

Bunker Hill Community College

Final Statistics Exam 2019-05-02

Exam ID 008

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) $P(\text{pig given blue}) = 0.228$
- (b) $P(\text{teal}) = 0.269$
- (c) $P(\text{pig or gray}) = 0.407$
- (d) $P(\text{horn}) = 0.185$
- (e) $P(\text{red given wheel}) = 0.212$
- (f) $P(\text{flower and red}) = 0.0663$
2. $P(\text{"not pig" given "yellow"}) = 0.316$
3. $P(69.67 < X < 70.23) = 0.785$
4. (a) $P(X = 37) = 0.0742$
- (b) $P(35 \leq X \leq 48) = 0.5848$
5. **(19, 22.1)**
6. (a) $H_0 : \mu_2 - \mu_1 = 0$
- (b) $H_0 : \mu_2 - \mu_1 \neq 0$
- (c) $t^* = 1.99$
- (d) $SE = 3.921$
- (e) $|t_{\text{obs}}| = 1.79$
- (f) $0.05 < p\text{-value} < 0.1$
- (g) **retain**
7. (a) **LB of p CI = 0.934 or 93.4%**
- (b) **UB of p CI = 0.938 or 93.8%**

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

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

(c) $z^* = 1.96$

(d) $SE = 0.024$

(e) $|z_{\text{obs}}| = 2.03$

(f) $p\text{-value} = 0.0424$

(g) **reject**

1. In a deck of strange cards, there are 966 cards. Each card has an image and a color. The amounts are shown in the table below.

	blue	gray	red	teal
flower	98	87	64	89
horn	27	48	16	88
pig	55	44	71	10
wheel	61	78	57	73

- (a) What is the probability a random card is a pig given it is blue?
- (b) What is the probability a random card is teal?
- (c) What is the probability a random card is either a pig or gray (or both)?
- (d) What is the probability a random card is a horn?
- (e) What is the probability a random card is red given it is a wheel?
- (f) What is the probability a random card is both a flower and red?

Solution

$$(a) P(\text{pig given blue}) = \frac{55}{98+27+55+61} = 0.228$$

$$(b) P(\text{teal}) = \frac{89+88+10+73}{966} = 0.269$$

$$(c) P(\text{pig or gray}) = \frac{55+44+71+10+87+48+44+78-44}{966} = 0.407$$

$$(d) P(\text{horn}) = \frac{27+48+16+88}{966} = 0.185$$

$$(e) P(\text{red given wheel}) = \frac{57}{61+78+57+73} = 0.212$$

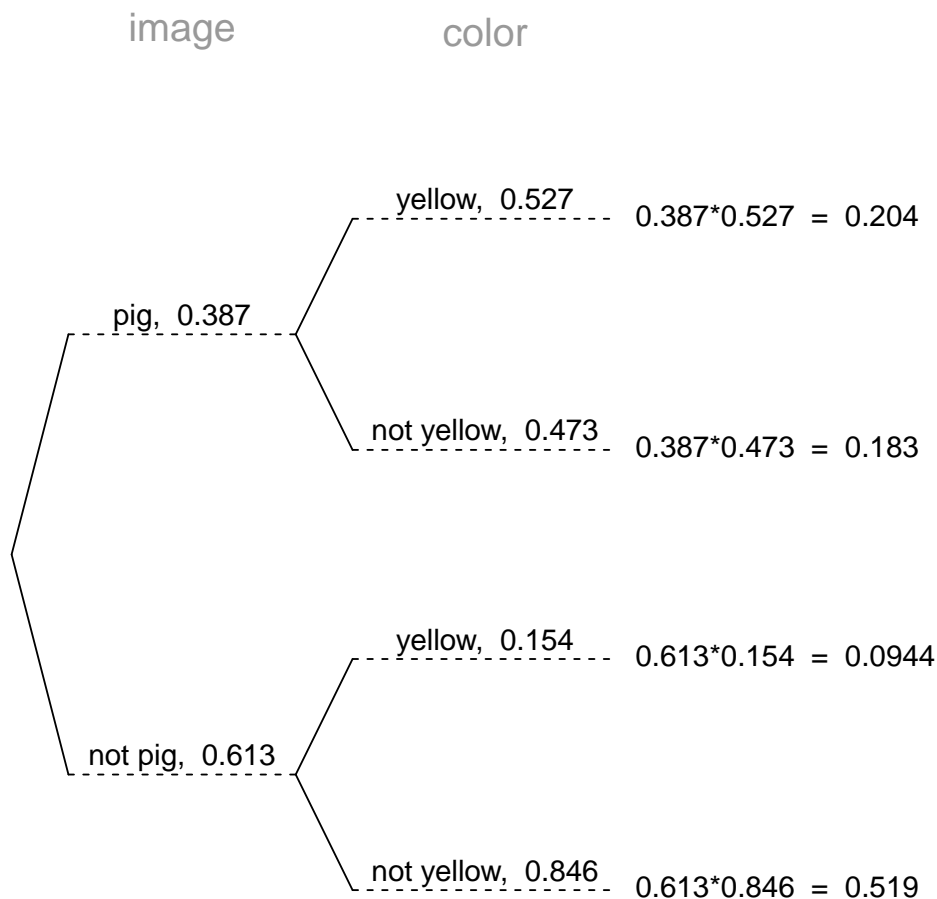
$$(f) P(\text{flower and red}) = \frac{64}{966} = 0.0663$$

