1. A survey of 20 gas stations showed the price of a gallon of gas (rounding the 0.9 cents up) as

2.28 2.29 2.33 2.33 2.34 2.34 2.34 2.35 2.35 2.37

2.37 2.39 2.39 2.39 2.39 2.39 2.39 2.39 2.41 2.55

1. What is the mode?
2. What is the median?
3. What is the mean?
4. What is the standard deviation?
5. What is the variance?
6. I collected the above data from all the gas stations I pass on my way to NASA. They are all within a 5 mile radius as the crow files of each other and are all in Prince George’s county, MD. What factors do you think causes the variation in prices?

Note: Use EXCEL or a calculator with statistical functions. Do not calculate by hand. Be sure you use the stdev.s function (for a sample) in EXCEL or the σn-1 function on the calculator.

1. Suppose I change the 2.28 entry to 2.18.
2. Will the mode go down, stay the same, or go up? (choose or circle one)
3. Will the median go down, stay the same, or go up? (choose or circle one)
4. Will the mean go down, stay the same, or go up? (choose or circle one)
5. Will the standard deviation go down, stay the same, or go up? (choose or circle one)
6. Will the variance go down, stay the same, or go up? (choose or circle one)
7. Will the skewness of the data shift to the left, stay the same or shift to the right (choose or circle one)
8. () Make a bar graph of the gasoline prices. Let the x-axis be the price of gasoline. Let the y-axis be the number of stations selling the gasoline at that price. NOTE: I will be grading your graphs not only on getting the data numerically correct, but also on how well it is presented. Be sure you have a title for each graph. Clearly label the axes. It should be obvious to a casual reader what you are trying to present in your graph.
9. Make a pie chart of the prices in problem 1 with pie slices using intervals of

$2.20 – 2.29

$2.30 – 2.39

$2.40 – 2.49

$2.50 – 2.59

Use a plotting program, make a title and CLEARLY label what you are trying to show.

1. Fill in the frequency distribution table below.

|  |  |  |
| --- | --- | --- |
| Price Interval | Frequency | Cumulative frequency |
| 2.28 – 2.29 |  |  |
| 2.30 – 2.31 |  |  |
| 2.32 – 2.33 |  |  |
| 2.34 – 2.35 |  |  |
| 2.36 – 2.37 |  |  |
| 2.38 – 2.39 |  |  |
| 2.40 – 2.41 |  |  |
| 2.42 – 2.43 |  |  |
| 2.44 – 2.45 |  |  |
| 2.46 – 2.47 |  |  |
| 2.48 – 2.49 |  |  |
| 2.50 – 2.51 |  |  |
| 2.52 – 2.53 |  |  |
| 2.54 – 2.55 |  | 20 |
| Total | 20 | n/a |

1. Use the data in problem 5 to make two plots. Be sure that everything in the plot is CLEARLY labeled. Include a title.
2. What is the difference between a vertical bar graph and a histogram? Choose one of the following answers.
3. There is no difference.
4. Histograms have defined bin sizes on the horizontal axis while bar graphs take on the value of whatever the data happens to be on the horizontal axis.
5. Make a histogram plot of the frequency distribution.
6. () Make a line plot of the cumulative frequency distribution.
7. () This past spring, McDonalds had a Monopoly game where you had a ¼ chance per sticker of winning a small prize. If you bought a regular meal deal consisting of a sandwich, fries, and drink, you would get three stickers and thus 3 possibilities of winning a prize. Using the binomial theorem:

1 = (p + q)n

where p is the chance of winning a prize, q = the chance of not winning a prize, and n is the number of possibilities you have to win,

1. What is n when buying a regular meal deal?
2. Write out all of the terms of (p + q)n using your value of n.

i.e. 1 = pn + ???p(n-1)q + … + ???qn

1. What is the value of p?
2. What is the value of q?
3. Let k be the number of prizes you win. What is the probability that you will not win any prize? P(k = 0)
4. What is the probability that you win at least 1 prize? P(k >=1)
5. What is the probability that you win exactly 1 prize? P(k = 1)

Note: for parts (e)-(g), you can use the binomial calculator applet or calculate the answers by hand. If you use this applet, attach a screen shot which shows how you used the applet to get the answer.

[Binomial Calculator](http://onlinestatbook.com/2/calculators/binomial_dist.html) (<http://onlinestatbook.com/2/calculators/binomial_dist.html>)

1. A test medicine is given to a group of 500 patients. It is estimated that 2% of the patients will suffer from side effects of the medicine. Let us use the binomial distribution to analyze the data,

1 = (p + q)n , μ = np, σ =

1. What is n?
2. What is p?
3. What is q?
4. What is the mean number of patients that will suffer the side effects?
5. What is the standard deviation of the number of patients which will suffer the side effects?
6. What is the probability that exactly 5 patients (i.e. k = 5) will suffer the side effects? Use the probability formula from the summary sheet to do this calculation, i.e.

P(k) = nCk pkq(n-k) = pkq(n-k)  (0 <= k <= n)

1. Check you answer by using the binomial calculator applet in problem 8. Attach a screen shot to show how you got your answer. Note: use option “Between 5 and 5 inclusive”. You should get P = 0.0371.
2. (s) Use the Z-score normal distribution table when doing this problem.

[Z table - Normal Distribution](http://davidmlane.com/hyperstat/z_table.html) (<http://davidmlane.com/hyperstat/z_table.html>)

Most of you have taken the SAT test. Test scores range from 200 to 800. Statistically, the Mean of the test is scaled to be 500 while the standard deviation (SD) is scaled to be 100. NOTE: you do not have to attach screen shots for this problem.

1. Using the Z-table applet, put in the Mean and the SD, choose the option “Area from a value”, and “Below 500”. Click on “Recalculate”. The “Results Area” should be 0.5, meaning that a score of 500 is in the 50th percentile of the scores, i.e. the person did better than 50% of the people taking the test. Did you get this result of 0.5? (yes or no)
2. Now suppose a person get 600. What percentile is this score?
3. What percentile is a score of 700?
4. What percentile is a score of 800?
5. What is the percentage of students who score between 400 and 600? (Note: percentage is a fraction, while percentile is a ranking. Do not confuse the two terms.)
6. Now choose the option “Value from an area”. Find the SAT score you need to be in the upper 25% of those taking the test by letting Area = 0.25. In “Results” choose “Above” before hitting “Recalculate”. What score did you get?
7. What score does a student have to get to be in the upper 10%?
8. What score does a student have t get to be in the upper 1%?