مفاهيم أساسية حول علم الاحصاء

أولا – الإحصاء الوصفي Statistique descriptive: وهو ذلك الفرع من الإحصاء الذي يهتم بتعميم النتائج بعد جمع جزء منها، وترتيبها وتلخيصها وعرضها في صورة جداول ثم تمثيلها بيانيا وتحليلها وتفسيرها لاتخاذ القرارات المناسبة.

أما الإحصاء الاستدلالي فهو يختص بإجراء التنبؤات والتقديرات والاستنتاجات.

ثانيا – مصادر جمع البيانات الإحصائية: يوجد مصدران لجمع البيانات هما:

1– مصدر تاريخي: وهو ما يؤخذ من السجلات المحفوظة.

مثل: – سجلات المواليد والوفيات.

- البيانات الواردة من رسائل الماستر والماجستير والدكتوراه.
- البيانات التي يتم نشرها من خلال المنظمات العالمية.

2– مصدر ميداني: ويتم فيه جمع البيانات مباشرة عن طريق إتصال الباحث بالوحدة محل الدراسة، ويتم جمع البيانات ميدانيا عن طريق الوسائل التالية:

أ– **المقابلات الشخصية**: يستطيع الباحث بهذه الطريقة تحقيق أعلى درجات الدقة في جمع البيانات. إلا أن هذه الطريقة وبالرغم مما تمتاز به من دقة المعلومات قد تكون مكلفة. وخاصة في حالة العينات كبيرة الحجم.

ب– **الاستمارة الاحصائية (الاستبيان)**: ويراعى في هذه الاستمارة الشروط التالية:

- أن تكون الأسئلة واضحة وسهلة.
- أن تكون الاستمارة قصيرة.
- يجب التأكيد على سرية البيانات للشخص محل الدراسة حتى لا تكون إجابته بعيدة عن الواقع.
- يجب أن تحقق الاستمارة الأهداف محل الدراسة.

ثالثا – التعريف بالمصطلحات والمفاهيم الإحصائية العامة:

1 – المجتمع Population: وهو المجموعة الكاملة من الافراد أو العناصر محل الدراسة والتي لها خصائص مشتركة.

وكل شخص من المجتمع يدعى وحدة احصائية (Unité Statistique) وهو العنصر الأساسي عند القيام بتجربة ما

مثال: لو كانت دراستنا حول: أطول طلبة المركز الجامعي ميلة.

في هذه الحالة المجتمع الاحصائي هو جميع طلبة المركز.

– قد يكون المجتمع الاحصائي محدودا حيث يمكن حصر عدد أفراده. **مثل**: عدد الطلاب في الكلية.

أو غير محدود **مثل**: عدد النجوم في السماء، عدد حبات القمح المحصودة في مزرعة ما.

2 – العينة Echantillon: وهي جزء من المجتمع التي يتم اختيارها وتجرى عليها القياسات. وبعد استخلاص النتائج تعمم على المجتمع الذي سحبت منه.

مثال: في المثال السابق إذا أخذنا طلبة معهد العلوم الاقتصادية. فيمكن اعتبار هذه المجموعة عينة من المجتمع الذي هو محل الدراسة.

ملاحظة: لماذا يتم اختيار العينة؟؟

يكون أخذ العينة أفضل من دراسة المجتمع كله للأسباب التالية: أقل وقت، أقل تكلفة، الحسابات تكون أدق، إمكانية الاتصال بكل مفردات العينة (عمال مثلا).

3 – الوحدة الاحصائية Unité Statistique: وهي الكائن الواحد أو الوحدة الأساسية التي يتكون منها المجتمع الاحصائي.

مثال: في المثال السابق: أطوال طلبة المركز الجامعي ميلة.

الوحدة الاحصائية هي الطالب الواحد

4 – الصفة أو الخاصية Caractère Statistique: وهي الخاصية التي تميز الفرد في المجتمع

(أو القاسم المشترك لكل وحدات المجتمع المدروس).

مثال: ماهي الخاصية المدروسة في المثال السابق؟

الخاصية المدروسة هي الطول.

