وزارة التعليم العالي والبحث العلمي
جامعة الموصل ـ كلية الزراعة والغابات
قسم المحاصيل الحقلية ـ المرحلة الاولى
احصاء زراعي

احصاء زراعي

المحاضرة الاولى

(المقدمة- تعريف علم الاحصاء- الرموز الاحصائية)

م.م. ظفر عبدالرزاق فرحان النجماوي

مبادئ الإحصاء

المحاضرة الاولى

علم الاحصاء 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 معينه تكون ممثلة لذلك المجتمع في حالة صعوبة دراسة جميع عناصر المجتمع وبالتالي تعمم نتائج العينة على المجتمع.

Statistical Notations إحصائية رموز

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 للمتغيرأوالمشاهدة الأولى و

$y_2=18$ أي القيمة الثانية للمتغيرأوالمشاهدة الثانية

وهكذا الى $y_n =16$ أي القيمة الاخيرة (n=5) للمتغيرأوالمشاهدة 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 و 1 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:

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

ويمكن ان يكتب Σy_i دون كتابة حدود المجموعة وذلك للاختصار والسهولة اذا لم يحدث هناك 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) $\Sigma y^2 = y_1^2+y_2^2+y_3^2+y_4^2$

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

$= 130$

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

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

$= (16)(20)$

$= 320$

المصادر:

1. الراوي، خاشع محمود. 1989. المدخل الى الإحصاء. وزارة التعليم العالي والبحث العلمي. جامعة الموصل.
2. طبيه، احمد عبدالسميع. 2007. مبادئ الإحصاء، عمان. دار البداية. ر.أ: (2007/6/1709) www.daralbedayah.com

وزارة التعليم العالي والبحث العلمي
جامعة الموصل ـ كلية الزراعة والغابات
قسم المحاصيل الحقلية ـ المرحلة الاولى
احصاء زراعي

احصاء زراعي

المحاضرة الثالثة

م.م. ظفر عبدالرزاق فرحان النجماوي

مبادئ الإحصاء

العرض الجدولي والتمثيل البياني

عند جمع بيانات حول ظاهرة معينة تبوب في جداول أو يعبر عنها برسوم بيانية لإعطاء الفكرة التي تتضمنها البيانات بأسلوب سريع وبسيط.

العرض الجدولي:

جدول التوزيع التكراري: هو جدول بسيط يتكون من عمودين:

الأول: يحتوي على قيم متغير أو تقسم فيه قيم المتغير الى اقسام او مجموعات تدعى الفئات Classes الثاني: يبين عدد مشاهدات كل قيمة من قيم المتغير او عدد المشاهدات التي تنتمي الى الفئة ويسمى **التكرار Frequency**

مثال لجدول توزيع تكراري يتكون من عمود للقيم وعمود للتكرارات، لمصنع تعليب ينتج علب بالأوزان التالية: 3 و4 و5 و6 كغم

التكرارات (عدد العلب المنتجة في يوم) **fi**	القيم (اوزان العلب بالكيلوغرام) **yi**
120	3
155	4
145	5
100	6

على سبيل المثال ان عدد العلب المنتجة في اليوم بوزن 3 كغم هو 120 علبة.

مثال لجدول توزيع تكراري يتكون من عمود للفئات وعمود للتكرارات لدرجات الطلبة في مادة الإحصاء.

عدد الطلبة، التكرار fi	فئات درجات الطلبة Classes
3	41-45
7	46-50
10	51-55
10	56-60
8	61-65
5	66-70
4	71-75
1	76-80

على سبيل المثال عدد الطلبة الذين حصلوا على درجات ضمن المدى 61 الى 65 هو 8 طلاب.

تبويب البيانات (انشاء جدول توزيع تكراري)

البيانات غير المبوبة: وهي البيانات الأولية او الاصلية التي جمعت ولم تبوب في جدول. البيانات المبوبة: وهي البيانات التي بوبت ونظمت في جدول توزيع تكراري.

مثال// البيانات التالية تمثل اطوال 80 نباتا من نباتات الذرة الصفراء، وهي بيانات غير مبوبة،
المطلوب تبويب البيانات في جدول توزيع تكراري (انشاء جدول توزيع تكراري يتكون من فئات وتكرارات).

