Bronx Community College of The City University of New York Department of Mathematics and Computer Science

Math 23 Review Sheet

- 1. Classify each of the following data according to the *level of measurement* (that is state whether it is nominal, ordinal, interval, or ratio):
 - (a) The telephone numbers in a telephone directory.
 - (b) The scores of a class in an exam.
 - (c) Absolute temperatures (that is temperatures measured in Kelvin degrees).
 - (d) Motion Picture Association of America ratings description (G, PG, PG-13, R, NC-17).
 - (e) Average monthly precipitation in inches for New York, NY.
 - (f) Average monthly temperature (in degrees Fahrenheit) for New York, NY.
- 2. A group of 25 people were observed regarding their TV habits and were found to spend the following number of hours per week watching television:

30	32	34	36	36
37	39	39	41	41
42	42	43	43	44
45	45	45	46	47
47	49	49	52	53

In order to display the data in clearer form,

- (a) determine the class width for four (4) classes,
- (b) construct a frequency distribution showing the class limits for the four classes,
- (c) in the table, show the class boundaries and the class marks,
- (d) construct a histogram, labeling the class boundaries.
- 3. A consumer testing service obtained the following mileage (in miles per gallon) in five test runs for three different types of compact cars:

	\mathbf{First}	Second	Third	Fourth	\mathbf{Fifth}
	\mathbf{Run}	Run	\mathbf{Run}	Run	Run
Car A	28	32	28	34	30
Car B	31	31	29	29	31
Car C	32	29	28	32	30

- (a) If the manufacturer of Car A wants to advertise that their car performed the best in this test, which measure of central tendency (mean, median or mode) should be used to support their claim?
- (b) Which measure should the manufacturer of Car B use to claim that their car performed best, mean median or mode?
- (c) Which measure should the manufacturer of Car C use to support a similar claim?
- 4. Florida's age distribution has mean value $\mu = 39.2$ and standard deviation $\sigma = 24.8$ (measured in years). Use Chebyshev's theorem to find an interval such that

- (a) the age in years of at least 75% of Florida's population is contained within that interval,
- (b) the age in years of at least 88.9% of Florida's population is contained within that interval,
- (c) the age in years of at least 93.8% of Florida's population is contained within that interval.
- 5. Calculate the range, mean, median, first and third quartiles, interquartile range, mode, variance, and standard deviation for the following population data.

 $47 \quad 59 \quad 50 \quad 56 \quad 56 \quad 51 \quad 53 \quad 57 \quad 52 \quad 49$

6. Find the mean, the range, and the standard deviation for the following set of sample data.

 $10 \quad 9 \quad 12 \quad 11 \quad 8 \quad 15 \quad 9 \quad 7 \quad 8 \quad 6$

7. Determine the range and the sample standard deviation of the following data:

x	f
10.3	7
22	12
38.5	5
43.2	2

- 8. The mean value of the scores in a Statistics exam was 85 with a standard deviation of 4. Find an interval that contains at least 75% of the scores in that exam.
- 9. The manager of a salmon cannery suspects that the demand for her product is closely related to the disposable income of her target region. To test out this hypothesis she collected the following data for five different target regions, where x represents the annual disposable income for a region in millions of dollars and y represents sales volume in thousands of cases.



- (c) Find and graph the least square line.
- (d) If a region has disposable annual income \$25,000,000 what is the predicted sales volume?

10. Match the appropriate statement about r and the scatter diagrams.



A. -1 < r < 0 B. r = 0. C. r = -1. D. 0 < r < 1

- 11. How many three-letter words can be formed of 21 consonants if
 - (a) Repetitions are allowed?
 - (b) Repetitions are not allowed?

