

Bunker Hill Community College

Final Statistics Exam 2019-05-02

Exam ID 010

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) $P(\text{green given wheel}) = 0.341$

(b) $P(\text{wheel or pink}) = 0.431$

(c) $P(\text{bike}) = 0.304$

(d) $P(\text{wheel and gray}) = 0.0169$

(e) $P(\text{indigo}) = 0.18$

(f) $P(\text{shovel given pink}) = 0.27$

2. $P(\text{"kite" given "not indigo"}) = 0.45$

3. $P(71.51 < X < 71.95) = 0.5183$

4. (a) $P(X = 83) = 0.0583$

(b) $P(70 \leq X \leq 77) = 0.0986$

5. **(64.3, 75.1)**

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

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

(c) $t^* = 2.2$

(d) $SE = 28.699$

(e) $|t_{\text{obs}}| = 2.09$

(f) $0.05 < p\text{-value} < 0.1$

(g) **retain**

7. (a) **LB of p CI = 0.671 or 67.1%**

(b) **UB of p CI = 0.705 or 70.5%**

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

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

(c) $z^* = 2.33$

(d) $SE = 0.028$

(e) $|z_{\text{obs}}| = 2.61$

(f) $p\text{-value} = 0.009$

(g) **reject**

1. In a deck of strange cards, there are 947 cards. Each card has an image and a color. The amounts are shown in the table below.

	gray	green	indigo	pink
bike	81	87	32	88
horn	53	80	44	25
shovel	39	58	65	78
wheel	16	74	29	98

- (a) What is the probability a random card is green given it is a wheel?
- (b) What is the probability a random card is either a wheel or pink (or both)?
- (c) What is the probability a random card is a bike?
- (d) What is the probability a random card is both a wheel and gray?
- (e) What is the probability a random card is indigo?
- (f) What is the probability a random card is a shovel given it is pink?

Solution

$$(a) P(\text{green given wheel}) = \frac{74}{16+74+29+98} = 0.341$$

$$(b) P(\text{wheel or pink}) = \frac{16+74+29+98+88+25+78+98-98}{947} = 0.431$$

$$(c) P(\text{bike}) = \frac{81+87+32+88}{947} = 0.304$$

$$(d) P(\text{wheel and gray}) = \frac{16}{947} = 0.0169$$

$$(e) P(\text{indigo}) = \frac{32+44+65+29}{947} = 0.18$$

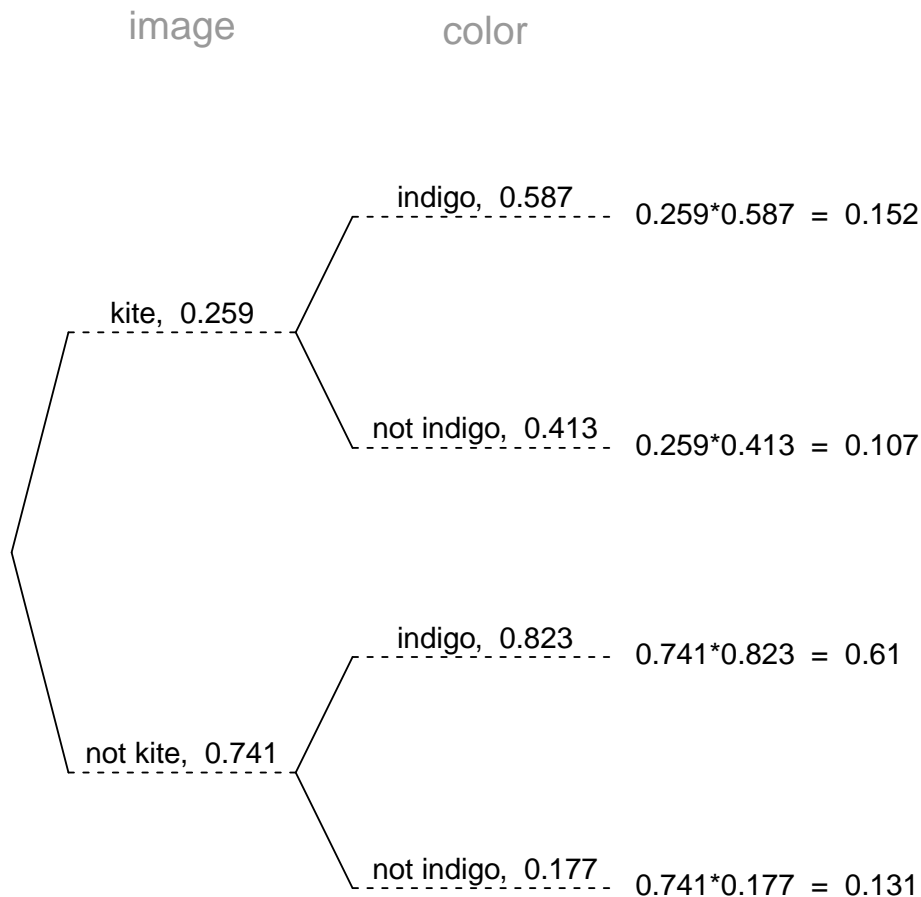
$$(f) P(\text{shovel given pink}) = \frac{78}{88+25+78+98} = 0.27$$

