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(**bike given black**) = 0.269

   (b) P(**bike and white**) = 0.0607

   (c) P(**shovel or violet**) = 0.458

   (d) P(**wheel**) = 0.206

   (e) P(**white given bike**) = 0.203

   (f) P(**white**) = 0.273

2. P(**"not shovel" given "not pink"**) = 0.832

3. $P(59.97 < X < 60.24) = 0.6641$

4. (a) $P(X = 107) = 0.0989$

   (b) $P(106 \leq X \leq 115) = 0.7905$

5. **(13.9, 17.4)**

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

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

   (c) $t^{\star} = 2.09$

   (d) $SE = 61.029$

   (e) $\mid t_{\textbf{obs}} \mid = 1.95$

   (f) $0.05 <$ p**-value** < 0.1

   (g) **retain**

7. (a) **LB of p CI = 0.165 or** 16.5%

   (b) **UB of p CI = 0.183 or** 18.3%

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

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

(c) $z^{\star} = 2.58$

(d) $SE = 0.049$

(e) $|\ z_{\mathbf{obs}}\ | = 2.68$

(f) p-**value** $= 0.0074$

(g) **reject**

1. In a deck of strange cards, there are 1021 cards. Each card has an image and a color. The amounts are shown in the table below.

	black	blue	violet	white
bike	86	58	99	62
gem	94	25	15	92
shovel	96	17	82	85
wheel	44	52	74	40

(a) What is the probability a random card is a bike given it is black?

(b) What is the probability a random card is both a bike and white?

(c) What is the probability a random card is either a shovel or violet (or both)?

(d) What is the probability a random card is a wheel?

(e) What is the probability a random card is white given it is a bike?

(f) What is the probability a random card is white?

Solution

(a) $P(\text{bike given black}) = \frac{86}{86+94+96+44} = 0.269$

(b) $P(\text{bike and white}) = \frac{62}{1021} = 0.0607$

(c) $P(\text{shovel or violet}) = \frac{96+17+82+85+99+15+82+74-82}{1021} = 0.458$

(d) $P(\text{wheel}) = \frac{44+52+74+40}{1021} = 0.206$

(e) $P(\text{white given bike}) = \frac{62}{86+58+99+62} = 0.203$

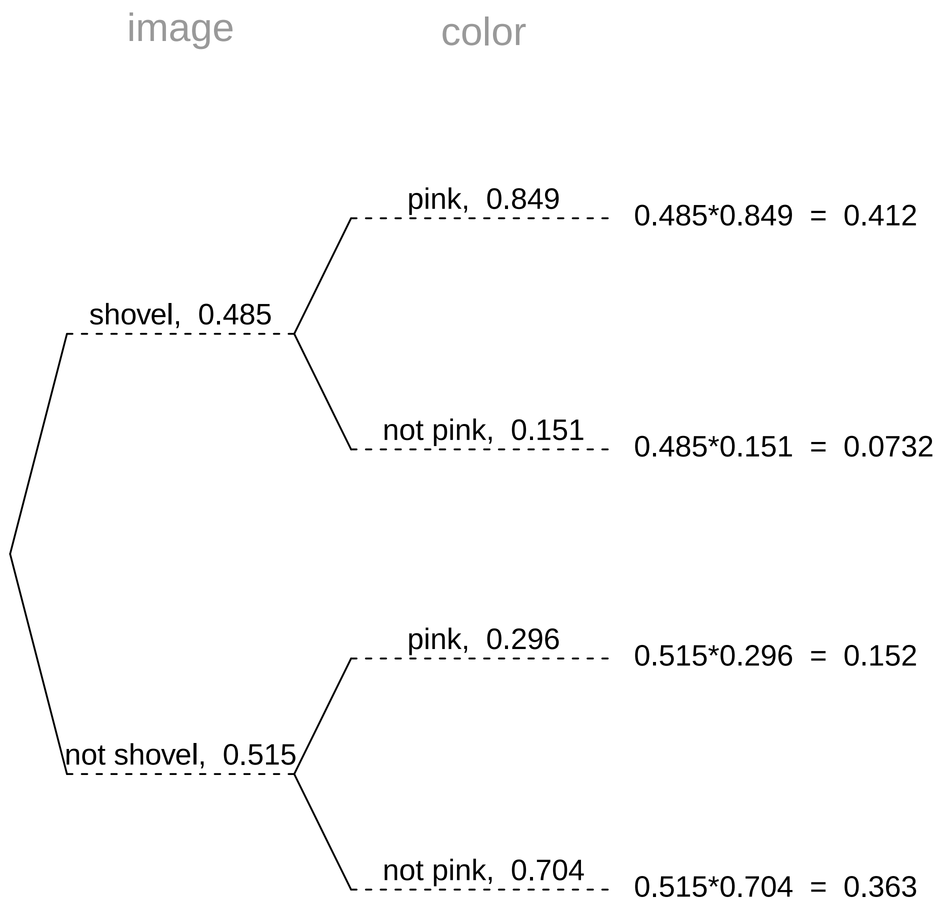
(f) $P(\text{white}) = \frac{62+92+85+40}{1021} = 0.273$

2. In a deck of strange cards, each card has an image and a color. The chance of drawing a shovel is 48.5%. If a shovel is drawn, there is a 84.9% chance that it is pink. If a card that is not a shovel is drawn, there is a 29.6% chance that it is pink.

