Bunker Hill Community College

Final Statistics Exam 2019-05-02 Exam ID 026

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) $P(\textbf{flower or indigo}) = 0.44$

   (b) $P(\textbf{indigo given flower}) = 0.278$

   (c) $P(\textbf{flower}) = 0.303$

   (d) $P(\textbf{flower and gray}) = 0.0266$

   (e) $P(\textbf{pig given green}) = 0.165$

   (f) $P(\textbf{green}) = 0.147$

2. $P(\textbf{"not horn" given "pink"}) = 0.83$

3. $P(62.06 < X < 62.47) = 0.8183$

4. (a) $P(X = 68) = 0.0508$

   (b) $P(61 \leq X \leq 82) = 0.8702$

5. **(9.44, 10.4)**

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

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

   (c) $t^{\star} = 2.74$

   (d) $SE = 0.054$

   (e) $\mid t_{\textbf{obs}} \mid = 2.84$

   (f) $0.005 < p\textbf{-value} < 0.01$

   (g) **reject**

7. (a) **LB of p CI = 0.677 or** 67.7%

   (b) **UB of p CI = 0.699 or** 69.9%

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

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

(c) $z^{\star} = 2.33$

(d) $SE = 0.083$

(e) $\mid z_{\mathbf{obs}} \mid = 2.47$

(f) p-**value** $= 0.0136$

(g) **reject**

1. In a deck of strange cards, there are 902 cards. Each card has an image and a color. The amounts are shown in the table below.

	blue	gray	green	indigo	teal
flower	65	24	48	76	60
pig	44	54	22	37	92
shovel	90	98	63	87	42

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

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

   (c) What is the probability a random card is a flower?

   (d) What is the probability a random card is both a flower and gray?

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

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

Solution

(a) $P(\text{flower or indigo}) = \frac{65+24+48+76+60+76+37+87-76}{902} = 0.44$

(b) $P(\text{indigo given flower}) = \frac{76}{65+24+48+76+60} = 0.278$

(c) $P(\text{flower}) = \frac{65+24+48+76+60}{902} = 0.303$

(d) $P(\text{flower and gray}) = \frac{24}{902} = 0.0266$

(e) $P(\text{pig given green}) = \frac{22}{48+22+63} = 0.165$

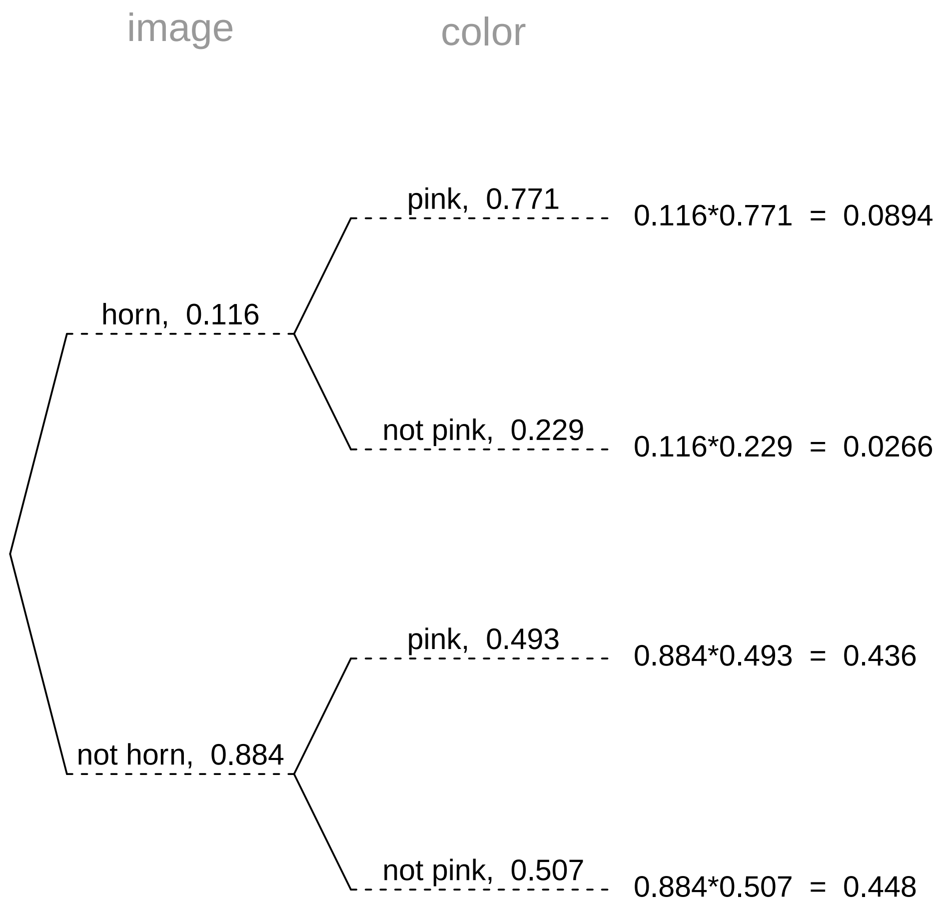
(f) $P(\text{green}) = \frac{48+22+63}{902} = 0.147$

