UNIVERSITY OF CALGARY

UNIVERSITY OF CALGARY
FACULTY OF SCIENCE
DEPARTMENT OF MATHEMATICS AND STATISTICS
MIDTERM TEST 1
STAT 213 L07 (de LEON) – FALL 2015

DATE: OCTOBER 15, 2015 **Time: 50 minutes**

Student ID Number	Last Name	First Name

EXAMINATION RULES

1. This is a closed-book-no-notes test.
2. You are allowed to use the attached formula sheet and normal table. You are also allowed use of MINITAB and/or your own non-programmable non-graphing calculator.
3. Indicate your selection clearly on the scantron sheet provided by shading one of the circled letters A, B, C or D corresponding to your answer.
4. Scantron sheets must be filled out during the exam time limit. No additional time will be granted to fill in your scantron form.
5. The use of personal electronic or communication devices is prohibited.
6. A University of Calgary Student ID card is required to write the midterm test. If adequate ID isn't presented, the student must complete an Identification Form.
7. Students late in arriving will not be permitted after one-half hour of the test time has passed.
8. No student will be permitted to leave the room during the first 30 minutes, nor during the last 15 minutes of the test. Students must stop writing and hand in their test immediately when time expires.
9. All inquiries and requests must be addressed to your lab TA.
10. Students are strictly cautioned against:
    a. communicating with other students;
    b. leaving answer papers exposed to view;
    c. attempting to read other students' test papers
11. Once the test has been handed in for marking, a student cannot request that the test be cancelled. Retroactive withdrawals from the course will be denied.
12. Failure to comply with these regulations will result in rejection of the test paper.

No. of Questions	Points per Question
25	4
Percent	

Use the following to answer Questions 1-3:

Suppose that the blood alcohol context (BAC) of students who drink 5 beers varies from student to student according to a normal distribution with mean 0.08 and standard deviation 0.01.

1. The middle 95% of students who drink 5 beers have BACs between

   A) 0.07 and 0.09.
   B) 0.06 and 0.10.
   C) 0.05 and 0.11.
   D) 0.04 and 0.12.

2. What percent of students who drink 5 beers have BACs above 0.08 (the legal limit for driving in most U. S. states)?

   A) 2.5%
   B) 5%
   C) 16%
   D) 50%

3. What percent of students who drink 5 beers have BACs above 0.10 (the legal limit for driving in some U. S. states)?

   A) 2.5%
   B) 5%
   C) 16%
   D) 50%

4. If the heights of 99.7% of Canadian men are between 5 feet and 7 feet, what is your estimate of the standard deviation of the heights of Canadian men? Assume the heights of Canadian men are approximately normally distributed.

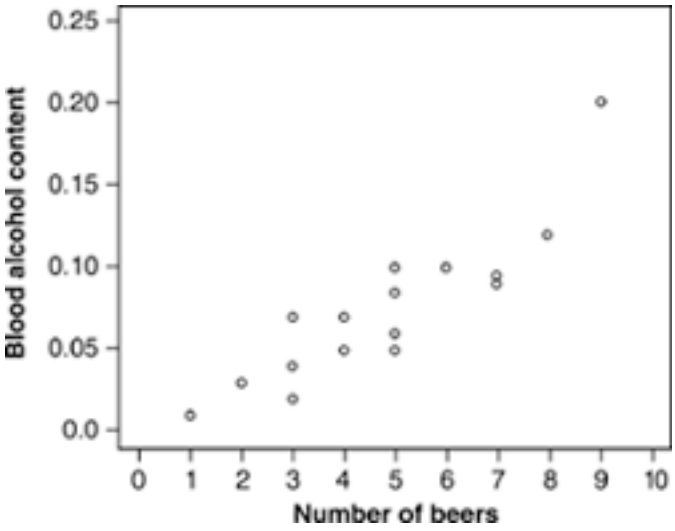
   A) 1 inch
   B) 3 inches
   C) 4 inches
   D) 6 inches

5. Scores of the adult population on a common "IQ" test are approximately normally distributed with mean 110 and standard deviation 15. About 40% of the scores are then between

   A) 80 and 140.
   B) 65 and 155.
   C) the 25th and 75th percentiles.
   D) the 30th and 70th percentiles.

