Bunker Hill Community College

Final Statistics Exam 2019-05-02 Exam ID 008

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(**black given flower**) = 0.273

   (b) P(**blue**) = 0.15

   (c) P(**bike and teal**) = 0.0735

   (d) P(**bike or gray**) = 0.504

   (e) P(**bike given teal**) = 0.399

   (f) P(**wheel**) = 0.35

2. P(**"pig" given "yellow"**) = 0.246

3. $P(70.9 < X < 71.15) = 0.8892$

4. (a) $P(X = 132) = 0.0604$

   (b) $P(120 \leq X \leq 126) = 0.1968$

5. **(17.7, 19.8)**

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

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

   (c) $t^\star = 2.49$

   (d) $SE = 0.565$

   (e) $\mid t_{\textbf{obs}} \mid = 2.37$

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

   (g) **retain**

7. (a) **LB of p CI = 0.565 or** 56.5%

   (b) **UB of p CI = 0.579 or** 57.9%

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

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

(c) $z^\star = 1.28$

(d) $SE = 0.039$

(e) $\mid z_{\mathbf{obs}} \mid = 1.37$

(f) p-**value** = 0.1706

(g) **reject**

1. In a deck of strange cards, there are 966 cards. Each card has an image and a color. The amounts are shown in the table below.

	black	blue	gray	red	teal
bike	80	25	54	79	71
flower	87	43	86	61	42
wheel	85	77	92	19	65

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

(b) What is the probability a random card is blue?

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

(d) What is the probability a random card is either a bike or gray (or both)?

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

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

Solution

(a) $P(\text{black given flower}) = \frac{87}{87+43+86+61+42} = 0.273$

(b) $P(\text{blue}) = \frac{25+43+77}{966} = 0.15$

(c) $P(\text{bike and teal}) = \frac{71}{966} = 0.0735$

(d) $P(\text{bike or gray}) = \frac{80+25+54+79+71+54+86+92-54}{966} = 0.504$

(e) $P(\text{bike given teal}) = \frac{71}{71+42+65} = 0.399$

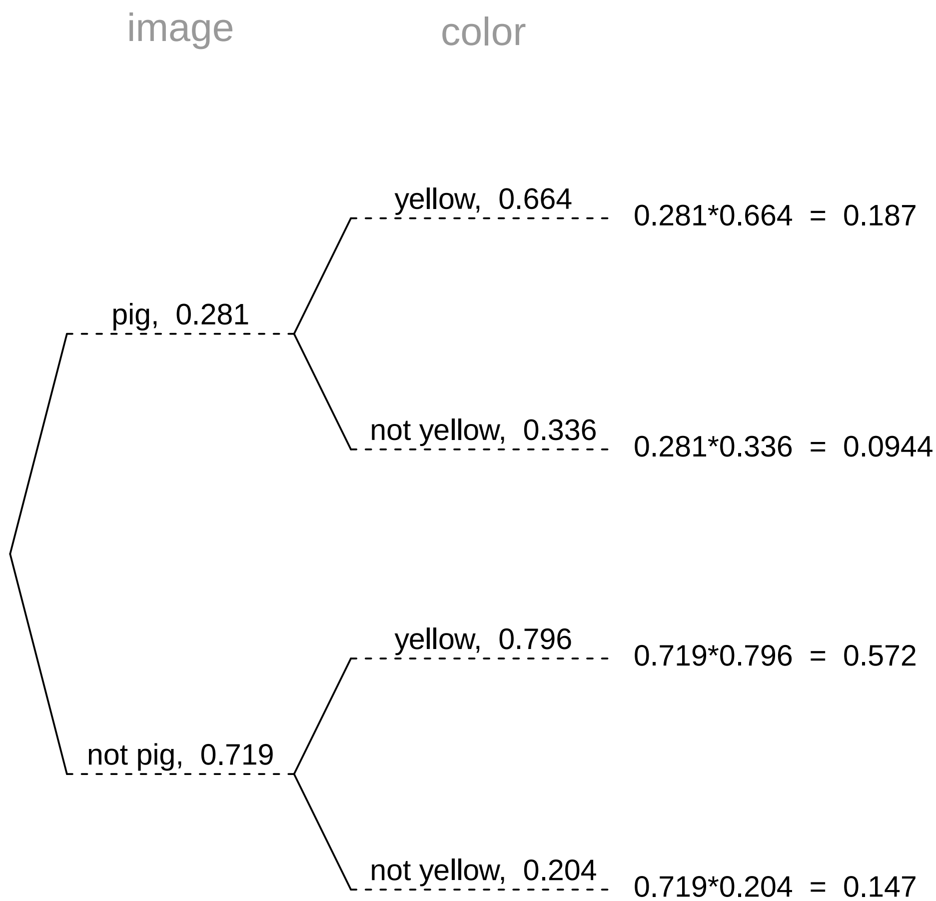
(f) $P(\text{wheel}) = \frac{85+77+92+19+65}{966} = 0.35$

