Bunker Hill Community College

Final Statistics Exam 2019-05-02 — Exam ID 001

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(**wheel given red**) = 0.61

   (b) P(**wheel or gray**) = 0.499

   (c) P(**violet given wheel**) = 0.103

   (d) P(**orange**) = 0.217

   (e) P(**cat**) = 0.296

   (f) P(**cat and orange**) = 0.0628

2. P(**"not kite" given "gray"**) = 0.881

3. $P(68.39 < X < 68.91) = 0.7856$

4. (a) $P(X = 19) = 0.1032$

   (b) $P(13 \leq X \leq 25) = 0.8758$

5. **(17.1, 20.4)**

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

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

   (c) $t^{\star} = 2.72$

   (d) $SE = 8.005$

   (e) $\mid t_{\textbf{obs}} \mid = 2.5$

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

   (g) **retain**

7. (a) **LB of p CI = 0.709 or** 70.9%

   (b) **UB of p CI = 0.763 or** 76.3%

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

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

(c) $z^{\star} = 2.58$

(d) $SE = 0.029$

(e) $| \; z_{\mathbf{obs}} \; | = 2.76$

(f) p-**value** $= 0.0058$

(g) **reject**

1. In a deck of strange cards, there are 764 cards. Each card has an image and a color. The amounts are shown in the table below.

	black	gray	orange	red	violet
cat	45	97	48	19	17
dog	82	32	54	27	91
wheel	43	47	64	72	26

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

(b) What is the probability a random card is either a wheel or gray (or both)?

(c) What is the probability a random card is violet given it is a wheel?

(d) What is the probability a random card is orange?

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

(f) What is the probability a random card is both a cat and orange?

Solution

(a) $P(\text{wheel given red}) = \frac{72}{19+27+72} = 0.61$

(b) $P(\text{wheel or gray}) = \frac{43+47+64+72+26+97+32+47-47}{764} = 0.499$

(c) $P(\text{violet given wheel}) = \frac{26}{43+47+64+72+26} = 0.103$

(d) $P(\text{orange}) = \frac{48+54+64}{764} = 0.217$

(e) $P(\text{cat}) = \frac{45+97+48+19+17}{764} = 0.296$

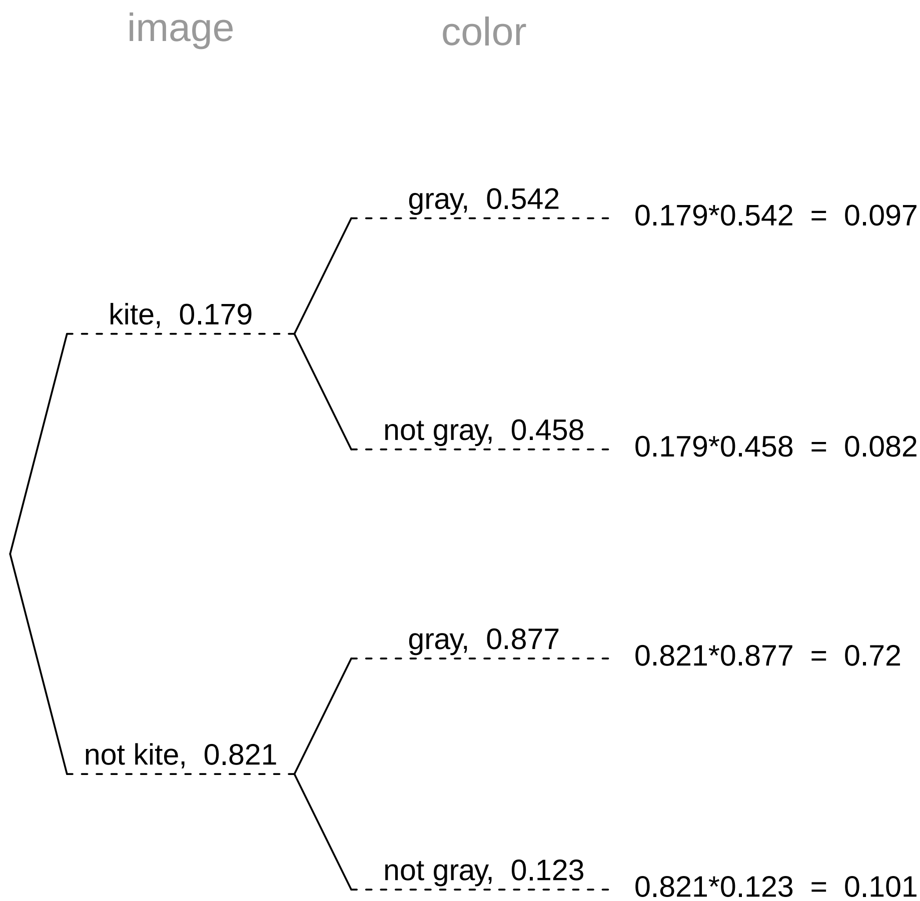
(f) $P(\text{cat and orange}) = \frac{48}{764} = 0.0628$

