4. Tumbleweed, commonly found in the western United States, is the dried structure of certain plants that are blown by the wind. Kochia, a type of plant that turns into tumbleweed at the end of the summer, is a problem for farmers because it takes nutrients away from soil that would otherwise go to more beneficial plants. Scientists are concerned that kochia plants are becoming resistant to the most commonly used herbicide, glyphosate. In 2014, 19.7 percent of 61 randomly selected kochia plants were resistant to glyphosate. In 2017, 38.5 percent of 52 randomly selected kochia plants were resistant to glyphosate. Do the data provide convincing statistical evidence, at the level of $\alpha = 0.05,$ that there has been an increase in the proportion of all kochia plants that are resistant to glyphosate?

5. A company that manufactures smartphones developed a new battery that has a longer life span than that of a traditional battery. From the date of purchase of a smartphone, the distribution of the life span of the new battery is approximately normal with mean 30 months and standard deviation 8 months. For the price of $50, the company offers a two-year warranty on the new battery for customers who purchase a smartphone. The warranty guarantees that the smartphone will be replaced at no cost to the customer if the battery no longer works within 24 months from the date of purchase.

   (a) In how many months from the date of purchase is it expected that 25 percent of the batteries will no longer work? Justify your answer.

   (b) Suppose one customer who purchases the warranty is selected at random. What is the probability that the customer selected will require a replacement within 24 months from the date of purchase because the battery no longer works?

   (c) The company has a gain of $50 for each customer who purchases a warranty but does not require a replacement. The company has a loss (negative gain) of $150 for each customer who purchases a warranty and does require a replacement. What is the expected value of the gain for the company for each warranty purchased?

GO ON TO THE NEXT PAGE.

Question 4

Intent of Question:

The primary goals of this question were to assess a student's ability to perform an appropriate hypothesis test to address a particular question. More specific goals were to assess students' ability to state appropriate hypotheses, identify the appropriate statistical test procedure, check appropriate assumptions/conditions for inference; calculate a correct test statistic and p-value; and draw a correct conclusion, with justification, in the context of the study.

Solution

Section 1:

Let p_{14} represent the proportion of the population of kochia plants in the western United States that were resistant to glyphosate in 2014. Let p_{17} represent the proportion of the population of kochia plants in the western United States that were resistant to glyphosate in 2017.

The null hypothesis $H_0 : p_{17} - p_{14} = 0$ is to be tested against the alternative hypothesis $H_a : p_{17} - p_{14} > 0$.

An appropriate inference procedure is a two-sample z-test for a difference in proportions. The formula for the test statistic is:

$$z = \frac{\hat{p}_{17} - \hat{p}_{14}}{\sqrt{\left(\dfrac{\hat{p}_c (1 - \hat{p}_c)}{n_{17}} + \dfrac{\hat{p}_c (1 - \hat{p}_c)}{n_{14}} \right)}}$$

where $\hat{p}_c = \dfrac{n_{14} \hat{p}_{14} + n_{17} \hat{p}_{17}}{n_{14} + n_{17}}$ is a pooled estimate of the proportion of resistant plants for 2014 and 2017 combined.

Section 2:

The first condition for applying the test is that the data are gathered from independent random samples from the populations of kochia plants in the western United States in 2014 and 2017. The question indicates that a random sample of 61 kochia plants was taken in 2014 and a second random sample of 52 kochia plants was taken in 2017. It is reasonable to assume that the 2017 sample of plants was in no way influenced by the 2014 sample of plants.

The second condition is that the sampling distribution of the test statistic is approximately normal. This condition is satisfied because the expected counts under the null hypothesis are all greater than 10. The pooled estimate of the proportion of resistant plants is $\hat{p}_c = \dfrac{(61)(0.197) + (52)(0.385)}{61 + 52} \approx 0.2835$. The estimates of the expected counts are
$61(0.2835) \approx 17.29, \ 61(1 - 0.2835) \approx 43.71, \ 52(0.2835) \approx 14.74, \ 52(1 - 0.2835) \approx 37.26,$ all of which are greater than 10.

Because sampling must have been done without replacement, the independence condition for each sample should be checked. Information on the population sizes of kochia plants is not given for either 2014 or 2017, but it is reasonable to assume that each population has millions of plants. Therefore it is reasonable to assume that the sample sizes are less than 10 percent of the respective population sizes.

