Bunker Hill Community College

Final Statistics Exam 2019-05-02 Exam ID 003

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) $P(\textbf{gray}) = 0.277$

   (b) $P(\textbf{flower}) = 0.314$

   (c) $P(\textbf{cat and yellow}) = 0.0432$

   (d) $P(\textbf{shovel given pink}) = 0.152$

   (e) $P(\textbf{green given pig}) = 0.47$

   (f) $P(\textbf{pig or yellow}) = 0.345$

2. $P(\textbf{"pig" given "not teal"}) = 0.101$

3. $P(69.26 < X < 70.28) = 0.8288$

4. (a) $P(X = 55) = 0.0437$

   (b) $P(49 \leq X \leq 66) = 0.7549$

5. $\mathbf{(19.3, 22)}$

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

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

   (c) $t^\star = 2.68$

   (d) $SE = 0.058$

   (e) $\mid t_{\textbf{obs}} \mid = 2.76$

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

   (g) **reject**

7. (a) **LB of p CI = 0.257 or** 25.7%

   (b) **UB of p CI = 0.271 or** 27.1%

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

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

(c) $z^\star = 1.96$

(d) $SE = 0.107$

(e) $\mid z_{\mathbf{obs}} \mid = 2.02$

(f) p-**value** = 0.0434

(g) **reject**

1. In a deck of strange cards, there are 1158 cards. Each card has an image and a color. The amounts are shown in the table below.

	gray	green	pink	yellow
cat	94	46	36	50
flower	85	96	88	95
gem	74	75	21	89
pig	13	63	44	14
shovel	55	54	34	32

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

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

(c) What is the probability a random card is both a cat and yellow?

(d) What is the probability a random card is a shovel given it is pink?

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

(f) What is the probability a random card is either a pig or yellow (or both)?

Solution

(a) $P(\text{gray}) = \frac{94+85+74+13+55}{1158} = 0.277$

(b) $P(\text{flower}) = \frac{85+96+88+95}{1158} = 0.314$

(c) $P(\text{cat and yellow}) = \frac{50}{1158} = 0.0432$

(d) $P(\text{shovel given pink}) = \frac{34}{36+88+21+44+34} = 0.152$

(e) $P(\text{green given pig}) = \frac{63}{13+63+44+14} = 0.47$

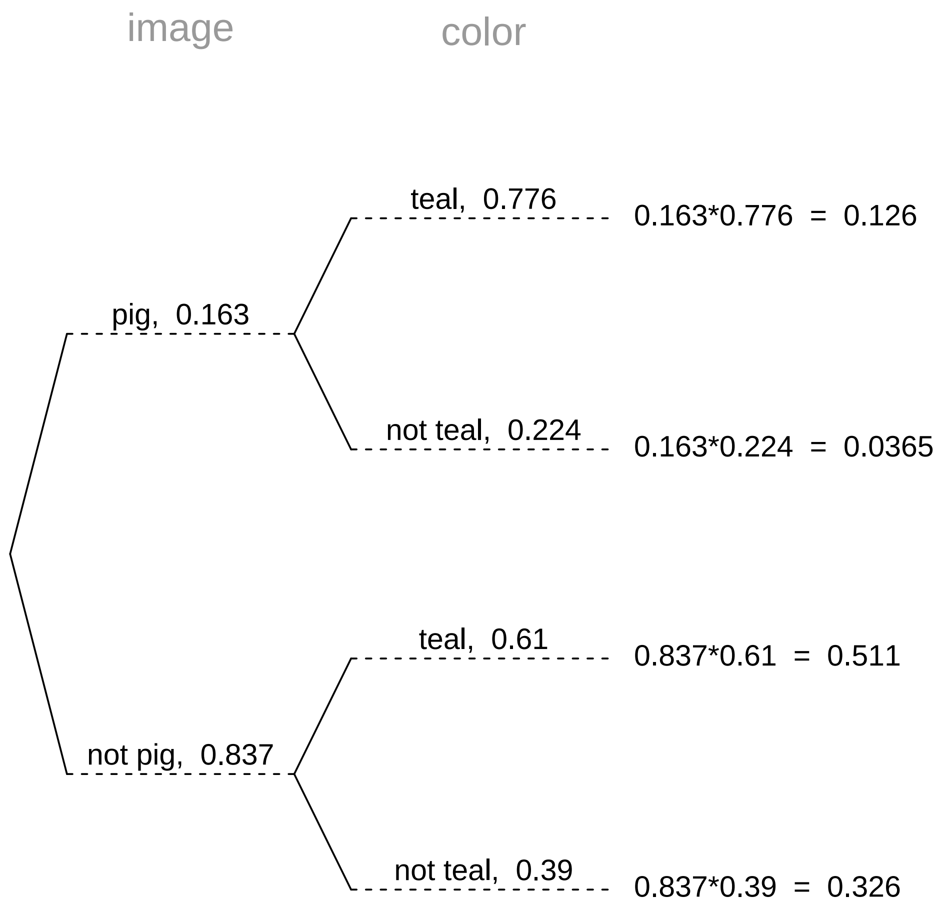
(f) $P(\text{pig or yellow}) = \frac{13+63+44+14+50+95+89+14+32-14}{1158} = 0.345$

