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(**horn given black**) = 0.4

   (b) P(**pig or teal**) = 0.614

   (c) P(**teal given pig**) = 0.358

   (d) P(**horn and indigo**) = 0.0707

   (e) P(**black**) = 0.269

   (f) P(**pig**) = 0.367

2. P(**"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) $H_0: \quad \mu_2 - \mu_1 = 0$

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

   (c) $t^\star = 2.45$

   (d) $SE = 4.179$

   (e) $\mid t_{\mathbf{obs}} \mid = 2.39$

   (f) $0.02 <$ p**-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) $H_0: \quad p_2 - p_1 = 0$

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

(c) $z^{\star} = 2.05$

(d) $SE = 0.056$

(e) $|\, z_{\mathbf{obs}} \,| = 2.31$

(f) p**-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) $P(\text{black}) = \frac{51+64+45}{594} = 0.269$

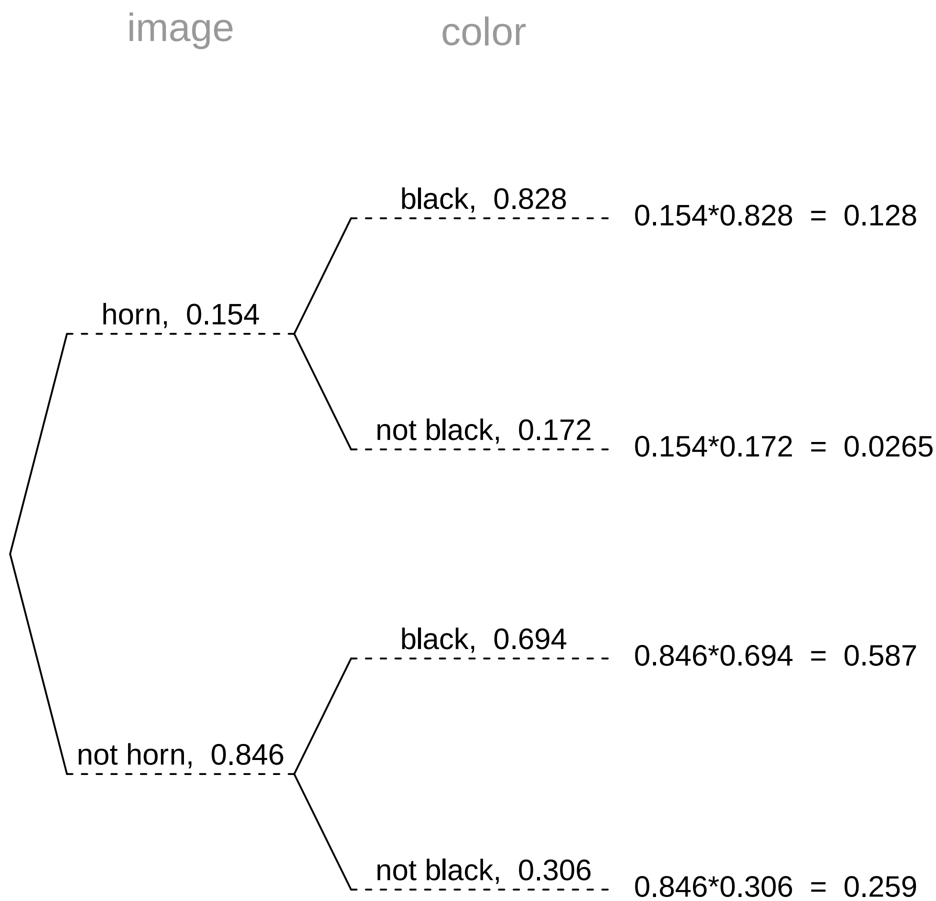
(f) $P(\text{pig}) = \frac{45+32+63+78}{594} = 0.367$

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 probabilites. 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.

$$\mu = 71.49$$
$$\sigma = 2.84$$
$$n = 169$$
$$\bar{x}_{\text{lower}} = 71.36$$
$$\bar{x}_{\text{upper}} = 71.65$$

