Bunker Hill Community College

Final Statistics Exam 2019-05-02 Exam ID 011

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(**tree**) = 0.26

   (b) P(**shovel or indigo**) = 0.327

   (c) P(**shovel and red**) = 0.0816

   (d) P(**yellow given dog**) = 0.0676

   (e) P(**pig given pink**) = 0.2

   (f) P(**pink**) = 0.201

2. P(**"not wheel" given "green"**) = 0.311

3. $P(68.15 < X < 68.51) = 0.7666$

4. (a) $P(X = 25) = 0.0815$

   (b) $P(23 \leq X \leq 34) = 0.8863$

5. **(9.89, 10.2)**

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

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

   (c) $t^\star = 2.66$

   (d) $SE = 3.192$

   (e) $\mid t_{\mathbf{obs}} \mid = 2.51$

   (f) $0.01 <$ p-**value** < 0.02

   (g) **retain**

7. (a) **LB of p CI = 0.872 or** 87.2%

   (b) **UB of p CI = 0.884 or** 88.4%

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) $SE = 0.053$

(e) $\mid z_{\mathbf{obs}} \mid = 2.33$

(f) p-**value** = 0.0198

(g) **reject**

1. In a deck of strange cards, there are 895 cards. Each card has an image and a color. The amounts are shown in the table below.

	green	indigo	pink	red	yellow
dog	16	26	86	65	14
pig	34	18	36	68	90
shovel	31	19	45	73	41
tree	84	40	13	20	76

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

(b) What is the probability a random card is either a shovel or indigo (or both)?

(c) What is the probability a random card is both a shovel and red?

(d) What is the probability a random card is yellow given it is a dog?

(e) What is the probability a random card is a pig given it is pink?

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

Solution

(a) $P(\text{tree}) = \frac{84+40+13+20+76}{895} = 0.26$

(b) $P(\text{shovel or indigo}) = \frac{31+19+45+73+41+26+18+19+40-19}{895} = 0.327$

(c) $P(\text{shovel and red}) = \frac{73}{895} = 0.0816$

(d) $P(\text{yellow given dog}) = \frac{14}{16+26+86+65+14} = 0.0676$

(e) $P(\text{pig given pink}) = \frac{36}{86+36+45+13} = 0.2$

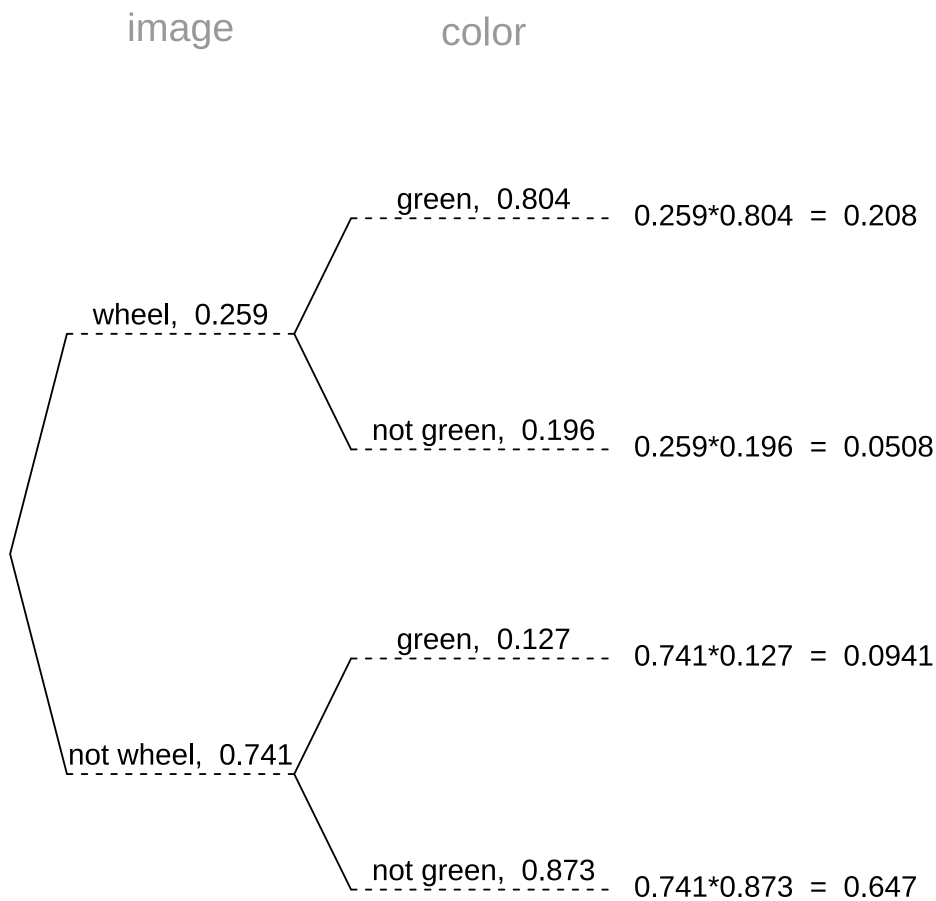
(f) $P(\text{pink}) = \frac{86+36+45+13}{895} = 0.201$