2. In a deck of strange cards, each card has an image and a color. The chance of drawing a pig is 16.3%. If a pig is drawn, there is a 77.6% chance that it is teal. If a card that is not a pig is drawn, there is a 61% chance that it is teal.

   Now, someone draws a random card and reveals it is not teal. 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.

$$P(\text{"pig" given "not teal"}) = \frac{0.0365}{0.0365 + 0.326} = 0.101$$

3. In a very large pile of toothpicks, the mean length is 69.6 millimeters and the standard deviation is 3.62 millimeters. If you randomly sample 120 toothpicks, what is the chance the sample mean is between 69.26 and 70.28 millimeters?

Solution

Label the given information.

$$\mu = 69.6$$
$$\sigma = 3.62$$
$$n = 120$$
$$\bar{x}_{\text{lower}} = 69.26$$
$$\bar{x}_{\text{upper}} = 70.28$$

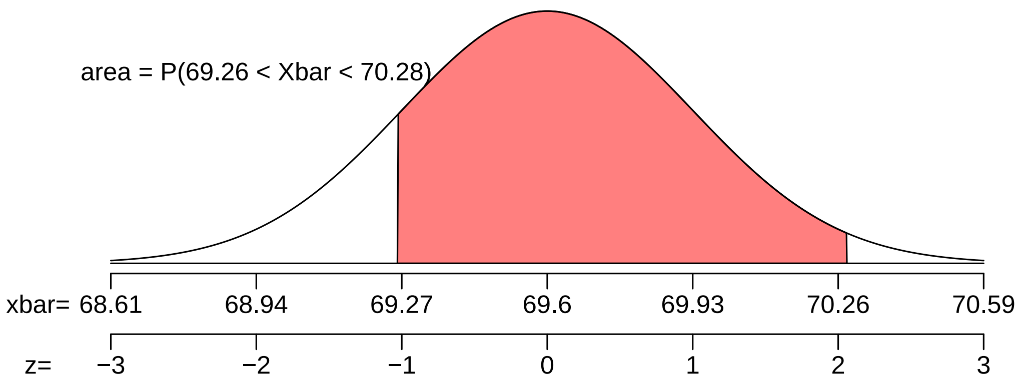
Find the standard error.

$$SE = \frac{\sigma}{\sqrt{n}} = \frac{3.62}{\sqrt{120}} = 0.33$$

Describe the sampling distribution.

$$\bar{X} \sim \mathcal{N}(69.6,\ 0.33)$$

Draw a sketch.



