

8. Which of the following statement is INCORRECT about a binomial Distribution random variable  $X$ , with mean 4 and variance 2?
- A. The number of trial is 10
  - B. The probability of failure is the same as the probability of success
  - C.  $P(X = 0) = 0.003906$
  - D.  $P(X \geq 0) = 1$
  - E.  $P(X \geq 15) = 0$

9. If the mean of a poison Distribution random variable,  $X$ , is 5, then
- A. Its variance is 25
  - B.  $P(X = 0) = 0.5$
  - C. Its variance is 5
  - D.  $P(X \geq 1) = 1$
  - E. C and D

10. Which one of the following does not belong to random sampling?
- A. Cluster sampling
  - B. Convenience sampling
  - C. Systematic sampling
  - D. Stratified sampling
  - E. None of the above

11. The office of dean of students wanted the study the opinion of students regarding the service of the cafeteria. It selected 70 students from their meal card number list of 2500 by taking every fifteenth number after selecting the first item randomly. What type of sampling technique did they use?
- A. Simple random sampling
  - B. Systematic random sampling
  - C. Stratified random sampling
  - D. Cluster random sampling
  - E. None of the above

12. The probability that a patient with a certain disease will be successfully treated with a medical treatment is  $p = 0.80$ . Suppose that the treatment is used on 40 patients. What is "expected value" of the number of patients who are successfully treated?
- A. 40
  - B. 20
  - C. 8
  - D. 32
  - E. 44

13. The Vitamin C content of a particular brand of vitamin supplement pills is normally distributed with mean 490 mg and standard deviation 12 mg. What is the probability that a randomly selected pill contains at least 500mg of Vitamin C?
- A. 0.7967
  - B. 0.8333
  - C. 0.0525
  - D. 0.1123
  - E. 0.2033

14. In a study of possible correlation between the height in cm (X) and weight in kg (Y) chimpanzees, a sample of 40 chimpanzee produces a regression line with equation  $\hat{Y} = 19.5 + 0.34X$ . What is the expected weight of an 80 cm chimpanzee?
- A. 46.7 kg
  - B. 177.9 kg
  - C. 24.0 kg
  - D. 34.8 kg
  - E. 57.1 kg

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(A|B)}{P(B)}$$

Part II: Fill in the Blank Space with correct and precise Answer (each carries 1.5 points)

1. In a standard Normal Distribution table if the area between 0 and Z is  $\alpha$ , find the following areas in terms of  $\alpha$ .
  - a) Area between  $-Z$  and  $Z$ :  $2\alpha$
  - b) Area less than  $-Z$ :  $\alpha - \frac{\alpha}{2}$
2. Given the following probability distribution for a random variable X representing the number of days a student is absent from statistics class. 1.5 pts

X	0	1	2	3	4
P(X=x)	0.2	0.1	P	0.4	0.2

- a) The probability that a student is absent for two days from statistics class is:  $0.1$
- b) The average number of days a student is absent from statistics class is:  $2$
- c) The Standard deviation of number of days a student is absent from statistics class is:  $\sqrt{2}$
3. An error that occurs in both sample and census survey is called ~~Sampling~~ ~~and~~ ~~Sampling~~
4. Arithmetic mean and variance of observations for a variable X are found to be 30 and 90 respectively. what is the coefficient of variation of the new observations Y obtained by linear transformation  $Y=2X-47$ ?  $CV = 31.6\%$
5. The life time of machine, say X, is assumed to be a continuous random variable with probability density function (pdf):

$$f(x) = \begin{cases} ke^{-x/3} & \text{for } x>0 \\ 0 & \text{Otherwise} \end{cases}$$

- a) Determine k =  $\frac{1}{3}$
- b) Find  $E(X) = 9$  and  $\text{Var}(X) = 18$
- c) Calculate  $\text{var}(kX) = K^2 \cdot 9$

6. List the advantages of sampling over census  
+ve sampling :-
  - ✓ it save time & cost
  - ✓ it prevent destruction
  - ✓ sampling having accurate than census

Part III: workout; show clearly and neatly all necessary steps on the space provided

1. A human gene carries a certain disease from the mother to the child with a probability rate of 35%. That is, there is a 35% chance that the child becomes infected with the disease. Suppose a female carrier of the gene has five children. Assume that the infections of the five children are independent of one another. (2+2=4 Points)
- Find the probability that at least two of the children get the disease from their mother.
  - Find mean and standard deviation of the children get the disease from their mother.