2. In a deck of strange cards, each card has an image and a color. The chance of drawing a horn is 11.6%. If a horn is drawn, there is a 77.1% chance that it is pink. If a card that is not a horn is drawn, there is a 49.3% chance that it is pink.

   Now, someone draws a random card and reveals it is pink. 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.

$$P(\text{"not horn" given "pink"}) = \frac{0.436}{0.436 + 0.0894} = 0.83$$

3. In a very large pile of toothpicks, the mean length is 62.17 millimeters and the standard deviation is 1.32 millimeters. If you randomly sample 125 toothpicks, what is the chance the sample mean is between 62.06 and 62.47 millimeters?

Solution

Label the given information.

$$\mu = 62.17$$
$$\sigma = 1.32$$
$$n = 125$$
$$\bar{x}_{\text{lower}} = 62.06$$
$$\bar{x}_{\text{upper}} = 62.47$$

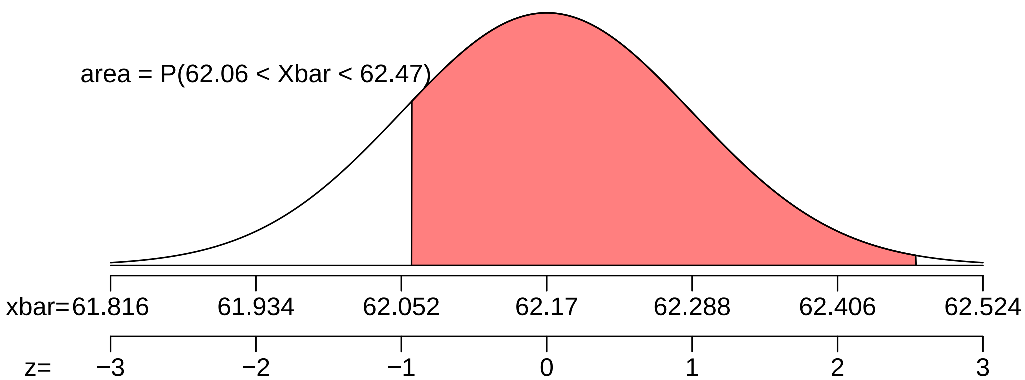
Find the standard error.

$$SE = \frac{\sigma}{\sqrt{n}} = \frac{1.32}{\sqrt{125}} = 0.118$$

Describe the sampling distribution.

$$\bar{X} \sim \mathcal{N}(62.17, 0.118)$$

Draw a sketch.



