Bunker Hill Community College

Final Statistics Exam 2019-05-02 — Exam ID 014

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(**cat and indigo**) = 0.0818

   (b) P(**black**) = 0.347

   (c) P(**pig or indigo**) = 0.471

   (d) P(**pig given indigo**) = 0.373

   (e) P(**pig**) = 0.244

   (f) P(**black given flower**) = 0.65

2. P(**"not shovel" given "violet"**) = 0.247

3. $P(72.76 < X < 73.89) = 0.8933$

4. (a) $P(X = 26) = 0.1302$

   (b) $P(21 \leq X \leq 26) = 0.5714$

5. **(14.5, 16.3)**

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

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

   (c) $t^{\star} = 2.67$

   (d) $SE = 0.682$

   (e) $\mid t_{\mathbf{obs}} \mid = 2.81$

   (f) $0.005 < p$-**value** < 0.01

   (g) **reject**

7. (a) **LB of p CI = 0.931 or** 93.1%

   (b) **UB of p CI = 0.949 or** 94.9%

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

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

(c) $z^\star = 1.28$

(d) $SE = 0.044$

(e) $\mid z_{\mathbf{obs}} \mid = 1.55$

(f) p-**value** = 0.1212

(g) **reject**

1. In a deck of strange cards, there are 599 cards. Each card has an image and a color. The amounts are shown in the table below.

	black	indigo	red
cat	12	49	35
flower	89	16	32
pig	54	81	11
shovel	53	71	96

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

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

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

(d) What is the probability a random card is a pig given it is indigo?

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

(f) What is the probability a random card is black given it is a flower?

Solution

(a) $P(\text{cat and indigo}) = \frac{49}{599} = 0.0818$

(b) $P(\text{black}) = \frac{12+89+54+53}{599} = 0.347$

(c) $P(\text{pig or indigo}) = \frac{54+81+11+49+16+81+71-81}{599} = 0.471$

(d) $P(\text{pig given indigo}) = \frac{81}{49+16+81+71} = 0.373$

(e) $P(\text{pig}) = \frac{54+81+11}{599} = 0.244$

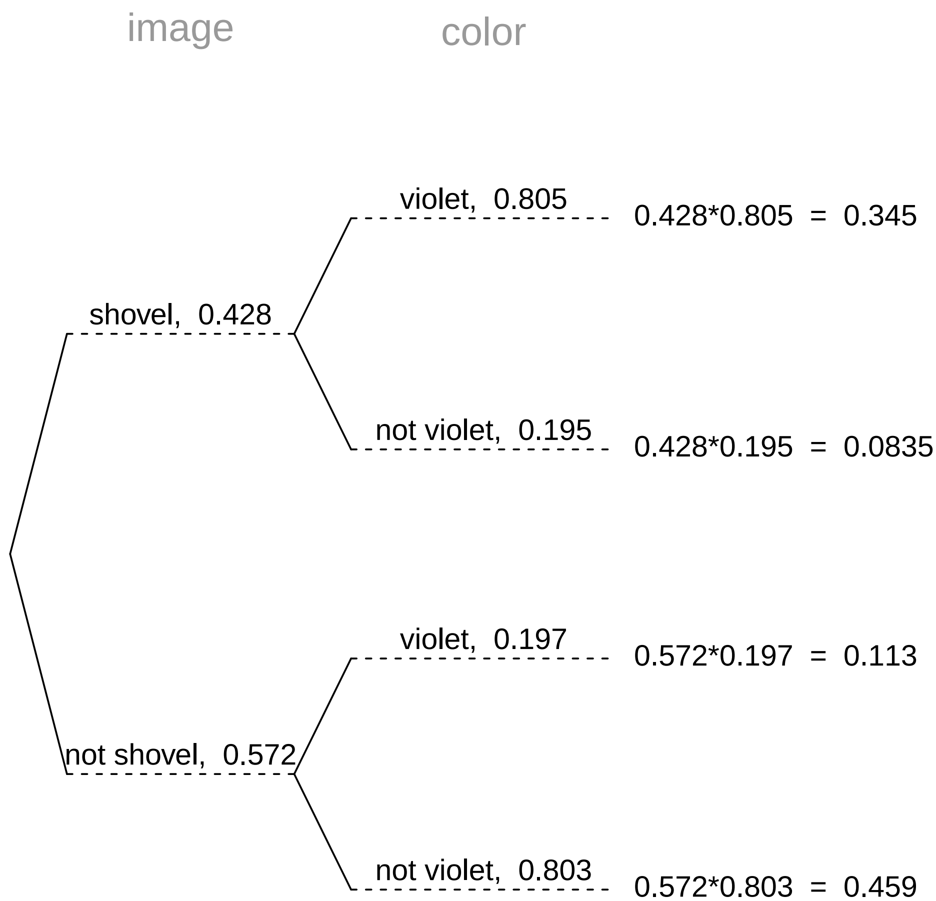
(f) $P(\text{black given flower}) = \frac{89}{89+16+32} = 0.65$

