Pharmaceutical statistics

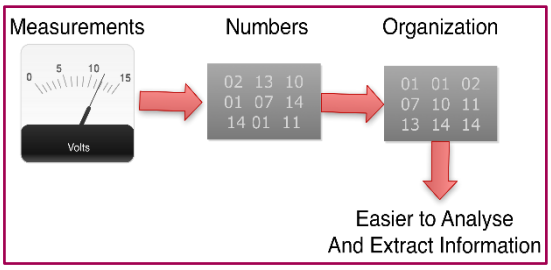
2025-2024

Dr. Rola Aqeel

ARKAN ACADEMY

Graphical Presentation of Data

- ***Statistics***: collection, organization, summarization and analysis of data, and the drawing of inferences about a body of data when only a part of the data is observed.
  - ***Raw data:*** measurements that have not been organized, summarized or otherwise manipulated.
  - Unless the number of observations is extremely small, it will be unlikely that these raw data will impart much information until they have been put into some kind of order.
  - It is *always easier to analyze organized data*



- ***The ordered array:***
  - The preparation of the ordered array is the ***first step*** in organizing data.
  - It is a listing of the values of a collection (either population or sample) ***from the smallest value to the largest value.***
  - The ordered array enables one to determine quickly the value of the smallest measurement, the value of the largest measurement and the general trends in the data.

Raw data	13	3	17	9	5	7	15	11
Organized	3	5	7	9	11	13	15	17

- ***Grouped data:***
  - *The Frequency Distribution*
    - ✓ Although a set of observations can be made more comprehensible and meaningful by means of an ordered array, further useful ***summarization*** may be achieved by ***grouping*** the data.
    - ✓ To group a set of observations, we select a set of non-overlapping intervals such that (each value in the data set of observations can be placed in one, ***and only one***, interval.
    - ✓ These intervals are usually referred to as Class Intervals
    - ✓ Usually, class intervals are ordered from smallest to largest.
    - ✓ ***Interval width*** = ***(UL-LL+1)*** OR ***(UL_2 -UL_1)*** OR ***(LL_2 -LL_1)***
      - In this table the Interval width = 19-10+1=10 **or** 20-10=10
    - ✓ ***Mid interval*** is the best value that represents each interval, and it Equals: $\left(\frac{UL+LL}{2}\right)$

      Mid interval = $\left(\frac{10+19}{2}\right)$ = ***14.5***

Class Interval	Frequency
10-19	4
20-29	66
30-39	47
40-49	36
50-59	12
60-69	4
Total	169

  - ***Construction of Frequency Distribution Table***
    - ✓ How many intervals should we use? (0-100 years)
      - Too few intervals are undesirable because of the resulting loss of information. (0-50, 51-100) two intervals.
      - Too many intervals, on the other hand, will not meet the objective of summarization. (0-1, 2-3,4-5, …….99-100)!!
      - A commonly used rule is there should be ***no fewer than six intervals and no more than 15***. (6-15 is optimal)

✓ To know how many class intervals to employ we use ***Sturge's rule***

- ***k = 1 + 3.322 * (log10 n)*** ⟶ *Rounded to nearest integer*
  k: Number of intervals, ***n*** is number of observations
- The rule is just used as guidance and should not ***be applied strictly***.
- The size of the class interval is often selected as 5, 10, 15 or 20 etc.

✓ ***The width of class intervals:***

- Class intervals should be generally of the same width.
- The width may be obtained by dividing the range by k, the number of class intervals.
- ***Interval width = Range/k***

★ Example 1:

Tablet hardness values range between 50 and 120 N, calculate the recommended number of intervals and the interval width for data contains **60 values** of tablet hardness?? **n= 60**

- **Range** = largest – smallest = 120-50 = **70**
- # intervals=1+3.329 (log n) =1+3.329 (log 60) = **7**
- Interval width=k/n=70/7=**10**

★ Example 2:

Frequency distribution *non-overlapping* of ages of 169 subjects.

non-overlapping **intervals**

How many subjects are there in each class interval?

