

Math 221

Hypothetical Exam 1, Wi2008, (Chapter 1-5 in Moore, 4th)

April 3, 2063

S. K. Hyde, S. Barton, P. Hurst, K. Yan

Name: _____

Show all your work to receive credit. All answers must be justified to get full credit.



1234567

These questions are intended to give students in Math 221 some idea of the types of questions which could be asked on an exam. They may not cover all of the topics which will be on your exam (and they may cover more topics than are on your exam). The length of your exam may be shorter than this practice exam. Working these problems is not a substitute for studying your notes and reading the book.

Multiple Choice

Circle the letter corresponding to the best answer for each of the problems below (4 pts each)

1. There are three children in a room—ages 3, 4, and 5. If a four-year-old child enters the room, the
- A. mean age will stay the same but the standard deviation will increase
 - B. mean age will stay the same but the standard deviation will decrease
 - C. mean age and standard deviation will stay the same
 - D. mean age and standard deviation will increase.

For the following description, answer questions 2 to 5.

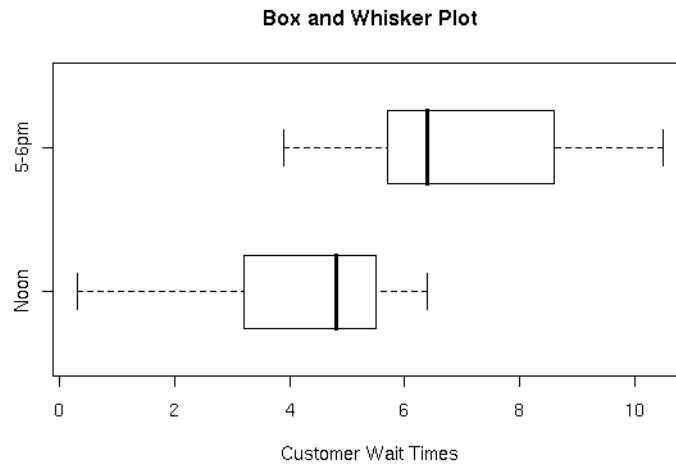
During the early part of the 1994 baseball season, many sports fans and baseball players noticed that the number of home runs being hit seemed to be unusually large. Below are the team-by-team statistics on home runs hit through Friday, June 3, 1994 (from the Columbus Dispatch Sports Section, Sunday, June 5, 1994). They are given as separate stemplots for the number of home runs by American and National League teams.

American League		National League	
2		2	9
3	5	3	1
4	0 3 9	4	2 6 7 8 8
5	1 4 7 8 8	5	3 5 5 5
6	4 8 8	6	3 3 7
7	5 7	7	

2. The median for the number of home runs hit through Friday, June 3, 1994 for the American League teams is
- A. 46
 - B. 50
 - C. 57.5
 - D. 68
 - E. Lower than the National League
3. The mean of the number of home runs hit through Friday, June 3, 1994 by National League teams
- A. 48.62
 - B. 50.14
 - C. 51.50
 - D. 48
 - E. Cannot be determine from the given information
4. The five-number summary of the number of home runs hit through Friday, June 3, 1994 by American League teams is
- A. 35, 49, 57.5, 68, 77
 - B. 35, 50, 57.5, 68, 77
 - C. 35, 50, 57.5, 66, 77
 - D. 35, 49, 57.5, 66, 77
 - E. Cannot be determined from the given information
5. Which of the following is a correct statement?
- A. The American League plot is reasonably symmetric.
 - B. The National League plot is slightly skewed to the left.
 - C. The median number of home runs hit by American League teams was higher than by National League teams.
 - D. All of the above.
 - E. None of the above.

For the following description, answer questions 6 to 8.

A local bank manager wanted to examine the relationship between customer wait times and the time of day they used the bank. The top boxplot displays the customer wait times in the evening hour from 5-6 pm. The bottom boxplot displays wait times for the noon hour.