80, 87, 98, 81, 74, 48, 79, 80, 78, 82, 93, 91, 70, 90, 80, 84, 73, 74, 81, 56,
65, 92, 70, 71, 86, 83, 93, 65, 51, 85, 68, 72, 68, 86, 43, 74, 73, 83, 90, 35,
75, 67, 72, 90, 71, 76, 92, 93, 81, 88, 91, 97, 72, 61, 80, 91, 77, 71, 59, 80,
95, 99, 70, 74, 63, 89, 67, 60, 82, 83, 63, 60, 75, 79, 88, 66, 70, 88, 76, 63

الحل// نتبع الخطوات التالية لتبويب البيانات في جدول توزيع تكراري

1. استخراج المدى The Range

المدى = اعلى قيمة – اقل قيمة

= 99 — 35

= 64

2. اختيار وتحديد عدد الفئات Number of classes

وهناك عدة طرق حسابية تقريبية لايجاد عدد الفئات أهمها:

طريقة Sturges

$$\text{Number of classes} = 1 + (3.3 * \log n)$$

n: عدد المفردات

وطريقة Yule

$$\text{Number of classes} = 2.5 * \sqrt[4]{n}$$

ولكل من الطريقتين ميزات وعيوب ولن نستعمل أيا منها هنا بل سنختار عدد الفئات اختيارا على ان لا تقل عن 5 ولا تزيد عن 15 فئة وذلك تبعا لطبيعة البيانات وعدد مفرداتها ومدى التغير فيها.

ولنفرض اننا اخترنا 7 فئات.

3. إيجاد طول الفئة Class length

يجب ان لا يقل طول الفئة عن $\frac{\text{مدى التغير}}{\text{عدد الفئات}}$ مقربة ~~الى اقرب~~ عدد صحيح اكبر

$$\text{طول الفئة} = \frac{\text{المدى}}{\text{عدد الفئات}}$$

$$= \frac{64}{7}$$

$$= 9.14$$

لذا يستحسن ان يكون طول الفئة = 10 ويفضل ان تكون اطوال الفئات متساوية.

4. كتابة حدود الفئات Class limits

يجب كتابة حدود الفئات بحيث ان جميع قيم المتغير تقع بين الحد الأدنى للفئة الأولى والحد الأعلى للفئة الأخيرة. يمكن ان يكون الحد الأدنى هو قيمة اقل مفردة او اقل منها بقليل، لذا من الممكن ان يكون الحد الأدنى للفئة الأولى 31 وبما ان طول الفئة =10 لذا فان الحد الاعلى للفئة الأولى هو 40، اذآ الفئة الأولى هي من 31 الى 40 والفئة الثانية من 41 الى 50 والفئة السابعة (الأخيرة) من 91 الى 100 يجب ان يكون الحد الأدنى للفئة الأولى والحد الأعلى للفئة الأخيرة يحوي كافة قيم المتغير.

5. استخراج عدد التكرارات لكل فئة Class frequency

يتم ذلك بتسجيل القيم الاصلية واحدة تلو الأخرى في الفئة الخاصة بها على شكل علامات أولا ثم ترجمتها الى ارقام كما مبين في الجدول ادناه:

الفئات Classes	التكرار بالعلامات fi	التكرار رقمآfi
31-40	I	1
41-50	II	2
51-60	IIIII	5
61-70	IIIII IIIII IIIII	15
71-80	IIIII IIIII IIIII IIIII IIIII	25
81-90	IIIII IIIII IIIII IIIII	20
91-100	IIIII IIIII II	12
المجموع		80

ويجب ان يكون المجموع الكلي للتكرارات يساوي العدد الكلي لقيم المتغير (80).

الفئات Classes: وهي المجاميع التي قسمت اليها قيم المتغير، وهي عبارة عن مدى معين من قيم المتغير.

حدود الفئة Class Limits: لكل فئة حدان حد ادنى وحد اعلى، فمثلا في الجدول السابق الفئة الأولى حدها الأدنى هو 31 وحدها الأعلى هو 40.

مركز الفئة: وهو عبارة عن منتصف المدى بين حدي الفئة، ويمكن حسابه عن طريق العلاقة الاتية:

$$\text{مركز الفئة} = \frac{\text{الحد الادنى+الحد الاعلى}}{2}$$

طول الفئة = (الحد الأعلى للفئة – الحد الأدنى للفئة)+1

في المثال السابق: طول الفئة الأولى = (40-31)+1

= 10

وكذلك باقي الفئات طول كل منها =10

الحدود الحقيقية:

الحد الأدنى الحقيقي للفئة = مركز الفئة - $\frac{1}{2}$ طول الفئة

الحد الأعلى الحقيقي للفئة = مركز الفئة + $\frac{1}{2}$ طول الفئة

طول الفئة = الحد الحقيقي الأعلى للفئة – الحد الحقيقي الأدنى للفئة

مبادئ الإحصاء

مثال// أكمل الجدول التالي

Classes الفئات	التكرار fi	مركز الفئة	الحدود الحقيقية
31-40	1		
41-50	2		
51-60	5		
61-70	15		
71-80	25		
81-90	20		
91-100	12		

الحل//

مركز الفئة= $\frac{\text{الحد الادنى+الحد الاعلى}}{2}$

مركز الفئة الأولى = $\frac{40+31}{2}$

$= 35.5$

مركز الفئة الثانية = $\frac{50+41}{2}$

$= 45.5$

وهكذا لبقية الفئات.