2. In a deck of strange cards, each card has an image and a color. The chance of drawing a shovel is 42.8%. If a shovel is drawn, there is a 80.5% chance that it is violet. If a card that is not a shovel is drawn, there is a 19.7% chance that it is violet.

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

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.

$$P(\text{"not shovel" given "violet"}) = \frac{0.113}{0.113 + 0.345} = 0.247$$

3. In a very large pile of toothpicks, the mean length is 73.52 millimeters and the standard deviation is 3.5 millimeters. If you randomly sample 144 toothpicks, what is the chance the sample mean is between 72.76 and 73.89 millimeters?

Solution

Label the given information.

$$\mu = 73.52$$
$$\sigma = 3.5$$
$$n = 144$$
$$\bar{x}_{\text{lower}} = 72.76$$
$$\bar{x}_{\text{upper}} = 73.89$$

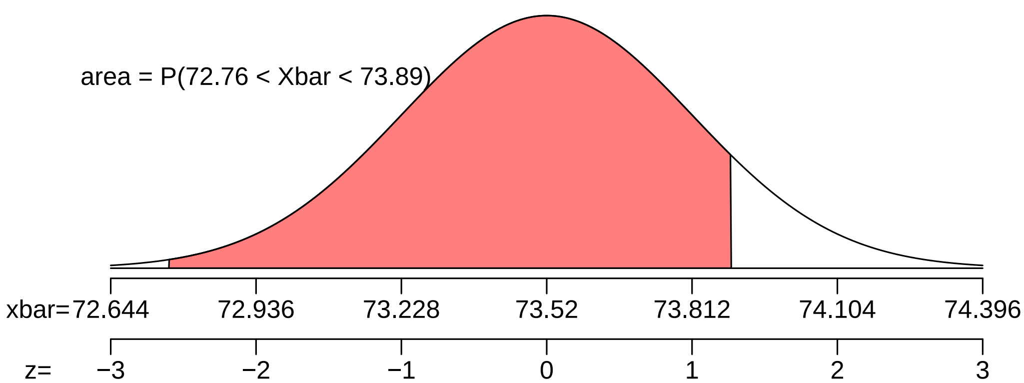
Find the standard error.

$$SE = \frac{\sigma}{\sqrt{n}} = \frac{3.5}{\sqrt{144}} = 0.292$$

Describe the sampling distribution.

$$\bar{X} \sim \mathcal{N}(73.52,\ 0.292)$$

Draw a sketch.