2. In a deck of strange cards, each card has an image and a color. The chance of drawing a pig is 28.1%. If a pig is drawn, there is a 66.4% chance that it is yellow. If a card that is not a pig is drawn, there is a 79.6% chance that it is yellow.

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

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.

$$P(\text{"pig" given "yellow"}) = \frac{0.187}{0.187 + 0.572} = 0.246$$

3. In a very large pile of toothpicks, the mean length is 71.04 millimeters and the standard deviation is 1.09 millimeters. If you randomly sample 200 toothpicks, what is the chance the sample mean is between 70.9 and 71.15 millimeters?

Solution

Label the given information.

$$\mu = 71.04$$
$$\sigma = 1.09$$
$$n = 200$$
$$\bar{x}_{\text{lower}} = 70.9$$
$$\bar{x}_{\text{upper}} = 71.15$$

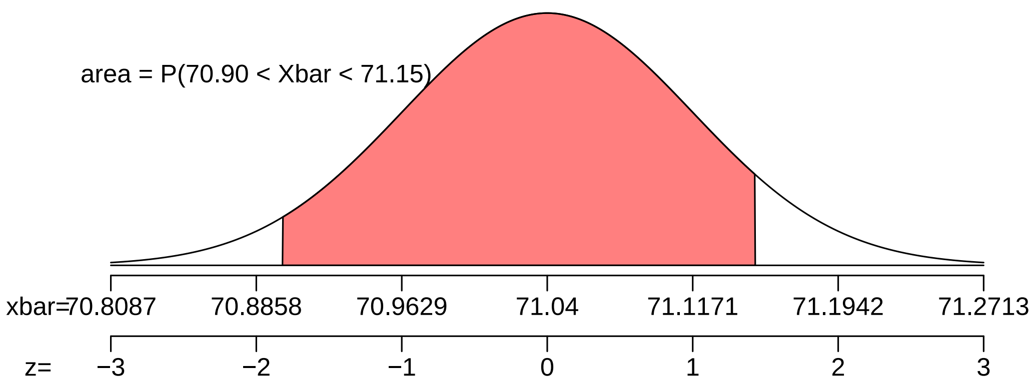
Find the standard error.

$$SE = \frac{\sigma}{\sqrt{n}} = \frac{1.09}{\sqrt{200}} = 0.0771$$

Describe the sampling distribution.

$$\bar{X} \sim \mathcal{N}(71.04,\ 0.0771)$$

Draw a sketch.



Calculate a z scores.

$$z_{\text{lower}} = \frac{x_{\text{lower}} - \mu}{SE} = \frac{70.9 - 71.04}{0.0771} = -1.82$$

$$z_{\text{upper}} = \frac{x_{\text{upper}} - \mu}{SE} = \frac{71.15 - 71.04}{0.0771} = 1.43$$

Determine the probability.

$$\begin{aligned} P(70.9 < X < 71.15) &= \Phi(z_{\text{upper}}) - \Phi(z_{\text{lower}}) \\ &= \Phi(1.43) - \Phi(-1.82) \\ &= 0.8892 \end{aligned}$$

4. In a game, there is a 67% chance to win a round. You will play 196 rounds.

   (a) What is the probability of winning exactly 132 rounds?

