تقنيات الاستقصاء

عموميات حول الإحصاء

الإحصاء

علم مكون من مجموعة من الأساليب والطرق الرياضية التي تستخدم في جمع وتنظيم وعرض وتحليل البيانات ثم الحصول على استدلالات واستنتاجات تستخدم في اتخاذ القرار.

يصنف إلى:

إحصاء وصفي يتناول طرق جمع، تنظيم، تلخيص، وعرض البيانات في صورة مبسطة.

إحصاء استدلالي ويهتم بطرق الوصول إلى نتائج أو توقعات عن المجتمعات المدروسة من خلال دراسة عينة من هذه المجتمعات

المجتمع

هو جميع العناصر المشتركة في الصفة التي تهم الباحث في دراسته

يقسم المجتمع إلى:

مجتمع محدود عدد محدود من الأشياء أو الأفراد

مجتمع غير محدود عدد من الأشياء غير منته

المجتمع المستهدف (المجتمع النظري)

يتمثل في جميع العناصر التي تحمل الخاصية التي تهم البحث، بغض النظر عن إمكانية الوصول إليهم أم لا

المجتمع المتاح (مجتمع الدراسة)

جزء من المجتمع النظري الذي يمكن للباحث الوصول إليه فعليا، ويختار منه العينة

إطار المعاينة قائمة تحتوي على جميع عناصر المجتمع تستخدم في اختيار عينة للدراسة

العينة جزء من المجتمع يتم اختيارها بحيث تمثل جميع أفراد المجتمع تمثيلا جيدا

مثال إذا كنا نرغب في دراسة سلوك القراءة بين الطلبة الجامعيين في بلد ما:

المجتمع المستهدف :جميع الطلبة الجامعيين في ذلك البلد

مجتمع الدراسة :الطلبة في جامعة معينة يمكن الوصول إليهم، مثل جامعة بسكرة.

إطار المعاينة :قائمة الطلبة المسجلين في جامعة بسكرة، التي يمكننا استخدامها لاختيار عينة منهم لإجراء البحث.

المعاينة : هي الطريقة التي يتم بها اختيار عينة مناسبة تسمح بتحديد خصائص أو مواصفات معينة عن مجتمع الدراسة .

على من تريد التعميم

المجتمع المستهدف
Target population

ما هو المجتمع الذي يمكنك الوصول اليه

المجتمع المتاح
Accessible Population

كيف يمكنك الوصول الى ذلك

اطار المعاينة
Sampling Frame

من سيكون في دراستك

العينة
Sample

العنصر هو عبارة عن عضو من أعضاء مجتمع الدراسة.

المفردة هي عبارة عن عضو من أعضاء العينة.

الظاهرة هي صفة لعناصر تختلف من عنصر لآخر في الشكل أو النوع أو الكمية

المعلمة قيمة رقمية (المتوسط الحسابي مثلا) لمتغير ما يتم الحصول عليها من دراسة شاملة للمجتمع. في الغالب تكون غير معروفة

الإحصاءة قيمة رقمية (المتوسط الحسابي مثلا) لمتغير ما يتم الحصول عليها من دراسة العينة المختارة.