Calculate a z scores.

$$z_{\text{lower}} = \frac{x_{\text{lower}} - \mu}{SE} = \frac{72.76 - 73.52}{0.292} = -2.6$$

$$z_{\text{upper}} = \frac{x_{\text{upper}} - \mu}{SE} = \frac{73.89 - 73.52}{0.292} = 1.27$$

Determine the probability.

$$\begin{aligned} P(72.76 < X < 73.89) &= \Phi(z_{\text{upper}}) - \Phi(z_{\text{lower}}) \\ &= \Phi(1.27) - \Phi(-2.6) \\ &= 0.8933 \end{aligned}$$

4. In a game, there is a 64% chance to win a round. You will play 40 rounds.
   (a) What is the probability of winning exactly 26 rounds?
   (b) What is the probability of winning at least 21 but at most 26 rounds?

Solution

We use the formula for binomial probabilities.

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

$$P(X = 26) = \binom{40}{26}(0.64)^{26}(1-0.64)^{40-26}$$

$$P(X = 26) = \binom{40}{26}(0.64)^{26}(0.36)^{14}$$

$$P(X = 26) = 0.1302$$

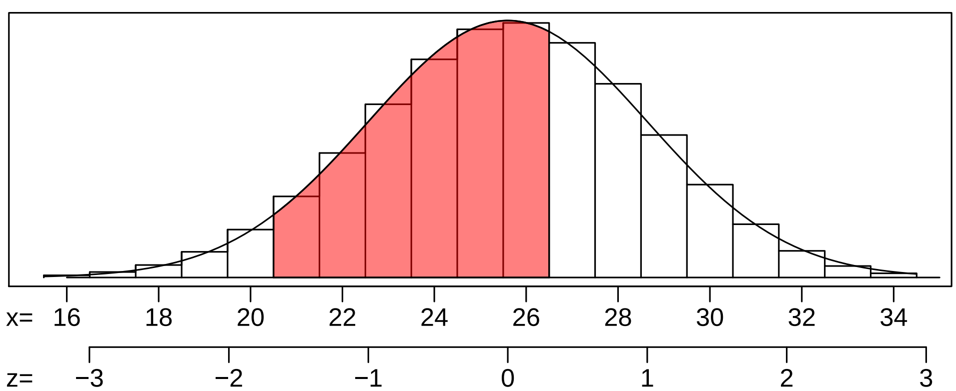
Find the mean.

$$\mu = np = (40)(0.64) = 25.6$$

Find the standard deviation.

$$\sigma = \sqrt{np(1-p)} = \sqrt{(40)(0.64)(1-0.64)} = 3.0358$$

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



Find the z scores.

$$z_1 = \frac{20.5 - 25.6}{3.0358} = -1.68$$

$$z_2 = \frac{26.5 - 25.6}{3.0358} = 0.3$$

Calculate the probability.

$$P(21 \leq X \leq 26) = \Phi(0.3) - \Phi(-1.68) = 0.5714$$

(a) $P(X = 26) = 0.1302$

(b) $P(21 \leq X \leq 26) = 0.5714$

5. As an ornithologist, you wish to determine the average body mass of *Passerina cyanea*. You randomly sample 30 adults of *Passerina cyanea*, resulting in a sample mean of 15.38 grams and a sample standard deviation of 2.38 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.

$$
\begin{aligned}
n &= 30 \\
\bar{x} &= 15.38 \\
s &= 2.38 \\
CL &= 0.95
\end{aligned}
$$

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

$$df = n - 1 = 29$$

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

$$t^\star = 2.05$$

Calculate the standard error.

$$SE = \frac{s}{\sqrt{n}} = \frac{2.38}{\sqrt{30}} = 0.435$$

We want to make an inference about the population mean.

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

Determine the bounds.

$$
\begin{aligned}
CI &= (\bar{x} - t^\star SE,\ \bar{x} + t^\star SE) \\
&= (15.38 - 2.05 \times 0.435,\ 15.38 + 2.05 \times 0.435) \\
&= (14.5,\ 16.3)
\end{aligned}
$$

We are 95% confident that the population mean is between 14.5 and 16.3.

6. A treatment group of size 38 has a mean of 11.5 and standard deviation of 2.46. A control group of size 30 has a mean of 9.58 and standard deviation of 3.03. 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 55.

