- (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.

$$\begin{aligned}n_1 &= 35 \\ \bar{x}_1 &= 1.03 \\ s_1 &= 0.188 \\ n_2 &= 38 \\ \bar{x}_2 &= 1.15 \\ s_2 &= 0.235 \\ \alpha &= 0.02 \\ df &= 69\end{aligned}$$

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.

$$t^* = 2.38$$

Calculate the standard error.

$$\begin{aligned}SE &= \sqrt{\frac{(s_1)^2}{n_1} + \frac{(s_2)^2}{n_2}} \\ &= \sqrt{\frac{(0.188)^2}{35} + \frac{(0.235)^2}{38}} \\ &= 0.05\end{aligned}$$

Determine the test statistic.

$$\begin{aligned}t_{\text{obs}} &= \frac{(\bar{x}_2 - \bar{x}_1) - (\mu_2 - \mu_1)_0}{SE} \\ &= \frac{(1.15 - 1.03) - (0)}{0.05} \\ &= 2.42\end{aligned}$$

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)  $t^* = 2.38$
- (d)  $SE = 0.05$
- (e)  $|t_{\text{obs}}| = 2.42$
- (f)  $0.01 < p\text{-value} < 0.02$
- (g) reject the null

7. From a very large population, a random sample of 770 individuals was taken. In that sample, 47.7% were green. 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 condifence interval.

**Solution**

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

$$z^* = 2.05$$

Calculate the standard error.

$$SE = \sqrt{\frac{\hat{p}(1 - \hat{p})}{n}} = \sqrt{\frac{(0.477)(1 - 0.477)}{770}} = 0.018$$

Calculate the margin of error.

$$ME = z^* SE = (2.05)(0.018) = 0.0369$$

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.

$$(0.44, 0.514)$$

We are 96% confident that the true population proportion is between 44% and 51.4%.

- (a) The lower bound = 0.44, which can also be expressed as 44%.
- (b) The upper bound = 0.514, which can also be expressed as 51.4%.

8. An experiment is run with a treatment group of size 198 and a control group of size 175. The results are summarized in the table below.

|                 | treatment | control |
|-----------------|-----------|---------|
| fluorescent     | 32        | 43      |
| not fluorescent | 166       | 132     |

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

- (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.

$$\hat{p}_1 = \frac{32}{198} = 0.162$$

$$\hat{p}_2 = \frac{43}{175} = 0.246$$

Determine the difference of sample proportions.

$$\hat{p}_2 - \hat{p}_1 = 0.246 - 0.162 = 0.084$$

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

$$\hat{p} = \frac{32 + 43}{198 + 175} = 0.201$$

Determine the standard error.

$$\begin{aligned} SE &= \sqrt{\frac{\hat{p}(1 - \hat{p})}{n_1} + \frac{\hat{p}(1 - \hat{p})}{n_2}} \\ &= \sqrt{\frac{(0.201)(0.799)}{198} + \frac{(0.201)(0.799)}{175}} \\ &= 0.0416 \end{aligned}$$

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

$$H_0 : \hat{P}_2 - \hat{P}_1 \sim \mathcal{N}(0, 0.0416)$$

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.

$$\begin{aligned} z &= \frac{(\hat{p}_2 - \hat{p}_1) - (p_2 - p_1)_0}{SE} \\ &= \frac{(0.246 - 0.162) - 0}{0.0416} \\ &= 2.02 \end{aligned}$$

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)  $SE = 0.0416$
- (e)  $|z_{\text{obs}}| = 2.02$
- (f)  $p\text{-value} = 0.0434$
- (g) reject the null

# Bunker Hill Community College

Final Statistics Exam 2019-05-02

Exam ID 023

**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)  $P(\text{cat}) = 0.296$

(b)  $P(\text{bike or green}) = 0.427$

(c)  $P(\text{cat and green}) = 0.0193$

(d)  $P(\text{green}) = 0.326$

(e)  $P(\text{green given bike}) = 0.451$

(f)  $P(\text{shovel given orange}) = 0.245$

2.  $P(\text{"not wheel" given "blue"}) = 0.943$

3.  $P(65.93 < X < 66.24) = 0.7574$

4. (a)  $P(X = 85) = 0.1117$

(b)  $P(84 \leq X \leq 92) = 0.7435$

5. **(10.1, 11.1)**

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

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

(c)  $t^* = 2.66$

(d)  $SE = 61.965$

(e)  $|t_{\text{obs}}| = 2.9$

(f)  $0.005 < p\text{-value} < 0.01$

(g) **reject**

7. (a) **LB of p CI = 0.806 or 80.6%**

(b) **UB of p CI = 0.864 or 86.4%**

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

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

(c)  $Z^* = 2.33$

(d)  $SE = 0.047$

(e)  $|Z_{\text{obs}}| = 2.46$

(f)  $p\text{-value} = 0.0138$

(g) **reject**

1. In a deck of strange cards, there are 724 cards. Each card has an image and a color. The amounts are shown in the table below.

|        | green | orange | teal | white |
|--------|-------|--------|------|-------|
| bike   | 60    | 50     | 10   | 13    |
| cat    | 14    | 52     | 71   | 77    |
| flower | 73    | 18     | 63   | 16    |
| shovel | 89    | 39     | 54   | 25    |

- (a) What is the probability a random card is a cat?
- (b) What is the probability a random card is either a bike or green (or both)?
- (c) What is the probability a random card is both a cat and green?
- (d) What is the probability a random card is green?
- (e) What is the probability a random card is green given it is a bike?
- (f) What is the probability a random card is a shovel given it is orange?

**Solution**

(a)  $P(\text{cat}) = \frac{14+52+71+77}{724} = 0.296$

(b)  $P(\text{bike or green}) = \frac{60+50+10+13+60+14+73+89-60}{724} = 0.427$

(c)  $P(\text{cat and green}) = \frac{14}{724} = 0.0193$

(d)  $P(\text{green}) = \frac{60+14+73+89}{724} = 0.326$

(e)  $P(\text{green given bike}) = \frac{60}{60+50+10+13} = 0.451$

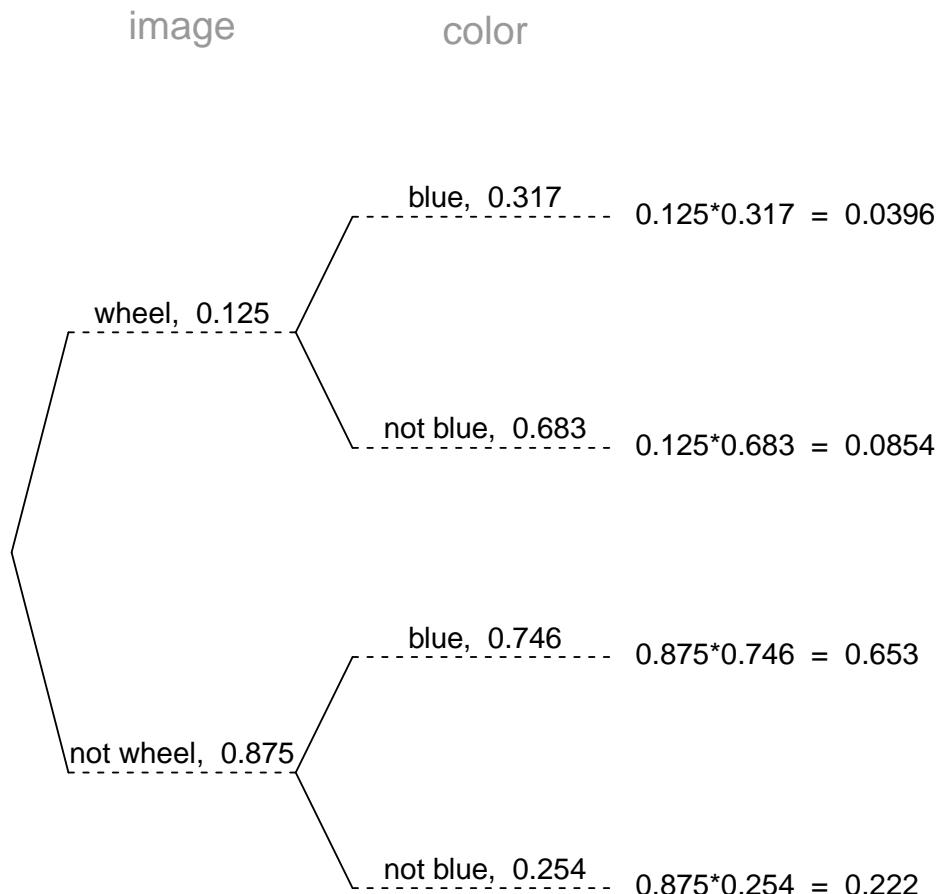
(f)  $P(\text{shovel given orange}) = \frac{39}{50+52+18+39} = 0.245$

2. In a deck of strange cards, each card has an image and a color. The chance of drawing a wheel is 12.5%. If a wheel is drawn, there is a 31.7% chance that it is blue. If a card that is not a wheel is drawn, there is a 74.6% chance that it is blue.

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

**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.

$$P(\text{"not wheel" given "blue"}) = \frac{0.653}{0.653 + 0.0396} = 0.943$$

3. In a very large pile of toothpicks, the mean length is 66.17 millimeters and the standard deviation is 1.07 millimeters. If you randomly sample 121 toothpicks, what is the chance the sample mean is between 65.93 and 66.24 millimeters?

**Solution**

Label the given information.

$$\mu = 66.17$$

$$\sigma = 1.07$$

$$n = 121$$

$$\bar{x}_{\text{lower}} = 65.93$$

$$\bar{x}_{\text{upper}} = 66.24$$

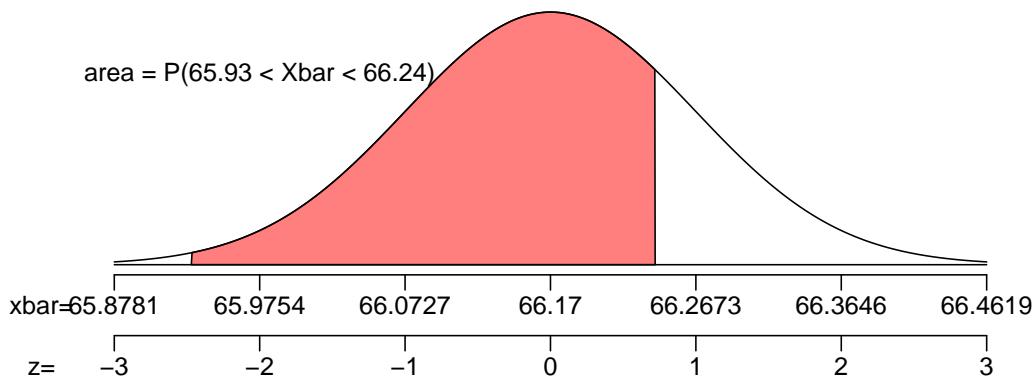
Find the standard error.

$$SE = \frac{\sigma}{\sqrt{n}} = \frac{1.07}{\sqrt{121}} = 0.0973$$

Describe the sampling distribution.

$$\bar{X} \sim \mathcal{N}(66.17, 0.0973)$$

Draw a sketch.



Calculate a  $z$  scores.

$$z_{\text{lower}} = \frac{\bar{x}_{\text{lower}} - \mu}{SE} = \frac{65.93 - 66.17}{0.0973} = -2.47$$

$$z_{\text{upper}} = \frac{\bar{x}_{\text{upper}} - \mu}{SE} = \frac{66.24 - 66.17}{0.0973} = 0.72$$

Determine the probability.

$$\begin{aligned} P(65.93 < \bar{X} < 66.24) &= \Phi(z_{\text{upper}}) - \Phi(z_{\text{lower}}) \\ &= \Phi(0.72) - \Phi(-2.47) \\ &= 0.7574 \end{aligned}$$

4. In a game, there is a 89% chance to win a round. You will play 97 rounds.
- (a) What is the probability of winning exactly 85 rounds?
  - (b) What is the probability of winning at least 84 but at most 92 rounds?

**Solution**

We use the formula for binomial probabilities.

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

$$P(X = 85) = \binom{97}{85} (0.89)^{85} (1 - 0.89)^{97-85}$$

$$P(X = 85) = \binom{97}{85} (0.89)^{85} (0.11)^{12}$$

$$P(X = 85) = 0.1117$$

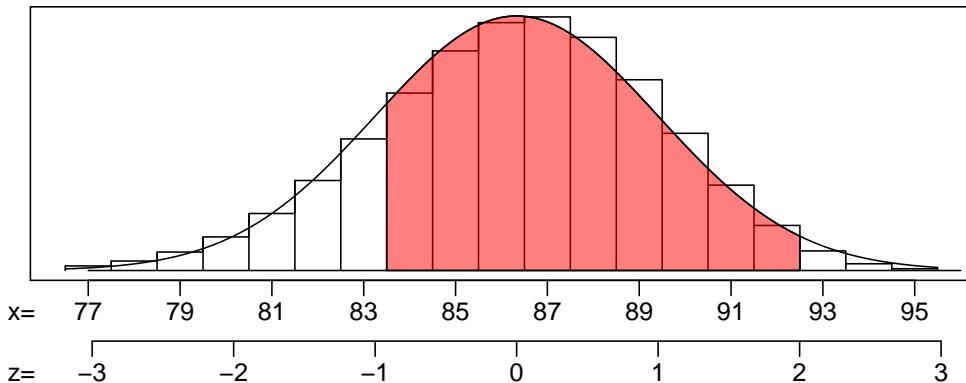
Find the mean.

$$\mu = np = (97)(0.89) = 86.33$$

Find the standard deviation.

$$\sigma = \sqrt{np(1 - p)} = \sqrt{(97)(0.89)(1 - 0.89)} = 3.0816$$

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



Find the z scores.

$$z_1 = \frac{83.5 - 86.33}{3.0816} = -0.76$$

$$z_2 = \frac{92.5 - 86.33}{3.0816} = 1.84$$

Calculate the probability.

$$P(84 \leq X \leq 92) = \Phi(1.84) - \Phi(-0.76) = 0.7435$$

(a)  $P(X = 85) = 0.1117$

(b)  $P(84 \leq X \leq 92) = 0.7435$

5. As an ornithologist, you wish to determine the average body mass of *Geothlypis trichas*. You randomly sample 32 adults of *Geothlypis trichas*, resulting in a sample mean of 10.6 grams and a sample standard deviation of 1.09 grams. Determine a 99% confidence interval of the true population mean.

**Solution**

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

$$n = 32$$

$$\bar{x} = 10.6$$

$$s = 1.09$$

$$CL = 0.99$$

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

$$df = n - 1 = 31$$

Determine the critical  $t$  value,  $t^*$ , such that  $P(|T| < t^*) = 0.99$ .

$$t^* = 2.74$$

Calculate the standard error.

$$SE = \frac{s}{\sqrt{n}} = \frac{1.09}{\sqrt{32}} = 0.193$$

We want to make an inference about the population mean.