Soln

Let  $x$  represent number of children

$$x \sim B(5, 0.35)$$

$$\begin{aligned} a, P(x \geq 2) &= 1 - [P(x=0) + P(x=1)] \\ &= 1 - \left[ \binom{5}{0} (0.35)^0 (1-0.35)^5 + \binom{5}{1} (0.35)^1 (1-0.35)^4 \right] \\ &= 1 - [0.116 + 0.312] \\ &= \underline{\underline{0.572}} \quad \checkmark \text{ (2)} \end{aligned}$$

$$b, E(x) = np = 3.75 \quad \checkmark$$

$$\text{standard deviation} = \sqrt{\text{variance}}$$

$$\begin{aligned} \text{variance} (V(x)) &= np(1-p) \\ &= (3.75)(0.65) \\ &= 2.4375 \end{aligned}$$

(2)

$$\text{standard deviation} = \sqrt{2.4375} = \underline{\underline{1.5665}}$$

2. A soft drink machine is regulated so that it discharges an average of 200 milliliters per cup. If the amount of drink is normally distributed with a standard deviation equal to 15 milliliters (1.5 points each= 6 Points)

- a) What fraction of the cups will contain more than 224 milliliters?  
 b) What is the probability that a cup contains between 191 and 209 milliliters?  
 c) How many cups over will probably overflow if 230-milliliter cups are used for the next 100 drinks?

d) Below what value do we get the smallest 25% of the drinks?