Calculate a z scores.

$$z_{\text{lower}} = \frac{x_{\text{lower}} - \mu}{SE} = \frac{62.06 - 62.17}{0.118} = -0.93$$

$$z_{\text{upper}} = \frac{x_{\text{upper}} - \mu}{SE} = \frac{62.47 - 62.17}{0.118} = 2.54$$

Determine the probability.

$$\begin{aligned} P(62.06 < X < 62.47) &= \Phi(z_{\text{upper}}) - \Phi(z_{\text{lower}}) \\ &= \Phi(2.54) - \Phi(-0.93) \\ &= 0.8183 \end{aligned}$$

4. In a game, there is a 33% chance to win a round. You will play 217 rounds.

   (a) What is the probability of winning exactly 68 rounds?

   (b) What is the probability of winning at least 61 but at most 82 rounds?

Solution

We use the formula for binomial probabilities.

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

$$P(X = 68) = \binom{217}{68}(0.33)^{68}(1 - 0.33)^{217-68}$$

$$P(X = 68) = \binom{217}{68}(0.33)^{68}(0.67)^{149}$$

$$P(X = 68) = 0.0508$$

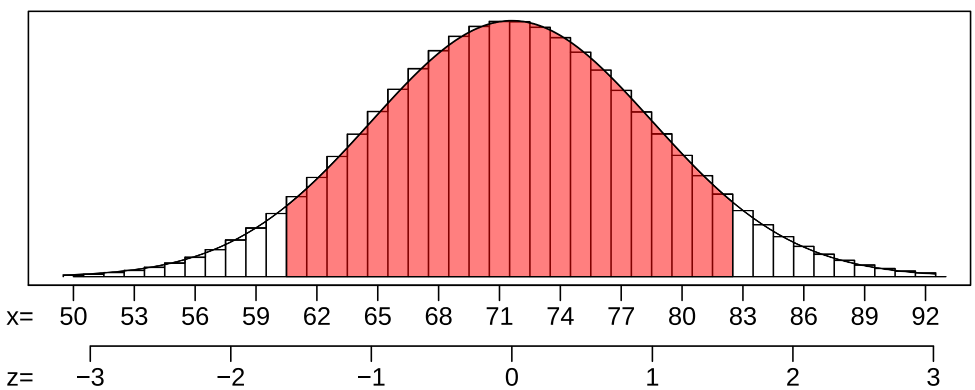
Find the mean.

$$\mu = np = (217)(0.33) = 71.61$$

Find the standard deviation.

$$\sigma = \sqrt{np(1 - p)} = \sqrt{(217)(0.33)(1 - 0.33)} = 6.9267$$

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



Find the z scores.

$$z_1 = \frac{60.5 - 71.61}{6.9267} = -1.53$$

$$z_2 = \frac{82.5 - 71.61}{6.9267} = 1.5$$

Calculate the probability.

$$P(61 \leq X \leq 82) = \Phi(1.5) - \Phi(-1.53) = 0.8702$$

(a) $P(X = 68) = 0.0508$

(b) $P(61 \leq X \leq 82) = 0.8702$

5. As an ornithologist, you wish to determine the average body mass of *Cistothorus palustris*. You randomly sample 26 adults of *Cistothorus palustris*, resulting in a sample mean of 9.9 grams and a sample standard deviation of 1.14 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.

$$\begin{aligned} n &= 26 \\ \bar{x} &= 9.9 \\ s &= 1.14 \\ CL &= 0.95 \end{aligned}$$

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

$$df = n - 1 = 25$$

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

$$t^\star = 2.06$$

Calculate the standard error.

$$SE = \frac{s}{\sqrt{n}} = \frac{1.14}{\sqrt{26}} = 0.224$$

We want to make an inference about the population mean.

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

Determine the bounds.

$$\begin{aligned} CI &= (\bar{x} - t^\star SE, \ \bar{x} + t^\star SE) \\ &= (9.9 - 2.06 \times 0.224, \ 9.9 + 2.06 \times 0.224) \\ &= (9.44, \ 10.4) \end{aligned}$$

We are 95% confident that the population mean is between 9.44 and 10.4.

6. A treatment group of size 19 has a mean of 1.1 and standard deviation of 0.141. A control group of size 18 has a mean of 0.947 and standard deviation of 0.183. 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 31.

