5. A recent report stated that less than 35 percent of the adult residents in a certain city will be able to pass a physical fitness test. Consequently, the city's Recreation Department is trying to convince the City Council to fund more physical fitness programs. The council is facing budget constraints and is skeptical of the report. The council will fund more physical fitness programs only if the Recreation Department can provide convincing evidence that the report is true.

The Recreation Department plans to collect data from a sample of 185 adult residents in the city. A test of significance will be conducted at a significance level of $\alpha = 0.05$ for the following hypotheses.

$$H_0 : p = 0.35$$
$$H_a : p < 0.35,$$

where p is the proportion of adult residents in the city who are able to pass the physical fitness test.

(a) Describe what a Type II error would be in the context of the study, and also describe a consequence of making this type of error.

(b) The Recreation Department recruits 185 adult residents who volunteer to take the physical fitness test. The test is passed by 77 of the 185 volunteers, resulting in a p-value of 0.97 for the hypotheses stated above. If it was reasonable to conduct a test of significance for the hypotheses stated above using the data collected from the 185 volunteers, what would the p-value of 0.97 lead you to conclude?

(c) Describe the primary flaw in the study described in part (b), and explain why it is a concern.

Solution

Part (a):

In the context of the study, a Type II error means failing to reject the null hypothesis that 35 percent of adult residents in the city are able to pass the test when, in reality, less than 35 percent are able to pass the test. The consequence of this error is that the council would not fund the program, and the city would continue to have a smaller proportion of physically fit residents than the council would like.

Part (b):

Because the p-value of 0.97 is larger than $\alpha = 0.05$, we fail to reject the null hypothesis. There is not convincing evidence that the proportion of adult residents in the city who are able to pass the physical fitness test is less than 0.35. After all, the sample proportion of $\hat{p} = 0.416$ is actually higher than 0.35, which is in the opposite direction of the alternative hypothesis.

Part (c):

This is not a randomly selected sample because the sample was selected by recruiting volunteers. It seems reasonable to think that volunteers would be more physically fit than the population of city adults as a whole. Therefore, the sample proportion will likely overestimate the population proportion of adult residents in the city who are able to pass the physical fitness test.

5. A large pet store buys the identical species of adult tropical fish from two different suppliers—Buy-Rite Pets and Fish Friends. Several of the managers at the pet store suspect that the lengths of the fish from Fish Friends are consistently greater than the lengths of the fish from Buy-Rite Pets. Random samples of 8 adult fish of the species from Buy-Rite Pets and 10 adult fish of the same species from Fish Friends were selected and the lengths of the fish, in inches, were recorded, as shown in the table below.

	Length of Fish	Mean	Standard Deviation
Buy-Rite Pets ($n_B = 8$)	3.4 2.7 3.3 4.1 3.5 3.4 3.0 3.8	3.40	0.434
Fish Friends ($n_F = 10$)	3.3 2.9 4.2 3.1 4.2 4.0 3.4 3.2 3.7 2.6	3.46	0.550

Do the data provide convincing evidence that the mean length of the adult fish of the species from Fish Friends is greater than the mean length of the adult fish of the same species from Buy-Rite Pets?

Solution

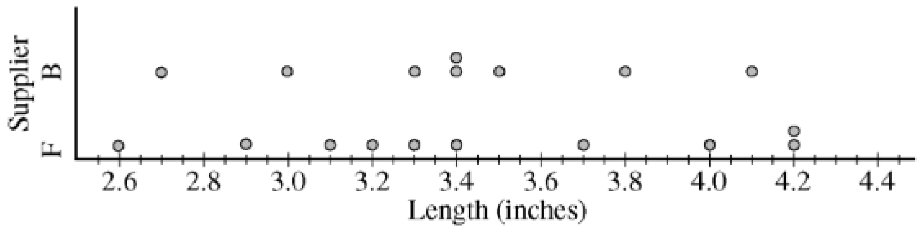
Step 1: States a correct pair of hypotheses

Let μ_B represent the population mean length of all adult fish of this species from Buy-Rite Pets, and
let μ_F represent the population mean length of all adult fish of this species from Fish Friends.