الحدود الحقيقية:

الحد الأدنى الحقيقي للفئة الاولى = مركز الفئة الاولى - $\frac{1}{2}$ طول الفئة

$= 35.5 - \frac{1}{2}(10)$

$= 30.5$

الحد الاعلى الحقيقي للفئة الاولى = مركز الفئة الاولى + $\frac{1}{2}$ طول الفئة

$= 35.5 + \frac{1}{2}(10)$

$= 40.5$

الحد الأدنى الحقيقي للفئة الثانية = مركز الفئة الثانية - $\frac{1}{2}$ طول الفئة

$= 45.5 - \frac{1}{2}(10)$

$= 40.5$

الحد الاعلى الحقيقي للفئة الثانية = مركز الفئة الثانية + $\frac{1}{2}$ طول الفئة

$= 45.5 + \frac{1}{2}(10)$

$= 50.5$

وهكذا لبقية الفئات.

مبادئ الإحصاء

لاكمال الجدول كالآتي

Classes الفئات	التكرار fi	مركز الفئة	الحدود الحقيقية
31-40	1	35.5	30.5-40.5
41-50	2	45.5	40.5-50.5
51-60	5	55.5	50.5-60-5
61-70	15	65.5	60.5-70.5
71-80	25	75.5	70.5-80.5
81-90	20	85.5	80.5-90.5
91-100	12	95.5	90.5-100.5

جدول التوزيع التكراري النسبي

هو جدول يبين الأهمية النسبية لكل فئة ويحسب:

التكرار النسبي لأي فئة = $\frac{\text{تكرار تلك الفئة}}{\text{مجموع التكرارات}} = \frac{fi}{\sum fi}$

وقد يعبر عن التكرار النسبي كنسبة مئوية بالضرب في 100

التكرار النسبي المئوي لأي فئة = $\frac{\text{تكرار تلك الفئة}}{\text{مجموع التكرارات}} \times 100$

مثال// أكمل الجدول

Classes الفئات	التكرار fi	التكرار النسبي	التكرار النسبي المئوي
31-40	1		
41-50	2		
51-60	5		
61-70	15		
71-80	25		
81-90	20		
91-100	12		

$\frac{\text{تكرار الفئة}}{\text{مجموع التكرارات}}$

1

—

=

التكرار النسبي للفئة الاولى =

= 0.0125

التكرار النسبي المئوي للفئة الأولى = 0.0125×100

= 1.25

وهكذا لبقية الفئات

مبادئ الإحصاء

Classes الفئات	التكرار fi	التكرار النسبي	التكرار النسبي المئوي
31-40	1	0.0125	1.25
41-50	2	0.0250	2.50
51-60	5	0.0625	6.25
61-70	15	0.1875	18.75
71-80	25	0.3125	31.25
81-90	20	0.2500	25.00
91-100	12	0.1500	15.00
المجموع	80	1	100

التوزيعات المتجمعة

1. التكرار التجميعي التصاعدي: ويرمز له UCF التكرار التجميعي التصاعدي للفئة الأولى = تكرار الفئة الأولى :

$$UCF\ 1=f1$$

التكرار التجميعي التصاعدي للفئة الثانية = تكرار الفئة الثانية + تكرار الفئة الأولى:

$$UCF\ 2=f2+f1$$

التكرار التجميعي التصاعدي للفئة الثالثة= تكرار الفئة الثالثة +تكرار الفئة الثانية +تكرار الفئة الأولى:

$$UCF\ 3=f3+f2+f1$$

وهكذا، أي ان التكرار التجميعي التصاعدي لاي فئة = تكرار تلك الفئة + تكرار ما قبلها من الفئات.

2. التكرار التجميعي التنازلي: ويرمز له LCF

التكرار التجميعي التنازلي للفئة الأولى = مجموع التكرارات :

$$LCF\ 1=\sum fi$$

التكرار التجميعي التنازلي للفئة الثانية = مجموع التكرارات – تكرار الفئة الأولى:

$$LCF\ 2=\sum fi-f1$$

التكرار التجميعي التنازلي للفئة الثالثة = مجموع التكرارات – تكرار الفئة الثانية- تكرار الفئة الأولى:

$$LCF\ 3=\sum fi-f2-f2-f1$$

مثال// اكمل الجدول

Classesالفئات	fiالتكرار	التكرار التجميعي التصاعديUCF	التكرار التجميعي التنازليLCF
31-40	1		
41-50	2		
51-60	5		
61-70	15		
71-80	25		
81-90	20		
91-100	12		

الحل//

التكرار التجميعي التصاعدي للفئة الأولى = تكرار الفئة الأولى

1 =

التكرار التجميعي التصاعدي للفئة الثانية = تكرار الفئة الثانية + تكرار الفئة الأولى

1+2 =

3=

التكرار التجميعي التصاعدي للفئة الثالثة= تكرار الفئة الثالثة +تكرار الفئة الثانية +تكرار الفئة الأولى

1+2+5 =

8 =

التكرار التجميعي التصاعدي للفئة الرابعة

1+2+5+15 =

23 =

وهكذا.