12. Given
$$P(E^C) = 0.3$$
, $P(F) = 0.35$, and $P(F|E) = 0.25$ find

- (a) P(E and F)
- (b) P(E or F)
- (c) P(E|F).
- 13. Two dice are rolled. Find the probability of the following events:
 - (a) Both numbers are 6.
 - (b) The first dice gives 5 and the second 6.
 - (c) There is one 5 and one 6.
 - (d) The sum is equal to 10.
 - (e) Both are 6 or the sum 10.
 - (f) The sum is more than 5 but less than 8.
 - (g) Both numbers are even.
 - (h) One number is even and one number is odd.
- 14. Calculate by hand (without a calculator). Show all work:
 - (a) 5!
 - (b) C(12,3).
- 15. An urn contains three yellow, four green, and five blue balls. Two balls are randomly drawn without replacement. Find the probability of the following events:
 - (a) Both balls are blue.
 - (b) The first ball is green and the second yellow.
 - (c) There is one green and one yellow ball.
- 16. Repeat the previous exercise but now assume that the balls are drawn with replacement.
- 17. Three cards are randomly drawn from a standard 52 card deck without replacement. Find the probability of the following events:
 - (a) All cards are red.
 - (b) There are two red and one black card.

- (c) All cards are spades.
- (d) There is one spade, one club, and one diamond.
- (e) All cards are aces.
- (f) Two cards are aces and one card is a king.
- 18. Most of the time, a medical test is able to correctly indicate if a person has a condition. However, some of the time, there are false positives (it indicates the condition is present when it is not) or false negatives (it indicates the condition is not present when it is there). Use the table below to determine the probabilities for a randomly selected person from the population.

	condition present	condition not present	row total
Test Result $+$	125	10	135
Test Result $-$	15	50	65
column total	140	60	200

- (a) What is the probability of a false positive?
- (b) What is the probability of either a false positive or a false negative?
- (c) What is the probability of a positive test result given that the condition is present?
- (d) What is the probability that the condition is present given a positive test result?
- (e) What is the probability of either a negative test result or the condition is not present?
- 19. One college found that during one semester 1,259 students in its four most popular majors had the following class distributions. Use the table below to determine the probabilities for a randomly selected student in this group.

	first year	sophomore	junior	senior	row total
Business	115	90	105	111	421
Psycology	88	95	91	96	370
Nursing	85	81	79	76	321
Biology	63	45	25	14	147
column total	351	311	300	297	1259

- (a) What is the probability of being a business major?
- (b) What is the probability of not being a biology major?
- (c) What is the probability of being a senior and majoring in psychology?
- (d) What is the probability of being a senior or sophomore?
- (e) What is the probability of being a senior or biology major?
- (f) What is the probability of being a junior, given being a nursing major?
- (g) What is the probability of being a nursing major, given being a junior?
- 20. An island is a habitat for 208 species of birds. 82 of these species are found only on this particular island. 75 species are seabirds. 12 are a species of seabird and are found only on this particular island. One species of bird is chosen at random.
 - (a) What is the probability it is a seabird or unique to this island?
 - (b) What is the probability it is neither a seabird nor unique to this island?
- 21. (a) A company is looking hire more sales staff. The human resources department accepts only the 45% of the submitted resumes that meet the hiring criteria. The managers then select 20% of the applicants with accepted resumes to come in for an interview. What is the probability that an applicant selected at random will have her resume accepted and be granted an interview?

- (b) In one high school, the athletic director found that 4% of the varsity athletes had concussions while playing at the school and 18% had severe sprains and 1% had experienced both. What is the probability that a randomly selected varsity athlete has either had a concussion or a severe sprain?
- 22. Consider the following discrete probability distribution:

x	2	3	4	5	6
P(x)	.25	.1	.3	.2	.15

Sketch the graph of this distribution and calculate its expected value and standard deviation.

23. Find the expected value and the standard deviation of the probability distribution whose graph is shown:



- 24. A fair coin is tossed 7 times. Sketch the graph of the resulting binomial distribution.
- 25. Alice and Bob play the following game: two cards are randomly drawn (with replacement) from a standard 52-card deck, if they are both red Alice wins otherwise Bob wins. If they play these game 16 times what is the probability that Alice will win at most 4 times?
- 26. If 30% of the people in a community use the Library in one year, find the probability that in a random sample of 15 people
 - (a) At most 7 use the Library,
 - (b) Exactly 7 use the Library,
 - (c) At least 5 use the Library,
 - (d) No more than 2 use the Library,
 - (e) Not less than 10 use the Library.
- 27. A basket ball player makes 70% of the free throws he shoots. What is the probability that he will make more than 7 throws
 - (a) If he tries 15 free throws?
 - (b) If he tries 10 free throws?
- 28. Approximately 5% of the eggs in a store are cracked. Suppose you buy a dozen eggs from the store.
 - (a) What is the probability that no more than one of your eggs is cracked?