The hypotheses to be tested are H_0: $\mu_B = \mu_F$ versus H_a: $\mu_B < \mu_F$.

Step 2: Identifies a correct test procedure (by name or by formula) and checks appropriate conditions

The appropriate test is a two-sample *t*-test. The first condition is that the samples are independent random samples from the two populations. This was stated in the question. The second condition is that the population distributions of fish lengths are normal. The following dotplots reveal no obvious departures from normality, so it appears reasonable to proceed with the two-sample *t*-test.



Step 3: Demonstrates correct mechanics, including the value of the test statistic, df and *p*-value (or rejection region)

The test statistic is: $t = \dfrac{\overline{x}_B - \overline{x}_F}{\sqrt{\dfrac{s_B^2}{n_B} + \dfrac{s_F^2}{n_F}}} = \dfrac{3.40 - 3.46}{\sqrt{\dfrac{0.434^2}{8} + \dfrac{0.550^2}{10}}} \approx -0.259$

With df = 15.99999, *p*-value = 0.3996.

Step 4: States a correct conclusion in the context of the problem, using the result of the statistical test

Because this *p*-value is larger than any conventional significance level (such as $\alpha = 0.10$ or $\alpha = 0.05$), we fail to reject H_0. The sample data do not provide convincing evidence to conclude that the mean length of the adult fish of the species from Fish Friends is greater than the mean length of the adult fish of the same species from Buy-Rite Pets.

4. A husband and wife, Mike and Lori, share a digital music player that has a feature that randomly selects which song to play. A total of 2,384 songs were loaded onto the player, some by Mike and the rest by Lori. Suppose that when the player was in the random-selection mode, 13 of the first 50 songs selected were songs loaded by Lori.

   (a) Construct and interpret a 90 percent confidence interval for the proportion of songs on the player that were loaded by Lori.

   (b) Mike and Lori are unsure about whether the player samples the songs with replacement or without replacement when the player is in random-selection mode. Explain why this distinction is not important for the construction of the interval in part (a).

Solution

Part (a):

The sample proportion of songs that were loaded by Lori is $\hat{p} = \frac{13}{50} = 0.26$. The conditions for constructing a confidence interval are satisfied because: (1) the problem states that the 50 songs in the sample were randomly selected, and (2) $n \times \hat{p} = 13$ and $n \times (1 - \hat{p}) = 37$ are both at least 10. A 90 percent confidence interval for the population proportion *p*, the actual proportion of all songs on the player that were loaded by Lori, is:

$$0.26 \pm 1.645\sqrt{\frac{.26 \times (1 - .26)}{50}} = 0.26 \pm 0.102 = (0.158, 0.362).$$

We can be 90 percent confident that for the population of all songs on the digital music player, the proportion of songs that were loaded by Lori is between 0.158 and 0.362.

Part (b):

The sample size of 50 is quite small compared with the population size of 2,384. The usual criterion for checking whether one can disregard the distinction between sampling with or without replacement is to check whether the ratio of the population size to the sample size is large, such as at least 10 or at least 20. In this case the ratio is $\frac{2{,}384}{50} = 47.7$, so the criterion is clearly met, and the confidence interval procedure in part (a) is valid.

4. One of the two fire stations in a certain town responds to calls in the northern half of the town, and the other fire station responds to calls in the southern half of the town. One of the town council members believes that the two fire stations have different mean response times. Response time is measured by the difference between the time an emergency call comes into the fire station and the time the first fire truck arrives at the scene of the fire.

   Data were collected to investigate whether the council member's belief is correct. A random sample of 50 calls selected from the northern fire station had a mean response time of 4.3 minutes with a standard deviation of 3.7 minutes. A random sample of 50 calls selected from the southern fire station had a mean response time of 5.3 minutes with a standard deviation of 3.2 minutes.

