« أحصاء نظري »   « المحاضرة الأولى »

* (علم الأحصاء) :- هو ذلك العلم الذي يعمل على أستخدام الاسلوب العلمي في طرق جمع البيانات وتبويبها وتلخيصها وعرضها وتحليلها بهدف الوصول منها الى استنتاجات وقرارات مناسبة.

ويمكن تقسيم علم الاحصاء بصورة عامة الى قسمين رئيسين :-

(1) - الأحصاء الوصفي (Descriptive Statistics)

ويشتمل على الطرق الاحصائية المستعملة في وصف مجموعة معينة من البيانات ، وتتضمن هذه الطرق الاحصائية اساليب جمع البيانات (Collection of data) في صورة قياسات رقمية (Numerical measurements) ثم تبويبها او تنظيمها (Organizing) وتلخيصها (Summarizing) وعرضها (Presenting) وحساب بعض المقاييس الاحصائية المختلفة لها.

(2)- الأحصاء الأستنتاجي أو الأستدلالي (Statistical Inference)

ويشمل الطرق الأحصائية التي تهدف الى عمل استنتاجات او استدلالات حول المصدر التي جمعت منه البيانات ويضم هذا القسم فرعين رئيسين :-

(أ)- التقدير (Estimation) :- ويهتم بإيجاد قيم تقديرية للاستدلال منها على القيم الحقيقية لمصدر جمع البيانات. وهذه القيم التقديرية اما ان يكون تقديراً محدداً اي عند نقطة معينة (Point estimation) او تقديراً في فترة او مدى (Interval estimation).

(ب)- أختبار الفرضيات (Test of Hypotheses) :-

ويتضمن أختبار الفرضيات التي توضع كتفسير اولي للظاهرة المراد دراستها للوصول منها الى قرار بقبولها او رفضها.

تاريخ علم الإحصاء (History of Statistics)

* ان كلمة الإحصاء في الماضي كانت تهدف الى العد والحصر حتى سمي الإحصاء سابقاً بعلم العد (The science of counting) ، كما ان لفظ إحصاء باللغة الانكليزية (Statistics) كانت تستعمل في بلاد أوربا للدلالة على أعمال وحسابات الدولة في شؤون الحرب والضرائب وعدد السكان والمواليد والوفيات والانتاج الزراعي ....الخ .

* اما الآن فقد تطور الإحصاء كثيراً وخاصة في القرن العشرين واصبح علماً مستقلاً له اهميته كوسيلة واداة في البحث العلمي لجميع العلوم .

* ومن اهم العلماء الذين درسوا وطبقوا علم الإحصاء هم العالم Laplace (١٧٤٩ - ١٨٢٧) الذي قام بتطبيق علم الإحصاء في علم الفلك . كما تم تطبيق علم الإحصاء من قبل الجيولوجي Charles Lyell (١٧٩٧ - ١٨٧٥) والبايلوجي Charles Darwin (١٨٠٩ - ١٨٨٢) وعالم النبات Johom Gregor Mendel (١٨٢٢ - ١٨٨٤) بالرغم من كونهم غير إحصائيين . وفي القرن التاسع عشر اشتهر العالم البلجيكي Adolph Quetelet (١٧٩٤ - ١٨٧٤) بتطبيق علم الإحصاء بشكل فعال في علميا الاجتماع والتعليم . ثم جاء العالم Francis Galton (١٨٢٢ - ١٩١١) والذي اشتهر بتطبيق علم الإحصاء في علوم الوراثة والتطور . اما العالم الرياضي الفيزيائي Karl Pearson فقد اشترك مع Galton في إيجاد نظرية الارتباط Correlation والانحدار Regression .

* أما أشهر علماء القرن العشرين فهو العالم R.A. Fisher (١٨٩٠ - ١٩٦٢) الذي طور علم الإحصاء وطبقه في علوم كثيرة كالزراعة والوراثة والاقتصاد ووضع اسس تصميم وتحليل التجارب . ومن العلماء الآخرين الذين اسهموا في تطوير علم الإحصاء هم :-

Smirnov , Kolmogorov , A. Wald , J. Neyman , E.S. Pearson وغيرهم .

** بعض معاني الرموز الإحصائية

ت	المعنى	الرمز
١	أكبر من	$>$
٢	أكبر من أو يساوي	$\geq$
٣	أصغر من	$<$
٤	أصغر من أو يساوي	$\leq$
٥	قيمة مطلقة	$\mid \; \mid$
٦	مجموع	Σ
٧	التكرار	f_i
٨	مضروب n (أو عاملي)	$n!$
٩	المدى	R
١٠	قيمة مشاهدة أو مفردة (أو مركز فئة)	y_i
١١	الوسط الحسابي للعينة	$\bar{y}$
١٢	الوسط الهندسي	$\bar{G}$
١٣	الوسط التوافقي	$\bar{H}$
١٤	الوسط التربيعي	$\bar{Q}$
١٥	الوسيط	$\bar{M}_e$
١٦	المنوال	$\bar{M}_o$
١٧	الوسط الحسابي للمجتمع	μ
١٨	تباين المجتمع	σ^2
١٩	الانحراف القياسي للمجتمع	σ
٢٠	تباين العينة	S^2

ت	المعنى	الرمز
٢١	الانحراف القياسي للعينة	S
٢٢	تباين الوسط الحسابي	$S^2_{\bar{y}}$
٢٣	الخطأ القياسي	$S_{\bar{y}}$
٢٤	التباين المجمع	S^2_p
٢٥	الانحراف المتوسط	$M.D.$
٢٦	معامل الاختلاف	$C.V.$
٢٧	تباديل r من n	nP_r
٢٨	توافيق r من n	$nC_r=\binom{n}{r}$
٢٩	معامل الانحدار للعينة	b
٣٠	معامل الارتباط للعينة	r
٣١	مستوى المعنوية (أو مستوى الاحتمال)	α
٣٢	فرضية العدم	H_0
٣٣	الفرضية البديلة	H_1
٣٤	مجموع مربعات الانحرافات	SS
٣٥	احتمال النجاح	P
٣٦	احتمال الفشل	q
٣٧	مالانهاية	∞
٣٨	قيمة مشاهدة	O_i
٣٩	القيمة المتوقعة	$E(y)$
٤٠	معامل الانحدار للمجتمع (معامل التفاضل)	B

« المحاضرة الثانية »

طبيعة البيانات الاحصائية :-

عند جمع بيانات حول ظاهرة ما فاننا نرمز للظاهرة بالرمز (Y) وكل مفردة او مشاهدة من هذه الظاهرة نرمز لها بالرمز (Y_i) .

* مثلاً عند دراسة أطوال الطلبة لجامعة ما فاننا نرمز لصفة الطول بالرمز (Y) ونرمز لطول اي طالب بالرمز (Y_i) وتسمى المشاهدة او المفردة (Observation) وان قيمة (Y_i) قد تختلف من طالب الى آخر ولهذا نقول بان (Y) متغير (Variable) .

((* إذن المتغير هو اي ظاهرة تظهر إختلافات بين مفرداتها ويرمز له بالرمز Y أو اي رمز آخر مثل X او Z ......))

* والمتغيرات Variables تنقسم الى قسمين :-

(1)- متغيرات وصفية أو نوعية (Qualitative Variables)

وهي تلك الظواهر أو الصفات التي لا يمكن قياسها مباشرة بالارقام العددية مثل صفة لون العيون (أزرق - اسود - بني) والحالة الاجتماعية (غني - متوسط الحال - فقير) والجنس (ذكر - انثى) .... الخ

(2)- متغيرات كمية (Quantitative Variables)

وهي تلك الظواهر او الصفات التي يمكن قياسها مباشرة بارقام عددية مثل صفة الطول والوزن والعمر وكمية الحصول ..... الخ . وتنقسم الى قسمين :-

(P) - متغيرات مستمرة (او متصلة) Continuous Variables

فالمتغير المستمر هو المتغير الذي تأخذ المشاهدة او المفردة فيه اي قيمة رقمية في مدى معين .

* مثلاً لو فرضنا ان أطوال طلبة جامعة ما تتراوح بين 130,5 سم و 170 سم فنقول بان :- $(130{,}5 \leqslant Y \leqslant 170)$

أي ان المتغير Y يمكن أن يأخذ أي قيمة بين 130,5 سم و 170 سم .

