

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

$$(a) P(\text{gem or orange}) = \frac{22+49+46+49+16+48-49}{447} = 0.405$$

$$(b) P(\text{gem}) = \frac{22+49+46}{447} = 0.262$$

$$(c) P(\text{teal}) = \frac{46+90+33}{447} = 0.378$$

$$(d) P(\text{wheel and gray}) = \frac{98}{447} = 0.219$$

$$(e) P(\text{gem given gray}) = \frac{22}{22+45+98} = 0.133$$

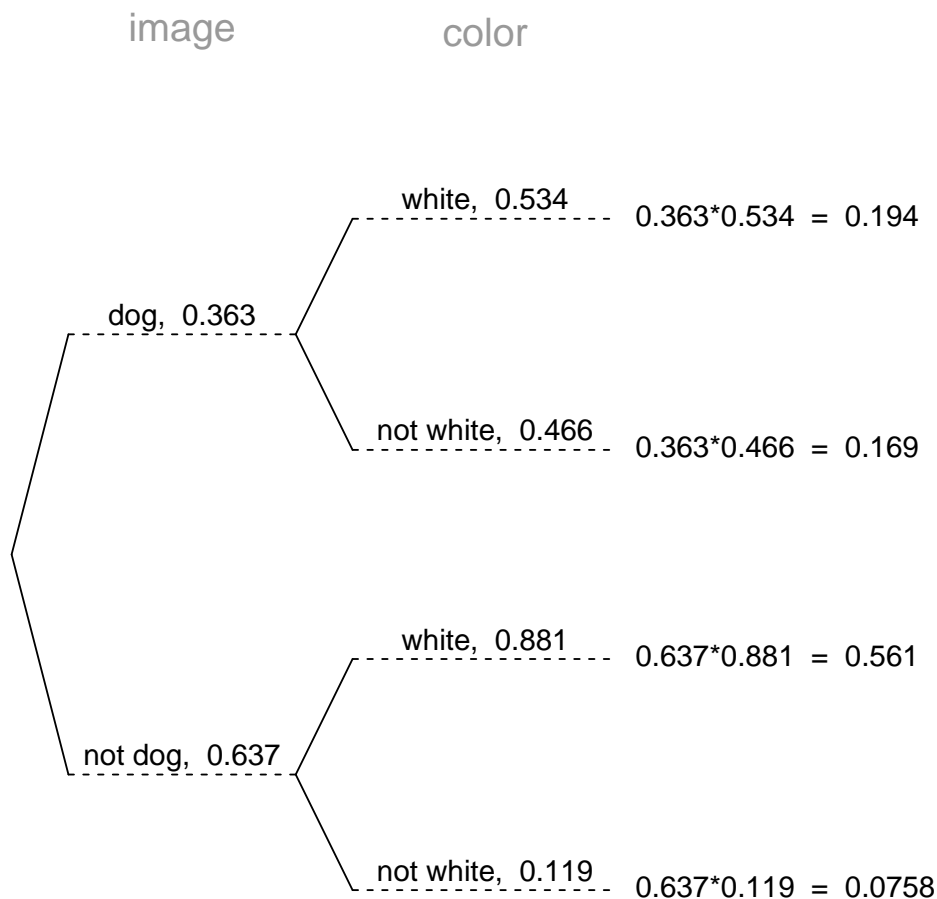
$$(f) P(\text{teal given wheel}) = \frac{33}{98+48+33} = 0.184$$

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.

