BNAD 276. STATISTICAL INFERENCE IN MANAGEMENT

Presession. Summer 2017.

TEST 1. DATE: 19 MAY 2017

Time: 75 minutes

NOTE:

1. The exam has total 31 questions for a possible total of 33 points.

2. You can use a calculator. Your calculator may be programmable but you are not allowed to use the programmable tools.

3. I will not answer any questions related to the content of the questions. Do your best.

Good luck!

----

1. (3 points) You are given 12 formulas/descriptions/notations on the left and 15 statistical concepts on the right column. Please match each formulas/description to its corresponding statistical concepts. Notice there are **3 redundant concepts**, i.e. 3 concepts that do not have the corresponding descriptions.

Hence, you need to **match 12 formulas/descriptions to the correct 12 concepts**. Each correct match is worth 0.25 point.

1. $\dfrac{s_{xy}}{s_x s_y}$	A. Sample mean
	B. Sample variance
2. $s_{xy} = \dfrac{\sum(x_i - \bar{x})(y_i - \bar{y})}{n-1}$	C. Standard deviation
	D. Sample covariance
3. $A \cup B$	E. Correlation coefficient
4. Largest Value – Smallest Value	F. Mode
	G. Interquartile Range
5. $\dfrac{\sum x_i}{n}$	H. Range
	I. Union event
6. $A \cap B$	J. Intersection event
7. $P(A \cap B) = 0$	K. Mutually exclusive events
8. $P(A^c) = 1 - P(A)$	L. Independent events
	M. Complement event
9. $P(A|B) = P(A)$	N. Combination rule
10. Choose k objects out of N objects, order matters, without replacement	O. Permutation rule
11. $s = \sqrt{s^2}$	
12. $\dfrac{\sum(x_i - \bar{x})^2}{n-1}$	

(Note, for descriptions 7 and 8, please pick the corresponding most relevant concepts.)

Please write down your answers (by simply listing the relevant letter such as 1A, 2B, …) below:

1.	7.
2.	8.
3.	9.
4.	10.
5.	11.
6.	12.

2.   (1 point)

The measure of dispersion which is not measured in the same units as the original data is the

   a. median
   b. standard deviation
   c. coefficient of determination
   d. variance

3. (1 point)

The interquartile range is used as a measure of variability to overcome what difficulty of the range?

   a. the sum of the range variances is zero
   b. the range is difficult to compute
   c. the range depends too much on extreme values
   d. the range is negative

4.   (0.5 point) The maximum possible grade on a test is 10, i.e. the test grade scale is $1 - 10$. The result shows that the median grade in the class is 7. Which of the following statements is correct?

   a.   Because the midpoint of the grade is $10/2 = 5$, the variance is $7 - 5 = 2$
   b.   Because the midpoint of the grade is $10/2 = 5$, the standard deviation is $7 - 5 = 2$
   c.   At least half of the class has the grade above or equal to 7 and at least half of the class has the grade below or equal to 7
   d.   Because the midpoint of the grade is $10/2 = 5$, the median grade is greater than the mean grade.

5.   (0.5 point) The maximum possible grade on a test is 10, i.e. the test grade scale is $1 - 10$, and the class has 10 students. The result shows that the median grade in the class is 7. Which of the following statements is correct?

   a.   Because the midpoint of the grade and the midpoint of the class (size) is $10/2 = 5$, the variance is $7 - 5 = 2$
   b.   Because the midpoint of the grade and the midpoint of the class (size) is $10/2 = 5$, the standard deviation is $7 - 5 = 2$
   c.   At least half of the class (5 students) has the grade above or equal to 7 and at least half of the class (5 students) has the grade below or equal to 7.
   d.   Because the midpoint of the grade and the midpoint of the class (size) is $10/2 = 5$, the median grade is greater than the mean grade.

The following data set is used to answer questions 6, 7, 8

After graduation ceremonies at a university, six graduates were asked whether they were in favor of (identified by 1) or against (identified by 0) legalized drugs. Some information about these graduates is shown below.