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

Determine the bounds.

$$\begin{aligned} CI &= (\bar{x} - t^* SE, \bar{x} + t^* SE) \\ &= (10.6 - 2.74 \times 0.193, 10.6 + 2.74 \times 0.193) \\ &= (10.1, 11.1) \end{aligned}$$

We are 99% confident that the population mean is between 10.1 and 11.1.

6. A treatment group of size 28 has a mean of 1210 and standard deviation of 204. A control group of size 35 has a mean of 1030 and standard deviation of 287. If you decided to use a significance level of 0.01, 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 60.

- (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.

$$n_1 = 28$$

$$\bar{x}_1 = 1210$$

$$s_1 = 204$$

$$n_2 = 35$$

$$\bar{x}_2 = 1030$$

$$s_2 = 287$$

$$\alpha = 0.01$$

$$df = 60$$

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  $P(|T| > t^*) = 0.01$  by using a  $t$  table.

$$t^* = 2.66$$

Calculate the standard error.

$$\begin{aligned} SE &= \sqrt{\frac{(s_1)^2}{n_1} + \frac{(s_2)^2}{n_2}} \\ &= \sqrt{\frac{(204)^2}{28} + \frac{(287)^2}{35}} \\ &= 61.965 \end{aligned}$$

Determine the test statistic.

$$\begin{aligned} t_{\text{obs}} &= \frac{(\bar{x}_2 - \bar{x}_1) - (\mu_2 - \mu_1)_0}{SE} \\ &= \frac{(1030 - 1210) - (0)}{61.965} \\ &= -2.9 \end{aligned}$$

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

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

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

$$0.005 < p\text{-value} < 0.01$$

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)  $t^* = 2.66$
- (d)  $SE = 61.965$
- (e)  $|t_{\text{obs}}| = 2.9$
- (f)  $0.005 < p\text{-value} < 0.01$
- (g) reject the null

7. From a very large population, a random sample of 650 individuals was taken. In that sample, 83.5% were super. 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 condifence interval.

**Solution**

Determine  $z^*$  such that  $P(|Z| < z^*) = 0.95$ .

$$z^* = 1.96$$

Calculate the standard error.

$$SE = \sqrt{\frac{\hat{p}(1 - \hat{p})}{n}} = \sqrt{\frac{(0.835)(1 - 0.835)}{650}} = 0.0146$$

Calculate the margin of error.

$$ME = z^* SE = (1.96)(0.0146) = 0.0286$$

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.

$$(0.806, 0.864)$$

We are 95% confident that the true population proportion is between 80.6% and 86.4%.

- (a) The lower bound = 0.806, which can also be expressed as 80.6%.
- (b) The upper bound = 0.864, which can also be expressed as 86.4%.

8. An experiment is run with a treatment group of size 221 and a control group of size 234. The results are summarized in the table below.

|          | treatment | control |
|----------|-----------|---------|
| cold     | 103       | 136     |
| not cold | 118       | 98      |

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

- (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.

$$\hat{p}_1 = \frac{103}{221} = 0.466$$

$$\hat{p}_2 = \frac{136}{234} = 0.581$$

Determine the difference of sample proportions.

$$\hat{p}_2 - \hat{p}_1 = 0.581 - 0.466 = 0.115$$

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

$$\hat{p} = \frac{103 + 136}{221 + 234} = 0.525$$

Determine the standard error.

$$\begin{aligned} SE &= \sqrt{\frac{\hat{p}(1 - \hat{p})}{n_1} + \frac{\hat{p}(1 - \hat{p})}{n_2}} \\ &= \sqrt{\frac{(0.525)(0.475)}{221} + \frac{(0.525)(0.475)}{234}} \\ &= 0.0468 \end{aligned}$$

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

$$H_0 : \hat{P}_2 - \hat{P}_1 \sim \mathcal{N}(0, 0.0468)$$

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.

$$\begin{aligned} z &= \frac{(\hat{p}_2 - \hat{p}_1) - (p_2 - p_1)_0}{SE} \\ &= \frac{(0.581 - 0.466) - 0}{0.0468} \\ &= 2.46 \end{aligned}$$

Determine the  $p$ -value.

$$\begin{aligned} p\text{-value} &= 2 \cdot \Phi(-|z|) \\ &= 2 \cdot \Phi(-2.46) \\ &= 0.0138 \end{aligned}$$

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)  $SE = 0.0468$
- (e)  $|Z_{\text{obs}}| = 2.46$
- (f)  $p\text{-value} = 0.0138$
- (g) reject the null

# Bunker Hill Community College

Final Statistics Exam 2019-05-02

Exam ID 024

**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)  $P(\text{yellow given gem}) = 0.209$

(b)  $P(\text{dog and red}) = 0.0757$

(c)  $P(\text{orange}) = 0.297$

(d)  $P(\text{dog or red}) = 0.487$

(e)  $P(\text{dog given orange}) = 0.396$

(f)  $P(\text{gem}) = 0.383$

2.  $P(\text{"cat" given "red"}) = 0.708$

3.  $P(72.13 < X < 72.47) = 0.8444$

4. (a)  $P(X = 12) = 0.083$

(b)  $P(7 \leq X \leq 24) = 0.9775$

5. **(12.7, 14.8)**

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

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

(c)  $t^* = 1.72$

(d)  $SE = 3.284$

(e)  $|t_{\text{obs}}| = 1.52$

(f)  $0.1 < p\text{-value} < 0.2$

(g) **retain**

7. (a) **LB of p CI = 0.31 or 31%**

(b) **UB of p CI = 0.318 or 31.8%**

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

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

(c)  $Z^* = 1.28$

(d)  $SE = 0.058$

(e)  $|Z_{\text{obs}}| = 1.54$

(f)  $p\text{-value} = 0.1236$

(g) **reject**

1. In a deck of strange cards, there are 713 cards. Each card has an image and a color. The amounts are shown in the table below.

|      | green | orange | red | yellow |
|------|-------|--------|-----|--------|
| dog  | 11    | 84     | 54  | 26     |
| gem  | 99    | 34     | 83  | 57     |
| tree | 38    | 94     | 89  | 44     |

- (a) What is the probability a random card is yellow given it is a gem?
- (b) What is the probability a random card is both a dog and red?
- (c) What is the probability a random card is orange?
- (d) What is the probability a random card is either a dog or red (or both)?
- (e) What is the probability a random card is a dog given it is orange?
- (f) What is the probability a random card is a gem?

**Solution**

(a)  $P(\text{yellow given gem}) = \frac{57}{99+34+83+57} = 0.209$

(b)  $P(\text{dog and red}) = \frac{54}{713} = 0.0757$

(c)  $P(\text{orange}) = \frac{84+34+94}{713} = 0.297$

(d)  $P(\text{dog or red}) = \frac{11+84+54+26+54+83+89-54}{713} = 0.487$

(e)  $P(\text{dog given orange}) = \frac{84}{84+34+94} = 0.396$

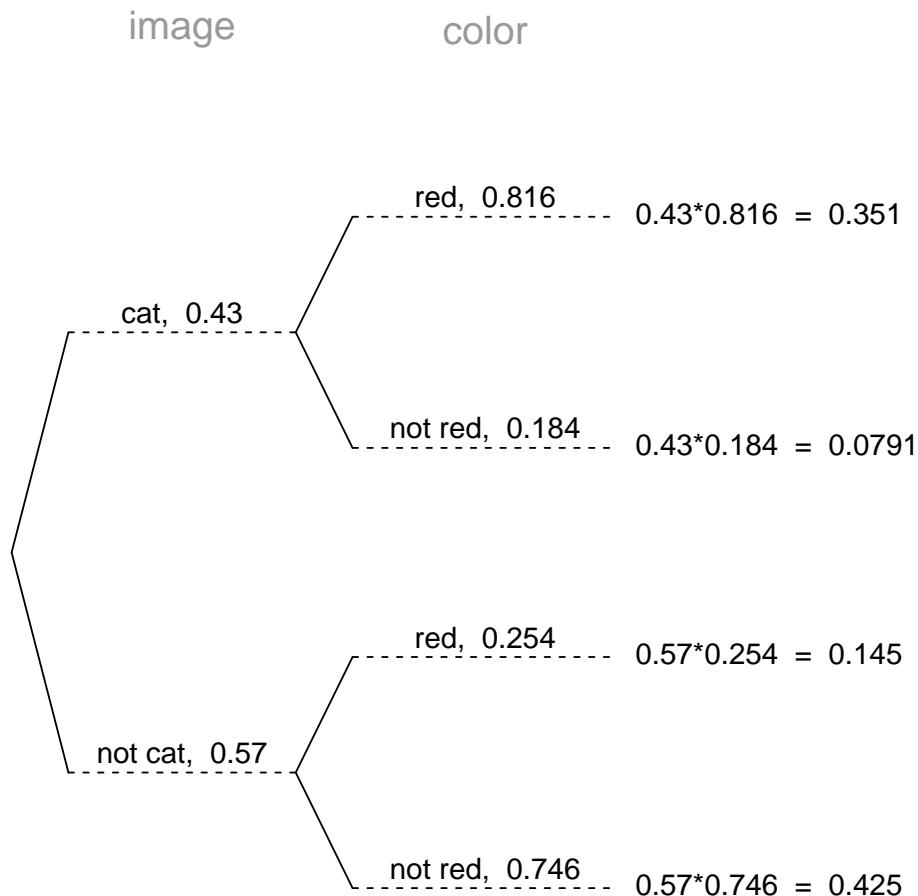
(f)  $P(\text{gem}) = \frac{99+34+83+57}{713} = 0.383$

2. In a deck of strange cards, each card has an image and a color. The chance of drawing a cat is 43%. If a cat is drawn, there is a 81.6% chance that it is red. If a card that is not a cat is drawn, there is a 25.4% chance that it is red.

Now, someone draws a random card and reveals it is red. What is the chance the card is a cat?

**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.

$$P(\text{"cat" given "red"}) = \frac{0.351}{0.351 + 0.145} = 0.708$$

3. In a very large pile of toothpicks, the mean length is 72.37 millimeters and the standard deviation is 1.36 millimeters. If you randomly sample 200 toothpicks, what is the chance the sample mean is between 72.13 and 72.47 millimeters?

**Solution**

Label the given information.

$$\mu = 72.37$$

$$\sigma = 1.36$$

$$n = 200$$

$$\bar{x}_{\text{lower}} = 72.13$$

$$\bar{x}_{\text{upper}} = 72.47$$

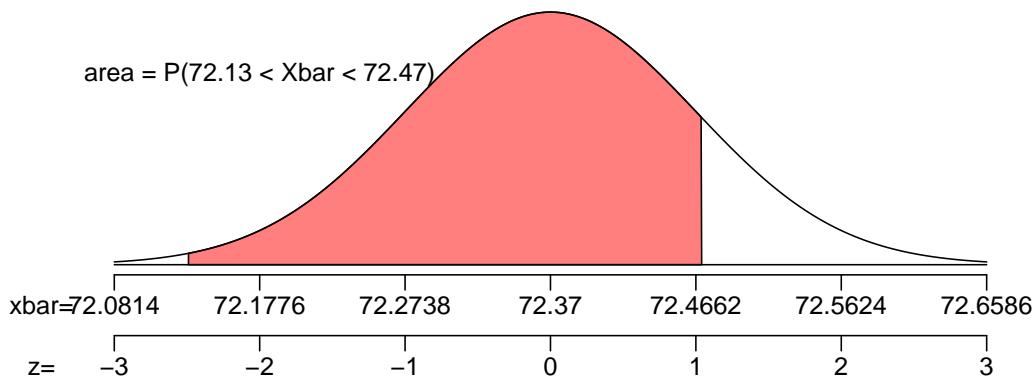
Find the standard error.

$$SE = \frac{\sigma}{\sqrt{n}} = \frac{1.36}{\sqrt{200}} = 0.0962$$

Describe the sampling distribution.

$$\bar{X} \sim \mathcal{N}(72.37, 0.0962)$$

Draw a sketch.



Calculate a  $z$  scores.

$$z_{\text{lower}} = \frac{x_{\text{lower}} - \mu}{SE} = \frac{72.13 - 72.37}{0.0962} = -2.49$$

$$z_{\text{upper}} = \frac{x_{\text{upper}} - \mu}{SE} = \frac{72.47 - 72.37}{0.0962} = 1.04$$

Determine the probability.

$$\begin{aligned} P(72.13 < \bar{X} < 72.47) &= \Phi(z_{\text{upper}}) - \Phi(z_{\text{lower}}) \\ &= \Phi(1.04) - \Phi(-2.49) \\ &= 0.8444 \end{aligned}$$

4. In a game, there is a 9% chance to win a round. You will play 167 rounds.
- (a) What is the probability of winning exactly 12 rounds?
  - (b) What is the probability of winning at least 7 but at most 24 rounds?

**Solution**

We use the formula for binomial probabilities.

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

$$P(X = 12) = \binom{167}{12} (0.09)^{12} (1 - 0.09)^{167-12}$$

$$P(X = 12) = \binom{167}{12} (0.09)^{12} (0.91)^{155}$$

$$P(X = 12) = 0.083$$

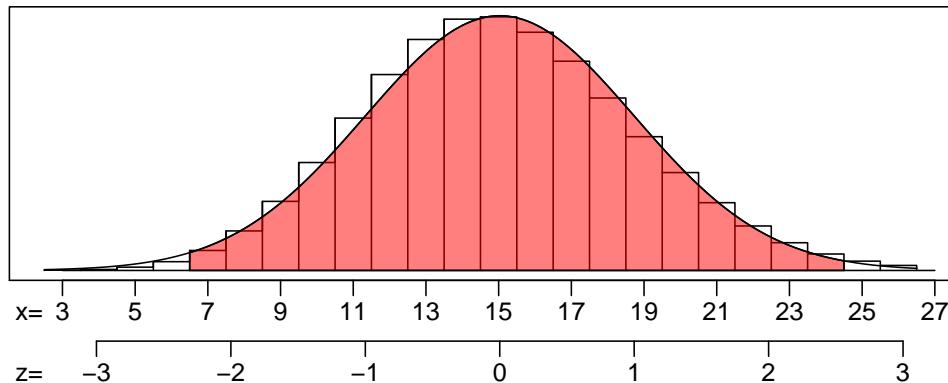
Find the mean.

$$\mu = np = (167)(0.09) = 15.03$$

Find the standard deviation.

$$\sigma = \sqrt{np(1 - p)} = \sqrt{(167)(0.09)(1 - 0.09)} = 3.6983$$

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



Find the z scores.

$$z_1 = \frac{6.5 - 15.03}{3.6983} = -2.17$$

$$z_2 = \frac{24.5 - 15.03}{3.6983} = 2.43$$

Calculate the probability.

$$P(7 \leq X \leq 24) = \Phi(2.43) - \Phi(-2.17) = 0.9775$$

(a)  $P(X = 12) = 0.083$

(b)  $P(7 \leq X \leq 24) = 0.9775$

5. As an ornithologist, you wish to determine the average body mass of *Protonotaria citrea*. You randomly sample 21 adults of *Protonotaria citrea*, resulting in a sample mean of 13.75 grams and a sample standard deviation of 1.64 grams. Determine a 99% confidence interval of the true population mean.