التكرار التجميعي التنازلي للفئة الأولى = مجموع التكرارات:

80=

التكرار التجميعي التنازلي للفئة الثانية = مجموع التكرارات – تكرار الفئة الأولى:

1-80=

79=

التكرار التجميعي التنازلي للفئة الثالثة = مجموع التكرارات – تكرار الفئة الثانية- تكرار الفئة الأولى:

1-2-

80=

77=

وهكذا

ليكتمل الجدول كما يلي

Classes الفئات	fi التكرار	التكرار التجميعي التصاعديUCF	التكرار التجميعي التنازليLCF
31-40	1	1	80
41-50	2	3	79
51-60	5	8	77
61-70	15	23	72
71-80	25	48	57
81-90	20	68	32
91-100	12	80	12

مبادئ الإحصاء

التمثيل البياني لجدول التوزيع التكراري

لتمثيل الجدول التالي بيانيا بالمدرج التكراري والمضلع التكراري والمنحنى التكراري.

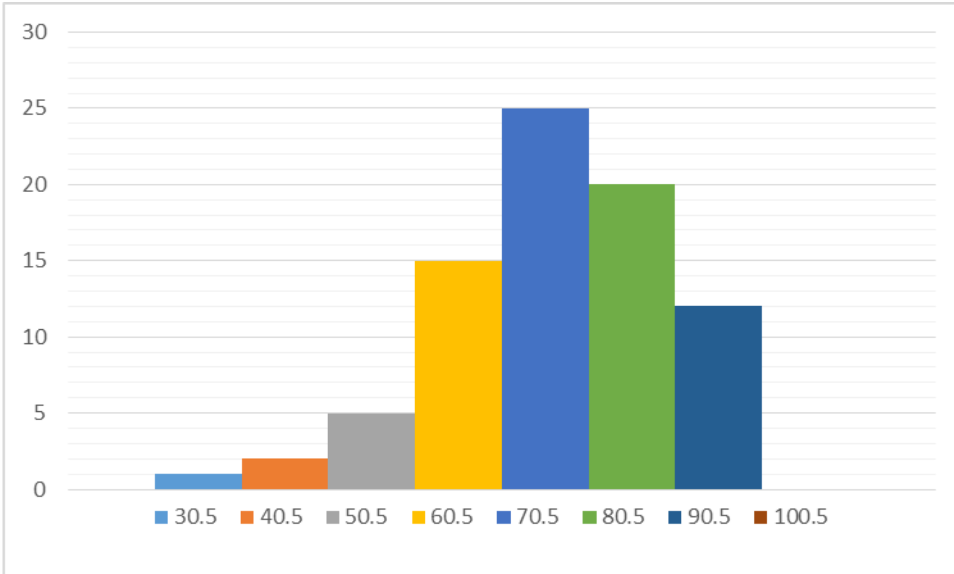
Classes الفئات	fi التكرار
31-40	1
41-50	2
51-60	5
61-70	15
71-80	25
81-90	20
91-100	12

1. المدرج التكراري Histogram

- إيجاد الحدود الحقيقية للفئات.

Classesالفئات	fiالتكرار	الحدود الحقيقية
31-40	1	30.5-40.5
41-50	2	40.5-50.5
51-60	5	50.5-60-5
61-70	15	60.5-70.5
71-80	25	70.5-80.5
81-90	20	80.5-90.5
91-100	12	90.5-100.5

- رسم المحورين الافقي x والعمودي y
- تقسيم المحور الافقي الى اقسام متساوية بمقياس رسم مناسب بحيث يشمل جميع الحدود الحقيقية للفئات.
- يقسم المحور العمودي الى اقسام متساوية بحيث يشمل اعلى تكرار.
- يرسم لكل فئة مستطيل تمثل قاعدته طول الفئة وارتفاعه يمثل تكرار الفئة.