2. In a deck of strange cards, each card has an image and a color. The chance of drawing a kite is 17.9%. If a kite is drawn, there is a 54.2% chance that it is gray. If a card that is not a kite is drawn, there is a 87.7% chance that it is gray.

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

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.

$$P(\text{"not kite" given "gray"}) = \frac{0.72}{0.72 + 0.097} = 0.881$$

3. In a very large pile of toothpicks, the mean length is 68.61 millimeters and the standard deviation is 2.92 millimeters. If you randomly sample 200 toothpicks, what is the chance the sample mean is between 68.39 and 68.91 millimeters?

Solution

Label the given information.

$$\mu = 68.61$$
$$\sigma = 2.92$$
$$n = 200$$
$$\bar{x}_{\text{lower}} = 68.39$$
$$\bar{x}_{\text{upper}} = 68.91$$

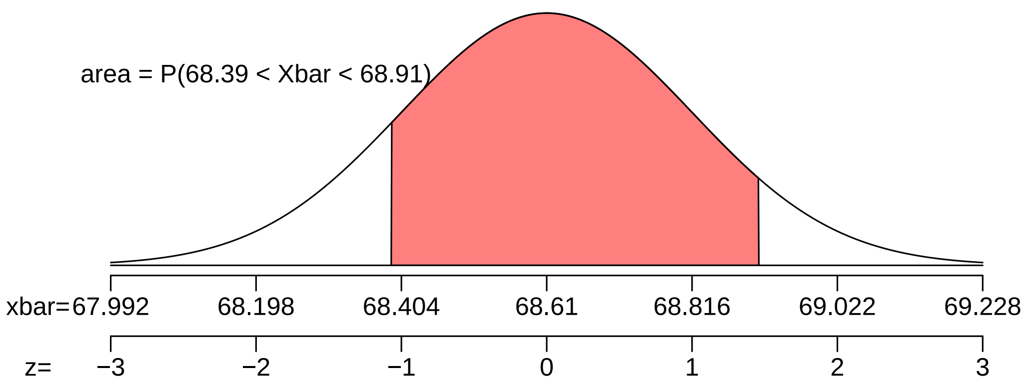
Find the standard error.

$$SE = \frac{\sigma}{\sqrt{n}} = \frac{2.92}{\sqrt{200}} = 0.206$$

Describe the sampling distribution.

$$\bar{X} \sim \mathcal{N}(68.61,\ 0.206)$$

Draw a sketch.



