Bunker Hill Community College

Final Statistics Exam 2019-05-02 — Exam ID 020

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(**teal**) = 0.205

   (b) P(**horn given red**) = 0.241

   (c) P(**bike or black**) = 0.363

   (d) P(**teal given wheel**) = 0.165

   (e) P(**bike**) = 0.24

   (f) P(**bike and teal**) = 0.0685

2. P(**"gem" given "black"**) = 0.265

3. $P(72.47 < X < 72.81) = 0.7557$

4. (a) $P(X = 20) = 0.1113$

   (b) $P(24 \leq X \leq 29) = 0.2023$

5. $\mathbf{(9.35, 10.4)}$

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

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

   (c) $t^\star = 2.11$

   (d) $SE = 0.009$

   (e) $\mid t_{\textbf{obs}} \mid = 2.21$

   (f) $0.02 < p$-**value** < 0.04

   (g) **reject**

7. (a) **LB of p CI = 0.522 or** 52.2%

   (b) **UB of p CI = 0.548 or** 54.8%

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

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

(c) $z^{\star} = 1.64$

(d) $SE = 0.042$

(e) $\mid z_{\mathbf{obs}} \mid = 1.76$

(f) p-**value** = 0.0784

(g) **reject**

1. In a deck of strange cards, there are 1271 cards. Each card has an image and a color. The amounts are shown in the table below.

	black	green	pink	red	teal
bike	34	85	68	31	87
dog	71	24	47	19	32
flower	20	69	55	88	35
horn	14	26	65	72	63
wheel	51	40	42	89	44

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

(b) What is the probability a random card is a horn given it is red?

(c) What is the probability a random card is either a bike or black (or both)?

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

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

(f) What is the probability a random card is both a bike and teal?

Solution

(a) $P(\text{teal}) = \frac{87+32+35+63+44}{1271} = 0.205$

(b) $P(\text{horn given red}) = \frac{72}{31+19+88+72+89} = 0.241$

(c) $P(\text{bike or black}) = \frac{34+85+68+31+87+34+71+20+14+51-34}{1271} = 0.363$

(d) $P(\text{teal given wheel}) = \frac{44}{51+40+42+89+44} = 0.165$

(e) $P(\text{bike}) = \frac{34+85+68+31+87}{1271} = 0.24$

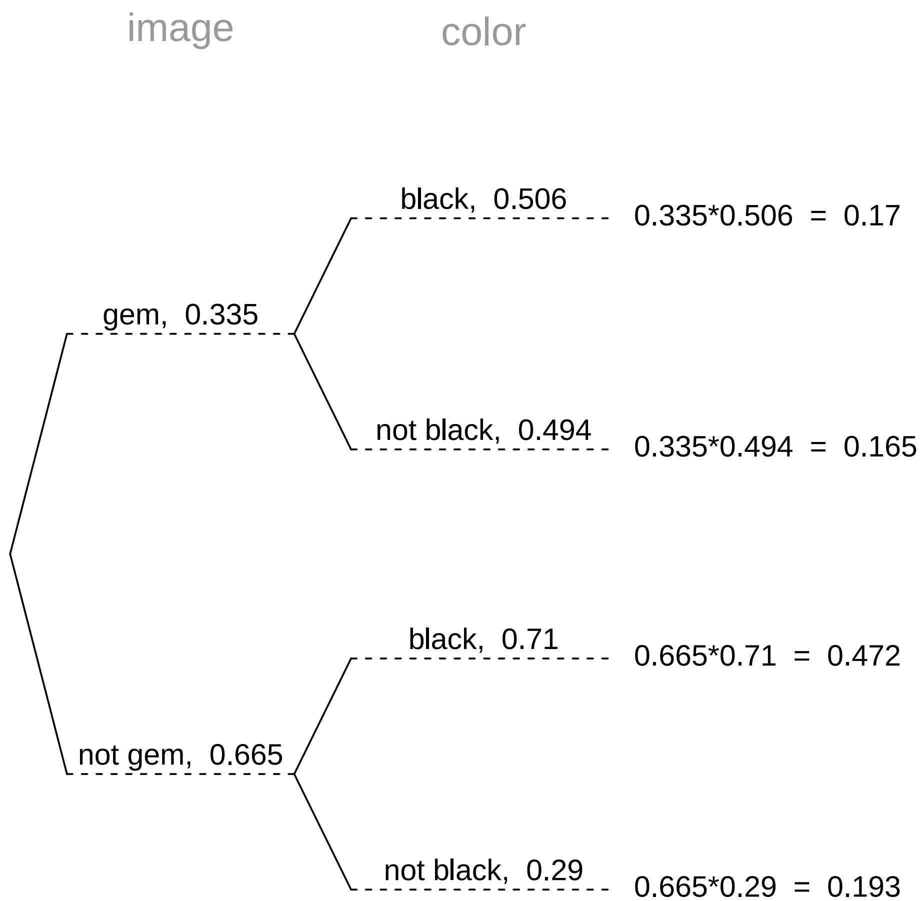
(f) $P(\text{bike and teal}) = \frac{87}{1271} = 0.0685$