رابعا– أنواع البيانات (المتغيرات) الإحصائية

تنقسم المتغيرات الإحصائية إلى نوعين:

1 – متغيرات وصفية (نوعية) Variable qualitative: وهي البيانات التي تكون لها صفات معينة، لايمكن قياسها ولا يعبر عنها بصورة عددية. وتنقسم المتغيرات النوعية إلى قسمين هما:

أ – متغيرات وصفية إسمية qualitative nominale : وهي بيانات غير رقمية لا تتأثر بأي ترتيب منطقي.

أمثلة: – الجنسية (جزائري، تونسي،...) وصفي إسمي

- الجنس (ذكر، أنثى) وصفي إسمي ذو فرعين
- اللون (أبيض، أسود،...)
- الحالة العائلية (متزوج، أعزب، أرمل، مطلق)

ب – متغيرات وصفية ترتيبية qualitative ordinale : وهي بيانات غير رقمية، تتكون من مستويات تتبع ترتيبا منطقيا معد مسبقا أو متفق عليه.

أمثلة: – المستوى التعليمي (ابتدائي، متوسط، ثانوي، جامعي)

- مقياس اللباس (s.m.l.xl) وصفي ترتيبي

2 – متغيرات كمية (عددية) Variable quantitatives: وهي بيانات رقمية، يمكن التعبير عنها في صورة عددية. مثل: الطول، الوزن، عدد الأفراد في الأسرة...الخ.

وتنقسم بدورها إلى قسمين هما:

أ – متغيرات كمية مستمرة (متصلة) quantitative Continue: وهي بيانات يمكن أن تأخذ أي قيمة عددية في مدى معين (تأخذ قيما غير محدودة وغير منتهية).

أمثلة عن المتغيرات الكمية المستمرة: الطول، الوزن، السرعة، كميات الأمطار...الخ.

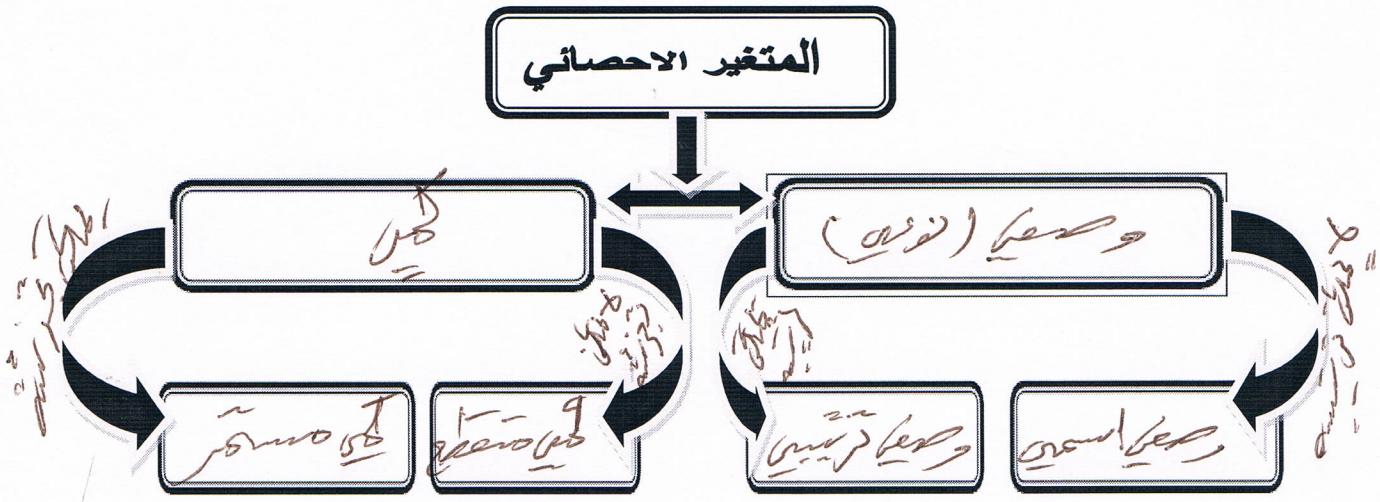
فإذا كان × هو متغير الطول فإن × يمكن أن يأخذ القيم 15متر، 11.3 متر، 14.75 متر، أي أن المتغير × يمكن أن يأخذ أي قيمة في فترة زمنية معينة.

ب – متغيرات كمية متقطعة (منفصلة) quantitative discret: وهي البيانات التي تأخذ أعداد صحيحة، لايمكن تجزئتها، وهي محدودة يمكن عدها وحصرها بالعد.