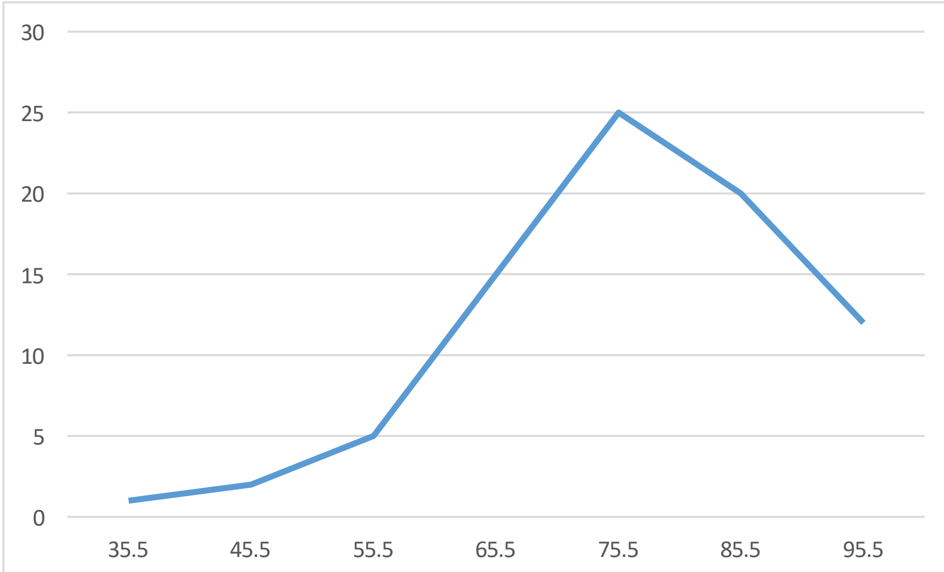


2. المضلع التكراري Frequency Polygon

- إيجاد مراكز الفئات.

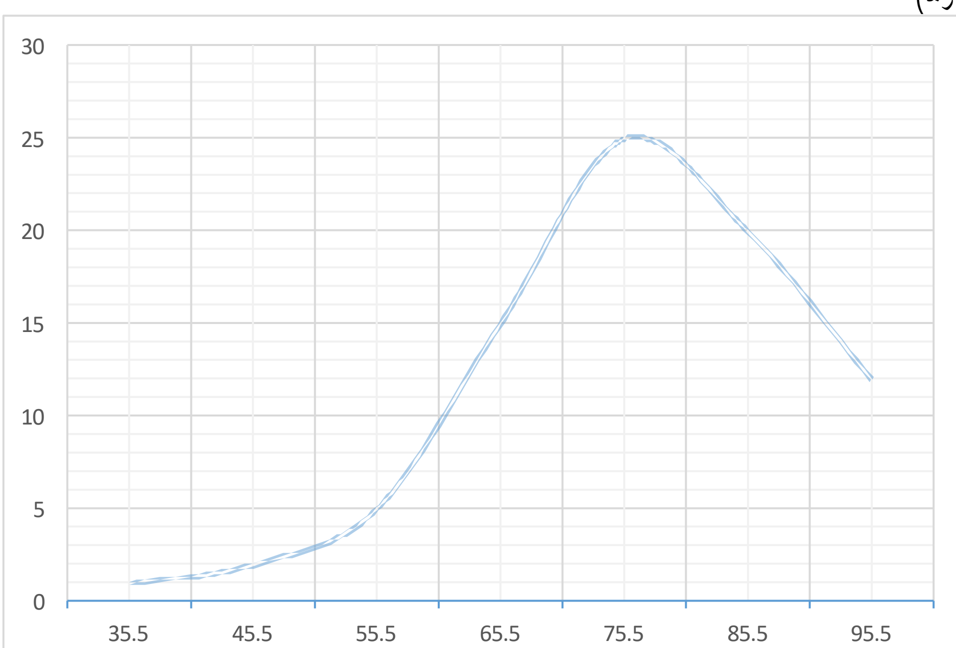
Classesالفئات	fiالتكرار	مركز الفئة
31-40	1	35.5
41-50	2	45.5
51-60	5	55.5
61-70	15	65.5
71-80	25	75.5
81-90	20	85.5
91-100	12	95.5

- رسم المحورين الافقي x والعمودي y
- تدريج المحور الافقي الى اقسام متساوية بمقياس رسم مناسب بحيث يشمل جميع مراكز الفئات.
- يقسم المحور العمودي الى اقسام متساوية بحيث يشمل اعلى تكرار.
- يرسم فوق مركز كل فئة نقطة ارتفاعها بقدر تكرار تلك الفئة (نقاط التقاء مركز الفئة مع تكرارها)
- توصيل النقاط بخطوط مستقيمة.



3. المنحنى التكراري Frequency Curve

وهو عبارة عن منحني يمر بمعظم النقاط الواقعة على مراكز الفئات والتي ارتفاعها يمثل تكرار تلك الفئة (نقاط التقاء مركز الفئة مع تكرارها)



المصادر:

1. الراوي، خاشع محمود. 1989. المدخل الى الإحصاء. وزارة التعليم العالي والبحث العلمي. جامعة الموصل.
2. طبيه، احمد عبدالسميع. 2007. مبادئ الإحصاء، عمان. دار البداية. ر.أ: (2007/6/1709) www.daralbedayah.com
3. David, M. Lane. Introduction to Statistics. Online Edition.

