Bunker Hill Community College

Final Statistics Exam 2019-05-02 Exam ID 025

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) $P($**horn and yellow**$) = 0.0524$

   (b) $P($**dog**$) = 0.19$

   (c) $P($**flower given black**$) = 0.198$

   (d) $P($**yellow**$) = 0.331$

   (e) $P($**teal given gem**$) = 0.384$

   (f) $P($**flower or yellow**$) = 0.425$

2. $P($**"not ring" given "not teal"**$) = 0.474$

3. $P(68.3 < X < 69.42) = 0.8663$

4. (a) $P(X = 88) = 0.0627$

   (b) $P(87 \leq X \leq 103) = 0.6487$

5. **(23.7, 25.4)**

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

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

   (c) $t^{\star} = 1.71$

   (d) $SE = 17.9$

   (e) $\mid t_{\textbf{obs}} \mid = 2.01$

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

   (g) **reject**

7. (a) **LB of p CI = 0.56 or** 56%

   (b) **UB of p CI = 0.648 or** 64.8%

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) $SE = 0.049$

(e) $\mid z_{\mathbf{obs}} \mid = 1.38$

(f) p-**value** = 0.1676

(g) **reject**

1. In a deck of strange cards, there are 993 cards. Each card has an image and a color. The amounts are shown in the table below.

	black	teal	yellow
dog	78	60	51
flower	64	29	66
gem	73	96	81
horn	13	99	52
tree	95	57	79

(a) What is the probability a random card is both a horn and yellow?

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

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

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

(e) What is the probability a random card is teal given it is a gem?

(f) What is the probability a random card is either a flower or yellow (or both)?

Solution

(a) $P(\text{horn and yellow}) = \frac{52}{993} = 0.0524$

(b) $P(\text{dog}) = \frac{78+60+51}{993} = 0.19$

(c) $P(\text{flower given black}) = \frac{64}{78+64+73+13+95} = 0.198$

(d) $P(\text{yellow}) = \frac{51+66+81+52+79}{993} = 0.331$

(e) $P(\text{teal given gem}) = \frac{96}{73+96+81} = 0.384$

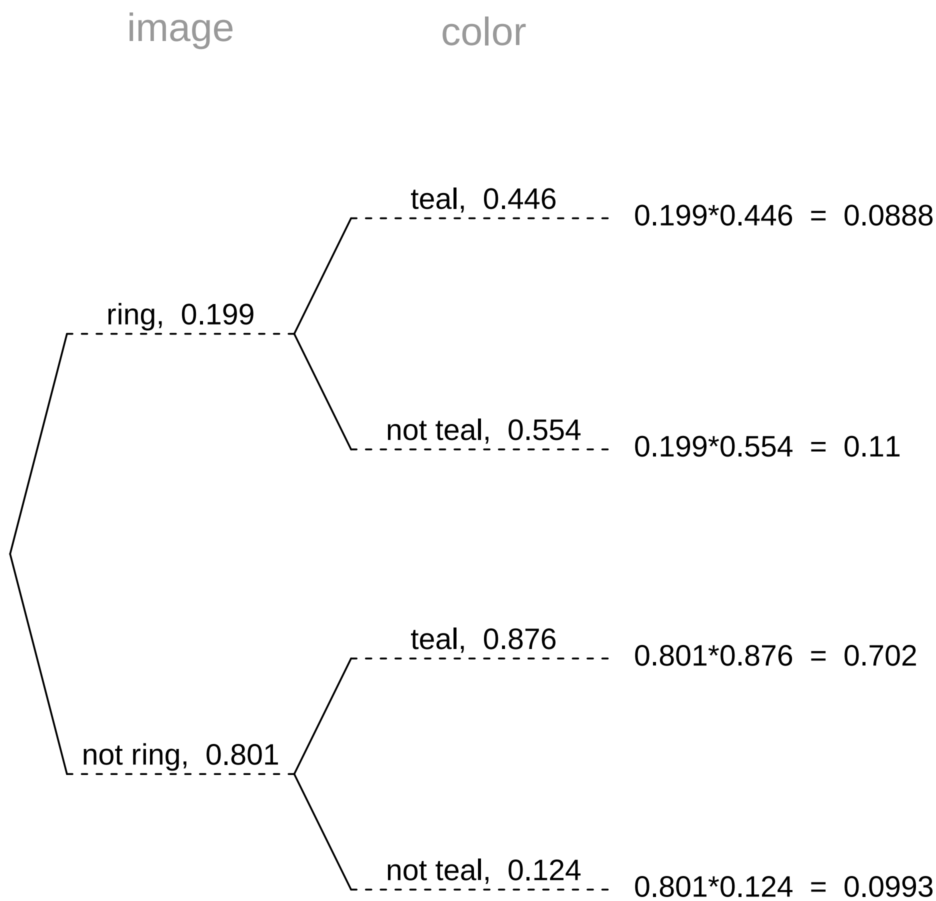
(f) $P(\text{flower or yellow}) = \frac{64+29+66+51+66+81+52+79-66}{993} = 0.425$