**Solution**

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

$$n = 21$$

$$\bar{x} = 13.75$$

$$s = 1.64$$

$$CL = 0.99$$

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

$$df = n - 1 = 20$$

Determine the critical  $t$  value,  $t^*$ , such that  $P(|T| < t^*) = 0.99$ .

$$t^* = 2.85$$

Calculate the standard error.

$$SE = \frac{s}{\sqrt{n}} = \frac{1.64}{\sqrt{21}} = 0.358$$

We want to make an inference about the population mean.

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

Determine the bounds.

$$\begin{aligned} CI &= (\bar{x} - t^* SE, \bar{x} + t^* SE) \\ &= (13.75 - 2.85 \times 0.358, 13.75 + 2.85 \times 0.358) \\ &= (12.7, 14.8) \end{aligned}$$

We are 99% confident that the population mean is between 12.7 and 14.8.

6. A treatment group of size 13 has a mean of 106 and standard deviation of 10.3. A control group of size 38 has a mean of 101 and standard deviation of 9.99. If you decided to use a significance level of 0.1, 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 20.

- (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.

$$n_1 = 13$$

$$\bar{x}_1 = 106$$

$$s_1 = 10.3$$

$$n_2 = 38$$

$$\bar{x}_2 = 101$$

$$s_2 = 9.99$$

$$\alpha = 0.1$$

$$df = 20$$

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  $P(|T| > t^*) = 0.1$  by using a  $t$  table.

$$t^* = 1.72$$

Calculate the standard error.

$$\begin{aligned} SE &= \sqrt{\frac{(s_1)^2}{n_1} + \frac{(s_2)^2}{n_2}} \\ &= \sqrt{\frac{(10.3)^2}{13} + \frac{(9.99)^2}{38}} \\ &= 3.284 \end{aligned}$$

Determine the test statistic.

$$\begin{aligned} t_{\text{obs}} &= \frac{(\bar{x}_2 - \bar{x}_1) - (\mu_2 - \mu_1)_0}{SE} \\ &= \frac{(101 - 106) - (0)}{3.284} \\ &= -1.52 \end{aligned}$$

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

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

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

$$0.1 < p\text{-value} < 0.2$$

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)  $t^* = 1.72$
- (d)  $SE = 3.284$
- (e)  $|t_{\text{obs}}| = 1.52$
- (f)  $0.1 < p\text{-value} < 0.2$
- (g) retain the null

7. From a very large population, a random sample of 71000 individuals was taken. In that sample, 31.4% were floating. Determine a 98% confidence interval of the population proportion.

- (a) Find the lower bound of the confidence interval.
- (b) Find the upper bound of the condifence interval.

**Solution**

Determine  $z^*$  such that  $P(|Z| < z^*) = 0.98$ .

$$z^* = 2.33$$

Calculate the standard error.

$$SE = \sqrt{\frac{\hat{p}(1 - \hat{p})}{n}} = \sqrt{\frac{(0.314)(1 - 0.314)}{71000}} = 0.00174$$

Calculate the margin of error.

$$ME = z^* SE = (2.33)(0.00174) = 0.00405$$

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.

$$(0.31, 0.318)$$

We are 98% confident that the true population proportion is between 31% and 31.8%.

- (a) The lower bound = 0.31, which can also be expressed as 31%.
- (b) The upper bound = 0.318, which can also be expressed as 31.8%.

8. An experiment is run with a treatment group of size 104 and a control group of size 149. The results are summarized in the table below.

|           | treatment | control |
|-----------|-----------|---------|
| sorry     | 35        | 37      |
| not sorry | 69        | 112     |

Using a significance level of 0.2, 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  $P(|Z| > z^*) = 0.2$ .

$$z^* = \Phi^{-1} \left( 1 - \frac{\alpha}{2} \right) = 1.28$$

Determine the sample proportions.

$$\hat{p}_1 = \frac{35}{104} = 0.337$$

$$\hat{p}_2 = \frac{37}{149} = 0.248$$

Determine the difference of sample proportions.

$$\hat{p}_2 - \hat{p}_1 = 0.248 - 0.337 = -0.089$$

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

$$\hat{p} = \frac{35 + 37}{104 + 149} = 0.285$$

Determine the standard error.

$$\begin{aligned} SE &= \sqrt{\frac{\hat{p}(1 - \hat{p})}{n_1} + \frac{\hat{p}(1 - \hat{p})}{n_2}} \\ &= \sqrt{\frac{(0.285)(0.715)}{104} + \frac{(0.285)(0.715)}{149}} \\ &= 0.0577 \end{aligned}$$

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

$$H_0 : \hat{P}_2 - \hat{P}_1 \sim \mathcal{N}(0, 0.0577)$$

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.

$$\begin{aligned} z &= \frac{(\hat{p}_2 - \hat{p}_1) - (p_2 - p_1)_0}{SE} \\ &= \frac{(0.248 - 0.337) - 0}{0.0577} \\ &= -1.54 \end{aligned}$$

Determine the  $p$ -value.

$$\begin{aligned} p\text{-value} &= 2 \cdot \Phi(-|z|) \\ &= 2 \cdot \Phi(-1.54) \\ &= 0.1236 \end{aligned}$$

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)  $z^* = 1.28$
- (d)  $SE = 0.0577$
- (e)  $|Z_{\text{obs}}| = 1.54$
- (f)  $p\text{-value} = 0.1236$
- (g) reject the null

# 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**

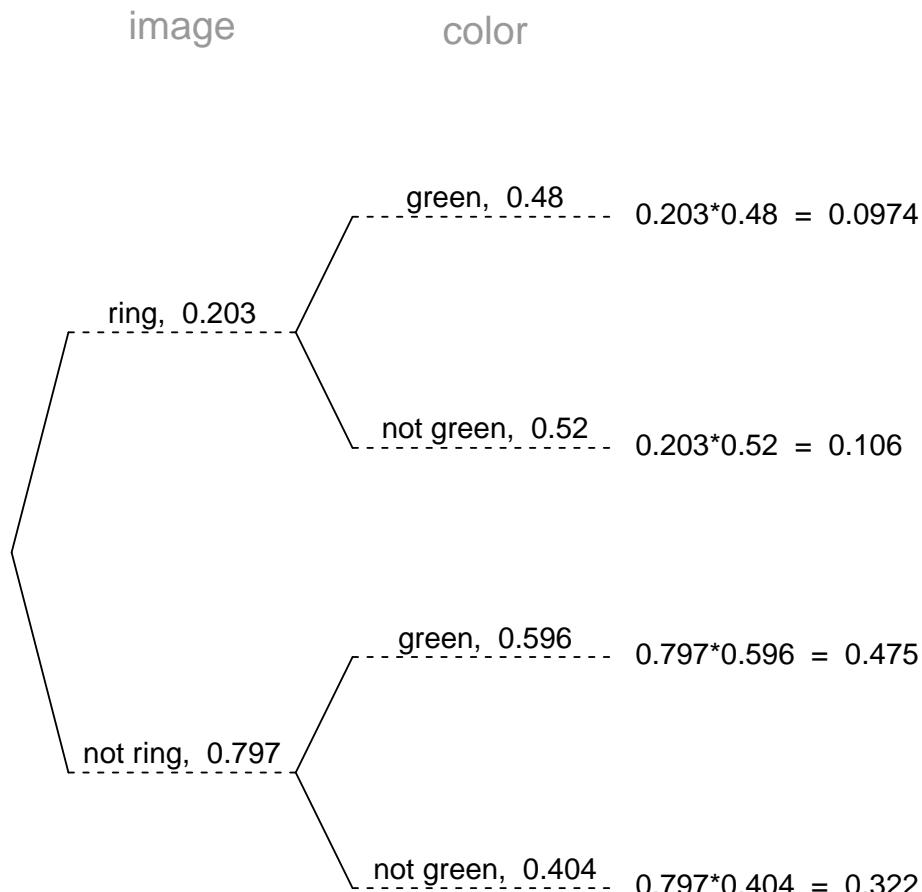
- (a)  $P(\text{wheel and teal}) = \frac{46}{375} = 0.123$
- (b)  $P(\text{violet given pig}) = \frac{26}{10+35+26} = 0.366$
- (c)  $P(\text{pig given black}) = \frac{10}{10+32+25} = 0.149$
- (d)  $P(\text{wheel}) = \frac{25+46+68}{375} = 0.371$
- (e)  $P(\text{black}) = \frac{10+32+25}{375} = 0.179$
- (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.

$$\mu = 66.17$$

$$\sigma = 1.17$$

$$n = 175$$

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

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

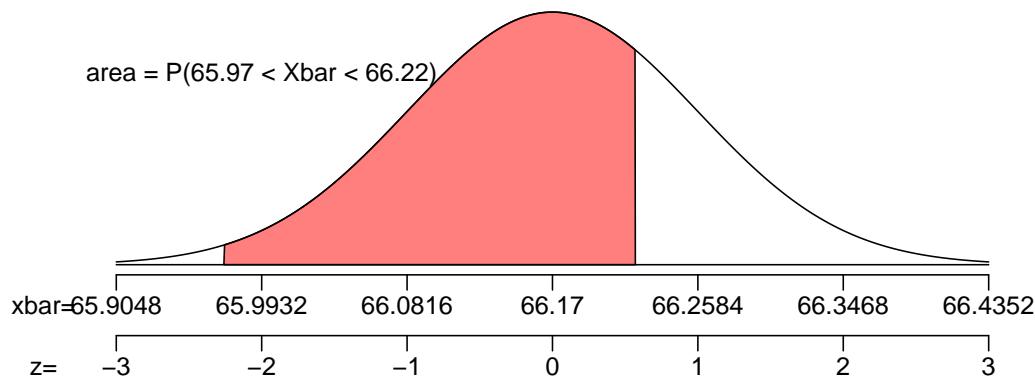
Find the standard error.

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

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.

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

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

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

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

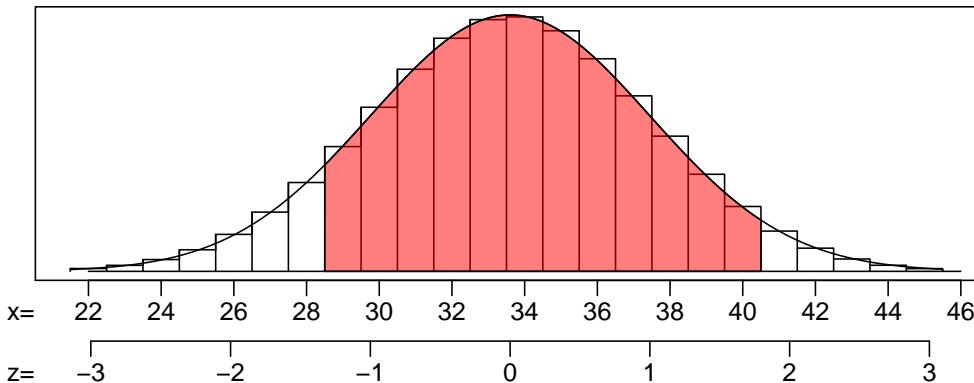
Find the mean.

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

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 condifence 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

# Bunker Hill Community College

Final Statistics Exam 2019-05-02

Exam ID 026

**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)  $P(\text{flower or indigo}) = 0.44$

(b)  $P(\text{indigo given flower}) = 0.278$

(c)  $P(\text{flower}) = 0.303$

(d)  $P(\text{flower and gray}) = 0.0266$

(e)  $P(\text{pig given green}) = 0.165$

(f)  $P(\text{green}) = 0.147$

2.  $P(\text{"not horn" given "pink"}) = 0.83$

3.  $P(62.06 < X < 62.47) = 0.8183$

4. (a)  $P(X = 68) = 0.0508$

(b)  $P(61 \leq X \leq 82) = 0.8702$

5.  $(9.44, 10.4)$

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

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

(c)  $t^* = 2.74$

(d)  $SE = 0.054$

(e)  $|t_{\text{obs}}| = 2.84$

(f)  $0.005 < p\text{-value} < 0.01$

(g) **reject**

7. (a) **LB of p CI = 0.677 or 67.7%**

(b) **UB of p CI = 0.699 or 69.9%**

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

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

(c)  $Z^* = 2.33$

(d)  $SE = 0.083$

(e)  $|Z_{\text{obs}}| = 2.47$

(f)  $p\text{-value} = 0.0136$

(g) **reject**

1. In a deck of strange cards, there are 902 cards. Each card has an image and a color. The amounts are shown in the table below.

|        | blue | gray | green | indigo | teal |
|--------|------|------|-------|--------|------|
| flower | 65   | 24   | 48    | 76     | 60   |
| pig    | 44   | 54   | 22    | 37     | 92   |
| shovel | 90   | 98   | 63    | 87     | 42   |

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

**Solution**

(a)  $P(\text{flower or indigo}) = \frac{65+24+48+76+60+76+37+87-76}{902} = 0.44$

(b)  $P(\text{indigo given flower}) = \frac{76}{65+24+48+76+60} = 0.278$

(c)  $P(\text{flower}) = \frac{65+24+48+76+60}{902} = 0.303$

(d)  $P(\text{flower and gray}) = \frac{24}{902} = 0.0266$

(e)  $P(\text{pig given green}) = \frac{22}{48+22+63} = 0.165$

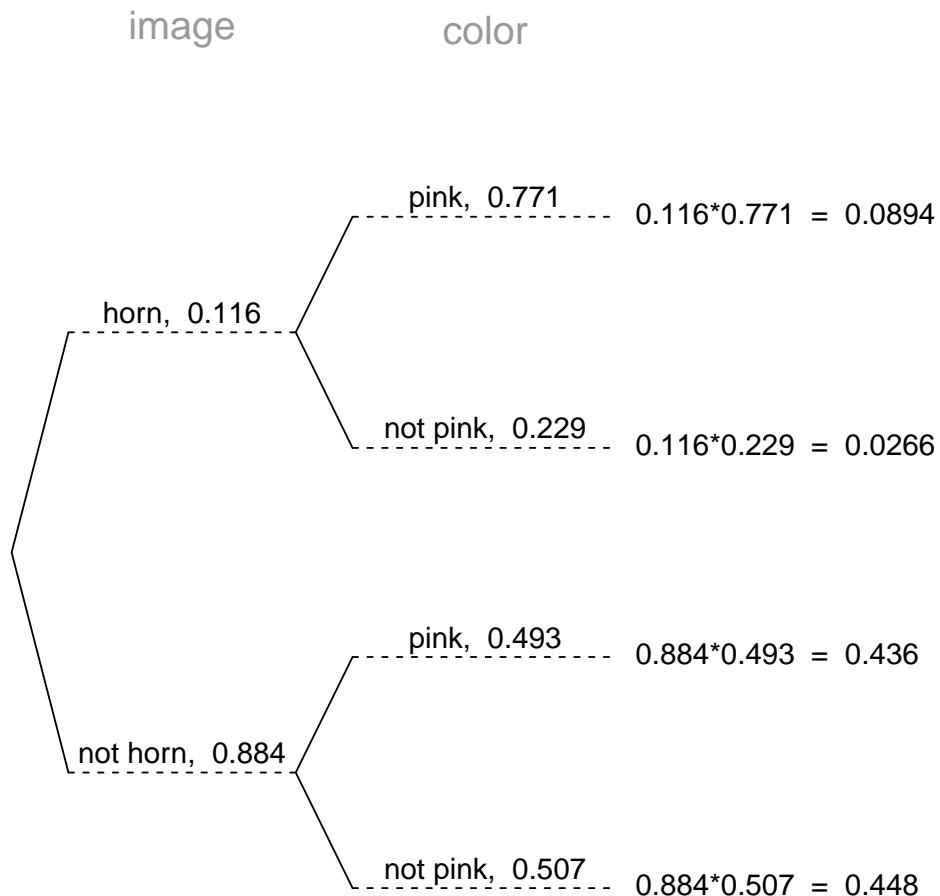
(f)  $P(\text{green}) = \frac{48+22+63}{902} = 0.147$

2. In a deck of strange cards, each card has an image and a color. The chance of drawing a horn is 11.6%. If a horn is drawn, there is a 77.1% chance that it is pink. If a card that is not a horn is drawn, there is a 49.3% chance that it is pink.