2. In a deck of strange cards, each card has an image and a color. The chance of drawing a pig is 38.7%. If a pig is drawn, there is a 52.7% chance that it is yellow. If a card that is not a pig is drawn, there is a 15.4% chance that it is yellow.

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



Determine the appropriate conditional probability.

$$P(\text{"not pig" given "yellow"}) = \frac{0.0944}{0.0944 + 0.204} = 0.316$$

3. In a very large pile of toothpicks, the mean length is 69.91 millimeters and the standard deviation is 2.89 millimeters. If you randomly sample 169 toothpicks, what is the chance the sample mean is between 69.67 and 70.23 millimeters?

Solution

Label the given information.

$$\mu = 69.91$$

$$\sigma = 2.89$$

$$n = 169$$

$$\bar{x}_{\text{lower}} = 69.67$$

$$\bar{x}_{\text{upper}} = 70.23$$

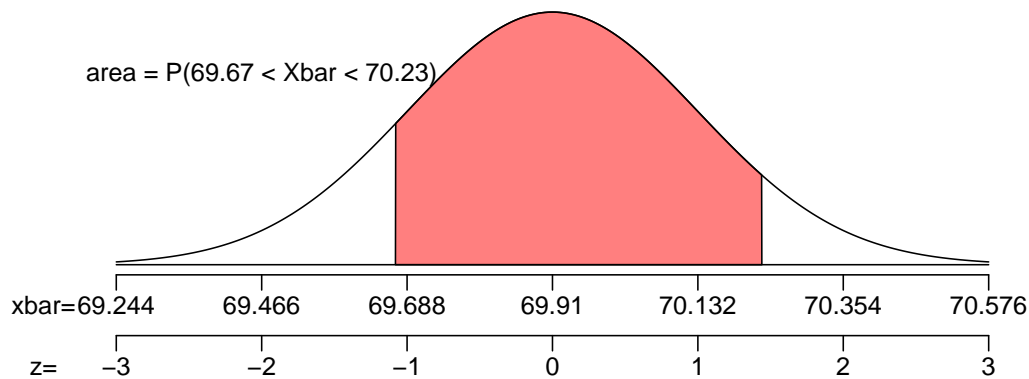
Find the standard error.

$$SE = \frac{\sigma}{\sqrt{n}} = \frac{2.89}{\sqrt{169}} = 0.222$$

Describe the sampling distribution.

$$\bar{X} \sim \mathcal{N}(69.91, 0.222)$$

Draw a sketch.