* امثلة أخرى على المتغيرات المستمرة (المتصلة) هي :- (الوزن - الزمن - كمية الحاصل - درجة الحرارة) لأنه يمكن قياسها بأجزاء صغيرة جداً وتأخذ اي قيمة تقع في حدود معينة .

* بصورة عامة كل البيانات التي تقاس (Measurements) تعتبر بيانات لمتغير مستمر .

(ب)- متغيرات غير مستمرة (أو منفصلة) Discrete variables

المتغير المنفصل هو المتغير الذي تأخذ المشاهدة او المفردة فيه قيما متباعدة او متقطعة غير مستمرة .

* مثلاً لو فرضنا ان عدد افراد الأسرة في أربعة عوائل هي :- (٢، ٣، ٤، ٥) فنقول بان $(y = 2, 3, 4, 5)$.

* امثلة أخرى على المتغيرات غير المستمرة او المنفصلة هي :- عدد الثمار على النباتات وعدد الوحدات الانتاجية لمصنع ما - عدد الطلبة في الصفوف الأولى لجامعة ما ... وفي الغالب تكون أعداد صحيحة .

* بصورة عامة كل البيانات التي نحصل عليها من العد Countings تعتبر بيانات لمتغير منفصل

* المجتمع والعينة (Population and sample).

(١) - المجتمع :- (Population) .
عبارة عن جميع القيم او المفردات التي يمكن أن يأخذها المتغير

* فمثلاً اذا كانت الدراسة متعلقة بأطوال جامعة ما فان المجتمع في هذه الحالة هو أطوال جميع الطلبة في تلك الجامعة .

* والمجتمع اما ان يكون :-

(أ) - مجتمعاً محدوداً (Finite Population) :
اي يمكن حصر عدد مفرداته كما هو الحال في اطوال جامعة الموصل مثلاً او عدد الوحدات الأنتاجية لمصنع ما في يوم معين

(ب)- مجتمعاً غير محدود (Infinite population)

وهو المجتمع الذي من الصعب او المستحيل حصر عدد مفرداته مثل مجتمع نوع سمك معين في نهر دجلة او عدد البكتريا في حقل ما.

(ج)- العينة :- (Sample)

العينة هي جزء من المجتمع.
فالعينة عبارة عن مجموعة من المشاهدات اختيرت بطريقة ما من المجتمع.

* ان دراسة المجتمع ككل قد يكون صعباً ويحتاج الى جهد ووقت ومال لذا فقد استعيض عن دراسة المجتمع بدراسة العينة وصفاتها ومن خلالها نستطيع ان نستنتج خواص المجتمع الأصلي الذي أخذت منه هذه العينة.

* الرموز الاحصائية (Statistical notations)

كما ذكرنا سابقاً سنرمز للمتغير بالرمز y ولكل قيمة له بالرمز y_i

* فمثلاً لو كان عندنا أعمار (٥ طلاب) كالآتي :-

$$y_i = 20, 18, 22, 16, 17$$

أي ان :- $(y_1 = 20)$ أي القيمة الأولى للمتغير او المشاهدة الأولى
$(y_2 = 18)$ أي القيمة الثانية للمتغير او المشاهدة الثانية
$(y_3 = 22)$ أي القيمة الثالثة للمتغير أو المشاهدة الثالثة
$(y_4 = 16)$ أي القيمة الرابعة للمتغير أو المشاهدة الرابعة
$(n = 5) \Leftarrow (y_5 = y_n = 17)$ أي القيمة الأخيرة (الخامسة) للمتغير او المشاهدة الأخيرة

* ملاحظة :- ① - n = دائماً تمثل عدد المشاهدات
② - y_n = تمثل المشاهدة الأخيرة

* يرمز عادة لمجموع قيم المتغير (المشاهدات) بالرمز $\left(\sum_{i=1}^{n} y_i\right)$

فالرمز $\sum$ هو حرف إغريقي يمثل أو يعبر عن المجموع ويسمى (Sigma) أو (Summation of) والرقمان (1 و n) هما حدا المجموع من أول مشاهدة الى آخر مشاهدة .

$$\sum_{i=1}^{n} y_i = y_1 + y_2 + y_3 + y_4 + y_5$$
$$= 20 + 18 + 22 + 16 + 17$$

* وللاختصار والسهولة تكتب الرمز السابق $\left(\sum_{i=1}^{n} y_i\right)$ بدون ذكر حدي المجموع أي $\left(\sum y_i\right)$ فقط إذ لم يكن هناك خوف من الالتباس .

*أي إن :-

$$\sum_{i=1}^{n} y_i = \sum y_i = y_1 + y_2 + y_3 + y_4 + y_5$$

---

* وهناك مجموع جزئي مثل $\left(\sum_{i=3}^{5} y_i\right)$ أي مجموع المشاهدة الثالثة والرابعة والخامسة فقط .

$$\sum_{i=3}^{5} y_i = y_3 + y_4 + y_5$$
$$22 + 16 + 17$$

---

* ويرمز لمجموع مربعات جميع المشاهدات بالرمز $\left(\sum_{i=1}^{n} y_i^2\right)$

أي أن :

$$\sum_{i=1}^{n} y_i^2 = y_1^2 + y_2^2 + y_3^2 + y_4^2 + y_5^2$$
$$= (20)^2 + (18)^2 + (22)^2 + (16)^2 + (17)^2$$

* ويرمز لمربع مجموع المشاهدات بالرمز $\left(\sum_{i=1}^{n} y_i\right)^2$.

أي ان :-

$$\left(\sum_{i=1}^{n} y_i\right)^2 = (y_1+y_2+y_3+y_4+y_5)^2$$
$$=(20+18+22+16+17)^2 = (93)^2$$

* ويرمز لمجموع حاصل ضرب قيم متغيرين X و Y بالرمز $\left(\sum x_i y_i\right)$

أي ان :-

$$\sum x_i y_i = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n$$

* ويرمز لحاصل ضرب مجموعين لقيم متغيرين بالرمز $(\sum x_i)(\sum y_i)$

أي ان :-

$$(\sum x_i)(\sum y_i) = (x_1+x_2+\cdots+x_n)(y_1+y_2+\cdots+y_n)$$

*(مثال) * نفرض ان قيمة المتغير Y هي كالاتي :-

$$y_i = 3, 9, 6, 2$$

وان قيمة المتغير X هي كالاتي :-

$$x_i = 4, 2, 3, 7$$

أوجد قيمة كل ما يأتي :-

ⓐ- $\sum_{i=1}^{n} y_i$

الحل $\sum_{i=1}^{n} y_i = y_1+y_2+y_3+y_4$
$= 3+9+6+2 = 20$

ⓑ- $\sum_{i=2}^{3} y_i$

الحل $\sum_{i=2}^{3} y_i = y_2+y_3 = 9+6 = 15$

(c) - $\sum y_i^2$

الحل $\sum y_i^2 = y_1^2 + y_2^2 + y_3^2 + y_4^2 = (3)^2 + (9)^2 + (6)^2 + (2)^2 = 130$

(d) - $(\sum y_i)^2$

الحل $(\sum y_i)^2 = (y_1 + y_2 + y_3 + y_4)^2 = (3 + 9 + 6 + 2)^2 = (20)^2 = 400$