- (b) What is the probability that fewer than 3 eggs are cracked?
- (c) Find the expected value and standard deviation of the number of cracked eggs.
- 29. A surgery has a success rate of 75%. Suppose that the surgery is performed on six patients. Find the expected value and the standard deviation of the number of successes.
- 30. One-third of all deaths are caused by heart attacks. If three deaths are chosen randomly, find the probability that none resulted from heart attack.
- 31. Let z have the standard normal distribution. For each of the following probabilities, draw an appropriate diagram, shade the appropriate region and then determine the value:
 - (a) P(0 < z < 1.74)
 - (b) P(0.62 < z < 2.48)
 - (c) P(z > 2.1)
 - (d) P(-1.31 < z < 1.07).
- 32. Let z have the standard normal distribution. For each of the following probabilities, draw an appropriate diagram, shade the appropriate region and then determine the value of z_c :
 - (a) $P(0 < z < z_c) = 0.4573$
 - (b) $P(z_c < z < 0) = 0.3790$
 - (c) $P(z < z_c) = 0.1190$
 - (d) $P(-z_c < z < z_c) = 0.8030.$
- 33. Let x be a normally distributed random variable with $\mu = 70$ and $\sigma = 8$. For each of the following probabilities, draw an appropriate diagram, shade the appropriate region and then determine the value:
 - (a) P(70 < x < 80.4)
 - (b) P(61.2 < x < 85.2)
 - (c) P(x < 58)
 - (d) P(x > 76).
 - (e) $P(68 < \bar{x} < 72)$, if a random sample of size n = 49 is drawn.
 - (f) $P(\bar{x} > 71)$, if a random sample of size n = 81 is drawn.
- 34. Find z so that:
 - (a) 98% of the area under the standard normal curve lies between -z and z.
 - (b) 97.5% of the area under the standard normal curve lies to the left of z.
 - (c) 46% of the area under the standard normal curve lies to the right of z.
- 35. Find the area under the standard normal curve
 - (a) between z = -2.74 and z = 2.33.
 - (b) between z = -2.47 and z = 1.03.
- 36. The lifetime of a certain type TV tube has a normal distribution with a mean of 80.0 and a standard deviation of 6.0 months. What portion of the tubes lasts between 62.0 and 95.0 months?
- 37. The scores in a standardized test are normally distributed with $\mu = 100$ and $\sigma = 15$.
 - (a) Find the percentage of scores that will fall below 112.
 - (b) A random sample of 10 tests is taken. What is the probability that their mean score \bar{x} is below 112?

- 38. The weights (in pounds) of metal discarded in one week by households are normally distributed with a mean of 2.22 lb. and a standard deviation of 1.09 lb.
 - (a) If one household is randomly selected, find the probability that it discards more than 2.00 lb. of metal in a week.
 - (b) Find a weight p_{30} so that the weight of metal discarded by 70% of the houses is above x.
- 39. If the salary of computer technicians in the United States is normally distributed with the mean of 32,550 and the standard deviation of 2,000, find the probability for a randomly selected technician to earn
 - (a) More than 35,000.
 - (b) Between 31,500 and 35,000.
 - (c) What is the probability that the mean salary of a random sample of 4 technicians is more than 35,000?
- 40. The lifetime of a AAA battery is normally distributed with mean $\mu = 28.5$ hours and standard deviation $\sigma = 5.3$ hours.
 - (a) For a battery selected at random, what is the probability that the lifetime will be more than 30 hours.
 - (b) For a sample of three batteries, what is the chance that all three last more than 30 hours?
 - (c) For a sample of three batteries, what is the probability that their mean lifetime \bar{x} is more than 30 hours?
 - (d) What is the probability that the mean lifetime \bar{x} of batteries from a package of 12 will be less than 27 hours?
- 41. In Jennifer's Fall 2014 history class, 14 of 34 students passed the class. If you assume a professor's passing rates are constant, would it be appropriate to use a normal curve approximation to the binomial distribution to estimate the mean passing rate for the same professor's Spring 2015 semester class of 28 students? Explain your answer.
- 42. According to the Vision Council of America, 75 percent of the U.S. adult population wears some form of glasses to correct their vision. In a random sample of 950 adults, what is the probability that fewer than 700 people wear glasses?
- 43. An environmental group did a study of recycling habits in a California community. It found that 70 percent of aluminum cans sold in the area were recycled. If 400 cans are sold in one day, what is the probability that between 260 and 300 will be recycled?
- 44. The weekly amount a family spends on groceries follows (approximately) a normal distribution with mean $\mu = 200 and a standard deviation $\sigma = 15 .
 - (a) If \$220 is budgeted for next week's groceries what is the probability that the actual cost will exceed the budget?
 - (b) How much should be budgeted for weekly grocery shopping so that the probability that the budgeted amount will be exceeded is only 0.05?
- 45. A study is being planned to estimate the mean number of semester hours taken by students at a college. The population standard deviation is assumed to be $\sigma = 4.7$ hours. How many students should be included in the sample to be 99% confident that the sample mean \bar{x} is within one semester hour of the population mean μ for all students at this college?
- 46. To determine the mileage of a new model automobile, a random sample of 36 cars was tested. A sample with a mean of 32.6 mpg and a standard deviation of 1.6 mpg was obtained. Construct the 90% confidence interval for the actual mean mpg of the population of this model automobile.