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

**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.

$$P(\text{"not horn" given "pink"}) = \frac{0.436}{0.436 + 0.0894} = 0.83$$

3. In a very large pile of toothpicks, the mean length is 62.17 millimeters and the standard deviation is 1.32 millimeters. If you randomly sample 125 toothpicks, what is the chance the sample mean is between 62.06 and 62.47 millimeters?

**Solution**

Label the given information.

$$\mu = 62.17$$

$$\sigma = 1.32$$

$$n = 125$$

$$\bar{x}_{\text{lower}} = 62.06$$

$$\bar{x}_{\text{upper}} = 62.47$$

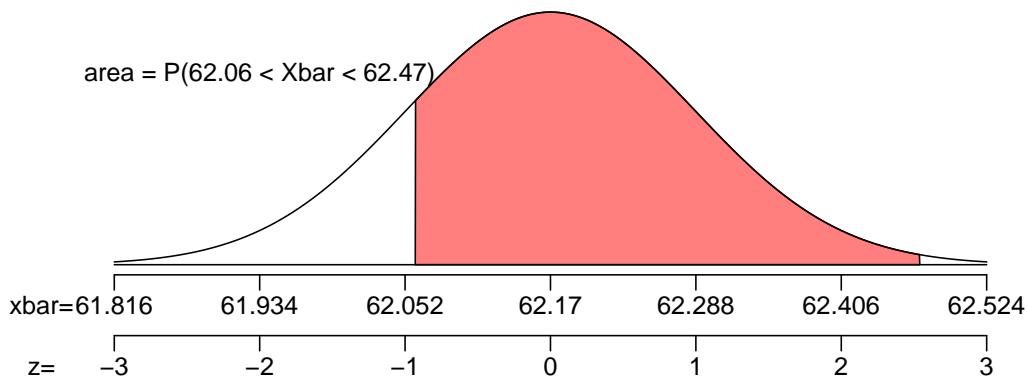
Find the standard error.

$$SE = \frac{\sigma}{\sqrt{n}} = \frac{1.32}{\sqrt{125}} = 0.118$$

Describe the sampling distribution.

$$\bar{X} \sim \mathcal{N}(62.17, 0.118)$$

Draw a sketch.



Calculate a  $z$  scores.

$$z_{\text{lower}} = \frac{x_{\text{lower}} - \mu}{SE} = \frac{62.06 - 62.17}{0.118} = -0.93$$

$$z_{\text{upper}} = \frac{x_{\text{upper}} - \mu}{SE} = \frac{62.47 - 62.17}{0.118} = 2.54$$

Determine the probability.

$$\begin{aligned} P(62.06 < X < 62.47) &= \Phi(z_{\text{upper}}) - \Phi(z_{\text{lower}}) \\ &= \Phi(2.54) - \Phi(-0.93) \\ &= 0.8183 \end{aligned}$$

4. In a game, there is a 33% chance to win a round. You will play 217 rounds.
- (a) What is the probability of winning exactly 68 rounds?
  - (b) What is the probability of winning at least 61 but at most 82 rounds?

**Solution**

We use the formula for binomial probabilities.

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

$$P(X = 68) = \binom{217}{68} (0.33)^{68} (1 - 0.33)^{217-68}$$

$$P(X = 68) = \binom{217}{68} (0.33)^{68} (0.67)^{149}$$

$$P(X = 68) = 0.0508$$

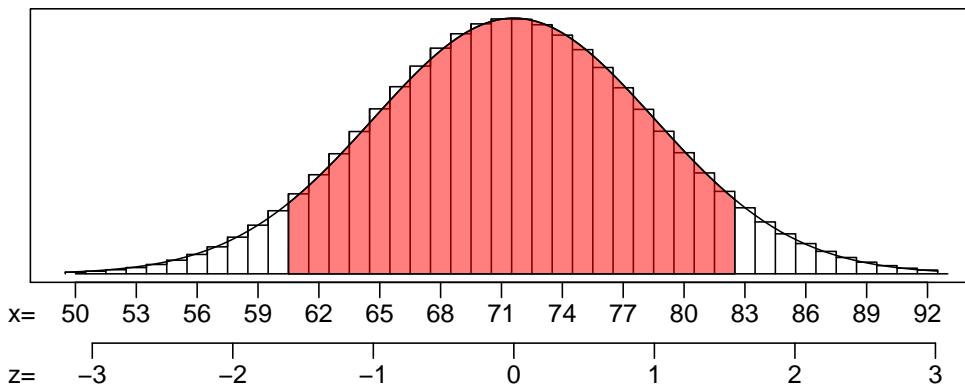
Find the mean.

$$\mu = np = (217)(0.33) = 71.61$$

Find the standard deviation.

$$\sigma = \sqrt{np(1 - p)} = \sqrt{(217)(0.33)(1 - 0.33)} = 6.9267$$

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



Find the z scores.

$$z_1 = \frac{60.5 - 71.61}{6.9267} = -1.53$$

$$z_2 = \frac{82.5 - 71.61}{6.9267} = 1.5$$

Calculate the probability.

$$P(61 \leq X \leq 82) = \Phi(1.5) - \Phi(-1.53) = 0.8702$$

(a)  $P(X = 68) = 0.0508$

(b)  $P(61 \leq X \leq 82) = 0.8702$

5. As an ornithologist, you wish to determine the average body mass of *Cistothorus palustris*. You randomly sample 26 adults of *Cistothorus palustris*, resulting in a sample mean of 9.9 grams and a sample standard deviation of 1.14 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.

$$n = 26$$

$$\bar{x} = 9.9$$

$$s = 1.14$$

$$CL = 0.95$$

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

$$df = n - 1 = 25$$

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

$$t^* = 2.06$$

Calculate the standard error.

$$SE = \frac{s}{\sqrt{n}} = \frac{1.14}{\sqrt{26}} = 0.224$$

We want to make an inference about the population mean.

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

Determine the bounds.

$$\begin{aligned} CI &= (\bar{x} - t^* SE, \bar{x} + t^* SE) \\ &= (9.9 - 2.06 \times 0.224, 9.9 + 2.06 \times 0.224) \\ &= (9.44, 10.4) \end{aligned}$$

We are 95% confident that the population mean is between 9.44 and 10.4.

6. A treatment group of size 19 has a mean of 1.1 and standard deviation of 0.141. A control group of size 18 has a mean of 0.947 and standard deviation of 0.183. If you decided to use a significance level of 0.01, 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 31.

- (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.

$$\begin{aligned}n_1 &= 19 \\ \bar{x}_1 &= 1.1 \\ s_1 &= 0.141 \\ n_2 &= 18 \\ \bar{x}_2 &= 0.947 \\ s_2 &= 0.183 \\ \alpha &= 0.01 \\ df &= 31\end{aligned}$$

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  $P(|T| > t^*) = 0.01$  by using a  $t$  table.

$$t^* = 2.74$$

Calculate the standard error.

$$\begin{aligned}SE &= \sqrt{\frac{(s_1)^2}{n_1} + \frac{(s_2)^2}{n_2}} \\ &= \sqrt{\frac{(0.141)^2}{19} + \frac{(0.183)^2}{18}} \\ &= 0.054\end{aligned}$$

Determine the test statistic.

$$\begin{aligned}t_{\text{obs}} &= \frac{(\bar{x}_2 - \bar{x}_1) - (\mu_2 - \mu_1)_0}{SE} \\ &= \frac{(0.947 - 1.1) - (0)}{0.054} \\ &= -2.84\end{aligned}$$

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

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

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

$$0.005 < p\text{-value} < 0.01$$

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)  $t^* = 2.74$
- (d)  $SE = 0.054$
- (e)  $|t_{\text{obs}}| = 2.84$
- (f)  $0.005 < p\text{-value} < 0.01$
- (g) reject the null

7. From a very large population, a random sample of 6700 individuals was taken. In that sample, 68.8% were blue. 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 condifence interval.

**Solution**

Determine  $z^*$  such that  $P(|Z| < z^*) = 0.95$ .

$$z^* = 1.96$$

Calculate the standard error.

$$SE = \sqrt{\frac{\hat{p}(1 - \hat{p})}{n}} = \sqrt{\frac{(0.688)(1 - 0.688)}{6700}} = 0.00566$$

Calculate the margin of error.

$$ME = z^* SE = (1.96)(0.00566) = 0.0111$$

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.

$$(0.677, 0.699)$$

We are 95% confident that the true population proportion is between 67.7% and 69.9%.

- (a) The lower bound = 0.677, which can also be expressed as 67.7%.
- (b) The upper bound = 0.699, which can also be expressed as 69.9%.

8. An experiment is run with a treatment group of size 24 and a control group of size 68. The results are summarized in the table below.

|             | treatment | control |
|-------------|-----------|---------|
| organic     | 7         | 6       |
| not organic | 17        | 62      |

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

- (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.

$$\hat{p}_1 = \frac{7}{24} = 0.292$$

$$\hat{p}_2 = \frac{6}{68} = 0.0882$$

Determine the difference of sample proportions.

$$\hat{p}_2 - \hat{p}_1 = 0.0882 - 0.292 = -0.2038$$

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

$$\hat{p} = \frac{7 + 6}{24 + 68} = 0.141$$

Determine the standard error.

$$\begin{aligned} SE &= \sqrt{\frac{\hat{p}(1 - \hat{p})}{n_1} + \frac{\hat{p}(1 - \hat{p})}{n_2}} \\ &= \sqrt{\frac{(0.141)(0.859)}{24} + \frac{(0.141)(0.859)}{68}} \\ &= 0.0826 \end{aligned}$$

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

$$H_0 : \hat{P}_2 - \hat{P}_1 \sim \mathcal{N}(0, 0.0826)$$

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.

$$\begin{aligned} z &= \frac{(\hat{p}_2 - \hat{p}_1) - (p_2 - p_1)_0}{SE} \\ &= \frac{(0.0882 - 0.292) - 0}{0.0826} \\ &= -2.47 \end{aligned}$$

Determine the  $p$ -value.

$$\begin{aligned} p\text{-value} &= 2 \cdot \Phi(-|z|) \\ &= 2 \cdot \Phi(-2.47) \\ &= 0.0136 \end{aligned}$$

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)  $SE = 0.0826$
- (e)  $|z_{\text{obs}}| = 2.47$
- (f)  $p\text{-value} = 0.0136$
- (g) reject the null

# Bunker Hill Community College

Final Statistics Exam 2019-05-02

Exam ID 027

**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)  $P(\text{flower}) = 0.318$

(b)  $P(\text{teal given shovel}) = 0.329$

(c)  $P(\text{teal}) = 0.373$

(d)  $P(\text{dog or red}) = 0.408$

(e)  $P(\text{shovel given red}) = 0.517$

(f)  $P(\text{shovel and teal}) = 0.147$

2.  $P(\text{"not kite" given "orange"}) = 0.577$

3.  $P(67.5 < X < 67.67) = 0.7121$

4. (a)  $P(X = 73) = 0.0704$

(b)  $P(73 \leq X \leq 81) = 0.2949$

5.  $(14.3, 17.7)$

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

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

(c)  $t^* = 2.02$

(d)  $SE = 0.029$

(e)  $|t_{\text{obs}}| = 2.08$

(f)  $0.04 < p\text{-value} < 0.05$

(g) **reject**

7. (a) **LB of p CI = 0.629 or 62.9%**

(b) **UB of p CI = 0.653 or 65.3%**

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

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

(c)  $Z^* = 2.05$

(d)  $SE = 0.095$

(e)  $|Z_{\text{obs}}| = 2.33$

(f)  $p\text{-value} = 0.0198$

(g) **reject**

1. In a deck of strange cards, there are 566 cards. Each card has an image and a color. The amounts are shown in the table below.

|        | gray | red | teal |
|--------|------|-----|------|
| dog    | 44   | 54  | 36   |
| flower | 69   | 19  | 92   |
| shovel | 91   | 78  | 83   |

- (a) What is the probability a random card is a flower?
- (b) What is the probability a random card is teal given it is a shovel?
- (c) What is the probability a random card is teal?
- (d) What is the probability a random card is either a dog or red (or both)?
- (e) What is the probability a random card is a shovel given it is red?
- (f) What is the probability a random card is both a shovel and teal?

