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(**wheel and violet**) = 0.0324

   (b) P(**violet**) = 0.185

   (c) P(**tree or black**) = 0.314

   (d) P(**horn given pink**) = 0.249

   (e) P(**pink given tree**) = 0.137

   (f) P(**flower**) = 0.203

2. P(**"horn" given "teal"**) = 0.144

3. $P(66.88 < X < 67.91) = 0.895$

4. (a) $P(X = 103) = 0.0616$

   (b) $P(87 \leq X \leq 105) = 0.7548$

5. **(37.5, 40.9)**

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

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

   (c) $t^{\star} = 1.69$

   (d) $SE = 31.782$

   (e) $\mid t_{\mathbf{obs}} \mid = 1.89$

   (f) $0.05 <$ p**-value** < 0.1

   (g) **reject**

7. (a) **LB of p CI = 0.437 or** 43.7%

   (b) **UB of p CI = 0.455 or** 45.5%

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

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

(c) $z^{\star} = 1.64$

(d) $SE = 0.042$

(e) $|\, z_{\mathbf{obs}} \,| = 1.81$

(f) p-**value** $= 0.0702$

(g) **reject**

1. In a deck of strange cards, there are 1387 cards. Each card has an image and a color. The amounts are shown in the table below.

	black	orange	pink	violet	white
dog	41	16	32	24	68
flower	67	85	43	77	10
horn	79	97	63	96	76
tree	52	56	29	14	60
wheel	38	72	86	45	61

   (a) What is the probability a random card is both a wheel and violet?

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

   (c) What is the probability a random card is either a tree or black (or both)?

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

   (e) What is the probability a random card is pink given it is a tree?

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

Solution

(a) $P(\text{wheel and violet}) = \frac{45}{1387} = 0.0324$

(b) $P(\text{violet}) = \frac{24+77+96+14+45}{1387} = 0.185$

(c) $P(\text{tree or black}) = \frac{52+56+29+14+60+41+67+79+52+38-52}{1387} = 0.314$

(d) $P(\text{horn given pink}) = \frac{63}{32+43+63+29+86} = 0.249$

(e) $P(\text{pink given tree}) = \frac{29}{52+56+29+14+60} = 0.137$

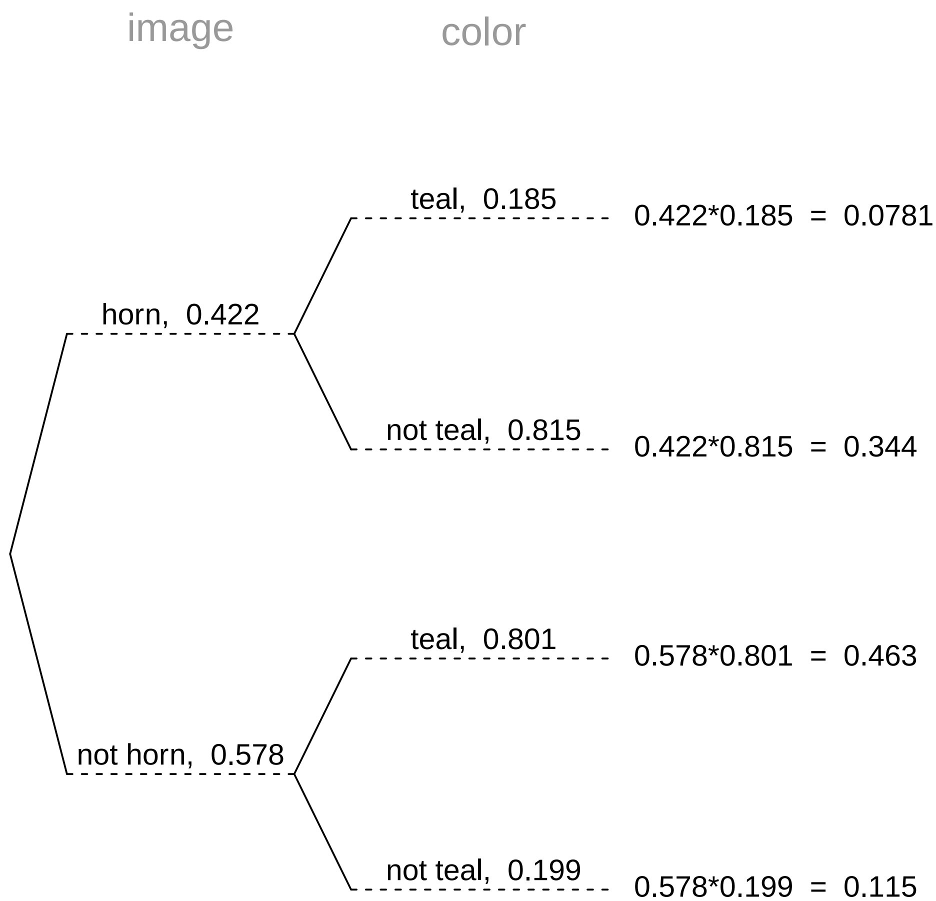
(f) $P(\text{flower}) = \frac{67+85+43+77+10}{1387} = 0.203$