- 47. A random sample of 12 employees was taken and the number of days each was absent for sickness was recorded (during a one-year period). If the sample had a mean \bar{x} of 5.03 days and standard deviation s of 3.48 days, create a 95% confidence interval for the population mean days absent for sickness, assuming the distribution of absences is normal.
- 48. Computer Depot is a large store that sells and repairs computers. A random sample of 110 computer repair jobs took technicians an average of $\bar{x} = 93.2$ minutes per computer. Assume that σ is known to be 16.9 minutes. Find a 99% confidence interval for the population mean time μ for computer repairs.
- 49. The following data represent a sample of the number of home fires started by candles. Assuming that the number of home fires started by candles is approximately normally distributed find a 95% confidence interval for mean number of home fires started by candles each year.
 - 5400 5860 6070 6210 7360 8450 9960
- 50. Leonor decides to run for political office. In order for her name to appear on the ballot, she must collect 7,500 valid signatures from registered voters. After she collects 10,000 signatures, she decides to check what proportion of the ones she collected are valid. She takes a random sample of 150 of the signatures she collected and brings them to the Board of Elections to verify them. It turns out that of the sample of 150, only 87 are valid. Construct a 95 percent confidence interval for the proportion of valid signatures she has collected.
- 51. In a Gallup poll, 1025 randomly selected adults were surveyed. 400 of them said that they shopped on the internet at least a few times per year. Construct a 99 percent confidence interval to estimate the percentage of all adults who shop on the internet several times per year.
- 52. A random sample of 41 NBA players gave a standard deviation s = 3.32 inches for their height. How many more NBA players have to be included in the sample to make 95% sure that the sample mean \bar{x} of their height is within 0.75 inch of the mean μ of the height of the population of all NBA players.
- 53. Gregor Mendel was a pioneer in the theory of genetics. His idea was to assign probabilities to significant population traits of plants or animals, like eye color, based on "dominant" or "recessive" traits. For example, he studied peas with green pods (a dominant trait) or yellow pods (a recessive trait). He predicted that the probability that a hybrid ("offsping") of a green pea with a yellow pea will have a yellow pod is p = 0.25.

Mendel conducted an experiment of green-yellow hybrids. In one experiment, 428 offspring had green pods and 152 offspring had yellow pods.

Use a level of significance of $\alpha = 0.01$ to test the claim that Mendel's claim that p = 0.25 is wrong.

- 54. A teacher has developed a new technique for teaching which he wishes to check by statistical methods. If the mean of a class test turns out to be 60 (or less), the results will be considered unsuccessful. Alternatively, if the mean is greater than 60, the results will be considered successful. The results of the test with a class of 36 students had a mean $\bar{x} = 66.2$ with a standard deviation of s = 24.0. Test whether the results were successful at the $\alpha = 5\%$ level of significance. (Use 1-tail test.) State the null and the alternate hypothesis and include diagrams.
- 55. The average annual salary of employees at a retail store was 28,750 last year. This year the company opened another store. Suppose a random sample of 18 employees had an average annual salary of $\bar{x} = \$25,\10 with sample standard deviation of s = \$4230. Use a level of significance $\alpha = 1\%$ to test the claim that the average annual salary for all employees is different from last year average salary. Assume salaries are normally distributed.
- 56. A machine in the lodge at a ski resort dispenses a hot chocolate drink. The average cup of hot chocolate is supposed to contain $\mu = 7.75$ ounces. We may assume that x has a normal distribution with $\sigma = 0.3$ ounces. A random sample of 16 cups of hot chocolate from this machine had a mean content of $\bar{x} = 7.62$ ounces. Use a $\alpha = 0.05$ level of significance and test whether the mean amount of liquid is different than 7.75 ounces.