**Solution**

(a)  $P(\text{flower}) = \frac{69+19+92}{566} = 0.318$

(b)  $P(\text{teal given shovel}) = \frac{83}{91+78+83} = 0.329$

(c)  $P(\text{teal}) = \frac{36+92+83}{566} = 0.373$

(d)  $P(\text{dog or red}) = \frac{44+54+36+54+19+78-54}{566} = 0.408$

(e)  $P(\text{shovel given red}) = \frac{78}{54+19+78} = 0.517$

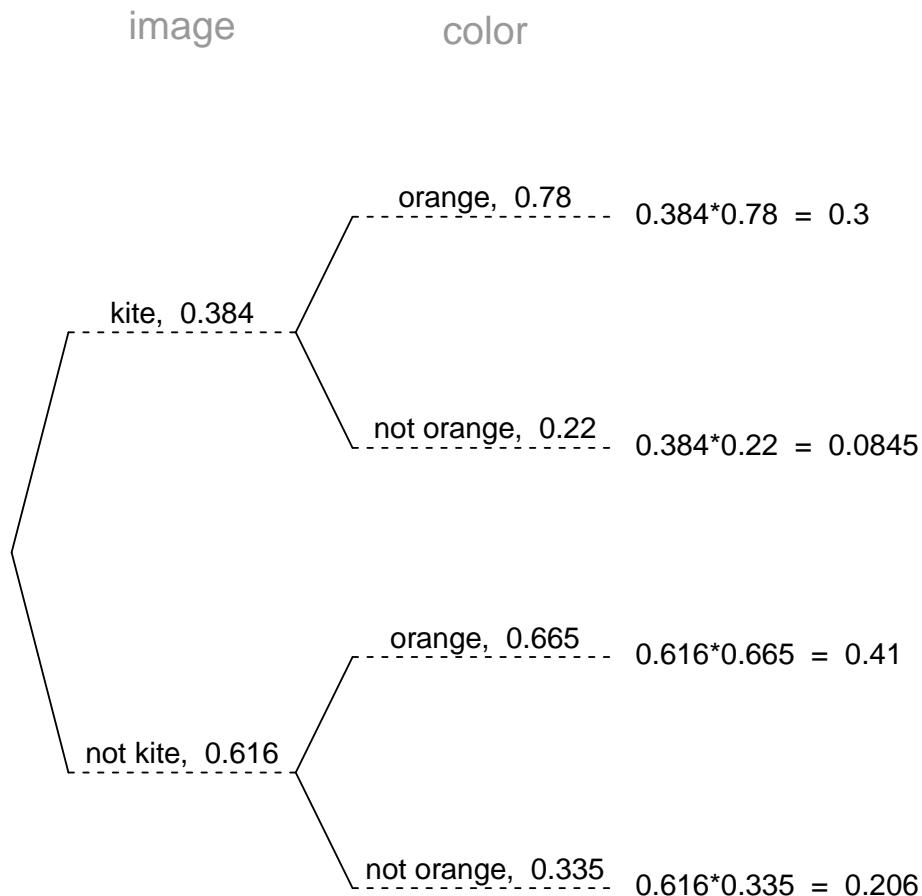
(f)  $P(\text{shovel and teal}) = \frac{83}{566} = 0.147$

2. In a deck of strange cards, each card has an image and a color. The chance of drawing a kite is 38.4%. If a kite is drawn, there is a 78% chance that it is orange. If a card that is not a kite is drawn, there is a 66.5% chance that it is orange.

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

**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.

$$P(\text{"not kite" given "orange"}) = \frac{0.41}{0.41 + 0.3} = 0.577$$

3. In a very large pile of toothpicks, the mean length is 67.58 millimeters and the standard deviation is 1.2 millimeters. If you randomly sample 225 toothpicks, what is the chance the sample mean is between 67.5 and 67.67 millimeters?

**Solution**

Label the given information.

$$\mu = 67.58$$

$$\sigma = 1.2$$

$$n = 225$$

$$\bar{x}_{\text{lower}} = 67.5$$

$$\bar{x}_{\text{upper}} = 67.67$$

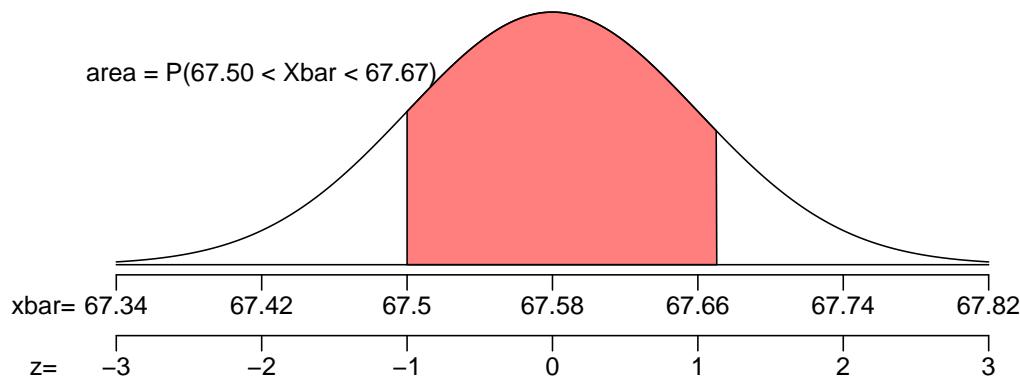
Find the standard error.

$$SE = \frac{\sigma}{\sqrt{n}} = \frac{1.2}{\sqrt{225}} = 0.08$$

Describe the sampling distribution.

$$\bar{X} \sim \mathcal{N}(67.58, 0.08)$$

Draw a sketch.



Calculate a  $z$  scores.

$$z_{\text{lower}} = \frac{x_{\text{lower}} - \mu}{SE} = \frac{67.5 - 67.58}{0.08} = -1$$

$$z_{\text{upper}} = \frac{x_{\text{upper}} - \mu}{SE} = \frac{67.67 - 67.58}{0.08} = 1.13$$

Determine the probability.

$$\begin{aligned} P(67.5 < \bar{X} < 67.67) &= \Phi(z_{\text{upper}}) - \Phi(z_{\text{lower}}) \\ &= \Phi(1.13) - \Phi(-1) \\ &= 0.7121 \end{aligned}$$

4. In a game, there is a 63% chance to win a round. You will play 112 rounds.
- (a) What is the probability of winning exactly 73 rounds?
  - (b) What is the probability of winning at least 73 but at most 81 rounds?

**Solution**

We use the formula for binomial probabilities.

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

$$P(X = 73) = \binom{112}{73} (0.63)^{73} (1 - 0.63)^{112-73}$$

$$P(X = 73) = \binom{112}{73} (0.63)^{73} (0.37)^{39}$$

$$P(X = 73) = 0.0704$$

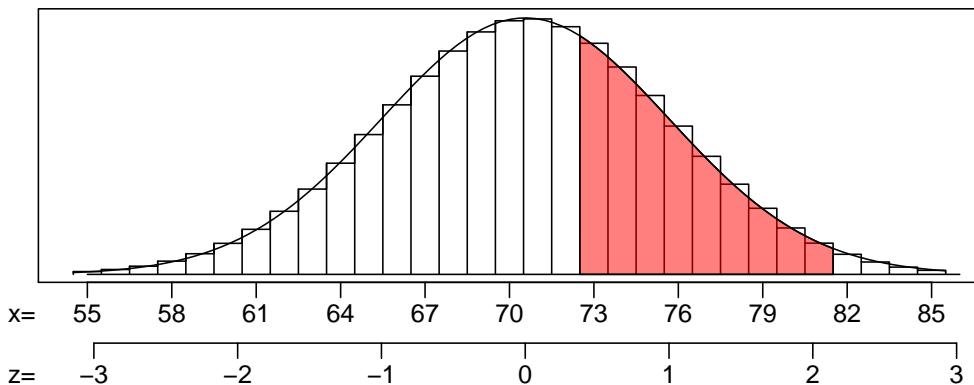
Find the mean.

$$\mu = np = (112)(0.63) = 70.56$$

Find the standard deviation.

$$\sigma = \sqrt{np(1 - p)} = \sqrt{(112)(0.63)(1 - 0.63)} = 5.1095$$

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



Find the z scores.

$$z_1 = \frac{72.5 - 70.56}{5.1095} = 0.48$$

$$z_2 = \frac{81.5 - 70.56}{5.1095} = 2.04$$

Calculate the probability.

$$P(73 \leq X \leq 81) = \Phi(2.04) - \Phi(0.48) = 0.2949$$

(a)  $P(X = 73) = 0.0704$

(b)  $P(73 \leq X \leq 81) = 0.2949$

5. As an ornithologist, you wish to determine the average body mass of *Helmitheros vermivorus*. You randomly sample 23 adults of *Helmitheros vermivorus*, resulting in a sample mean of 16.01 grams and a sample standard deviation of 2.66 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.

$$n = 23$$

$$\bar{x} = 16.01$$

$$s = 2.66$$

$$CL = 0.995$$

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

$$df = n - 1 = 22$$

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

$$t^* = 3.12$$

Calculate the standard error.

$$SE = \frac{s}{\sqrt{n}} = \frac{2.66}{\sqrt{23}} = 0.555$$

We want to make an inference about the population mean.

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

Determine the bounds.

$$\begin{aligned} CI &= (\bar{x} - t^* SE, \bar{x} + t^* SE) \\ &= (16.01 - 3.12 \times 0.555, 16.01 + 3.12 \times 0.555) \\ &= (14.3, 17.7) \end{aligned}$$

We are 99.5% confident that the population mean is between 14.3 and 17.7.

6. A treatment group of size 28 has a mean of 1.06 and standard deviation of 0.108. A control group of size 18 has a mean of 1 and standard deviation of 0.0868. If you decided to use a significance level of 0.05, 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 41.

- (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.

$$n_1 = 28$$

$$\bar{x}_1 = 1.06$$

$$s_1 = 0.108$$

$$n_2 = 18$$

$$\bar{x}_2 = 1$$

$$s_2 = 0.0868$$

$$\alpha = 0.05$$

$$df = 41$$

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  $P(|T| > t^*) = 0.05$  by using a  $t$  table.

$$t^* = 2.02$$

Calculate the standard error.

$$\begin{aligned} SE &= \sqrt{\frac{(s_1)^2}{n_1} + \frac{(s_2)^2}{n_2}} \\ &= \sqrt{\frac{(0.108)^2}{28} + \frac{(0.0868)^2}{18}} \\ &= 0.029 \end{aligned}$$

Determine the test statistic.

$$\begin{aligned} t_{\text{obs}} &= \frac{(\bar{x}_2 - \bar{x}_1) - (\mu_2 - \mu_1)_0}{SE} \\ &= \frac{(1 - 1.06) - (0)}{0.029} \\ &= -2.08 \end{aligned}$$

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

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

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

$$0.04 < p\text{-value} < 0.05$$

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)  $t^* = 2.02$
- (d)  $SE = 0.029$
- (e)  $|t_{\text{obs}}| = 2.08$
- (f)  $0.04 < p\text{-value} < 0.05$
- (g) reject the null

7. From a very large population, a random sample of 8100 individuals was taken. In that sample, 64.1% were angry. Determine a 98% confidence interval of the population proportion.

- (a) Find the lower bound of the confidence interval.
- (b) Find the upper bound of the condifence interval.

**Solution**

Determine  $z^*$  such that  $P(|Z| < z^*) = 0.98$ .

$$z^* = 2.33$$

Calculate the standard error.

$$SE = \sqrt{\frac{\hat{p}(1 - \hat{p})}{n}} = \sqrt{\frac{(0.641)(1 - 0.641)}{8100}} = 0.00533$$

Calculate the margin of error.

$$ME = z^* SE = (2.33)(0.00533) = 0.0124$$

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.

$$(0.629, 0.653)$$

We are 98% confident that the true population proportion is between 62.9% and 65.3%.

- (a) The lower bound = 0.629, which can also be expressed as 62.9%.
- (b) The upper bound = 0.653, which can also be expressed as 65.3%.

8. An experiment is run with a treatment group of size 72 and a control group of size 32. The results are summarized in the table below.

|                | treatment | control |
|----------------|-----------|---------|
| omnivorous     | 47        | 28      |
| not omnivorous | 25        | 4       |

Using a significance level of 0.04, determine whether the treatment causes an effect on the proportion of cases that are omnivorous.

- (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  $P(|Z| > z^*) = 0.04$ .

$$z^* = \Phi^{-1} \left( 1 - \frac{\alpha}{2} \right) = 2.05$$

Determine the sample proportions.

$$\hat{p}_1 = \frac{47}{72} = 0.653$$

$$\hat{p}_2 = \frac{28}{32} = 0.875$$

Determine the difference of sample proportions.

$$\hat{p}_2 - \hat{p}_1 = 0.875 - 0.653 = 0.222$$

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

$$\hat{p} = \frac{47 + 28}{72 + 32} = 0.721$$

Determine the standard error.

$$\begin{aligned} SE &= \sqrt{\frac{\hat{p}(1 - \hat{p})}{n_1} + \frac{\hat{p}(1 - \hat{p})}{n_2}} \\ &= \sqrt{\frac{(0.721)(0.279)}{72} + \frac{(0.721)(0.279)}{32}} \\ &= 0.0953 \end{aligned}$$

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

$$H_0 : \hat{P}_2 - \hat{P}_1 \sim \mathcal{N}(0, 0.0953)$$

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.

$$\begin{aligned} z &= \frac{(\hat{p}_2 - \hat{p}_1) - (p_2 - p_1)_0}{SE} \\ &= \frac{(0.875 - 0.653) - 0}{0.0953} \\ &= 2.33 \end{aligned}$$

Determine the  $p$ -value.

$$\begin{aligned} p\text{-value} &= 2 \cdot \Phi(-|z|) \\ &= 2 \cdot \Phi(-2.33) \\ &= 0.0198 \end{aligned}$$

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)  $z^* = 2.05$
- (d)  $SE = 0.0953$
- (e)  $|Z_{\text{obs}}| = 2.33$
- (f)  $p\text{-value} = 0.0198$
- (g) reject the null

# Bunker Hill Community College

Final Statistics Exam 2019-05-02

Exam ID 028

**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)  $P(\text{gem and blue}) = 0.0854$

(b)  $P(\text{orange given dog}) = 0.114$

(c)  $P(\text{dog}) = 0.183$

(d)  $P(\text{yellow}) = 0.154$

(e)  $P(\text{gem given red}) = 0.366$

(f)  $P(\text{flower or orange}) = 0.316$

2.  $P(\text{"not shovel" given "pink"}) = 0.75$

3.  $P(64.77 < X < 65.05) = 0.728$

4. (a)  $P(X = 53) = 0.0613$

(b)  $P(45 \leq X \leq 66) = 0.8542$

5.  $(15.4, 17)$

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

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

(c)  $t^* = 2.65$

(d)  $SE = 0.426$

(e)  $|t_{\text{obs}}| = 2.54$

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

(g) **retain**

7. (a) **LB of p CI = 0.27 or 27%**

(b) **UB of p CI = 0.284 or 28.4%**

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

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

(c)  $Z^* = 1.28$

(d)  $SE = 0.027$

(e)  $|Z_{\text{obs}}| = 1.36$

(f)  $p\text{-value} = 0.1738$

(g) **reject**

1. In a deck of strange cards, there are 1054 cards. Each card has an image and a color. The amounts are shown in the table below.

|        | black | blue | orange | red | yellow |
|--------|-------|------|--------|-----|--------|
| dog    | 46    | 70   | 22     | 23  | 32     |
| flower | 49    | 36   | 55     | 65  | 26     |
| gem    | 62    | 90   | 67     | 96  | 31     |
| tree   | 80    | 40   | 13     | 78  | 73     |

- (a) What is the probability a random card is both a gem and blue?
- (b) What is the probability a random card is orange given it is a dog?
- (c) What is the probability a random card is a dog?
- (d) What is the probability a random card is yellow?
- (e) What is the probability a random card is a gem given it is red?
- (f) What is the probability a random card is either a flower or orange (or both)?

