ا.م.دحسن هادي فرج
م.م.سارة علي حسين

علم الإحصاء :Statistics
هو العلم الذي يهتم بدراسة طرق جمع البيانات و تبويبها و تلخيصها بشكل يمكن الاستفادة منها في وصف البيانات وتحليلها للوصول الى قرارت سليمة في ظل ظروف عدم التاكد .

وظائف علم الاحصاء :

اولا: وصف البيانات:
تعتبر طريقة جمع البيانات وتبويبها وتلخيصها من اهم وظائف علم الاحصاء ، اذ لايمكن الاستفادة من البيانات الخام الا اذا تم جمع البيانات وعرضها على شكل جدولي او بياني من ناحية ، وحساب بعض المؤشرات الاحصائية البسيطة التي تدلنا على طبيعة البيانات من ناحية اخرى .

ثانيا : الاحصاء الاستدلالي :
يستند الاستدلال الاحصائي على فكرة اختيار جزء من المجتمع يسمى عينة بطريقة علمية مناسبة . لغرض استخدام بيانات هذه العينة في التوصل الى نتائج يمكن تعميمها على مجتمع الدراسة . يهتم الاحصاء الاستدلالي بـــ
١- التقدير Estimate :ويتم فيه حساب مؤشرات المجتمع (معالم المجتمع) المحسوبة من بيانات العينة في تقدير المدى الذي يقع داخله (معلمة من معالم المجتمع) باحتمال معين .
٢- اختبار الفروض : Tests Hypotheses وفيه يستخدم بيانات العينة للوصول الى قرار علمي سليم بخصوص الفرضيات المحددة حول معالم المجتمع .

ثالثا : التنبؤ :
وفيه يتم استخدام نتائج الاستدلال الاحصائي ، والتي تدلنا على سلوك الظاهرة في الماضي في معرفة ما يمكن ان يحدث لها في الحاضر والمستقبل . وهناك العديد من الاساليب الاحصائية المعروفة التي تستخدم في التنبؤ.

انواع البيانات:
١- البيانات الوصفية : وهي بيانات غير رقمية او بيانات رقمية مرتبة في شكل مستويات او في شكل فئات رقمية . تقاس البيانات الوصفية بمعيارين :
أ- بيانات وصفية مقاسة بمعيار اسمي : وهي بيانات غير رقمية تتكون من مجموعات متنافية ، كل مجموعة لها خصائص تميزها عن المجموعة الاخرى كما ان هذه المجموعة لا يمكن المفاضلة بينها ، ومن الامثلة على ذلك :

- الحالة الاجتماعية : متغير وصفي تقاس بياناته بمعيار اسمي (متزوج - اعزب - ارمل - مطلق).
- اصناف التمور : متغير وصفي تقاس بياناته بمعيار اسمي (برحي - شكر - ....) .
- الجنس : متغير وصفي تقاس بياناته بمعيار اسمي (ذكر - انثى) .
- الجنسية : متغير وصفي تقاس بياناته بمعيار اسمي (عراقي - مصري - .....) .

ب- بيانات وصفية مقاسة بمعيار ترتيبي : وتتكون من مستويات او فئات يمكن ترتيبها تصاعديا او تنازليا ومن الامثلة على ذلك :

- تقدير الطالب : متغير وصفي تقاس بياناته بمعيار ترتيبي (ضعيف - مقبول - متوسط - جيد - جيد جدا - امتياز) .
- المستوى التعليمي : متغير وصفي تقاس بياناته بمعيار ترتيبي (امي - يقرأ ويكتب - ابتدائية - متوسطة - ثانوية - جامعية - اعلى من الجامعية).

٢- البيانات الكمية : هي بيانات يعبر عنها بارقام عددية تمثل القيمة الفعلية للظاهرة ، وتنقسم الى قسمين هما :

أ- بيانات فترة :وهي بيانات رقمية تقاس بمقدار بعدها عن الصفر اي ان للصفر دلالة على وجود الظاهرة مثل :

- درجة الحرارة : متغير كمي تقاس بياناته بمعيار بعدي ، حيث ان درجة الحرارة (٠ المئوي) ليس معناه انعدام الظاهرة ، ولكنه يدل على وجود الظاهرة .
- درجة الطالب في الامتحان : متغير كمي تقاس بياناته بمعيار بعدي ، حيث حصول الطالب على الدرجة (٠) لايعني انعدام مستوى الطالب .

ب- بيانات نسبية : وهي متغيرات كمية ، تدل قيمة (٠) على عدم وجود الظاهرة ومن الامثلة على ذلك :

- انتاجية الارض الزراعية: بالطن / هكتار .
- المساحة المستزرعة بالاعلاف بالدونم .
- كمية الالبان التي تنتجها البقرة في اليوم الواحد .
- عدد مرات استخدام المزرعة لنوع معين من الاسمدة .

