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) $P(\textbf{tree}) = 0.232$

   (b) $P(\textbf{dog given blue}) = 0.108$

   (c) $P(\textbf{pink}) = 0.227$

   (d) $P(\textbf{orange given dog}) = 0.0966$

   (e) $P(\textbf{wheel or blue}) = 0.354$

   (f) $P(\textbf{wheel and green}) = 0.0408$

2. $P(\textbf{"wheel" given "not violet"}) = 0.267$

3. $P(65.66 < X < 66.57) = 0.7185$

4. (a) $P(X = 42) = 0.0546$

   (b) $P(26 \leq X \leq 49) = 0.981$

5. $\mathbf{(19.4, 23)}$

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

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

   (c) $t^{\star} = 2.11$

   (d) $SE = 19.933$

   (e) $\mid t_{\textbf{obs}} \mid = 2.01$

   (f) $0.04 < p\textbf{-value} < 0.05$

   (g) **retain**

7. (a) **LB of p CI = 0.909 or** 90.9%

   (b) **UB of p CI = 0.927 or** 92.7%

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

(b) $H_{\mathbf{A}}: \quad p_2 - p_1 \neq 0$

(c) $z^\star = 2.33$

(d) $SE = 0.047$

(e) $\mid z_{\mathbf{obs}} \mid = 2.56$

(f) p**-value** $= 0.0104$

(g) **reject**

1. In a deck of strange cards, there are 1152 cards. Each card has an image and a color. The amounts are shown in the table below.

	blue	green	orange	pink	white
dog	20	19	17	56	64
shovel	80	65	99	89	97
tree	29	63	90	37	48
wheel	57	47	10	79	86

(a) What is the probability a random card is a tree?

(b) What is the probability a random card is a dog given it is blue?

(c) What is the probability a random card is pink?

(d) What is the probability a random card is orange given it is a dog?

(e) What is the probability a random card is either a wheel or blue (or both)?

(f) What is the probability a random card is both a wheel and green?

Solution

(a) $P(\text{tree}) = \frac{29+63+90+37+48}{1152} = 0.232$

(b) $P(\text{dog given blue}) = \frac{20}{20+80+29+57} = 0.108$

(c) $P(\text{pink}) = \frac{56+89+37+79}{1152} = 0.227$

(d) $P(\text{orange given dog}) = \frac{17}{20+19+17+56+64} = 0.0966$

(e) $P(\text{wheel or blue}) = \frac{57+47+10+79+86+20+80+29+57-57}{1152} = 0.354$

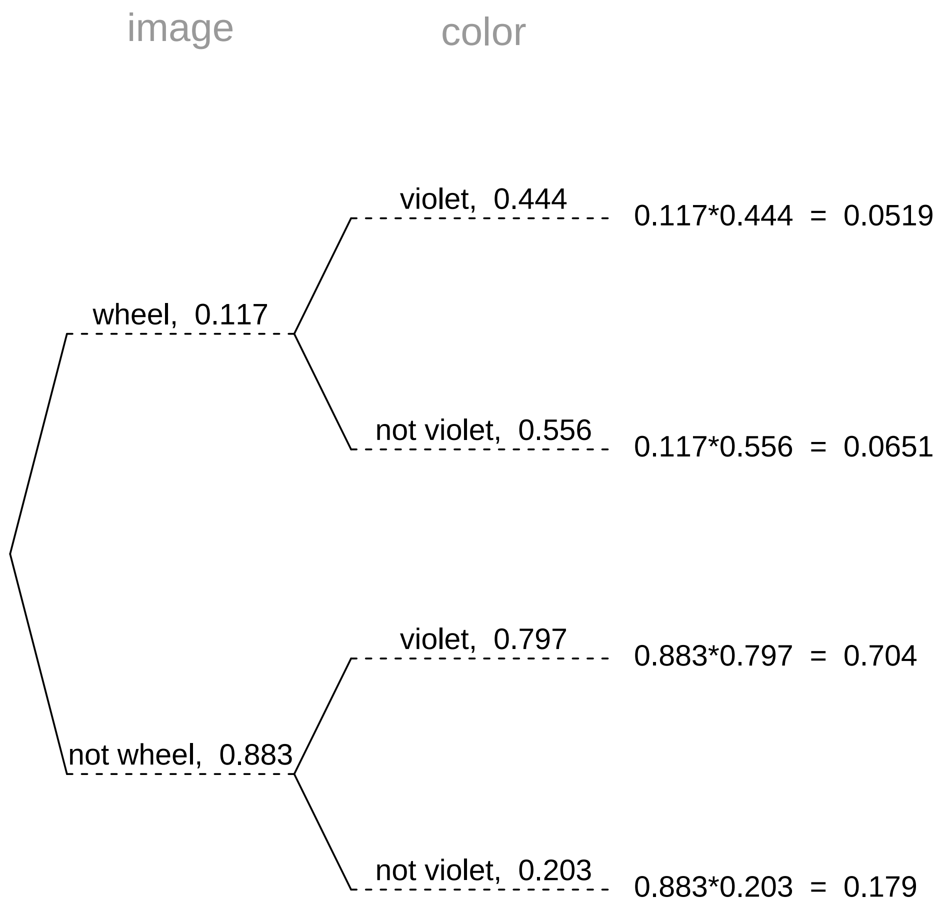
(f) $P(\text{wheel and green}) = \frac{47}{1152} = 0.0408$