**Solution**

(a)  $P(\text{gem and blue}) = \frac{90}{1054} = 0.0854$

(b)  $P(\text{orange given dog}) = \frac{22}{46+70+22+23+32} = 0.114$

(c)  $P(\text{dog}) = \frac{46+70+22+23+32}{1054} = 0.183$

(d)  $P(\text{yellow}) = \frac{32+26+31+73}{1054} = 0.154$

(e)  $P(\text{gem given red}) = \frac{96}{23+65+96+78} = 0.366$

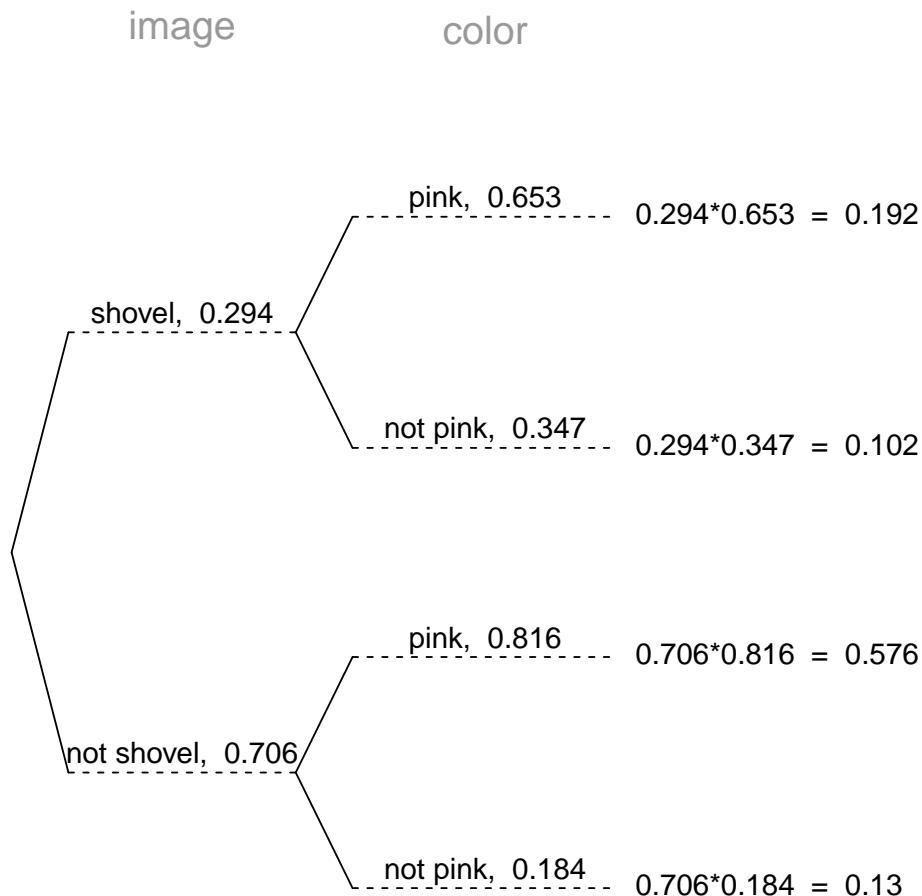
(f)  $P(\text{flower or orange}) = \frac{49+36+55+65+26+22+55+67+13-55}{1054} = 0.316$

2. In a deck of strange cards, each card has an image and a color. The chance of drawing a shovel is 29.4%. If a shovel is drawn, there is a 65.3% chance that it is pink. If a card that is not a shovel is drawn, there is a 81.6% chance that it is pink.

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

**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.

$$P(\text{"not shovel" given "pink"}) = \frac{0.576}{0.576 + 0.192} = 0.75$$

3. In a very large pile of toothpicks, the mean length is 64.94 millimeters and the standard deviation is 1.74 millimeters. If you randomly sample 196 toothpicks, what is the chance the sample mean is between 64.77 and 65.05 millimeters?

**Solution**

Label the given information.

$$\mu = 64.94$$

$$\sigma = 1.74$$

$$n = 196$$

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

$$\bar{x}_{\text{upper}} = 65.05$$

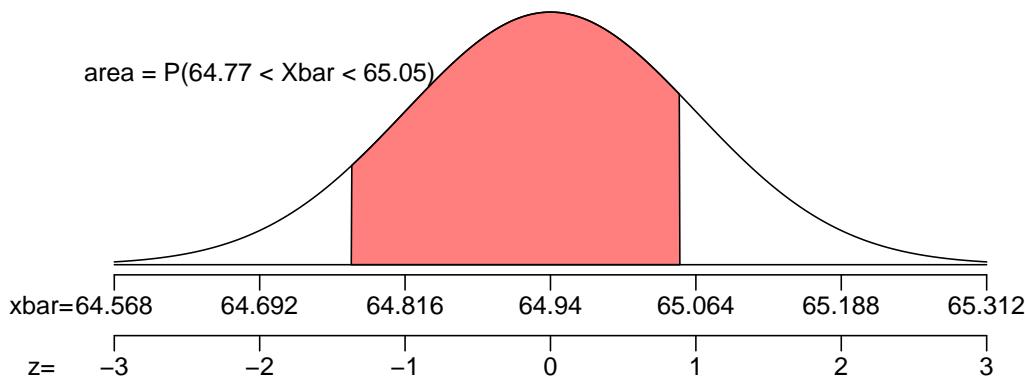
Find the standard error.

$$SE = \frac{\sigma}{\sqrt{n}} = \frac{1.74}{\sqrt{196}} = 0.124$$

Describe the sampling distribution.

$$\bar{X} \sim \mathcal{N}(64.94, 0.124)$$

Draw a sketch.



Calculate a  $z$  scores.

$$z_{\text{lower}} = \frac{\bar{x}_{\text{lower}} - \mu}{SE} = \frac{64.77 - 64.94}{0.124} = -1.37$$

$$z_{\text{upper}} = \frac{\bar{x}_{\text{upper}} - \mu}{SE} = \frac{65.05 - 64.94}{0.124} = 0.89$$

Determine the probability.

$$\begin{aligned} P(64.77 < \bar{X} < 65.05) &= \Phi(z_{\text{upper}}) - \Phi(z_{\text{lower}}) \\ &= \Phi(0.89) - \Phi(-1.37) \\ &= 0.728 \end{aligned}$$

4. In a game, there is a 21% chance to win a round. You will play 249 rounds.
- (a) What is the probability of winning exactly 53 rounds?
  - (b) What is the probability of winning at least 45 but at most 66 rounds?

**Solution**

We use the formula for binomial probabilities.

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

$$P(X = 53) = \binom{249}{53} (0.21)^{53} (1 - 0.21)^{249-53}$$

$$P(X = 53) = \binom{249}{53} (0.21)^{53} (0.79)^{196}$$

$$P(X = 53) = 0.0613$$

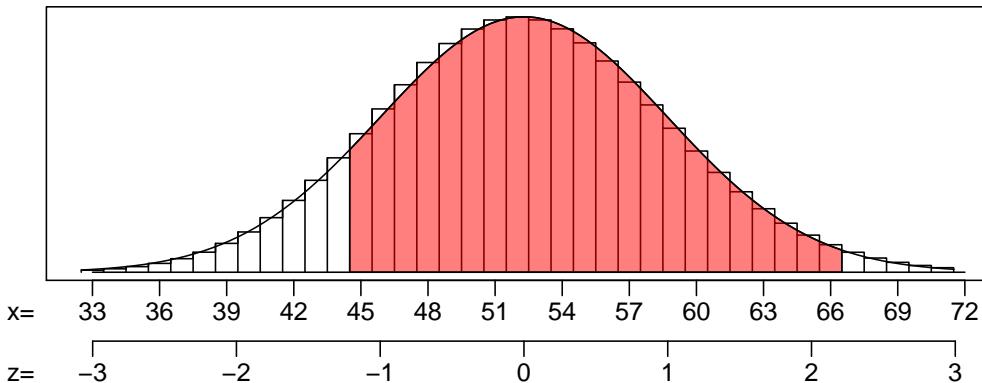
Find the mean.

$$\mu = np = (249)(0.21) = 52.29$$

Find the standard deviation.

$$\sigma = \sqrt{np(1 - p)} = \sqrt{(249)(0.21)(1 - 0.21)} = 6.4272$$

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



Find the z scores.

$$z_1 = \frac{44.5 - 52.29}{6.4272} = -1.13$$

$$z_2 = \frac{66.5 - 52.29}{6.4272} = 2.13$$

Calculate the probability.

$$P(45 \leq X \leq 66) = \Phi(2.13) - \Phi(-1.13) = 0.8542$$

(a)  $P(X = 53) = 0.0613$

(b)  $P(45 \leq X \leq 66) = 0.8542$

5. As an ornithologist, you wish to determine the average body mass of *Helmitheros vermivorus*. You randomly sample 18 adults of *Helmitheros vermivorus*, resulting in a sample mean of 16.21 grams and a sample standard deviation of 2.5 grams. Determine a 80% confidence interval of the true population mean.

**Solution**

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

$$n = 18$$

$$\bar{x} = 16.21$$

$$s = 2.5$$

$$CL = 0.8$$

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

$$df = n - 1 = 17$$

Determine the critical  $t$  value,  $t^*$ , such that  $P(|T| < t^*) = 0.8$ .

$$t^* = 1.33$$

Calculate the standard error.

$$SE = \frac{s}{\sqrt{n}} = \frac{2.5}{\sqrt{18}} = 0.589$$

We want to make an inference about the population mean.

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

Determine the bounds.

$$\begin{aligned} CI &= (\bar{x} - t^* SE, \bar{x} + t^* SE) \\ &= (16.21 - 1.33 \times 0.589, 16.21 + 1.33 \times 0.589) \\ &= (15.4, 17) \end{aligned}$$

We are 80% confident that the population mean is between 15.4 and 17.

6. A treatment group of size 40 has a mean of 11 and standard deviation of 1.89. A control group of size 30 has a mean of 9.92 and standard deviation of 1.66. If you decided to use a significance level of 0.01, 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 66.

- (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.

$$n_1 = 40$$

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

$$s_1 = 1.89$$

$$n_2 = 30$$

$$\bar{x}_2 = 9.92$$

$$s_2 = 1.66$$

$$\alpha = 0.01$$

$$df = 66$$

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  $P(|T| > t^*) = 0.01$  by using a  $t$  table.

$$t^* = 2.65$$

Calculate the standard error.

$$\begin{aligned} SE &= \sqrt{\frac{(s_1)^2}{n_1} + \frac{(s_2)^2}{n_2}} \\ &= \sqrt{\frac{(1.89)^2}{40} + \frac{(1.66)^2}{30}} \\ &= 0.426 \end{aligned}$$

Determine the test statistic.

$$\begin{aligned} t_{\text{obs}} &= \frac{(\bar{x}_2 - \bar{x}_1) - (\mu_2 - \mu_1)_0}{SE} \\ &= \frac{(9.92 - 11) - (0)}{0.426} \\ &= -2.54 \end{aligned}$$

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 retain the null hypothesis.

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

- (b)  $H_A : \mu_2 - \mu_1 \neq 0$
- (c)  $t^* = 2.65$
- (d)  $SE = 0.426$
- (e)  $|t_{\text{obs}}| = 2.54$
- (f)  $0.01 < p\text{-value} < 0.02$
- (g) retain the null

7. From a very large population, a random sample of 6400 individuals was taken. In that sample, 27.7% were tasty. 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 condifence interval.

**Solution**

Determine  $z^*$  such that  $P(|Z| < z^*) = 0.8$ .

$$z^* = 1.28$$

Calculate the standard error.

$$SE = \sqrt{\frac{\hat{p}(1 - \hat{p})}{n}} = \sqrt{\frac{(0.277)(1 - 0.277)}{6400}} = 0.00559$$

Calculate the margin of error.

$$ME = z^* SE = (1.28)(0.00559) = 0.00716$$

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.

$$(0.27, 0.284)$$

We are 80% confident that the true population proportion is between 27% and 28.4%.

- (a) The lower bound = 0.27, which can also be expressed as 27%.
- (b) The upper bound = 0.284, which can also be expressed as 28.4%.

8. An experiment is run with a treatment group of size 172 and a control group of size 157. The results are summarized in the table below.

|          | treatment | control |
|----------|-----------|---------|
| sick     | 14        | 7       |
| not sick | 158       | 150     |

Using a significance level of 0.2, determine whether the treatment causes an effect on the proportion of cases that are sick.

- (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  $P(|Z| > z^*) = 0.2$ .

$$z^* = \Phi^{-1} \left( 1 - \frac{\alpha}{2} \right) = 1.28$$

Determine the sample proportions.

$$\hat{p}_1 = \frac{14}{172} = 0.0814$$

$$\hat{p}_2 = \frac{7}{157} = 0.0446$$

Determine the difference of sample proportions.

$$\hat{p}_2 - \hat{p}_1 = 0.0446 - 0.0814 = -0.0368$$

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

$$\hat{p} = \frac{14 + 7}{172 + 157} = 0.0638$$

Determine the standard error.

$$\begin{aligned} SE &= \sqrt{\frac{\hat{p}(1 - \hat{p})}{n_1} + \frac{\hat{p}(1 - \hat{p})}{n_2}} \\ &= \sqrt{\frac{(0.0638)(0.9362)}{172} + \frac{(0.0638)(0.9362)}{157}} \\ &= 0.027 \end{aligned}$$

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

$$H_0 : \hat{P}_2 - \hat{P}_1 \sim \mathcal{N}(0, 0.027)$$

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.

$$\begin{aligned} z &= \frac{(\hat{p}_2 - \hat{p}_1) - (p_2 - p_1)_0}{SE} \\ &= \frac{(0.0446 - 0.0814) - 0}{0.027} \\ &= -1.36 \end{aligned}$$

Determine the  $p$ -value.

$$\begin{aligned} p\text{-value} &= 2 \cdot \Phi(-|z|) \\ &= 2 \cdot \Phi(-1.36) \\ &= 0.1738 \end{aligned}$$

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)  $z^* = 1.28$
- (d)  $SE = 0.027$
- (e)  $|Z_{\text{obs}}| = 1.36$
- (f)  $p\text{-value} = 0.1738$
- (g) reject the null

# Bunker Hill Community College

Final Statistics Exam 2019-05-02

Exam ID 029

**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)  $P(\text{bike or yellow}) = 0.376$

(b)  $P(\text{blue}) = 0.21$

(c)  $P(\text{pig}) = 0.201$

(d)  $P(\text{wheel and blue}) = 0.0604$

(e)  $P(\text{teal given gem}) = 0.599$

(f)  $P(\text{wheel given blue}) = 0.288$

2.  $P(\text{"kite" given "green"}) = 0.0572$

3.  $P(61.1 < X < 61.83) = 0.7739$

4. (a)  $P(X = 30) = 0.0558$

(b)  $P(24 \leq X \leq 41) = 0.9097$

5.  $(29.6, 41)$

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

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

(c)  $t^* = 2.4$

(d)  $SE = 0.062$

(e)  $|t_{\text{obs}}| = 2.1$

(f)  $0.04 < p\text{-value} < 0.05$

(g) **retain**

7. (a) **LB of p CI = 0.138 or 13.8%**

(b) **UB of p CI = 0.252 or 25.2%**

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

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

(c)  $Z^* = 1.64$

(d)  $SE = 0.09$

(e)  $|Z_{\text{obs}}| = 1.92$

(f)  $p\text{-value} = 0.0548$

(g) **reject**

1. In a deck of strange cards, there are 1043 cards. Each card has an image and a color. The amounts are shown in the table below.