مبادئ الإحصاء

وزارة التعليم العالي والبحث العلمي
جامعة الموصل - كلية الزراعة والغابات
قسم المحاصيل الحقلية – المرحلة الاولى
أحصاء زراعي

أحصاء الزراعي

المحاضرة الرابعة

(مقاييس التشتت او الاختلاف)

م.م. ظفر عبدالرزاق فرحان النجماوي

مبادئ الإحصاء

مقاييس التشتت أو الاختلاف

Measures of Dispersion or Variation

مقاييس التشتت او الاختلاف هي مؤشرات لمدى التباعد او التقارب بين قيم مشاهدات متغير ما، وتضم مقاييس التشتت:

أولا: مقاييس التشتت المطلق:

وهي التي تكون وحداتها نفس وحدات القيم الاصلية وأهمها:

1. المدى The Range
2. الانحراف المتوسط The Mean Deviation
3. التباين والانحراف القياسي The Variance and The Standard Deviation

ثانيا: مقاييس التشتت النسبي:

وهي المقاييس الخالية من وحدات القياس وأهمها:

معامل الاختلاف Coefficient of Variation

The Range المدى

المدى هو الفرق بين اعلى قيمة واقل قيمة ويرمز له R

$$R = y_{max} - y_{min}$$

مثال// جد المدى لكل من المجموعات التالية:

a = 5,7,2,1, 9, 6

b= 98, 104, 102, 100, 103, 99

c = 3, 2,6,4,5,210

الحل//

(a) R= 9-1

= 8

(B) R= 104 -98

= 6

(c) R = 210 -2

= 208

ان المدى أحيانا يكون مضللا ولا يعطي فكرة واضحة عن طبيعة البيانات لأنه يعتمد على القيمتين الصغرى والكبرى اللتين كثيرا ما تكون شاذة.

مبادئ الإحصاء

الانحراف المتوسط

The Mean Deviation

يعرف الانحراف المتوسط بانه معدل الانحرافات المطلقة عن وسطها الحسابي، ويرمز له .M.D اذا هو يمثل معدل تشتت القيم عن وسطها الحسابي ويحسب:

للعينة $$M.\ D. = \frac{\sum |x_i - \bar{x}|}{n}$$

حيث ان:

$\bar{X}$= متوسط القيم X_i= القيم

n= عدد القيم

$|x_i - \bar{x}|$ = مجموع الانحرافات المطلقة عن المتوسط الحسابي، اي دون الاخذ بنظر الاعتبار الاشارة السالبة اوالموجبة وانما تكون جميع القيم موجبة ويسمى بالفرق المطلق. وتجدر الاشارة الى ان مجموع انحرافات القيم عن وسطها الحسابي = 0 دائماً ، لذلك تؤخذ القيمة المطلقة لحساب الانحراف المتوسط.

للمجتمع $$M.\ D. = \frac{\sum |x_i - \mu|}{N}$$

مثال: جد الانحراف المتوسط للبيانات التالية، والذي يمثل أوزان 5 دجاجات.

X_i=2, 1, 2.5, 1.5, 3 kg

أستخرج المتوسط الحسابي ($\bar{X}$) للقيم

$$\bar{X} = \frac{\sum x_i}{n}$$

$$= \frac{2+1+2+2.5+1.5+3}{5}$$

$$= \frac{10}{5}$$

$$=2$$

مبادئ الإحصاء

تسلسل القيم	Xi	الفرق بين القيمة والمتوسط الحسابي $Xi - \bar{X}$	الفرق المطلق $\lvert xi - \bar{x} \rvert$
1	3	3-2 =1	1
2	1.5	1.5 -2 = -0.5	0.5
3	2.5	2.5-2 = 0.5	0.5
4	1	1-2 = -1	1
5	2	2-2 = 0	0
المجموع	**10**	**0**	**3**

$$M.D = \frac{\sum \lvert xi - \bar{x} \rvert}{N}$$

$$= \frac{3}{5}$$

$$= 0.6 \quad kg$$

وهذا يعني ان معدل انحراف الوزن عن المتوسط = 0.6 كغم

التباين Variance

هو من اهم مقاييس التشتت و يعبر عن مجموع مربعات انحرافات القيم عن وسطها الحسابي ويرمز له (δ^2) للمجتمع و (S^2) للعينة

$$\delta^2 = \frac{\sum (xi - \mu)^2 \ N}{\ }$$

μ= المتوسط الحسابي للمجتمع N= حجم المجتمع

ويمكن كتابة القانون اعلاه بصيغة اخرى:

$$\delta^2 = \frac{\sum xi^2 - \frac{(\sum xi)^2}{N}}{N}$$

أما تباين العينة

$$S^2 = \frac{\sum (xi - \bar{x})^2 \ n}{-1}$$

$\bar{x}$ = المتوسط الحسابي للعينة n = حجم العينة

ويمكن ان تكتب المعادلة بالصيغة التالية:

$$S^2 = \frac{\sum xi^2 - \frac{(\sum xi)^2}{n}}{n - 1}$$

مبادئ الإحصاء

المحاضرة الرابعة

مثال: جد قيمة التباين للقيم التالية التي تمثل وزن اللحم الصافي لعينة من ستة رؤوس من الغنم.

35, 34, 40, 38, 37, 32 kg

الحل//

- إيجاد مجموع القيم
- إيجاد مجموع مربعات القيم

X_i	X_i^2
35	1225
34	1156
40	1600
38	1444
37	1369
32	1024
$\sum X_i = 216$	$\sum X_i^2 = 7818$

$$s^2 = \frac{\sum x_i^2 - \frac{(\sum x_i)^2}{n}}{n-1}$$

$$= \frac{7818 - \frac{(216)2}{6}}{6-1}$$

$$= \frac{7818 - \frac{46656}{6}}{6-1}$$

$$= \frac{42}{5}$$

$$= 8.4\ Kg^2$$

وهذا يعني ان متوسط تباين كل قيمة عن المتوسط الحسابي للمجتمع هو 8.4 كغم^2 وبما ان هذه الوحدات غير متداولة في الحياة العامة او غير مألوفة (كغم^2) لهذا يمكن التعبير عن التشتت بوحدات قياس اعتيادية وذلك عن طريق استخدام مقياس تشتت يطلق علية بالانحراف القياسي او المعياري.

مبادئ الإحصاء

الانحراف القياسي

Standard Deviation

الانحراف القياسي (S) هو عبارة عن الجذر التربيعي للتباين (S^2) وعلية فان الانحراف القياسي يمكن حسابة من المعادلات الآتية بالنسبة للمجتمع او العينة:

$$\text{للمجتمع } \delta = \sqrt{\frac{\sum x^2 - \frac{(\sum xi)^2}{N}}{N}}$$

$$\text{للعينة } S = \sqrt{\frac{\sum x^2 - \frac{(\sum xi)^2}{n}}{n-1}}$$

مثال: جد قيمة الانحراف المعياري للبيانات المعطاة في المثال السابق؟

$$S = \sqrt{\frac{\sum x^2 - \frac{(\sum xi)^2}{N}}{N-1}} = \sqrt{\frac{7818 - \frac{(216)^2}{6}}{5}} = \sqrt{8.4} = 2.89 \text{ kg}$$

ويبدو واضحا ان الانحراف المعياري يأخذ بنظر الاعتبار جميع القيم دون استثناء ويعطي التشتت مقاسا بوحدات قياس متداولة ومتعارف عليها. ولهذه الاسباب وغيرها فان الانحراف المعياري يعتبر أكثر مقاييس التشتت شيوعا سواء بصيغته المباشرة او عن طريق المقاييس المشتقة منه كمعامل الاختلاف.

معامل الاختلاف

Coefficient of Variation

أن درجة تشتت مفردات مجموعة معينة قد تختلف عن درجة تشتت مجموعة أخرى، وقد يكون هذا الاختلاف كبيرا او صغيرا. وبناء على ذلك، يستخدم معامل الاختلاف كوسيلة لمقارنة درجات التشتت بين مجموعات مختلفة.

ويرمز لمعامل الاختلاف بالرمز .C. V ويحسب وفق المعادلة الاتية:

$$C.\,V. = \frac{S_X}{\bar{x}} \times 100$$

S_X هو الانحراف القياسي

$\bar{x}$ هو الوسط الحسابي

مبادئ الإحصاء

مثال// أجريت تجربة لدراسة طول النبات (سم) والحاصل (كغم دونم) لمحصول الذرة الصفراء، وتم الحصول على النتائج المبينة في الجدول ادناه:

	الطول	كمية الحاصل
الوسط الحسابي	200	800
الانحراف القياسي	16	36

قارن بين تشتت الطول والحاصل (ايهما أكثر تشتتاً)

الحل//

$$C.V. = \frac{S_x}{\bar{x}} \times 100$$

$$C.V. = \frac{16}{200} \times 100$$ بالنسبة للطول

$$= 8\ \%$$

$$C.V. = \frac{36}{800} \times 100$$ بالنسبة للحاصل

$$= 4.5\ \%$$

نستنتج ان الطول كان أكثر تشتتاً

نلاحظ انه لو قارنا التشتت بمقياس الانحراف القياسي لكان التشتت أكبر في الحاصل.