Calculate a z scores.

$$z_{\text{lower}} = \frac{\bar{x}_{\text{lower}} - \mu}{SE} = \frac{69.67 - 69.91}{0.222} = -1.08$$

$$z_{\text{upper}} = \frac{\bar{x}_{\text{upper}} - \mu}{SE} = \frac{70.23 - 69.91}{0.222} = 1.44$$

Determine the probability.

$$\begin{aligned} P(69.67 < \bar{X} < 70.23) &= \Phi(z_{\text{upper}}) - \Phi(z_{\text{lower}}) \\ &= \Phi(1.44) - \Phi(-1.08) \\ &= 0.785 \end{aligned}$$

4. In a game, there is a 23% chance to win a round. You will play 158 rounds.
- (a) What is the probability of winning exactly 37 rounds?
 - (b) What is the probability of winning at least 35 but at most 48 rounds?

Solution

We use the formula for binomial probabilities.

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

$$P(X = 37) = \binom{158}{37} (0.23)^{37} (1 - 0.23)^{158-37}$$

$$P(X = 37) = \binom{158}{37} (0.23)^{37} (0.77)^{121}$$

$$P(X = 37) = 0.0742$$

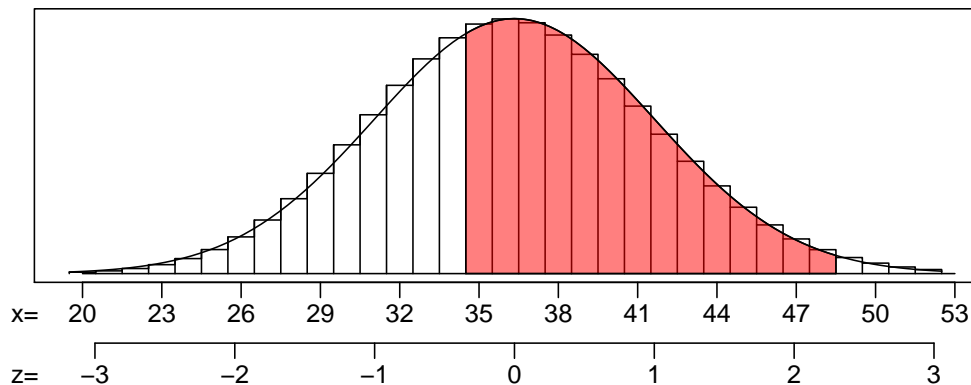
Find the mean.

$$\mu = np = (158)(0.23) = 36.34$$

Find the standard deviation.

$$\sigma = \sqrt{np(1 - p)} = \sqrt{(158)(0.23)(1 - 0.23)} = 5.2898$$

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



Find the z scores.

$$z_1 = \frac{34.5 - 36.34}{5.2898} = -0.25$$

$$z_2 = \frac{48.5 - 36.34}{5.2898} = 2.2$$

Calculate the probability.

$$P(35 \leq X \leq 48) = \Phi(2.2) - \Phi(-0.25) = 0.5848$$

(a) $P(X = 37) = 0.0742$

(b) $P(35 \leq X \leq 48) = 0.5848$

5. As an ornithologist, you wish to determine the average body mass of *Seiurus noveboracensis*. You randomly sample 33 adults of *Seiurus noveboracensis*, resulting in a sample mean of 20.51 grams and a sample standard deviation of 3.64 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.

$$n = 33$$

$$\bar{x} = 20.51$$

$$s = 3.64$$

$$CL = 0.98$$

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

$$df = n - 1 = 32$$

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

$$t^* = 2.45$$

Calculate the standard error.

$$SE = \frac{s}{\sqrt{n}} = \frac{3.64}{\sqrt{33}} = 0.634$$

We want to make an inference about the population mean.

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

Determine the bounds.

$$\begin{aligned} CI &= (\bar{x} - t^* SE, \bar{x} + t^* SE) \\ &= (20.51 - 2.45 \times 0.634, 20.51 + 2.45 \times 0.634) \\ &= (19, 22.1) \end{aligned}$$

We are 98% confident that the population mean is between 19 and 22.1.

6. A treatment group of size 33 has a mean of 102 and standard deviation of 15.4. A control group of size 40 has a mean of 109 and standard deviation of 18.1. If you decided to use a significance level of 0.05, 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 70.

