Bunker Hill Community College

Final Statistics Exam 2019-05-02 Exam ID 004

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(**flower**) = 0.382

   (b) P(**gem and red**) = 0.108

   (c) P(**flower or red**) = 0.557

   (d) P(**violet given bike**) = 0.12

   (e) P(**bike given violet**) = 0.167

   (f) P(**indigo**) = 0.251

2. P(**"gem" given "not green"**) = 0.41

3. $P(70.3 < X < 70.48) = 0.5709$

4. (a) $P(X = 21) = 0.0893$

   (b) $P(12 \leq X \leq 29) = 0.9456$

5. $\mathbf{(22, 22.9)}$

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

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

   (c) $t^{\star} = 2.1$

   (d) $SE = 6.868$

   (e) $\mid t_{\mathbf{obs}} \mid = 2.18$

   (f) $0.02 < p$**-value** < 0.04

   (g) **reject**

7. (a) **LB of p CI = 0.849 or** 84.9%

   (b) **UB of p CI = 0.885 or** 88.5%

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

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

(c) $z^{\star} = 2.05$

(d) $SE = 0.072$

(e) $\mid z_{\mathbf{obs}} \mid = 2.22$

(f) p-**value** $= 0.0264$

(g) **reject**

1. In a deck of strange cards, there are 752 cards. Each card has an image and a color. The amounts are shown in the table below.

	blue	indigo	red	violet	white
bike	13	36	51	24	76
flower	43	65	33	91	55
gem	19	88	81	29	48

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

(b) What is the probability a random card is both a gem and red?

(c) What is the probability a random card is either a flower or red (or both)?

(d) What is the probability a random card is violet given it is a bike?

(e) What is the probability a random card is a bike given it is violet?

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

Solution

(a) $P(\text{flower}) = \frac{43+65+33+91+55}{752} = 0.382$

(b) $P(\text{gem and red}) = \frac{81}{752} = 0.108$

(c) $P(\text{flower or red}) = \frac{43+65+33+91+55+51+33+81-33}{752} = 0.557$

(d) $P(\text{violet given bike}) = \frac{24}{13+36+51+24+76} = 0.12$

(e) $P(\text{bike given violet}) = \frac{24}{24+91+29} = 0.167$

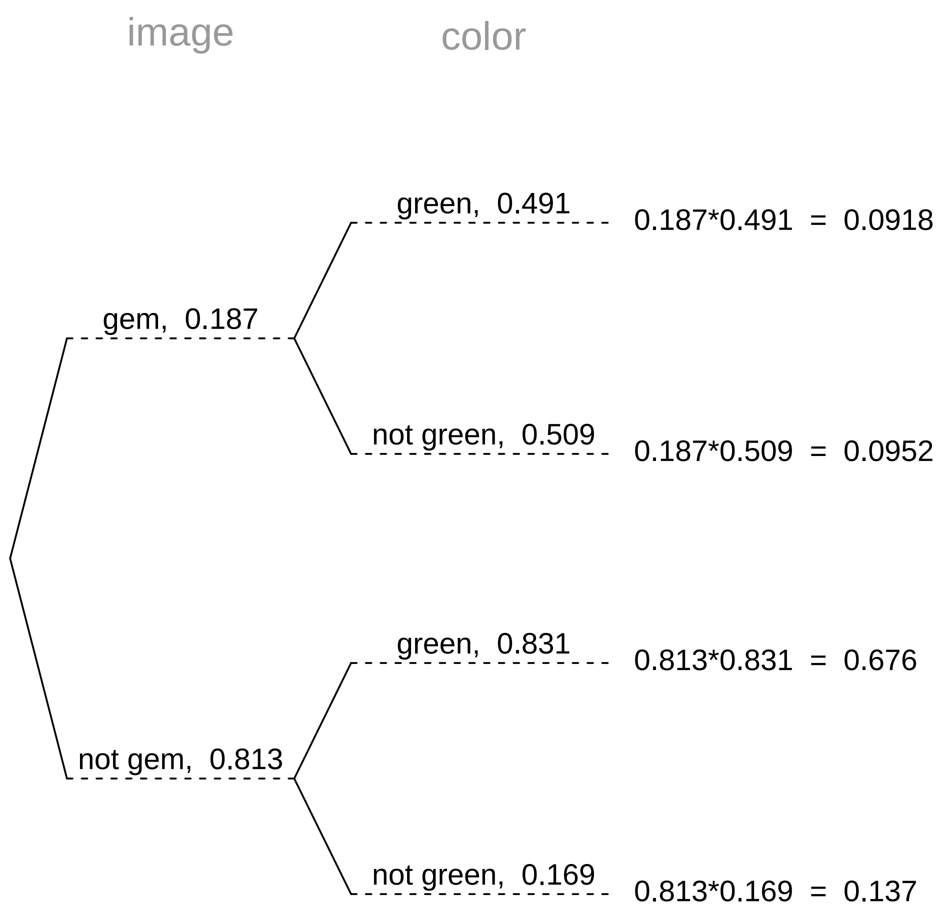
(f) $P(\text{indigo}) = \frac{36+65+88}{752} = 0.251$