الصفر المطلق(الحقيقي) هو الصفر الذي يعني الانعدام

الصفر الاعتباري(النسبي) يستخدم كنقطة بدء ، ولا يعني الانعدام

المتغير

هو الصفة تحت الدراسة

هو الشيء الذي يمكن أن يأخذ قيما مختلفة في الظروف المختلفة

أنواع المتغيرات 1 – متغيرات نوعية (كيفية يعبر عنها في شكل صفات

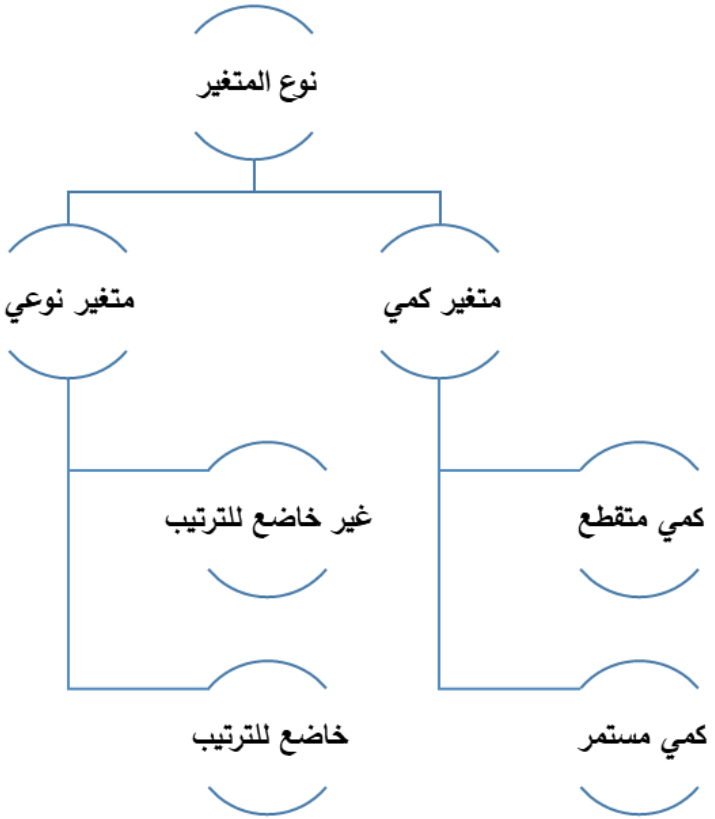
متغير نوعي خاضع للترتيب

متغير نوعي غير خاضع للترتيب

2. متغيرات كمية (عددية) يعبر عنها في صورة عددية وتنقسم إلى:

متغير متقطع يأخذ أعداد صحيحة

متغير متصل يمكن أن يأخذ أي قيمة بين قيمتين صحيحتين



مقاييس النزعة المركزية

- عدد من المتوسطات للتعبير عن القيمة المتوسطة للمتغيرات
- تختلف باختلاف الغرض الذي تستخدم فيه، وطبيعة البيانات المحسوبة منها
- تصف الظاهرة المدروسة مثل الجداول الإحصائية، إلا أنها أكثر اختصارا وأكثر فائدة
- تمكننا من المقارنة بين مجموعة من القيم ومجموعة أخرى، أو بين ظاهرة وأخرى.
- أهم هذه المتوسطات وأكثرها استخداما: المتوسط الحسابي، الوسيط، والمنوال.

1– المتوسط الحسابي

$$\overline{X} = \frac{x_1 + x_2 + x_3 + ... + x_n}{n} = \frac{\sum_{i=1}^{n} Xi}{n}$$

المتوسط الحسابي في حالة بيانات مكررة

$$\overline{X} = \frac{n_1x_1 + n_2x_2 + n_3x_3 + ...n_nx_n}{n_1 + n_2 + n_3 + n_4 + ...n_n} = \frac{\sum_{i=1}^{n} n_i x_i}{\sum n_i}$$

المتوسط الحسابي في حالة توزيع تكراري

$$\overline{X} = \frac{\sum_{i=1}^{n} n_i x_i}{\sum_{i=1}^{n}}$$

حيث

xi = مراكز الفئات
ni = التكرارات
∑ni = مجموع التكرارات

ملاحظة تستخدم نفس العلاقة لحساب المتوسط الحسابي الموزون حيث

حيث

xi = القيم
ni = الأوزان
∑ni = مجموع الأوزان

الوسيط

- إذا كانت بيانات الظاهرة تحتوي على قيم متطرفة في الصغر أو الكبر فإن النتيجة التي يعطيها المتوسط الحسابي تكون غير واقعية

- الوسيط هو القيمة التي تقسم مجموع البيانات إلى قسمين بحيث يكون نصف عدد البيانات أكبر منه ونصف عدد البيانات أصغر

1- إذا كان عدد البيانات فردي فإن الوسيط هو القيمة التي ترتيبها $(n+1)/2$

2 – إذا كان عدد البيانات زوجي فإن الوسيط هو متوسط القيمتين اللتين ترتيبهما $n/2$ و $n/2 +1$

مثال: البيانات التالية تمثل الدرجات التي تحصل عليها 10 طلبة في امتحان معين:
16، 17، 17، 15، 14، 16، 15، 13، 4، 3.