   (a) Construct and interpret a 95 percent confidence interval for the difference in mean response times between the two fire stations.

   (b) Does the confidence interval in part (a) support the council member's belief that the two fire stations have different mean response times? Explain.

Solution

Part (a):

Step 1: Identify the appropriate confidence interval by name or formula and check for appropriate conditions.

The two-sample *t* interval for $\mu_N - \mu_S$, the difference in population mean response times, is

$$(\bar{x}_N - \bar{x}_S) \pm t^* \sqrt{\frac{s_N^2}{n_N} + \frac{s_S^2}{n_S}}$$

where μ_N denotes the mean response for calls from the northern fire station and μ_S denotes the mean response for calls from the southern fire station.

Conditions: 1. Independent random samples
2. Large samples or normal population distributions

A random sample of 50 calls was selected from the northern fire station, independent of the random sample of 50 calls selected from the southern fire station.

The use of the two-sample *t* interval is reasonable because both sample sizes are large ($n_N = 50 > 30$ and $n_S = 50 > 30$), and by the central limit theorem, the sampling distributions for the two sample means are approximately normal. Therefore the sampling distribution of the difference of the sample means $\bar{x}_N - \bar{x}_S$ is approximately normal.

Step 2: Correct Mechanics

Unequal variances: Degrees of freedom = 96.

$$(4.3 - 5.3) \pm 1.985\sqrt{\frac{3.7^2}{50} + \frac{3.2^2}{50}}$$

$$-1.0 \pm 1.985 \times .6918$$

$$(-2.37, 0.37)$$

Step 3: Interpretation

Based on these samples, one can be 95 percent confident that the difference in the population mean response times (northern - southern) is between -2.37 minutes and 0.37 minutes.

Part (b):

Zero is within the 95 percent confidence interval of plausible values for the difference in population means. Therefore this confidence interval does not support the council member's belief that there i difference in mean response times for the two fire stations.

5. For many years, the medically accepted practice of giving aid to a person experiencing a heart attack was to have the person who placed the emergency call administer chest compression (CC) plus standard mouth-to-mouth resuscitation (MMR) to the heart attack patient until the emergency response team arrived. However, some researchers believed that CC alone would be a more effective approach.

   In the 1990s a study was conducted in Seattle in which 518 cases were randomly assigned to treatments: 278 to CC plus standard MMR and 240 to CC alone. A total of 64 patients survived the heart attack: 29 in the group receiving CC plus standard MMR, and 35 in the group receiving CC alone. A test of significance was conducted on the following hypotheses.

   H_0: The survival rates for the two treatments are equal.

   H_a: The treatment that uses CC alone produces a higher survival rate.

   This test resulted in a *p*-value of 0.0761.

   (a) Interpret what this *p*-value measures in the context of this study.

   (b) Based on this *p*-value and study design, what conclusion should be drawn in the context of this study? Use a significance level of $\alpha = 0.05$.

   (c) Based on your conclusion in part (b), which type of error, Type I or Type II, could have been made? What is one potential consequence of this error?

Solution

Part (a):

The p-value of 0.0761 measures the chance of observing a difference between the two sample proportions $(\hat{p}_{CC} - \hat{p}_{CC+MMR})$ as large as or larger than the one observed, if the survival rates for the two treatments (CC alone and CC + MMR) are in fact the same.

Part (b):

Because the p-value of 0.0761 is greater than 0.05, the null hypothesis should not be rejected. That is, there is not sufficient evidence to conclude that the treatment "CC alone" produces a higher survival rate than the standard treatment "CC + MMR."

Part (c):

Because the null hypothesis was not rejected, a Type II error could have occurred. A possible consequence is that CC + MMR would continue as the accepted practice when, in fact, CC alone would result in a higher survival rate.