2. In a deck of strange cards, each card has an image and a color. The chance of drawing a gem is 33.5%. If a gem is drawn, there is a 50.6% chance that it is black. If a card that is not a gem is drawn, there is a 71% chance that it is black.

   Now, someone draws a random card and reveals it is black. What is the chance the card is a gem?

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.

$$P(\text{"gem" given "black"}) = \frac{0.17}{0.17 + 0.472} = 0.265$$

3. In a very large pile of toothpicks, the mean length is 72.61 millimeters and the standard deviation is 2.14 millimeters. If you randomly sample 225 toothpicks, what is the chance the sample mean is between 72.47 and 72.81 millimeters?

Solution

Label the given information.

$$\mu = 72.61$$
$$\sigma = 2.14$$
$$n = 225$$
$$\bar{x}_{\text{lower}} = 72.47$$
$$\bar{x}_{\text{upper}} = 72.81$$

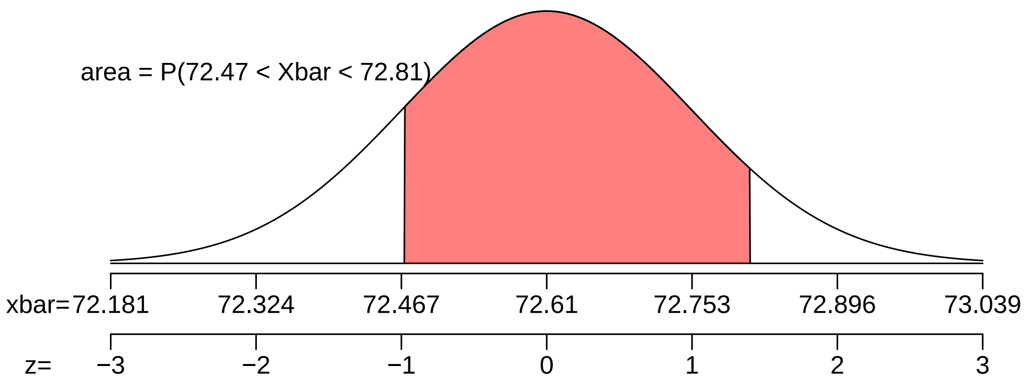
Find the standard error.

$$SE = \frac{\sigma}{\sqrt{n}} = \frac{2.14}{\sqrt{225}} = 0.143$$

Describe the sampling distribution.

$$\bar{X} \sim \mathcal{N}(72.61, 0.143)$$

Draw a sketch.