اسلوب جمع البيانات :

١- اسلوب الحصر الشامل : يستخدم هذا الأسلوب إذا كان الغرض من البحث هوحصر جميع مفردات المجتمع ، وفي هذه الحالة يتم جمع بيانات عن كل مفردة من مفردات المجتمع بلا استثناء، كحصر جميع المزارع التي تنتج التمور، أو حصر البنوك الزراعية في العراق او مثل التعداد السكاني

مميزاته : الحصر الشامل للبيانات وعدم التحيز ودقة النتائج

عيوبه : يحتاج الى الوقت والمجهود ، التكلفة العالية .

٢- اسلوب المعاينة : يعتمد هذا الاسلوب على معاينة جزء من المجتمع محل الدراسة ، يتم اختياره بطريقة علمية سليمة ، ودراسته ثم تعميم النتائج على المجتمع .

مميزاته :

أ- تقليل الوقت والجهد .

ب- تقليل الكلفة .

ت- الحصول على بيانات اكثر تفصيلا ، خاصة اذا جمعت البيانات من خلال استمارة الاستبيان .

ث- كما ان اسلوب المعاينة يفضل في بعض الحالات التي يصعب فيها اجراء حصر شامل .

عيوبه :

ان النتائج التي تعتمد هذا الاسلوب اقل دقة من نتائج اسلوب الحصر الشامل ، وخاصة اذا كانت العينة المختارة لا تمثل المجتمع تمثيلا حقيقيا .

المجتمع : Population

هو مجموعة من المفردات التي تشترك في صفات وخصائص محددة ، ومجتمع الدراسة ، مثل مجتمع مزارع انتاج الدواجن ، او مجتمع طلاب الصف الثالث المتوسط .

العينة : Sample

هو جزء من المجتمع يتم اختياره بطرق مختلفة بغرض دراسة هذا المجتمع . ويتوقف نجاح استخدام

اسلوب المعاينة على عدة عوامل هي :

١- كيفية تحديد حجم العينة .

٢- طريقة اختيار مفردات العينة

٣- نوع العينية المختارة .

انواع العينات : type of sample

١- العينات الاحتمالية : هي العينات التي يتم اختيار مفرداتها وفقا لقواعد الاحتمالات بمعنى ان يتم اختيار مفرداتها من المجتمع بطريقة عشوائية بهدف تجنب التحيز الناتج عن اختيار المفردات ومن اهم انواع العينات الاحتمالية :

أ- العينة العشوائية البسيطة Simple Random Sample .

ب- العينة العشوائية الطبقية Stratified Random Sample .

ت- العينة العشوائية المنتظمة Systematic Random Sample .

ث- العينات العنقودية او المتعددة المراحل Cluster Sample .

٢- العينات غير الاحتمالية : هي التي يتم اختيار مفرداتها بطريقة غير عشوائية مثل اختيار المزارع التي تنتج نوع معين من التمور . واهم هذه العينات غير الاحتمالية :

أ- العينة العمدية Judgmental Sample .

ب- العينة الحصصية Quota Sample .

الرموز الاحصائية : Statistical Notation

نظرا لاستخدام كافة المراجع العلمية العالمية الرموز والمعادلات الرياضية اللاتينية وذلك لكونها رموزا عالمية متفق عليها من جهة ولعدم وجود اتفاق تام بالوقت الحاضر على تعريبها من جهة اخرى لذا سوف نستخدم تلك الرموز وذلك لسهولة الاستفادة من المراجع الاجنبية والعلمية الاخرى .

يرمز للمتغير الرمز X او Y او Z ولكل قيمة له بالرمز X_i او Y_i او Z_i ويمثل الرمز i رقم المفردة للمتغير .

فلو كانت اطوال اربعة اشجار نخيل هي : ٦ ،٤، ٣ ،٥ فان

$X_i = 6, 4, 3, 5$

اي ان $X_1 = 6$ القيمة الاولى للمتغير X

$X_2 = 4$ القيمة الثانية للمتغير X

$X_3 = 3$ القيمة الثالثة للمتغير X

$X_4 = 5$ القيمة الرابعة للمتغير X

ويرمز لمجموع قيم المتغير بالرمز $\sum_{i=1}^{n} x_i$

الرمز Σ وهو حرف اغريقي يسمى Sigma ويعني مجموع .. والرقمان i و n هما حدا المجموع وعليه

فالرمز $\sum_{i=1}^{n} x_i$ يقرا مجموع قيم X من المشاهدة الاولى وحتى المشاهدة الاخيرة

$$\sum_{i=1}^{n} x_i = x_1 + x_2 + x_3 + ... + x_n$$

اما الرمز $\sum_{i=3}^{5} x_i = x_3 + x_4 + x_5$ فيقرا مجموع قيم X الثالثة والرابعة والخامسة .

