1. A factory manager selected a random sample of parts produced on an old assembly line and a random sample of parts produced on a new assembly line. The difference between the sample proportion of defective parts made on the old assembly line and the sample proportion of defective parts made on the new assembly line (old minus new) was 0.006. Under the assumption that all conditions for inference were met, a hypothesis test was conducted with the alternative hypothesis being the proportion of defective parts made on the old assembly line is greater than that of the new assembly line. The p-value of the test was 0.018.

Which of the following is the correct interpretation of the p-value?

(A) If there is a difference of 0.018 in the proportions of all defective parts made on the two assembly lines, the probability of observing that difference is 0.006.

(B) If there is a difference of 0.006 in the proportions of all defective parts made on the two assembly lines, the probability of observing that difference is 0.018.

(C) If there is no difference in the proportions of all defective parts made on the two assembly lines, the probability of observing a difference equal to 0.006 is 0.018.

(D) If there is no difference in the proportions of all defective parts made on the two assembly lines, the probability of observing a difference of at least 0.006 is 0.018. ✓

(E) If there is no difference in the proportions of all defective parts made on the two assembly lines, the probability of observing a difference of at most 0.006 is 0.018.

Answer D

Correct. The direction of the test is to the right of 0.006, because the alternative hypothesis is that the proportion of defective parts on the old line is greater than the proportion of defective parts on the new line. The test is in the direction of the right tail. The p-value is the probability of observing a sample difference of at least 0.006 if the null hypothesis of no difference is true.

2. Market researchers interviewed a random sample of 60 men and a random sample of 55 women about their preferences for different color designs for the packaging of a certain product. Of those interviewed, 23 men and 28 women preferred color design X.

Which of the following is the correct test statistic for a two-sample z-test for a difference in population proportions for men and women (men minus women) in their preference for color design X ?

(A) $z = \frac{0.51-0.38}{\sqrt{(0.44)(0.56)\left(\frac{1}{55}+\frac{1}{60}\right)}}$

(B) $z = \frac{23-28}{\sqrt{(0.38)(0.51)\left(\frac{1}{60+55}\right)}}$

(C) $z = \frac{23-28}{\sqrt{(0.44)(0.56)\left(\frac{1}{60}+\frac{1}{55}\right)}}$

(D) $z = \frac{0.38-0.51}{\sqrt{(0.44)(0.56)\left(\frac{1}{60+55}\right)}}$

(E) $z = \frac{0.38-0.51}{\sqrt{(0.44)(0.56)\left(\frac{1}{60}+\frac{1}{55}\right)}}$ ✓

Answer E

Correct. The sample proportion for men is $\frac{23}{60} \approx 0.38$, the sample proportion for women is $\frac{28}{55} \approx 0.51$, and the combined (or pooled) proportion is $\frac{23+28}{60+55} \approx 0.44$. The test statistic for the difference in population proportions is $\frac{0.38-0.51}{\sqrt{(0.44)(0.56)\left(\frac{1}{60}+\frac{1}{55}\right)}}$.

3. To investigate whether there is a difference in opinion on a certain proposal between two voting districts, A and B, two independent random samples were taken. From district A, 35 of the 50 voters selected were in favor of the proposal, and from district B, 36 of the 60 voters selected were in favor of the proposal.

Which of the following is the test statistic for the appropriate test to investigate whether there is a difference in the proportion of voters who are in favor of the proposal between the two districts (district A minus district B)?