Using the pooled estimate of the proportion of resistant plants, $\hat{p}_c \approx 0.2835,$ the value of the test statistic is:

$$z = \frac{0.385 - 0.197}{\sqrt{\left(\frac{(0.2835)(0.7165)}{61} + \frac{(0.2835)(0.7165)}{52}\right)}} \approx 2.21$$

The *p*-value is 0.0135.

Section 3:

Because the *p*-value is less than $\alpha = 0.05,$ there is convincing statistical evidence to conclude that the proportion of resistant plants in the 2017 population of kochia plants is greater than the proportion of resistant plants in the 2014 population of kochia plants.

Scoring

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

Section 1 is scored as follows:

Essentially correct (E) if the response satisfies components 1 and 4 *AND* at least one of the remaining components:
1. Hypotheses imply equality of proportions in the null hypothesis and correct direction in the alternative hypothesis, which utilize an appropriate population parameter in words or symbols.
2. Identifies parameters that are population proportions.
3. Both parameters are correctly defined as proportions of resistant plants in 2014 and 2017.
4. The two-sample *z*-test for proportions is identified by name or formula

Partially correct (P) if the response does not meet the requirement for E, but at least two of the components are satisfied.

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

Notes:
- Correct ways to state the <u>null</u> hypothesis that satisfy component 1:
    - $H_0 : p_{17} = p_{14}$ or $H_0 : p_{17} - p_{14} = 0$
    - $H_0 : p_{17} \leq p_{14}$ or $H_0 : p_{17} - p_{14} \leq 0$
    - $H_0 : p_{14} \geq p_{17}$ or $H_0 : p_{14} - p_{17} \geq 0$
    Correct ways to state the alternative hypothesis that satisfy component 1:
    - $H_a : p_{17} > p_{14}$ or $H_a : p_{17} - p_{14} > 0$
    - $H_a : p_{14} < p_{17}$ or $H_a : p_{14} - p_{17} < 0$
    Incorrect ways to state the <u>null</u> hypothesis that do not satisfy component 1:
    - $H_0 : p_{17} < p_{14}$ or $H_0 : p_{17} - p_{14} < 0$
    - $H_0 : p_{14} > p_{17}$ or $H_0 : p_{14} - p_{17} > 0$
    Incorrect ways to state the <u>alternative</u> hypothesis that do not satisfy component 1:
    - $H_a : p_{17} \neq p_{14}$ or $H_a : p_{17} - p_{14} \neq 0$
    - $H_a : p_{17} < p_{14}$ or $H_a : p_{17} - p_{14} < 0$

- Examples for components 2 and 3:
  - o Satisfies both components 2 and 3:
    - p_{17} is the proportion of resistant plants

      p_{14} is the proportion of resistant plants
  - o Satisfies component 2 but not component 3:
    - p_1 is the proportion of resistant plants

      p_2 is the proportion of resistant plants
    - p_{17} is the proportion of plants

      p_{14} is the proportion of plants
    - p_{17}, p_{14}
    - p_1, p_2
- If the test is correctly identified by name, but then an incorrect formula is stated, this is considered to be a parallel response and component 4 is not satisfied.
- If the test identifies an unpooled two sample z-test for a difference in proportions as the correct test or formula, component 4 is satisfied.

Section 2 is scored as follows:

Essentially correct (E) if the response satisfies components 1 and 2 *AND* at least two of the remaining components:
1. Notes that the use of random samples in 2014 and 2017 satisfies the randomness condition.
2. Checks for approximate normality of the test statistic by showing that the expected numbers of resistant and non-resistant kochia plants are both larger than some commonly accepted criterion (e.g. 5 or 10) for both samples.
3. Notes that the populations of kochia plants must be extremely large in both years, thus satisfies the independence (10%) conditions.
4. Reports a correct value of the z-test statistic.
5. Reports a p-value that is consistent with the stated alternative hypothesis and reported test statistic.

Partially correct (P) if the response does not meet the criteria for E, but at least two of the five components are satisfied.

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

Notes:
- For the randomness component it is minimally acceptable to say "random samples—check" or "SRSs—check." The important concept is that the study used two independent random samples. Although it is not known if a SRS was taken versus another type of random sample, it is minimally acceptable to indicate SRSs since the sampling method is unknown. If the response implies that random assignment was used, the randomness component is not satisfied.