Graduate	Sex	Age	Drugs Issue	Class Rank
Catherine	F	24	1	1
Jason	M	22	1	2
Wendy	F	41	0	3
Edward	M	38	0	20
Jennifer	F	25	1	4
Tim	M	19	0	8

6. How many variables are in the above data set? What are they? (1 point)

7. Which variable(s) is/are categorical? Which variable(s) is/are quantitative? (1 point)

8. Show the frequency distribution for the variable Drugs Issue. (1 point)

The following data is used for questions 9, 10, 11, 12, 13

The closing stock price of MNM Corporation for the last 7 trading days is shown below.

Day	Stock Price
1	84
2	87
3	84
4	88
5	85
6	90
7	91

9. (1 point) What is the type of this data set? (Choose a or b or c or d)

   a. Cross-sectional data
   b. Time-series data
   c. Panel data
   d. Longitudinal data

10. (1 point) What is the **mean** value of the stock price?

11. (1 point) What is the **median** value of the stock price?

12. (1 point) What is the **range** of the stock price?

13. (1 point) What is the **mode** value of the stock price? Hint: the value that occurs …?.. frequently

The following data set is used for questions 14, 15, 16, 17, 18, 19

The Michael Painting Company has purchased paint from several suppliers. The purchase price per gallon and the number of gallons purchased are shown below.

Supplier	Price Per Gallon ($)	Number of Gallons
A	15	40
B	22	30
C	30	10
D	27	20

14. What is the **sample variance** of the **Price Per Gallon**? (1 point) Hint: You are given a mysterious table below, which may be useful.

15		-8.5	72.25
22		-1.5	2.25
30		6.5	42.25
27		3.5	12.25

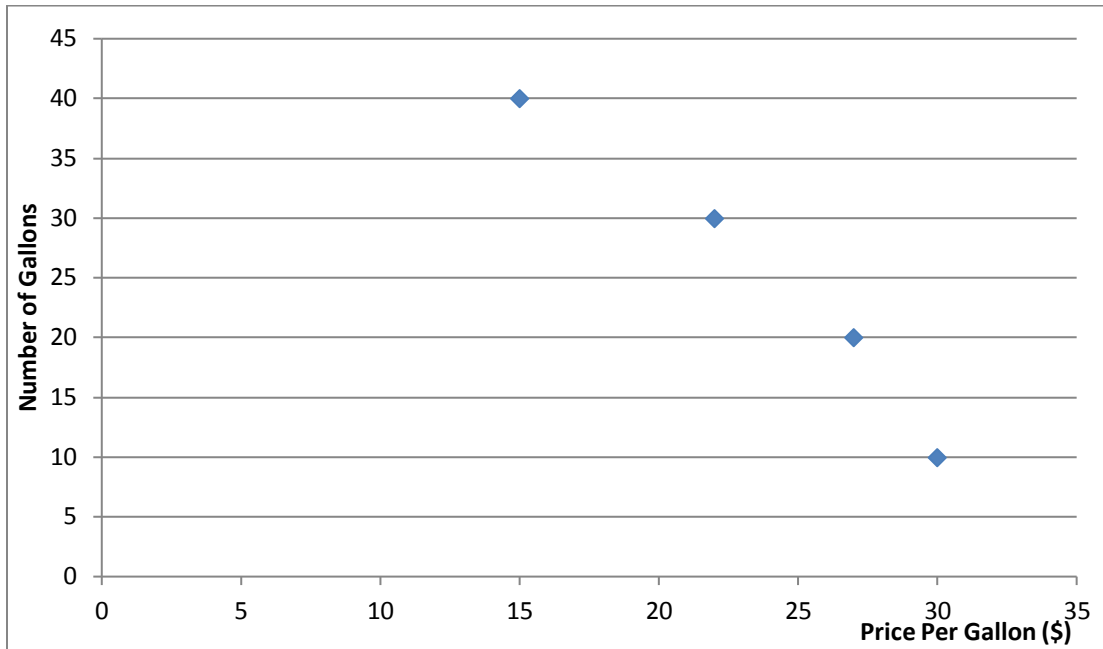15. Calculate the **sample standard deviation** of the Price Per Gallon? (1 point)

16. Calculate **the sample covariance** of the Price Per Gallon and the Number of Gallons? (1 point).
    Hint: You are given a mysterious table below, which may be useful.