أمثلة عن المتغيرات الكمية المتقطعة: عدد الأفراد في الأسرة، عدد الغرف في المسكن، عدد الأهداف المسجلة، عدد السيارات المباعة...الخ.

فمثلا إذا كان × متغير يمثل عدد أفراد الأسرة، فإنه يمكن أن يأخذ القيم 1، 2، 3، 4، 5 ... ولا يمكن أن يأخذ × القيم 1.5، 3.25، 5.17.

مخطط يبين أنواع البيانات أو المتغيرات



تمرين: حدد كلا من المجتمع الاحصائي، الوحدة الاحصائية، الخاصية المدروسة ونوعيتها بدقة لكل من:

1- الأجور السنوية لأساتذة الجامعة.

2- عدد الأولاد في كل عائلة لأحد الأحياء السكنية.

3- كمية الأمطار المتساقطة خلال فصول السنة في ولاية ميلة.

4- جنسية مجموعة من السياح.

5- وزن اللاعبين.

6- توزيع عينة من 30 فرد حسب المستوى التعليمي.

المحور الأول
تنظيم وعرض البيانات

تمهيد:

يعتبر تنظيم وعرض البيانات الاحصائية أول مرحلة للتحليل الاحصائي، وتتقيد طريقة تنظيم وعرض البيانات على نوع هذه البيانات سواء كانت وصفية (نوعية) أو كمية. وعليه يمكن تنظيم وعرض البيانات إما عن طريق تصميم التوزيعات أو الجداول التكرارية أو باستعمال الرسومات البيانية. والهدف من كل ذلك تقديم البيانات بطريقة مبسطة ومختصرة.

أولا– التوزيعات والجداول التكرارية:

يعتبر تبويب (تفريغ) وعرض البيانات في جداول من الخطوات الأساسية للحصول على المعلومات.

الهدف من استخدام الجداول هو تلخيص البيانات وتسهيل فهمها ودراستها، لذا يجب أن يكون الجدول المسمى بجدول التوزيعات التكرارية بسيط وغير غامض بقدر الإمكان.

– حسب نوعية المتغير (وصفي أو كمي) يمكن تصميم الجداول التالية:

1– الجداول التكرارية للبيانات الوصفية (النوعية): وهو عبارة عن جدول يحتوي في صورته البسيطة على العناصر التالية:

أ– نوع المتغير: يتم رصد كل صيفات المتغير في العمود الأول.

ب– التكرار المطلق: وهو يمثل عدد المرات التي يتكرر فيها نفس المتغير ونرمز له بالرمز n_i ويكون في العمود الثاني.

– مع مراعاة مجموع الأعداد الموجودة في عمود التكرارية يساوي عدد المفردات أو مجموع التكرارات

$\Sigma n_i = n_1+n_2+n_3+.....+n_K$

المتغير	التكرار المطلق n_i
الصفة 1	n_1
الصفة 2	n_2
. . .	
الصفة 3	n_K
المجموع Σ	Σn_i

ج– التوزيع التكراري النسبي والمئوي:

– **التكرار النسبي**: ونرمز له بالرمز fi

وهو حاصل قسمة تكرار كل صفة أو فئة على مجموع التكرارات أي $f_i = \frac{n_i}{\Sigma n_i}$

ملاحظة: دائما مجموع التكرار النسبي يساوي 1 $\Sigma f_i = 1$

– **التكرار النسبي المئوي**: ونرمز له بالرمز $f_{i\%}$

$$f_{i\%} = f_i \times 100$$

أو

$$f_{i\%} = \frac{n_i}{\Sigma n_i} \times 100$$

– **التوزيع التكراري التجميعي الصاعد**: ونرمز له بالرمز $Ni \nearrow$

ويحسب كالآتي:

$N_1 \nearrow = n_1$

$N_2 \nearrow = n_1 + n_2$ » $N_2 \nearrow = N_1 \nearrow + n_2$

.

