Edhesive AP Statistics **Unit 9 Test – Solutions**

**Multiple Choice:** Choose the best answer choice for the following problems.

*Questions 1-3 apply to the following situation:*

Most people, when they touch their thumb and pinky fingers together, can observe a vestigial tendon from our evolutionary history. Known as the palmaris longus, it is present in around 86% of the population. A random sample of 200 people from Iceland were surveyed, as were 200 people from Australia. The results are shown in the table below:

|  |  |  |
| --- | --- | --- |
|  | Tendon present | Tendon absent |
| Iceland | 171 | 29 |
| Australia | 175 | 25 |

Let be the proportion of Icelanders with palmaris longus, and the proportion of Australians with palmaris longus.

1. The researchers are trying to determine if the proportion of people who retain palmaris longus differs between the two groups. What is the appropriate pair of hypotheses to test?

The null hypothesis for this test is that there is no difference between the populations, thus the proportion of people who present palmaris longus should be the same. The alternative is that there is a difference between these populations, ie that the proportions are not equal.

1. Which of the following would violate the conditions necessary to carry out the test described in question 1?
2. If or were too small to ensure reasonable accuracy for Normal calculations
3. If the total population of Australia or Iceland is less than 2,000
4. If the sample of Australians or the sample of Icelanders was not random
   1. I only
   2. II only
   3. I and II only
   4. II and III only
   5. I, II and III would each violate a necessary condition

Option I violates the requirement to have a large enough sample, II violates the small population proportion requirement, and III violates the random sample requirement.

1. The test from question 1 results in a z-value of . If the researchers wanted to reach a significance level of what is the correct conclusion to draw from this study?
   1. P-value = 0.044, we can reject the null hypothesis. Significant evidence exists to suggest a difference between the two populations.
   2. P-value = 0.044, we cannot reject the null hypothesis. Significant evidence does not exist to support a difference between the two populations.
   3. P-value = 0.09, we can reject the null hypothesis. Significant evidence exists to suggest a difference between the two populations.
   4. P-value = 0.09, we cannot reject the null hypothesis. Significant evidence does not exist to support a difference between the two populations.
   5. P-value = 0.4564, we cannot reject the null hypothesis. Significant evidence does not exist to support a difference between the two populations.

As this is a two-tailed test, a z-value of -1.71 gives a z-table value of 0.4564. Thus the P-value is 1-2\*0.4564=0.09. Statistical significance is not met at the level and we therefore do not have evidence that there is a statistically significant difference between the two populations.

1. The American Heart Association wants to assess the impact of marathon training on heart health. The researchers record the resting heart rate of 10 individuals before they begin their training program and measure them again at the end of the training program right before the marathon. Below are the results:

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Before training | 68 | 76 | 74 | 71 | 63 | 82 | 88 | 80 | 79 | 77 |
| After training | 65 | 73 | 67 | 62 | 55 | 72 | 72 | 69 | 65 | 67 |

Which test would be appropriate to determine if marathon training lowers a person’s resting heart rate?

* 1. Two sample t test with 9 degrees of freedom
  2. Two sample t test with 20 degrees of freedom
  3. Two proportion z test
  4. Paired t test with 9 degrees of freedom
  5. Paired t test with 10 degrees of freedom

We want a paired t-test because we are sampling the same population before/after some event to test for a change in some parameter. In this case if the resting heart rate lowers as a result of marathon training. Given a sample size of n=10, the degrees of freedom are df=n-1=9.

1. A new type of food for goldfish is said to increase their lifespan by 5%. The hypotheses to be tested are:

The fish food does not increase the lifespan of the fish

The fish food does increase the lifespan of the fish

If the study results in a Type II error, which of the following is true?

1. The study rejects the null hypothesis and in reality the food does increase lifespan
2. The study rejects the null hypothesis but in reality the food does not increase lifespan
3. The study fails to reject the null hypothesis but in reality the food does increase lifespan
4. The study fails to reject the null hypothesis and in reality the food does not increase lifespan
5. The study is inconclusive due to too small a sample size

A type II error is the incorrect acceptance of a false null hypothesis. In this context that means that the food does in fact increase lifespan, but the test concludes that it does not.

