## Math Logic Test

## Partial Answers to Practice Questions for Test 1

**1.** Categorize the following data according to level: nominal, ordinal, interval, or ratio.

(a) Time of first class.

**ANS:** Interval---differences in time are meaningful, but ratios are not. For example, a first class at 3:00pm is not 1.5 times later than a first class at 2:00pm.

(b) Length of time to complete an exam.

**ANS:** Ratio---differences in time make sense as do ratios. For example, if Student A takes 50 minutes and Student B takes 100 minutes, it makes sense to say that Student B took twice as long as Student A.

(c) Course evaluation scale: poor, acceptable, good.

**ANS:** Ordinal---the categories can be ranked, but differences between ranks do not make sense.

**2.** To estimate the average GPA of all La Sierra Students, President Geraty computed the average GPA obtained in his Advanced Hebrew Grammar class.

(a) Identify the variable?

**ANS:** GPA

(b) What is the implied population?

**ANS:** The GPA's of all La Sierra University Students.

(c) What is the sample?

**ANS:** The GPA's of all students in Dr. Geraty's Advanced Hebrew Grammar class.

(d) What type of sample was this?

**ANS:** This is a sample of convenience.

**3.** (a) If your instructor were to compute the class mean of this test when it is graded, and use it to estimate the average for all tests taken by this class this quarter, would this be an example of descriptive or inferential statistics? Explain.

**ANS:** Inferential – using a sample mean (one test) to estimate the population (all tests, quizzes and assignments) mean.

(b) A study on attitudes about smoking is conducted at a college. The students are divided by class, and then a random sample is selected from each class. What type of sampling technique is this (e.g. simple random, convenient, stratified, systematic, cluster)? Explain why this type of sample is not a simple random sample.

**ANS:** Stratified – a random sample is taken from each class (strata). This cannot be a simple random sample because it requires elements from each class. A simple random sample need not have representation from each class.

(c) A politician wishes to determine the reading level of 5th graders in her State. She does not have funding to test all 5th graders in her state, so she randomly selects some of the schools in her state and tests every 5th grader in those schools. What type of a sample is this?

**ANS:** Cluster—the population is divided into groups, some of the groups are selected randomly and everyone in those groups is tested.

**4.** Explain how you could use the table of random numbers in your text to help design a true false test of 10 questions so that the pattern of answers is random.

**ANS:** Randomly select a starting spot in the table. If the digit is odd, make a question with a false answer, if the digit is even make a question with a true answer. Proceed along the row for 10 such digits. For example, if the starting point had been the beginning of the 3rd row, the digits are: 59654 71966 which leads to answers of F F T F T F F F T T

**5.** How would you use a randomized two-treatment experiment in the following setting. Is placebo being used or not? Explain.

A veterinarian wants to test a strain of antibiotic on calves to determine their resistance to common infection. In a pasture are 22 newborn calves. There is enough vaccine for 10 calves. However, blood tests to determine resistance to infection can be done on all calves.

**ANS:** There is no placebo being used, as a ``fake" vaccine is not administered to the control group calves. To randomize, number the calves 1 to 22 and put those numbers in a hat, and select the calves corresponding to the first ten numbers chosen; alternatively, use a suitable random number generator to select the ten numbers.

**6.** Consider the data (which are systolic blood pressures of 50 subjects):

100,102,104,108,108,110,110,112,112,112,115,116,116,118,118,

118,118,120,120,126,126,126,128,128,128,130,130,130,130,130,

132,132,134,134,136,136,138,140,140,146,148,152,152,152,156,

160,190,200,208,208

(a) What class width should be chosen if you would like to have 6 classes.

**ANS:** (208 – 100)/6 = 18. Go to the next higher whole number to ensure that all of the data is covered. Thus a class width of 19 would be suitable.

(b) Suppose you don’t want a class width of 19, but would like a class width of 15 irrespective of how many classes that would give you. Complete the following table for this data.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| LowerLimit | UpperLimit | Lower Boundary | Upper Boundary | Midpoint | Frequency | CumulativeFrequency | Relative Frequency |
| 100 | 114 | 99.5 | 114.5 | 107 | 10 | 10 | .20 |
| 115 | 129 | 114.5 | 129.5 | 122 | 15 | 25 | .30 |
| 130 | 144 | 129.5 | 144.5 | 137 | 14 | 39 | .28 |
| 145 | 159 | 144.5 | 159.5 | 152 | 6 | 45 | .12 |
| 160 | 174 | 159.5 | 174.5 | 167 | 1 | 46 | .02 |
| 175 | 189 | 174.5 | 189.5 | 182 | 0 | 46 | .00 |
| 190 | 204 | 189.5 | 204.5 | 197 | 2 | 48 | .04 |
| 205 | 219 | 204.5 | 219.5 | 212 | 2 | 50 | .04 |

(c) Draw a frequency histogram using the table in (b).

(d) Draw a frequency polygon using the table in (b).

(e) Draw an Ogive using the table in (b).

**7.** For this question, consider the ogive that is displayed below for winning times for the Kentucky Derby.

(a) How many races had winning times under 2:05.15?

**ANS:** 75

(b) What percentages of the races had winning times between 2:01.15 and 2:03.15?

**ANS:** There were a total of 48-12 = 36 races with such winning times, hence 36 out of 101 totals gives us 35.64%.

(c) How many races had winning times of more than 2:03.15?

**ANS:** 101 – 48 = 53.

**8.** The following represent the high temperatures in Celsius in Minneapolis over a 12 day period in January: -10, -8,-7,-7,-7,-4,-3,-3,-2,0,1,5

Find the mean, median and mode of these temperatures. Note ∑ x = -45.