2. In a deck of strange cards, each card has an image and a color. The chance of drawing a ring is 19.9%. If a ring is drawn, there is a 44.6% chance that it is teal. If a card that is not a ring is drawn, there is a 87.6% chance that it is teal.

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

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.

$$P(\text{"not ring" given "not teal"}) = \frac{0.0993}{0.0993 + 0.11} = 0.474$$

3. In a very large pile of toothpicks, the mean length is 68.73 millimeters and the standard deviation is 3.51 millimeters. If you randomly sample 100 toothpicks, what is the chance the sample mean is between 68.3 and 69.42 millimeters?

Solution

Label the given information.

$$\mu = 68.73$$
$$\sigma = 3.51$$
$$n = 100$$
$$\bar{x}_{\text{lower}} = 68.3$$
$$\bar{x}_{\text{upper}} = 69.42$$

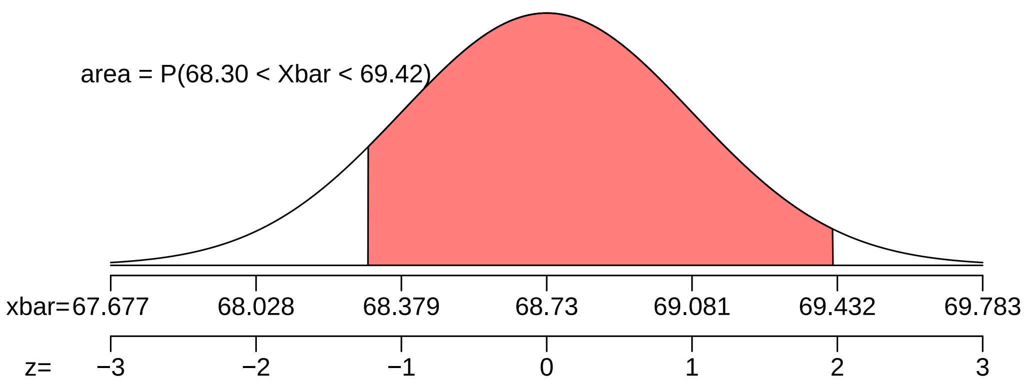
Find the standard error.

$$SE = \frac{\sigma}{\sqrt{n}} = \frac{3.51}{\sqrt{100}} = 0.351$$

Describe the sampling distribution.

$$\bar{X} \sim \mathcal{N}(68.73,\ 0.351)$$

Draw a sketch.