(e) - $\sum x_i y_i$

الحل $\sum x_i y_i = x_1 y_1 + x_2 y_2 + x_3 y_3 + x_4 y_4$

$= 4 \times 3 + 2 \times 9 + 3 \times 6 + 7 \times 2 = 62$

(f) - $(\sum x_i)(\sum y_i)$

الحل $(\sum x_i)(\sum y_i) = (x_1 + x_2 + x_3 + x_4)(y_1 + y_2 + y_3 + y_4)$

$= (4 + 2 + 3 + 7)(3 + 9 + 6 + 2)$

$= 16 \times 20 = 320$

* وفيما يلي بعض القواعد المفيدة في عملية الجمع :-

* قاعدة (1) :- اذا كانت (c) أي عدد ثابت (قيمة ثابتة) فان

$$\sum_{i=1}^{n} c = nc$$

حيث ان n = عدد المشاهدات

* قاعدة (2) :- اذا كانت (c) اي عدد ثابت فان

$$\sum c y_i = c \sum y_i$$

* قاعدة (3) :- جمع قيم متغيرين أو اكثر هو مجموع مجموعهم اي ان :

$$\sum (x_i + y_i) = \sum x_i + \sum y_i$$

*قاعدة (٤) :-

$$\sum (x_i + y_i)^2 = (x_1+y_1)^2 + (x_2+y_2)^2 + \cdots + (x_n+y_n)^2$$

هذا ويجب التفريق بين بعض الرموز الأحصائية مثل

$$\sum \frac{x_i}{y_i} = \frac{x_1}{y_1} + \frac{x_2}{y_2} + \cdots + \frac{x_n}{y_n}$$

$$\frac{\sum x_i}{\sum y_i} = \frac{x_1+x_2+\cdots+x_n}{y_1+y_2+\cdots+y_n}$$

$$\sum (x_i - 3) = \sum x_i - n(3)$$

$$\sum x_i - 3$$

*(مثال) * اذا علمت بان قيم كل من المتغيرين x و y هي كالاتي:-

$$x_i = 2, 6, 3, 1 \quad , \quad y_i = 3, 9, 6, 2$$

أوجد قيمة كل مما يأتي :-

ⓐ- $\sum (y_i - x_i)^2$

الحل

$$\sum (y_i - x_i)^2 = (y_1-x_1)^2 + (y_2-x_2)^2 + (y_3-x_3)^2 + (y_4-x_4)^2$$
$$= (3-2)^2 + (9-6)^2 + (6-3)^2 + (2-1)^2$$
$$= 20$$

ⓑ- $\sum (x_i-3)(y_i-5)$

الحل

$$\sum (x_i-3)(y_i-5) = (x_1-3)(y_1-5) + (x_2-3)(y_2-5) + (x_3-3)(y_3-5) + (x_4-3)(y_4-5)$$
$$= (2-3)(3-5) + (6-3)(9-5) + (3-3)(6-5) + (1-3)(2-5)$$
$$= 20$$

(C)- $\sum x_i y_i^2$

الحل $\sum x_i y_i^2 = x_1 y_1^2 + x_2 y_2^2 + x_3 y_3^2 + x_4 y_4^2$

$= (2)(3)^2 + (6)(9)^2 + (3)(6)^2 + (1)(2)^2 = 616$

(d)- $\sum (y_i - 3)$

الحل $\sum (y_i - 3) = \sum y_i - n(3)$

$= \sum y_i - 4(3) = 20 - 12 = 8$

(E)- $\sum y_i - 3$

الحل $\sum y_i - 3 = 20 - 3 = 17$

(f)- $\sum \frac{x_i + 2}{y_i}$

الحل $\sum \frac{x_i+2}{y_i} = \frac{x_1+2}{y_1} + \frac{x_2+2}{y_2} + \frac{x_3+2}{y_3} + \frac{x_4+2}{y_3}$

$= \frac{2+2}{3} + \frac{6+2}{9} + \frac{3+2}{6} + \frac{1+2}{2} = \frac{164}{36}$

(G)- $\frac{\sum (x_i + 2)}{\sum y_i}$

الحل $\frac{\sum (x_i+2)}{\sum y_i} = \frac{\sum x_i + n(2)}{\sum y_i} = \frac{12+8}{20} = 1$

(H)- $\sum y_i^2 - \frac{(\sum y_i)^2}{n}$

الحل $\sum y_i^2 - \frac{(\sum y_i)^2}{n} = (y_1^2 + y_2^2 + y_3^2 + y_4^2) - \frac{(y_1+y_2+y_3+y_4)^2}{4}$

$= (3)^2 + (9)^2 + (6)^2 + (2)^2 - \frac{(3+9+6+2)^2}{4} = 30$

(I)- $\sum x_i y_i - \frac{(\sum x_i)(\sum y_i)}{n}$

الحل $\sum x_i y_i - \frac{(\sum x_i)(\sum y_i)}{n} = (x_1 y_1 + x_2 y_2 + x_3 y_3 + x_4 y_4) - \frac{(\sum x_i)(\sum y_i)}{n}$

$= (2\times3 + 6\times9 + 3\times6 + 1\times2 - \frac{(12)(20)}{4}$

$= 80 - \frac{(12)(20)}{4} = 20$

العرض الجدولي والتمثيل البياني :-

عند جمع البيانات الأولية (Raw data) الخاصة بدراسة ظاهرة ما فانه عادة لا يمكن الاستفادة منها وهي بهذه الصورة . لذلك فغالباً ما توضع في جداول مبسطة او يعبر عنها في صورة اشكال ورسوم بيانية لكي يسهل دراستها وتحليلها .

① - العرض الجدولي (Tabular presentation)

هناك نوعان من الجداول الأحصائية هما :-

(أ)- الجدول البسيط :- وهو الجدول التي توزع فيه البيانات حسب صفة واحدة ويتألف عادة من عمودين :- الأول يمثل تقسيم الصفة او الظاهرة الى فئات او مجموعات والثاني يبين عدد المفردات التابعة لكل فئة او مجموعة .(جدول ١) .

* جدول (١) : يوضح توزيع عدد من طلبة جامعة ما حسب أوزانهم (كغم)

فئات الوزن (كغم)	عدد الطلبة
٦٠ - ٦٢	٥
٦٣ - ٦٥	١٥
٦٦ - ٦٨	٤٥
٦٩ - ٧١	٢٧
٧٢ - ٧٤	٨
المجموع	١٠٠

(ب)- الجدول المركب :- وهو الجدول التي توزع فيه البيانات حسب صفتين او ظاهرتين أو اكثر في نفس الوقت .
ويتألف من (الصفوف) التي تمثل فئات او مجاميع احدى الصفتين (والاعمدة) التي تمثل فئات او مجاميع الصفة الأخرى .
اما المربعات التي تقابل الصفوف او الأعمدة فتحتوي على عدد المفردات أو التكرارات المشتركة في فئات ومجاميع كلتا الصفتين . (جدول ٢)

*جدول (١): يوضح توزيع عدد من طلبة جامعة ما حسب صفتي الطول بـ (سم) والوزن (كغم).

الطول (سم) / الوزن (كغم)	٥١-٦٠	٦١-٧٠	٧١-٨٠	المجموع
١٢١-١٤٠	٢٠	٦	٤	٣٠
١٤١-١٦٠	٢	٤٠	١٠	٥٢
١٦١-١٨٠	٢	٦	١٠	١٨
المجموع	٢٤	٥٢	٢٤	١٠٠

٢- جدول التوزيع التكراري: (Frequency Table)

هو جدول بسيط يتكون من عمودين :-

*(الأول) وتقسم فيه قيم المتغير الى اقسام او مجموعات تسمى الفئات (Classes).

*(الثاني) يبين مفردات كل فئة ويسمى بالتكرار (Frequency) كما في (جدول ٢).

*جدول (٢) :- يوضح توزيع تكراري لأطوال (٨٠ نبات) من القطن بـ (م)

فئات الطول (م)	التكرار (عدد النباتات)
٢١-٤٠	١
٤١-٥٠	٢
٥١-٦٠	٥
٦١-٧٠	١٥
٧١-٨٠	٢٥
٨١-٩٠	٢٠
٩١-١٠٠	١٢
المجموع	٨٠

* بعض التعاريف المهمة :-

١- البيانات غير المبوبة :- (Ungrouped date) .

وهي البيانات الأصلية أو الأولية التي جمعت ولم تبوب .

٢- البيانات المبوبة :- (Grouped date) .

وهي البيانات التي بوبت ونظمت في جدول توزيع تكراري .

٣- الفئات :- (Classes) .

وهي المجاميع التي قسمت اليها قيم المتغير . وكل فئة تأخذ مدى معين من قيم المتغير . مثلاً جدول (٢) السابق يحوي على سبع فئات .

٤- حدود الفئات :- (Class limits) -

لكل فئة حدان . حد أعلى وحد أدنى .

٥- الحدود الحقيقية للفئات :- (True class limits) .

لكل فئة حدان حقيقيان . حد اعلى حقيقي وحد أدنى حقيقي .

٦- طول الفئة :- (Class length) .

وهو مقدار المدى بين حدي الفئة . هذا ويستحسن ان تكون أطوال الفئات متساوية لتسهيل العملية الحسابية . ويرمز لطول الفئة بالرمز (C) .