Calculate a z scores.

$$z_{\text{lower}} = \frac{x_{\text{lower}} - \mu}{SE} = \frac{68.39 - 68.61}{0.206} = -1.07$$

$$z_{\text{upper}} = \frac{x_{\text{upper}} - \mu}{SE} = \frac{68.91 - 68.61}{0.206} = 1.46$$

Determine the probability.

$$\begin{aligned} P(68.39 < X < 68.91) &= \Phi(z_{\text{upper}}) - \Phi(z_{\text{lower}}) \\ &= \Phi(1.46) - \Phi(-1.07) \\ &= 0.7856 \end{aligned}$$

4. In a game, there is a 31% chance to win a round. You will play 65 rounds.
   (a) What is the probability of winning exactly 19 rounds?
   (b) What is the probability of winning at least 13 but at most 25 rounds?

Solution

We use the formula for binomial probabilities.

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

$$P(X = 19) = \binom{65}{19}(0.31)^{19}(1-0.31)^{65-19}$$

$$P(X = 19) = \binom{65}{19}(0.31)^{19}(0.69)^{46}$$

$$P(X = 19) = 0.1032$$

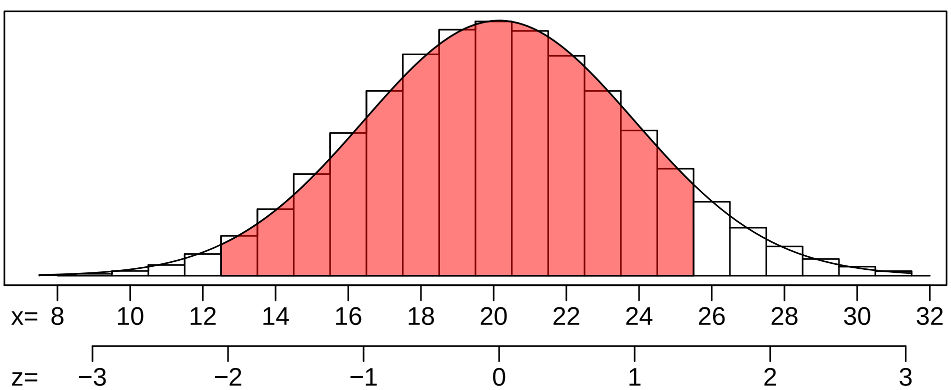
Find the mean.

$$\mu = np = (65)(0.31) = 20.15$$

Find the standard deviation.

$$\sigma = \sqrt{np(1-p)} = \sqrt{(65)(0.31)(1-0.31)} = 3.7287$$

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



Find the z scores.

$$z_1 = \frac{12.5 - 20.15}{3.7287} = -1.92$$

$$z_2 = \frac{25.5 - 20.15}{3.7287} = 1.3$$

Calculate the probability.

$$P(13 \leq X \leq 25) = \Phi(1.3) - \Phi(-1.92) = 0.8758$$

(a) $P(X = 19) = 0.1032$

(b) $P(13 \leq X \leq 25) = 0.8758$

5. As an ornithologist, you wish to determine the average body mass of *Vireo olivaceus*. You randomly sample 20 adults of *Vireo olivaceus*, resulting in a sample mean of 18.75 grams and a sample standard deviation of 2.6 grams. Determine a 99% confidence interval of the true population mean.

Solution

We are given the sample size, sample mean, sample standard deviation, and confidence level.

$$\begin{aligned} n &= 20 \\ \bar{x} &= 18.75 \\ s &= 2.6 \\ CL &= 0.99 \end{aligned}$$

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 $P(|T| < t^{\star}) = 0.99$.

$$t^{\star} = 2.86$$

Calculate the standard error.

$$SE = \frac{s}{\sqrt{n}} = \frac{2.6}{\sqrt{20}} = 0.581$$

We want to make an inference about the population mean.

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

Determine the bounds.

$$\begin{aligned} CI &= (\bar{x} - t^{\star} SE,\ \bar{x} + t^{\star} SE) \\ &= (18.75 - 2.86 \times 0.581,\ 18.75 + 2.86 \times 0.581) \\ &= (17.1,\ 20.4) \end{aligned}$$

We are 99% confident that the population mean is between 17.1 and 20.4.

6. A treatment group of size 15 has a mean of 123 and standard deviation of 22.6. A control group of size 32 has a mean of 103 and standard deviation of 31. 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 36.

