

DILLA UNIVERSITY
COLLEGE OF NATURAL AND COMPUTATIONAL SCIENCES
DEPARTMENT OF STATISTICS
WORK SHEET

1. Classify the following sentences as belonging to the area of descriptive statistics or inferential statistics.
 - a) As a result of recent cutbacks by oil-producing nations, we can expect the price of gasoline to double in the next year.
 - b) At least 5% of all killings reported last year in city X were due to tourists.
 - c) Of all patients who received this particular type of drug at a clinic Y, 75% later developed significant side effect.
 - d) Adane concludes that his chance of passing the first year this academic year is at least 80% based on the statistics that 75% of the freshmen passed last year.
2. Classify each of the following first as qualitative or quantitative and second as nominal (categorical) ordinal, interval, or ratio measure.
 - a) Times for swimmers to complete a 50-meter race
 - b) Months of the year Meskerm, Tikimit...
 - c) Socioeconomic status of a family when classified as low, middle and upper classes.
 - d) Blood type of individuals, A, B, AB and O.
 - e) Pollen counts provided as numbers between 1 and 10 where 1 implies there is almost no pollen and 10 that it is rampant, but for which the values do not represent an actual counts of grains of pollen.
 - f) Regions numbers of Ethiopia (1, 2, 3 etc.)
 - g) The number of students in a college;
 - h) The net wages of a group of workers;
 - i) The height of the men in the same town;
3. For each of the hypothetical research situations described, identify the scale of measurement and indicate whether the data are discrete or continuous.
 - a) Number of complaint letters received by a firm's administrator
 - b) You measure the body lengths (in inches) of 10 full term infants as birth and record the following:

17.5	19.5	17.5	19	20
21	18	18	19.5	10.75
 - c) Number of tomato's on each plant in a field
 - d) Number of students taking ESLCE each year
 - e) Miles driven by city -bus drivers each day
 - f) Number of cartons of corn oil manufactured each day
 - g) Weights of newborn children at Black Lion hospital
 - h) Geologists have a "hardness scale" for identifying different rocks, called Mohs' scale. The hardest rock (diamond) has a value of 10 and will scratch all others. The second hardest will scratch all but diamond, and so on, down to a rock such as talc, with a value of one, which can be scratched by any other rock
 - i) The volumes of three different cubes are 40, 64, and 65 cubic inches.
 - j) Three different high ways are identified by their umbers:40,64,and 65
 - k) Republicans, Democrats, Independents, and others are identified on the voters' list with numbers 1, 2, 3, &4
 - l) The pages of a book are numbered from 1 through 150

- m) The winner of the Miss America was Miss California; runners-up were Miss Ohio and Miss Pennsylvania
 - n) The prices on three items were \$3.00, \$10.00, and \$12.00.
 - o) She earned three degrees: B.SC., M.SC., and PhD.
4. Find the class boundaries, class marks and class widths for the following intervals (classes).
- a) 6-13 b) (-8)-(-1) c) 0.378 – 0.677 d) 68.99 – 88.97
5. Which of the diagrams is most appropriate for each of the following data set? Draw the diagrams.
- a) Students enrolled to a certain department from year 1 to 3

<u>No. of students</u>		
Year	Male	Female
1	50	20
2	45	15
3	40	10

- b) Net profit (in million birr) earned by different companies in 1980

Company	Net profit
A	20
B	-15
C	30

6. Construct a frequency distribution of weights of miniature poodles if the class marks are 6.5, 8.5, 10.5, 12.5 and 14.5 kgs with corresponding frequencies 8, 12, 22, 17 and 3.
7. The investigator was interested in studying the marital status, which is often grouped as Single(S), Married (M), Divorced (D), and Widowed (W) of people in a certain town. The following data were obtained.
- D S D D S W S D S S D D W M M S D D D W M S S W M D D M D W D S S W D D S D S M W M D S D W D M S S D W W S S S W S D M W S S
- a. To which scale of measurement do these data belong?
 - b. Summarize the data by constructing the appropriate frequency distribution
 - c. Present the data using the appropriate Graph/Diagrams.
8. Suppose data collected for heights (in cms) of 390 cows were tabulated in a frequency distribution and the following results were obtained.

f_i : 6, 25, 48, 72, 116, 60, 38, 22, 3
 $CM_1 = 112, CM_2 = 117$ where CM_i is i^{th} class mark

Determine:

- a. the class interval size (class width)
- b. the class limits
- c. class boundaries
- d. class marks
- e. the less than cumulative frequency distribution
- f. the class intervals having the highest frequency
- g. Above which height do we find 50% of the cows?
- h. Below which height do we get 25% of the cows?