2. In a deck of strange cards, each card has an image and a color. The chance of drawing a horn is 42.2%. If a horn is drawn, there is a 18.5% chance that it is teal. If a card that is not a horn is drawn, there is a 80.1% chance that it is teal.

   Now, someone draws a random card and reveals it is teal. What is the chance the card is a horn?

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.

$$P(\text{"horn" given "teal"}) = \frac{0.0781}{0.0781 + 0.463} = 0.144$$

3. In a very large pile of toothpicks, the mean length is 67.57 millimeters and the standard deviation is 2.91 millimeters. If you randomly sample 121 toothpicks, what is the chance the sample mean is between 66.88 and 67.91 millimeters?

Solution

Label the given information.

$$\mu = 67.57$$
$$\sigma = 2.91$$
$$n = 121$$
$$\bar{x}_{\text{lower}} = 66.88$$
$$\bar{x}_{\text{upper}} = 67.91$$

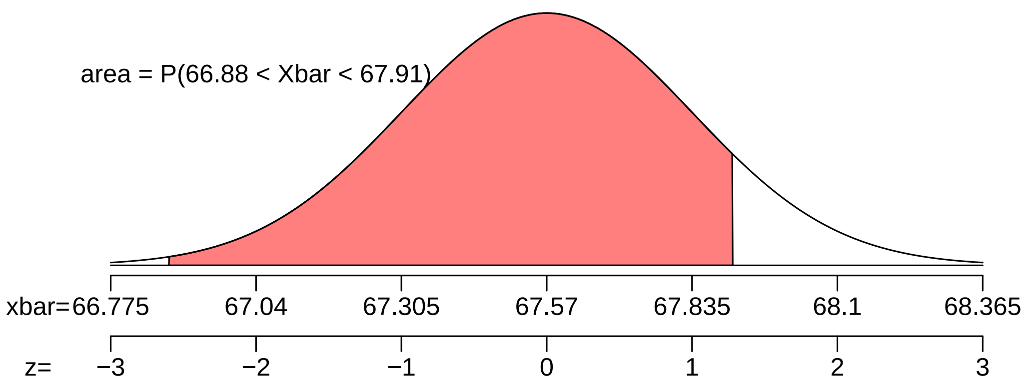
Find the standard error.

$$SE = \frac{\sigma}{\sqrt{n}} = \frac{2.91}{\sqrt{121}} = 0.265$$

Describe the sampling distribution.

$$\bar{X} \sim \mathcal{N}(67.57,\ 0.265)$$

Draw a sketch.



