**Haramaya University**

**College of Education and Behavioural Sciences**

**Department of Psychology**

**Prepared by**

**Gemechu Abera (Assistant Professor)**

**Working sheet Statistical Methods in Education**

**Part I:** The chart shows the number of job-related injuries for each of the transportation industries in Ethiopia for 2018 (individual-8%)

|  |  |  |
| --- | --- | --- |
| No | Industry | Number of injuries |
| 1 | Railroad | 5520 |
| 2 | Intercity bus | 6100 |
| 3 | Subway | 6850 |
| 4 | Trucking | 7144 |
| 5 | Airline | 2000 |

a. What are the variables under study?

b. Categorize each variable as quantitative or qualitative.

c. Categorize each quantitative variable as discrete or continuous.

d. Identify the level of measurement for each variable.

e. The Airline is shown as the safest transportation industry. Does that mean Airline have fewer accidents than the other industries? Explain.

f. What factors other than safety inﬂuence a person’s choice of transportation?

g. From the information given, comment on the relationship between the variables.

h. A computerized data file consists of several columns that correspond to values for variables. If

one such column represents sex and consists of zeros and ones, where 0 represents a female and 1 represents a male, what scale of measurement is represented?

**2**. **Given below are the marks obtained a batch of 78 students at the Haramaya University, College of Education and Behavioural Sciences, CEBS, Statistical Methods in Education final exam out of 60%in 2018?**

34 43 32 57 35 56 54 10 52 19 48 17 24

43 56 40 54 54 44 50 13 18 49 57 21 55

57 32 21 52 40 35 57 43 45 44 55 39 19

14 45 17 51 35 27 47 22 14 22 15 14 23

35 14 31 21 52 48 10 22 12 12 15 40 39

30 42 27 17 44 19 10 30 16 19 31 25 33

i. Construct a frequency distribution with appropriate class intervals in both discrete and continuous variables (5%)- **individual task.**

ii. Table 2.7 gives the grades on a quiz for a class of 40 students. (a) Arrange these grades (raw

data set) into an array from the lowest grade to the highest grade. (b) Construct a table showing

class intervals and class midpoints and the absolute, relative, and cumulative frequencies for each grade. (c) Present the data in the form of a histogram, relative-frequency histogram, frequency polygon, and ogive. (10%)- ***Individual Task***

Table 2.7: Grades on a Quiz for a Class of 40 Students

75 62 87 67 39 10 45 54 67 82

35 67 98 24 79 46 78 36 79 78

**Part II: Activities on Central Tendencies**

**Nb. The following questions(1,3,5,7) should be an individual work whereas questions (2,4,6,8) are group works (10%) for both cases,**

1. Two items 60 and 70 were left out at a time of calculating the arithmetic mean of 98 items. The arithmetic mean was 50.What would be the correct arithmetic mean of these items?
2. If the mean of x, x-3, 10, x+3 and x-5 is 15, find x
3. The mean of 20 observations is 85 but it was later found that the two of the observation were wrongly read as 75 and 70 instead of 57 and 60. Find the correct mean,
4. The mean weight of 150 students in certain class is 60kg. The mean weight of boys in the class is 66kg and that of girls is 55kg. Find the number of boys and girls in the class.
5. The mean age of the combined group of men and women is 30 yrs. If the mean age of the group men is 32 and that of women is 27, find the percentage of the men and women in the group.
6. Worku Debebe last year scored “A” in a four credit hours course, a “B” in another four credit hours course, an “A” in a three credit hours course and a “B” in another three credit hrs course, a “C” in a two credit hrs course and a “B” in another two credit hrs course. Compute his GPA of the semester.
7. A man climbs up a slope at a speed of 5km/hr and descends it at a speed of 8km/hr. If a distance covered is 10km, find the average speed for the entire journey?
8. The population of India in 1951 and 1961 were 361 and 439 million, respectively

i. what was the average increase per annum during these period?

ii. If the average rate of increase of 1961 to 1971 remains the same, what would be the population in 1971?

**Part III Activities on Variability (All are group works -10% -three marks each)**

**Show all the necessary step clearly and legibly**

1. The arithmetic mean and standard deviation of a set of 9 items are 43 and 5. If an item of value 63 is added to the set, find mean and standard deviation of 10 items.
2. The mean of 5 observations is 4.4 and their variance is 8.24. If three of the observations 1, 2 and 6, find the values of the other two.
3. The arithmetic mean and standard deviation of a series of 20 items were calculated by a student as 20cm and 5cm respectively, but while calculating them an item 13 was misread as 30. Find the correct arithmetic mean and standard deviation.
4. For group containing 100 observations, the arithmetic mean and standard deviation respectively. For 50 observation selected from these 100 observations, the mean and standard deviation are 10 and 2 respectively. Calculate values of the mean and standard deviation for other half.
5. Find the missing information from the following

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Group1 | Group2 | Group3 | Combined |
| Number | 50 | x | 90 | 200 |
| SD | 6 | 7 | a | 7.776 |
| Mean | 113 | y | 115 | 116 |

**Part Iv: Measures of Relationship (Group Work -25%)**

1. A study on the relation between rejection and depression in adolescents conducted by one of the authors (Nolan, Flynn, & Garber, 2003) also collected data on externalizing behaviors (e.g., acting out in negative ways, such as causing fights) and anxiety. We wondered whether externalizing behaviors were related to feelings of anxiety. Some of the data are presented in the accompanying table.

Externalizing Anxiety Externalizing Anxiety

9 37 6 33

7 23 2 26

7 26 6 35

3 21 6 23

11 42 9 28

a. Create a scatter plot of these data.

b. Calculate the partial correlation between these variables?

c. Is there statistically a significant relationship among them at α = 0.05?

d. What is the t-value of this correlation coefficient?

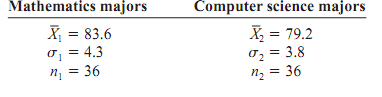
e. Would it be appropriate to calculate a Pearson correlation coefficient? Explain your answer.

f. Calculate the coefficient of Correlation and interpreter what it explains for you,

2. As an aid for improving students’ study habits, nine students were randomly selected to attend a seminar on the importance of education in life. The table shows the number of hours each student studied per week before and after the seminar. At α = 0.05



Based on this information, calculate the

1. Standard deviation
2. Standard error of the distribution
3. The 95% confidential interval
4. Did attending the seminar increase the number of hours the students studied per week?
5. Two groups of students are given a problem-solving test, and the results are compared. 

Based on this data, find

1. The 95% conﬁdence interval of the true difference in means.
2. What kind test is it?
3. Is there statistically a significant mean difference between the two cases?

4. using the following data, find the lines of regression equation and from them compute the Karl Pearson’s coefficients of correlations.

∑xi = 250, ∑yi = 300, ∑xiyi = 7,900, ∑xi2= 6,500 , ∑yi = 10,000, n = 10

5. A pharmaceutical company has developed a drug that is expected to reduce hunger. To test the drug, three samples of rats are selected, with n = 10 in each sample. The first sample receives the drug every day. The second sample is given the drug once a week and the third sample receives no drug at all. The dependent variable is the amount of food eaten by each rat over a 1-month period. These data are analyzed by an analysis of variance, and the results are reported in the following summary table.

a. Fill in all missing values in the table

b. Compute η2 to measure the size of the effect (the percentage of variance accounted for).

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Source of Variation (SV) | SS | df | MS | F |
| Between treatments |  |  |  | 12 |
| Within treatments | 54 |  |  |  |
| Total |  |  |  |  |

1. A book publisher wishes to determine whether there is a difference in the type of book selected by males and females for recreational reading. A random sample provides the data given here. At α = 0.05, test the claim that the type of book selected is independent of the gender of the individual. Use the P-value method.