Answers

- 1. A. Nominal. B. Ratio. C. Ratio. D. Ordinal. E. Ratio. F. Interval.
- 2. The class width has to be 6. We then have the following frequency table.

Class Limits	Class Boundaries		Class Marks
Lower-Upper	Lower-Upper	Frequency	(midpoints)
30 - 35	29.5 - 35.5	3	32.5
36 - 41	36.5 - 41.5	7	38.5
42 - 47	41.5 - 47.5	11	44.5
48 - 53	48.5 - 53.5	4	50.5

And we have the following histogram: Figure ??



Figure 1: The histogram of problem 2

- 3. A. Mean. B. Median. C. Mode.
- 4. A. [0, 88] B. [0, 113.6] C. [0, 138.4].
- 5. The range is 12, the mode is 56, the mean is $\mu = 53$, the standard variation is $\sigma = 3.69$, the variance is $\sigma^2 = 13.6$. The quartiles are $Q_1 = 50$, the median $Q_2 = 52.5$, and $Q_3 = 56$ while the interquartile range is 6.
- 6. Mean is $\bar{x} = 9.5$, range is 9, sample standard deviation is s = 2.64.
- 7. The range is 32.9. The standard deviation is s = 11.22.
- 8. [77, 93]
- 9. The correlation coefficient is r = 0.9. The line of least squares is $\hat{y} = 0.3 + 0.09x$. For a region with disposable annual income of \$25,000,000 the model predicts sale of 2,550 cases. The scatter graph and the plot of the line are shown in Figure ??.
- 10. A. (d) B. (b) C. (a) D. (c).
- 11. A. 9261 B. 7980.
- 12. A. 0.175 B. 0.875 C. 0.5.

13. A. $\frac{1}{36}$ B. $\frac{1}{36}$ C. $\frac{1}{18}$ D. $\frac{1}{12}$ E. $\frac{1}{9}$ F. $\frac{11}{36}$ G. $\frac{1}{4}$ H. $\frac{1}{2}$. 14. A. 120 B. 220. 15. A. $\frac{5}{33}$ B. $\frac{1}{11}$ C. $\frac{2}{11}$.



Figure 2: The scatter plot and the regression line of problem 9

16. A.
$$\frac{25}{144}$$
 B. $\frac{1}{12}$ C. $\frac{1}{6}$.
17. A. $\frac{2}{17}$ B. $\frac{13}{34}$ C. $\frac{11}{850}$ D. $\frac{169}{1700}$ E. $\frac{1}{5225}$ F. $\frac{6}{5525}$.
18. A. $\frac{1}{20}$ B. $\frac{1}{8}$ C. $\frac{25}{28}$ D. $\frac{25}{27}$ E. $\frac{3}{8}$.
19. A. $\frac{421}{1259}$ B. $\frac{1112}{1259}$ C. $\frac{96}{1259}$ D. $\frac{608}{1259}$ E. $\frac{430}{1259}$ F. $\frac{79}{321}$ G. $\frac{79}{300}$.
20. A. $\frac{145}{208}$ B. $\frac{63}{208}$.
21. A. 0.09 B. 0.21.

