

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

$$(a) P(\text{blue}) = \frac{61+95+78+86+80}{1027} = 0.389$$

$$(b) P(\text{red given gem}) = \frac{68}{86+77+68} = 0.294$$

$$(c) P(\text{wheel and red}) = \frac{59}{1027} = 0.0574$$

$$(d) P(\text{gem or blue}) = \frac{86+77+68+61+95+78+86+80-86}{1027} = 0.531$$

$$(e) P(\text{gem given red}) = \frac{68}{32+46+85+68+59} = 0.234$$

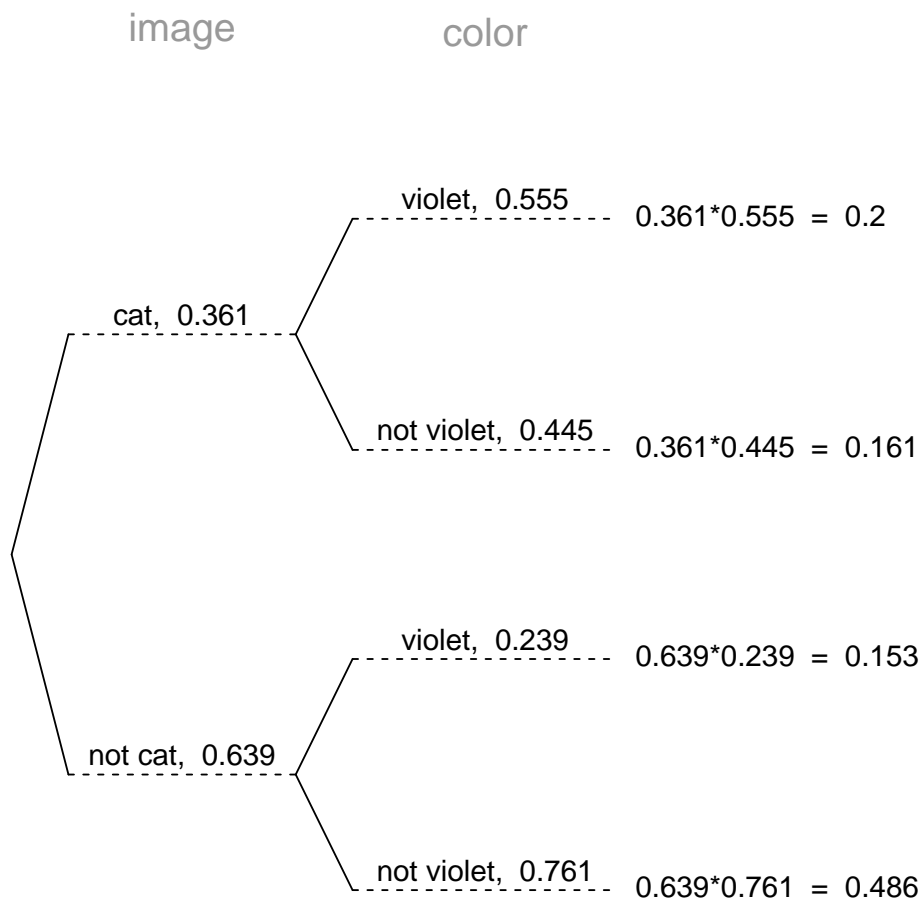
$$(f) P(\text{cat}) = \frac{95+57+46}{1027} = 0.193$$

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.

$$\mu = 61.02$$

$$\sigma = 1.96$$

$$n = 196$$

$$\bar{x}_{\text{lower}} = 60.84$$

$$\bar{x}_{\text{upper}} = 61.37$$

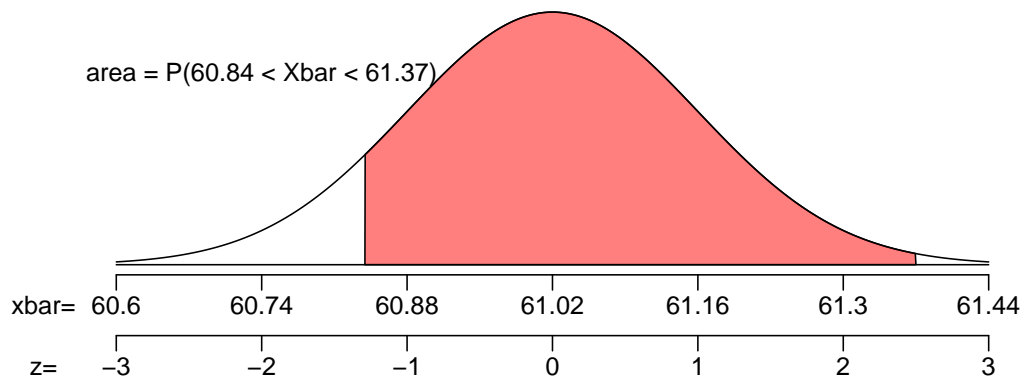
Find the standard error.

$$SE = \frac{\sigma}{\sqrt{n}} = \frac{1.96}{\sqrt{196}} = 0.14$$

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.

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

$$P(X = 49) = \binom{199}{49} (0.22)^{49} (1 - 0.22)^{199-49}$$

$$P(X = 49) = \binom{199}{49} (0.22)^{49} (0.78)^{150}$$

$$P(X = 49) = 0.0444$$

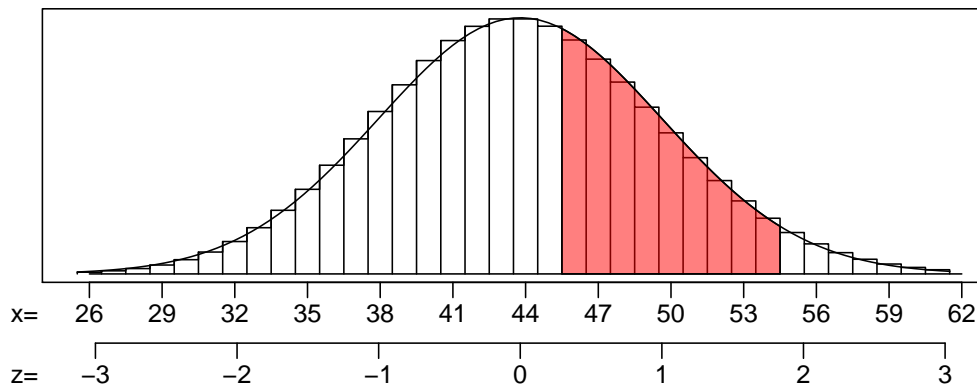
Find the mean.

$$\mu = np = (199)(0.22) = 43.78$$

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 σ 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_{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.

$$n_1 = 19$$

$$\bar{x}_1 = 1120$$

$$s_1 = 158$$

$$n_2 = 31$$

$$\bar{x}_2 = 999$$

$$s_2 = 180$$

$$\alpha = 0.01$$

$$df = 42$$

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 α .

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

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 confidence 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_{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.

$$H_0 : p_2 - p_1 = 0$$

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

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