Calculate a z scores.

$$z_{\text{lower}} = \frac{x_{\text{lower}} - \mu}{SE} = \frac{68.3 - 68.73}{0.351} = -1.23$$

$$z_{\text{upper}} = \frac{x_{\text{upper}} - \mu}{SE} = \frac{69.42 - 68.73}{0.351} = 1.97$$

Determine the probability.

$$\begin{aligned} P(68.3 < X < 69.42) &= \Phi(z_{\text{upper}}) - \Phi(z_{\text{lower}}) \\ &= \Phi(1.97) - \Phi(-1.23) \\ &= 0.8663 \end{aligned}$$

4. In a game, there is a 56% chance to win a round. You will play 159 rounds.
   (a) What is the probability of winning exactly 88 rounds?
   (b) What is the probability of winning at least 87 but at most 103 rounds?

Solution
We use the formula for binomial probabilities.

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

$$P(X = 88) = \binom{159}{88}(0.56)^{88}(1-0.56)^{159-88}$$

$$P(X = 88) = \binom{159}{88}(0.56)^{88}(0.44)^{71}$$

$$P(X = 88) = 0.0627$$

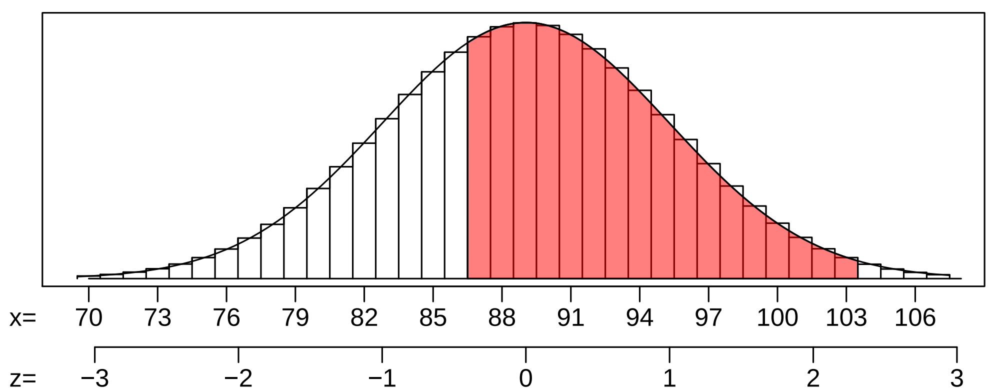
Find the mean.

$$\mu = np = (159)(0.56) = 89.04$$

Find the standard deviation.

$$\sigma = \sqrt{np(1-p)} = \sqrt{(159)(0.56)(1-0.56)} = 6.2592$$

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



Find the z scores.

$$z_1 = \frac{86.5 - 89.04}{6.2592} = -0.41$$

$$z_2 = \frac{103.5 - 89.04}{6.2592} = 2.31$$

Calculate the probability.

$$P(87 \le X \le 103) = \Phi(2.31) - \Phi(-0.41) = 0.6487$$

(a) $P(X = 88) = 0.0627$

(b) $P(87 \le X \le 103) = 0.6487$

5. As an ornithologist, you wish to determine the average body mass of *Zonotrichia albicollis*. You randomly sample 24 adults of *Zonotrichia albicollis*, resulting in a sample mean of 24.51 grams and a sample standard deviation of 1.93 grams. Determine a 96% confidence interval of the true population mean.

Solution

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

$$\begin{aligned} n &= 24 \\ \bar{x} &= 24.51 \\ s &= 1.93 \\ CL &= 0.96 \end{aligned}$$

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

$$df = n - 1 = 23$$

Determine the critical t value, $t^\star$, such that $P(|T| < t^\star) = 0.96$.

$$t^\star = 2.18$$

Calculate the standard error.

$$SE = \frac{s}{\sqrt{n}} = \frac{1.93}{\sqrt{24}} = 0.394$$

We want to make an inference about the population mean.

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

Determine the bounds.

$$\begin{aligned} CI &= (\bar{x} - t^\star SE,\ \bar{x} + t^\star SE) \\ &= (24.51 - 2.18 \times 0.394,\ 24.51 + 2.18 \times 0.394) \\ &= (23.7,\ 25.4) \end{aligned}$$

We are 96% confident that the population mean is between 23.7 and 25.4.