6. Which of the following is correct?
 - A. The distribution of wait times for the evening hour (5-6pm) is roughly symmetric.
 - B. The distribution of wait times for the noon hour is skewed right.
 - C. The distribution of wait times for the evening hour (5-6pm) is skewed left.
 - D. The distribution of wait times for the noon hour is skewed left.
7. Which of the following can you conclude?
 - A. The mean of the noon hour wait times is probably lower than the median wait time.
 - B. The median of the noon hour wait times is probably lower than the mean wait time.
 - C. The mean and the median for the evening wait times should be approximately equal.
 - D. Because of the shape of the distribution of noon hour wait times, the mean and median should be the same.
8. Which of the following is NOT true?
 - A. The middle 50% of noon hour wait times were less spread out than the middle 50% of the evening wait times.
 - B. More than 75% of the noon hour wait times were less than six minutes in length.
 - C. More than half of the evening wait times were longer than six minutes in length.
 - D. Approximately half of the noon hour wait times were less than three minutes in length.
9. Performance ratings of stocks are approximately Normal with mean 100 and standard deviation 20. “Uber” stocks are classified as those stocks that have a performance rating of 140 OR MORE. What percentage of stocks are “Uber” stocks?

A. 2.5%
B. 5%
C. 90%
D. 95%

10. A particular stock has a performance rating of 124. Based on the distribution information in problem 9, what is the z -score for this stock?
- A. 18 B. 7.87 C. 1.2 D. 12.4
11. The scores of adults on an IQ test are approximately Normal with mean 100 and standard deviation 15. Susan scores 127 on such a test. She scores higher than what percent of all adults?
- A. about 4% B. about 88% C. about 92% D. about 96%
12. Assume the average number of minutes per day women cell-phone owners spend talking on their cell phones approximately follows the $N(75, 30)$ distribution. Approximately what percent of such women talk on their cell phones an average of BETWEEN 60 minutes and 100 minutes per day?
- A. 31% B. 95% C. 48% D. 64%
13. A local park district is planning to build a recreation center. The park district conducted a poll to find out the types of physical activities the local population would be interested in. The poll was based on telephone responses from 1013 randomly selected adults. The table shows the percentages of people who expressed interest in various activities.

Activity	Percent
Swimming	77
Running/Walking	46
Weight Training	34
Biking	30
Aerobics	18

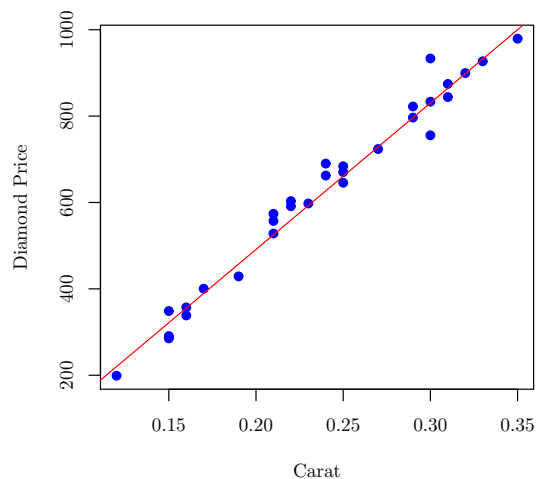
What percentage of people expressed that they are interested in biking or swimming?

- A. 77 percent D. 48 percent
 B. 46 percent E. not enough information to determine the
 C. 107 percent percentage precisely
14. Suppose x is a variable describing the heights of a group of 200 male professional athletes, half of whom are horse racing jockeys and half of whom are basketball players. Which of the following will be the best description of the distribution of this variable.
- A. The distribution would likely have one peak and be symmetric.
 B. The distribution would likely be uniform.
 C. The distribution would likely have one peak and be slightly skewed right.
 D. The distribution would likely have two peaks and be slightly skewed right.
15. Suppose x is a variable describing the age of a high school student in a high school that has 3000 students. In this high school, seniors, sophomores, juniors and freshman all have an approximately equal class size. Which of the following will be the best description of the distribution of the variable x .
- A. The distribution would likely be uniform.
 B. The distribution would likely be one peaked.
 C. The distribution would likely have two peaks.
 D. The distribution would likely be one peaked and strongly right skewed.
16. Which of the following measures the direction and strength of the linear association between X and Y ?
- A. IQR B. Median C. Correlation r D. Mean

17. For which of the following situations would it be appropriate to calculate r , the correlation coefficient?
- The age of a person and their income.
 - Income for government employees and the city they reside in.
 - Eye color and hair color on a driver's license.
 - Hours studied for this exam and their respective countries.
18. A researcher wants to know if taking increasing amounts of ginkgo biloba will result in increased capacities of memory ability for different students. He administers it to the students in doses of 250 milligrams, 500 milligrams, and 1000 milligrams. What is the explanatory variable in this study?
- Size of the student's brain.
 - Whether the student takes the ginkgo biloba.
 - Amount of ginkgo biloba given to each student.
 - Change in memory ability.