22. The expected value of the distribution is $\mu = 3.9$ and the standard deviation is $\sigma = 1.37$. The graph of the distribution is



Figure 3: The graph of the probability distribution of problem 23

23. The expected value is $\mu = 3.05$ and the standard deviation $\sigma = 1.58$.

24. First compute the probabilities (you can also get these values from the tables in the appendix of the textbook):

x	0	1	2	3	4	5	6	7
P(x)	.008	.055	.164	.273	.273	.164	.055	.008

The graph is:



Figure 4: The graph of the binomial distribution of problem 24

- 25. $P(0 \le r \le 4) \approx 0.63.$
- 26. A. $P(0 \le r \le 7) = 0.951$ B. P(r = 7) = .081 C. $P(5 \le r \le 15) = 0.485$ D. $P(0 \le r \le 2) = 0.128$ E. $P(10 \le r \le 15) = 0.004$.
- 27. A. $P(7 < r \le 15) = 0.951$. Why is this answer the same as the answer for 26 (a)? B. $P(7 < r \le 10) = 0.382$.
- 28. A. $P(0 \le r \le 1) = 0.881$ B. $P(0 \le r < 3) = 0.98$ C. The expected number $\mu = 0.6$ and the standard deviation $\sigma = 0.755$.
- 29. The expected number $\mu = 4.5$ and the standard deviation $\sigma = 1.061$.





37. A. 78.81%B. 0.9943.38. A. P(x > 2.00) = 0.58B. 1.65 lb.39. A. 0.1093B. 0.5926C. 0.0071.

- 40. A. 0.3897 B. 0.0592 C. 0.3121 D. 0.1635.
- 41. A binomial distribution can be approximated by a normal distribution if both np > 5 and nq > 5. In Fall 2014 the passing rate was p = 0.41 with np = 14 > 5 and nq = 20 > 5 so it would be appropriate to assume a normal distribution for the next semester as well. As in Spring 2015 n = 28, using the normal distribution approximation would be assuming that $0.18 \le p \le 0.82$; p = 0.41 is within this range.
- $42. \ 0.1660$
- $43. \ 0.9750$
- 44. A. 0.0918 B. \$224.67.
- 45. 148.
- 46. [32.16, 33.04].
- 47. [2.82, 7.24].
- 48. [89.04, 97.36].
- 49. [5518.54, 8570.04].
- 50. 0.50
- 51. 0.35
- 52. 35 more players need to be included.
- 53. Partial solution: $H_0: p = 0.25, H_a: p \neq 0.25, n = 428 + 152$ so $\hat{p} = 0.26$. The sample test statistic is

$$z = \frac{0.26 - 0.25}{\sqrt{\frac{(0.25)(0.75)}{580}}} = \frac{0.01}{0.01798} = 0.56$$

$$P$$
-value = $2 \cdot P(z \le -0.56) = 0.5754 > 0.01 = \alpha$

Conclusion: Do not reject H_0 . The results were statistically significant at the 1% level of significance. Based on the sample data, we think that probability that a pea hybrid will have a yellow pod is 0.25.

54. Partial solution: $H_0: \mu = 60$ (or $\mu \le 60$), $H_a: \mu > 60$. The critical z-value is $z_c = 1.645$. Then

$$z = \frac{66.2 - 60.0}{\frac{24.0}{\sqrt{36}}} = \frac{6.2}{4.0} = 1.55 < z_c$$

Conclusion: Do not reject H_0 . The results were statistically unsuccessful at the 5% level of significance. (That is the results could not be distinguished from a random sample from a normal population with mean $\mu = 60$ and standard deviation $\sigma = 24.0$.) 55. Partial solution: $H_0: \mu = 28,750, H_a: \mu \neq 28,750$. The test statistic is

$$t = \frac{25810 - 28750}{\frac{4230}{\sqrt{18}}} = -\frac{2940}{997.02} = -2.949.$$

For d.f. = 17, the test statistic t = -2.949 is in the interval

$$2.898 < |t| < 3.965.$$

Thus a 2-tail test shows

$$0.010 > P$$
-value > 0.001 .

Conclusion: Reject H_0 because the *P*-value $< \alpha = 0.01$. At the 1% level of significance, the evidence is sufficient to reject H_0 . Based on the sample data, we think that the mean annual salary is different from that of the previous year.

56. Partial solution: $H_0: \mu = 7.75, H_a: \mu \neq 7.75$. Critical value $\pm z_c = \pm 1.96$. Then

$$z = \frac{7.62 - 7.75}{\frac{0.3}{\sqrt{16}}} = -1.73 < -z_0.$$

Conclusion: Reject H_0 .

August 2006, NEA, revised November 2014 & October 2016 by the MTH 23 Committee. To report errors email evangelia.antonakos@bcc.cuny.edu.