- (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.

$$\begin{aligned}n_1 &= 33 \\ \bar{x}_1 &= 102 \\ s_1 &= 15.4 \\ n_2 &= 40 \\ \bar{x}_2 &= 109 \\ s_2 &= 18.1 \\ \alpha &= 0.05 \\ df &= 70\end{aligned}$$

State the hypotheses.

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

We are using a two-tail test. Find t^* such that $P(|T| > t^*) = 0.05$ by using a t table.

$$t^* = 1.99$$

Calculate the standard error.

$$\begin{aligned}SE &= \sqrt{\frac{(s_1)^2}{n_1} + \frac{(s_2)^2}{n_2}} \\ &= \sqrt{\frac{(15.4)^2}{33} + \frac{(18.1)^2}{40}} \\ &= 3.921\end{aligned}$$

Determine the test statistic.

$$\begin{aligned}t_{\text{obs}} &= \frac{(\bar{x}_2 - \bar{x}_1) - (\mu_2 - \mu_1)_0}{SE} \\ &= \frac{(109 - 102) - (0)}{3.921} \\ &= 1.79\end{aligned}$$

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) $t^* = 1.99$
- (d) $SE = 3.921$
- (e) $|t_{\text{obs}}| = 1.79$
- (f) $0.05 < p\text{-value} < 0.1$
- (g) retain the null

7. From a very large population, a random sample of 49000 individuals was taken. In that sample, 93.6% were asleep. 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 confidence interval.

Solution

Determine z^* such that $P(|Z| < z^*) = 0.96$.

$$z^* = 2.05$$

Calculate the standard error.

$$SE = \sqrt{\frac{\hat{p}(1 - \hat{p})}{n}} = \sqrt{\frac{(0.936)(1 - 0.936)}{49000}} = 0.00111$$

Calculate the margin of error.

$$ME = z^* SE = (2.05)(0.00111) = 0.00228$$

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.

$$(0.934, 0.938)$$

We are 96% confident that the true population proportion is between 93.4% and 93.8%.

- (a) The lower bound = 0.934, which can also be expressed as 93.4%.
- (b) The upper bound = 0.938, which can also be expressed as 93.8%.

8. An experiment is run with a treatment group of size 204 and a control group of size 243. The results are summarized in the table below.

	treatment	control
glossy	185	232
not glossy	19	11

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

- (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 $P(|Z| > z^*) = 0.05$.

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

Determine the sample proportions.

$$\hat{p}_1 = \frac{185}{204} = 0.907$$

$$\hat{p}_2 = \frac{232}{243} = 0.955$$

Determine the difference of sample proportions.

$$\hat{p}_2 - \hat{p}_1 = 0.955 - 0.907 = 0.048$$

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

$$\hat{p} = \frac{185 + 232}{204 + 243} = 0.933$$

Determine the standard error.

$$\begin{aligned} SE &= \sqrt{\frac{\hat{p}(1 - \hat{p})}{n_1} + \frac{\hat{p}(1 - \hat{p})}{n_2}} \\ &= \sqrt{\frac{(0.933)(0.067)}{204} + \frac{(0.933)(0.067)}{243}} \\ &= 0.0237 \end{aligned}$$

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

$$H_0 : \hat{P}_2 - \hat{P}_1 \sim \mathcal{N}(0, 0.0237)$$

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.

$$\begin{aligned} z &= \frac{(\hat{p}_2 - \hat{p}_1) - (p_2 - p_1)_0}{SE} \\ &= \frac{(0.955 - 0.907) - 0}{0.0237} \\ &= 2.03 \end{aligned}$$

Determine the p -value.

$$\begin{aligned} p\text{-value} &= 2 \cdot \Phi(-|z|) \\ &= 2 \cdot \Phi(-2.03) \\ &= 0.0424 \end{aligned}$$

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) $z^* = 1.96$

(d) $SE = 0.0237$

(e) $|z_{\text{obs}}| = 2.03$

(f) $p\text{-value} = 0.0424$

(g) reject the null