2. In a deck of strange cards, each card has an image and a color. The chance of drawing a kite is 25.9%. If a kite is drawn, there is a 58.7% chance that it is indigo. If a card that is not a kite is drawn, there is a 82.3% chance that it is indigo.

Now, someone draws a random card and reveals it is not indigo. What is the chance the card is 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 probabilities. The results (products) are joint probabilities.



Determine the appropriate conditional probability.

$$P(\text{"kite" given "not indigo"}) = \frac{0.107}{0.107 + 0.131} = 0.45$$

3. In a very large pile of toothpicks, the mean length is 71.75 millimeters and the standard deviation is 3.42 millimeters. If you randomly sample 120 toothpicks, what is the chance the sample mean is between 71.51 and 71.95 millimeters?

Solution

Label the given information.

$$\mu = 71.75$$

$$\sigma = 3.42$$

$$n = 120$$

$$\bar{x}_{\text{lower}} = 71.51$$

$$\bar{x}_{\text{upper}} = 71.95$$

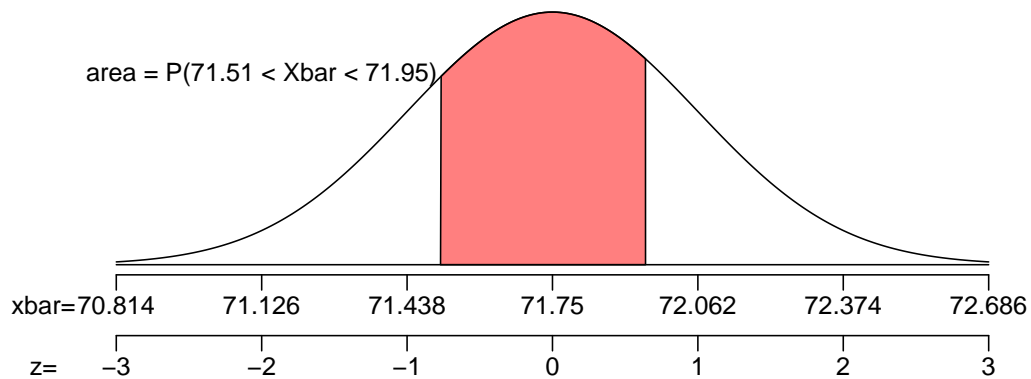
Find the standard error.

$$SE = \frac{\sigma}{\sqrt{n}} = \frac{3.42}{\sqrt{120}} = 0.312$$

Describe the sampling distribution.

$$\bar{X} \sim \mathcal{N}(71.75, 0.312)$$

Draw a sketch.



Calculate a z scores.

$$z_{\text{lower}} = \frac{\bar{x}_{\text{lower}} - \mu}{SE} = \frac{71.51 - 71.75}{0.312} = -0.77$$

$$z_{\text{upper}} = \frac{\bar{x}_{\text{upper}} - \mu}{SE} = \frac{71.95 - 71.75}{0.312} = 0.64$$

Determine the probability.

$$\begin{aligned} P(71.51 < \bar{X} < 71.95) &= \Phi(z_{\text{upper}}) - \Phi(z_{\text{lower}}) \\ &= \Phi(0.64) - \Phi(-0.77) \\ &= 0.5183 \end{aligned}$$

4. In a game, there is a 50% chance to win a round. You will play 170 rounds.
- (a) What is the probability of winning exactly 83 rounds?
 - (b) What is the probability of winning at least 70 but at most 77 rounds?