SOL: lot x rock soft rock

$$\mu = 200$$

$$\sigma = 15 \sim \frac{224 - \mu}{\sigma}$$

$$P(x > \frac{224 - 200}{15}) = P(x > 1.6)$$

$$P(x > 1.6) = 0.5 - P(0 \leq x \leq 1.6)$$

$$= 0.5 - 0.458 = 0.058$$

$$P(\frac{191 - 200}{15} \leq x \leq \frac{209 - 200}{15})$$

$$= P(-0.6 \leq x \leq 0.6) = P(0 \leq x \leq 0.6)$$

$$P(x < -0.6) = P(x > 0.6) = 0.5$$

$$= 0.451 = 0.5$$

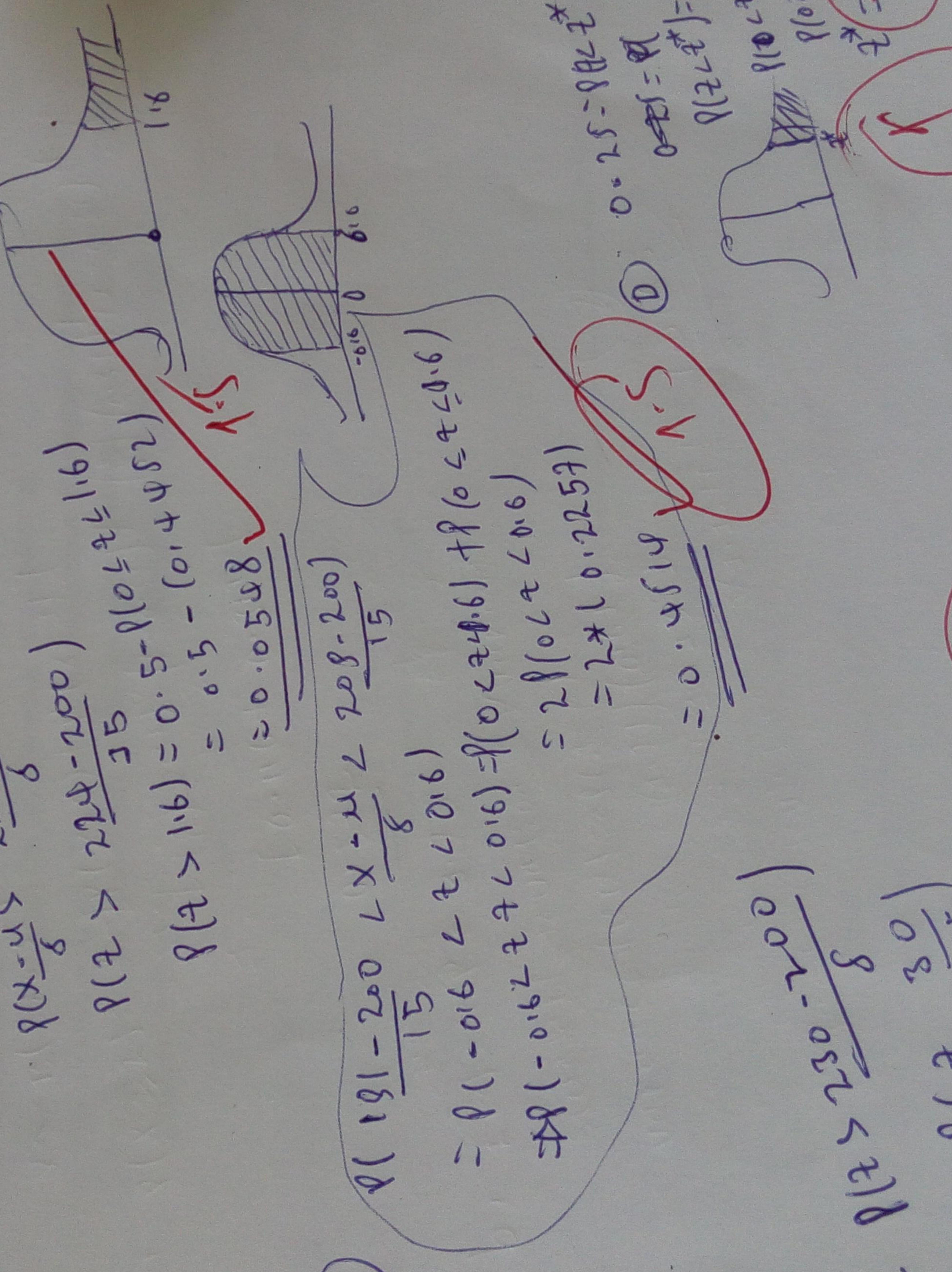
$$P(x > \frac{230 - 200}{15})$$

$$= P(x > 2) = \frac{3}{5}$$

$$= P(x > 2) = 0.477$$

$$P(x > 2) = 0.5 - P(x \leq 2) = 0.5 - 0.477 = 0.023$$

$$P(x > 2) = 0.5 - P(x \leq 2) = 0.5 - 0.477 = 0.023$$




3. The number of pounds of steam used per month by a chemical plant is thought to be related to the average ambient temperature ( $^{\circ}F$ ) for that month. The past year's usage and temperature are shown in the following table:

Month	Jan.	Feb.	Mar.	Apr.	May	Jun.	Jul.	Aug.	Sep.	Oct.	Nov.	Dec.
Temp.	21	24	32	47	50	59	68	74	62	50	41	30
Usage (1000)	18	21	28	42	45	54	62	67	56	45	37	27

Using the above information

- a) Draw the scatter plot and comment the relationship among variables.  
 b) Assuming that a simple linear regression model is appropriate; fit the regression model relating steam usage ( $y$ ) to the average temperature ( $x$ ).  
 c) What is the estimate of expected steam usage when the average temperature is  $55^{\circ}F$ ?  
 d) Compute simple correlation coefficient between steam usage & average temperature and interpret the result.



0 Answer

Part I: For the following questions, choose the best alternative and circle on the letter you choose. 1 pts

1. A committee of 6 people is to be formed from a group of 20 people. The committee needs to have the number of women double that of the number of men. If there are 12 men in the group, in how many ways can this committee be formed?  
a. 13860 b. 4620 c. 1836 d. 9240 e. none
2. The collection of all outcomes of an experiment is called  
a. event b. Sample space c. Sample point d. probability e. None

Part III: attempt the following questions by showing your necessary steps

1. The probability that a cellular phone company NOKIA sells  $X$  number of new phone contracts per day is shown below. (3 points)

X	4	5	6	8	10
P(X)	0.4	0.3	0.1	0.15	0.05

- i. What is the probability that they will sell 6 or more contracts three days in a row? (1 pts)
- ii. Find the mean, variance, and standard deviation for this probability distribution. (2 pts)

Soln: ①

$$P(6 \text{ or more}) = P(6) + P(8) + P(10)$$
$$= 0.1 + 0.15 + 0.05$$

$$\text{mean} = \sum x \cdot p(x) = (4 \cdot 0.4) + (5 \cdot 0.3) + (6 \cdot 0.1) + (8 \cdot 0.15) + (10 \cdot 0.05)$$

$$E(x) = \mu = 5.0$$

$$\text{variance } V(x) = \sigma^2 = (x - \mu)^2 \cdot p(x) = (4 - 5)^2 \cdot 0.4 + (5 - 5)^2 \cdot 0.3 + (6 - 5)^2 \cdot 0.1 + (8 - 5)^2 \cdot 0.15 + (10 - 5)^2 \cdot 0.05$$

$$V(x) = 4.045$$

$$\text{standard deviation } \delta = \sqrt{V(x)} = \sqrt{4.045} = 2.01$$

1. Let A and B be events such that  $P(A \cup B) = 0.75$ ,  $P(A \cap B) = 0.25$  and  $P(\bar{A}) = 0.6$ .

a) What is the  $P(A)$ ?  $P(A) = 1 - P(\bar{A}) = 1 - 0.6 = 0.4$

b) What is the  $P(B)$ ?  $P(A \cup B) = P(A) + P(B) - P(A \cap B) \Rightarrow 0.4 + P(B) = 0.75 \Rightarrow P(B) = 0.35$

c) What is the  $P(A|B)$ ?  $P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{0.25}{0.35} = 0.71$

d) Are A and B mutually exclusive? (yes/no) NO

4

$\frac{600}{1000} \text{ true}$   
 v.v  
 $\frac{600}{1000}$   
 v.v  
 0.006

2. Suppose it is known that the probability of recovery for a certain disease is 0.4. If random sample of 10 people who are stricken with the disease are selected, what is the probability that:
- Exactly 5 of them will recover? 0.2
  - At most 9 of them will recover? 0.95

a Let  $x = \text{# of recoveries}$  if 5 recovery of 5/10

$$\begin{aligned}
 P(x=5) &= \binom{10}{5} (0.4)^5 (1-0.4)^{10-5} \\
 &= (252)(0.01024)(0.6)^5 \\
 &= (2.58)(0.077) \\
 &= 0.2 \quad \text{∴ the probability is exactly 5 will recover}
 \end{aligned}$$

b Let  $x = \text{# of 9 will recover}$   
 If max one or zero will recover

$$\begin{aligned}
 P(x \leq 9) &= 1 - [P(x=0) + P(x=1)] \\
 &= 1 - [\binom{10}{0}(0.4)^0(0.6)^{10} + \binom{10}{1}(0.4)^1(0.6)^9] \\
 &= 1 - [0.006 + 0.04] \\
 &= 0.95
 \end{aligned}$$

3. The diameter of electric cables, say  $X$ , is assumed to be a continuous random variable with probability density function (pdf):

$$f(x) = \begin{cases} kx(1-x), & 0 \leq x \leq 1 \\ 0 \text{ otherwise} \end{cases}$$

