*Module 2, Week 1, Paper and Pencil Assignment 3*

1. You have the following random sample containing information on master’s students in Data Analytics programs in the United States. The population standard deviation is unknown.

|  |  |  |
| --- | --- | --- |
| Student | Undergraduate GPA | International |
| A | 3.3 | No |
| B | 2.9 | No |
| C | 3.0 | No |
| D | 3.1 | Yes |
| E | 2.7 | Yes |
| F | 2.6 | No |
| G | 4.0 | No |
| H | 3.8 | Yes |
| I | 2.8 | Yes |
| J | 3.6 | No |

1. Estimate the mean GPA for master’s students in Data Analytics programs.
2. Calculate a 95% confidence interval for your estimate from a).
3. Calculate a 90% confidence interval for your estimate from a).
4. Which confidence interval is wider? Why?
5. If you drew a new random sample would you expect your confidence intervals from b) and c) to change? Why or why not?
6. With more observations, would you expect your confidence intervals to become wider or narrower?
7. What are the primary differences between a standard normal distribution and a Student’s t-distribution?
8. Estimate the proportion of international students in Data Analytics master’s programs.
9. Use a normal approximation to the binominal distribution to calculate a 95% confidence interval for your estimate from h).
10. Use a normal approximation to the binominal distribution to calculate a 99% confidence interval for your estimate from h).
11. How would you describe a confidence interval to someone with no background in statistics?

2. You have the following random sample containing information on customers at auto-repair shops. The population standard deviation is unknown.

|  |  |  |
| --- | --- | --- |
| Customer | $ spend on repair | Provided rental car |
| A | 90 | No |
| B | 290 | No |
| C | 560 | Yes |
| D | 430 | Yes |
| E | 900 | Yes |
| F | 770 | Yes |
| G | 130 | No |
| H | 850 | No |
| I | 600 | No |
| J | 120 | No |
| K | 390 | No |

1. Test the hypothesis that the mean spending on repairs is not equal to $500. Use a 5% level of significance.
2. Test the hypothesis that the mean spending on repairs is not equal to $600. Use a 10% level of significance.
3. Test the hypothesis that the mean spending on repairs is greater than $450. Use a 5% level of significance.
4. Test the hypothesis that the proportion of customers provided a rental is different from 0.30. Use a 5% level of significance.
5. Test the hypothesis that the proportion of customers provided a rental is greater than 0.25. Use a 10% level of significance.
6. How would you describe hypothesis testing to someone with no background in statistics?
7. Describe what a p-value represents.