أوجد متوسط ووسيط درجات الطلبة

الحل: أ) متوسط الدرجات = 13

ب) الوسيط: = 15

نلاحظ أن المتوسط الحسابي لم ينصف أغلب الطلبة الذين تحصلوا على درجات أكبر بكثير من هذا المتوسط وأنحاز ناحية نتيجة الطالبين اللذين تحصلا على نتائج سيئة، في حين أن وسيط هذه الدرجات قسم نتائج الطلبة إلى قسمين بحيث نصف عدد الطلبة تحصلوا على درجات أعلى منه ونصف عدد الطلبة تحصلوا على درجة أقل منه وهو هنا أصدق من المتوسط الحسابي

الوسيط في حالة بيانات متكررة

$$M_e = L_1 + \frac{\frac{N}{2} - N_0}{n_e}.K$$

حيث: M_e = الوسيط.

L_1 = الحد الأدنى للفئة الوسيطية.

N = مجموع التكرارات.

N_0 = التكرار المتجمع الصاعد للفئة قبل الفئة الوسيطية.

n_e = تكرار الفئة الوسطية.

المنوال: هو القيمة الأكثر انتشارا

المنوال لبيانات مبوبة في جداول توزيع تكراري

$$M_0 = L_1 + \frac{d_1}{d_1 + d_2}.K$$

حيث

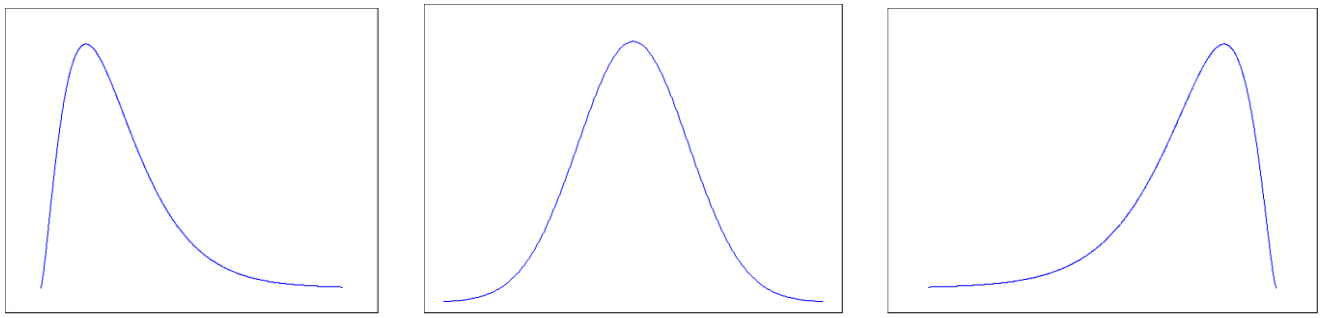
- ✓ d_1 = الفرق بين تكرار الفئة المنوالية وتكرار الفئة التي قبلها
- ✓ d_2 الفرق بين تكرار الفئة المنوالية و التي بعدها
- ✓ K طول الفئة المنوالية
- ✓ L_1 الحد الأدنى للفئة المنوالية

العلاقة بين المتوسط الحسابي والوسيط والمنوال:

1 – تنطبق هذه المقاييس على بعضها وتتساوى في حالة التوزيع التكراري المعتدل وعندها يكون شكل منحنى التوزيع على شكل جرس.

