أولا : التوزيعات التكرارية

مدخل: تتكون المعلومات التي يتم جمعها عن طريق التجارب المخبرية أوعن طريق المعاينة الميدانية في أغلب الأحيان من مجموعة أرقام، هذه الأرقام يصعب علينا التعبير عنها إذا لم نكن على دراية كافية بحقائقها وبالتالي ينبغي معالجة ومناقشة ترتيب هذه المعلومات في صورة توزيع تكراري وتمثيل بياني. فالتوزيع التكراري يهدف إلى تبسيط العمليات الإحصائية وذلك بتبويب البيانات في صورة مناسبة تيسر معالجتها بسرعة ودقة , كما يهدف أيضا إلى إعادة صياغة البيانات العددية صياغة علمية , ولهذا تعتبر التوزيعات التكرارية نقطة البدء في كل العمليات الإحصائية وتسمى بالتوزيعات التكرارية لأنها تشكل بحساب عدد المرات التي تتكرر فيها نفس القيمة. **مثال :** من خلال رمينا لقطعة نقود عشر مرات تحصلنا على النتائج التالية :

H, H, T, H, T, H, H, H, T, H (حيث: H تمثل الصورة وT تمثل الكتابة)أي تكررت الصورة 7مرات والكتابة 3 مرات وهو يدل على أن الصورة كانت أكثر شيوعا في الظهور،ويمكن تلخيص ذلك في جدول كما يلي :

وجه العملة (x_i)	التكرار (f_i)
H	7
T	3
المجمــوع	10

وتنظيم المعومات بهذه الطريقة يمكن تسميتها بالتوزيع التكراري .

1/ الفئات التكرارية : لنفرض أن لدينا علامات مقياس الإحصاء الوصفي لـ 120 طالبا من قسم علم الاجتماع، نريد أن نحصل من هذه العلامات على فكرة دقيقة دقة معقولة تبين لنا كيفية توزيع هذه النتائج, سنواجه صعوبة كبيرة في ذلك ما لم نقم بتكثيفها نوعا ما، وذلك بتبويب نتائج الامتحان على شكل مجموعات، عندئذ يمكن رسم منحنى التوزيع والذي يعرفنا بالكثير من الحقائق عن المجموعة الأصلية، وهذا التكثيف يفيدنا أيضا في تبسيط حساب قيم بعض المقاييس الإحصائية المهمة. ولا تواجهنا مسالة تبويب البيانات والأرقام على شكل فئات إلا في حالة المتغيرات المتصلة كالوزن والطول ودرجة الحرارة والزمن وغيرها من المتغيرات القادرة على أخذ أية قيمة في مجال ما من القيم. ففي المثال المتعلق بعلامات 120 طالبا نحتاج إلى إختيار مجموعة من الفئات للعلامات ثم وضع كل علامة في فئتها المناسبة، فمثلا إذا اخترنا مجال الفئة بطول مداه 5علامات، فكل العلامات التي

تنحصر بين 10 و15 شاملة العلامة الأولى دون الثانية توضع كلها في نفس الفئة، وتسمى القيمتان 10 و15 حدي هذه الفئة. وبعد تعيين حدود الفئات يسهل استخراج قائمة الفئات لتغطي في مجملها على كل القياسات، ومتى تم ذلك تقول أننا قد بوبنا البيانات في جدول تكراري. ومن المستحسن عند تبويب البيانات استخدام ما بين 5 و 15 فئة، كما يفضل أن تكون هذه الفئات متساوية في الطول لتسهيل المقارنة وكذا حساب بعض المقاييس الإحصائية.

2/ المدى العام: يرمز له بالرمز E_t (Etendue) ويمثل الفرق بين أكبر قيمة وأصغر قيمة ضمن البيانات ونكتب رياضيا: $E_t = x_n - x_1$ حيث: x_1, x_n أكبر قيمة وأصغر قيمة ضمن البيانات.

3/ مدى الفئة: يرمز له بالرمز C أو K ويحسب وفقا للعبارة الرياضية التالية:

$C = \frac{E_t}{N}$ E_t: المدى العام وN: عدد الفئات المقترح.

وقد كان للعالم الإحصائي ستورجز السبق في إيجاد نموذج من قوانين تحديد مدى الفئة لما يكون عدد البيانات (الأرقام) ـ المراد تبويبها على شكل فئات ـ محصورا بين 100 و 1000. ويعطى قانون ستورجز في تحديد مدى الفئة بالشكل التالي: $C = \frac{E_t}{1 + (3,322)\log n}$ حيث n:عدد البيانات (الأرقام)، و $N = 1+(3,322)\log n$:عدد الفئات.

ملاحظة: يقدم لنا ستورجز في هذه الحالة العدد النموذجي للفئات الذي ينبغي التقيد به في بعض الحالات.

مثال تطبيقي لقانون ستورجز : نفرض أن أدنى علامة سجلت لدى مجموعة الـ 120 طالب السابقة الذكر هي 0 و أعلاها هي 20، فيكون مدى الفئة هو :

$C = E_t/1+(3,322)\log n$
$= (X_n - X_1) / 1+(3,322)\log n$
$= (20-0) /1+(3,322)\log 120$
$= 1.18 (=1)$

4/ مركز الفئة : يرمز له بالرمز X_C وهو منتصف الفئة يعبر عنها أو يمثلها أو يلخصها ويلجأ إليه لتسهيل العمليات الحسابية ويحسب كما يلي :

$X_c = \frac{L_o + L_1}{2}$ حيث: L_0 :الحد الأدنى للفئة و L_1 :الحد الأعلى لها

كما يحسب بالطريقة التالية: $X_c = L_O + \frac{C}{2}$ أو $X_C = L_1 - \frac{C}{2}$ حيث C: يمثل مدى الفئة.

5/ التكرار المتجمع الصاعد والنازل:

يرمز إلى التكرار المتجمع الصاعد(المطلق) بالرمزF.c. ، ويساوي التكرار المتجمع الصاعد (المطلق) لكل فئة إلى مجموع تكرارات تلك الفئة وتكرارات بقية الفئات التي تسبقها. أما التكرار النسبي المتجمع الصاعد لكل فئة فيرمز له بالرمز c.c.r.F ويحسب بقسمة التكرار المتجمع الصاعد (المطلق) لتلك الفئة على المجموع الكلي للتكرارات وضرب الحاصل في 100٪ ونكتب : $F.r.c.c = \frac{F.c.c}{\sum f_i} \times 100\%$

وبنفس الطريقة يمكن حساب التكرارات المتجمعة النازلة المطلقة والنسبية حيث نحسب الأولى لكل فئة وتكون مساوية لمجموع تكرارات تلك الفئة مع تكرارات الفئات التي تأتي بعدها ونكتب :

$$F.r.c.d = \frac{F.c.d}{\sum f_i} \times 100\%$$

*** ملاحظة:**

من أهداف إعداد جدول التكرار المتجمع الصاعد (المطلق والنسبي) أنه يحدد لنا عدد المفردات التي تقل قيمة المتغير المراد دراسته لديهم عتبة معينة وكذا نسبة هذه المفردات إلى العدد الكلي.

وأيضا من أهداف إعداد جدول التكرار المتجمع النازل (المطلق والنسبي) أنه يحدد لنا عدد المفردات التي تزيد قيمة المتغير (الرقم) المراد دراسته لديهم عتبة معينة وكذا نسبة هذه المفردات إلى العدد الكلي