|        | blue | orange | teal | yellow |
|--------|------|--------|------|--------|
| bike   | 16   | 22     | 92   | 34     |
| gem    | 28   | 17     | 82   | 10     |
| pig    | 51   | 52     | 24   | 83     |
| shovel | 61   | 38     | 60   | 39     |
| wheel  | 63   | 86     | 89   | 96     |

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

**Solution**

$$(a) P(\text{bike or yellow}) = \frac{16+22+92+34+34+10+83+39+96-34}{1043} = 0.376$$

$$(b) P(\text{blue}) = \frac{16+28+51+61+63}{1043} = 0.21$$

$$(c) P(\text{pig}) = \frac{51+52+24+83}{1043} = 0.201$$

$$(d) P(\text{wheel and blue}) = \frac{63}{1043} = 0.0604$$

$$(e) P(\text{teal given gem}) = \frac{82}{28+17+82+10} = 0.599$$

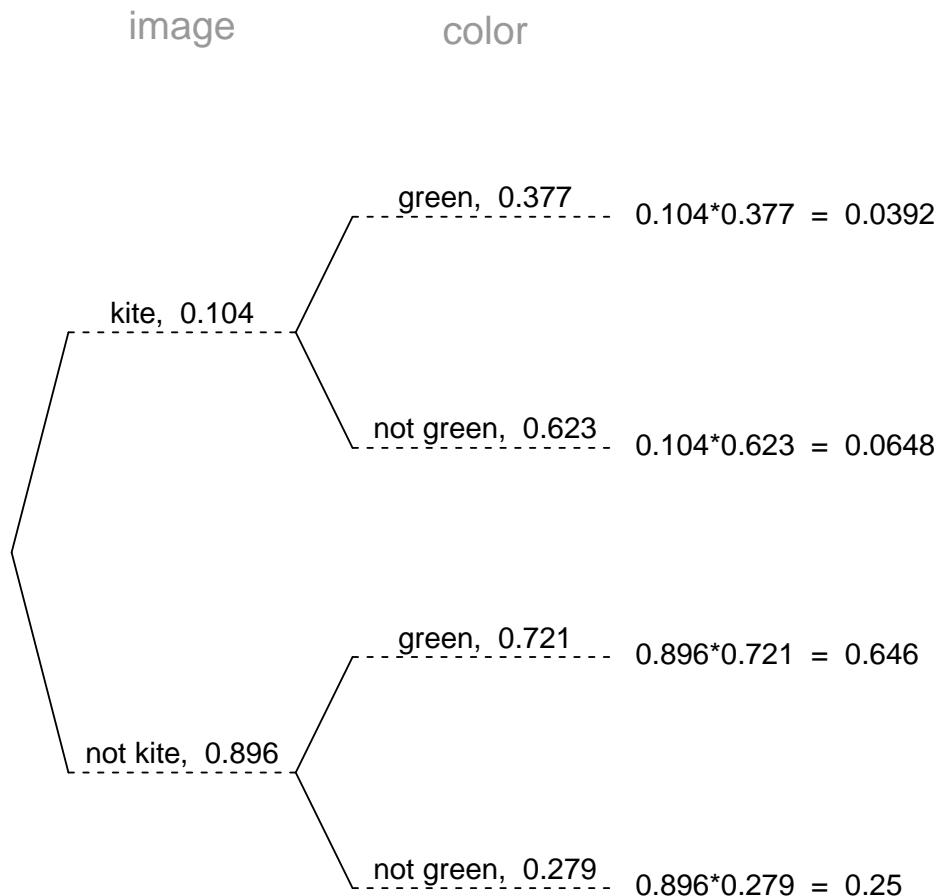
$$(f) P(\text{wheel given blue}) = \frac{63}{16+28+51+61+63} = 0.288$$

2. In a deck of strange cards, each card has an image and a color. The chance of drawing a kite is 10.4%. If a kite is drawn, there is a 37.7% chance that it is green. If a card that is not a kite is drawn, there is a 72.1% chance that it is green.

Now, someone draws a random card and reveals it is green. What is the chance the card is a kite?

**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.

$$P(\text{"kite" given "green"}) = \frac{0.0392}{0.0392 + 0.646} = 0.0572$$

3. In a very large pile of toothpicks, the mean length is 61.42 millimeters and the standard deviation is 3.94 millimeters. If you randomly sample 175 toothpicks, what is the chance the sample mean is between 61.1 and 61.83 millimeters?

**Solution**

Label the given information.

$$\mu = 61.42$$

$$\sigma = 3.94$$

$$n = 175$$

$$\bar{x}_{\text{lower}} = 61.1$$

$$\bar{x}_{\text{upper}} = 61.83$$

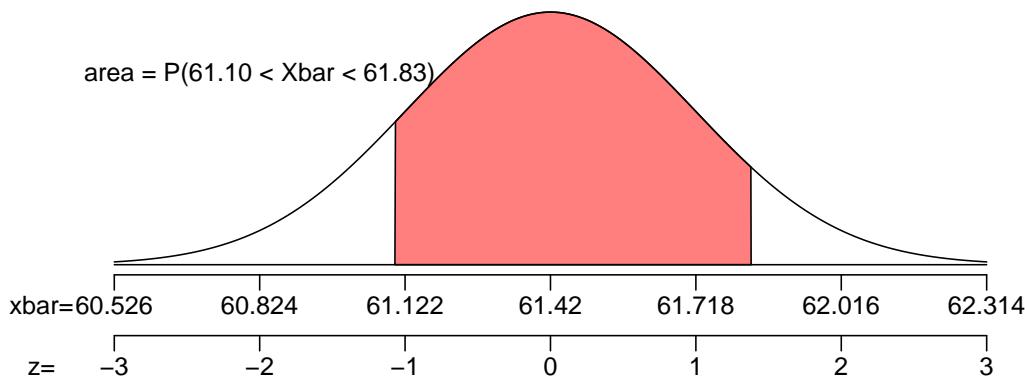
Find the standard error.

$$SE = \frac{\sigma}{\sqrt{n}} = \frac{3.94}{\sqrt{175}} = 0.298$$

Describe the sampling distribution.

$$\bar{X} \sim \mathcal{N}(61.42, 0.298)$$

Draw a sketch.



Calculate a  $z$  scores.

$$z_{\text{lower}} = \frac{\bar{x}_{\text{lower}} - \mu}{SE} = \frac{61.1 - 61.42}{0.298} = -1.07$$

$$z_{\text{upper}} = \frac{\bar{x}_{\text{upper}} - \mu}{SE} = \frac{61.83 - 61.42}{0.298} = 1.38$$

Determine the probability.

$$\begin{aligned} P(61.1 < \bar{X} < 61.83) &= \Phi(z_{\text{upper}}) - \Phi(z_{\text{lower}}) \\ &= \Phi(1.38) - \Phi(-1.07) \\ &= 0.7739 \end{aligned}$$

4. In a game, there is a 37% chance to win a round. You will play 93 rounds.
- (a) What is the probability of winning exactly 30 rounds?
  - (b) What is the probability of winning at least 24 but at most 41 rounds?

**Solution**

We use the formula for binomial probabilities.

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

$$P(X = 30) = \binom{93}{30} (0.37)^{30} (1 - 0.37)^{93-30}$$

$$P(X = 30) = \binom{93}{30} (0.37)^{30} (0.63)^{63}$$

$$P(X = 30) = 0.0558$$

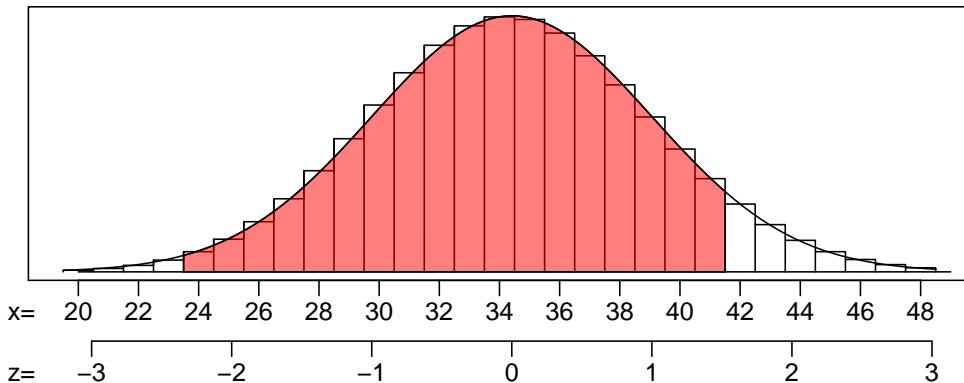
Find the mean.

$$\mu = np = (93)(0.37) = 34.41$$

Find the standard deviation.

$$\sigma = \sqrt{np(1 - p)} = \sqrt{(93)(0.37)(1 - 0.37)} = 4.656$$

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



Find the z scores.

$$z_1 = \frac{23.5 - 34.41}{4.656} = -2.24$$

$$z_2 = \frac{41.5 - 34.41}{4.656} = 1.42$$

Calculate the probability.

$$P(24 \leq X \leq 41) = \Phi(1.42) - \Phi(-2.24) = 0.9097$$

(a)  $P(X = 30) = 0.0558$

(b)  $P(24 \leq X \leq 41) = 0.9097$

5. As an ornithologist, you wish to determine the average body mass of *Dolichonyx orizivorus*. You randomly sample 20 adults of *Dolichonyx orizivorus*, resulting in a sample mean of 35.29 grams and a sample standard deviation of 8.9 grams. Determine a 99% confidence interval of the true population mean.

**Solution**

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

$$n = 20$$

$$\bar{x} = 35.29$$

$$s = 8.9$$

$$CL = 0.99$$

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

$$df = n - 1 = 19$$

Determine the critical  $t$  value,  $t^*$ , such that  $P(|T| < t^*) = 0.99$ .

$$t^* = 2.86$$

Calculate the standard error.

$$SE = \frac{s}{\sqrt{n}} = \frac{8.9}{\sqrt{20}} = 1.99$$

We want to make an inference about the population mean.

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

Determine the bounds.

$$\begin{aligned} CI &= (\bar{x} - t^* SE, \bar{x} + t^* SE) \\ &= (35.29 - 2.86 \times 1.99, 35.29 + 2.86 \times 1.99) \\ &= (29.6, 41) \end{aligned}$$

We are 99% confident that the population mean is between 29.6 and 41.

6. A treatment group of size 23 has a mean of 1.14 and standard deviation of 0.209. A control group of size 31 has a mean of 1.01 and standard deviation of 0.245. 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 50.

- (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.

$$\begin{aligned}n_1 &= 23 \\ \bar{x}_1 &= 1.14 \\ s_1 &= 0.209 \\ n_2 &= 31 \\ \bar{x}_2 &= 1.01 \\ s_2 &= 0.245 \\ \alpha &= 0.02 \\ df &= 50\end{aligned}$$

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.

$$t^* = 2.4$$

Calculate the standard error.

$$\begin{aligned}SE &= \sqrt{\frac{(s_1)^2}{n_1} + \frac{(s_2)^2}{n_2}} \\ &= \sqrt{\frac{(0.209)^2}{23} + \frac{(0.245)^2}{31}} \\ &= 0.062\end{aligned}$$

Determine the test statistic.

$$\begin{aligned}t_{\text{obs}} &= \frac{(\bar{x}_2 - \bar{x}_1) - (\mu_2 - \mu_1)_0}{SE} \\ &= \frac{(1.01 - 1.14) - (0)}{0.062} \\ &= -2.1\end{aligned}$$

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

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

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

$$0.04 < p\text{-value} < 0.05$$

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)  $t^* = 2.4$
- (d)  $SE = 0.062$
- (e)  $|t_{\text{obs}}| = 2.1$
- (f)  $0.04 < p\text{-value} < 0.05$
- (g) retain the null

7. From a very large population, a random sample of 380 individuals was taken. In that sample, 19.5% were purple. Determine a 99.5% confidence interval of the population proportion.

- (a) Find the lower bound of the confidence interval.
- (b) Find the upper bound of the condifence interval.

**Solution**

Determine  $z^*$  such that  $P(|Z| < z^*) = 0.995$ .

$$z^* = 2.81$$

Calculate the standard error.

$$SE = \sqrt{\frac{\hat{p}(1 - \hat{p})}{n}} = \sqrt{\frac{(0.195)(1 - 0.195)}{380}} = 0.0203$$

Calculate the margin of error.

$$ME = z^* SE = (2.81)(0.0203) = 0.057$$

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.

$$(0.138, 0.252)$$

We are 99.5% confident that the true population proportion is between 13.8% and 25.2%.

- (a) The lower bound = 0.138, which can also be expressed as 13.8%.
- (b) The upper bound = 0.252, which can also be expressed as 25.2%.

8. An experiment is run with a treatment group of size 72 and a control group of size 52. The results are summarized in the table below.

|             | treatment | control |
|-------------|-----------|---------|
| organic     | 36        | 35      |
| not organic | 36        | 17      |

Using a significance level of 0.1, determine whether the treatment causes an effect on the proportion of cases that are organic.