Calculate a z scores.

$$z_{\text{lower}} = \frac{x_{\text{lower}} - \mu}{SE} = \frac{72.47 - 72.61}{0.143} = -0.98$$

$$z_{\text{upper}} = \frac{x_{\text{upper}} - \mu}{SE} = \frac{72.81 - 72.61}{0.143} = 1.4$$

Determine the probability.

$$\begin{aligned} P(72.47 < X < 72.81) &= \Phi(z_{\text{upper}}) - \Phi(z_{\text{lower}}) \\ &= \Phi(1.4) - \Phi(-0.98) \\ &= 0.7557 \end{aligned}$$

4. In a game, there is a 51% chance to win a round. You will play 42 rounds.

   (a) What is the probability of winning exactly 20 rounds?

   (b) What is the probability of winning at least 24 but at most 29 rounds?

Solution

We use the formula for binomial probabilities.

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

$$P(X = 20) = \binom{42}{20}(0.51)^{20}(1-0.51)^{42-20}$$

$$P(X = 20) = \binom{42}{20}(0.51)^{20}(0.49)^{22}$$

$$P(X = 20) = 0.1113$$

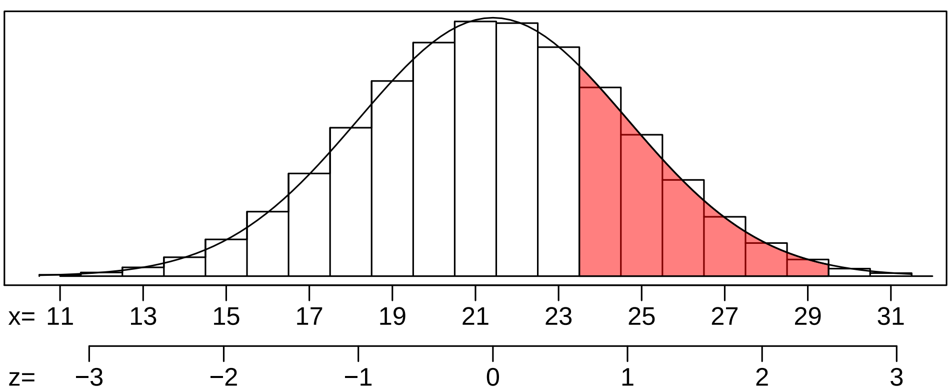
Find the mean.

$$\mu = np = (42)(0.51) = 21.42$$

Find the standard deviation.

$$\sigma = \sqrt{np(1-p)} = \sqrt{(42)(0.51)(1-0.51)} = 3.2397$$

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



Find the z scores.

$$z_1 = \frac{23.5 - 21.42}{3.2397} = 0.8$$

$$z_2 = \frac{29.5 - 21.42}{3.2397} = 2.34$$

Calculate the probability.

$$P(24 \leq X \leq 29) = \Phi(2.34) - \Phi(0.8) = 0.2023$$

(a) $P(X = 20) = 0.1113$

(b) $P(24 \leq X \leq 29) = 0.2023$

5. As an ornithologist, you wish to determine the average body mass of *Dendroica dominica*. You randomly sample 27 adults of *Dendroica dominica*, resulting in a sample mean of 9.87 grams and a sample standard deviation of 1.59 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.

$$\begin{aligned} n &= 27 \\ \bar{x} &= 9.87 \\ s &= 1.59 \\ CL &= 0.9 \end{aligned}$$

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

$$df = n - 1 = 26$$

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

$$t^\star = 1.71$$

Calculate the standard error.

$$SE = \frac{s}{\sqrt{n}} = \frac{1.59}{\sqrt{27}} = 0.306$$

We want to make an inference about the population mean.

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

Determine the bounds.

$$\begin{aligned} CI &= (\bar{x} - t^\star SE, \ \bar{x} + t^\star SE) \\ &= (9.87 - 1.71 \times 0.306, \ 9.87 + 1.71 \times 0.306) \\ &= (9.35, \ 10.4) \end{aligned}$$

We are 90% confident that the population mean is between 9.35 and 10.4.