2. In a deck of strange cards, each card has an image and a color. The chance of drawing a gem is 18.7%. If a gem is drawn, there is a 49.1% chance that it is green. If a card that is not a gem is drawn, there is a 83.1% chance that it is green.

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

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.

$$P(\text{"gem" given "not green"}) = \frac{0.0952}{0.0952 + 0.137} = 0.41$$

3. In a very large pile of toothpicks, the mean length is 70.42 millimeters and the standard deviation is 1.21 millimeters. If you randomly sample 120 toothpicks, what is the chance the sample mean is between 70.3 and 70.48 millimeters?

Solution

Label the given information.

$$\mu = 70.42$$
$$\sigma = 1.21$$
$$n = 120$$
$$\bar{x}_{\text{lower}} = 70.3$$
$$\bar{x}_{\text{upper}} = 70.48$$

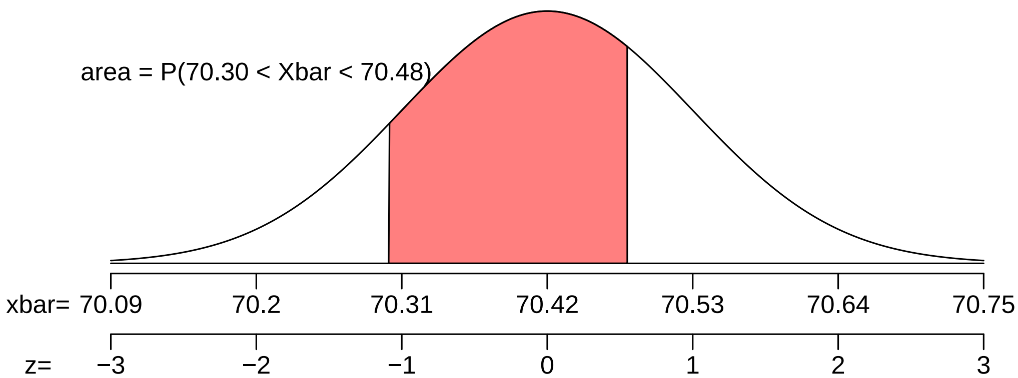
Find the standard error.

$$SE = \frac{\sigma}{\sqrt{n}} = \frac{1.21}{\sqrt{120}} = 0.11$$

Describe the sampling distribution.

$$\bar{X} \sim \mathcal{N}(70.42,\ 0.11)$$

Draw a sketch.