2. In a deck of strange cards, each card has an image and a color. The chance of drawing a wheel is 25.9%. If a wheel is drawn, there is a 80.4% chance that it is green. If a card that is not a wheel is drawn, there is a 12.7% chance that it is green.

   Now, someone draws a random card and reveals it is green. What is the chance the card is not a wheel?

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.

$$P(\text{"not wheel" given "green"}) = \frac{0.0941}{0.0941 + 0.208} = 0.311$$

3. In a very large pile of toothpicks, the mean length is 68.4 millimeters and the standard deviation is 1.99 millimeters. If you randomly sample 225 toothpicks, what is the chance the sample mean is between 68.15 and 68.51 millimeters?

Solution

Label the given information.

$$\mu = 68.4$$
$$\sigma = 1.99$$
$$n = 225$$
$$\bar{x}_{\text{lower}} = 68.15$$
$$\bar{x}_{\text{upper}} = 68.51$$

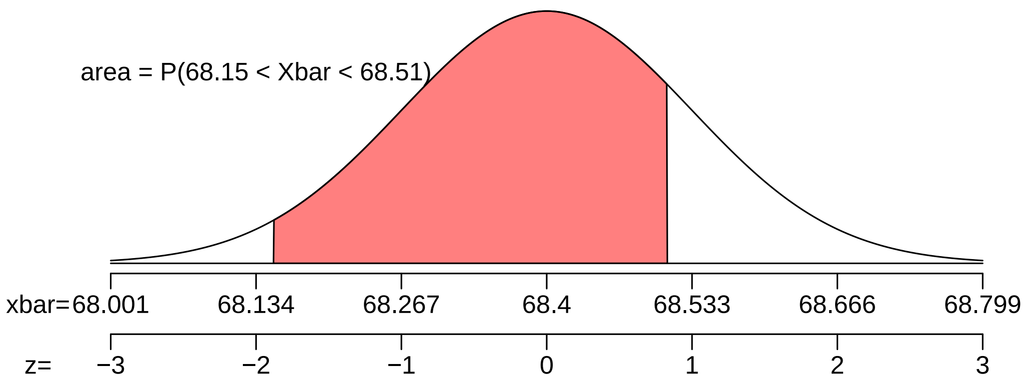
Find the standard error.

$$SE = \frac{\sigma}{\sqrt{n}} = \frac{1.99}{\sqrt{225}} = 0.133$$

Describe the sampling distribution.

$$\bar{X} \sim \mathcal{N}(68.4, 0.133)$$

Draw a sketch.