Use the following to answer Questions 6-7:

How well does the number of beers a student drinks predict his or her BAC? Sixteen student volunteers at Mount Royal University drank a randomly assigned number of cans of beer. Thirty minutes later, a police officer measured their BACs. A scatterplot of the data appears below.



6. One student drank 9 beers. You see from the scatterplot that his BAC was about

A) 0.19.
B) 9.
C) 19.
D) 0.05.

7. A plausible value of the correlation between number of beers and blood alcohol content, based on the scatterplot, is

A) $r = -0.9$. B) $r = -0.3$. C) $r = 0.3$. D) $r = 0.9$.

8. The correlation between height (in inches) and weight (in pounds) among first-grade students at St. Patrick Elementary School is exactly 0.57. If heights are converted to centimeters and weights are converted to kilograms, what happens to the correlation between height and weight among the first-graders? (1 inch=2.54 cm.; 1 lb.=0.394 kg.)

A) The correlation is still 0.57.
B) The correlation gets stronger.
C) The correlation gets weaker.
D) It is not possible to tell how the correlation will change without further information.

9. Which of the following statements about correlation r is false?

A) r describes how tightly the points on a scatterplot cluster about a straight line.
B) r measures the proportion of the variance of one variable that can be explained by straight line dependence on the other variable.
C) r can never take a value larger than 1.
D) The value of r is heavily influenced by outliers.

10. Which of these statements about the standard deviation s is true?

A) s is always 0 or positive.
B) s should be used to measure spread only when the mean $\bar{x}$ is used to measure center.
C) s is a number that has no units of measurement.
D) Both (A) and (B), but not (C).

11. The United Nations also has data on the percent of adult females who are illiterate in each of these 142 countries. The correlation between male illiteracy rate and female illiteracy rate is r =0.945. This tells us that

A) countries with high male illiteracy tend to also have high female illiteracy, and the relationship is very strong.
B) countries with high male illiteracy tend to also have high female illiteracy, but the two are only weakly related.
C) countries with high male illiteracy tend to have low female illiteracy, and the relationship is very strong.
D) countries with high male illiteracy tend to have low female illiteracy, but the two are only weakly related.

12. If 30 is added to every value on a data set, the only one of the following that is *not* changed is

A) the mean.
B) the 75^{th} percentile.
C) the median.
D) the standard deviation.

Use the following to answer Questions 13-14:

The five-number summary for scores on a statistics exam is {11, 35, 61, 70, 79}. In all, 380 students took the test.

13. About how many students had scores between 35 and 61?

A) 35
B) 61
C) 95
D) 190.

14. For the distribution of 380 exam scores,

A) the mean is very likely to be greater than the median.
B) the median is very likely to be greater than the mean.
C) the mean and the median are probably about equal.
D) we have no indication of whether the mean is larger than, smaller than, or equal to the median.

Use the following to answer Questions 15-16:

The stock market did well during the 1990s. Here are the percent total returns (change in price plus dividends paid) for the Standard & Poor's 500 stock index:

Year	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998
Return	31.7	–3.1	30.5	7.6	10.1	1.3	37.6	23.0	33.4	28.6

15. The median return during this period is

A) 5.5.
B) 20.07.
C) 23.
D) 25.8.

16. The third quartile of these returns is

A) 7.6.
B) 30.5.
C) 31.1.
D) 31.7.

17. The least-squares regression line is $\hat{y} = -0.013 + 0.018x$ for predicting BAC from number of beers. The slope 0.018 of this line tells us that

A) the correlation between number of beers and BAC is 0.018.
B) the predicted BAC increases by 0.018 for each additional beer a student drinks.
C) a student who drinks no beer will still have a BAC of 0.018.
D) the average BAC of all the students in the study was 0.018.

18. The correlation between two variables x and y is 0.5. If we used a regression line to predict y using x, what percent of the variation in y would be explained?

A) 50%.
B) 25%
C) 75%
D) 2.23%

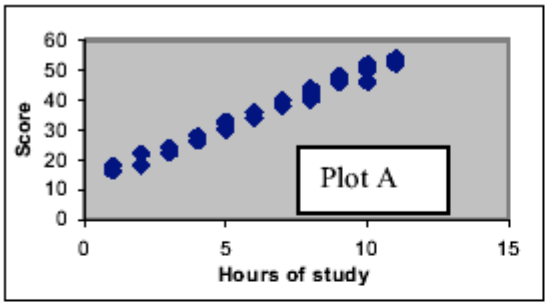
19. The correlation between the heights of fathers and the heights of their (grownup) sons is $r = 0.52$. This shows the fact that fathers have different heights

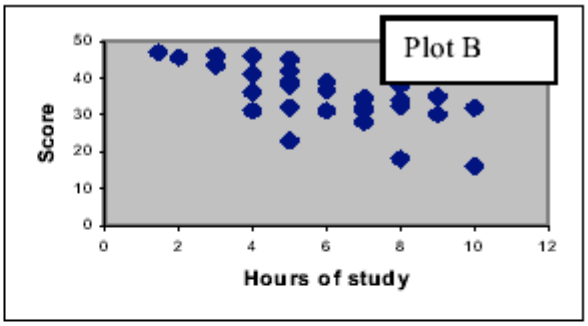
A) explains about 27% of the observed variation in their sons' heights.
B) explains about 52% of the observed variation in their sons' heights.
C) explains about 73% of the observed variation in their sons' heights.
D) explains about 95% of the observed variation in their sons' heights.

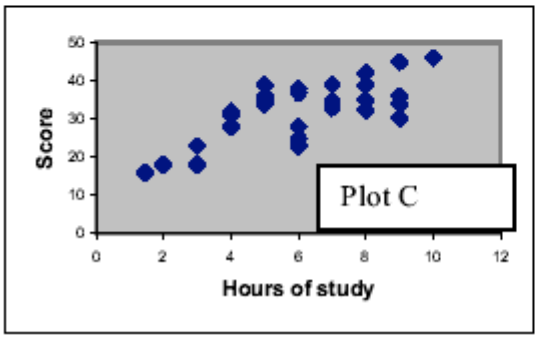
20. Suppose the mean number of goals per year in the first 10 years that Hockey Player A has played is 30 goals per year, with a standard deviation of 5 goals per year. If he scores 30 goals in his eleventh year, then his mean number of goals and the standard deviation for the 11-year period

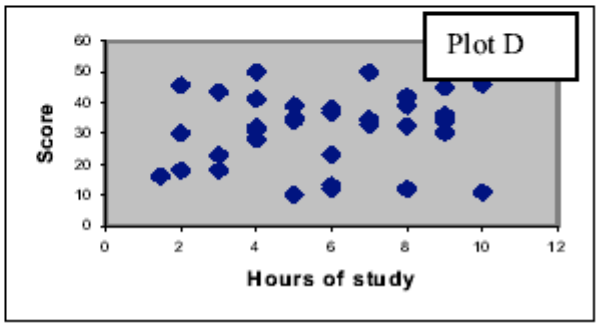
A) increases and decreases, respectively.
B) remains the same and decreases, respectively.
C) both remain the same.
D) remains the same and increases, respectively.

21.









A study of the relationship between the number of hours of study in the week immediately preceding an exam and the scores received by students in a small class showed that, in general, students who spent more hours studying for the exam obtained higher scores. The number of hours of study, when used to predict a student's score on the exam, was able to explain 55% of the variation in scores among the students. One of the following scatterplots is that one for the scores against the number of hours of study. Which one is it?

A) Plot A
B) Plot B
C) Plot C
D) Plot D

Use the following to answer Questions 22-23:

A well-known maker of jams and jellies packages its jams in jars labelled "250 ml". The process used to fill the jars is known to dispense an amount of jam that follows the normal distribution with μ=252 ml and σ=0.9 ml.

22. What proportion of the jars filled by the process contain less than 250 ml?

A) 0.5
B) 0.9868
C) 0.0068
D) 0.0132

23. What proportion of jars from this filling process contain no more than 253.5 ml?

A) 0.9678
B) 0.9525
C) 0.9973
D) 0.0027

24. There is a close relationship between the correlation r and the slope b of the least-squares regression line. In particular, it is true that

A) r and b both always take values between -1 and 1.
B) r and b always have the same sign, which shows whether the variables are positively or negatively associated.
C) the slope b is always at least as large as the correlation r.
D) the slope b is always equal to r^2, the square of the correlation.

25. The ages (in years) of students in a class are as follows:

Age	18	19	20	21	22	23	24
# students	14	15	68	12	18	2	2

What is the first quartile of the ages of students in this class?

A) It must be over 21.
B) It must be 20.
C) It must be 21.
D) It could be any age between 22 and 23.

FORMULAS

CHAPTER 1 Examining Distributions

- Remember overall pattern (shape, center, spread) and deviations.
- Mean (use a calculator):

$$\bar{x} = \frac{x_1 + x_2 + \cdots + x_n}{n} = \frac{1}{n}\sum x_i$$

- Standard deviation (use a calculator):

$$s = \sqrt{\frac{1}{n-1}\sum (x_i - \bar{x})^2}$$

- Median: Arrange all observations from smallest to largest. The median M is located $(n + 1)/2$ observations from the beginning of this list.
- Quartiles: The first quartile Q_1 is the median of the observations whose position in the ordered list is to the left of the location of the overall median. The third quartile Q_3 is the median of the observations to the right of the location of the overall median.
- Five-number summary:

$$\text{Minimum},\ Q_1,\ M,\ Q_3,\ \text{Maximum}$$

- Standardized observation from x:

$$z = \frac{x - \mu}{\sigma}$$

CHAPTER 2 Examining Relationships

- Remember overall pattern (form, direction, strength) and deviations.
- Correlation (use a calculator):

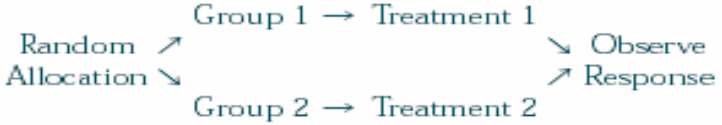
$$r = \frac{1}{n-1}\sum \left(\frac{x_i - \bar{x}}{s_x}\right)\left(\frac{y_i - \bar{y}}{s_y}\right)$$

- Least-squares regression line (use a calculator): $\hat{y} = a + bx$ with slope $b = rs_y/s_x$ and intercept $a = \bar{y} - b\bar{x}$
- Residuals:

$$\begin{aligned}\text{residual} &= \text{observed } y - \text{predicted } y \\ &= y - \hat{y}\end{aligned}$$

CHAPTER 3 Producing Data

- Simple random sample: Choose an SRS by giving every individual in the population a numerical label and using Table B of random digits to choose the sample.
- Randomized comparative experiments:



CHAPTER 4 Probability and Sampling Distributions

- Probability rules:
  - Any probability satisfies $0 \le P(A) \le 1$.
  - The sample space S has probability $P(S) = 1$.
  - For any event A, $P(A \text{ does not occur}) = 1 - P(A)$.
  - If events A and B are disjoint, $P(A \text{ or } B) = P(A) + P(B)$.
- Sampling distribution of a sample mean:
  - $\bar{x}$ has mean μ and standard deviation $\sigma/\sqrt{n}$.
  - $\bar{x}$ has a normal distribution if the population distribution is normal.
  - Central limit theorem: $\bar{x}$ is approximately normal when n is large.

Table entry for z is the area under the standard normal curve left of z.

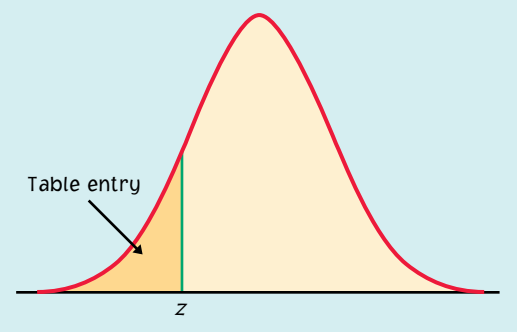


TABLE A Standard normal probabilities

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

STAT 213 L07
Midterm Test 1
Answer Key

1. B
2. D
3. A
4. C
5. D
6. A
7. D
8. A
9. B
10. D
11. A
12. D
13. C
14. B
15. D
16. D
17. B
18. B
19. A
20. B
21. C
22. D
23. B
24. B
25. B