lecture: 1 / Page: 3 / Section: 1-1 / /

Statistics: is the science of conducting studies to collect, organize, summarize, analyze, and draw conclutions from data.

Variable: is a characteristic or attribute that can assume different values.

Random variable: Variables whose value are determined by chance

Data: Are the values (measurements or observations) that the variable can assume.

" / Page: 4 / "

Statistic

Descriptive

Consist of the collection, organization, Summarization, and presentation of data.

Inferential

Consist of generalizing from Samples to population, performing estimation and hypothesis test, determining relationships among variable, and making predictions.

Population: Consists of all subjects (human or otherwise) that are being studied.

Sample: A group of subjects selected from a population.

" -Page: 6 - Section: 1-2

/ /

البيانات
Data

الوصفية (النوعية)
Qualitative variable
Are variable that can be placed into distinict categories, according to some characteristic or attribute.

الكمية
Quantitative variable
Are numerical and can be ordered or ranked.

Continuous
can assum an infinite number of values between any two specific values. They are obtained by measuring. They often include fractions and decimals.

Discrete
assum values that can be counted.

lecture: 2 , Page: 7 , Section: 1-2

The nominal level of measurement: Classifies data into mutually exclusive (non over-lapping) categories in which no order or ranking can be imposed on the data.

البيانات الاسمية: هي البيانات التي لا يوجد بها ترتيب أو تسلسل.
Exp: السكن الحديني، الرقم الجامعي، الجنس، ذكر أم أنثى

" , Page: 8 , "

The Ordinal level of measurement: Classifies data into categories that can be ranked, however precise differences between the ranks don't exist.

البيانات الترتيبية: هي البيانات التي يوجد بها ترتيب أو تسلسل.
Exp: درجة الجامعة أو الكلية، الـ

« Table 1-2 »

lecture: 2 , Page: 9 , Section: 1-3

Data can be collected in variety of ways. One of the most Common methods is through the use of Surveys. Surveys can be done by using a variety of methods. Three of the most Common methods are the telephone Survey, the mailed questionnaire, and the personal interview.

" , Page: 10 , "

---

1. Random Samples: Are selected by using chance methods or random methods.

Exp: 1. وضع الأسماء في أوراق ثم وضعها في صندوق وسحب أي ورقة.
2. ترقيم الطلاب من 1-60 وجعل الحاسب يختار من بينهم

" , Page: 11 , "

---

2. Systematic Sampling: by numbering each Subject of the population and then selecting every Kth Subject.

Exp: 1. المطلوب 6 طلاب من 60، تقسيمهم مجموعات من 10 طلاب، ثم اختيار اثنين من كل مجموعة.
2. تفتيش المرور: كل 13 سيارة يوقفها.

---

lecture: 3 , Page: 12 , Section: "

---

3. Stratified Sampling: by dividing the population into groups (Called Strata), then Sampling from each group.

Exp: الاختيار داخل المجموعات (العشوائية)

4. Cluster Sampling: " divided " " " " (" Clusters), then randomly selects some of these clusters and use all members of the Selected Clusters such as the subject of the samples.

Exp: الاختيار بين المجموعات

" , Page: 13 , Section: 1-4

Study

Observation Study — الدراسة في هذه الحالة

Experimental Study — الدراسة في هذه الحالة

1. Observation Study: the researcher merely observes what is happenning or what has happend in the past and tries to draw conclusions based on these observations.

Exp: 1. الحوادث المرورية

2. قطع الإنارة

lecture: 3 , Page: 14 , Section: 1-4

2. Experimental Study: the researcher manipulates one of the variable and tries to determine how the mainpulation influences other variables.

Exp: في الكلية : نوع الكلية مستقل ، نسبة الكلوسترول تابع

1. Independent variable: Is the one that manipulated by the researcher (explanatory variable). — المستقل / المتغير المستكشف

2. The resultant variable is called the dependent variable or the Outcome variable. — المتغير التابع / الناتج أو الحاصل / المتغير

١. المتغير المستقل: هو المتغير الذي يتحكم أو يأثر به من قبل الباحث (المتغير المستكشف).

٢. " التابع: " " " يظهر به أثر التحكم في المتغير المستقل (المتغير الحاصل أو الناتج)

" , Page: 37 , Section: 2-1

When the data are in original form, they are called Raw Data. — بياناتها الأصلية / بياناتها الأولية

Each raw data value is placed into quantitative or qualitative category called a class. The frequency of a class is the number of data values contained in specific class. — كمي / وصفي / فئة

/ /

Frequency distribution: Is the organization of raw data in table form, using classes and frequencies.

lecture: 3 , Page: 38 , Section: 2-1

~~Three type of frequency distribution~~

Three type of frequency distribution:

1. Categorical frequency distributions
2. Grouped " "
3. Ungrouped " "

1. Categorical frequency distributions: Is ~~used~~ used for data that can be placed in specific categories, such as nominal- or ordinal-level data.