3. A French study was conducted in the 1990s to compare the effectiveness of using an instrument called a cardiopump with the effectiveness of using traditional cardiopulmonary resuscitation (CPR) in saving lives of heart attack victims. Heart attack patients in participating cities were treated with either a cardiopump or CPR, depending on whether the individual's heart attack occurred on an even-numbered or an odd-numbered day of the month. Before the start of the study, a coin was tossed to determine which treatment, a cardiopump or CPR, was given on the even-numbered days. The other treatment was given on the odd-numbered days. In total, 754 patients were treated with a cardiopump, and 37 survived at least one year; while 746 patients were treated with CPR, and 15 survived at least one year.

   (a) The conditions for inference are satisfied in the study. State the conditions and indicate how they are satisfied.

   (b) Perform a statistical test to determine whether the survival rate for patients treated with a cardiopump is significantly higher than the survival rate for patients treated with CPR.

Solution

Let A represent the cardiopump treatment, and let B represent the CPR treatment.

Let p_A = proportion of patients who will survive at least one year if treated with the cardiopump.
Let p_B = proportion of patients who will survive at least one year if treated with CPR.

Part (a):

Step 1: State the conditions for inference.

The conditions required for a two-sample z test of equal proportions for an experiment are:
1. Random assignment of treatments to subjects
2. Sufficiently large sample sizes

Step 2: Check the conditions.

1. If we assume that the relevant characteristics of people who have heart attacks on even-numbered and odd-numbered days are comparable, randomly assigning one treatment to be given on even-numbered days and the other to be given on odd-numbered days is a reasonable approximation to randomly assigning the two treatments to the available subjects.
2. The large sample condition is met because all of the following are at least 5 (or 10):

$n_A\hat{p}_A = 37 \geq 5$ or 10, $n_A(1-\hat{p}_A) = 717 \geq 5$ or 10
$n_B\hat{p}_B = 15 \geq 5$ or 10, $n_B(1-\hat{p}_B) = 731 \geq 5$ or 10

Part (b):

Step 1: State a correct pair of hypotheses.

$H_0: p_A - p_B = 0$ (or $p_A = p_B$)
$H_a: p_A - p_B > 0$ (or $p_A > p_B$)

Step 2: Identify a correct test by name or by formula.

Two-sample z test for proportions

OR

$$z = \frac{\hat{p}_A - \hat{p}_B}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_A} + \frac{1}{n_B}\right)}} = \frac{\hat{p}_A - \hat{p}_B}{\sqrt{\frac{\hat{p}(1-\hat{p})}{n_A} + \frac{\hat{p}(1-\hat{p})}{n_B}}} \text{ where } \hat{p} = \frac{n_A\hat{p}_A + n_B\hat{p}_B}{n_A + n_B}.$$

Step 3: Correct mechanics, including the value of the test statistic and p-value (or rejection region).

$$\hat{p}_A = \frac{37}{754} \approx 0.049 \qquad \hat{p}_B = \frac{15}{746} \approx 0.020 \qquad \hat{p} = \frac{37+15}{754+746} = \frac{52}{1500} \approx 0.035$$

$$z = \frac{\frac{37}{754} - \frac{15}{746}}{\sqrt{\frac{52}{1500}\left(1 - \frac{52}{1500}\right)\left(\frac{1}{754} + \frac{1}{746}\right)}} \approx 3.066$$

The p-value is 0.0011.

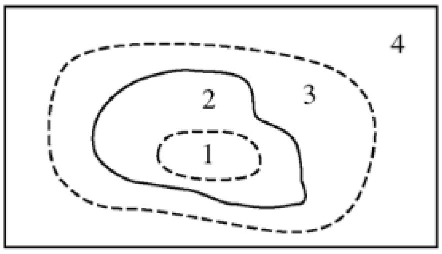
Step 4: State a correct conclusion in the context of the problem, using the result of the statistical test.