Calculate a z scores.

$$z_{\text{lower}} = \frac{x_{\text{lower}} - \mu}{SE} = \frac{70.3 - 70.42}{0.11} = -1.09$$

$$z_{\text{upper}} = \frac{x_{\text{upper}} - \mu}{SE} = \frac{70.48 - 70.42}{0.11} = 0.55$$

Determine the probability.

$$\begin{aligned} P(70.3 < X < 70.48) &= \Phi(z_{\text{upper}}) - \Phi(z_{\text{lower}}) \\ &= \Phi(0.55) - \Phi(-1.09) \\ &= 0.5709 \end{aligned}$$

4. In a game, there is a 11% chance to win a round. You will play 200 rounds.
   (a) What is the probability of winning exactly 21 rounds?
   (b) What is the probability of winning at least 12 but at most 29 rounds?

Solution
We use the formula for binomial probabilities.

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

$$P(X = 21) = \binom{200}{21}(0.11)^{21}(1-0.11)^{200-21}$$

$$P(X = 21) = \binom{200}{21}(0.11)^{21}(0.89)^{179}$$

$$P(X = 21) = 0.0893$$

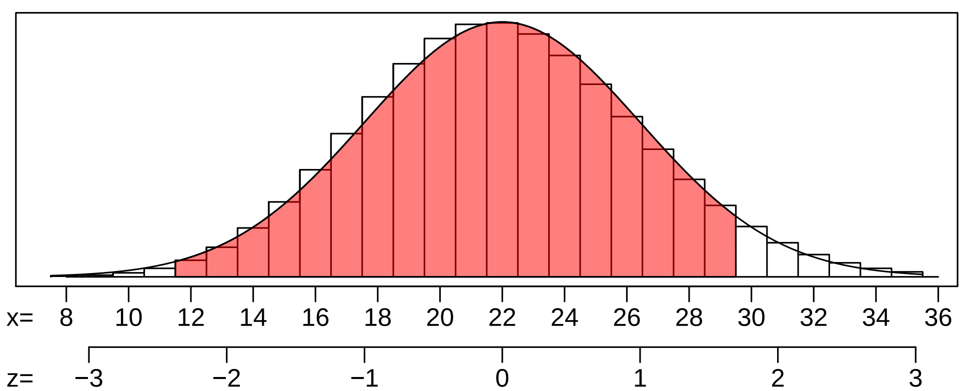
Find the mean.

$$\mu = np = (200)(0.11) = 22$$

Find the standard deviation.

$$\sigma = \sqrt{np(1-p)} = \sqrt{(200)(0.11)(1-0.11)} = 4.4249$$

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



Find the z scores.

$$z_1 = \frac{11.5 - 22}{4.4249} = -2.37$$

$$z_2 = \frac{29.5 - 22}{4.4249} = 1.69$$

Calculate the probability.

$$P(12 \leq X \leq 29) = \Phi(1.69) - \Phi(-2.37) = 0.9456$$

(a) $P(X = 21) = 0.0893$

(b) $P(12 \leq X \leq 29) = 0.9456$

5. As an ornithologist, you wish to determine the average body mass of *Ammodramus maritimus*. You randomly sample 29 adults of *Ammodramus maritimus*, resulting in a sample mean of 22.42 grams and a sample standard deviation of 1.15 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.

$$n = 29$$
$$\bar{x} = 22.42$$
$$s = 1.15$$
$$CL = 0.95$$

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

$$df = n - 1 = 28$$

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

$$t^\star = 2.05$$

Calculate the standard error.

$$SE = \frac{s}{\sqrt{n}} = \frac{1.15}{\sqrt{29}} = 0.214$$

We want to make an inference about the population mean.

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

Determine the bounds.

$$\begin{aligned} CI &= (\bar{x} - t^\star SE,\ \bar{x} + t^\star SE) \\ &= (22.42 - 2.05 \times 0.214,\ 22.42 + 2.05 \times 0.214) \\ &= (22,\ 22.9) \end{aligned}$$

We are 95% confident that the population mean is between 22 and 22.9.

6. A treatment group of size 37 has a mean of 120 and standard deviation of 33.7. A control group of size 31 has a mean of 105 and standard deviation of 22.6. If you decided to use a signficance 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 63.