(A) $\frac{35-36}{\sqrt{\frac{35}{50}+\frac{36}{60}}}$

(B) $\frac{35-36}{\sqrt{\frac{0.7}{50}+\frac{0.6}{60}}}$

(C) $\frac{0.7-0.6}{\sqrt{(0.65)(0.35)\left(\frac{1}{50}+\frac{1}{60}\right)}}$ ✓

(D) $\frac{0.7-0.6}{\sqrt{(0.7)(0.6)\left(\frac{1}{50+60}\right)}}$

(E) $\frac{0.7-0.6}{\left(0.7\right)\left(0.6\right)\sqrt{\frac{1}{50}+\frac{1}{60}}}$

Answer C

Correct. The sample proportion from district A is $\frac{35}{50} = 0.7$, the sample proportion from district B is $\frac{36}{60} = 0.6$, and the combined (or pooled) proportion is $\frac{35+36}{50+60} \approx 0.65$. The test statistic for the

difference in population proportions is $\frac{0.7-0.6}{\sqrt{(0.65)(0.35)\left(\frac{1}{50}+\frac{1}{60}\right)}}$.

4. A sociologist studying teen behavior took independent random samples of students from two high schools, F and G. Of the 80 students selected from High School F, 36 indicated they had seen a certain movie. Of the 72 students selected from High School G, 18 indicated they had seen the movie.

Which of the following is the test statistic for the appropriate test to investigate whether there is a difference in population proportions (High School F minus High School G) ?

(A) $\frac{0.45-0.25}{\sqrt{(0.36)(0.64)\left(\frac{1}{80}+\frac{1}{72}\right)}}$ ✓

(B) $\frac{0.45-0.25}{\sqrt{(0.36)(0.64)\left(\frac{1}{80+72}\right)}}$

(C) $\frac{0.45-0.25}{\left(0.45\right)\left(0.25\right)\sqrt{\frac{1}{80}+\frac{1}{72}}}$

(D) $\frac{36-18}{\sqrt{\frac{0.45}{80}+\frac{0.25}{72}}}$

(E) $\frac{36-18}{\sqrt{\frac{0.45+0.25}{80+72}}}$

Answer A

Correct. The sample proportion from High School F is $\frac{36}{80}=0.45$, the sample proportion from High School G is $\frac{18}{72}=0.25$, and the combined (or pooled) proportion is $\frac{36+18}{80+72}\approx 0.36$. The test statistic for the difference in population proportions is $\frac{0.45-0.25}{\sqrt{(0.36)(0.64)\left(\frac{1}{80}+\frac{1}{72}\right)}}$.

5. Because library books are read many times, glue is often applied to the spine of a book to keep the pages tight. A glue is considered successful if a book lasts at least 6 months before needing to be reglued. Two brands of glue, G and K, were tested to determine whether there was a difference in the proportion of books lasting at least 6 months. Let p_{G} represent the proportion of books lasting at least 6 months when glued with G, and let p_{K} represent the proportion of books lasting at least 6 months when glued with K. The following hypothesis test was conducted at the significance level of $\alpha = 0.01$.

$\text{H}_0 : p_{\text{G}} = p_{\text{K}}$
$\text{H}_\text{a} : p_{\text{G}} \neq p_{\text{K}}$

All conditions for inference were met, and the resulting p-value was 0.006. Which of the following is the correct decision for the test?

(A) The p-value is less than α, and the null hypothesis is rejected. There is not convincing evidence to support the claim that the proportion of books lasting at least 6 months when glued with G is different from the proportion of books lasting at least 6 months when glued with K.

(B) The p-value is less than α, and the null hypothesis is rejected. There is convincing evidence to support the claim that the proportion of books lasting at least 6 months when glued with G is different from the proportion of books lasting at least 6 months when glued with K. ✓

(C) The p-value is less than α, and the null hypothesis is not rejected. There is not convincing evidence to support the claim that the proportion of books lasting at least 6 months when glued with G is different from the proportion of books lasting at least 6 months when glued with K.

(D) The p-value is greater than α, and the null hypothesis is rejected. There is convincing evidence to support the claim that the proportion of books lasting at least 6 months when glued with G is different from the proportion of books lasting at least 6 months when glued with K.

(E) The p-value is greater than α, and the null hypothesis is not rejected. There is not convincing evidence to support the claim that the proportion of books lasting at least 6 months when glued with G is different from the proportion of books lasting at least 6 months when glued with K.