Variables range = 69-10 + 1 = 60

Interval width

Class Interval	Frequency
10-19	4
20-29	66
30-39	47
40-49	36
50-59	12
60-69	4
Total	169

- ***The relative frequency distribution***

➢ It may be useful sometimes to know the ***proportion*** rather than the number of values falling between a particular class interval.

➢ We obtain this information by dividing the number of values in the particular class interval by the ***total*** number of values.

➢ We refer to the proportion of values falling within a class interval as the *relative frequency* of values in that interval (we call it probability distribution).

✓ The total of relative frequency for all intervals ***must be 1***

Class Interval	Frequency	Relative Frequency	Cumulative Frequency	Cumulative Relative Frequency
10-19	4	0.0237	4	0.0237
20-29	66	0.3905	70	0.4142
30-39	47	0.2781	117	0.6923
40-49	36	0.2130	153	0.9053
50-59	12	0.0710	165	0.9763
60-69	4	0.0237	169	1.0000
Total	169	1.0000		

- We may sum (***cumulate***) the frequencies and relative frequencies to facilitate obtaining information regarding frequency or relative frequency of values within two or more contiguous class intervals.
- ***Cumulative frequency***: number of subjects with values <u>below</u> the UL.
- <u>The total of cumulative relative frequency for all intervals</u> ***must be 1***
- The percentage of each interval = % fr = ***relative frequency* 100%***

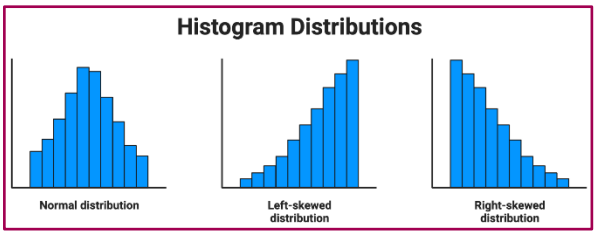
✓ We use ***true limits*** to fill the gaps between intervals for a continuous variable.
- Using true limits is very essential to calculate statistics (range, median, etc.) of grouped data.
- ***Upper true limit*** = upper class value + 0.5
- ***Lower true limit*** = lower class value - 0.5

Intervals	True limits	frequency
10--19	9.5-19.5	4
20-29	19.5-29.5	66
30-39	29.5-39.5	47
40-49	39.5-49.5	36
50-59	49.5-59.5	12
60-69	59.5-69.5	4

- ***Histogram:***

➢ We may display a frequency distribution (or a relative frequency distribution) graphically in the form of a histogram, which is a special type of bar graph.

➢ When we construct a histogram, the ***variables*** under consideration are represented by the horizontal ***(x) axis***, while the ***frequency*** (or relative frequency) of occurrence is the ***(y) axis.***



★ Example:

A school nurse weighed 30 students in Year 10. Their weights (in kg) were recorded as follows:

50	52	53	54	55	65	60	70	48	63
74	**40**	46	59	68	44	47	56	49	58
63	66	68	61	57	58	62	52	56	58

1. Use the data above to construct a frequency table.
   - ✓ **Range** = 74-40=34
   - ✓ Let **width** of class interval = 5
   - ✓ **# intervals** =34/5=7 There are 7 class intervals
2. Complete the table to calculate:
   - cumulative frequency
   - relative frequencies
   - cumulative relative frequencies

Class interval x(weight in kg)	Tally	Frequency f
40 - 44	II	2
45 - 49	IIII	4
50 - 54	~~IIII~~	5
55 - 59	~~IIII~~ III	8
60 - 64	~~IIII~~	5
65 - 69	IIII	4
70 - 74	II	2
		30

علم في كل مكان

Arkan academy
Arkanacademy
www.arkan-academy.com
+962 790408805