*في حالة عدم ذكر حدي المجموع (i و n) فيعني اخذ جميع القيم من ١ الى n *

اما الرمز $\sum_{i=1}^{n} x_i^{\,2}$ فيعني مجموع مربعات قيم X من المشاهدة الاولى وحتى الاخيرة

$$\sum_{i=1}^{n} x_i^2 = x_1^2 + x_2^2 + ... + x_n^2$$

والرمز $(\sum_{i=1}^{n} x_i)^2$ يعني مربع مجموع المشاهدات $(\sum_{i=1}^{n} x_i)^2 = (x_1 + x_2 + x_3 + ... + x_n)^2$

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

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

مثال /// اذا كانت قيم المتغير X كالاتي $x_i = 6,2,3,5,4$

وقيم المتغير y كالاتي $y_i = 1,3,7,8,6$

اوجد $\sum_{i=1}^{n} x_i$ و $\sum_{2}^{4} y_i$ و $\sum_{i=1}^{n} x_i^{\,2}$ و $(\sum y_i)^2$

و $(\sum x_i)(\sum y_i)$ و $\sum x_i y_i$

الحـــل

١-
$$\sum_{i=1}^{n} x_i = x_1 + x_2 + x_3 + ... + x_n$$
$$= 6 + 2 + 3 + 5 + 4 = 20$$

٢-
$$\sum_{2}^{4} y_i = y_2 + y_3 + y_4$$
$$= 3 + 7 + 8 = 18$$

٣-
$$\sum_{i=1}^{n} x_i^2 = x_1^2 + x_2^2 + ... + x_n^2$$
$$= 6^2 + 2^2 + 3^3 + 5^2 + 4^2 = 90$$

٤-
$$(\sum y_i)^2 = (y_1 + y_2 + y_3 + y_4 + y_5)^2$$
$$= (1 + 3 + 7 + 8 + 6)^2$$
$$= (25)^2$$
$$= 625$$

٥-
$$(\sum x_i)(\sum y_i) = (x_1 + x_2 + ... + x_n)(y_1 + y_2 + ... + y_n)$$
$$= (6 + 2 + 3 + 5 + 4)(1 + 3 + 7 + 8 + 6)$$
$$= (20)(25) = 500$$

٦-
$$\sum x_i y_i = x_1 y_1 + x_2 y_2 + x_3 y_3 + x_4 y_4 + x_5 y_5$$
$$= (6 \times 1) + (2 \times 3) + (3 \times 7) + (5 \times 8) + (4 \times 6)$$
$$= 97$$

تمرين //// اذا علمت بان قيم المتغيرين X و Y هما كالتالي :

X=5,3,6,2

Y=4,7,9,5

جد ما يلي : ١- $\sum(x_i - 5)$ ٢- $\sum(y_i - 2)$ ٣- $\sum(x_i + y_{i)}$

٤- $\sum(x_i - 1)(y_i - 2)$

طرق عرض البيانات

١- عرض البيانات جدوليا :

أ- عرض البيانات جدوليا لمتغير وصفي :

اذا كنا بصدد دراسة ظاهرة ما تحتوي على متغير وصفي واحد فانه يمكن عرض بياناته في شكل جدول تكراري بسيط ، وهو جدول يتكون من عمودين احدهما توضع فيه مستويات (مجموعات) المتغير والثاني يوضع فيه عدد المفردات (التكرارات) لكل مستوى مجموعة .

والمثال التالي يوضح لنا كيفية تبويب البيانات الوصفية :

مثال // فيما يلي بيانات عينة من ٤٠ مزرعة من نوع التمر الذي تنتجه :

شكـــر	استعمران	زهـــدي	خستاوي	زهـدي	زهـدي	شكـر	خستاوي
خستاوي	شكـــر	زهـــدي	خستاوي	شكـــر	استعمران	زهدي	شكـــر
استعمران	زهـــدي	استعمران	خستاوي	زهـــدي	خستاوي	خستاوي	شكـــر
زهـــدي	خستاوي	برحـــي	استعمران	استعمران	زهـــدي	زهـدي	زهـــدي
استعمران	زهـــدي	زهـــدي	زهـــدي	خستاوي	زهـــدي	برحـي	خستاوي

المطلوب :

١- ما هو نوع المتغير ؟ وما هو المعيار المستخدم في قياس البيانات ؟
٢- اعرض البيانات في جدول توزيع تكراري.
٣- كون التوزيع التكراري النسبي.
٤- كون التوزيع التكراري المئوي.
٥- علق على النتائج.