Draw

- A. histogram
- B. a frequency polygon
- C. a less than ogive for the above data

9. The following is the distribution of weights of 125 mineral specimen collected on the field trip.

Weight (Grams)	No.of Specimen
0.0-19.9	19
20.0-39.9	36
40.0-59.9	35
60.0-79.9	17
80.0-99.9	11
100-119.9	3
120-139.9	2
Total	123

- i) If possible, find the number of specimen, which weigh
 a) at most 40.0 gram b) 40.0 gram or more c) Less than 40.0 gram
 d) at least 120.0 gram e) Exactly 20.0 gram
 ii) Draw an appropriate Diagram / Graph for the distribution.

10. What are the objectives of studying averages?
 11. What do measures of central tendency represent? When is an average a typical average?
 12. If you have an open-end continuous series, which measures of central tendency do you recommend representing this series? Why?
 13. If the arithmetic mean and the geometric mean of two observations are 6.5 and 6 respectively, find the values of the observations.

14. A person purchased the following quantities of cement (kg) with rates shown at 4 stores:

Store	Quantity purchased (kg)	Rate of purchase (kg/birr)
A	15	20
B	20	16
C	25	12
D	40	10

Determine the average amount of cement the person purchased for one birr.

15. a) Determine the missing frequencies of the following distribution given that the median is 33.5 and the mode is 34.0.

Class limits	0-9	10-19	20-29	30-39	40-49	50-59	60-69	Total
Freq.	4	16	f_3	f_4	f_5	6	4	230

- b) Compute the arithmetic mean.
 c) Compute the value below which 25% of the observations lie.
 d) Compute the value above which 25% of the observations lie

16. The following data on income in the form of cumulative frequency distribution is given:

INCOME	NO.OF PERSONS
100---200	15
100---300	33
100---400	63
100---500	83
100---600	100

- Find (a) construct a complete Grouped Frequency table
 (b) Draw the appropriate graph
 (c) The mean
 (d) The median and all quartiles
 (e) The 2nd and the 8th deciles
 (f) The 40th and the 90th percentiles

17. Some characteristics of annually family income distribution (in Birr) in two regions is as follows:

Region	Mean	Median	Standard Deviation
A	6250	5100	960
B	6980	5500	940

- a) For which region is the income more consistent?
 b) If a household X from region A has an annual family income of 5800 birr and a household Y from region B has an annual family income of 6200 birr. Relatively speaking, which household earns higher annual income?

18. The number of employees in two hospitals A and B, the average wage of employees and the variances of wages are given below:

	Hospital A	Hospital B
No. of employees	50	100
Mean wage	120	85
Variance	9	6

- a) In which hospital is the variation in the distribution of wages greater?
- b) Suppose the wage of an employee in hospital B was wrongly noted as 120 instead of 250. What will be the corrected mean for hospital B?
19. a) Explain the meaning of dispersion
 b) What purpose does a measure of dispersion serve?
 c) Why are measures of central tendency not sufficient to study a frequency distribution?
 d) Distinguish between absolute and relative measures of dispersion.
20. In a final exam in a Statistics course, the mean mark of a group of 150 students is 78 and the standard deviation is 8. In Mathematics exam of the same group, the mean is 73 and the standard deviation was 7.6.
- a) In which exam was there a greater dispersion?
 b) Suppose a student scored 75 in Statistics and 71 in Mathematics. In which exam was his relative standing better?
 c) Suppose Student "A" scored 80 in Statistics and 75 in Mathematics while student "B" scored 75 in Statistics and 80 in mathematics, whose performance is better A or B ?
21. A newly married couple is planning to have three children. List the elements of the sample space
- a) using M for male and F for female
 b) if the sample points in the sample space represents the number of females
22. Four married couples have bought 8 seats in a row for a show. In how many different ways can they be seated
- a) If each couple is to sit together?
 b) If all the women sit together?
 c) If all the women sit together to the right of all the men?
23. If a multiple choice test consists of 4 questions each with 4 possible answers of which one is correct
- a) In how many different ways can a student check off one answer to each question?
 b) In how many different ways can a student check off one answer for each question and get all the questions wrong?
24. Four men and three women are to be seated at a lunch counter that has only five stools.
- a) In how many ways can the customers 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?
25. 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) Husband and wife can not serve in the same committee?
26. One bag contains 4 white balls and 3 black balls, and the second bag contains 3 white balls and 5 black balls. One ball is drawn at random from the second bag and placed unseen in the first bag. What is the probability that a ball now drawn from the first bag is white?