19. The graph to the right shows the linear relationship between diamond size and price for diamonds size 0.35 carats or less. Using this relationship to predict the price of a diamond that is 1 carat is considered

- Extrapolation.
- An influential observation.
- the IQR
- A lurking variable



20. Using data on the appraised value of homes, a real estate agent computes the least-squares regression line for predicting a home's value in 2002 from its value in 1992. The equation of the least-squares regression line is

$$y = \$22,000 + 1.6x$$

where y represents a home's value in 2002 and x is the value in 1992. Which of the following is true of the least-squares regression line?

- The slope is the change in a home's value in 2002 that would be predicted by a \$1 increase in a home's value in 1992.
- It always passes through the point, $(\bar{x}, \bar{y})$, the means of the values of home's (that were used to compute the least-squares regression line) in 1992 and 2002, respectively.
- It will only pass through all the data points if $r = \pm 1$.
- All of the above.

Show Your Work

Show all work clearly and neatly. No work shown means no credit will be given. Use correct notation to get full credit. Reserve scratch paper work for scratch paper, which means only include necessary work on the exam. Erase all mistakes neatly. Keep it neat!

21. A sample was taken of the verbal GRE scores of 20 applicants to graduate school at a large midwestern university. Below are the scores.

520	310	340	350	370	410	420	420	420	720
490	510	280	520	610	670	600	770	750	470

- (a) (5 pts) The mean applicant score is
- (b) (5 pts) The standard deviation for the applicant's is
- (c) (5 pts) The median applicant score is
- (d) (5 pts) The five number summary for the applicant data is
- (e) (5 pts) Are there any outliers for the applicant data? Explain why or why not.

22. (10 pts) Performance ratings of stocks are approximately Normal with mean 100 and standard deviation 20. The stocks with the LOWEST 10% performance ratings are classified as “Bleh” stocks. What performance rating separates the “Bleh” stocks from the other stocks?
23. (10 pts) The seller and the buyer both review the prices of the last 5 textbooks sold. The seller makes a case that the “average” price is 20 dollars, while the buyer makes a case that the “average” price of the same 5 textbooks is only 15 dollars. Construct an example of 5 textbook prices that could explain how both sides may be correct, identifying the measure of center you think each side is using.

24. The Leaning Tower of Pisa is an architectural wonder. Engineers concerned about the tower's stability have done extensive studies of its increasing tilt. Measurements of the lean of the tower over time provide much useful information. The following table gives measurements for the years 1975 to 1987. The variable "lean" represents the difference between where a point on the tower would be if the tower were straight and where it actually is. The lean is measured in meters.

	Year	Lean
(a) (5 pts) Find the correlation r of these data. Use your calculator. State only the answer.	1975	2.9642
	1976	2.9644
	1977	2.9656
	1978	2.9667
(b) (5 pts) Find the least-squares regression line for predicting y from x . Use your calculator. State only the answer.	1979	2.9673
	1980	2.9688
	1981	2.9696
	1982	2.9698
	1983	2.9713
(c) (5 pts) Is the relationship strong? If not, describe how good the relationship is.	1984	2.9717
	1985	2.9725
	1986	2.9742
	1987	2.9757

- (d) (10 pts) Make a scatterplot and draw the regression line on the plot.

- (e) (5 pts) Using the regression line, predict how far the lean will be in 2008.

- (f) (5 pts) Are there any reasons that the prediction may not be valid? Explain why or why not.