- Find constant K.
- find  $E(X)$
- Find  $V(X)$
- standard deviation of X

a  $\int kx - kx^2 \stackrel{1}{=} 1, \quad 0 \leq x \leq 1$

$$\begin{aligned}
 &= \int (kx - kx^2) dx \\
 &= \left[ \frac{kx^2}{2} - \frac{kx^3}{3} \right]_0^1 \\
 &= \frac{k}{2} - \frac{k}{3} - (0) \\
 &= k\left(\frac{1}{2} - \frac{1}{3}\right) = 1 \\
 &\frac{3-2}{6} = \frac{1}{6} \quad \therefore k = 1
 \end{aligned}$$

b  $E(X) = \int x(x-x^2) dx$

$$\begin{aligned}
 &= \int_0^1 (x^2 - x^3) dx \\
 &= \frac{1}{3}(x^3 - \frac{x^4}{4}) \Big|_0^1 \\
 &= \frac{1}{3}(1 - \frac{1}{4}) = \frac{1}{3} \cdot \frac{3}{4} = \frac{1}{4}
 \end{aligned}$$

c  $E(X) = \mu = \frac{1}{2}$

d  $f_X(x) = \frac{1}{6}(x-x^2)$

b  $E(X) = \int x^2 dx - \bar{x}^2$

$$\begin{aligned}
 &= \int_0^1 x^2 dx - \left(\frac{1}{2}\right)^2 \\
 &= \left[\frac{x^3}{3}\right]_0^1 - \frac{1}{4} \\
 &= \frac{1}{3} - \frac{1}{4} = \frac{1}{12} \\
 &\sigma^2 = \text{var} = \frac{1}{12} - \frac{1}{4} = \frac{5}{48} \\
 &\sigma = \sqrt{\text{var}} = \sqrt{\frac{5}{48}} = \frac{\sqrt{5}}{4\sqrt{3}} = \frac{\sqrt{15}}{12}
 \end{aligned}$$

**Part I: For the following questions, choose the best alternative and write it on the blank space. 1 pts**

- C 1. Which of the following is not true about statistics?
- A. statistics can be used aid in decision making based on data.
  - B. Statistics deals with only quantitative information.
  - C. statistics deals with information pertaining to individual objects and no importance attached to aggregate facts.
  - D. All of the Above E none
- C 2. Is a branch of modern statistics which facilitates estimation the characteristics of population or making decisions concerning a population on the basis of sample results.
- A. Inferential Statistics B. descriptive statistics C. A and B D. none
- A 3. If you have data on continuous variable organized in the form of grouped frequency distribution, which graphic presentation is appropriate to present the data?
- A: Histogram B: Bar chart C Pie chart D. All of the Above E none
- A 4. Which of the following is different from the others?
- A. Document review B. Interview C. telephone survey D. Observational E. none
- A 5. Which of the following is property of mode?
- A:  0 B:  5 C:  10 D:  15 E:  20
- A 6. A study was conducted to investigate the effect of a detergent factory upon the water quality of a river. As part of an environmental impact study, some fishes were captured tagged, and released. The following information is recorded for each fish: Sex(0=male, 1=female), length(cm), maturation(0=young, 1= adult), weight (g). Which of the following is true?
- A. The scale of measurement used is census.
  - B. The data collection method used with ratio scale of measurements.
  - C. all are wrong
  - D. Maturity is ordinal scale.
  - E. all are wrong

- The class width in a histogram is the difference between
- A. Two consecutive classes
  - B. Three consecutive classes
  - C. Three times of two consecutive classes
  - D. Any consecutive class marks
  - E. All of the above

Q. Which of the following method of data representation is not appropriate for presenting data

- A. Histogram
- B. Bar chart
- C. Frequency polygon
- D. A and B
- E. none

Q. 2. The mode of a series is

- A. Half way between the first and third quartiles.
- B. Always unique
- C. The most frequent value
- D. Always determinable
- E. A value which divides the series into four equal parts

A. II. Which of the following is not true?

- A. Time to complete marathon is a discrete variable
- B. Number of students in class are discrete variable
- C. Life time of electronic device is a continuous variable
- D. Colour of computer is a nominal variable
- E. All of the above

Q. 3. In the following question please write the short (most simplified) answer on the

- ~~Part III for the following question: please write the short (most simplified) answer on the right side of each question~~
- Q. 1. What should be the value of "A" be in order to obtain the smallest value of  $\sum_{i=0}^n (x_i - A)^2$  by considering 4, 5, 6, 20 as data set.
- ~~Q. 2. If  $s^2$  is a descriptive measure which is computed from population or summary value calculated from a population.~~

Test

Good luck!