Solution

We use the formula for binomial probabilities.

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

$$P(X = 83) = \binom{170}{83} (0.5)^{83} (1 - 0.5)^{170-83}$$

$$P(X = 83) = \binom{170}{83} (0.5)^{83} (0.5)^{87}$$

$$P(X = 83) = 0.0583$$

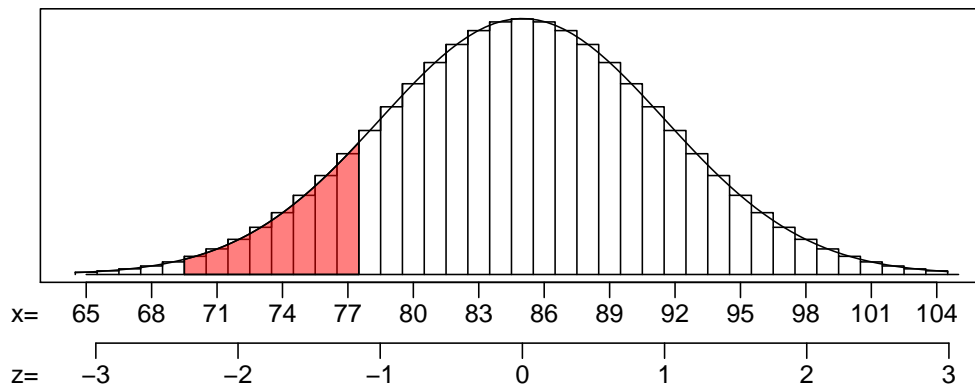
Find the mean.

$$\mu = np = (170)(0.5) = 85$$

Find the standard deviation.

$$\sigma = \sqrt{np(1 - p)} = \sqrt{(170)(0.5)(1 - 0.5)} = 6.5192$$

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



Find the z scores.

$$z_1 = \frac{69.5 - 85}{6.5192} = -2.3$$

$$z_2 = \frac{77.5 - 85}{6.5192} = -1.23$$

Calculate the probability.

$$P(70 \leq X \leq 77) = \Phi(-1.23) - \Phi(-2.3) = 0.0986$$

(a) $P(X = 83) = 0.0583$

(b) $P(70 \leq X \leq 77) = 0.0986$

5. As an ornithologist, you wish to determine the average body mass of *Porzana carolina*. You randomly sample 27 adults of *Porzana carolina*, resulting in a sample mean of 69.7 grams and a sample standard deviation of 11.3 grams. Determine a 98% confidence interval of the true population mean.

Solution

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

$$n = 27$$

$$\bar{x} = 69.7$$

$$s = 11.3$$

$$CL = 0.98$$

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

$$df = n - 1 = 26$$

Determine the critical t value, t^* , such that $P(|T| < t^*) = 0.98$.

$$t^* = 2.48$$

Calculate the standard error.

$$SE = \frac{s}{\sqrt{n}} = \frac{11.3}{\sqrt{27}} = 2.17$$

We want to make an inference about the population mean.

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

Determine the bounds.

$$\begin{aligned} CI &= (\bar{x} - t^* SE, \bar{x} + t^* SE) \\ &= (69.7 - 2.48 \times 2.17, 69.7 + 2.48 \times 2.17) \\ &= (64.3, 75.1) \end{aligned}$$

We are 98% confident that the population mean is between 64.3 and 75.1.

6. A treatment group of size 12 has a mean of 1090 and standard deviation of 63.8. A control group of size 11 has a mean of 1030 and standard deviation of 73. If you decided to use a significance level of 0.04, 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 19.