   (b) What is the probability of winning at least 120 but at most 126 rounds?

Solution
We use the formula for binomial probabilities.

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

$$P(X = 132) = \binom{196}{132}(0.67)^{132}(1-0.67)^{196-132}$$

$$P(X = 132) = \binom{196}{132}(0.67)^{132}(0.33)^{64}$$

$$P(X = 132) = 0.0604$$

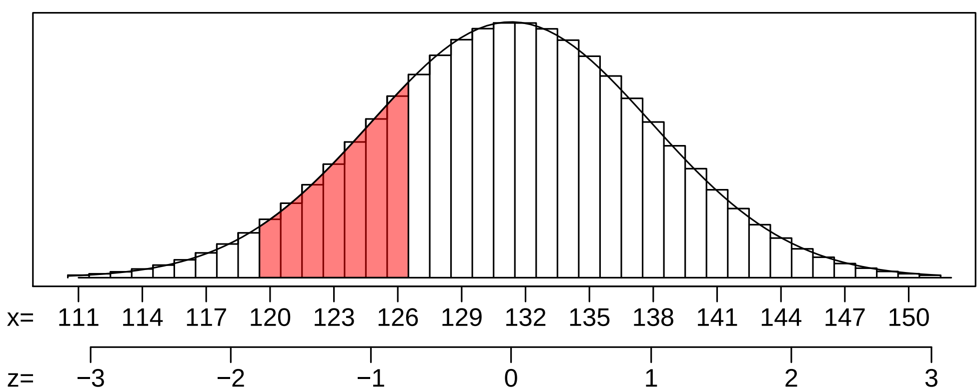
Find the mean.

$$\mu = np = (196)(0.67) = 131.32$$

Find the standard deviation.

$$\sigma = \sqrt{np(1-p)} = \sqrt{(196)(0.67)(1-0.67)} = 6.583$$

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



Find the z scores.

$$z_1 = \frac{119.5 - 131.32}{6.583} = -1.8$$

$$z_2 = \frac{126.5 - 131.32}{6.583} = -0.73$$

Calculate the probability.

$$P(120 \le X \le 126) = \Phi(-0.73) - \Phi(-1.8) = 0.1968$$

(a) $P(X = 132) = 0.0604$

(b) $P(120 \le X \le 126) = 0.1968$

5. As an ornithologist, you wish to determine the average body mass of *Vireo olivaceus*. You randomly sample 36 adults of *Vireo olivaceus*, resulting in a sample mean of 18.75 grams and a sample standard deviation of 3.01 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.

$$\begin{aligned} n &= 36 \\ \bar{x} &= 18.75 \\ s &= 3.01 \\ CL &= 0.95 \end{aligned}$$

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

$$df = n - 1 = 35$$

Determine the critical t value, $t^\star$, such that $P(|T| < t^\star) = 0.95$.

$$t^\star = 2.03$$

Calculate the standard error.

$$SE = \frac{s}{\sqrt{n}} = \frac{3.01}{\sqrt{36}} = 0.502$$

We want to make an inference about the population mean.

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

Determine the bounds.

$$\begin{aligned} CI &= (\bar{x} - t^\star SE,\ \bar{x} + t^\star SE) \\ &= (18.75 - 2.03 \times 0.502,\ 18.75 + 2.03 \times 0.502) \\ &= (17.7,\ 19.8) \end{aligned}$$

We are 95% confident that the population mean is between 17.7 and 19.8.

6. A treatment group of size 27 has a mean of 9.86 and standard deviation of 1.7. A control group of size 13 has a mean of 11.2 and standard deviation of 1.66. 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 24.