Because the p-value of 0.0011 is very small, that is, less than any reasonable significance level such as $\alpha = 0.01$, or $\alpha = 0.05$, we reject the null hypothesis. We have strong evidence to support the conclusion that the proportion of patients who survive when treated with the cardiopump is higher than the proportion of patients who survive when treated with CPR; that is, the survival rate is higher for patients treated with the cardiopump. (*OR*, If all of these patients had been assigned the cardiopump, we have strong evidence that the survival rate would be higher than if all of these patients had been assigned CPR.)

5. A study was conducted to determine where moose are found in a region containing a large burned area. A map of the study area was partitioned into the following four habitat types.

(1) Inside the burned area, not near the edge of the burned area,
(2) Inside the burned area, near the edge,
(3) Outside the burned area, near the edge, and
(4) Outside the burned area, not near the edge.

The figure below shows these four habitat types.



Note: Figure not drawn to scale.

The proportion of total acreage in each of the habitat types was determined for the study area. Using an aerial survey, moose locations were observed and classified into one of the four habitat types. The results are given in the table below.

Habitat Type	Proportion of Total Acreage	Number of Moose Observed
1	0.340	25
2	0.101	22
3	0.104	30
4	0.455	40
Total	1.000	117

(a) The researchers who are conducting the study expect the number of moose observed in a habitat type to be proportional to the amount of acreage of that type of habitat. Are the data consistent with this expectation? Conduct an appropriate statistical test to support your conclusion. Assume the conditions for inference are met.

(b) Relative to the proportion of total acreage, which habitat types did the moose seem to prefer? Explain.

Solution

Part (a):

Step 1: States a correct pair of hypotheses.

H_0 : Moose have no preference for habitat type.
H_a : Moose have a preference for habitat type.

OR

H_0 : The number of moose in each habitat type is proportional to the amount of acreage of that habitat type.
H_a : The number of moose in at least one habitat type is not proportional to the amount of acreage of that habitat type.

OR

$H_0 : p_1 = 0.340, p_2 = 0.101, p_3 = 0.104, p_4 = 0.455$, where p_i = the proportion of moose in habitat type *i*.
H_a : At least one of these proportions is incorrect.

Step 2: Identifies a correct test (by name or formula) and checks appropriate conditions.

- Chi-square goodness-of-fit test (or test for more than two proportions)

$$\chi^2 = \sum \frac{(\text{observed} - \text{expected})^2}{\text{expected}}$$

- The stem of the problem stated that conditions for inference are met.

Step 3: Correct mechanics, including the value of the test statistic, df, and *p*-value (or rejection region).

- The test statistic, with df = 4 − 1 = 3, is

$$\chi^2 = \frac{(25-39.780)^2}{39.780} + \frac{(22-11.817)^2}{11.817} + \frac{(30-12.168)^2}{12.168} + \frac{(40-53.235)^2}{53.235} = 43.6893.$$

- The *p*-value is $P(\chi_3^2 \geq 43.6893) < 0.0005$ (a calculator gives the *p*-value as 1.7569×10^{-9}).

Step 4: States a correct conclusion in the context of the problem, using the result of the statistical test.

The data are not consistent with the researchers' expectation. Because the *p*-value is less than $\alpha = 0.05$, we reject H_0. There is strong evidence that moose have a preference for habitat type.

OR

The data are not consistent with the researchers' expectation. If the null hypothesis is true and the numb of moose in each of the habitat types is proportional to the acreage in that habitat type, then we would observe a test statistic of 43.69 or one more extreme less than 0.05 percent of the time. There is strong evidence that moose have a preference for habitat type.

Part (b):

The moose seem to prefer habitat types 2 and 3. Relative to the proportion of total acreage, a higher proportion of moose were observed in each of these habitat types than expected. In habitat types 1 and 4, th observed proportion of moose was less than the expected proportion of moose, indicating that these two habitat types are less desirable.

OR

Habitat type 3 seems to be the most preferred—it has a positive difference between the observed (30) and expected (12.168) counts of moose and the largest contribution to the chi-square statistic (26.1325). Alternatively, habitat type 3 has the largest positive difference between the observed proportion of moose (0.256) and the expected proportion of moose (0.104).