$N_k \nearrow = \Sigma n_i$

ملاحظة: التكرار التجميعي الصاعد الأول دائما يساوي التكرار المطلق الأول $N_1 \nearrow = n_1$ والتكرار التجميعي الصاعد الأخير دائما يساوي مجموع التكرارات $N_k \nearrow = \Sigma n_i$

– **التوزيع التكراري التجميعي النسبي الصاعد**: ونرمز له بالرمز $Fi \nearrow$

ويحسب بنفس الطريقة أو $Fi \nearrow = \frac{N_i \nearrow}{\Sigma n_i}$

– **التوزيع التكراري التجميعي النسبي المئوي الصاعد**: ونرمز له بالرمز $Fi_{\%} \nearrow$

$$Fi_{\%} \nearrow = Fi \nearrow \times 100$$

– **التوزيع التكراري التجميعي النازل**: ونرمز له بالرمز $Ni \swarrow$

ويحسب كالآتي:

$N_1 \swarrow = \Sigma n_i$

$N_2 \swarrow = \Sigma n_i - n_1$ » $N_2 \swarrow = N_1 \swarrow - n_1$

.

.

$N_k \swarrow = n_k$

ملاحظة: التكرار التجميعي النازل الأول دائما يساوي مجموع التكرارات $N_1 \swarrow = \Sigma n_i$ والتكرار التجميعي النازل الأخير دائما يساوي التكرار المطلق الأخير $N_k \swarrow = n_k$

– **التوزيع التكراري التجميعي النسبي النازل**: ونرمز له بالرمز $Fi \swarrow$

– **التوزيع التكراري التجميعي النسبي المئوي النازل**: ونرمز له بالرمز $Fi_{\%} \swarrow$

ويحسب كالآتي: $Fi_{\%} \swarrow = Fi \swarrow \times 100$

تمرين: سئل 20 شاب عن رأيهم في آداء مقدم البرنامج التلفزيوني الجديد وكان على كل شاب أن يختار واحد من خمسة إجابات وهي – ضعيف، – تحت لمتوسط، – متوسط، – فوق المتوسط، – ممتاز.

فكانت الإجابات كالآتي:

ممتاز، تحت المتوسط، متوسط، متوسط، فوق المتوسط، ممتاز، ضعيف، فوق المتوسط، ممتاز، تحت المتوسط، فوق المتوسط، ممتاز، فوق المتوسط، تحت المتوسط، فوق المتوسط، متوسط، تحت المتوسط، ضعيف، ممتاز، فوق المتوسط.

المطلوب: – أعرض البيانات في شكل جدول تكراري.

- كون التكرار النسبي الصاعد والنازل.
- كون التكرار التجميعي الصاعد والنازل.
- كون التكرار التجميعي النسبي الصاعد والنازل.
- حساب النسبة المئوية للشباب الذي كان رأيهم في أن الآداء كان أعلى من درجة متوسط (أي فوق المتوسط+ ممتاز).

الحل:

المتغير	التكرار المطلق n_i	التكرار النسبي fi	التكرار النسبي المئوي $fi_{\%}$	التكرار التجميعي الصاعد Ni ↗	التكرار التجميعي النازل Ni ↙	التكرار التجميعي النسبي الصاعد Fi ↗
ضعيف	2	0,10	10	2	20	0,10
تحت المتوسط	4	0,20	20	6	18	0,30
متوسط	3	0,15	15	9	14	0,45
فوق المتوسط	6	0,30	30	15	11	0,75
ممتاز	5	0,25	25	20	5	1
المجموع	20	1	100	—	—	—

$fi = \frac{n_i}{\Sigma n_i}$

$f_1 = \frac{2}{20} = 0,10$

$fi_{\%} = fi \times 100$

$fi_{\%} = 0,10 \times 100 = 10$

- النسبة المئوية للشباب الذي كان رأيهم في أن الآداء كان أعلى من درجة متوسط هي $30+25=55\%$

2- الجداول التكرارية للبيانات الكمية:

أ- الجدول التكراري للبيانات الكمية المتقطعة (المنفصلة): يكون الشكل العام للجدول التكراري في حالة البيانات الكمية المتقطعة كالآتي:

- في العمود الأول: نضع قيم المتغير X بصورة فردية ومرتبة ترتيبا تصاعديا.
- العمود الثاني: يمثل عدد المرات التي تتكرر فيها نفس القيمة.

قيم المتغير	التكرار المطلق n_i
X_1	n_1
X_2	n_2
. . .	. . .
X_K	n_K
المجموع Σ	Σn_i