Calculate a z scores.

$$z_{\text{lower}} = \frac{x_{\text{lower}} - \mu}{SE} = \frac{66.88 - 67.57}{0.265} = -2.6$$

$$z_{\text{upper}} = \frac{x_{\text{upper}} - \mu}{SE} = \frac{67.91 - 67.57}{0.265} = 1.28$$

Determine the probability.

$$\begin{aligned} P(66.88 < X < 67.91) &= \Phi(z_{\text{upper}}) - \Phi(z_{\text{lower}}) \\ &= \Phi(1.28) - \Phi(-2.6) \\ &= 0.895 \end{aligned}$$

4. In a game, there is a 62% chance to win a round. You will play 163 rounds.

   (a) What is the probability of winning exactly 103 rounds?

   (b) What is the probability of winning at least 87 but at most 105 rounds?

Solution
We use the formula for binomial probabilities.

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

$$P(X = 103) = \binom{163}{103}(0.62)^{103}(1-0.62)^{163-103}$$

$$P(X = 103) = \binom{163}{103}(0.62)^{103}(0.38)^{60}$$

$$P(X = 103) = 0.0616$$

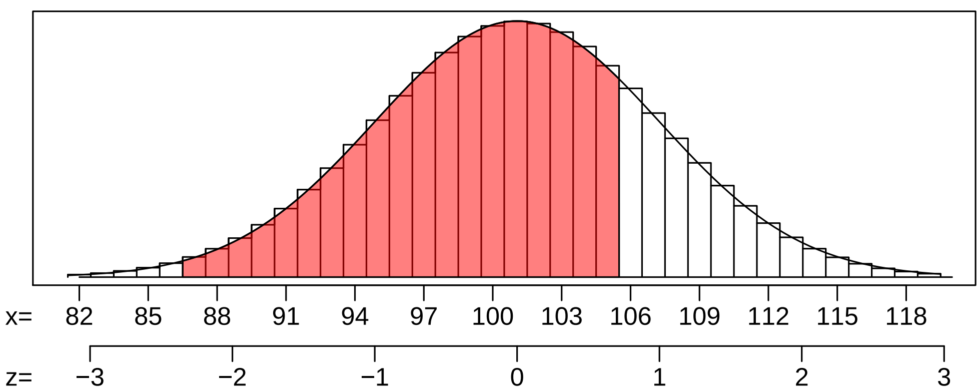
Find the mean.

$$\mu = np = (163)(0.62) = 101.06$$

Find the standard deviation.

$$\sigma = \sqrt{np(1-p)} = \sqrt{(163)(0.62)(1-0.62)} = 6.197$$

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



Find the z scores.

$$z_1 = \frac{86.5 - 101.06}{6.197} = -2.35$$

$$z_2 = \frac{105.5 - 101.06}{6.197} = 0.72$$

Calculate the probability.

$$P(87 \le X \le 105) = \Phi(0.72) - \Phi(-2.35) = 0.7548$$

(a) $P(X = 103) = 0.0616$
(b) $P(87 \le X \le 105) = 0.7548$

5. As an ornithologist, you wish to determine the average body mass of *Piranga olivacea*. You randomly sample 30 adults of *Piranga olivacea*, resulting in a sample mean of 39.2 grams and a sample standard deviation of 5.49 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.

$$n = 30$$
$$\bar{x} = 39.2$$
$$s = 5.49$$
$$CL = 0.9$$

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 $P(|T| < t^{\star}) = 0.9$.

$$t^{\star} = 1.7$$

Calculate the standard error.

$$SE = \frac{s}{\sqrt{n}} = \frac{5.49}{\sqrt{30}} = 1$$

We want to make an inference about the population mean.

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

Determine the bounds.

$$\begin{aligned} CI &= (\bar{x} - t^{\star}SE,\ \bar{x} + t^{\star}SE) \\ &= (39.2 - 1.7 \times 1,\ 39.2 + 1.7 \times 1) \\ &= (37.5,\ 40.9) \end{aligned}$$

We are 90% confident that the population mean is between 37.5 and 40.9.