٧- مركز الفئة : (Class mark or class mid-point) .

لكل فئة مركز وسنرمز له بـ (y_i) وهو عبارة عن منتصف المدى بين حدي الفئة .

٨- تكرار الفئة :- (Class frequency) وهي عدد المفردات أو القيم التي تقع في مدى تلك الفئة وسنرمز لها بـ (f_i) . هذا ومجموع التكرارات يجب ان يكون دائماً مساوياً للعدد الكلي لقيم الظاهرة . كما في جدول (٢) السابق عدد التكرارات يساوي المجموع الكلي وهو (٨٠) .

والجدول (٤) يوضح ما سبقا شرحه بالتفصيل .

*جدول (٤) : جدول توزيع تكراري لأطوال نباتات القطن مبينا فيه الحدود الحقيقية ومراكز الفئات .

تسلسل الفئات	الفئات	الحدود الحقيقية للفئات	مركز الفئة (ﺯi)	التكرار (fi)
١	٢١ - ٤٠	٢٠٫٥ - ٤٠٫٥	٣٥٫٥	١
٢	٤١ - ٥٠	٤٠٫٥ - ٥٠٫٥	٤٥٫٥	٢
٣	٥١ - ٦٠	٥٠٫٥ - ٦٠٫٥	٥٥٫٥	٥
٤	٦١ - ٧٠	٦٠٫٥ - ٧٠٫٥	٦٥٫٥	١٥
٥	٧١ - ٨٠	٧٠٫٥ - ٨٠٫٥	٧٥٫٥	٢٥
٦	٨١ - ٩٠	٨٠٫٥ - ٩٠٫٥	٨٥٫٥	٢٠
٧	٩١ - ١٠٠	٩٠٫٥ - ١٠٠٫٥	٩٥٫٥	١٢
	المجموع			٨٠

* كيفية حساب طول الفئة والحدود الحقيقية للفئة ومركز الفئة .

* مثلاً نأخذ الفئة الرابعة من جدول (٤) أعلاه :- وتساوي = (٦١ - ٧٠) ؛ فالحد الأدنى للفئة الرابعة = ٦١ ، والحد الأعلى للفئة الرابعة = ٧٠ .

* يتم حساب الطول للفئة الرابعة بإحدى الطرق التالية

* (١) - عندما تكون حدود الفئات أعداد صحيحة فقط .

$$\boxed{\text{طول الفئة} = \text{الحد الأعلى} - \text{الحد الأدنى} + ١}$$

= ٧٠ - ٦١ + ١ = ١٠

* (٢) - طول الفئة = الحد الحقيقي الأعلى - الحد الحقيقي الأدنى لتلك الفئة

= ٧٠٫٥ - ٦٠٫٥ = ١٠

* (٣)- طول الفئة = الفرق بين الحدين الأدنى او الحدين الأعلى لفئتين متتاليتين .

* الفرق بين الحدين الأدنى = ٧١ - ٦١ = ١٠

* الفرق بين الحدين الأعلى = ٨٠ - ٧٠ = ١٠

* (٤)- طول الفئة = الفرق بين الحدين الحقيقين الأدنى أو الأعلى لفئتين متتاليتين .

* الفرق بين الحدين الحقيقين الأدنى = ٧٠٫٥ - ٦٠٫٥ = ١٠

* الفرق بين الحدين الحقيقين الأعلى = ٨٠٫٥ - ٧٠٫٥ = ١٠

* (٥)- طول الفئة = الفرق بين مركزي فئتين متتاليتين .

= ٧٥٫٥ - ٦٥٫٥ = ١٠

* اما طرق حساب (الحدود الحقيقية) لأي فئة كما يلي :.

* (١) - الحد (الأدنى) الحقيقي لأي فئة = مركز تلك الفئة - $\frac{1}{2}$ (طول تلك الفئة)

∴ الحد الأدنى للفئة الرابعة = ٦٥٫٥ - $\frac{1}{2}$ (١٠) = ٦٠٫٥

اما الحد (الأعلى) الحقيقي لأي فئة = مركز تلك الفئة + $\frac{1}{2}$ (طول تلك الفئة)

∴ الحد الأعلى للفئة الرابعة = ٦٥٫٥ + $\frac{1}{2}$ (١٠) = ٧٠٫٥

* (٢)- الحد (الأدنى) الحقيقي لأي فئة = (الحد الأدنى لتلك الفئة + الحد الأعلى للفئة السابقة) ÷ ٢

∴ الحد الأدنى للفئة الرابعة = $\frac{٦١ + ٦٠}{٢}$ = ٦٠٫٥

اما الحد (الأعلى) الحقيقي لأي فئة = (الحد الأعلى لتلك الفئة + الحد الأدنى للفئة التي تليها) ÷ ٢

∴ الحد الأعلى للفئة الرابعة = $\frac{٧٠ + ٧١}{٢}$ = ٧٠٫٥

* ملاحظة :- اذا كانت حدود الفئات أعداد صحيحة فان :-

② - الحد (الأدنى) الحقيقي لأي فئة = الحد الادنى لتلك الفئة - ٠٫٥

مثلا للفئة الرابعة = ٦١ - ٠٫٥ = ٦٠٫٥

اما الحد (الأعلى) الحقيقي لأي فئة = الحد الاعلى لتلك الفئة + ٠٫٥

مثلا للفئة الرابعة = ٧٠ + ٠٫٥ = ٧٠٫٥

* اما طرق حساب (مركز الفئة) فهي كالاتي :

* ① مركز الفئة = $\frac{\text{الحد الادنى + الحد الأعلى}}{٢}$

اذن مركز الفئة الرابعة = $\frac{٦١ + ٧٠}{٢}$ = ٦٥٫٥

* ② مركز الفئة = $\frac{\text{الحد الادنى الحقيقي + الحد الاعلى الحقيقي}}{٢}$

إذن مركز الفئة الرابعة = $\frac{٧٠٫٥ + ٦٠٫٥}{٢}$ = ٦٥٫٥

* ملاحظة * عندما نقول ان تكرار الفئة الرابعة = ١٥
هذا يعني ان هناك (١٥ قيمة) من قيم المتغير
واقعة في المدى (٦١ - ٧٠) .

* الخطوات العامة في إنشاء جداول التوزيع التكرارية :-

* لانشاء جدول توزيع تكراري يجب إتباع الخطوات التالية :-

① - إستخراج مدى المتغير .
② - إختيار وتحديد عدد الفئات .
③ - إيجاد طول مدى الفئة .
④ - كتابة حدود الفئات .
⑤ - إستخراج عدد التكرارات لكل فئة .

والمثال التالي يوضح كيفية إنشاء جدول توزيع تكراري (جدول ٥) لنباتات القطن .

* مثال :- القيم التالية تمثل أطوال (٨٠) نبات من القطن والمطلوب إنشاء جدول توزيع تكراري لأطوال هذه النباتات

* جدول (٥) :- يوضح أطوال (٨٠) نبات من القطن بـ(سم)

٨٠	٧٩	٤٨	٧٤	٨١	٩٨	٨٧	٨٠
٨٤	٨٠	٩٠	٧٠	٩١	٩٢	٨٣	٧٨
٧١	٧٠	٩٣	٦٥	٥٦	٨١	٧٤	٧٢
٧٣	٦٨	٨٥	٥١	٦٥	٩٢	٨٢	٨٦
٢٥	٩٠	٨٢	٧٢	٧٤	٤٢	٨٦	٦٨
٩٢	٩٣	٧٦	٧١	٩٠	٧٣	٦٧	٧٥
٩١	٨٠	٦١	٧٣	٩٧	٩١	٨٨	٨١
٧٤	٧٠	٩٩	٩٥	٨٠	٥٩	٧١	٧٧
٦٠	٦٢	٨٢	٨٣	٦٠	٦٧	٨٩	٦٢
٦٢	٧٦	٨٨	٧٠	٦٦	٨٨	٧٩	٧٥

الحل :* (١)- استخراج المدى (مدى المتغير) .

المدى = (اعلى قيمة - اقل قيمة = ٩٩ - ٢٥ = ٦٤ سم

* (٢)- أختيار وتحديد عدد الفئات .