   (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.

$$\begin{aligned} n_1 &= 15 \\ \bar{x}_1 &= 123 \\ s_1 &= 22.6 \\ n_2 &= 32 \\ \bar{x}_2 &= 103 \\ s_2 &= 31 \\ \alpha &= 0.01 \\ df &= 36 \end{aligned}$$

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.

$$t^\star = 2.72$$

Calculate the standard error.

$$\begin{aligned} SE &= \sqrt{\frac{(s_1)^2}{n_1} + \frac{(s_2)^2}{n_2}} \\ &= \sqrt{\frac{(22.6)^2}{15} + \frac{(31)^2}{32}} \\ &= 8.005 \end{aligned}$$

Determine the test statistic.

$$\begin{aligned} t_{\text{obs}} &= \frac{(\bar{x}_2 - \bar{x}_1) - (\mu_2 - \mu_1)_0}{SE} \\ &= \frac{(103 - 123) - (0)}{8.005} \\ &= -2.5 \end{aligned}$$

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) $t^\star = 2.72$

(d) $SE = 8.005$

(e) $|t_{\text{obs}}| = 2.5$

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

(g) retain the null

7. From a very large population, a random sample of 1100 individuals was taken. In that sample, 73.6% were special. Determine a 96% 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 $P(|Z| < z^{\star}) = 0.96$.

$$z^{\star} = 2.05$$

Calculate the standard error.

$$SE = \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = \sqrt{\frac{(0.736)(1-0.736)}{1100}} = 0.0133$$

Calculate the margin of error.

$$ME = z^{\star} SE = (2.05)(0.0133) = 0.0273$$

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.

$$(0.709\,,\ 0.763)$$

We are 96% confident that the true population proportion is between 70.9% and 76.3%.

(a) The lower bound = 0.709, which can also be expressed as 70.9%.

(b) The upper bound = 0.763, which can also be expressed as 76.3%.

8. An experiment is run with a treatment group of size 147 and a control group of size 170. The results are summarized in the table below.

	treatment	control
green	130	164
not green	17	6

Using a significance level of 0.01, determine whether the treatment causes an effect on the proportion of cases that are green.

(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 $P(|Z| > z^\star) = 0.01$.

$$z^\star = \Phi^{-1}\left(1 - \frac{\alpha}{2}\right) = 2.58$$

Determine the sample proportions.

$$\hat{p}_1 = \frac{130}{147} = 0.884$$
$$\hat{p}_2 = \frac{164}{170} = 0.965$$

Determine the difference of sample proportions.

$$\hat{p}_2 - \hat{p}_1 = 0.965 - 0.884 = 0.081$$

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

$$\hat{p} = \frac{130 + 164}{147 + 170} = 0.927$$

Determine the standard error.

$$\begin{aligned} SE &= \sqrt{\frac{\hat{p}(1-\hat{p})}{n_1} + \frac{\hat{p}(1-\hat{p})}{n_2}} \\ &= \sqrt{\frac{(0.927)(0.073)}{147} + \frac{(0.927)(0.073)}{170}} \\ &= 0.0293 \end{aligned}$$

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

$$H_0: \quad \hat{P}_2 - \hat{P}_1 \sim \mathcal{N}(0,\ 0.0293)$$

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.

$$\begin{aligned} z &= \frac{(\hat{p}_2 - \hat{p}_1) - (p_2 - p_1)_0}{SE} \\ &= \frac{(0.965 - 0.884) - 0}{0.0293} \\ &= 2.76 \end{aligned}$$

Determine the p-value.

$$\begin{aligned} p\text{-value} &= 2 \cdot \Phi(-|z|) \\ &= 2 \cdot \Phi(-2.76) \\ &= 0.0058 \end{aligned}$$

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) $z^\star = 2.58$

(d) $SE = 0.0293$

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

(f) p-value = 0.0058

(g) reject the null