Calculate a z scores.

$$z_{\text{lower}} = \frac{x_{\text{lower}} - \mu}{SE} = \frac{68.15 - 68.4}{0.133} = -1.88$$

$$z_{\text{upper}} = \frac{x_{\text{upper}} - \mu}{SE} = \frac{68.51 - 68.4}{0.133} = 0.83$$

Determine the probability.

$$\begin{aligned} P(68.15 < X < 68.51) &= \Phi(z_{\text{upper}}) - \Phi(z_{\text{lower}}) \\ &= \Phi(0.83) - \Phi(-1.88) \\ &= 0.7666 \end{aligned}$$

4. In a game, there is a 58% chance to win a round. You will play 48 rounds.
   (a) What is the probability of winning exactly 25 rounds?
   (b) What is the probability of winning at least 23 but at most 34 rounds?

Solution
We use the formula for binomial probabilities.

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

$$P(X = 25) = \binom{48}{25}(0.58)^{25}(1-0.58)^{48-25}$$

$$P(X = 25) = \binom{48}{25}(0.58)^{25}(0.42)^{23}$$

$$P(X = 25) = 0.0815$$

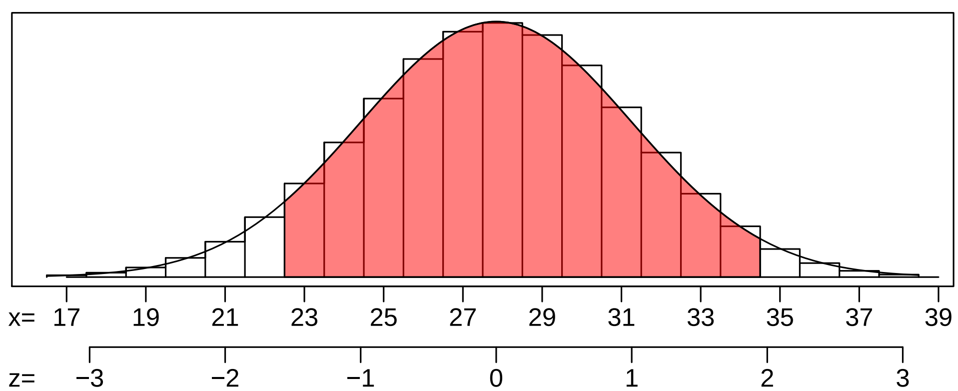
Find the mean.

$$\mu = np = (48)(0.58) = 27.84$$

Find the standard deviation.

$$\sigma = \sqrt{np(1-p)} = \sqrt{(48)(0.58)(1-0.58)} = 3.4195$$

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



Find the z scores.

$$z_1 = \frac{22.5 - 27.84}{3.4195} = -1.42$$

$$z_2 = \frac{34.5 - 27.84}{3.4195} = 1.8$$

Calculate the probability.

$$P(23 \leq X \leq 34) = \Phi(1.8) - \Phi(-1.42) = 0.8863$$

(a) $P(X = 25) = 0.0815$

(b) $P(23 \leq X \leq 34) = 0.8863$

5. As an ornithologist, you wish to determine the average body mass of *Vireo griseus*. You randomly sample 33 adults of *Vireo griseus*, resulting in a sample mean of 10.04 grams and a sample standard deviation of 0.653 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.

$$n = 33$$
$$\bar{x} = 10.04$$
$$s = 0.653$$
$$CL = 0.8$$

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

$$df = n - 1 = 32$$

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

$$t^\star = 1.31$$

Calculate the standard error.

$$SE = \frac{s}{\sqrt{n}} = \frac{0.653}{\sqrt{33}} = 0.114$$

We want to make an inference about the population mean.

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

Determine the bounds.

$$\begin{aligned} CI &= (\bar{x} - t^\star SE,\ \bar{x} + t^\star SE) \\ &= (10.04 - 1.31 \times 0.114,\ 10.04 + 1.31 \times 0.114) \\ &= (9.89,\ 10.2) \end{aligned}$$

We are 80% confident that the population mean is between 9.89 and 10.2.

6. A treatment group of size 33 has a mean of 109 and standard deviation of 14. A control group of size 29 has a mean of 101 and standard deviation of 11.1. If you decided to use a signficance level of 0.01, 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 59.