$$\mu = 60.51$$

$$\sigma = 1.44$$

$$n = 121$$

$$\bar{x}_{\text{lower}} = 60.24$$

$$\bar{x}_{\text{upper}} = 60.69$$

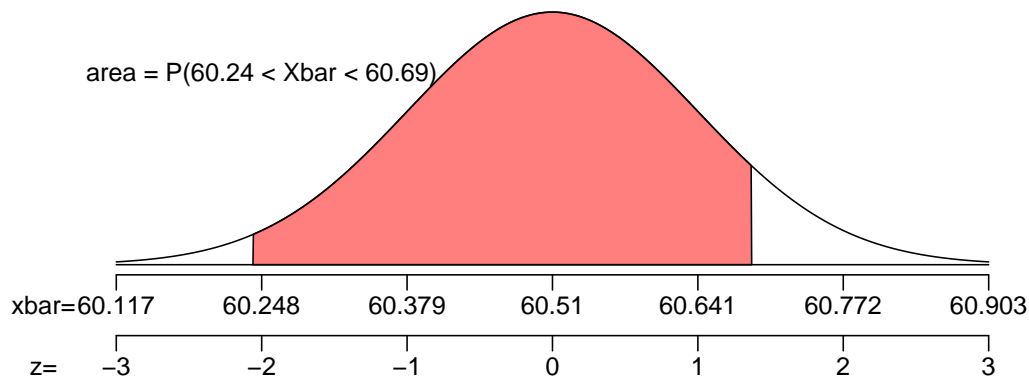
Find the standard error.

$$SE = \frac{\sigma}{\sqrt{n}} = \frac{1.44}{\sqrt{121}} = 0.131$$

Describe the sampling distribution.

$$\bar{X} \sim \mathcal{N}(60.51, 0.131)$$

Draw a sketch.



Calculate a z scores.

$$z_{\text{lower}} = \frac{x_{\text{lower}} - \mu}{SE} = \frac{60.24 - 60.51}{0.131} = -2.06$$

$$z_{\text{upper}} = \frac{x_{\text{upper}} - \mu}{SE} = \frac{60.69 - 60.51}{0.131} = 1.37$$

Determine the probability.

$$\begin{aligned} P(60.24 < X < 60.69) &= \Phi(z_{\text{upper}}) - \Phi(z_{\text{lower}}) \\ &= \Phi(1.37) - \Phi(-2.06) \\ &= 0.895 \end{aligned}$$

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.

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

$$P(X = 73) = \binom{115}{73} (0.6)^{73} (1 - 0.6)^{115-73}$$

$$P(X = 73) = \binom{115}{73} (0.6)^{73} (0.4)^{42}$$

$$P(X = 73) = 0.0575$$

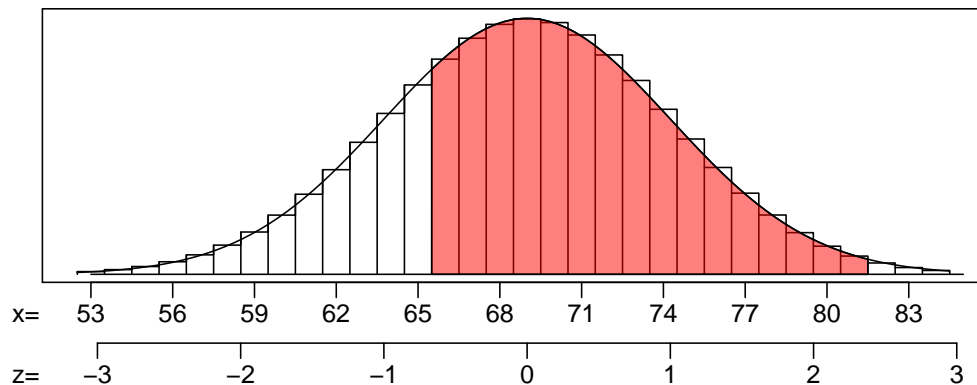
Find the mean.

$$\mu = np = (115)(0.6) = 69$$

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 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 $P(|T| < t^*) = 0.95$.

$$t^* = 2.18$$

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_{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.

$$n_1 = 35$$

$$\bar{x}_1 = 1.03$$

$$s_1 = 0.188$$

$$n_2 = 38$$

$$\bar{x}_2 = 1.15$$

$$s_2 = 0.235$$

$$\alpha = 0.02$$

$$df = 69$$

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.

$$t^* = 2.38$$

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 α .

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

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 confidence 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_{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.

$$\hat{p}_1 = \frac{32}{198} = 0.162$$

$$\hat{p}_2 = \frac{43}{175} = 0.246$$

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