سوف يتم تحديد عدد الفئات من السؤال ولنفرض انها تساوي (٧) .

* (٣)- ايجاد طول الفئة = مدى المتغير / عدد الفئات (مقربة الى اقرب عدد صحيح أكبر)

اذن طول الفئة = ٦٤ / ٧ = ٩ ١/٧ = تقرب الى ١٠

اذن طول الفئة = ١٠

* (٤)- كتابة حدود الفئات . يجب كتابة حدود الفئات بحيث إن جميع قيم المتغير تقع بين الحد الادنى للفئة الأولى والحد الاعلى للفئة الأخيرة .

ويستحسن ان نبدأ بكتابة الحد الأدنى للفئة بقيمة اقل من اصغر مفردة بقليل وننتهي بالحد الأعلى للفئة الأخيرة بقيمة اكبر بقليل من اكبر مفردة . فمثلاً أصغر قيمة من قيم أطوال النبات (جدول ٥) هي (٢٢٫٥) لذا كان من الممكن أن يكون الرقم (٢١) يمثل الحد الادنى للفئة الأولى وبما أن طول الفئة هي (١٠) لذا كان حدي الفئة الأولى هما (٢١ـ٤٠) والفئة الثانية تبدأ من (٤١ـ٥٠) بينما الفئة السابعة (الأخيرة) هي (٩١ـ١٠٠) . بحيث ان الحد الادنى للفئة الأولى (٢١) والحد الأعلى للفئة الأخيرة (١٠٠) تحوي على كافة قيم المتغير .
وهذا موضح في الجدول (٦) التالي .

الفئات	التكرار بالعلامات	التكرار رقماً
٢١ـ٤٠	𝍷	١
٤١ـ٥٠	𝍷𝍷	٢
٥١ـ٦٠	𝍸	٥
٦١ـ٧٠	𝍸 𝍸 𝍸	١٥
٧١ـ٨٠	𝍸 𝍸 𝍸 𝍸 𝍸	٢٥
٨١ـ٩٠	𝍸 𝍸 𝍸 𝍸	٢٠
٩١ـ١٠٠	𝍷𝍷 𝍸 𝍸	١٢
المجموع		٨٠

*⑤ - استخراج عدد التكرارات لكل صنف . ويتم ذلك بتسجيل القيم الأصلية واحدة بعد الأخرى (جدول ٥) في الفئة الخاصة بها على شكل إشارات او علامات ثم تحويلها الى ارقام كما في (الجدول ٦) اعلاه .

* هذا ويجب التأكد بأن المجموع الكلي للتكرارات يجب ان يساوي العدد الكلي لقيم المتغير .

② – جدول التوزيع التكراري النسبي :- Relative Frequency Distribution

وهو جدول يبين الأهمية النسبية لكل فئة . ويحسب التكرار النسبي لكل فئة بالطريقة التالية .

* التكرار النسبي لأي فئة = $\frac{\text{تكرار تلك الفئة}}{\text{المجموع الكلي للتكرارات}} = \frac{f_i}{\sum f_i}$

* فمثلاً . التكرار النسبي للفئة الرابعة في (الجدول ٧) أدناه

$= \frac{\text{تكرار الفئة الرابعة}}{\text{المجموع الكلي للتكرارات}} = \frac{١٥}{٨٠} =$ ٫١٨٧٥

وعادة يوضع التكرار النسبي لأي فئة كنسبة مئوية وذلك بضرب كل تكرار نسبي لكل فئة × ١٠٠ (جدول ٧) .

فمثلاً للفئة الرابعة = ٫١٨٧٥ × ١٠٠ = ١٨٫٧٥

* جدول (٧) :- يوضح جدول التوزيع التكراري النسبي والمئوي لأطوال نباتات القمح

الفئات	التكرار	التكرار النسبي	التكرار النسبي المئوي
٢١ – ٤٠	١	٫٠١٢٥	١٫٢٥
٤١ – ٥٠	٢	٫٠٢٥٠	٢٫٥٠
٥١ – ٦٠	٥	٫٠٦٢٥	٦٫٢٥
٦١ – ٧٠	١٥	٫١٨٧٥	١٨٫٧٥
٧١ – ٨٠	٢٥	٫٣١٢٥	٣١٫٢٥
٨١ – ٩٠	٢٠	٫٢٥٠٠	٢٥٫٠٠
٩١ – ١٠٠	١٢	٫١٥٠٠	١٥٫٠٠
المجموع	٨٠	١٫٠٠٠٠	١٠٠٫٠٠

٤ التوزيعات المتجمعة : (Cumulative Distribution)

هناك نوعان من هذه الجداول للتوزيعات المتجمعة هي :-

* ٦ جدول التوزيع التكراري التجميعي التصاعدي .

(Less than cumulative distribution)

وهو ذلك الجدول الذي يعطينا عدد المفردات التي تقل قيمتها عن الحد الأدنى لفئة معينة . وسنرمز للتكرار المتجمع لأي فئة ب (F_i) ويتكون جدول التوزيع التكراري المتجمع التصاعدي من عمودين :-

* (العمود الأول) ونكتب فيه حدود الفئات (جدول ٨) التالي
* (العمود الثاني) ونكتب فيه التكرار التجميعي التصاعدي بالشكل التالي . (جدول ٨)
  * تكرار ماقبل الفئة الأولى = F_0 = صفر
  * تكرار الفئة الأولى = $F_1 = f_1$
  * تكرار الفئة الثانية = $F_2 = f_1+f_2$
  * تكرار الفئة الثالثة = $F_3 = f_1+f_2+f_3$
* وهكذا بحيث ان التكرار التجميعي التصاعدي للفئة الاخيرة = $\sum f_i = f_n$
* جدول (٨) :- يوضح التوزيع التكراري التجميعي التصاعدي لأطوال نباتات القطن الموضح تكرارها في (جدول ٧)

التكرار التجميعي التصاعدي	حدود الفئات
٠	اقل من ٢١
١	اقل من ٤١
٢	اقل من ٥١
٨	اقل من ٦١
٢٢	اقل من ٧١
٤٨	اقل من ٨١
٦٨	اقل من ٩١
٨٠	اقل من ١٠١

*(ب) جدول التوزيع التكراري التجميعي التنازلي

(More than cumulation distribution)

وهو الجدول الذي يعطينا عدد المفردات التي تزيد قيمتها عن الحد الأدنى لفئة معينة . ويتألف أيضاً من عمودين :-

*(العمود الأول) وتكتب فيه حدود الفئات (جدول ٩) التالي

*(العمود الثاني) تكتب فيه التكرارات التجميعية التنازلية بالطريقة التالية (جدول ٩) .

* تكرار الفئة الأولى = $\sum f_i = F_1$

* تكرار الفئة الثانية = F_2 = مجموع التكرارات - تكرار الفئة الأولى

* تكرار الفئة الثالثة = F_3 = مجموع التكرارات - تكرار الفئة الأولى - تكرار الفئة الثانية .

* وهكذا كما مبين في الجدول (٩) أدناه .

*جدول (٩) : يوضح التوزيع التكراري التجميعي التنازلي لأطوال نباتات القطن الموضح تكرارها في (جدول ٧) .

حدود الفئات	التكرار التجميعي التنازلي
٢١ فأكثر	٨٠
٤١ فأكثر	٧٩
٥١ فأكثر	٧٧
٦١ فأكثر	٧٢
٧١ فأكثر	٥٧
٨١ فأكثر	٢٢
٩١ فأكثر	١٢
١٠١ فأكثر	٠

*ملاحظة * أحيانا يعبر عن التكرار التجميعي التصاعدي او التنازلي بشكل تكرار تجميعي نسبي أو مئوي . وفي هذه الحالة .

* التكرار التجميعي النسبي لاي فئة = التكرار التجميعي لتلك الفئة / المجموع الكلي للتكرارات

* اما التكرار التجميعي المئوي = التكرار التجميعي النسبي X ١٠٠

⑤ التمثيل البياني : (Graphical Presentation)

ويشمل (P) - التمثيل البياني لجدول التوزيع التكراري * ويشمل

① - المدرج التكراري (Histogram)

وهو عبارة عن مستطيلات رأسية تمتد قواعدها على المحور الأفقي لتمثل أطوال الفئات (الحدود الحقيقية للفئات) بينما ارتفاعاتها على المحور العمودي لتمثل تكرارات الفئات .

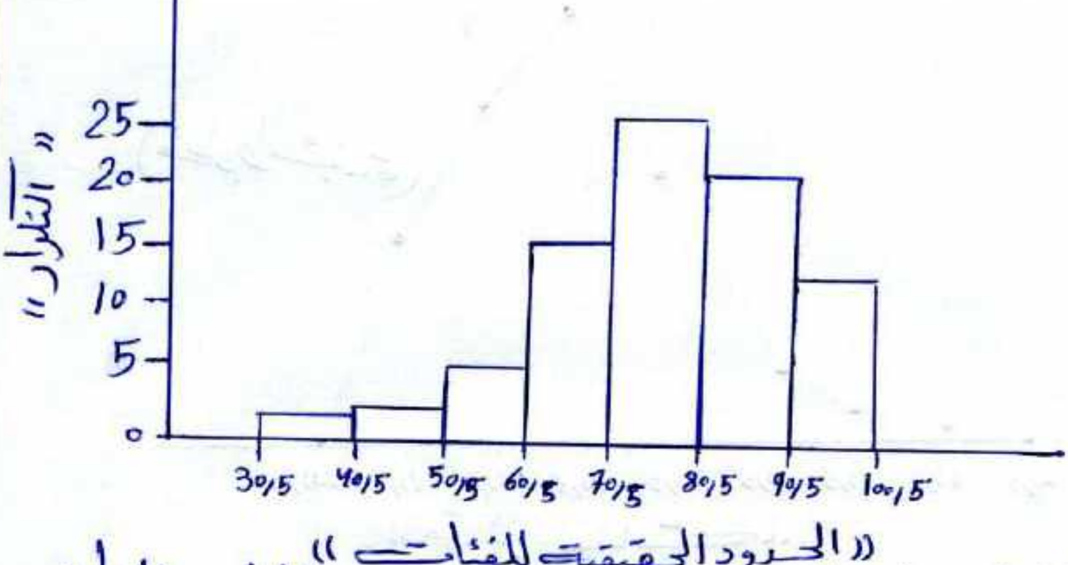
* ولرسم مدرج تكراري يتبع الخطوات التالية :-

١- رسم المحور الأفقي والمحور العمودي :

٢- تدريج المحور الأفقي الى أقسام متساوية بمقياس رسم مناسب بحيث يشمل جميع الحدود الحقيقية للفئات ويفضل ترك مسافة صغيرة بين نقطة الصفر والحد الادنى للفئة الأولى (فيما اذا كانت بداية الفئة الأولى لا تساوي صفر) . ويقسم المحور العمودي الى اقسام متساوية بحيث تشتمل على أكبر التكرارات .

٣- يرسم على كل فئة مستطيل رأسيا تمثل قاعدته طول تلك الفئة وارتفاعه يمثل تكرار تلك الفئة .

* والشكل (١) يوضح المدرج التكراري لجدول (٤) الموجود ص ١٦



* شكل (١) يوضح المدرج التكراري لأطوال نباتات القطن لجدول (٤) ص ١٦

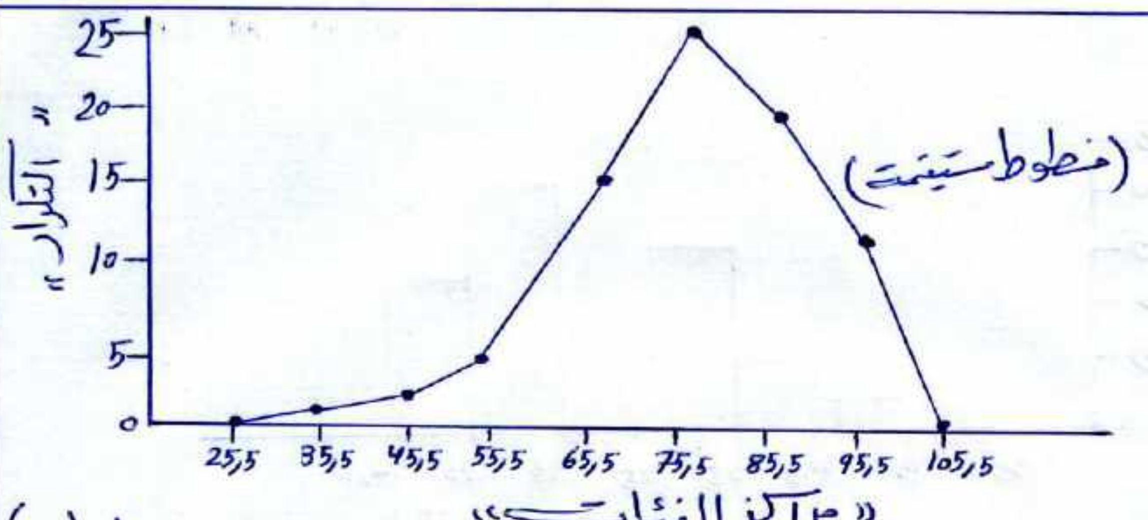
٢- المضلع التكراري (Frequency Polygon)

وهو عبارة عن نقاط ناتجة من مراكز الفئات التي تمثل المحور الأفقي و التكرارات التي تمثل المحور العمودي . ثم توصل هذه النقاط بخطوط مستقيمة منكسرة.

* ولرسم المضلع التكراري يتبع الخطوات التالية :-

١- رسم المحور الافقي والمحور العمودي
٢- تدرج المحور الأفقي الى اقسام متساوية بحيث يشمل على جميع مراكز الفئات . ويقسم المحور العمودي الى أقسام متساوية بحيث تشمل على اكبر التكرارات .
٣- وضع نقطة أمام مركز كل فئة ارتفاعها يعادل تكرار تلك الفئة .
٤- توصيل تلك النقاط بخطوط مستقيمة .

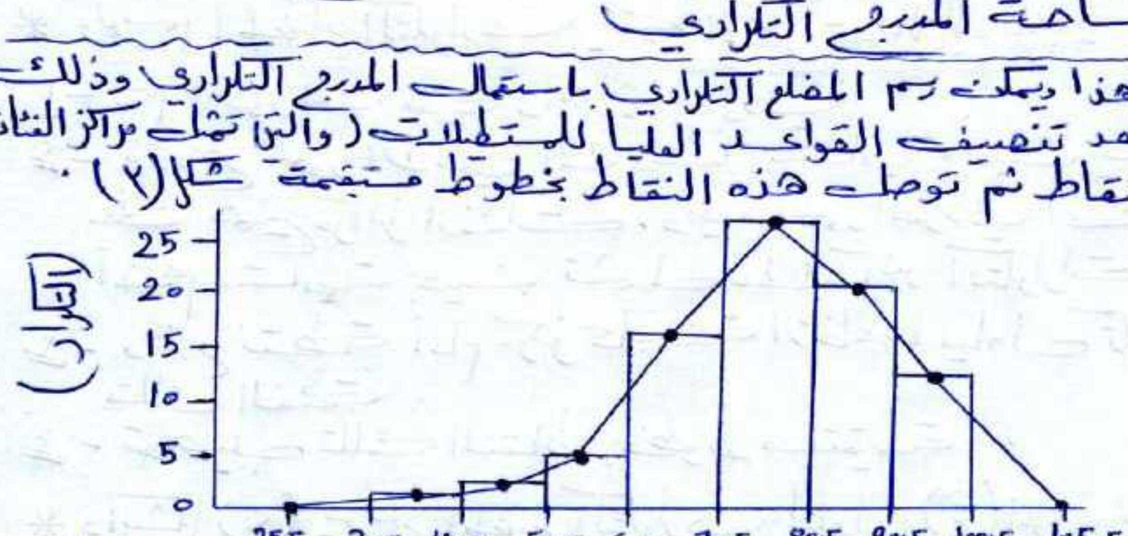
* والشكل (٢) يوضح المضلع التكراري لجدول (٤) الموجود ص ١٦



* شكل (١) المضلع التكراري لأطوال نباتات القطن لجدول (٤) ص ١٦

* ملاحظة :- عادة يقفل المضلع بأن نصل بداية المضلع بالمحور الأفقي بمركز فئة (خيالية) واقعة الى يسار أول فئة تكرارها صفراً . ونصل نهاية المضلع بالمحور الافقي بمركز فئة (خيالية) واقعة الى يمين آخر فئة تكرارها أيضاً صفراً . وبذلك يكون مساحة المضلع التكراري مساوية لمساحة المدرج التكراري