Calculate a z scores.

$$z_{\text{lower}} = \frac{x_{\text{lower}} - \mu}{SE} = \frac{69.26 - 69.6}{0.33} = -1.03$$

$$z_{\text{upper}} = \frac{x_{\text{upper}} - \mu}{SE} = \frac{70.28 - 69.6}{0.33} = 2.06$$

Determine the probability.

$$\begin{aligned} P(69.26 < X < 70.28) &= \Phi(z_{\text{upper}}) - \Phi(z_{\text{lower}}) \\ &= \Phi(2.06) - \Phi(-1.03) \\ &= 0.8288 \end{aligned}$$

4. In a game, there is a 27% chance to win a round. You will play 224 rounds.
   (a) What is the probability of winning exactly 55 rounds?
   (b) What is the probability of winning at least 49 but at most 66 rounds?

Solution

We use the formula for binomial probabilities.

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

$$P(X = 55) = \binom{224}{55}(0.27)^{55}(1-0.27)^{224-55}$$

$$P(X = 55) = \binom{224}{55}(0.27)^{55}(0.73)^{169}$$

$$P(X = 55) = 0.0437$$

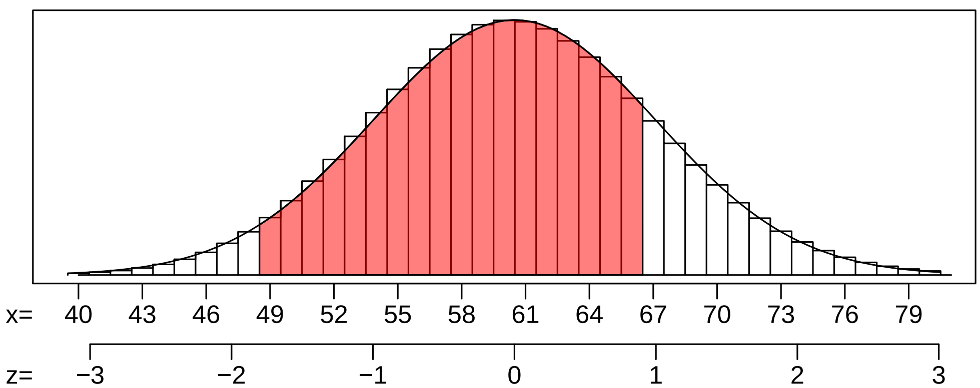
Find the mean.

$$\mu = np = (224)(0.27) = 60.48$$

Find the standard deviation.

$$\sigma = \sqrt{np(1-p)} = \sqrt{(224)(0.27)(1-0.27)} = 6.6446$$

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



Find the z scores.

$$z_1 = \frac{48.5 - 60.48}{6.6446} = -1.73$$

$$z_2 = \frac{66.5 - 60.48}{6.6446} = 0.83$$

Calculate the probability.

$$P(49 \leq X \leq 66) = \Phi(0.83) - \Phi(-1.73) = 0.7549$$

(a) $P(X = 55) = 0.0437$

(b) $P(49 \leq X \leq 66) = 0.7549$

5. As an ornithologist, you wish to determine the average body mass of *Seiurus noveboracensis*. You randomly sample 20 adults of *Seiurus noveboracensis*, resulting in a sample mean of 20.67 grams and a sample standard deviation of 3.48 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.

$$\begin{aligned} n &= 20 \\ \bar{x} &= 20.67 \\ s &= 3.48 \\ CL &= 0.9 \end{aligned}$$

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

$$df = n - 1 = 19$$

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

$$t^\star = 1.73$$

Calculate the standard error.

$$SE = \frac{s}{\sqrt{n}} = \frac{3.48}{\sqrt{20}} = 0.778$$

We want to make an inference about the population mean.

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

Determine the bounds.

$$\begin{aligned} CI &= (\bar{x} - t^\star SE,\ \bar{x} + t^\star SE) \\ &= (20.67 - 1.73 \times 0.778,\ 20.67 + 1.73 \times 0.778) \\ &= (19.3,\ 22) \end{aligned}$$

We are 90% confident that the population mean is between 19.3 and 22.

6. A treatment group of size 28 has a mean of 1.21 and standard deviation of 0.245. A control group of size 23 has a mean of 1.05 and standard deviation of 0.168. 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 47.