Part I: Choose the best answer from among the alternatives and circle on for each question.

(each carries 1 point)

(3-3) - 3-

1. Which of the following is false
  - A. The numbers 3, 3, 3 have a standard deviation of 0.
  - B. The numbers 3, 4, 5 have the same standard deviation as 1003, 1004, 1005.
  - C. The standard deviation is a measure of spread around the centre of the data.
  - D. The numbers 1, 5, 9 have a smaller standard deviation than 101, 105, 109.
  - E. The standard deviation can only be computed for interval or ratio scaled data.
2. The coefficient of variation of two series, A & B are 60% and 80%. Which one of is true about series A and series B?
  - A. Series A is more scattered than series B.
  - B. Series A is less scattered than series B.
  - C. The values in series B are more unlike.
  - D. The variation in both series is the same.
  - E. B and C
3. Suppose that  $P(A/B) = 0.3$  and  $P(B) = 0.4$ ; then  $P(A' \cap B)$  is
  - A. 0.12
  - B. 0.28
  - C. 0.4
  - D. 0.5
  - E. None of the above
4. Suppose that for two events T and S: If  $P(T) = 0.69$ ,  $P(S) = 0.5$ ,  $P(S/T) = 0.5$ , then the two S and T are:
  - A. mutually exclusive
  - B. Independent
  - C. Mutually exclusive and independent
  - D. Neither mutually exclusive nor independent
  - E. None of the above
5. Which of the following is NOT considered as a characteristic for a Binomial Distribution?
  - A. A constant probability of success
  - B. Only two possible outcomes
  - C. Equally likely outcomes
  - D. None of the above
  - E. All of the above
6. Which of the following are characteristics of the normal distribution?
  - A. It is symmetric distribution
  - B. It is a bell-shaped distribution
  - C. The curve of a normal distribution is an asymptotic to the X-axis
  - D. All of the above
  - E. All except C
7. In which of the following distributions, the probability of success is usually small?
  - A. Binomial
  - B. Poisson
  - C. Random variable
  - D. Normal
  - E. All of the above

$$CV = \frac{S}{\bar{x}} \times 100\%$$

$$\frac{1}{6} = 16$$

13



5

Logic 1<sup>st</sup> test (17%) Time Allotted: 40 minutes

III CIVILICS Dep't Comp. Eng.

Time limit: 70 minutes

INSTRUCTIONS: The following selections relate to distinguishing arguments from non-arguments and identifying conclusions. Select the best answer for each. (1 point each)

1. ~~C~~ The Chinese 'economization of politics' is benefiting the authoritarian and semi-autoritarian states of Africa. This economic diplomacy of China asserts that the country has no intention to intervene in the internal politics of other states. Those African states having poor records on human and democratic rights seem to have making a divorce from the western counterparts. These countries opened their doors wide for numerous Chinese companies for intensive investments. In return, they allowed a huge amount of loans and aid from the Chinese government.
- A. Argument: They allowed .... Chinese government.
  - B. Argument: Those African states having .... western counterparts.
  - C. Argument: The Chinese 'economization of .... States of Africa.
  - D. Argument: Those countries opened ... intensive investments.
  - E. Non-Argument.
2. ~~E~~ Philosophy is an oldest and infamous discipline. Some regards as it's a queen of all sciences. It's devoted to the systematic examination of basic concepts such as truth, existence, reality, causality, and freedom. Metaphysics, sometimes called philosophy of being, is one among the many branches of philosophy concerns about examining the nature of being, existence, time and space, and causality.
- A. Argument: Metaphysics, sometimes .... and causality.
  - B. Argument: Some regards... all sciences.
  - C. Argument: Philosophy ... discipline.
  - D. Argument: It's devoted... and freedom.
  - E. Non-Argument
3. ~~C~~ the deliberate negligence of social sciences inevitably causes a detrimental effect in the long run. It's the prime task of the social sciences to study about the people in a society and how they relate to one another and to the group which they belong. There is a unanimous conviction about the production of a holistic personality is the ultimate goal of any education policy. However, a holistic personality is hard to meet in a situation where hard sciences and technology are favored at the expense of the social sciences.
- A. Argument: However, a holistic....the social sciences.
  - B. Argument: The deliberate....in the long run.
  - C. Argument: There is a unanimous...education policy.
  - D. Argument: It's the prime ...they belong.
  - E. Non-Argument