Exp. 2-1 (Blood Type)

Class	frequency	percentage $\frac{f}{n} \cdot 100\%$
A	5	$\frac{5}{25} \times 100 = 25\%$
B	7	
C	9	
AB	4	
Σf	25	

" , Page: 39 , "

2. Grouped frequency distributions: The data must be grouped into classes that are more than one unit in width, when the range of the data is large.

- The lower class limit: It represents the smallest data value that can be included in the class.

- The upper class limit: " " " largest " " " " " " " " " " .

Class boundaries: Used to separate the classes, So, that there are no gaps in the frequency distribution. By Subtracting -0.5 from (The lower class limit) and adding +0.5 to (The upper class limit).

Exp:

class limit	C boun...	
24 - 30	23.5 - 30.5	30.2
31 - 37	30.5 - 37.5	30.5
38 - 44	37.5 - 44.5	30.7

Class Width: Can be found by ~~two~~ methods:
one of these

1. طرح الحد الأدنى من الحد الأدنى للفئة التالية
2. " " الأعلى " " الأعلى "
3. " " الأدنى Bound من الحد الأدنى Bound للفئة التالية.
4. " " الأعلى " " " الأعلى "
5. " " الأدنى من الحد الأعلى لنفس الفئة لهما وزن = الأوزان المطلوبين.
6. " " Bound من " Bound " " " لنفس الفئة.

ملاحظة: طريقة واحدة تكفي.

lecture: 3 , Page: 40 , Section: 2-1

نقطة وسط الفئة
Class midpoint X_m:

1. Obtained by adding lower and upper boundaries and diving by 2 $X_m = \frac{LCB + UCB}{2}$

or Adding the lower and upper limits and diving by 2.

$$X_m = \frac{LC + UC}{2}$$

" , Page: 41 , "

Exp. 2-2

The range (المدى): R = highest value - lowest value = H-L

/ /

lecture: 3 , Page: 42 , Section: 2-1

Cumulative frequency distribution: التكرارات التراكمية

التكرارات التراكمية: تكرار كل فئة زائد تكرارات الفئات السابقة.

3. Ungrouped frequency distribution: When the range of the data values is relatively small. A frequency distribution can be constructed using single data values for each class.

Exp. 2-3

lecture: 3 , Page: 51 , Section: 2-2

The three most commonly used graph in research are:

1. The histogram.
2. The frequency polygon
3. The cumulative frequency graph or ogive.

1. The histograms: المدرج التكراري Is a graph that displays the data by using contiguous vertical bars (Unless the frequency of a class is 0) of various highest to represent the frequencies of the classes.

, Page: 53-54 , Section: "

2. The frequency polygon: المضلع التكراري Is a graph that displays data by using lines that connect points plotted for the frequencies at the midpoints of the classes. The frequencies are represented by the highest of the points.

Exp. 2-5

Figure. 2-3

/ /

lecture: 3 , Page: 54-56 , Section: 2-2

The Ogive (Cumulative frequency): Is a graph that represents the cumulative ~~frequency~~ frequencies for the classes in a frequency distribution.

Exp. 2-6

Figure 2-6

---

lecture: 4 , Page: 69 -70 -72 -73 , Section: 2-3

Bar Graphs when the data are ~~qualitative~~ qualitative or categorical.

~~Bar Graph~~ Bar graph: Is ~~represented~~ represents ~~the~~ the data by using vertical or ~~horizontal~~ horizontal bars whose heights or lenghts represent the frequencies of the data.

Exp. 2-8

Pareto Charts: Is used to represent a frequency distribution for a categorical variable, and frequencies are displayed by the heights of vertical bars, which are arranged in order from highest to lowest.

Exp. 2-9

Time series graph: represent data that ~~occur~~ occur over a specific period of time.

Exp. 2-10

Figure: 2-12

The pie graph: Is a circle that is divided into sections or wedges according to the percentage of frequencies in each category of the distribution.

lecture: 4 , Page: 74 - 75 , Section: 2-3

1. Degree $= \frac{f}{n} \cdot 360 \quad \therefore \frac{f}{\Sigma f} \cdot 360°$

2. [redacted] $\% = \frac{f}{n} \cdot 100 \quad \therefore \frac{f}{\Sigma f} \cdot 100\%$

Percentage

Exp. 2-12

lecture: 5 , Page: 106 - 107 - 109 , Section: 3-1

Statistic: Is a charcteristic or measure obtained by using the data values from Sample

Parameter: " " " " " " " " all the data values from a specific Population.