   (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.

$$\begin{aligned} n_1 &= 28 \\ \bar{x}_1 &= 1.21 \\ s_1 &= 0.245 \\ n_2 &= 23 \\ \bar{x}_2 &= 1.05 \\ s_2 &= 0.168 \\ \alpha &= 0.01 \\ df &= 47 \end{aligned}$$

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 $P(|T| > t^\star) = 0.01$ by using a t table.

$$t^\star = 2.68$$

Calculate the standard error.

$$\begin{aligned} SE &= \sqrt{\frac{(s_1)^2}{n_1} + \frac{(s_2)^2}{n_2}} \\ &= \sqrt{\frac{(0.245)^2}{28} + \frac{(0.168)^2}{23}} \\ &= 0.058 \end{aligned}$$

Determine the test statistic.

$$\begin{aligned} t_{\text{obs}} &= \frac{(\bar{x}_2 - \bar{x}_1) - (\mu_2 - \mu_1)_0}{SE} \\ &= \frac{(1.05 - 1.21) - (0)}{0.058} \\ &= -2.76 \end{aligned}$$

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) $t^\star = 2.68$

(d) $SE = 0.058$

(e) $|t_{obs}| = 2.76$

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

(g) reject the null

7. From a very large population, a random sample of 7300 individuals was taken. In that sample, 26.4% were bitter. 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.

$$SE = \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = \sqrt{\frac{(0.264)(1-0.264)}{7300}} = 0.00516$$

Calculate the margin of error.

$$ME = z^{\star}SE = (1.28)(0.00516) = 0.0066$$

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.

$$(0.257\,,\ 0.271)$$

We are 80% confident that the true population proportion is between 25.7% and 27.1%.

(a) The lower bound = 0.257, which can also be expressed as 25.7%.

(b) The upper bound = 0.271, which can also be expressed as 27.1%.

8. An experiment is run with a treatment group of size 27 and a control group of size 39. The results are summarized in the table below.

	treatment	control
special	17	33
not special	10	6

Using a significance level of 0.05, determine whether the treatment causes an effect on the proportion of cases that are special.

(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 $P(|Z| > z^\star) = 0.05$.

$$z^\star = \Phi^{-1}\left(1 - \frac{\alpha}{2}\right) = 1.96$$

Determine the sample proportions.

$$\hat{p}_1 = \frac{17}{27} = 0.63$$

$$\hat{p}_2 = \frac{33}{39} = 0.846$$

Determine the difference of sample proportions.

$$\hat{p}_2 - \hat{p}_1 = 0.846 - 0.63 = 0.216$$

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

$$\hat{p} = \frac{17 + 33}{27 + 39} = 0.758$$

Determine the standard error.

$$SE = \sqrt{\frac{\hat{p}(1-\hat{p})}{n_1} + \frac{\hat{p}(1-\hat{p})}{n_2}}$$

$$= \sqrt{\frac{(0.758)(0.242)}{27} + \frac{(0.758)(0.242)}{39}}$$

$$= 0.107$$

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

$$H_0 : \quad \hat{P}_2 - \hat{P}_1 \sim \mathcal{N}(0, 0.107)$$

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.

$$\begin{aligned} z &= \frac{(\hat{p}_2 - \hat{p}_1) - (p_2 - p_1)_0}{SE} \\ &= \frac{(0.846 - 0.63) - 0}{0.107} \\ &= 2.02 \end{aligned}$$

Determine the p-value.

$$\begin{aligned} p\text{-value} &= 2 \cdot \Phi(-|z|) \\ &= 2 \cdot \Phi(-2.02) \\ &= 0.0434 \end{aligned}$$

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) $H_0: \ \ p_2 - p_1 = 0$

(b) $H_A: \ \ p_2 - p_1 \neq 0$

(c) $z^\star = 1.96$

(d) $SE = 0.107$

(e) $|z_{\text{obs}}| = 2.02$

(f) p-value = 0.0434

(g) reject the null