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



Determine the appropriate conditional probability.

$$P(\text{"not shovel" given "not pink"}) = \frac{0.363}{0.363 + 0.0732} = 0.832$$

3. In a very large pile of toothpicks, the mean length is 60.04 millimeters and the standard deviation is 1.35 millimeters. If you randomly sample 120 toothpicks, what is the chance the sample mean is between 59.97 and 60.24 millimeters?

Solution

Label the given information.

$$\mu = 60.04$$
$$\sigma = 1.35$$
$$n = 120$$
$$\bar{x}_{\text{lower}} = 59.97$$
$$\bar{x}_{\text{upper}} = 60.24$$

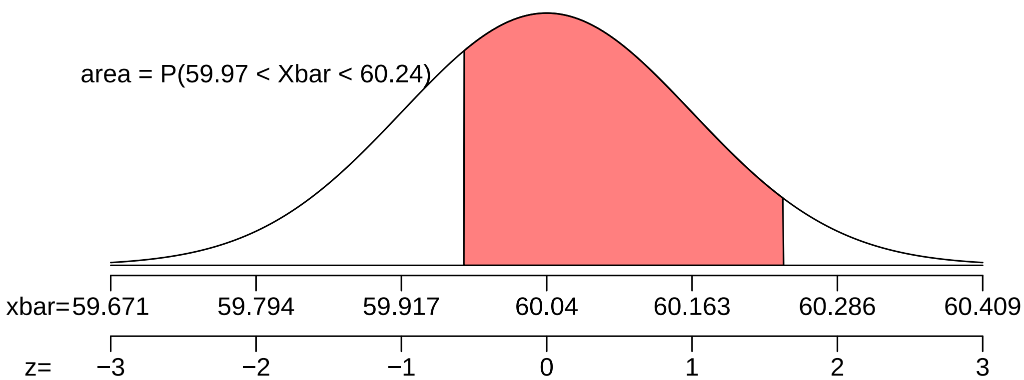
Find the standard error.

$$SE = \frac{\sigma}{\sqrt{n}} = \frac{1.35}{\sqrt{120}} = 0.123$$

Describe the sampling distribution.

$$\bar{X} \sim \mathcal{N}(60.04,\ 0.123)$$

Draw a sketch.



Calculate a z scores.

$$z_{\text{lower}} = \frac{x_{\text{lower}} - \mu}{SE} = \frac{59.97 - 60.04}{0.123} = -0.57$$

$$z_{\text{upper}} = \frac{x_{\text{upper}} - \mu}{SE} = \frac{60.24 - 60.04}{0.123} = 1.63$$

Determine the probability.

$$\begin{aligned} P(59.97 < X < 60.24) &= \Phi(z_{\text{upper}}) - \Phi(z_{\text{lower}}) \\ &= \Phi(1.63) - \Phi(-0.57) \\ &= 0.6641 \end{aligned}$$

4. In a game, there is a 89% chance to win a round. You will play 122 rounds.
   (a) What is the probability of winning exactly 107 rounds?
   (b) What is the probability of winning at least 106 but at most 115 rounds?

Solution

We use the formula for binomial probabilities.

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

$$P(X = 107) = \binom{122}{107}(0.89)^{107}(1-0.89)^{122-107}$$

$$P(X = 107) = \binom{122}{107}(0.89)^{107}(0.11)^{15}$$

$$P(X = 107) = 0.0989$$

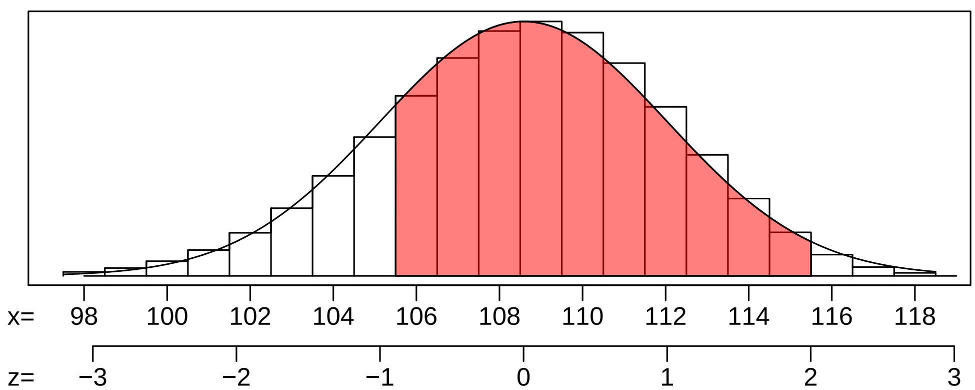
Find the mean.

$$\mu = np = (122)(0.89) = 108.58$$

Find the standard deviation.

$$\sigma = \sqrt{np(1-p)} = \sqrt{(122)(0.89)(1-0.89)} = 3.456$$

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



Find the z scores.