15	40	-8.5	72.25	15	225	-127.5
22	30	-1.5	2.25	5	25	-7.5
30	10	6.5	42.25	-15	225	-97.5
27	20	3.5	12.25	-5	25	-17.5

17. What is the name of the plot below? (1 point)

18. What does the sample covariance tell you about the relation between the two variables Price Per Gallon and Number of Gallons? (1 point)
(If you could not calculate the covariance, have a look at the plot in question 17, what can you tell about the relationship between the two variables?)

19. (1 point) What do you think about the **sign and the magnitude** of the sample **correlation coefficient** of the Price Per Gallon and the Number of Gallons? Specifically, is it near to - 1, 0, or 1? What does it mean? Hint: You can either calculate the correlation coefficient to answer this question, or you can look at the plot in question 17 to answer this question. In case you want to calculate the correlation coefficient, you may find it useful to look at the mysterious table in question 16.

Peter is conducting an experiment of rolling a die. Answer the questions 20, 21, 22

20. How many possible outcomes does the above experiment have? (1 point)

21. What is the probability of each outcome if Peter thinks the die is a fair die, i.e. the die does not have any faults? (1 point)

22. Suppose now, Peter is doing a second step in which he tosses a fair coin. How many possible outcomes, in total, does the new experiment have (the experiment includes both rolling a die and tossing the coin)? **Show the work (i.e. not just a final number) why/how you got your answer.** (1 point)

23. John is doing an experiment in which he chooses **3 balls out of 4** balls **at a time** (he does not care the order). How many possible outcomes does this experiment have? **Show the work (i.e. not just a final number) why/how you got your answer.** (1 point)

The following data set is used to answer questions 24, 25, 26, 27

A survey of 300 college seniors resulted in the following crosstabulation regarding their undergraduate major and whether or not they plan to go to graduate school.

Graduate School	Undergraduate Major		Total
	Business	Engineering	
Yes	70	30	100
No	110	90	200
Total	180	120	300

24. Show the joint probability distribution table. Include marginal probabilities in the table. (2 points)

25. What is the probability that we will see a student who majors in engineering and also plans to attend graduate school? (1 point)

26. What is the probability that we will see a student majoring in **business**? (1 point)

27. What is the probability that we will see a student going to the graduate school given that we know the student majors in engineering? (1 point)

Bank of America has conducted a survey on how frequently consumers use credit cards in making a purchase. The results of the study provide the following information. (The survey only interviews people who are above 18 years old.)

37% of consumers use a credit card when making a purchase.

Among those consumers who use a credit card, 60% are **under** 40 years old.

45% of people interviewed for the survey are **above** 40 years old.

Let Y (young) = event that a consumer is under 40 years old (and of course, above 18 years old.)

Let U (use a credit card) = event that a consumer uses a credit card in making a purchase.

Answer questions 28, 29, 30, 31 below:

28. Determine the value of P(U) (1 point)

29. Determine the value of P(Y) (1 point)

30. (1 point) What is the probability that a consumer uses a credit card given that the consumer is under 40 (and above 18) years old? Hint: conditional probabilities and the formulas in the handed out sheet are useful.

31. (1 point) What can you say about the two events Y and U? Hint: compare P(U) and P(U|Y), and recall something similar to 'relevant' or 'irrelevant'?

    If you can't calculate P(U|Y), think about the relation between Y and U in a real life situation; of course, you have to **name a correct statistical concept** to get credit.