2 – عندما يكون التوزيع التكراري المدروس غير متناظر من اليمن، أي عندما تكون البيانات الصغيرة كثيرة (أكثر تكرارا) تكون المقاييس بالشكل التالي $\overline{X} > M_e > M_0$ ويكون منحنى التوزيع ملتوي ناحية اليمين.

3 – عندما يكون التوزيع التكراري المدروس غير متناظر من اليسار، أي عندما تكون البيانات الكبيرة كثيرة فإن المقاييس تكون بالشكل التالي: $\overline{X} < M_e < M_0$



$\overline{X} > M_e > M_0$ $\overline{X} = M_e = M_0$ $\overline{X} < M_e < M_0$

يعتمد اختيار مقياس النزعة المركزية على نوع البيانات والسياق. المتوسط مفيد للبيانات المتوزعة بشكل طبيعي، بينما الوسيط يكون أفضل في حالة وجود قيم شاذة. المنوال يعتبر مفيدًا عند تحليل البيانات النوعية أو لتحديد القيم الأكثر شيوعًا.

مقاييس التشتت

لا تكفي مقاييس النزعة المركزية لوحدها لمعرفة الصفات الإحصائية اللازمة لوصف الظواهر، لأن الفروق بين قيم الظواهر قد تزداد أو تنقص رغم تساوي المتوسطات لهذه الظواهر

مثال

نفترض أن طالبين تحصلا على النتائج التالية في خمس مواد دراسية:
الطالب (X): 10،11،13،14،15.
الطالب (Y):8،9،13،15،18.

- متوسط درجات الطالب (X) يساوي 12،6 وكذلك متوسط درجات الطالب (Y) يساوي 12،6

- وسيط درجات الطالب (X) يساوي 13 وكذلك وسيط درجات الطالب (Y) يساوي 13

قد يفهم مما سبق أن الطالبين (X)و (Y) لهما نفس المستوى غير أن التمعن الجيد في الدرجات التي تحصل عليها الطالبين تبين أن الطالب (X) ناجح في كل المواد المدروسة في حين أن الطالب (Y) ناجح في ثلاث مواد فقط.

مقاييس النزعة المركزية لا تعطي فكرة وافية عن اختلاف قيم الظواهر لذلك فإن مقاييس النزعة المركزية لا بد أن تكون مصحوبة بمقاييس أخرى لقياس مدى تباعد أو تقارب البيانات من بعضها البعض أو من متوسطها، تسمى هذه المقاييس بمقاييس التشتت.

يقاس تشتت البيانات بعدة مقاييس أهمها:

المدى (المطلق)

المدى= أكبر قيمة – أصغر قيمة.

المدى في التوزيعات التكرارية

المدى = مركز الفئة الأخيرة – مركز الفئة الأولى

= الحد الأعلى لفئة الأخيرة – الحد الأدنى للفئة الأولى

التباين

عبارة عن المتوسط الحسابي لمربعات الفروق بين قيم المتغير الإحصائي ومتوسطها الحسابي

$$V_X = \frac{(X_1 - \overline{X})^2 + (X_2 - \overline{X})^2 + (X_3 - \overline{X})^2 + \ldots\ldots + (X_n - \overline{X})^2}{n}$$

التباين لبيانات متكررة أو مبوبة

$$V_X = \frac{\sum_{i=1}^{n} ni(X_1 - \overline{X})^2}{\sum ni}$$

الانحراف المعياري

الجذر ألتربيعي لمتوسط مجموع مربع انحراف القيم عن متوسطها، أي أنه الجذر التربيعي للتباين.

الانحراف المعياري لبيانات مفردة

$$Sx = \sqrt{\frac{\sum (X_i - \overline{X})^2}{n}}$$

الانحراف المعياري لبيانات متكررة أو مبوبة

$$Sx = \sqrt{\frac{\sum hi(X_i - \overline{X})^2}{\sum hi}}$$