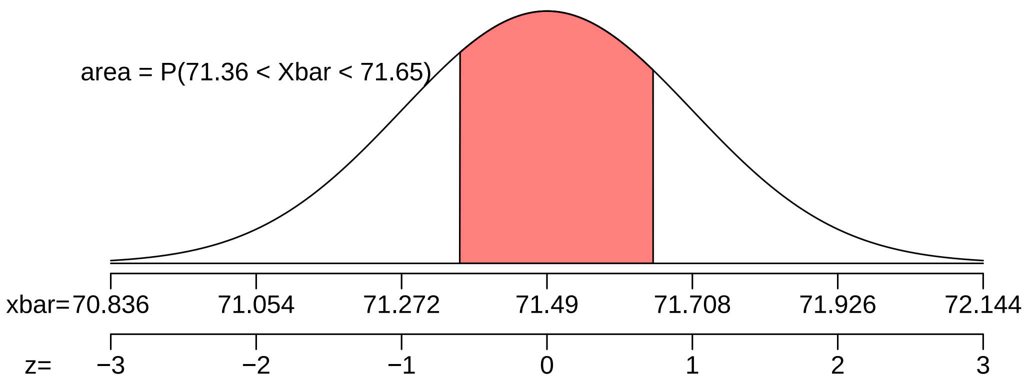
Find the standard error.

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

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.

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

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.

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

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

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

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

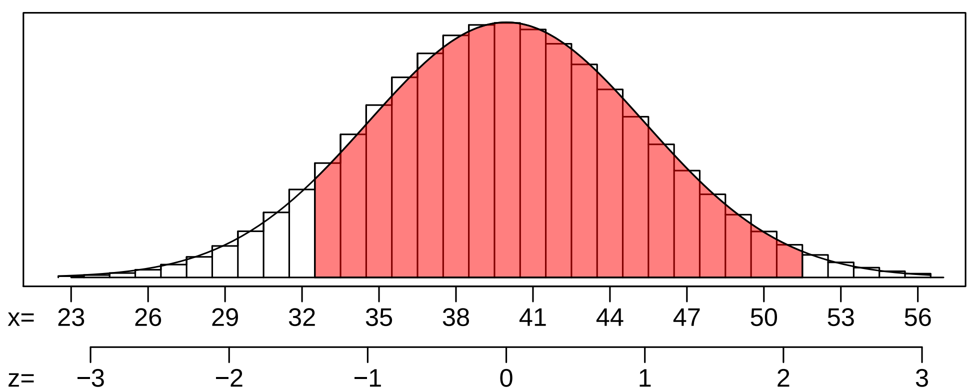
Find the mean.

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

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 σ and we are doing inference so we need to use the t distribution).

$$df = n - 1 = 25$$

Determine the critical t value, $t^{\star}$, such that $P(|T| < t^{\star}) = 0.8$.

$$t^{\star} = 1.32$$

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.

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

Determine the bounds.

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

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 signficance 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 $z^\star$ 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.

$$n_1 = 18$$
$$\bar{x}_1 = 110$$
$$s_1 = 14.7$$
$$n_2 = 30$$
$$\bar{x}_2 = 100$$
$$s_2 = 12.8$$
$$\alpha = 0.02$$
$$df = 32$$

State the hypotheses.

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

We are using a two-tail test. Find $t^\star$ such that $P(|T| > t^\star) = 0.02$ by using a t table.

$$t^\star = 2.45$$

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 $t^\star$.

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

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

$$0.02 < p\text{-value} < 0.04$$

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) $t^{\star} = 2.45$

(d) $SE = 4.179$

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

(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 $z^\star$ such that $P(|Z| < z^\star) = 0.99$.

$$z^\star = 2.58$$

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.

$$ME = z^\star SE = (2.58)(0.00313) = 0.00808$$

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 $z^\star$ 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 $P(|Z| > z^\star) = 0.04$.

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

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.

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

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 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) $z^\star = 2.05$

(d) $SE = 0.0555$

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

(f) p-value = 0.0208

(g) reject the null