   (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.

$$
\begin{aligned}
n_1 &= 33\\
\bar{x}_1 &= 109\\
s_1 &= 14\\
n_2 &= 29\\
\bar{x}_2 &= 101\\
s_2 &= 11.1\\
\alpha &= 0.01\\
df &= 59
\end{aligned}
$$

State the hypotheses.

$$
\begin{aligned}
H_0 &: \quad \mu_2 - \mu_1 = 0\\
H_A &: \quad \mu_2 - \mu_1 \neq 0
\end{aligned}
$$

We are using a two-tail test. Find $t^\star$ such that $P(|T| > t^\star) = 0.01$ by using a t table.

$$t^\star = 2.66$$

Calculate the standard error.

$$
\begin{aligned}
SE &= \sqrt{\frac{(s_1)^2}{n_1} + \frac{(s_2)^2}{n_2}}\\
&= \sqrt{\frac{(14)^2}{33} + \frac{(11.1)^2}{29}}\\
&= 3.192
\end{aligned}
$$

Determine the test statistic.

$$
\begin{aligned}
t_{\text{obs}} &= \frac{(\bar{x}_2 - \bar{x}_1) - (\mu_2 - \mu_1)_0}{SE}\\
&= \frac{(101 - 109) - (0)}{3.192}\\
&= -2.51
\end{aligned}
$$

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.

$$0.01 < p\text{-value} < 0.02$$

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) $t^{\star} = 2.66$

(d) $SE = 3.192$

(e) $|t_{\text{obs}}| = 2.51$

(f) $0.01 < p\text{-value} < 0.02$

(g) retain the null

7. From a very large population, a random sample of 5400 individuals was taken. In that sample, 87.8% were tasty. Determine a 80% 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 $P(|Z| < z^\star) = 0.8$.

$$z^\star = 1.28$$

Calculate the standard error.

$$SE = \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = \sqrt{\frac{(0.878)(1-0.878)}{5400}} = 0.00445$$

Calculate the margin of error.

$$ME = z^\star SE = (1.28)(0.00445) = 0.0057$$

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.

$$(0.872\,,\ 0.884)$$

We are 80% confident that the true population proportion is between 87.2% and 88.4%.

(a) The lower bound = 0.872, which can also be expressed as 87.2%.

(b) The upper bound = 0.884, which can also be expressed as 88.4%.

8. An experiment is run with a treatment group of size 165 and a control group of size 173. The results are summarized in the table below.

	treatment	control
folksy	110	94
not folksy	55	79

Using a significance level of 0.04, determine whether the treatment causes an effect on the proportion of cases that are folksy.

(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.

$$\hat{p}_1 = \frac{110}{165} = 0.667$$
$$\hat{p}_2 = \frac{94}{173} = 0.543$$

Determine the difference of sample proportions.

$$\hat{p}_2 - \hat{p}_1 = 0.543 - 0.667 = -0.124$$

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

$$\hat{p} = \frac{110 + 94}{165 + 173} = 0.604$$

Determine the standard error.

$$\begin{aligned} SE &= \sqrt{\frac{\hat{p}(1-\hat{p})}{n_1} + \frac{\hat{p}(1-\hat{p})}{n_2}} \\ &= \sqrt{\frac{(0.604)(0.396)}{165} + \frac{(0.604)(0.396)}{173}} \\ &= 0.0532 \end{aligned}$$

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

$$H_0: \quad \hat{P}_2 - \hat{P}_1 \sim \mathcal{N}(0,\ 0.0532)$$

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.

$$\begin{aligned} z &= \frac{(\hat{p}_2 - \hat{p}_1) - (p_2 - p_1)_0}{SE} \\ &= \frac{(0.543 - 0.667) - 0}{0.0532} \\ &= -2.33 \end{aligned}$$

Determine the p-value.

$$\begin{aligned} p\text{-value} &= 2 \cdot \Phi(-|z|) \\ &= 2 \cdot \Phi(-2.33) \\ &= 0.0198 \end{aligned}$$

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) $SE = 0.0532$
(e) $|z_{\text{obs}}| = 2.33$
(f) p-value = 0.0198
(g) reject the null