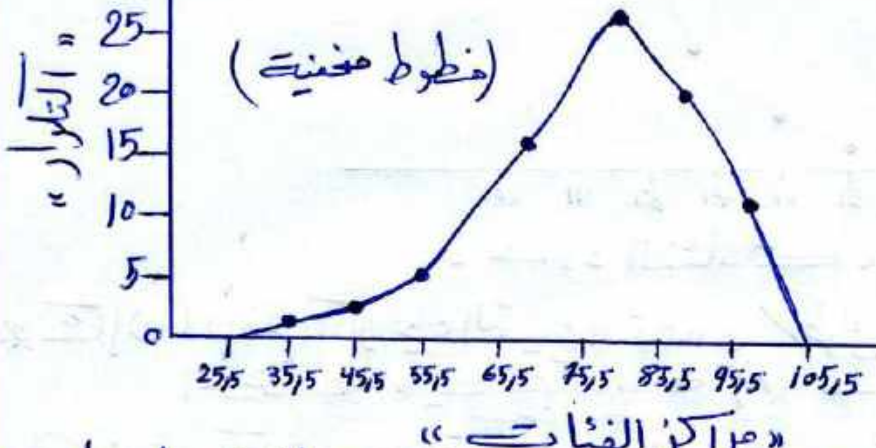
* هذا ويمكن رسم المضلع التكراري باستعمال المدرج التكراري وذلك بعد تنصيف القواعد العليا للمستطيلات (والتي تمثل مراكز الفئات) بنقاط ثم توصل هذه النقاط بخطوط مستقيمة شكل (٢) .



* شكل (٢) المدرج والمضلع التكراري لنباتات القطن

② المنحنى التكراري (Frequency curve).

وهو عبارة عن منحنى يمر بمعظم النقاط الواقعة على مراكز الفئات (المحور الأفقي) والتي ارتفاعها يمثل تكرارات تلك الفئات (المحور العمودي). وعادة يقفل المنحنى التكراري بأن نصل بدايته بالحد الأدنى للفئة الأولى ونهايته بالحد الأعلى للفئة الأخيرة. وهو مشابه من حيث الرسم للمضلع التكراري (شكل ٤).



* شكل (٤) المنحنى التكراري لأطوال نباتات القطن (جدول ٤) ص ١٦

* (ب) التمثيل البياني لجدول التوزيع التكراري التجميعي :- ويشمل

① - المضلع التكراري التجميعي التصاعدي :-

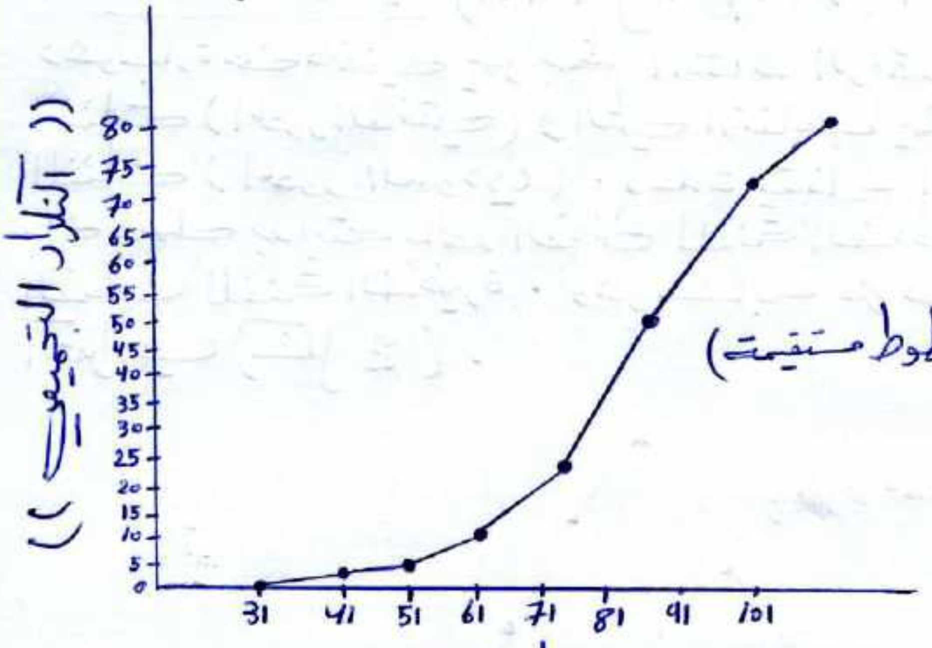
وهو عبارة عن خطوط مستقيمة منكسرة تصل بين نقاط واقعة ضمن حدود الفئات (المحور الأفقي) والتكرار التجميعي التصاعدي (المحور العمودي). ويتبع ما يلي عند الرسم :-

أ- رسم المحور الأفقي والمحور العمودي

ب- تدريج المحور الأفقي إلى أقسام متساوية تشمل على جميع حدود الفئات ويقسم المحور العمودي إلى أقسام متساوية تشمل على أكبر التكرارات التجميعية وهي المجموع الكلي للتكرارات.

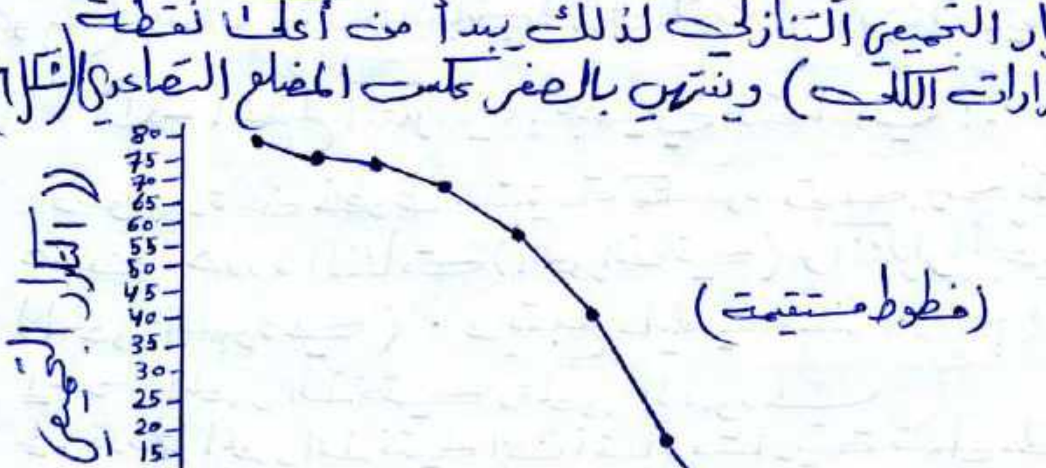
جـ- وضع نقطة أمام كل حد فئة ارتفاعها يعادل التكرار التجميعي التصاعدي لذلك الحد

٤- توصيل تلك النقاط بخطوط مستقيمة (شكل ٥)



* شكل (٥) المضلع التكراري التجميعي التصاعدي لأطوال نباتات القطن الموضحة في جدول (٨) ص ٢٢

٢) المضلع التكراري التجميعي التنازلي . ويرسم بنفس طريقة رسم المضلع التكراري التجميعي التصاعدي ماعدا كون إرتفاع النقاط هنا هو التكرار التجميعي التنازلي لذلك يبدأ من أعلى نقطة (مجموع التكرارات الكلية) وينتهي بالصفر عكس المضلع التصاعدي (شكل ٦)



* شكل (٦) المضلع التكراري التجميعي التنازلي لأطوال نباتات القطن الموضحة في جدول (٩) ص ٢٢

((مع تمنياتي

لكم

بالموفقية والنجاح))

((مدرسة المادة))

احصاء نظري

* مقاييس التمركز او التوسط. ((Measures of Central Tendency))

هي تلك المقاييس التي تبحث في تقدير قيمة تتمركز حولها أغلبية البيانات التابعة لظاهرة ما. وان هذه القيمة المتوسطة أو المتمركزة هي رقم واحد يعبر عن او يمثل جميع بيانات تلك الظاهرة التابعة لمجموعة ما.