6. A treatment group of size 27 has a mean of 994 and standard deviation of 46.3. A control group of size 16 has a mean of 1030 and standard deviation of 62.1. 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 24.

   (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.

$$
\begin{aligned}
n_1 &= 27 \\
\bar{x}_1 &= 994 \\
s_1 &= 46.3 \\
n_2 &= 16 \\
\bar{x}_2 &= 1030 \\
s_2 &= 62.1 \\
\alpha &= 0.1 \\
df &= 24
\end{aligned}
$$

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

$$t^\star = 1.71$$

Calculate the standard error.

$$
\begin{aligned}
SE &= \sqrt{\frac{(s_1)^2}{n_1} + \frac{(s_2)^2}{n_2}} \\
&= \sqrt{\frac{(46.3)^2}{27} + \frac{(62.1)^2}{16}} \\
&= 17.9
\end{aligned}
$$

Determine the test statistic.

$$
\begin{aligned}
t_{\text{obs}} &= \frac{(\bar{x}_2 - \bar{x}_1) - (\mu_2 - \mu_1)_0}{SE} \\
&= \frac{(1030 - 994) - (0)}{17.9} \\
&= 2.01
\end{aligned}
$$

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

(c) $t^\star = 1.71$

(d) $SE = 17.9$

(e) $|t_{obs}| = 2.01$

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

(g) reject the null

7. From a very large population, a random sample of 970 individuals was taken. In that sample, 60.4% were super. Determine a 99.5% 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 $P(|Z| < z^\star) = 0.995$.

$$z^\star = 2.81$$

Calculate the standard error.

$$SE = \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = \sqrt{\frac{(0.604)(1-0.604)}{970}} = 0.0157$$

Calculate the margin of error.

$$ME = z^\star SE = (2.81)(0.0157) = 0.0441$$

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.

$$(0.56\,,\ 0.648)$$

We are 99.5% confident that the true population proportion is between 56% and 64.8%.

(a) The lower bound = 0.56, which can also be expressed as 56%.

(b) The upper bound = 0.648, which can also be expressed as 64.8%.

8. An experiment is run with a treatment group of size 189 and a control group of size 217. The results are summarized in the table below.

	treatment	control
organic	103	133
not organic	86	84

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

(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.

$$\hat{p}_1 = \frac{103}{189} = 0.545$$
$$\hat{p}_2 = \frac{133}{217} = 0.613$$

Determine the difference of sample proportions.

$$\hat{p}_2 - \hat{p}_1 = 0.613 - 0.545 = 0.068$$

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

$$\hat{p} = \frac{103 + 133}{189 + 217} = 0.581$$

Determine the standard error.

$$\begin{aligned} SE &= \sqrt{\frac{\hat{p}(1-\hat{p})}{n_1} + \frac{\hat{p}(1-\hat{p})}{n_2}} \\ &= \sqrt{\frac{(0.581)(0.419)}{189} + \frac{(0.581)(0.419)}{217}} \\ &= 0.0491 \end{aligned}$$

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

$$H_0: \quad \hat{P}_2 - \hat{P}_1 \sim \mathcal{N}(0,\ 0.0491)$$

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.

$$\begin{aligned} z &= \frac{(\hat{p}_2 - \hat{p}_1) - (p_2 - p_1)_0}{SE} \\ &= \frac{(0.613 - 0.545) - 0}{0.0491} \\ &= 1.38 \end{aligned}$$

Determine the p-value.

$$\begin{aligned} p\text{-value} &= 2 \cdot \Phi(-|z|) \\ &= 2 \cdot \Phi(-1.38) \\ &= 0.1676 \end{aligned}$$

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) $SE = 0.0491$

(e) $|z_{\text{obs}}| = 1.38$

(f) p-value = 0.1676

(g) reject the null