الكلية	كلية العلوم
القسم	الكيمياء
المادة باللغة الانجليزية	Science of principles of statistics
المادة باللغة العربية	علم مبادئ الاحصاء
المرحلة الدراسية	الثانية
اسم التدريسي	م.د. محمد اسماعيل خلف حمادي
عنوان المحاضرة باللغة الانجليزية	Statistics
عنوان المحاضرة باللغة العربية	علم الإحصاء
رقم المحاضرة	1
المصادر والمراجع	الراوي، خاشع محمود. 1989. المدخل الى الإحصاء. وزارة التعليم العالي والبحث العلمي. جامعة الموصل.
	طبيه، احمد عبدالسميع. 2007. مبادئ الإحصاء، عمان. دار البداية. ر.أ: www.daralbedayah.com (2007/6/1709)

محتوى المحاضرة

علم الإحصاء Statistics: هو ذلك العلم الذي يعمل على استخدام الأسلوب العلمي في طرق جمع البيانات وتبويبها وتلخيصها وعرضها وتحليلها بهدف الوصول الى استنتاجات وقرارات مناسبة.

أهم وظائف علم الإحصاء:

1- **الوضوحية: اي عرض الحقائق والبيانات بصورة واضحة.**

2- **التكثيف: اي تلخيص البيانات الكثيرة بقيم قليلة.**

3- **صياغة الفرضيات واختبارها.**

4- **المقارنة: يساعد على وضع الأسس السليمة لمقارنة العوامل.**

5- التنبؤ والتكهن: حيث يساعد علم الاحصاء على التنبؤ والتكهن باتجاه قيمة معينة خلال فترة زمنية محددة.

6- يساعد علم الاحصاء على وضع الخطط واتخاذ القرارات المناسبة.

طبيعة البيانات الإحصائية:

عند جمع بيانات حول صفة او ظاهرة ما فإننا نرمز لها برمز مثل y وكل مفردة او مشاهدة منها بالرمز yi. فمثلا عند دراسة اطوال نباتات الذرة فإننا نرمز لصفة الطول بالرمز y وتسمى متغير Variable، وطول أي نبات بالرمز yi ويسمى مشاهدة Observation

مصطلحات إحصائية:

1. المتغير Variable هو أي صفة او ظاهرة لها مفردات مختلفة. ويرمز له برمز مثل x او y او z.

تقسم المتغيرات الى:

أ. متغيرات وصفية او نوعية Qualitative Variables وهي تلك الصفات او الظواهر التي لا يمكن قياسها مباشرة بقيم عددية مثل اللون (احمر، ازرق، اصفر) والجنس (ذكر وانثى)

ب. متغيرات كمية Quantitative Variables وهي الصفات او الظواهر التي يمكن قياسها مباشرة بقيم عددية مثل الطول والوزن والعمر.

تقسم المتغيرات الكمية الى قسمين:

i. المتغيرات المستمرة Continuous Variables: تأخذ جميع القيم داخل مدى معين، كالوزن او الطول، فمثلا وزن دجاجة 1.25 كغم، وطول شجرة 15.7 م.

ii. المتغيرات المتقطعة Discrete Variables: لا تأخذ جميع القيم داخل مدى معين، كعدد الثمار على النبات الذي يكون مثلا 3 او 4 ولا يكون 3.5 او 5.2

2. المشاهدة Observation: وهي مفردة من مجموعة مفردات لمتغير تشكل سجل رقمي لمتغير وتوضع تحت الدراسة وتعتبر المادة الأولية للبحث، فاذا أراد الباحث دراسة عدد الأوراق النبات لصنف معين من زهرة الشمس فانه يأخذ عدد من النبات ويشاهد عدد الأوراق لكل نبات منها، فلو كان عدد الأوراق لنبات معين هو 7 فان هذا العدد يمثل مشاهدة واحدة.

3. المجتمع Population: هو جميع القيم او المفردات التي يمكن ان يأخذها المتغير مثل عدد الوحدات الإنتاجية في اليوم لمصنع معين.

4. العينة Sample: هي جزء من المجتمع مأخوذ بطريقة عشوائية وتكون ممثلاً له تمثيلاً حقيقياً.

يعتمد اسلوب العينات على معاينة جزء من المجتمع محل الدراسة، يتم اختيارة بطريقة علمية سليمة، ودراسته ثم تعميم نتائج العينة على المجتمع، ويتميز هذا الاسلوب بالتالي:

- تقليل الوقت والجهد.
- تقليل التكلفة.
- الحصول على بيانات اكثر تفصيلا، وخاصة اذا جمعت البيانات من خلال استمارة استبيان.
- كما ان اسلوب المعاينة يفضل في بعض الحالات التي يصعب فيها اجراء حصر شامل مثل معاينة دم المريض، اعداد الاسماك في البحر.