Answer B

Correct. Since $0.006 < 0.01$, the null hypothesis is rejected. There is convincing statistical evidence that the proportion of books lasting at least 6 months when glued with G is different from the proportion of books lasting at least 6 months when glued with K.

6. Market researchers selected a random sample of people from region A and a random sample of people from region B. The researchers asked the people in the samples whether they had tried a new product. The difference between the sample proportions (B minus A) of people in the regions who indicated they had tried the new product was 0.15. Under the assumption that all conditions for inference were met, a hypothesis test was conducted with the alternative hypothesis being that the population proportion of B is greater than that of A. The p-value of the test was 0.34.

Which of the following is the correct interpretation of the p-value?

(A) If the difference in proportions of people who have tried the new product between the two populations is actually 0.15, the probability of observing that difference is 0.34.

(B) If the difference in proportions of people who have tried the new product between the two populations is actually 0.34, the probability of observing that difference is 0.15.

(C) If the proportions of all people who have tried the new product is the same for both regions, the probability of observing a difference of at least 0.15 is 0.34. ✓

(D) If the proportions of all people who have tried the new product is the same for both regions, the probability of observing a difference of at most 0.15 is 0.34.

(E) If the proportions of all people who have tried the new product is the same for both regions, the probability of observing a difference equal to 0.15 is 0.34.

Answer C

Correct. The direction of the test is to the right of 0.15, because the researchers hypothesize that the proportion in region B is greater than the proportion in region A. The p-value is the probability of observing a sample difference of at least 0.15 if the null hypothesis is true.

7. Show all your work. Indicate clearly the methods you use, because you will be scored on the correctness of your methods as well as on the accuracy and completeness of your results and explanations.

A random sample of 100 people from region A and a random sample of 100 people from region B were surveyed about their grocery-shopping habits. From the region A sample, 16 percent of the people indicated that they shop for groceries online. From the region B sample, 24 percent of the people indicated that they shop for groceries online.

At the significance level of $\alpha = 0.05$, do the data provide convincing statistical evidence that there is a difference between the two regions for the population proportions of people who shop online for groceries? Complete the appropriate inference procedure to support your answer.

3-part Inference scoring

The primary goal of this question is to assess a student's ability to identify, set up, perform, and interpret the results of an appropriate hypothesis test to address a particular question. More specific goals are to assess a student's ability to (1) identify the appropriate procedure and state appropriate hypotheses; (2) check appropriate conditions, calculate a test statistic, and calculate a *p*-value; and (3) draw an appropriate conclusion, with justification, in the context of the study.

Scoring

Scoring steps 1, 2, and 3 are each scored as essentially correct (E), partially correct (P), or incorrect (I).

0	1	2	3	4 ✓

All three scoring parts essentially correct

- ☐ **Scoring Step 1 - part (a)** Essentially correct
- ☐ **Scoring Step 1 - part (a)** Partially correct
- ☐ **Scoring Step 1 - part (a)** Incorrect
- ☐ **Scoring Step 2 - part (a)** Essentially correct
- ☐ **Scoring Step 2 - part (a)** Partially correct
- ☐ **Scoring Step 2 - part (a)** Incorrect

- ☐ **Scoring Step 3 - part (a)** Essentially correct
- ☐ **Scoring Step 3 - part (a)** Partially correct
- ☐ **Scoring Step 3 - part (a)** Incorrect

Solution

Scoring step 1: Appropriate test and hypothesis

The appropriate test is the two-sample *z*-test for a difference in proportions.

Let p_{A} represent the proportion of all people from region A who shop online for groceries.

Let p_{B} represent the proportion of all people from region B who shop online for groceries.

The hypotheses are as follows.

$\mathrm{H}_0 : p_{\mathrm{A}} = p_{\mathrm{B}}$ (or $p_{\mathrm{A}} - p_{\mathrm{B}} = 0$)
$\mathrm{H}_{\mathrm{a}} : p_{\mathrm{A}} \neq p_{\mathrm{B}}$ (or $p_{\mathrm{A}} - p_{\mathrm{B}} \neq 0$)

Scoring

Scoring step 1 is scored as follows:

Essentially correct (E) if the response satisfies the following three components.