The mean: is the Sum of the values, divided by the total number of values. The Symbol $\overline{X}$ represents the sample mean.

$$\overline{X} = \frac{X_1 + X_2 + X_3 + \cdots + X_n}{n} = \frac{\Sigma X}{n}$$

for Population: $\mu = \frac{X_1 + X_2 + X_3 + \cdots X_N}{N} = \frac{\Sigma X}{N}$

Exp. 3-1

Exp. 3-2

The median: Is the midPoint of the data array. The Symbol of the median is MD. (نقطة المنتصف بعد ترتيب البيانات)

" , Page: 110, 111 , "

Exp. 3-4 Exp. 3-6 Exp. 3-8

Exp. 3-5 Exp. 3-7 Exp. 3-9

/ /

lecture: 4 , Page: 111 Section: 3-1

المنوال — قيمة أو قيم — مجموعة البيانات

The mode: The value that occurs most often in data set.

1. A data set that has only one value that occurs with the greatest frequency is said to Unimodal. وحيد المنوال

2. A data set has two values that occurs with the greatest frequency, both values are considered to be the mode and the data set is said to be bimodal. ثنائي المنوال

3. A data set has more than two values that occur with the same greatest frequency, each value is used as the mode, and the data set is said to be multimodal. متعدد المنوال

4. When no data value occurs more than once, the data set is said to have no mode. لا يوجد منوال أو عديم المنوال

Exp: 3-9 , 3-10 , 3-11

Exp of Median-MD:

1. 60 (1) 90 (2) 80 (3) 70 (4) 50 (5)

Solution: A. n = odd

B. 50 60 (70) 80 90

C. MD = 70

2. 100 (1) 60 (2) 90 (3) 80 (4) 70 (5) 50 (6)

Solution: A. n = even

B. 50 60 (70 80) 90 100

C. $MD = \frac{70+80}{2} = \frac{150}{2} = 75$

80 90 100 = البيانات أعلى الوسيط

50 60 70 = " أسفل "

imp for MD: 1. ترتيب البيانات تصاعدياً أو تنازلياً

2. إيجاد القيمة الوسطى

Exp of Mode:

1. 70 (1) 50 (1) 70 (2) 90 80 70 (3) 50 (2)

Solution: Mode = 70 (unimodal).

2. $\overset{1}{50}\ \overset{1}{70}\ \overset{2}{50}\ \overset{2}{70}\ 90\ 80\ \overset{3}{70}\ \overset{3}{50}$

Solution: Mode = 50, 70 (bimodal).

3. $\overset{1}{50}\ \overset{1}{70}\ \overset{2}{50}\ \overset{2}{70}\ \overset{1}{80}\ 90\ \overset{2}{80}\ \overset{3}{70}\ \overset{3}{50}\ \overset{3}{80}$

Solution: Mode = 50, 70, 80 (Multimodal).

4. $\overset{1}{50}\ \overset{1}{70}\ \overset{2}{50}\ \overset{2}{70}\ \overset{1}{80}\ \overset{2}{80}\ \overset{3}{70}\ \overset{3}{50}\ \overset{3}{80}$

Solution: (no mode)

The MidRange: $\frac{min + max}{2} = MR$ = وسط المدى

Exp. 60 90 80 70 50

Solution: $MR = \frac{50 + 90}{2} = \frac{140}{2} = 70$

Exp. 3-15, 3-16

The Weighted Mean: $\bar{X} = \frac{w_1 X_1 + w_2 X_2 + w_3 X_3 + \dots + w_n X_n}{w_1 + w_2 + w_3 + \dots + w_n} = \frac{\Sigma WX}{\Sigma W}$

Exp: 3-17

Exp.

note: $\bar{x}$ المتوسط

X	W	
4 = A	3	A = 4
2 = C	3	B = 3
3 = B	4	C = 2
1 = D	2	D = 1
		F = 0

$$\bar{X} = \frac{(4)(3) + (2)(3) + (3)(4) + (1)(2)}{3 + 3 + 4 + 2} = \frac{12 + 6 + 12 + 2}{12}$$

$$= \frac{32}{12} = 2.7$$

Properties and Uses of central Tendency:

A. The mean:

1.

4.

lecture: 6 ,Page: 116 -117 ,Section: 3-1

6.

B. The median:

1.

2.

4.

C. The mode:

1.

2.

3.

4.

D. The midrange:

1.

2.

3.

Figure: 3-1

A. Positively skewed:

B. Symmetric distribution:

C. negatively skewed: