بسم الله الرحمن الرحيم

Biostatistics

Lecture one

Doctor:

Nedal Nawaiseh

Done by Doctor:

Abdullah daradkh

Mahmoud Otoom

(Second part L1)

Note :the single of data= datum

simple frequency distribution table

- **Qualitative Variable**
- **The left column is the variable, the right one is the frequency(number of observation for each category)**
- **Have a total at the bottom**
- **Every table should have a title which answers three questions (what, who and when).**
- **Source should be written below the table**
- Ligand (key) : includes a legend key for referencing the data. each legend key is a name identifying the data represented by the specific key
- **Example**

Distribution of the third year medical students according to their degree of success in may 2007 in Mutah university faculty of medicine

Degree of success	No of students
Excellent	120
Very good	200
Good	350
fair	180
total	850

Source: records of the faculty of medicine 2007

Quantitative variable (continuous or discrete)

1. **The smallest observation is chosen as the lower limits of the intervals.**
2. **The width or the size of the interval is 5,10 or 15**
3. **The number of interval counted to be in the permissible range of 4-12. to achieve the main advantage of tabular presentation namely summarization of a large mass of data.**
4. **The main disadvantage of tabular presentation is loss of precision in the presentation.**
5. **If less than 4 (loss of precision)we can reconstruct the table using a width less than 5**
6. **If > 12 (result in missing the summarization value of the table) we can reconstruct the table using a width > than 5**

Duplication of the limits of intervals should be avoided by one of the 4 methods

A (cont. & discrete)	B (discrete)	C (cont. &discrete)	D (cont.)
10 to less than 15	10-14	10-	10-14.9
15 to less than 20	15-19	15-	15-19.9
20 to less than 25	20-24	20	20-24.9
25 to less than 30	25-29	25-30	25-29.9

1. **Forms A, and C can be used in both continuous and discrete quantitative variables.**
2. **Form B can be used in discrete quantitative variables**
3. **While form D is only used in continuous quantitative variables.**

Example for continuous variable:

Distribution of 100 medical student (1985) according to their Hb %

Hb %	No of patients
65-	22
70-	15
75-	19
80-	11
85-	17
90-	12
95<100	4
total	100

Each Interval In Table Has The Following.

1. Lower limit
2. Upper limit
3. Width
4. Midpoint
5. Number of observation

First Interval:
Lower limit (L.L)= 65
Upper limit (U.L)= 69.999, practically= 70
Width(continuous quantitative variables)=
U.L - L.L= 70 – 65 = 5

Mid point = $\frac{\text{L.L + U.L}}{2} = \frac{65 + 70}{2} = 67.5$

Number of observations = 22

*Example for discrete quantitative variable:

Distribution of patients in hospital

(X) according to their family size in 1984.

Family size	No. of patients
3-4	28
5-6	55
7-8	32
9-10	18
11-12	12
Total	145

Last interval:

L.L =11

U.L =12

Width (discrete quantitative variables) = (U.L – L.L + 1) =

(12 – 11 + 1) = 2 (No. of counts within the interval e.g.

(11 & 12 = 2)

Midpoint = $\frac{11+12}{2}$ = 11.5

Number of observation = 12

N.B: in construction of the table, we should avoid open ended table I.e. lower limit of first interval and/or upper limit of last interval are missed or unknown

Table Of An Association Or Contingency Table.

A) Two by two table:

i.e Two columns by two rows, it is used to show relation between a condition and characteristic e.g relation between smoking and lung cancer

Example: relation between smoking and lung cancer

Smoking Status	Lung cancer		Total
	Yes	No	
Smoker	35	15	50
None smoker	65	85	150
Total	100	100	200

19

B) c x r table: i.e. more than two columns by two or more than two rows

Example: Different types of treatment of disease (X) and outcome.

Treatment	outcome			Total
	Cured	Improved	died	
A	35	12	3	50
B	30	10	10	50
C	33	12	5	50
Total	98	34	18	150

20

Comparing Frequency Distribution Table

i.e distribution of two different groups according to one variable.

N.B : For comparison, the total groups should have the same total frequencies otherwise calculate the percent of total for each frequency.

Examples: distribution of low birth weight (L.B.Wt) babies and normal babies according to their mother's age.

Age of mother (years)	L.B.Wt	Normal
20-	22	30
25-	13	39
30-	15	18
35-	20	8
40<45	30	5
Total	100	100

Two Way Table Or Two Way Classification

i.e One group is classified according to two variables e.g weight and height or age and blood pressure to find correlation between these two variables.

Example: Distribution study by weight and height

Weight (KG)	Height (cm)					Total
	150-	155-	160-	165-	170<175	
60-	18	5	2	-	-	25
65-	15	8	7	2	-	32
70-	11	15	8	2	2	38
75-	7	15	10	1	3	40
80<85	-	2	5	2	7	15
Total	**51**	**45**	**32**	**10**	**12**	**150**