- a) the mean and the standard deviation of the population;  
b) the sampling distribution of the mean  
c) the mean of the sampling distribution of the sample mean ( $\bar{X}$ );  
d) the standard deviation of the sampling distribution of the sample mean ( $\bar{X}$ );  
e) sketch the graph of the sampling distribution of  $\bar{X}$ .
17. Describe briefly the difference between:
- a) Census and sample survey
  - b) Sampling and non sampling error
  - c) Sampling frame and sampling units
  - d) Probability and non probability sampling
  - e) Simple random sampling and stratified random sampling
  - f) Cluster sampling and stratified random sampling
18. Why do researchers usually select sample elements from a given population?
19. An insurance company has insured 300,000 cars over the last six years. The company would like to know the number of cars involved in one or more accidents over this period. The manager selected 1000 cars from the files and made a record of cars that were involved in one or more accidents.
- a. What is the population?
  - b. What is the sample?
  - c. What is the sampling unit?
  - d. What would be the sampling frame?
  - e. What is the variable of interest to the insurance company?
20. Suppose that in a certain population of drug addicts, the mean duration of abuse is 5 years and the standard deviation is 3 years. What is the probability that a random sample of 36 people from this population will yield a mean duration of abuse between 4 and 6 years?
21. Suppose that a random sample of  $n = 25$  observations is selected from a population which is normally distributed with mean 1 and standard deviation 0.25. Calculate the standard error of the sample mean.

ADDIS ABABA SCIENCE AND TECHNOLOGY UNIVERSITY

COLLEGE OF NATURAL AND SOCIAL SCIENCE

DEPARTMENT OF STATISTICS

PROBABILITY AND STATISTICS ASSIGNMENT Max Mark 15%

---

## Instructions:

- Write legibly; unreadable answer will not be marked!
  - Show all the necessary steps!
- 

1. Four men and three women are to be seated at a lunch counter that has only five stools.

- a) In how many ways can these people be arranged at the counter?
- b) In how many ways can they be arranged at the counter if all the women are to be seated?
- c) In how many ways can they be arranged at the counter if all the women are to be seated and if men occupy the first and last stool?
- d) If customers take seats at random, what is the probability that all of the men are seated and that a woman occupies the middle stool?

2. In how many ways can a committee of three be chosen from 4 married couples if

- (a) all are equally eligible?
- (b) one particular man must be on the committee?
- (c) a husband and a wife cannot be in the same committee?

3. If the permutation of the word WHITE is selected at random, find the probability that the permutation.

- a) begins with a consonant,
- b) ends with a vowel,
- c) has a consonant and vowels alternating

4. A lot consists of 20 defective and 80 non-defective items from which two items are chosen without replacement. Events A & B are defined as A = the first item chosen is defective, B = the second item chosen is defective.
- What is the probability that both items are defective?
  - What is the probability that the second item is defective?
5. Let A and B be two events associated with an experiment and suppose that  $P(A) = 0.4$  while  $P(A \cap B) = 0.7$ . Let  $P(B) = P$
- For what choice of P are A and B mutually exclusive?
  - For what choice of P are A and B independent?
6. A medical research suggest that 20% of the general population suffers adverse side effects from a new drug. If a doctor prescribes the drug for 4 patients, what is the probability that:
- None will have side effects?
  - All will have side effects?
  - exactly 3 will have side effects?
  - Find the expected number of patients that will have side effects.
7. The following table is obtained from the personnel department of a company. If a

Age	Bachelor's degree only	Master's degree
Under 30	90	10
30 to 40	20	30
Over 40	40	10

staff is selected at random from the company, find:

- the probability that s/he has only a bachelor's degree;
  - the probability that s/he has a master's degree, given that s/he is over 40;
  - the probability that s/he is under 30, given that s/he has a bachelor's degree.
8. A chemical company currently has in stock 100 lb of a certain chemical, which it sells to customers in 5 lb packages. Let X be the r.v. denoting the number of packages ordered by a randomly chosen customer, and suppose that the p.d.f. of X is given by:

x	1	2	3	4
$P(X=x)$	$2k$	$4k$	$3k$	$k$

- a) Determine the value of  $k$ .
  - b) Compute the following quantities:  $E(X)$ ,  $E(X^2)$ , and  $\text{Var}(X)$ .
  - c) Compute the expected number of pounds left after the order of the customer in question has been shipped, as well as the s.d. of the number of pounds around the expected value.
9. For the r.v.  $X$  with p.d.f.  $f(x) = ke^{-3x}$ , for  $0 < x < \infty$ :
- a) Determine the value of  $k$ .
  - b) Compute  $P(1 < X < 2)$
  - c) Calculate  $E(X)$  and  $\text{Var}(X)$ .
10. A secretary makes 2 errors per page on the average. What is the probability that on the next page she makes
- a) 4 or more errors?
  - b) no errors at all?
11. All the pumps at a water treatment plant have been made to the same specifications by a single manufacturer. From tests made over 4-week period, it has been ascertained that there are on average two breakdowns during each period. A new plant manager assumes that the problem is not serious if there are no more than four breakdowns over a period of 4 weeks. What is the probability of such an occurrence?
12. From past data, an engineer has estimated a probability of  $p = 0.01$  that timber delivered at a construction site from a particular source is below specification. If 150 joists of timber are necessary for a particular construction job, determine the minimum number which should be ordered so that the chance of not having the required number of suitable joists is less than 10%.
13. The compressive strengths of concrete have an estimated mean of  $60.14 \text{ N/mm}^2$  and a standard deviation of  $5.02 \text{ N/mm}^2$  and are assumed to be normally distributed. What is the probability that in ten random tests the compressive strength will be in the range  $45\text{-}75 \text{ N/mm}^2$ ?
14. The scores for a statistics course seems to be normally distributed with mean 70 & standard deviation of 8. If the instructor wishes to give an A grade to 10% of the students, what should be the dividing line between an A grade and a B grade?
15. A normal distribution has mean  $\mu = 62.4$ . Find its standard deviation if 20% of the area under the curve lies to the right of 79.2.
16. A Population consists of five numbers 3, 4, 7, 9, and 12. Consider all possible samples of size 2 that can be drawn with replacement from the population and find:

- a) the mean and the standard deviation of the population;
- b) the sampling distribution of the mean
- c) the mean of the sampling distribution of the sample mean ( $\bar{X}$ );
- d) the standard deviation of the sampling distribution of the sample mean ( $\bar{X}$ );
- e) sketch the graph of the sampling distribution of  $\bar{X}$ .

17. Describe briefly the difference between:

- a) Census and sample survey
- b) Sampling and non sampling error
- c) Sampling frame and sampling units
- d) Probability and non probability sampling.
- d) Simple random sampling and stratified random sampling
- e) Cluster sampling and stratified random sampling

18. Why do researchers usually select sample elements from a given population?

19. An insurance company has insured 300,000 cars over the last six years. The company would like to know the number of cars involved in one or more accidents over this period. The manager selected 1000 cars from the files and made a record of cars that were involved in one or more accidents.

- a. What is the population?
- b. What is the sample?
- c) What is the sampling unit?
- d) What would be the sampling frame?
- c. What is the variable of interest to the insurance company?

20. Suppose that in a certain population of drug addicts, the mean duration of abuse is 5 years and the standard deviation is 3 years. What is the probability that a random sample of 36 people from this population will yield a mean duration of abuse between 4 and 6 years?

21. Suppose that a random sample of  $n = 25$  observations is selected from a population which is normally distributed with mean 1 and standard deviation 0.25. Calculate the standard error of the sample mean.

"Which would be worse - to  
live as a monster or to die  
as a good man?" DagmawiA