2. In a deck of strange cards, each card has an image and a color. The chance of drawing a wheel is 11.7%. If a wheel is drawn, there is a 44.4% chance that it is violet. If a card that is not a wheel is drawn, there is a 79.7% chance that it is violet.

   Now, someone draws a random card and reveals it is not violet. What is the chance the card is 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 probabilites. The results (products) are joint probabilities.



Determine the appropriate conditional probability.

$$P(\text{"wheel" given "not violet"}) = \frac{0.0651}{0.0651 + 0.179} = 0.267$$

3. In a very large pile of toothpicks, the mean length is 65.86 millimeters and the standard deviation is 3.54 millimeters. If you randomly sample 120 toothpicks, what is the chance the sample mean is between 65.66 and 66.57 millimeters?

Solution
Label the given information.

$$\mu = 65.86$$
$$\sigma = 3.54$$
$$n = 120$$
$$\bar{x}_{\text{lower}} = 65.66$$
$$\bar{x}_{\text{upper}} = 66.57$$

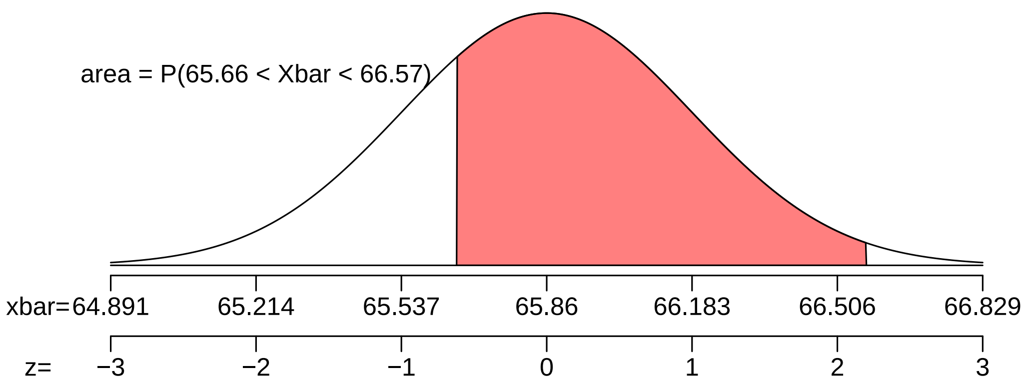
Find the standard error.

$$SE = \frac{\sigma}{\sqrt{n}} = \frac{3.54}{\sqrt{120}} = 0.323$$

Describe the sampling distribution.

$$\bar{X} \sim \mathcal{N}(65.86,\ 0.323)$$

Draw a sketch.