* واهم مقاييس التوسط هي :-

① - الوسط الحسابي (المتوسط). The Arithmetic Mean
② - الوسط الهندسي. The Geometric Mean
③ - الوسط التوافقي. The Harmonic Mean
④ - الوسط التربيعي. The Quadratic Mean
⑤ - الوسيط. The Median
⑥ - المنوال. The Mode

* ① الوسط الحسابي. (The Arithmetic Mean)

الوسط الحسابي او المتوسط لقيم متغير ما هو القيمة الناتجة من قسمة مجموع تلك القيم على عددها ويرمز له بالرمز ($\bar{y}$)

* طرق حساب الوسط الحسابي :-

* (أ) عندما تكون البيانات الاولية (غير مبوبة) يكون :-

الوسط الحسابي = $\bar{y} = \frac{\sum y_i}{n}$

(مثال)* البيانات التالية تمثل كمية المطر الساقطة سنوياً (بالمليتر) على مدينة الموصل خلال فترة خمس سنوات هي (٥٠، ٢٠، ٤٥، ٢٨، ٤٠ ملم) فما متوسط سقوط المطر خلال هذه الفترة.

الحل :-
$$\bar{y} = \frac{\sum y_i}{n} = \frac{520+350+450+380+400}{5} =$$
$$= 420 \ mm$$

* (ب) اذا كان لكل قيمة من المشاهدات (y_i) وزن خاص يتناسب مع أهميتها (w_i) فان الوسط الحسابي الموزون :- يساوي .

$$\bar{y} = \frac{\sum w_i y_i}{\sum w_i}$$

الوسط الحسابي الموزون

* (مثال) * . القيم التالية تمثل نتائج امتحان أحد الطلبة في درس الاحصاء . علماً بان لكل إمتحان وزناً أو أهمية او نسبة معينة .
* المطلوب إيجاد الوسط الحسابي أو معدل الطالب :-

الإمتحان	الدرجة (y_i)	اهميتها أو نسبتها أو وزنها (w_i)	($w_i y_i$)
الأول	٧٠	٪١٠	٧٠٠
الثاني	٦٠	٪٢٠	١٢٠٠
الثالث	٧٥	٪١٠	٧٥٠
الرابع	٥٥	٪٥٠	٢٧٥٠
		$\sum w_i = ١٠٠$	$\sum w_i y_i = ٦٠٠٠$

الحل :- الوسط الحسابي يساوي
$$\bar{y} = \frac{\sum w_i y_i}{\sum w_i} = \frac{6000}{100} = 60$$

* (ج) * عندما تكون البيانات الأولية (مبوبة) في جدول توزيع تكراري فان الوسط الحسابي يساوي :-

$$\bar{y} = \frac{\sum f_i y_i}{\sum f_i}$$

حيث إن $f_i =$ التكرار

$y_i =$ مركز الفئة

* خطوات إيجاد الوسط الحسابي من بيانات مبوبة هي كالاتي :-

١- تعيين مراكز الفئات (y_i) :
٢- ضرب مركز كل فئة بمقدار تكرارها ($f_i y_i$) .
٣- قسمة مجموع (حاصل ضرب مركز كل فئة × تكرارها) على مجموع التكرارات .

(مثال) إستخرج الوسط الحسابي لأطوال النباتات من جدول التوزيع التكراري التالي :-

الفئات	التكرار (f_i)	مراكز الفئات (y_i)	التكرار × مركز الفئات $y_i * f_i$
٢١ - ٤٠	١	٣٠/٥	٣٠/٥
٤١ - ٥٠	٢	٤٥/٥	٩١/٠
٥١ - ٦٠	٥	٥٥/٥	٢٧٧/٥
٦١ - ٧٠	١٥	٦٥/٥	٩٨٢/٥
٧١ - ٨٠	٢٥	٧٥/٥	١٨٨٧/٥
٨١ - ٩٠	٢٠	٨٥/٥	١٧١٠/٠
٩١ - ١٠٠	١٢	٩٥/٥	١١٤٦/٠
المجموع	$\sum f_i =$ ٨٠		$\sum f_i y_i =$ ٦١٣٠

*الحل

$$\bar{y} = \frac{\sum f_i y_i}{\sum f_i} = \frac{6130}{80} = 76.62 \text{ cm}$$

*((خواص الوسط الحسابي)) *

(أولاً) * مجموع إنحرافات القيم عن وسطها الحسابي = صفراً

أي إن :-

$$\sum (y_i - \bar{y}) = 0$$

* والمثال التالي يبرهن ذلك :-

y_i	$(y_i - \bar{y})$
6	$6-5=1$
4	$4-5=-1$
8	$8-5=3$
2	$2-5=-3$
$\sum y_i = 20$ $\therefore \bar{y} = \frac{20}{4} = 5$	$\sum (y_i - \bar{y}) = Zero$

(ثانيا) * مجموع مربعات الانحرافات عن الوسط الحسابي هي أقل ما يمكن (أي أقل من مجموع مربعات الانحرافات عن أي قيمة غير الوسط الحسابي نفسه) أي إن :-

$$\sum (y_i - \bar{y})^2 = \text{Less value} \text{ (أقل ما يمكن)}$$

* والمثال التالي يبرهن ذلك :-

y_i	$(y_i - \bar{y})$	$(y_i - \bar{y})^2$
6	$6-5=1$	1
4	$4-5=-1$	1
8	$8-5=3$	9
2	$2-5=-3$	9
$\sum y_i = 20$ $\bar{y} = 5$	$\sum (y_i - \bar{y}) = Zero$	$\sum (y_i - \bar{y})^2 = 20$ أقل ما يمكن

* فلو طرحنا مثلاً هذه القيم في الجدول السابق وهي (6 - 4 - 8 - 2) من أي قيمة غير الوسط الحسابي ($\bar{y} = 5$) فإن مجموع مربعات الانحرافات ستكون قيمتها أكبر. ((مثلاً لو طرحت من القيمة (٣) فيكون الناتج كالآتي)).

y_i	$(y_i - 3)$	$(y_i - 3)^2$
6	$6 - 3 = 3$	9
4	$4 - 3 = 1$	1
8	$8 - 3 = 5$	25
2	$2 - 3 = -1$	1
		$\sum (y_i - 3)^2 = 36$

∴ (36) أكبر من الناتج السابق (20)

(ثالثاً) * عند إضافة أو طرح عدد ثابت (C) إلى أو من قيم المشاهدات فإن :-

* الوسط الحسابي للقيم الجديدة = الوسط الحسابي للقيم الأصلية ($\mp$) العدد الثابت (C)

أي إن :-

$$X_i = y_i \mp C$$
$$\bar{X} = \bar{y} \mp C$$

* والمثال التالي يوضح ذلك في حالة إضافة مثلاً قيمة ثابتة ولتكن (٣) أي إن :-

.((C = 3))

y_i	C	X_i
4	3	4+3 = 7
5	3	5+3 = 8
4	3	4+3 = 7
7	3	7+3 = 10
$\bar{y} = 20/4 = 5$		$\bar{X} = 32/4 = 8$

$\bar{X} = 5+3=8$

أي $\bar{X} = \bar{y}+3=8$

* ونفس الحالة في حالة الطرح حيث بأن

$\bar{X} = \bar{y} - C$ حيث أن (C = قيمة ثابتة)

(رابعاً) * إذا ضربت أو قسمت كل قيمة من قيم المشاهدات على قيمة ثابتة (K) تلك فإن :-

* الوسط الحسابي للقيم الجديدة = الوسط الحسابي للقيم الأصلية × القيمة الثابتة (K)

أو * الوسط الحسابي للقيم الجديدة = الوسط الحسابي للقيم الأصلية ÷ القيمة الثابتة (K)

أي إن :-

$$\left(\begin{array}{l} X_i = y_i K \\ \bar{X} = \bar{y} K \end{array} \right)$$

«في حالة الضرب»

أو

$$\left(\begin{array}{l} X_i = y_i / K \\ \bar{X} = \bar{y} / K \end{array} \right)$$

«في حالة القسمة»

* فمثلاً عند ضرب كل قيمة من قيم المشاهدات بقيمة ثابتة ولتكن مثلاً (K=3) فالناتج يكون كما يلي :-