- (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  $P(|Z| > z^*) = 0.1$ .

$$z^* = \Phi^{-1} \left( 1 - \frac{\alpha}{2} \right) = 1.64$$

Determine the sample proportions.

$$\hat{p}_1 = \frac{36}{72} = 0.5$$

$$\hat{p}_2 = \frac{35}{52} = 0.673$$

Determine the difference of sample proportions.

$$\hat{p}_2 - \hat{p}_1 = 0.673 - 0.5 = 0.173$$

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

$$\hat{p} = \frac{36 + 35}{72 + 52} = 0.573$$

Determine the standard error.

$$\begin{aligned} SE &= \sqrt{\frac{\hat{p}(1 - \hat{p})}{n_1} + \frac{\hat{p}(1 - \hat{p})}{n_2}} \\ &= \sqrt{\frac{(0.573)(0.427)}{72} + \frac{(0.573)(0.427)}{52}} \\ &= 0.09 \end{aligned}$$

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

$$H_0 : \hat{P}_2 - \hat{P}_1 \sim \mathcal{N}(0, 0.09)$$

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.

$$\begin{aligned} z &= \frac{(\hat{p}_2 - \hat{p}_1) - (p_2 - p_1)_0}{SE} \\ &= \frac{(0.673 - 0.5) - 0}{0.09} \\ &= 1.92 \end{aligned}$$

Determine the  $p$ -value.

$$\begin{aligned} p\text{-value} &= 2 \cdot \Phi(-|z|) \\ &= 2 \cdot \Phi(-1.92) \\ &= 0.0548 \end{aligned}$$

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)  $z^* = 1.64$
- (d)  $SE = 0.09$
- (e)  $|z_{\text{obs}}| = 1.92$
- (f)  $p\text{-value} = 0.0548$
- (g) reject the null

# Bunker Hill Community College

Final Statistics Exam 2019-05-02

Exam ID 030

**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)  $P(\text{green given wheel}) = 0.179$

(b)  $P(\text{flower or teal}) = 0.43$

(c)  $P(\text{pink}) = 0.23$

(d)  $P(\text{tree}) = 0.289$

(e)  $P(\text{shovel given pink}) = 0.351$

(f)  $P(\text{tree and yellow}) = 0.091$

2.  $P(\text{"tree" given "not pink"}) = 0.427$

3.  $P(74.6 < X < 75) = 0.8912$

4. (a)  $P(X = 44) = 0.0581$

(b)  $P(33 \leq X \leq 38) = 0.2492$

5. **(13.8, 15.9)**

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

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

(c)  $t^* = 2.45$

(d)  $SE = 56.41$

(e)  $|t_{\text{obs}}| = 2.73$

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

(g) **reject**

7. (a) **LB of p CI = 0.624 or 62.4%**

(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)  $Z^* = 1.64$

(d)  $SE = 0.038$

(e)  $|Z_{\text{obs}}| = 1.85$

(f)  $p\text{-value} = 0.0644$

(g) **reject**

1. In a deck of strange cards, there are 758 cards. Each card has an image and a color. The amounts are shown in the table below.

|        | green | pink | teal | yellow |
|--------|-------|------|------|--------|
| flower | 50    | 21   | 57   | 28     |
| shovel | 24    | 61   | 27   | 47     |
| tree   | 22    | 49   | 79   | 69     |
| wheel  | 40    | 43   | 64   | 77     |

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

**Solution**

(a)  $P(\text{green given wheel}) = \frac{40}{40+43+64+77} = 0.179$

(b)  $P(\text{flower or teal}) = \frac{50+21+57+28+57+27+79+64-57}{758} = 0.43$

(c)  $P(\text{pink}) = \frac{21+61+49+43}{758} = 0.23$

(d)  $P(\text{tree}) = \frac{22+49+79+69}{758} = 0.289$

(e)  $P(\text{shovel given pink}) = \frac{61}{21+61+49+43} = 0.351$

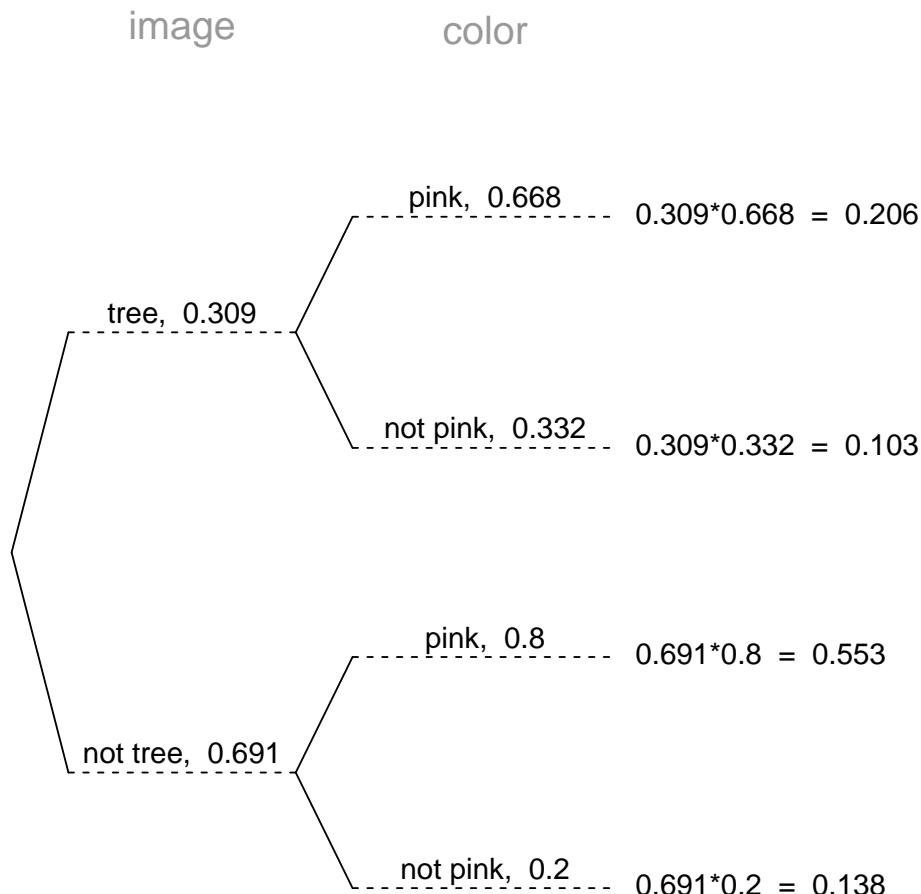
(f)  $P(\text{tree and yellow}) = \frac{69}{758} = 0.091$

2. In a deck of strange cards, each card has an image and a color. The chance of drawing a tree is 30.9%. If a tree is drawn, there is a 66.8% chance that it is pink. If a card that is not a tree is drawn, there is a 80% chance that it is pink.

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

**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.

$$P(\text{"tree" given "not pink"}) = \frac{0.103}{0.103 + 0.138} = 0.427$$

3. In a very large pile of toothpicks, the mean length is 74.82 millimeters and the standard deviation is 1.37 millimeters. If you randomly sample 125 toothpicks, what is the chance the sample mean is between 74.6 and 75 millimeters?

**Solution**

Label the given information.

$$\mu = 74.82$$

$$\sigma = 1.37$$

$$n = 125$$

$$\bar{x}_{\text{lower}} = 74.6$$

$$\bar{x}_{\text{upper}} = 75$$

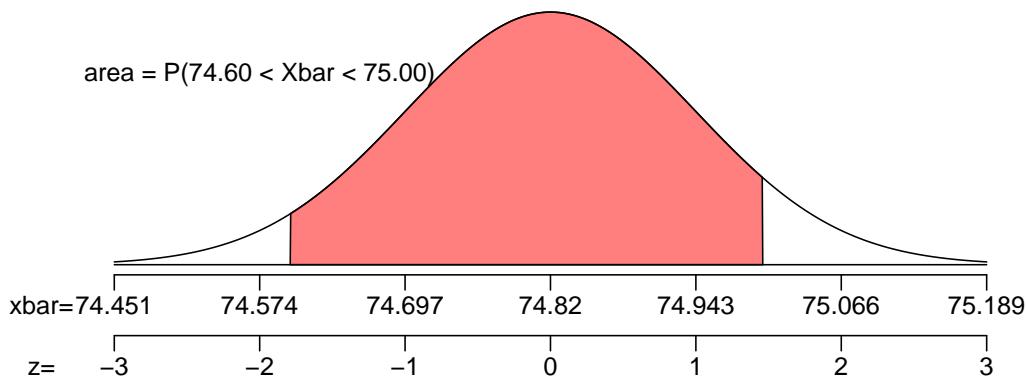
Find the standard error.

$$SE = \frac{\sigma}{\sqrt{n}} = \frac{1.37}{\sqrt{125}} = 0.123$$

Describe the sampling distribution.

$$\bar{X} \sim \mathcal{N}(74.82, 0.123)$$

Draw a sketch.



Calculate a  $z$  scores.

$$z_{\text{lower}} = \frac{\bar{x}_{\text{lower}} - \mu}{SE} = \frac{74.6 - 74.82}{0.123} = -1.79$$

$$z_{\text{upper}} = \frac{\bar{x}_{\text{upper}} - \mu}{SE} = \frac{75 - 74.82}{0.123} = 1.46$$

Determine the probability.

$$\begin{aligned} P(74.6 < \bar{X} < 75) &= \Phi(z_{\text{upper}}) - \Phi(z_{\text{lower}}) \\ &= \Phi(1.46) - \Phi(-1.79) \\ &= 0.8912 \end{aligned}$$

4. In a game, there is a 33% chance to win a round. You will play 122 rounds.
- (a) What is the probability of winning exactly 44 rounds?
  - (b) What is the probability of winning at least 33 but at most 38 rounds?

**Solution**

We use the formula for binomial probabilities.

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

$$P(X = 44) = \binom{122}{44} (0.33)^{44} (1 - 0.33)^{122-44}$$

$$P(X = 44) = \binom{122}{44} (0.33)^{44} (0.67)^{78}$$

$$P(X = 44) = 0.0581$$

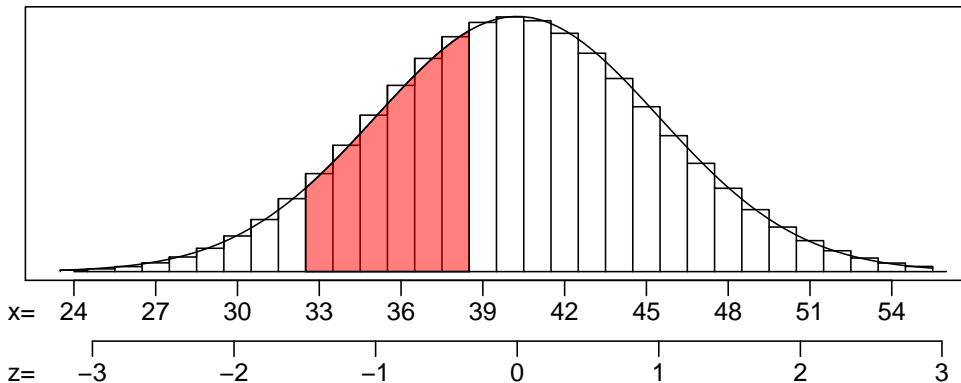
Find the mean.

$$\mu = np = (122)(0.33) = 40.26$$

Find the standard deviation.

$$\sigma = \sqrt{np(1 - p)} = \sqrt{(122)(0.33)(1 - 0.33)} = 5.1937$$

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



Find the z scores.

$$z_1 = \frac{32.5 - 40.26}{5.1937} = -1.4$$

$$z_2 = \frac{38.5 - 40.26}{5.1937} = -0.44$$

Calculate the probability.

$$P(33 \leq X \leq 38) = \Phi(-0.44) - \Phi(-1.4) = 0.2492$$

(a)  $P(X = 44) = 0.0581$

(b)  $P(33 \leq X \leq 38) = 0.2492$

5. As an ornithologist, you wish to determine the average body mass of *Passerina cyanea*. You randomly sample 31 adults of *Passerina cyanea*, resulting in a sample mean of 14.86 grams and a sample standard deviation of 2.37 grams. Determine a 98% confidence interval of the true population mean.

**Solution**

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

$$n = 31$$

$$\bar{x} = 14.86$$

$$s = 2.37$$

$$CL = 0.98$$

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

$$df = n - 1 = 30$$

Determine the critical  $t$  value,  $t^*$ , such that  $P(|T| < t^*) = 0.98$ .

$$t^* = 2.46$$

Calculate the standard error.

$$SE = \frac{s}{\sqrt{n}} = \frac{2.37}{\sqrt{31}} = 0.426$$

We want to make an inference about the population mean.

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

Determine the bounds.

$$\begin{aligned} CI &= (\bar{x} - t^* SE, \bar{x} + t^* SE) \\ &= (14.86 - 2.46 \times 0.426, 14.86 + 2.46 \times 0.426) \\ &= (13.8, 15.9) \end{aligned}$$

We are 98% confident that the population mean is between 13.8 and 15.9.

6. A treatment group of size 32 has a mean of 966 and standard deviation of 204. A control group of size 15 has a mean of 1120 and standard deviation of 168. 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 32.

- (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.

$$n_1 = 32$$

$$\bar{x}_1 = 966$$

$$s_1 = 204$$

$$n_2 = 15$$

$$\bar{x}_2 = 1120$$

$$s_2 = 168$$

$$\alpha = 0.02$$

$$df = 32$$

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  $P(|T| > t^*) = 0.02$  by using a  $t$  table.

$$t^* = 2.45$$

Calculate the standard error.

$$\begin{aligned} SE &= \sqrt{\frac{(s_1)^2}{n_1} + \frac{(s_2)^2}{n_2}} \\ &= \sqrt{\frac{(204)^2}{32} + \frac{(168)^2}{15}} \\ &= 56.41 \end{aligned}$$

Determine the test statistic.

$$\begin{aligned} t_{\text{obs}} &= \frac{(\bar{x}_2 - \bar{x}_1) - (\mu_2 - \mu_1)_0}{SE} \\ &= \frac{(1120 - 966) - (0)}{56.41} \\ &= 2.73 \end{aligned}$$

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)  $t^* = 2.45$
- (d)  $SE = 56.41$
- (e)  $|t_{\text{obs}}| = 2.73$
- (f)  $0.01 < p\text{-value} < 0.02$
- (g) reject the null

7. From a very large population, a random sample of 3600 individuals was taken. In that sample, 64% were glowing. 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 condifence interval.

**Solution**

Determine  $z^*$  such that  $P(|Z| < z^*) = 0.95$ .

$$z^* = 1.96$$

Calculate the standard error.

$$SE = \sqrt{\frac{\hat{p}(1 - \hat{p})}{n}} = \sqrt{\frac{(0.64)(1 - 0.64)}{3600}} = 0.008$$

Calculate the margin of error.

$$ME = z^* SE = (1.96)(0.008) = 0.0157$$

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.

$$(0.624, 0.656)$$

We are 95% confident that the true population proportion is between 62.4% and 65.6%.

- (a) The lower bound = 0.624, which can also be expressed as 62.4%.
- (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 195 and a control group of size 165. The results are summarized in the table below.

|                | treatment | control |
|----------------|-----------|---------|
| omnivorous     | 172       | 134     |
| not omnivorous | 23        | 31      |

Using a significance level of 0.1, determine whether the treatment causes an effect on the proportion of cases that are omnivorous.

- (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  $P(|Z| > z^*) = 0.1$ .

$$z^* = \Phi^{-1} \left( 1 - \frac{\alpha}{2} \right) = 1.64$$

Determine the sample proportions.

$$\hat{p}_1 = \frac{172}{195} = 0.882$$

$$\hat{p}_2 = \frac{134}{165} = 0.812$$

Determine the difference of sample proportions.

$$\hat{p}_2 - \hat{p}_1 = 0.812 - 0.882 = -0.07$$

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

$$\hat{p} = \frac{172 + 134}{195 + 165} = 0.85$$

Determine the standard error.

$$\begin{aligned} SE &= \sqrt{\frac{\hat{p}(1 - \hat{p})}{n_1} + \frac{\hat{p}(1 - \hat{p})}{n_2}} \\ &= \sqrt{\frac{(0.85)(0.15)}{195} + \frac{(0.85)(0.15)}{165}} \\ &= 0.0378 \end{aligned}$$

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

$$H_0 : \hat{P}_2 - \hat{P}_1 \sim \mathcal{N}(0, 0.0378)$$

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.

$$\begin{aligned} z &= \frac{(\hat{p}_2 - \hat{p}_1) - (p_2 - p_1)_0}{SE} \\ &= \frac{(0.812 - 0.882) - 0}{0.0378} \\ &= -1.85 \end{aligned}$$

Determine the  $p$ -value.

$$\begin{aligned} p\text{-value} &= 2 \cdot \Phi(-|z|) \\ &= 2 \cdot \Phi(-1.85) \\ &= 0.0644 \end{aligned}$$

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)  $z^* = 1.64$
- (d)  $SE = 0.0378$
- (e)  $|Z_{\text{obs}}| = 1.85$
- (f)  $p\text{-value} = 0.0644$
- (g) reject the null