Table entry for z is the area under the standard Normal curve to the left of z

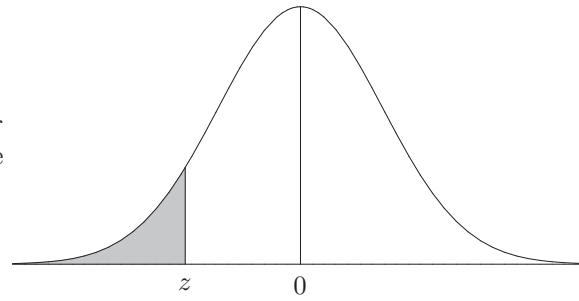


Table A: Standard Normal probabilities

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
-3.4	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0002
-3.3	.0005	.0005	.0005	.0004	.0004	.0004	.0004	.0004	.0004	.0003
-3.2	.0007	.0007	.0006	.0006	.0006	.0006	.0006	.0005	.0005	.0005
-3.1	.0010	.0009	.0009	.0009	.0008	.0008	.0008	.0008	.0007	.0007
-3.0	.0013	.0013	.0013	.0012	.0012	.0011	.0011	.0011	.0010	.0010
-2.9	.0019	.0018	.0018	.0017	.0016	.0016	.0015	.0015	.0014	.0014
-2.8	.0026	.0025	.0024	.0023	.0023	.0022	.0021	.0021	.0020	.0019
-2.7	.0035	.0034	.0033	.0032	.0031	.0030	.0029	.0028	.0027	.0026
-2.6	.0047	.0045	.0044	.0043	.0041	.0040	.0039	.0038	.0037	.0036
-2.5	.0062	.0060	.0059	.0057	.0055	.0054	.0052	.0051	.0049	.0048
-2.4	.0082	.0080	.0078	.0075	.0073	.0071	.0069	.0068	.0066	.0064
-2.3	.0107	.0104	.0102	.0099	.0096	.0094	.0091	.0089	.0087	.0084
-2.2	.0139	.0136	.0132	.0129	.0125	.0122	.0119	.0116	.0113	.0110
-2.1	.0179	.0174	.0170	.0166	.0162	.0158	.0154	.0150	.0146	.0143
-2.0	.0228	.0222	.0217	.0212	.0207	.0202	.0197	.0192	.0188	.0183
-1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233
-1.8	.0359	.0351	.0344	.0336	.0329	.0322	.0314	.0307	.0301	.0294
-1.7	.0446	.0436	.0427	.0418	.0409	.0401	.0392	.0384	.0375	.0367
-1.6	.0548	.0537	.0526	.0516	.0505	.0495	.0485	.0475	.0465	.0455
-1.5	.0668	.0655	.0643	.0630	.0618	.0606	.0594	.0582	.0571	.0559
-1.4	.0808	.0793	.0778	.0764	.0749	.0735	.0721	.0708	.0694	.0681
-1.3	.0968	.0951	.0934	.0918	.0901	.0885	.0869	.0853	.0838	.0823
-1.2	.1151	.1131	.1112	.1093	.1075	.1056	.1038	.1020	.1003	.0985
-1.1	.1357	.1335	.1314	.1292	.1271	.1251	.1230	.1210	.1190	.1170
-1.0	.1587	.1562	.1539	.1515	.1492	.1469	.1446	.1423	.1401	.1379
-0.9	.1841	.1814	.1788	.1762	.1736	.1711	.1685	.1660	.1635	.1611
-0.8	.2119	.2090	.2061	.2033	.2005	.1977	.1949	.1922	.1894	.1867
-0.7	.2420	.2389	.2358	.2327	.2296	.2266	.2236	.2206	.2177	.2148
-0.6	.2743	.2709	.2676	.2643	.2611	.2578	.2546	.2514	.2483	.2451
-0.5	.3085	.3050	.3015	.2981	.2946	.2912	.2877	.2843	.2810	.2776
-0.4	.3446	.3409	.3372	.3336	.3300	.3264	.3228	.3192	.3156	.3121
-0.3	.3821	.3783	.3745	.3707	.3669	.3632	.3594	.3557	.3520	.3483
-0.2	.4207	.4168	.4129	.4090	.4052	.4013	.3974	.3936	.3897	.3859
-0.1	.4602	.4562	.4522	.4483	.4443	.4404	.4364	.4325	.4286	.4247
-0.0	.5000	.4960	.4920	.4880	.4840	.4801	.4761	.4721	.4681	.4641

Table entry for z is the area under the standard Normal curve to the left of z

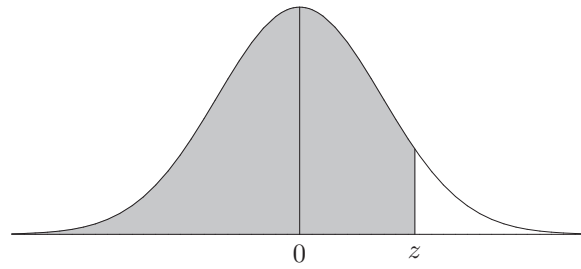


Table A: Standard Normal probabilities

[illegible]