   (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.

$$
\begin{aligned}
n_1 &= 19 \\
\bar{x}_1 &= 1.1 \\
s_1 &= 0.141 \\
n_2 &= 18 \\
\bar{x}_2 &= 0.947 \\
s_2 &= 0.183 \\
\alpha &= 0.01 \\
df &= 31
\end{aligned}
$$

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.

$$t^\star = 2.74$$

Calculate the standard error.

$$
\begin{aligned}
SE &= \sqrt{\frac{(s_1)^2}{n_1} + \frac{(s_2)^2}{n_2}} \\
&= \sqrt{\frac{(0.141)^2}{19} + \frac{(0.183)^2}{18}} \\
&= 0.054
\end{aligned}
$$

Determine the test statistic.

$$
\begin{aligned}
t_{\text{obs}} &= \frac{(\bar{x}_2 - \bar{x}_1) - (\mu_2 - \mu_1)_0}{SE} \\
&= \frac{(0.947 - 1.1) - (0)}{0.054} \\
&= -2.84
\end{aligned}
$$

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.

$$0.005 < p\text{-value} < 0.01$$

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) $t^{\star} = 2.74$

(d) $SE = 0.054$

(e) $|t_{obs}| = 2.84$

(f) $0.005 < p\text{-value} < 0.01$

(g) reject the null

7. From a very large population, a random sample of 6700 individuals was taken. In that sample, 68.8% were blue. Determine a 95% 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 $P(|Z| < z^\star) = 0.95$.

$$z^\star = 1.96$$

Calculate the standard error.

$$SE = \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = \sqrt{\frac{(0.688)(1-0.688)}{6700}} = 0.00566$$

Calculate the margin of error.

$$ME = z^\star SE = (1.96)(0.00566) = 0.0111$$

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.

$$(0.677\,,\ 0.699)$$

We are 95% confident that the true population proportion is between 67.7% and 69.9%.

(a) The lower bound = 0.677, which can also be expressed as 67.7%.

(b) The upper bound = 0.699, which can also be expressed as 69.9%.

8. An experiment is run with a treatment group of size 24 and a control group of size 68. The results are summarized in the table below.

	treatment	control
organic	7	6
not organic	17	62

Using a significance level of 0.02, determine whether the treatment causes an effect on the proportion of cases that are organic.

(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 $P(|Z| > z^\star) = 0.02$.

$$z^\star = \Phi^{-1}\left(1 - \frac{\alpha}{2}\right) = 2.33$$

Determine the sample proportions.

$$\hat{p}_1 = \frac{7}{24} = 0.292$$
$$\hat{p}_2 = \frac{6}{68} = 0.0882$$

Determine the difference of sample proportions.

$$\hat{p}_2 - \hat{p}_1 = 0.0882 - 0.292 = -0.2038$$

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

$$\hat{p} = \frac{7+6}{24+68} = 0.141$$

Determine the standard error.

$$\begin{aligned} SE &= \sqrt{\frac{\hat{p}(1-\hat{p})}{n_1} + \frac{\hat{p}(1-\hat{p})}{n_2}} \\ &= \sqrt{\frac{(0.141)(0.859)}{24} + \frac{(0.141)(0.859)}{68}} \\ &= 0.0826 \end{aligned}$$

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

$$H_0: \quad \hat{P}_2 - \hat{P}_1 \sim \mathcal{N}(0,\ 0.0826)$$

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.

$$\begin{aligned} z &= \frac{(\hat{p}_2 - \hat{p}_1) - (p_2 - p_1)_0}{SE} \\ &= \frac{(0.0882 - 0.292) - 0}{0.0826} \\ &= -2.47 \end{aligned}$$

Determine the p-value.

$$\begin{aligned} p\text{-value} &= 2 \cdot \Phi(-|z|) \\ &= 2 \cdot \Phi(-2.47) \\ &= 0.0136 \end{aligned}$$

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) $z^\star = 2.33$

(d) $SE = 0.0826$

(e) $|z_{\text{obs}}| = 2.47$

(f) p-value = 0.0136

(g) reject the null