1. In an attempt to reduce traffic accidents, a city is going to install new LED speed limit signs to attract motorists’ attention. Since they are expensive, the city wants to install them where people speed the most. The police randomly sample vehicles at two 35 mph zones in town and record the following data:

|  |  |  |  |
| --- | --- | --- | --- |
|  | n | Mean (mph) | St. Dev (mph) |
| Location A | 56 | 42.2 | 4.86 |
| Location B | 72 | 44.1 | 5.83 |

Which of the following is the correct equation to compute the test statistic for the hypothesis that the mean speeds at the two locations are equal?

We use the test statistic for a two-sample unpooled t-test where the expected difference is 0 so that

.

1. The concentrations of ozone in Florida vary with a known standard deviation of 21.3 (micrograms per cubic meter) in Tampa, and 34.4 in Miami. A random sample of 25 air quality measurements in Tampa and 47 in Miami yield concentration means of 221.4 and 243.1 respectively. In order to determine the likelihood that the two locations have the same ozone concentration means, which combination of test and hypotheses is appropriate? Let the subscripts M, T signify Miami and Tampa.
2. Two-sample unpooled t test
3. Paired t test
4. Two-sample z test
   1. I and IV
   2. I and VI
   3. II and V
   4. II and IV
   5. III and IV

With known standard deviations, a z-test is appropriate (III). To test whether or not the two locations have different ozone concentrations, the null hypothesis is that the means are equal, the alternative is that they are not (IV)

*Questions 8-10 relate to the following scenario:*

The state of Florida decides to reduce their ozone concentrations and sets a goal threshold of 200. To do this, a number of air scrubbers are distributed around the state. A random sampling of air quality was taken at 200 locations beforehand and 127 tested above the 200 threshold. After the scrubbers had been in operation for six months, another 150 locations were tested and 82 tested above the threshold.

1. The researchers hope to show that the scrubbers are succeeding in pulling the ozone concentration below the desired threshold. What is the correct test statistic for the hypotheses

?

We use the test statistic for the two-proportion z test, pooled for .

where . For this study then

1. Suppose the results of the scrubber study give . If the researchers want a significance level of , what is the correct conclusion to draw from the study?
   1. P-value < 0.05; the evidence does not suggest that the scrubbers are reducing ozone concentrations to below the goal threshold.
   2. P-value = 0.0495; the evidence suggests that the scrubbers are reducing ozone concentrations to below the goal threshold.
   3. P-value > 0.05; the evidence does not suggest that the scrubbers are reducing ozone concentrations to below the goal threshold.
   4. P-value = 0.099; the evidence suggests that the scrubbers are reducing ozone concentrations to below the goal threshold.
   5. P-value = 0.901; the evidence suggests that the scrubbers are reducing ozone concentrations to below the goal threshold.

We use a 1-sided test to get P-value = 0.5-0.4505=0.0495. Thus we have reached significance and the level and can reject the null hypothesis and conclude that the scrubbers are effective in reducing the ozone concentrations below the goal threshold.

1. Given Florida’s success with the scrubbers, Georgia conducts a similar study. They conduct a two-proportion pooled z-test of the hypotheses , which results in a P-value of 0.13. Which of the following are true?
2. A 75% confidence interval will include the value 0.
3. A 85% confidence interval will not include the value 0.
4. A 95% confidence interval will not contain the value 0.
5. A 99% confidence interval will contain the value 0.
   1. II and IV only
   2. I and II only
   3. III and IV only
   4. I only
   5. None of the above

A P-value of 0.13 means we have reached significance at a level of 87%. Thus a 75% or 85% confidence interval would not include 0 (I is false, II is true). At 95% and 99%, the confidence interval would include 0 as we were not able to reject the null hypothesis at this level (III is false, IV is true).