- (a) State the null hypothesis.
- (b) State the alternative hypothesis.
- (c) Evaluate the critical value. (The critical value is either z^* or t^* . 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.

$$n_1 = 12$$

$$\bar{x}_1 = 1090$$

$$s_1 = 63.8$$

$$n_2 = 11$$

$$\bar{x}_2 = 1030$$

$$s_2 = 73$$

$$\alpha = 0.04$$

$$df = 19$$

State the hypotheses.

$$H_0 : \mu_2 - \mu_1 = 0$$

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

We are using a two-tail test. Find t^* such that $P(|T| > t^*) = 0.04$ by using a t table.

$$t^* = 2.2$$

Calculate the standard error.

$$\begin{aligned} SE &= \sqrt{\frac{(s_1)^2}{n_1} + \frac{(s_2)^2}{n_2}} \\ &= \sqrt{\frac{(63.8)^2}{12} + \frac{(73)^2}{11}} \\ &= 28.699 \end{aligned}$$

Determine the test statistic.

$$\begin{aligned} t_{\text{obs}} &= \frac{(\bar{x}_2 - \bar{x}_1) - (\mu_2 - \mu_1)_0}{SE} \\ &= \frac{(1030 - 1090) - (0)}{28.699} \\ &= -2.09 \end{aligned}$$

Compare $|t_{\text{obs}}|$ and t^* .

$$|t_{\text{obs}}| < t^*$$

We can determine an interval for the p -value using the t table.

$$0.05 < p\text{-value} < 0.1$$

Compare p -value and α .

$$p\text{-value} > \alpha$$

We conclude that we should retain the null hypothesis.

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

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

(c) $t^* = 2.2$

(d) $SE = 28.699$

(e) $|t_{\text{obs}}| = 2.09$

(f) $0.05 < p\text{-value} < 0.1$

(g) retain the null

7. From a very large population, a random sample of 1900 individuals was taken. In that sample, 68.8% were cold. Determine a 90% confidence interval of the population proportion.
- (a) Find the lower bound of the confidence interval.
 - (b) Find the upper bound of the confidence interval.

Solution

Determine z^* such that $P(|Z| < z^*) = 0.9$.

$$z^* = 1.64$$

Calculate the standard error.

$$SE = \sqrt{\frac{\hat{p}(1 - \hat{p})}{n}} = \sqrt{\frac{(0.688)(1 - 0.688)}{1900}} = 0.0106$$

Calculate the margin of error.

$$ME = z^* SE = (1.64)(0.0106) = 0.0174$$

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.

$$(0.671, 0.705)$$

We are 90% confident that the true population proportion is between 67.1% and 70.5%.

- (a) The lower bound = 0.671, which can also be expressed as 67.1%.
- (b) The upper bound = 0.705, which can also be expressed as 70.5%.

8. An experiment is run with a treatment group of size 275 and a control group of size 226. The results are summarized in the table below.

	treatment	control
cold	237	211
not cold	38	15

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

- (a) State the null hypothesis.
- (b) State the alternative hypothesis.
- (c) Evaluate the critical value. (The critical value is either z^* or t^* . 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.

$$H_0 : p_2 - p_1 = 0$$

$$H_A : p_2 - p_1 \neq 0$$

Find z^* such that $P(|Z| > z^*) = 0.02$.

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

Determine the sample proportions.

$$\hat{p}_1 = \frac{237}{275} = 0.862$$

$$\hat{p}_2 = \frac{211}{226} = 0.934$$

Determine the difference of sample proportions.

$$\hat{p}_2 - \hat{p}_1 = 0.934 - 0.862 = 0.072$$

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

$$\hat{p} = \frac{237 + 211}{275 + 226} = 0.894$$

Determine the standard error.

$$\begin{aligned} SE &= \sqrt{\frac{\hat{p}(1 - \hat{p})}{n_1} + \frac{\hat{p}(1 - \hat{p})}{n_2}} \\ &= \sqrt{\frac{(0.894)(0.106)}{275} + \frac{(0.894)(0.106)}{226}} \\ &= 0.0276 \end{aligned}$$

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

$$H_0 : \hat{P}_2 - \hat{P}_1 \sim \mathcal{N}(0, 0.0276)$$

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.

$$\begin{aligned} z &= \frac{(\hat{p}_2 - \hat{p}_1) - (p_2 - p_1)_0}{SE} \\ &= \frac{(0.934 - 0.862) - 0}{0.0276} \\ &= 2.61 \end{aligned}$$

Determine the p -value.

$$\begin{aligned} p\text{-value} &= 2 \cdot \Phi(-|z|) \\ &= 2 \cdot \Phi(-2.61) \\ &= 0.009 \end{aligned}$$

Compare the p -value to the significance 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) $z^* = 2.33$

(d) $SE = 0.0276$

(e) $|z_{\text{obs}}| = 2.61$

(f) $p\text{-value} = 0.009$

(g) reject the null