ولكن يعاب على هذا الاسلوب بان النتائج المستحصلة بهذا الاسلوب تكون اقل دقة من نتائج اسلوب الحصر الشامل وخاصة اذا كانت العينة المختارة لا تمثل المجتمع تمثيلا جيدا.

شروط اختيار العينة

1. يجب ان لا تتسم العينة التي تم اختيارها بالتحيز او المحاباة بمعنى ان تأخذها من بين مفردات المجتمع الاصلي عشوائياً.
2. ان تكون الظاهرة المراد عمل معاينة لها سائدة ومنتشرة في المجتمع الاصلي وليست نادرة الحدوث.
3. يجب ان تكون العينة ممثلة لجميع المجتمع الاصلي.
4. ضرورة افتراض تجانس مفردات المجتمع الاصلي وفي حاله تعذر ذلك في المجتمعات غير المتجانسة يلجأ الباحث الى تقسيمها الى مجتمعات صغيرة متجانسة.

1. انواع البيانات الاحصائية: البيانات الإحصائية هي المواد الأولية للإحصاء، والتي من خلالها يمكن تبويبها وتصنيفها واجراء التحاليل عليها واستنباط التوصيات منها. وتعتمد طبيعة البيانات الإحصائية على الاهداف المطلوبة منها، لذا فان هذا الفصل سيتناول مصادر البيانات وطرق جمعها.

ويمكن تقسيم البيانات الاحصائية الى نوعين اساسيين من البيانات هما:

i. البيانات النوعية Qualitative data: وهي بيانات وصفية، تشمل الظواهر التي لا تخضع للقياسات الكميه، ويصعب التعبير عنها بصورة عددية. كما تعرف احياناً باسم (البيانات التصنيفية) لأنها تصنف البيانات حسب الصفات، سواء أكانت ذلك من حيث النوع، مثل تصنيف التربة الى طينية وغرينية ورملية، أم كان ذلك من حيث الدرجة مثل تصنيف القوى العاملة بحسب الدرجة التعليمية الى أمي، وملم بالقراءة والكتابة، وحاصل على الشهادة الابتدائية ...الخ، حيث تشتمل البيانات على عدد الافراد الذين ينتمون الى كل صفة من هذه الصفات.

ii. البيانات الكمية Quantitative data: وهي بيانات رقمية، تشمل الظواهر القابلة للقياسات الكمية، ويمكن التعبير عنها بصورة عددية مثل كميات الامطار، وأعمار السكان وكميات الانتاج ... الخ، وغير ذلك من البيانات التي تعكس القيم الفعلية للظواهر. ومن الجدير بالذكر ان معظم الاساليب الإحصائية تعنى بمعالجه البيانات الرقمية وهذا على خلاف البيانات النوعية، التي لا تستخدم فيها سوى بعض الاساليب الإحصائية المعدودة.

2. طرق واساليب جمع البيانات الاحصائية: تعتبر طريقة جمع البيانات من أهم المراحل التي يعتمد عليها البحث الاحصائي، كما ان جمع البيانات بأسلوب علمي صحيح، يترتب عليه الوصول الى نتائج دقيقة في التحليل، وقد تكون البيانات المطلوبة متوفرة سواء كانت منشورة او غير منشورة نتيجة لجمعها من قبل جهات رسمية او غير رسمية وفي هذه الحالة فان جمعها يتطلب الرجوع الى تلك المصادر. اما اذا كانت البيانات غير متوفرة فان الامر يتطلب اجراء استقصاءات او تجارب ميدانية لتوفيرها.

مصادر جمع البيانات المتوفرة

هناك مصدرين للحصول منها على البيانات هما:

1- المصادر الاولية Primary Source: وهي المصادر التي نحصل منها على البيانات بشكل مباشر، حيث يقوم الباحث نفسه بجمع البيانات من المفردة محل البحث مباشرة، فعندما يهتم الباحث بجمع بيانات عن الاسر الريفية يقوم بأجراء مقابلة مع رب الاسرة، ويتم الحصول منه مباشرة على بيانات خاصة بأسرته مثل بيانات المنطقة التابع لها وعدد افراد الاسرة واعمارهم والمستوى التعليمي،

ومساحه الملكية الزراعية، وانواع المحاصيل المزروعة واعداد وانواع الثروة الحيوانية التي يمتلكها، وانواع المكننة الزراعية التي يستخدمها ونوع ملكيتها (ملك شخصيه ام ايجار) والدخل الشهري وهكذا. او قد تقوم جهات رسمية او شبة رسمية بجمع انواع معينه من البيانات بشكل دوري او خلال فترة زمنية محددة كالجهاز المركزي للإحصاء او الهيئة العامة للأنواء الجوية وقد تصدر هذه الجهات نشرات إحصائية دورية.