· The correct test is identified.

· The null and alternative hypotheses are stated correctly.

· Context is explicitly stated or is implied through labeling of variables.

Partially correct (P) if the response satisfies only two of the three components.

Incorrect (I) if the response does not meet the criteria for E or P.

Note: To satisfy component 2, the hypotheses must be in terms of the population and use notation associated with the parameter (p) and not the sample ($\hat{p}$).

Solution

Scoring step 2: Conditions and calculations

Conditions:

· Data were collected from random samples as stated.

· The sampling method was probably conducted without replacement. It is reasonable to assume that the population of people in each region was greater than 10 times the sample size of 100.

· Both sample sizes are large enough:

The combined (or pooled) proportion is $\hat{p}_C = \frac{100(0.16)+100(0.24)}{100+100} = 0.20.$

$n_{\mathrm{A}}\hat{p}_{\mathrm{C}} = 100(0.20) \geq 10$
$n_{\mathrm{A}}(1-\hat{p}_{\mathrm{C}}) = 100(0.80) \geq 10$
$n_{\mathrm{B}}\hat{p}_{\mathrm{C}} = 100(0.20) \geq 10$
$n_{\mathrm{B}}(1-\hat{p}_{\mathrm{C}}) = 100(0.80) \geq 10$

Calculations:

The test statistic is $z = \frac{0.16-0.24}{\sqrt{(0.20)(0.80)(\frac{1}{100}+\frac{1}{100})}} \approx -1.41.$

The *p*-value is $2[P(z \leq -1.41)] \approx 0.1585.$

Scoring

Scoring step 2 is scored as follows:

Essentially correct (E) if the response satisfies the following three components.

· The three conditions are checked.

· The correct test statistic is computed.

· A *p*-value consistent with the computed test statistic.

Partially correct (P) if the response satisfies only two of the three components.

Incorrect (I) if the response does not meet the criteria for E or P.

Notes:

· The normality condition can be checked using a value of 5 instead of 10 (i.e that is, $n_A\hat{p}_C \geq 5$).

· The normality condition can be checked using the non-pooled proportions (i.e that is, $n_A\hat{p}_A \geq 10$ or $n_A\hat{p}_A \geq 5$).

· Component 2 can be satisfied with a test statistic of $z = 1.41$, which would occur when $\hat{p}_B - \hat{p}_A$ is used in the numerator.

Solution

Scoring step 3: Justification of conclusion

The *p*-value of 0.1585 is greater than the significance level of 0.05. We fail to reject the null hypothesis. The data do not provide sufficient evidence to conclude that the proportions of all people who shop for groceries online is different between the two regions A and B.

Scoring

Scoring step 3 is scored as follows:

Essentially correct (E) if the response satisfies the following three components.

· Explicitly compares the *p*-value to the significance level.

· Correct decision about the null hypothesis that is consistent with the computed *p*-value.

· Statement of conclusion in context.

Partially correct (P) if the response includes only two of the three conditions.

Incorrect (I) if the response does not meet the criteria for E or P.

Note: Any response that implies accepting the null hypothesis should be scored an I.

8. Maria has two routes, E and W, she can take when commuting to work. Both routes go through a railroad crossing, and sometimes she needs to stop at the crossing to allow trains to pass. She claims that the proportion of times she needs to stop when taking route E is different from the proportion of times she needs to stop when taking route W. She conducted the following hypothesis test at the significance level of $\alpha = 0.10$.

$H_0 : p_E = p_W$
$H_a : p_E \neq p_W$

In the hypotheses, p_E represents the proportion of times she needs to stop at the crossing when using route E, and p_W represents the proportion of times she needs to stop at the crossing when using route W.

All conditions for inference were met, and the resulting p-value was 0.37. Which of the following is the correct decision for the test?