   (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.

$$n_1 = 38$$
$$\bar{x}_1 = 11.5$$
$$s_1 = 2.46$$
$$n_2 = 30$$
$$\bar{x}_2 = 9.58$$
$$s_2 = 3.03$$
$$\alpha = 0.01$$
$$df = 55$$

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.

$$t^{\star} = 2.67$$

Calculate the standard error.

$$\begin{aligned} SE &= \sqrt{\frac{(s_1)^2}{n_1} + \frac{(s_2)^2}{n_2}} \\ &= \sqrt{\frac{(2.46)^2}{38} + \frac{(3.03)^2}{30}} \\ &= 0.682 \end{aligned}$$

Determine the test statistic.

$$\begin{aligned} t_{\text{obs}} &= \frac{(\bar{x}_2 - \bar{x}_1) - (\mu_2 - \mu_1)_0}{SE} \\ &= \frac{(9.58 - 11.5) - (0)}{0.682} \\ &= -2.81 \end{aligned}$$

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

(c) $t^\star = 2.67$

(d) $SE = 0.682$

(e) $|t_{\text{obs}}| = 2.81$

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

(g) reject the null

7. From a very large population, a random sample of 4800 individuals was taken. In that sample, 94% were happy. Determine a 99% 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 $P(|Z| < z^\star) = 0.99$.

$$z^\star = 2.58$$

Calculate the standard error.

$$SE = \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = \sqrt{\frac{(0.94)(1-0.94)}{4800}} = 0.00343$$

Calculate the margin of error.

$$ME = z^\star SE = (2.58)(0.00343) = 0.00885$$

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.

$$(0.931\ ,\ 0.949)$$

We are 99% confident that the true population proportion is between 93.1% and 94.9%.

(a) The lower bound = 0.931, which can also be expressed as 93.1%.

(b) The upper bound = 0.949, which can also be expressed as 94.9%.

8. An experiment is run with a treatment group of size 184 and a control group of size 137. The results are summarized in the table below.

	treatment	control
pink	29	31
not pink	155	106

Using a significance level of 0.2, determine whether the treatment causes an effect on the proportion of cases that are pink.

(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 $P(|Z| > z^\star) = 0.2$.

$$z^\star = \Phi^{-1}\left(1 - \frac{\alpha}{2}\right) = 1.28$$

Determine the sample proportions.

$$\hat{p}_1 = \frac{29}{184} = 0.158$$
$$\hat{p}_2 = \frac{31}{137} = 0.226$$

Determine the difference of sample proportions.

$$\hat{p}_2 - \hat{p}_1 = 0.226 - 0.158 = 0.068$$

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

$$\hat{p} = \frac{29 + 31}{184 + 137} = 0.187$$

Determine the standard error.

$$\begin{aligned} SE &= \sqrt{\frac{\hat{p}(1-\hat{p})}{n_1} + \frac{\hat{p}(1-\hat{p})}{n_2}} \\ &= \sqrt{\frac{(0.187)(0.813)}{184} + \frac{(0.187)(0.813)}{137}} \\ &= 0.044 \end{aligned}$$

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

$$H_0: \quad \hat{P}_2 - \hat{P}_1 \sim \mathcal{N}(0, 0.044)$$

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.

$$\begin{aligned} z &= \frac{(\hat{p}_2 - \hat{p}_1) - (p_2 - p_1)_0}{SE} \\ &= \frac{(0.226 - 0.158) - 0}{0.044} \\ &= 1.55 \end{aligned}$$

Determine the p-value.

$$\begin{aligned} p\text{-value} &= 2 \cdot \Phi(-|z|) \\ &= 2 \cdot \Phi(-1.55) \\ &= 0.1212 \end{aligned}$$

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) $z^\star = 1.28$

(d) $SE = 0.044$

(e) $|z_{obs}| = 1.55$

(f) p-value = 0.1212

(g) reject the null