$$z_1 = \frac{105.5 - 108.58}{3.456} = -0.89$$

$$z_2 = \frac{115.5 - 108.58}{3.456} = 2$$

Calculate the probability.

$$P(106 \leq X \leq 115) = \Phi(2) - \Phi(-0.89) = 0.7905$$

(a) $P(X = 107) = 0.0989$

(b) $P(106 \leq X \leq 115) = 0.7905$

5. As an ornithologist, you wish to determine the average body mass of *Oporornis formosus*. You randomly sample 15 adults of *Oporornis formosus*, resulting in a sample mean of 15.63 grams and a sample standard deviation of 3.84 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.

$$\begin{aligned} n &= 15 \\ \bar{x} &= 15.63 \\ s &= 3.84 \\ CL &= 0.9 \end{aligned}$$

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

$$df = n - 1 = 14$$

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

$$t^\star = 1.76$$

Calculate the standard error.

$$SE = \frac{s}{\sqrt{n}} = \frac{3.84}{\sqrt{15}} = 0.991$$

We want to make an inference about the population mean.

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

Determine the bounds.

$$\begin{aligned} CI &= (\bar{x} - t^\star SE,\ \bar{x} + t^\star SE) \\ &= (15.63 - 1.76 \times 0.991,\ 15.63 + 1.76 \times 0.991) \\ &= (13.9,\ 17.4) \end{aligned}$$

We are 90% confident that the population mean is between 13.9 and 17.4.

6. A treatment group of size 28 has a mean of 1100 and standard deviation of 206. A control group of size 9 has a mean of 981 and standard deviation of 141. If you decided to use a signficance 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 19.

   (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.

$$n_1 = 28$$
$$\bar{x}_1 = 1100$$
$$s_1 = 206$$
$$n_2 = 9$$
$$\bar{x}_2 = 981$$
$$s_2 = 141$$
$$\alpha = 0.05$$
$$df = 19$$

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

$$t^\star = 2.09$$

Calculate the standard error.

$$\begin{aligned} SE &= \sqrt{\frac{(s_1)^2}{n_1} + \frac{(s_2)^2}{n_2}} \\ &= \sqrt{\frac{(206)^2}{28} + \frac{(141)^2}{9}} \\ &= 61.029 \end{aligned}$$

Determine the test statistic.

$$\begin{aligned} t_{\text{obs}} &= \frac{(\bar{x}_2 - \bar{x}_1) - (\mu_2 - \mu_1)_0}{SE} \\ &= \frac{(981 - 1100) - (0)}{61.029} \\ &= -1.95 \end{aligned}$$

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.

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

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) $H_A: \mu_2 - \mu_1 \neq 0$

(c) $t^\star = 2.09$

(d) $SE = 61.029$

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

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

(g) retain the null

7. From a very large population, a random sample of 4500 individuals was taken. In that sample, 17.4% were frigid. 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 $z^\star$ such that $P(|Z| < z^\star) = 0.9$.

$$z^\star = 1.64$$

Calculate the standard error.

$$SE = \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = \sqrt{\frac{(0.174)(1-0.174)}{4500}} = 0.00565$$

Calculate the margin of error.

$$ME = z^\star SE = (1.64)(0.00565) = 0.00927$$

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.

$$(0.165\,,\ 0.183)$$

We are 90% confident that the true population proportion is between 16.5% and 18.3%.

(a) The lower bound = 0.165, which can also be expressed as 16.5%.

(b) The upper bound = 0.183, which can also be expressed as 18.3%.

8. An experiment is run with a treatment group of size 118 and a control group of size 111. The results are summarized in the table below.

	treatment	control
sick	106	85
not sick	12	26

Using a significance level of 0.01, 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 $z^{\star}$ 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 $P(|Z| > z^\star) = 0.01$.

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

Determine the sample proportions.

$$\hat{p}_1 = \frac{106}{118} = 0.898$$
$$\hat{p}_2 = \frac{85}{111} = 0.766$$

Determine the difference of sample proportions.

$$\hat{p}_2 - \hat{p}_1 = 0.766 - 0.898 = -0.132$$

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

$$\hat{p} = \frac{106 + 85}{118 + 111} = 0.834$$

Determine the standard error.

$$\begin{aligned} SE &= \sqrt{\frac{\hat{p}(1-\hat{p})}{n_1} + \frac{\hat{p}(1-\hat{p})}{n_2}} \\ &= \sqrt{\frac{(0.834)(0.166)}{118} + \frac{(0.834)(0.166)}{111}} \\ &= 0.0492 \end{aligned}$$

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

$$H_0: \quad \hat{P}_2 - \hat{P}_1 \sim \mathcal{N}(0,\ 0.0492)$$

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.

$$\begin{aligned} z &= \frac{(\hat{p}_2 - \hat{p}_1) - (p_2 - p_1)_0}{SE} \\ &= \frac{(0.766 - 0.898) - 0}{0.0492} \\ &= -2.68 \end{aligned}$$

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 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) $z^\star = 2.58$
(d) $SE = 0.0492$
(e) $|z_{\text{obs}}| = 2.68$
(f) p-value = 0.0074
(g) reject the null