**ANSWERS:**

Mean = -45/12 = **-3.75** Median = (-4 + -3)/2 = **-3.5** Mode = **- 7**

**9.** (a) Find the median of the sample: 5,6,8,9,11,1400,1401,2000,4837. **ANS:** **11** .

(b) In what place (or average of which places) would the median be given a sample size of:

(i) 1000 (ii) 2125

**ANS:** The average of the 500th and 501st **ANS:** The (2125+1)/2 = 1063rd data in the ordered data in the ordered list. list.

(c) Given a set of data of size at least 3, how does increasing the largest number affect the median? How would it affect the mean?

**ANS:** The median would not change, the mean would increase.

**10.** (a) Make a stem and leaf display for the following data.

58 52 68 86 72 66 97 89 84 91 91

92 66 68 87 86 73 61 70 75 72 73

85 84 90 57 77 76 84 93 58 47

**ANS:** (not exactly a stem and leaf display, but a similar table**)**

|  |  |
| --- | --- |
| 456789 | 72 7 8 81 6 6 8 80 2 2 3 3 5 6 74 4 4 5 6 6 7 90 1 1 2 3 7 4|7 = 47 |

(b) After making the display, find the median of the data.

**ANS:** The median is the average of the 16th and 17th places which is 75.5

**11.** Consider the two populations as follows:

Population 1: -5,-2,3,5,7,10,15,17,19,23,25,26,28,30,33,37,43,48,50, 56

Population 2: -9,4,8,13,18,22,25,30,38,39,42,51,58,70,78,79,97,106,108,119

(a) Which population do you think has a larger standard deviation? Why?

**ANS:** Population 2 because the data appear to be more spread out.

(b)Given that Σ x = 468 and Σ (x-μ)2= 5896.8 for Population 1, compute its standard deviation.

**ANS:** σ=(5896.8÷20)1/2 = 17.171

(c) Given that Σ x = 996 and Σ x2= 77052 for Population 2, compute its standard deviation.

**ANS:** σ2 = (77052 – 9962/20)/20 = 27451.2 ÷ 20 = 1372.56, therefore the standard deviation is

 σ = 37.048

(d) Assuming Population 1 is merely a sample, compute the standard deviation:

**ANS:** s2 = (77052 – 9962/20)/29 = 27451.2 ÷ 19 = 1444.8, therefore the standard deviation is

 s = 38.011

(e) According to Chebyshev’s theorem, what proportion of the data in Population 2 should lie within 3 standard deviations of the mean. Compute the interval for Population 2 that is within 3 standard deviations of its mean.

At least (1 – 1/8), i.e. at least 8/9 of data should lie within 3 standard deviations of the mean.

**12.** Consider the following data of 26 numbers.

8 35 47 48 51 57 60 64 64 65

66 70 72 76 78 80 82 84 85 89

90 90 93 94 96 111

(a) Find the median of the data.

**ANS:** Because 26 is even, the median is the average of the 26/2 = 13th place and the 14th place, therefore the median is (72+76)/2 = 74

(b) (4 pts) Find Q1, Q3 and the IQR. Construct a box and whisker plot for the data.

**ANS:** The first quartile is the median of the 13 numbers below the median 74 of the entire set. Hence the first quartile is Q1=60 and the third quartile is the median of the 13 numbers above 74. Therefore the third quartile is Q3=89.

The IQR is the inter quartile range, IQR = Q3 - Q1 = 89 – 60 = 29.

See section 3.4 in the text for construction of the box and whisker plot. .

Note the lower whisker will go down to 33, the bottom of the box will start at 60, the line in the box will be at 74, the top of the box will be at 89, the upper whisker will go to 97.

(c) Compute the interval (Q1 - 1.5⋅IQR, Q3 + 1.5⋅IQR). Data outside of this interval are identified as suspected outliers. Are there any suspected outliers in the above data?

**ANS:** The interval is(16.6,132.5), so 8 is a suspected outlier.

**13.** The following are ground water depths for a sample of 20 wells.

|  |  |
| --- | --- |
| Distance from ground to water level (ft), x |  12—14 15—17 18—20 21—23 24—26  |
| Number of wells, f |  1 3 8 2 6  |

Using the midpoints of the intervals, estimate the mean depth, the standard deviation, and the coefficient of variation. (This is Problem 4, on p. 119).

**ANS:** For the mean, use the formula  so we compute:

∑xf = (13⋅1 + 16⋅3+19⋅8 + 22⋅2+25⋅6) = 407.

Thus the mean is approximately 407/20 = 20.35.

For the standard deviation, we use formula (7) on p. 114.

SSx= 132⋅1 + 162⋅3 +192⋅8 +222⋅2 +252⋅6 - 4072÷20 = 260.55

Hence, the standard deviation is approximately the square root of (260.55/19) ≈ 3.703.

The coefficient of variation is approximately (3.703/20.35)⋅100% = 18.2%.

**14.**  (a) If you are among 8000 people that took a test and you scored at the 79th percentile, approximately how many people scored at your score or lower? Approximately how many people scored at your score or higher?

**ANS:** Approximately .79⋅8000 = 6320 scores were less than or equal to your score, and approximately 1680 were greater than or equal to your score.

(b) If you are among 4000 students taking the MCAT, and you wish to score at least the 95th percentile, what is the maximum number of students that can score at least as well or better than you?

**ANS:** Not more than 5% of 4000, or 200 can score as well or better than you if you are to achieve the 95th percentile, or better.