- To satisfy component 2, the response must include actual numbers, or a formula with numbers plugged in, as well as a clear indication of comparison of the four quantities to some standard criterion, such as 5 or 10, or the statement that each such quantity is large enough. If a formula with numbers is used, simplification is NOT required.

  Examples of acceptable quantities (comparison still must be made):
  - 12, 49, 20, 32
  - 12.017, 48.983, 20.02, 31.98
  - $61(0.197), 61(1 - 0.197), 52(0.385), 52(1 - 0.385)$

  Examples of unacceptable quantities:
  - $n_{17}\hat{p}_{17}, n_{17}(1 - \hat{p}_{17}), n_{14}\hat{p}_{14}, n_{14}(1 - \hat{p}_{14})$
  - $n_{17}p_{17}, n_{17}(1 - p_{17}), n_{14}p_{14}, n_{14}(1 - p_{14})$
  - $n_{17}\hat{p}_{c}, n_{17}(1 - \hat{p}_{c}), n_{14}\hat{p}_{c}, n_{14}(1 - \hat{p}_{c})$
  - $61\hat{p}_{c}, 61(1 - \hat{p}_{c}), 52\hat{p}_{c}, 52(1 - \hat{p}_{c})$

- The test statistics for the pooled and unpooled z-tests are 2.21 and 2.22 respectively, thus they are close to the same value. If the response provides the unpooled formula but then states a pooled test statistic, component 4 is satisfied. If the response provides the pooled formula but then states an unpooled test statistic, component 4 is satisfied.
- If the response uses a critical value approach rather than a p-value approach, then the correct critical value of -1.645 or 1.645, that is consistent with the alternative hypothesis, satisfies component 5.
- If the response did not satisfy component 1 in section 1 because a two-tailed alternative was stated or the direction of the alternative was incorrect, then the p-value in component 5 should be consistent with the stated alternative. If the response omits hypotheses or other incorrect hypotheses are stated, assume the correct alternative hypothesis is provided when scoring component 5.

Section 3 is scored as follows:

Essentially correct (E) if the response includes the following three components:
1. Provides justification of the conclusion based on a correct comparison between a stated p-value and an alpha value of 0.05.
2. Provides a correct conclusion consistent with the alternative hypothesis.
3. The conclusion is stated in context.

Partially correct (P)
   if the response satisfies components 1 and 2

   OR

   if the response satisfies components 2 and 3

   OR

   if the response satisfies components 1 and 3 *AND*, based on the p-value from section 2, either
      - the conclusion correctly rejects the null hypothesis but does not state that there is convincing evidence for the <u>alternative</u> hypothesis
         OR
      - the conclusion correctly fails to reject the null hypothesis but does not state there is not convincing evidence for the <u>alternative</u> hypothesis.

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

Notes:
- If the conclusion is consistent with a reasonable, but incorrect, *p*-value from section 2, and is presented in context with justification based on comparison of the *p*-value to the level of significance, then section 3 is scored E.
- If the response implies that the outcome of the hypothesis test is a "proof" of either a true or false null, the score is lowered one level (that is, from E to P, or from P to I).
- If an incorrect interpretation of the *p*-value is given, the score is lowered one level (that is, from E to P, or from P to I).
- If the response uses a critical value approach rather than a *p*-value approach, then the correct critical value of -1.645 or 1.645 replaces the *p*-value in section 2, and comparison of the test statistic from section 2 to the critical value (e.g. $2.21 > 1.645$) satisfies component 1.
- If the response clearly states a reasonable level of significance that differs from 0.05 and provides a justification and conclusion in context based on that justification, the response is scored E.
- If the response provides the incorrect comparison between the stated *p*-value and the level of significance, but the conclusion is consistent with the given comparison and the alternative hypothesis, then component 2 is satisfied.
- If the response did not satisfy component 1 in section 1 because a two-tailed alternative was stated or the direction of the alternative was incorrect, then the conclusion component 2 should be consistent with the stated alternative. If the response states other incorrect hypotheses or omits hypotheses, assume the correct alternative hypothesis is provided when scoring component 2.