Calculate a z scores.

$$z_{\text{lower}} = \frac{x_{\text{lower}} - \mu}{SE} = \frac{65.66 - 65.86}{0.323} = -0.62$$

$$z_{\text{upper}} = \frac{x_{\text{upper}} - \mu}{SE} = \frac{66.57 - 65.86}{0.323} = 2.2$$

Determine the probability.

$$\begin{aligned} P(65.66 < X < 66.57) &= \Phi(z_{\text{upper}}) - \Phi(z_{\text{lower}}) \\ &= \Phi(2.2) - \Phi(-0.62) \\ &= 0.7185 \end{aligned}$$

4. In a game, there is a 31% chance to win a round. You will play 122 rounds.
   (a) What is the probability of winning exactly 42 rounds?
   (b) What is the probability of winning at least 26 but at most 49 rounds?

Solution

We use the formula for binomial probabilities.

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

$$P(X = 42) = \binom{122}{42} (0.31)^{42} (1 - 0.31)^{122-42}$$

$$P(X = 42) = \binom{122}{42} (0.31)^{42} (0.69)^{80}$$

$$P(X = 42) = 0.0546$$

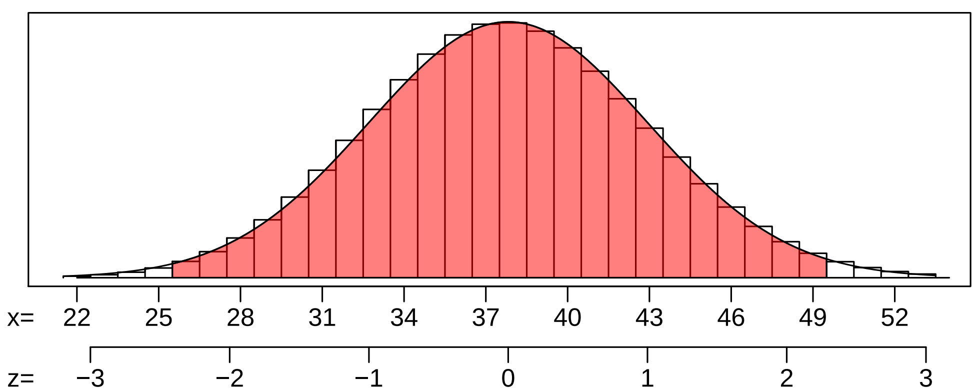
Find the mean.

$$\mu = np = (122)(0.31) = 37.82$$

Find the standard deviation.

$$\sigma = \sqrt{np(1 - p)} = \sqrt{(122)(0.31)(1 - 0.31)} = 5.1084$$

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



Find the z scores.

$$z_1 = \frac{25.5 - 37.82}{5.1084} = -2.41$$

$$z_2 = \frac{49.5 - 37.82}{5.1084} = 2.29$$

Calculate the probability.

$$P(26 \leq X \leq 49) = \Phi(2.29) - \Phi(-2.41) = 0.981$$

(a) $P(X = 42) = 0.0546$

(b) $P(26 \leq X \leq 49) = 0.981$

5. As an ornithologist, you wish to determine the average body mass of *Seiurus noveboracensis*. You randomly sample 27 adults of *Seiurus noveboracensis*, resulting in a sample mean of 21.2 grams and a sample standard deviation of 3.71 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.

$$\begin{aligned} n &= 27 \\ \bar{x} &= 21.2 \\ s &= 3.71 \\ CL &= 0.98 \end{aligned}$$

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

$$df = n - 1 = 26$$

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

$$t^\star = 2.48$$

Calculate the standard error.

$$SE = \frac{s}{\sqrt{n}} = \frac{3.71}{\sqrt{27}} = 0.714$$

We want to make an inference about the population mean.

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

Determine the bounds.

$$\begin{aligned} CI &= (\bar{x} - t^\star SE,\ \bar{x} + t^\star SE) \\ &= (21.2 - 2.48 \times 0.714,\ 21.2 + 2.48 \times 0.714) \\ &= (19.4,\ 23) \end{aligned}$$

We are 98% confident that the population mean is between 19.4 and 23.

6. A treatment group of size 28 has a mean of 1010 and standard deviation of 88. A control group of size 40 has a mean of 1050 and standard deviation of 69.5. If you decided to use a signficance 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 49.

   (a) State the null hypothesis.

   (b) State the alternative hypothesis.