**Free Response – Solutions**

1. The effects on an astronauts body from being in space are still being studied. One worrisome effect is the loss of bone mineral density (BMD) over time. While it is established that it happens, a study is being conducted to see if the rate of bone loss increases with the length of time spent in space. Over a number of years the National Aeronautics and Space Administration (NASA) records the decrease in bone mineral density for a number of astronauts who have spent either one or two months in space. Below are the results, where is the sample mean loss-rate in BMD over the given time period in grams per square centimeter per month.

|  |  |  |  |
| --- | --- | --- | --- |
|  | n |  | s |
| 1-month in space | 29 | 0.0129 | 0.0062 |
| 2-months in space | 17 | 0.0155 | 0.0107 |

Assume the necessary conditions have been met for this test.

* 1. Compute the test statistic for the data above using the appropriate test.

The appropriate test is the 2-sample pooled t-test with . The test statistic is where under the null hypothesis d=0. This gives t = 0.9158.

* 1. Compute the degrees of freedom for this test.

We use the formula to get

* 1. At what confidence level does this test suggest a change in the loss rate of BMD?

Using the t-table with df=22, t=0.9158 we find 0.25<P-value<0.10. Thus we have reached statistical significance at a confidence level of 75% but not 90%.

1. A supplier of sushi-grade fish conducts a survey of 1000 random people in order to better understand their market. The researchers asked the participants whether or not they lived within 250 miles of the ocean, and whether or not they liked sushi. Of those surveyed, 128 of the 210 who live near the coast say they like sushi while 393 of the 790 who live inland say they like sushi.

Construct a 95% confidence interval for the difference in the proportion of coastal and inland people who like sushi and offer an interpretation of the result.

We are interested in the difference in the proportion of people who like sushi and live near the coast to those who like sushi and live inland . Thus we are seeking a confidence interval for . We compute this using the pooled 2-sample z test (pooled since we are testing against the null hypothesis or ).

Therefore we compute and

The confidence interval is then given by which for a 95% confidence interval we use z=1.96. This results in a confidence interval of . The fact that 0 is not in this interval means that at a significance level of 95% we can reject the null hypothesis that the proportions are the same, and claim that people who live near the coast are more likely to like sushi. (In fact using this data we can compute a P-value of 0.0038, which confirms that we may reject the null hypothesis at the level ).

1. A construction company claiming to have developed stronger I-beams wishes to rebuild a bridge and raise the weight rating, allowing more and heavier traffic. To investigate the performance of the new versus the old beams, a random sample of each of the two types is stress tested. The hypotheses the construction company will test are vs where , are the old and new mean stress limits respectively.
2. Describe a Type I error in this context and explain how it may have dangerous consequences

A Type I error is the incorrect rejection of a true null hypothesis. In this context that means the true mean stress limit of the new beams is the same as the old beams but our study concludes that the stress limit is in fact higher.

This is potentially dangerous since raising the bridge’s weight rating on the false belief that the beams are stronger may over-stress the new bridge and put it in danger of collapse.

1. Concerned about the possibility of a Type I error, should the company conduct their test with a significance level of or ? Explain your reasoning.

A significance level gives the probability of obtaining the given results or more extreme results purely by chance under the assumption of the null hypothesis. Thus sets the maximum probability of a Type I error (for example if a test has and results in a P-value of 0.03, the null hypothesis is rejected with a 3% chance of a Type I error). Therefore setting a higher significance level will reduce the chances of committing a Type I error and the company should set .

1. A 90% confidence interval for the true difference in the stress limits between the new beams minus the old beams was found to be (0.45,1.28). Explain what this interval means in the context of this problem. Does it provide evidence that the new beams have significantly higher stress limits than the old beams?

A confidence interval is the range of values in which the true mean is likely to be. In this context it means that 90% of similarly conducted experiments would include the true mean. Since this interval represents the difference between the new and old stress limits and is purely positive, we can say at a confidence level of 90% that or that the stress limits in the new beams is significantly higher than the old beams.