Alternative Approach
Two-Sided Confidence Interval for Difference in Two Population Proportions

Section 1 is scored E, P, or I according to the guidelines in section 1 for a 2 sample *z*-test for proportions.

Notes:
- To satisfy component 4, the two-sample confidence interval for a difference in two proportions should be identified by name or formula by referring to the *z*-distribution and two proportions.

Section 2 is scored as follows:
Essentially correct (E) if the response satisfies components 1 and 2 *AND* at least one of the remaining components:

1. Notes that the use of random sample<u>s</u> in 2014 and 2017 satisfies the randomness condition.
2. Checks for approximate normality of the test statistic by showing that the observed numbers of resistant and non-resistant kochia plants are both larger than some commonly accepted criterion (say 5 or 10) for both samples.
3. Notes that the populations of kochia plants must be extremely large in both years, thus satisfies the independence (10%) conditions.
4. Reports the correct confidence interval that is consistent with the stated alternative hypothesis.

Partially correct (P) if the response does not meet the criteria for E, but at least two of the four components are satisfied.

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

Notes:
- Examples of correct 90% confidence intervals to address a one-sided alternative for $\alpha = 0.05$ are:
  $(-0.327, -0.049)$
  $(0.049, 0.327)$
- Examples of correct 95% confidence intervals to address a two-sided alternative for $\alpha = 0.05$ are:
  $(-0.353, -0.022)$
  $(0.022, 0.353)$

Section 3 is scored E, P or I according to the guidelines in section 3 for a 2 sample *z*-test for proportions.

Notes:
- Component 1 is satisfied if a confidence interval that is consistent with the alternative hypothesis is given and the appropriate interval endpoint(s) are compared to zero.

Overall Notes:
- If the response constructs two separate one-proportion *z*-intervals for 2014 and 2017, then sections 1 and 2 are scored as above, and section 3 is scored I.

Alternative Approach
Chi-square test for homogeneity

The value of the Pearson chi-square test statistic (uncorrected) is 4.8821 with 1 degree of freedom and a *p*-value of 0.02714. This value is the same as the square of the pooled *z*-statistic, so it is the same test, but the *p*-value is for a two-sided alternative. This *p*-value could be divided by 2 to obtain an appropriate *p*-value for the one-sided alternative, but the sample data needs to be examined to determine the correct direction of the alternative.

Section 1 is scored E, P, or I according to the guidelines in section 1 for a 2 sample *z*-test for proportions.

Notes:
- Examples of unacceptable hypotheses:
  $H_0 : p_{17}$ and p_{14} are independent or $H_0 : p_{17}$ and p_{14} have no association
  $H_a : p_{17}$ and p_{14} are dependent or $H_a : p_{17}$ and p_{14} have an association
  $H_a : p_{17} \neq p_{14}$ or $H_a : p_{17} \geq p_{14}$ or $H_a : p_{14} \leq p_{17}$
- To satisfy component 4, a chi-square test for homogeneity is identified by name or formula.

Section 2 is scored E, P, or I according to the guidelines in section 2 for a 2 sample *z*-test for proportions.

Notes:
- To satisfy component 4, the reported value of the chi-square test statistic is 4.8821.
- If a one sided alternative hypothesis is given in section 1, then to satisfy component 5, the reported value of the *p*-value is 0.01357.
- If a two sided alternative hypothesis is given in section 1, then to satisfy component 5, the reported *p*-value is 0.02714.

Section 3 is scored E, P or I according to the guidelines in section 3 for a 2 sample *z*-test for proportions.

Notes:
- If the response clearly indicates that the two sample proportions were used to determine the correct one-sided direction and a *p*-value of 0.0135 was used as justification, section 3 is scored E.
- If a correct conclusion is reached based on the *p*-value of 0.02714 and a two-sided alternative, then section 3 is scored at most P.
- If the final response justifies the conclusion based on the *p*-value of 0.0135 but does not explicitly indicate how the correct direction was determined, section 3 is scored at most P.

4   **Complete Response**

        Three sections essentially correct

3   **Substantial Response**

        Two sections essentially correct and one section partially correct

2   **Developing Response**

        Two sections essentially correct and no sections partially correct

OR

        One section essentially correct and one or two sections partially correct

OR

        Three sections partially correct

1   **Minimal Response**

        One section essentially correct

OR

        No section essentially correct and two sections partially correct