   (c) Evaluate the critical value. (The critical value is either $z^\star$ or $t^\star$. 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_{obs} or t_{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.

$$n_1 = 28$$
$$\bar{x}_1 = 1010$$
$$s_1 = 88$$
$$n_2 = 40$$
$$\bar{x}_2 = 1050$$
$$s_2 = 69.5$$
$$\alpha = 0.04$$
$$df = 49$$

State the hypotheses.

$$H_0: \quad \mu_2 - \mu_1 = 0$$
$$H_A: \quad \mu_2 - \mu_1 \neq 0$$

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

$$t^\star = 2.11$$

Calculate the standard error.

$$\begin{aligned} SE &= \sqrt{\frac{(s_1)^2}{n_1} + \frac{(s_2)^2}{n_2}} \\ &= \sqrt{\frac{(88)^2}{28} + \frac{(69.5)^2}{40}} \\ &= 19.933 \end{aligned}$$

Determine the test statistic.

$$\begin{aligned} t_{\text{obs}} &= \frac{(\bar{x}_2 - \bar{x}_1) - (\mu_2 - \mu_1)_0}{SE} \\ &= \frac{(1050 - 1010) - (0)}{19.933} \\ &= 2.01 \end{aligned}$$

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

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

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

$$0.04 < p\text{-value} < 0.05$$

Compare p-value and α.

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

We conclude that we should retain the null hypothesis.

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

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

(c) $t^{\star} = 2.11$

(d) $SE = 19.933$

(e) $|t_{\text{obs}}| = 2.01$

(f) $0.04 < p\text{-value} < 0.05$

(g) retain the null

7. From a very large population, a random sample of 1700 individuals was taken. In that sample, 91.8% were broken. 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 $z^{\star}$ such that $P(|Z| < z^{\star}) = 0.8$.

$$z^{\star} = 1.28$$

Calculate the standard error.

$$SE = \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = \sqrt{\frac{(0.918)(1-0.918)}{1700}} = 0.00665$$

Calculate the margin of error.

$$ME = z^{\star} SE = (1.28)(0.00665) = 0.00851$$

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.

$$(0.909\,,\ 0.927)$$

We are 80% confident that the true population proportion is between 90.9% and 92.7%.

(a) The lower bound = 0.909, which can also be expressed as 90.9%.

(b) The upper bound = 0.927, which can also be expressed as 92.7%.

8. An experiment is run with a treatment group of size 223 and a control group of size 234. The results are summarized in the table below.

	treatment	control
sorry	100	133
not sorry	123	101

Using a significance level of 0.02, 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 $z^{\star}$ or $t^{\star}$. 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_{obs} or t_{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.

$$H_0: \quad p_2 - p_1 = 0$$

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

Find $z^\star$ such that $P(|Z| > z^\star) = 0.02$.

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

Determine the sample proportions.

$$\hat{p}_1 = \frac{100}{223} = 0.448$$

$$\hat{p}_2 = \frac{133}{234} = 0.568$$

Determine the difference of sample proportions.

$$\hat{p}_2 - \hat{p}_1 = 0.568 - 0.448 = 0.12$$

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

$$\hat{p} = \frac{100 + 133}{223 + 234} = 0.51$$

Determine the standard error.

$$SE = \sqrt{\frac{\hat{p}(1-\hat{p})}{n_1} + \frac{\hat{p}(1-\hat{p})}{n_2}}$$

$$= \sqrt{\frac{(0.51)(0.49)}{223} + \frac{(0.51)(0.49)}{234}}$$

$$= 0.0468$$

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

$$H_0: \quad \hat{P}_2 - \hat{P}_1 \sim \mathcal{N}(0,\ 0.0468)$$

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.

$$\begin{aligned} z &= \frac{(\hat{p}_2 - \hat{p}_1) - (p_2 - p_1)_0}{SE} \\ &= \frac{(0.568 - 0.448) - 0}{0.0468} \\ &= 2.56 \end{aligned}$$

Determine the p-value.

$$\begin{aligned} p\text{-value} &= 2 \cdot \Phi(-|z|) \\ &= 2 \cdot \Phi(-2.56) \\ &= 0.0104 \end{aligned}$$

Compare the p-value to the signficance level.

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

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

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

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

(c) $z^{\star} = 2.33$

(d) $SE = 0.0468$

(e) $|z_{\text{obs}}| = 2.56$

(f) p-value = 0.0104

(g) reject the null