   (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.

$$n_1 = 37$$
$$\bar{x}_1 = 120$$
$$s_1 = 33.7$$
$$n_2 = 31$$
$$\bar{x}_2 = 105$$
$$s_2 = 22.6$$
$$\alpha = 0.04$$
$$df = 63$$

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 $P(|T| > t^\star) = 0.04$ by using a t table.

$$t^\star = 2.1$$

Calculate the standard error.

$$\begin{aligned} SE &= \sqrt{\frac{(s_1)^2}{n_1} + \frac{(s_2)^2}{n_2}} \\ &= \sqrt{\frac{(33.7)^2}{37} + \frac{(22.6)^2}{31}} \\ &= 6.868 \end{aligned}$$

Determine the test statistic.

$$\begin{aligned} t_{\text{obs}} &= \frac{(\bar{x}_2 - \bar{x}_1) - (\mu_2 - \mu_1)_0}{SE} \\ &= \frac{(105 - 120) - (0)}{6.868} \\ &= -2.18 \end{aligned}$$

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.

$$0.02 < p\text{-value} < 0.04$$

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) $t^\star = 2.1$

(d) $SE = 6.868$

(e) $|t_{obs}| = 2.18$

(f) $0.02 < p\text{-value} < 0.04$

(g) reject the null

7. From a very large population, a random sample of 610 individuals was taken. In that sample, 86.7% were broken. 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.

$$SE = \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = \sqrt{\frac{(0.867)(1-0.867)}{610}} = 0.0137$$

Calculate the margin of error.

$$ME = z^\star SE = (1.28)(0.0137) = 0.0175$$

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.

$$(0.849\,,\ 0.885)$$

We are 80% confident that the true population proportion is between 84.9% and 88.5%.

(a) The lower bound = 0.849, which can also be expressed as 84.9%.

(b) The upper bound = 0.885, which can also be expressed as 88.5%.

8. An experiment is run with a treatment group of size 32 and a control group of size 78. The results are summarized in the table below.

	treatment	control
happy	24	71
not happy	8	7

Using a significance level of 0.04, 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 $P(|Z| > z^\star) = 0.04$.

$$z^\star = \Phi^{-1}\left(1 - \frac{\alpha}{2}\right) = 2.05$$

Determine the sample proportions.

$$\hat{p}_1 = \frac{24}{32} = 0.75$$
$$\hat{p}_2 = \frac{71}{78} = 0.91$$

Determine the difference of sample proportions.

$$\hat{p}_2 - \hat{p}_1 = 0.91 - 0.75 = 0.16$$

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

$$\hat{p} = \frac{24 + 71}{32 + 78} = 0.864$$

Determine the standard error.

$$\begin{aligned} SE &= \sqrt{\frac{\hat{p}(1-\hat{p})}{n_1} + \frac{\hat{p}(1-\hat{p})}{n_2}} \\ &= \sqrt{\frac{(0.864)(0.136)}{32} + \frac{(0.864)(0.136)}{78}} \\ &= 0.072 \end{aligned}$$

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

$$H_0: \quad \hat{P}_2 - \hat{P}_1 \sim \mathcal{N}(0, 0.072)$$

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.

$$\begin{aligned} z &= \frac{(\hat{p}_2 - \hat{p}_1) - (p_2 - p_1)_0}{SE} \\ &= \frac{(0.91 - 0.75) - 0}{0.072} \\ &= 2.22 \end{aligned}$$

Determine the p-value.

$$\begin{aligned} p\text{-value} &= 2 \cdot \Phi(-|z|) \\ &= 2 \cdot \Phi(-2.22) \\ &= 0.0264 \end{aligned}$$

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) $z^\star = 2.05$

(d) $SE = 0.072$

(e) $|z_{\text{obs}}| = 2.22$

(f) p-value = 0.0264

(g) reject the null