ويتميز هذا النوع من المصادر بالدقة والثقة في البيانات، لان الباحث هو الذي يقوم بنفسه بجمع البيانات من المفردة محل البحث مباشرة، ولكن اهم ما يعاب عليها انها تحتاج الى وقت ومجهود كبير ومن ناحيه اخرى فأنها مكلفه من الناحية المادية.

2- المصادر الثانوية Secondary source: وهي المصادر التي نحصل منها على البيانات بشكل غير مباشر، بمعنى اخر يتم الحصول عليها بواسطه اشخاص اخرين أو اجهزة أو هيئات رسمية اخرى غير متخصصة بجمع البيانات الإحصائية، مثل نشرات منظمة الاغذية والزراعة الدولية ... وهكذا.

ومن مزايا هذا النوع من المصادر توفير الوقت والجهد والمال، الا انها تحتوي على تفاصيل اقل من المصادر الاولية، وان درجه ثقة الباحث فيها ليست بنفس الدرجة في حالة المصادر الاولية.

رموز إحصائية Statistical Notations

كما مر سابقا يرمز للمتغير بالرمز y ولكل قيمة له بالرمز y_i فلو كانت اعمار 5 طلاب هي: 16 و 22 و 24 و 18 و 20 فتكتب

$$y_i = 16,22,24,18,20$$

أي ان:

$y_1=20$ أي القيمة الأولى للمتغير أو المشاهدة الأولى

و $y_2=18$ أي القيمة الثانية للمتغير أو المشاهدة الثانية

وهكذا الى $y_n=16$ أي القيمة الاخيرة (n=5) للمتغير أو المشاهدة الأخيرة أو الخامسة.

ويرمز لمجموع قيم المتغير بالرمز

$$\sum_{i=1}^{n} yi$$

فالرمز Σ حرف اغريقي يسمى Sigma أي مجموع الـ.... أو Summation of و n و1 هما حدا المجموع، حيث n هو عدد المشاهدات. وعليه فالرمز

$$\sum_{i=1}^{n} yi$$

يقرأ: مجموع قيم y من المشاهدة الأولى وحتى الأخيرة ويكتب:

$$\sum_{i=1}^{n} yi = y1 + y2 + \cdots + yn$$

Σ دون كتابة حدود المجموعة وذلك للاختصار والسهولة اذا لم يحدث هناك y_i ويمكن ان يكتب التباس كما يلي:

$$\sum yi = 20 + 18 + 24 + 22 + 16$$

$$\sum yi = 100$$

وهناك مجموع جزئي مثل:

$$\sum_{i=3}^{5} yi$$

ويقرأ مجموع المشاهدات من المشاهدة الثالثة وحتى الخامسة ويكتب:

$$\sum_{i=3}^{5} yi = y3 + y4 + y5$$

$$\sum_{i=3}^{5} yi = 24 + 22 + 16$$

$$\sum_{i=3}^{5} yi = 62$$

ويرمز لمجموع مربعات المشاهدات بالرمز Σyi^2 أي

$$\Sigma yi^2 = y_1^2 + y_2^2 + \ldots + y_n^2$$

ويرمز لمربع مجموع المشاهدات بالرمز $(\Sigma yi)^2$ أي:

$$(\Sigma yi)^2 = (y_1 + y_2 + \ldots + y_n)^2$$

ويرمز لمجموع حاصل ضرب قيم متغيرين x و y بالرمز $\Sigma x_i y_i$ أي

$$\Sigma x_i y_i = x_1 y_1 + x_2 y_2 + \ldots + x_n y_n$$

مثال/ اذا كانت قيم المتغيرين x و y هي كالآتي:

x_i=4,2,3,7

y_i=3,9,6,2

جد

(a) Σy_i^2 (b) $(\Sigma y_i)^2$ (c) $\Sigma x_i y_i$

(d) $(\Sigma x_i)(\Sigma y_i)$

الحل

(a) $\Sigma y_i^2 = y_1^2+y_2^2+y_3^2+y_4^2$

$= 3^2+9^2+6^2+2^2$

$= 130$

(b) $(\Sigma y_i)^2 = (y_1+y_2+y_3+y_4)^2$

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

$= (20)2$

$= 400$

(c) $\Sigma x_i y_i = x_1y_1+x_2y_2+x_3y_3+x_4y_4$

$= 4*3+2*9+3*6+7*2$

$= 62$

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

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

$= (16)(20)$

$= 320$