6. A treatment group of size 16 has a mean of 1060 and standard deviation of 85.8. A control group of size 22 has a mean of 1000 and standard deviation of 110. If you decided to use a signficance level of 0.1, 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 35.

   (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.

$$n_1 = 16$$
$$\bar{x}_1 = 1060$$
$$s_1 = 85.8$$
$$n_2 = 22$$
$$\bar{x}_2 = 1000$$
$$s_2 = 110$$
$$\alpha = 0.1$$
$$df = 35$$

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 $P(|T| > t^\star) = 0.1$ by using a t table.

$$t^\star = 1.69$$

Calculate the standard error.

$$\begin{aligned} SE &= \sqrt{\frac{(s_1)^2}{n_1} + \frac{(s_2)^2}{n_2}} \\ &= \sqrt{\frac{(85.8)^2}{16} + \frac{(110)^2}{22}} \\ &= 31.782 \end{aligned}$$

Determine the test statistic.

$$\begin{aligned} t_{\text{obs}} &= \frac{(\bar{x}_2 - \bar{x}_1) - (\mu_2 - \mu_1)_0}{SE} \\ &= \frac{(1000 - 1060) - (0)}{31.782} \\ &= -1.89 \end{aligned}$$

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.

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

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) $t^{\star} = 1.69$

(d) $SE = 31.782$

(e) $|t_{\text{obs}}| = 1.89$

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

(g) reject the null

7. From a very large population, a random sample of 4900 individuals was taken. In that sample, 44.6% were tasty. 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.

$$SE = \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = \sqrt{\frac{(0.446)(1-0.446)}{4900}} = 0.0071$$

Calculate the margin of error.

$$ME = z^\star SE = (1.28)(0.0071) = 0.00909$$

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.

$$(0.437\,,\ 0.455)$$

We are 80% confident that the true population proportion is between 43.7% and 45.5%.

(a) The lower bound = 0.437, which can also be expressed as 43.7%.

(b) The upper bound = 0.455, which can also be expressed as 45.5%.

8. An experiment is run with a treatment group of size 258 and a control group of size 226. The results are summarized in the table below.

	treatment	control
happy	170	166
not happy	88	60

Using a significance level of 0.1, determine whether the treatment causes an effect on the proportion of cases that are happy.

(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 $P(|Z| > z^\star) = 0.1$.

$$z^\star = \Phi^{-1}\left(1 - \frac{\alpha}{2}\right) = 1.64$$

Determine the sample proportions.

$$\hat{p}_1 = \frac{170}{258} = 0.659$$

$$\hat{p}_2 = \frac{166}{226} = 0.735$$

Determine the difference of sample proportions.

$$\hat{p}_2 - \hat{p}_1 = 0.735 - 0.659 = 0.076$$

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

$$\hat{p} = \frac{170 + 166}{258 + 226} = 0.694$$

Determine the standard error.

$$\begin{aligned} SE &= \sqrt{\frac{\hat{p}(1-\hat{p})}{n_1} + \frac{\hat{p}(1-\hat{p})}{n_2}} \\ &= \sqrt{\frac{(0.694)(0.306)}{258} + \frac{(0.694)(0.306)}{226}} \\ &= 0.042 \end{aligned}$$

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

$$H_0: \quad \hat{P}_2 - \hat{P}_1 \sim \mathcal{N}(0, 0.042)$$

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.

$$\begin{aligned} z &= \frac{(\hat{p}_2 - \hat{p}_1) - (p_2 - p_1)_0}{SE} \\ &= \frac{(0.735 - 0.659) - 0}{0.042} \\ &= 1.81 \end{aligned}$$

Determine the p-value.

$$\begin{aligned} p\text{-value} &= 2 \cdot \Phi(-|z|) \\ &= 2 \cdot \Phi(-1.81) \\ &= 0.0702 \end{aligned}$$

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) $z^\star = 1.64$

(d) $SE = 0.042$

(e) $|z_{obs}| = 1.81$

(f) p-value = 0.0702

(g) reject the null