6. A treatment group of size 26 has a mean of 1 and standard deviation of 0.0315. A control group of size 29 has a mean of 1.02 and standard deviation of 0.0356. 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 52.

   (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.

$$\begin{aligned} n_1 &= 26 \\ \bar{x}_1 &= 1 \\ s_1 &= 0.0315 \\ n_2 &= 29 \\ \bar{x}_2 &= 1.02 \\ s_2 &= 0.0356 \\ \alpha &= 0.04 \\ df &= 52 \end{aligned}$$

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.

$$\begin{aligned} SE &= \sqrt{\frac{(s_1)^2}{n_1} + \frac{(s_2)^2}{n_2}} \\ &= \sqrt{\frac{(0.0315)^2}{26} + \frac{(0.0356)^2}{29}} \\ &= 0.009 \end{aligned}$$

Determine the test statistic.

$$\begin{aligned} t_{\text{obs}} &= \frac{(\bar{x}_2 - \bar{x}_1) - (\mu_2 - \mu_1)_0}{SE} \\ &= \frac{(1.02 - 1) - (0)}{0.009} \\ &= 2.21 \end{aligned}$$

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 reject the null hypothesis.

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

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

(c) $t^\star = 2.11$

(d) $SE = 0.009$

(e) $|t_{\text{obs}}| = 2.21$

(f) $0.02 < p\text{-value} < 0.04$

(g) reject the null

7. From a very large population, a random sample of 4000 individuals was taken. In that sample, 53.5% were broken. 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.

$$SE = \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = \sqrt{\frac{(0.535)(1-0.535)}{4000}} = 0.00789$$

Calculate the margin of error.

$$ME = z^\star SE = (1.64)(0.00789) = 0.0129$$

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.

$$(0.522\,,\ 0.548)$$

We are 90% confident that the true population proportion is between 52.2% and 54.8%.

(a) The lower bound = 0.522, which can also be expressed as 52.2%.

(b) The upper bound = 0.548, which can also be expressed as 54.8%.

8. An experiment is run with a treatment group of size 296 and a control group of size 284. The results are summarized in the table below.

	treatment	control
fluorescent	151	124
not fluorescent	145	160

Using a significance level of 0.1, 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 $z^\star$ 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 $P(|Z| > z^\star) = 0.1$.

$$z^\star = \Phi^{-1}\left(1 - \frac{\alpha}{2}\right) = 1.64$$

Determine the sample proportions.

$$\hat{p}_1 = \frac{151}{296} = 0.51$$
$$\hat{p}_2 = \frac{124}{284} = 0.437$$

Determine the difference of sample proportions.

$$\hat{p}_2 - \hat{p}_1 = 0.437 - 0.51 = -0.073$$

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

$$\hat{p} = \frac{151 + 124}{296 + 284} = 0.474$$

Determine the standard error.

$$\begin{aligned} SE &= \sqrt{\frac{\hat{p}(1-\hat{p})}{n_1} + \frac{\hat{p}(1-\hat{p})}{n_2}} \\ &= \sqrt{\frac{(0.474)(0.526)}{296} + \frac{(0.474)(0.526)}{284}} \\ &= 0.0415 \end{aligned}$$

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

$$H_0: \quad \hat{P}_2 - \hat{P}_1 \sim \mathcal{N}(0,\ 0.0415)$$

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.

$$\begin{aligned} z &= \frac{(\hat{p}_2 - \hat{p}_1) - (p_2 - p_1)_0}{SE} \\ &= \frac{(0.437 - 0.51) - 0}{0.0415} \\ &= -1.76 \end{aligned}$$

Determine the p-value.

$$\begin{aligned} p\text{-value} &= 2 \cdot \Phi(-|z|) \\ &= 2 \cdot \Phi(-1.76) \\ &= 0.0784 \end{aligned}$$

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) $z^\star = 1.64$
(d) $SE = 0.0415$
(e) $|z_{obs}| = 1.76$
(f) p-value = 0.0784
(g) reject the null