الحل :

١- نوع التمر (شكر- برحي - خستاوي – استعمران – زهدي) متغير وصفي ، تقاس بياناته بمعيار اسمي .
٢- لعرض البيانات في جدول توزيع تكراري ، نتبع الاتي :

- تكوين جدول لتفريغ البيانات : وهو جدول يحتوي على علامات احصائية كل علامة تبين تكرار المجموعة التي ينتمي اليها النوع . كما مبين ادناه:

نوع التمر Classes	العلامات الاحصائية Tally	عدد المزارع (التكرارات) Frequency
شكر		٦
خستاوي		١٠
زهدي		١٥
استعمران		٧
برحي		٢
Sum.		40

- تكوين الجدول التكراري .

وهو نفس الجدول السابق ، باستثناء العمود الثاني وياخذ الصورة التالية :

نوع التمر (الفئات) Classes	عدد المزارع (التكرارات) frequency	التوزيع التكراري النسبي	التكرار المئوي
شكر	٦	$\left(\frac{6}{40}\right) = 0.15$	15
خستاوي	١٠	$\left(\frac{10}{40}\right) = 0.25$	٢٥
زهدي	١٥	$\left(\frac{15}{40}\right) = 0.375$	٣٧.٥
استعمران	٧	$\left(\frac{7}{40}\right) = 0.175$	١٧.٥
برحي	٢	$\left(\frac{2}{40}\right) = 0.05$	٥
Sum.	40	1.00	١٠٠

٣- التوزيع التكراري النسبي و التكرار المئوي (في الجدول اعلاه) يحسبان كما يلي :

$$\text{التكرار النسبي} = \frac{\text{تكرار المجموعة}}{\text{مجموع التكرارات } (n)} = \left(\frac{f}{\sum f}\right)$$

التكرار المئوي = التكرار النسبي × ١٠٠

والعمود الثالث في الجدول اعلاه يعرض التكرار النسبي للمزارعين حسب نوع التمر .

٤- التعليق : من الجدول يلاحظ ان نسبة المزارع التي تنتج النوع زهدي في العينة هي ٣٧.٥% وهي اكبر نسبة مما يدل على ان النمط الشائع في انتاج التمور هو(الزهدي)، بينما نجد ان نسبة المزارع التي تنتج النوع (برحي) حوالي ٥% وهي اقل نسبة .

ب- عرض البيانات جدوليا لمتغير كمي :

بنفس الأسلوب السابق المتبع في تكوين جدول تكراري ، يمكن ايضا عرض بيانات المتغير الكمي في شكل جدول توزيع تكراري بسيط ويتكون هذا الجدول من عمودين ، الاول يحتوي على فئات (Classes) تصاعدية للقراءات التي ياخذها المتغير ، والثاني يشمل التكرارات او عدد المفردات التي تنتمي قرأتها للفئة المناسبة لها.

والمثال التالي يبين كيف يمكن عرض البيانات الكمية بيانيا .

مثال //فيما يلي بيانات درجات ٧٠ طالب في الاختبار النهائي لمقررمادة الاحصاء التطبيقي:

٥٦	٦٥	٧٠	٦٥	٥٥	٦٠	٦٦	٧٠	٧٥	٥٦
٦٠	٧٠	٦١	٦٧	٦١	٧١	٦٧	٦٢	٧١	٦٦
٦٨	٧٢	٥٧	٦٨	٧٢	٦٩	٥٧	٧١	٦٩	٧٥
٧٢	٦٢	٦٧	٧٣	٥٨	٦٣	٦٦	٧٣	٦٣	٦٥
٥٨	٧٣	٧٤	٧٦	٧٤	٨٠	٨١	٦٠	٧٤	٥٨
٧٦	٨٢	٧٧	٨٣	٧٧	٨٥	٩١	٧٨	٩٤	٧٢
٧٩	٦٤	٥٧	٧٩	٥٥	٧٨	٦٤	٨٨	٧٨	٦٢

المطلوب :

١- كون جدول التوزيع التكراري لدرجات الطلاب .
٢- كون جدول التوزيع التكراري النسبي .
٣- ما هي نسبة الطلاب الحاصلين على درجة ما بين ٧٠ الى اقل من ٨٠ ؟
٤- ما هي نسبة الطلاب الحاصلين على درجة اقل من ٧٠ درجة ؟
٥- ما هي نسبة الطلاب الحاصلين على درجة ٨٠ او اكثر ؟

الحـــــل//

١- درجة الطالب في الامتحان متغير كمي مستمر ، ولكي يتم تبويب البيانات في شكل جدول تكراري . يتم اتباع الاتي :

- حساب المدى Range (R)

Range = maximum – minimum

$$R = 94 - 55 = 39$$

- تحديد عدد الفئات

تتحدد عدد الفئات وفقا لاعتبارات منها : راي الباحث ، الهدف من البحث ، حجم البيانات ، ويرى كثير من الباحثين ان افضل عدد للفئات يجب ان يتراوح بين ٥ الى ١٥ يستخدم لاستخراج عدد الفئات عدة قوانين هي منها :