   (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.

$$\begin{aligned} n_1 &= 27 \\ \bar{x}_1 &= 9.86 \\ s_1 &= 1.7 \\ n_2 &= 13 \\ \bar{x}_2 &= 11.2 \\ s_2 &= 1.66 \\ \alpha &= 0.02 \\ df &= 24 \end{aligned}$$

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 $P(|T| > t^\star) = 0.02$ by using a t table.

$$t^\star = 2.49$$

Calculate the standard error.

$$\begin{aligned} SE &= \sqrt{\frac{(s_1)^2}{n_1} + \frac{(s_2)^2}{n_2}} \\ &= \sqrt{\frac{(1.7)^2}{27} + \frac{(1.66)^2}{13}} \\ &= 0.565 \end{aligned}$$

Determine the test statistic.

$$\begin{aligned} t_{\text{obs}} &= \frac{(\bar{x}_2 - \bar{x}_1) - (\mu_2 - \mu_1)_0}{SE} \\ &= \frac{(11.2 - 9.86) - (0)}{0.565} \\ &= 2.37 \end{aligned}$$

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

(c) $t^\star = 2.49$

(d) $SE = 0.565$

(e) $|t_{obs}| = 2.37$

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

(g) retain the null

7. From a very large population, a random sample of 40000 individuals was taken. In that sample, 57.2% were floating. Determine a 99.5% 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 $P(|Z| < z^\star) = 0.995$.

$$z^\star = 2.81$$

Calculate the standard error.

$$SE = \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = \sqrt{\frac{(0.572)(1-0.572)}{40000}} = 0.00247$$

Calculate the margin of error.

$$ME = z^\star SE = (2.81)(0.00247) = 0.00694$$

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.

$$(0.565\,,\ 0.579)$$

We are 99.5% confident that the true population proportion is between 56.5% and 57.9%.

(a) The lower bound = 0.565, which can also be expressed as 56.5%.

(b) The upper bound = 0.579, which can also be expressed as 57.9%.

8. An experiment is run with a treatment group of size 276 and a control group of size 270. The results are summarized in the table below.

	treatment	control
reclusive	77	90
not reclusive	199	180

Using a significance level of 0.2, determine whether the treatment causes an effect on the proportion of cases that are reclusive.

(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 $P(|Z| > z^\star) = 0.2$.

$$z^\star = \Phi^{-1}\left(1 - \frac{\alpha}{2}\right) = 1.28$$

Determine the sample proportions.

$$\hat{p}_1 = \frac{77}{276} = 0.279$$
$$\hat{p}_2 = \frac{90}{270} = 0.333$$

Determine the difference of sample proportions.

$$\hat{p}_2 - \hat{p}_1 = 0.333 - 0.279 = 0.054$$

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

$$\hat{p} = \frac{77 + 90}{276 + 270} = 0.306$$

Determine the standard error.

$$\begin{aligned} SE &= \sqrt{\frac{\hat{p}(1-\hat{p})}{n_1} + \frac{\hat{p}(1-\hat{p})}{n_2}} \\ &= \sqrt{\frac{(0.306)(0.694)}{276} + \frac{(0.306)(0.694)}{270}} \\ &= 0.0394 \end{aligned}$$

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

$$H_0: \quad \hat{P}_2 - \hat{P}_1 \sim \mathcal{N}(0,\ 0.0394)$$

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.

$$\begin{aligned} z &= \frac{(\hat{p}_2 - \hat{p}_1) - (p_2 - p_1)_0}{SE} \\ &= \frac{(0.333 - 0.279) - 0}{0.0394} \\ &= 1.37 \end{aligned}$$

Determine the p-value.

$$\begin{aligned} p\text{-value} &= 2 \cdot \Phi(-|z|) \\ &= 2 \cdot \Phi(-1.37) \\ &= 0.1706 \end{aligned}$$

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) $z^{\star} = 1.28$

(d) $SE = 0.0394$

(e) $|z_{\text{obs}}| = 1.37$

(f) p-value = 0.1706

(g) reject the null