Statistics For Behavioral And Social Sciences

Essay 1.

The experiment seeks to address the following research question; which vaccine between the shot spray and the nasal spray is more effective for the prevention of getting the flu?

H0= there is no significant difference in the effectiveness of the shot and nasal spray vaccines.

H1= there is a significant difference in the effectiveness of the shot and nasal spray vaccines.

The results of the test are statistically significant. This is because the P-value is less than the level of significance (0.008<0.05) and statistical significance is attained when the p-value is less than the level of significance and therefore, we reject the test's null hypothesis ("Statistical and practical significance - Minitab", 2016).

The researchers would reject the test's null hypothesis and conclude in the favor of alternative hypothesis. The results do provide sufficient evidence to support the alternative hypothesis. This is because the p-value is very small (less than 1%) and that is a suggestion that the test provides enough evidence to reject the null hypothesis and support the alternative hypothesis (Frost, 2016).

The sample size was appropriate for the study. Considering that the population is all people, it is considered irrelevant, a sample 1000 individuals is sufficient to administer the experiment because they are a representative of the population.

The experiment had some limitations. First, it was not taken into consideration that among the sample, there would have been some individuals who had taken a vaccine earlier before the study and that is a possible reason why they did not get the flu. This would result in the wrong conclusion being formed about which drug is better. The characteristics of individuals like health and age in the sample were not considered. A vaccine will work best in healthy adults and will fail in the sick and the elderly, and such factors were not considered.

In conducting the follow-up study, I would seek to determine which vaccine is more effective in children aged 4-12 years. I would take a sample of 1000 children, administer 500 with the shot spray and the other 500 with the nasal spray and collect the number of children who contract the flu after administration of each vaccine type. I would consider children because they are the most vulnerable to getting the flu, have not received the vaccine before, and therefore, the effectiveness of each vaccine will be accurately determined.

Statistical significance refers to the unlikelihood that mean differences observed in a sample have occurred due to sampling error (University of Guelph Numeracy project, 2008) meaning the results are unlikely due to chance alone while practical significance assesses whether the difference is large enough to be of value in a practical sense or a real world perspective (University of Guelph Numeracy project, 2008).

Essay 2.

The correlation is positive (+0.75) meaning that there is a strong relationship between IQ and GPA i.e. when the IQ is increasing, the GPA on average is also increasing. However, this does not imply that individuals with high IQ have high GPA. Correlation only shows the degree of relationship and does not imply a causation or functional relationship (Gupta & Gupta, 2010) therefore, a conclusion cannot be formed that individuals with high IQ have high GPA.

This correlation does not provide evidence that high IQ causes GPA to go higher meaning that there are other variables that influence the relationship. One of them is the amount of study hours and the higher the hours spent by an individual studying, the higher the GPA. The amount of academic guidance given to individuals affects the GPA in that the more academic guidance an individual is given, the GPA will be higher. Therefore, the conclusion that high IQ causes high GPA cannot be formed because there exist other factors apart from IQ which affects GPA.

Correlation describes the degree of relationship between two or more variables while causation indicates that one event is the result of the occurrence of the other effect and it describes a cause and effect relationship ("Statistical Language - Correlation and Causation", 2016). Correlation between two variables does not mean causation, for instance, smoking is correlated with alcoholism, but it does not cause alcoholism, but if there is a causal relationship between two variables, they must be correlated.

The following factors affect the size of this correlation:

i.   The amount of variability in X or Y. The value of correlation will be greater if there is more variability among the variables (Goodwin & Leech, 2006).

ii.   Lack of linearity. If the relationship between X and Y is not linear, the value of r is usually very low (Goodwin & Leech, 2006). However, this correlation is linear as evidenced by the strong positive correlation coefficient.

iii.   Characteristics of the sample like the level of study of participants and age. IQ and GPA usually decrease as the age increases and as one move to a higher level of education.

Correlation is not the best tool for predicting GPA. A researcher should instead use regression analysis as a test for predicting GPA. This is because regression analysis can help the researcher predict the GPA given the IQ and this will give the cause and effect relationship between the two variables.

Essay 3.

Task 1.

Group 1: 2.2, 2.5, 2.7, 2.9, 3.1, 3.5, 4.1, 4.3, 4.7, 4.8

	Group 1	Group 2
sum	34.8	92.4
mean	3.48	9.24
mode	none	9.5
median	3.3	8.85
standard deviation	0.893085	2.122357
range	2.6	7.9
skew	0.190223	2.443916
kurtosis	-1.5709	6.86883

Group 2: 7.3, 7.6, 8.1, 8.2, 8.5, 9.2, 9.3, 9.5, 9.5, 15.2

Differences in the two groups.

i.   Group 1 has a smaller standard deviation meaning that most of the data points are clustered around the mean (close to the mean) while group 2 has a higher standard deviation meaning that the data points are spread out over a wider range of values.

ii.   Group 2 is skewed more to the right, that is, it has a longer right tail compared to group 1 because it has a higher degree of skew (both are positively skewed but 2 has a longer right tail).

iii. Group 1 is platykurtic because kurtosis<3 meaning that its tail is shorter and thinner and its peak is lower and broader while group 2 is leptokurtic because kurtosis>3 meanings its tail is longer and broader, and its central peak higher and sharper.

Group 1 has no outlier. 15.2 is the outlier in group 2 because it lies beyond the upper limit of 11.4. Outliers affect the mean mostly because all the values in the sample are given the same weight in the calculation of the mean. As a result, the outlier increases the standard deviation because the

value is far away from the mean leading to the skewing of results. There is a little significant effect on the mode and median.

	Group 1	Group 2
sum	69.6	184.8
mean	3.48	9.24
mode	2.2	9.5
median	3.3	8.9
Standard deviatio	0.8930	2.12235
range	2.6	7.9
skew	0.17372	2.23189
kurtosis	-1.49733	4.72757

Yes, the sum, the degree of skew, and the degree of kurtosis changed in both groups.

i.   The increase in sample size increases the sum. For group 1, the sum doubled 34.8 to 69.6, and for group 2, it doubled from 92.4 to 184.8.

ii. The sample size reduces the degree of skew. For group 1, it reduces from 0.19 to 0.17 while for group 2, it reduces from 2.4 to 2.3.

iii. The sample size increases the degree of kurtosis for group 1, where it increases from -1.5 to -1.4 while for group 2, it reduces from 6.86 to 4.72.

References

Frost, J. (2016). *How to Correctly Interpret P Values*. *Blog.minitab.com*. Retrieved 7 September 2016, from http://blog.minitab.com/blog/adventures-in-statistics/how-to-correctly-interpret-p-values

Goodwin, L. & Leech, N. (2006). Understanding Correlation: Factors That Affect the Size of r. *The Journal Of Experimental Education*, *74*(3).

Gupta, S. & Gupta, M. (2010). *Business Statistics* (16th ed., p. 200). New Delhi: Sultan Chand & Sons.

Statistical and practical significance - Minitab. (2016). *Support.minitab.com*. Retrieved 7 September 2016, from http://support.minitab.com/en-us/minitab/17/topic-library/basic-statistics-and-graphs/introductory-concepts/p-value-and-significance-level/practical-significance/

Statistical Language - Correlation and Causation. (2016). *Abs.gov.au*. Retrieved 7 September 2016, from http://www.abs.gov.au/websitedbs/a3121120.nsf/home/statistical+language+-+correlation+and+causation

University of Guelph Numeracy project. (2008) (pp. 1-5). Retrieved from https://atrium.lib.uoguelph.ca/xmlui/bitstream/handle/10214/1869/A_Statistical_versus_Practical_Significance.pdf?sequence=7