(A) The p-value is less than α, and the null hypothesis is rejected. There is convincing evidence to support the claim that the proportion of times she needs to stop at the crossing is different for the different routes.

(B) The p-value is greater than α, and the null hypothesis is not rejected. There is convincing evidence to support the claim that the proportion of times she needs to stop at the crossing is the same for the different routes.

(C) The p-value is greater than α, and the null hypothesis is rejected. There is not convincing evidence to support the claim that the proportion of times she needs to stop at the crossing is different for the different routes.

(D) The p-value is greater than α, and the null hypothesis is not rejected. There is not convincing evidence to support the claim that the proportion of times she needs to stop at the crossing is different for the different routes. ✓

(E) The p-value is greater than α, and the null hypothesis is rejected. There is convincing evidence to support the claim that the proportion of times she needs to stop at the crossing is different for the different routes.

Answer D

Correct. The claim is not supported. Since $0.37 > 0.10$, the null hypothesis is not rejected. There is not

convincing statistical evidence that the proportion of times she needs to stop at the crossing is different for the different routes.

9. Biologists were studying the proportions of cats that had spotted markings on their fur in two populations of cats, C and F. An independent random sample of cats was taken from each population, and the difference between the sample proportions of cats with the spotted markings (C minus F) was 0.62. Under the assumption that all conditions for inference were met, a hypothesis test was conducted with the alternative hypothesis being that the population proportions are not equal. The p-value of the test was 0.01.

Which of the following is the correct interpretation of the p-value?

(A) If the proportions of all cats with spotted markings is the same for both populations, the probability of observing a sample difference of at least 0.62 or at most -0.62 is 0.01. ✓

(B) If the proportions of all cats with spotted markings is the same for both populations, the probability of observing a sample difference of at least 0.62 is 0.01.

(C) If the proportions of all cats with spotted markings is the same for both populations, the probability of observing a sample difference of at most -0.62 is 0.01.

(D) If the difference in proportions of cats with spotted markings between the two populations is actually 0.62, the probability of observing that difference is 0.01.

(E) If the difference in proportions of cats with spotted markings between the two populations is actually 0.01, the probability of observing that difference is 0.62.

Answer A

Correct. The test is two-sided, so the p-value refers to the union of area in both tails. It is the probability of observing a sample difference of at least 0.62 or of at most -0.62 if the null hypothesis is true.

10. A political scientist claims that negative advertising on television affects younger voters more than it affects older voters. To test this claim, the scientist obtained data from two random samples of voters categorized into two age-groups, older and younger. The null hypothesis was that there was no difference in the proportions of voters in the two age-groups who would be affected by negative ads. The alternative hypothesis was that the proportion of younger voters affected would be greater than the proportion of older voters affected.

Assuming all conditions for inference were met, the scientist conducted the test at a significance level of $\alpha = 0.05$. The resulting p-value was 0.206. Which of the following is the correct decision for the test?

(A) The p-value is less than α, and the null hypothesis is rejected. There is convincing evidence to support the claim that younger voters are more affected by negative ads than are older voters.

(B) The p-value is less than α, and the null hypothesis is rejected. There is not convincing evidence to support the claim that younger voters are more affected by negative ads than are older voters.

(C) The p-value is greater than α, and the null hypothesis is not rejected. There is convincing evidence to support the claim that younger voters are more affected by negative ads than are older voters.

(D) The p-value is greater than α, and the null hypothesis is rejected. There is convincing evidence to support the claim that younger voters are more affected by negative ads than are older voters.

(E) The p-value is greater than α, and the null hypothesis is not rejected. There is not convincing evidence to support the claim that younger voters are more affected by negative ads than are older voters. ✓

Answer E

Correct. The claim is not supported. Since $0.206 > 0.05$, the null hypothesis is not rejected. There is not convincing statistical evidence that the proportion of younger voters affected by negative ads is greater than the proportion of older voters affected by negative ads.