27. In a large graduating class of 100 students 54 studied mathematics, 69 studied library science, and 35 studied both mathematics and library science. If one of these student is selected at random, find the probability that
- The student takes mathematics or library science
 - The student does not take either of these subjects
 - The student takes library science but not mathematics
28. Let A and B be two events associated with an experiment and suppose that $P(A)=0.4$ while $P(A\cup B)=0.7$. 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?
29. The personnel department of a company has records which show the following analysis of its 200 accountants.
- | Age | Bachelor's degree only | Master's degree |
|----------|------------------------|-----------------|
| Under 30 | 90 | 10 |
| 30 to 40 | 20 | 30 |
| Over 40 | 40 | 10 |
- If one accountant is selected at random from the company, find
- The probability he has only a bachelor's degree
 - The probability he has a master's degree, given that he is over 40
 - The probability he is under 30, given that he has a bachelor's degree
30. Two dice are rolled. Let X be a random variable denoting the sum of the numbers on the two dice.
- Give the probability distribution of X
 - Compute the expected value of X and its variance
31. State the conditions (assumptions) under which random variable can have a binomial distribution.
32. Let the probability that a freshman entering DU (Science Faculty) will survive first semester is 0.92. Assuming this pattern remain unchanged over the subsequent years, what is the probability that among 100 randomly selected freshmen in first semester,
- None will survive?
 - Exactly 97 will survive?
 - At least three will survive?
33. A multiple-choice test consists of six questions and three choices to each question (of which only one is correct). If a student answers each question by rolling a balanced die and checking the first answer if he gets a 1 or a 2, the second answer if he gets a 3 or a 4, and the third answer if he gets a 5 or a 6, find the probabilities of getting
- Exactly three answers
 - No correct answers
 - At least five correct answers.
34. A secretary makes 2 errors per page on average. What is the probability that on the next page she makes
- 4 or more errors?
 - No errors?
35. It is known that bacteria of a certain kind occur in water at a rate 2 bacteria per cubic centimeter of water. Assuming that the phenomenon obeys a Poisson probability law, what is the probability that a sample of two cube centimeters of water will contain
- No bacteria?
 - At least one bacterias?
36. Women's shoes are manufactured in sizes 2, 3, 4, ..., 8. Size 5 is suitable for a foot of length ranging from 9.25 inches to 9.5 inches. If length of women's foot are normally distributed with mean 9.4 inches and standard deviation 0.25 inches, how many pairs of size 5 are required out of every 10,000 pairs manufactured?(Hint: Use expected value)

37. The scores for a mathematics 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?
38. A manufacturer of sports equipment has developed a new synthetic fishing line that he claims has a mean breaking strength of 8 kilograms with standard deviation of 0.5 kilograms. Test the hypothesis that the mean breaking strength is greater if a random sample of 49 liners is tested and found to have a mean breaking strength of 8.2 kilograms. Use the 0.05 level of significance.
39. A law student, who wants to check a professor’s claim that convicted embezzlers spend on the average 12.3 months in jail, takes a random sample of 16 such cases and obtains an average of 11.5 months with a standard deviation of 3.8 months. A) Construct a 99% CI for the mean months spent in jail by embezzlers? B) Can the sample result be taken as evidence that the professor’s claim was an exaggeration?
40. The weight of adults in a certain city has a mean of 130 lbs, with standard deviation of 41 lbs. A sample of 25 adults living in one of the *kebeles* of the city is found to have a mean weight of 129 lbs. Does this mean the residents of the *kebele* weigh significantly less on the basis of the 1% level of significance? (Assume normal population.)
41. A major truck has kept extensive records on various transactions with its customers. If a random sample of 16 of these records shows average sales of 290 liters of diesel fuel with a standard deviation of 12 liters, construct a 95% confidence interval for the mean of the population sampled.
42. The following is the results of a sample survey regarding the employment of adult males and females.

	Sex	
	Male	Female
Employed	160	440
Not employed	40	360

Does this survey reveal any association b/n the sex of an individual and his/her employment? Use $\alpha = 0.05$.

43. Results of a sample poll on television ownership were as follows:

	Men	Women
Want TV	8	17
Do not want TV	12	13

Is the drive to win a television set unaffected by the sex of the respondent? (Take $\alpha = 0.10$)

44. The following data were collected from a certain household on the monthly income (X) and consumption (Y) for the past 10 months.

X: 650 654 720 456 536 853 735 650 536 666
 Y: 450 523 235 398 500 632 500 635 450 360

- a) Fit a linear regression equation and comment on the result.
 - b) Calculate the correlation coefficient and comment on it.
 - c) Estimate the amount of consumption for monthly income amount of 900.
45. It has been observed that the amount of soil eroded (in Kg) per day (Y) is determined by the wind velocity (X) in that day (Km/Hr). Data obtained from the ministry of agriculture for a certain area gave the ff summary statistics. $\Sigma Y = 80.9$, $\Sigma X^2 = 117,123.86$, $\Sigma Y^2 = 412.81$
 Regression line: $Y = -4.54 + 0.1123X$
 - a. What change in amount of soil erosion would be associated with 1 Km/Hr change in wind velocity?
 - b. What amount of soil erosion would you predict for a wind velocity of 90 Km/Hr?

46. Ten accounting majors were recently interviewed by the management of a firm. The following represent the personnel director's ranking of the candidates and his assessment of the quality of the candidate's college performance (rank one is best).

Candidate	Interview Rank	Rank of College
1	10	6
2	3	7.5
3	2	3
4	9	9
5	5	1
6	6	7.5
7	1	2
8	4	4.5
9	8	10
10	7	4.5

Compute the rank correlation coefficient and comment on the value.