المصادر//

1. الراوي، خاشع محمود. 1989. المدخل الى الإحصاء. وزارة التعليم العالي والبحث العلمي. جامعة الموصل.
2. طبيه، احمد عبدالسميع. 2007. مبادئ الإحصاء، عمان. دار البداية. ر.أ:

www.daralbedayah.com (2007/6/1709)

3. David, M. Lane. Introduction to Statistics. Online Edition.

وزارة التعليم العالي والبحث العلمي
جامعة الموصل ـ كلية الزراعة والغابات
قسم المحاصيل الحقلية ـ المرحلة الاولى
أحصاء زراعي

أحصاء زراعي

المحاضرة الأولى

(الارتباط)

م.م. ظفر عبدالرزاق فرحان النجماوي

مبادئ الإحصاء

الارتباط

الارتباط البسيط Simple Correlation

وهو مقياس لتحديد طبيعة العلاقة وقوتها بين متغيرين مستقلين مثل x و y، وتسمى العلاقة طردية او موجبة اذا كانت زيادة قيم أحد المتغيرين يصحبها زيادة قيم المتغير الاخر، واذا اخذت القيم اتجاهين متعاكسين تسمى العلاقة عكسية او سالبة.

تقدير معامل الارتباط Estimation of Correlation Coefficient

ان معامل الارتباط والذي يرمز له بالرمز r لعينة عشوائية حجمها n لمتغيرين (x وy)

(x1,y1),(x2,y2),….(xn,yn)

يحسب من المعادلة التالية

$$r = \frac{\sum x_i y_i - \frac{(\sum x_i)(\sum y_i)}{n}}{\sqrt{\left(\sum x_i^2 - \frac{(\sum x_i)^2}{n}\right)\left(\sum y_i^2 - \frac{(\sum y_i)^2}{n}\right)}}$$

تتراوح قيمة معامل الارتباط (r) بين +1 و -1 أو أي قيمة بينهما (-1 ≤ r ≤ 1)، وهتين القيمتين هما درجتا الارتباط التام، الطردي التام (+1) والعكسي التام (-1)

وتجدر الإشارة الى ان معامل الارتباط بين متغيرين هو مقياس وصفي للترابط الخطي بينهما، وعندما تكون r=0 فان ذلك يعني عدم وجود ترابط خطي بينهما وليس بالضرورة عدم وجود علاقة بينهما.

قسم المحاصيل الحقلية **الإحصاء مبادئ**

المحاضرة الخامسة

مثال//

احسب معامل الارتباط للبيانات التالية والتي تمثل طول وعرض الورقة لنبات ما.

عرض الورقة (x)	13	19	13	18	14	17	14	17	15	16
طول الورقة (y)	15	22	13	20	13	20	15	19	15	18

الحل: لايجاد معامل الارتباط (r) حسب المعادلة السابقة نرتب البيانات في جدول كما يلي:

xi	yi	xiyi	x_i^2	y_i^2
13	15	195	169	225
19	22	418	361	484
13	13	169	169	169
18	20	360	324	400
14	13	182	196	169
17	20	340	289	400
14	15	210	196	225
17	19	323	289	361
15	15	225	225	225
16	18	288	256	324
∑xi= 156	∑yi= 170	∑xiyi= 2710	$\sum x_i^2 = 2474$	$\sum y_i^2 = 2982$

$$r = \frac{\sum x_i y_i - \frac{(\sum x_i)(\sum y_i)}{n}}{\sqrt{(\sum x_i^2 - \frac{(\sum x_i)^2}{n})(\sum y_i^2 - \frac{(\sum y_i)^2}{n})}}$$

$$r = \frac{2710 - \frac{(156)(170)}{10}}{\sqrt{(2474 - \frac{(156)^2}{10})(2982 - \frac{(170)^2}{10})}}$$

$$= 0.95$$

نوع العلاقة: طردية (موجبة)

اختبار معنوية الارتباط:

ان الحكم على المعنوية الإحصائية لمعامل الارتباط يعتمد على اختبارات إحصائية محددة، وذلك بمقارنة القيمة المطلقة لـ r المحسوبة مع قيمة r الجدولية عند درجة حرية n-2 ، حيث n = عدد ازواج المشاهدات، وعند مستوى الاحتمالية المطلوب، فاذا كانت قيمة r المحسوبة اكبر من او تساوي

قسم المحاصيل الحقلية

الإحصاء مبادئ

المحاضرة الخامسة

r الجدولية فان هذا يعني ان الارتباط معنوي اما اذا كانت قيمة r المحسوبة اقل من r الجدولية فان هذا يعني ان الارتباط غير معنوي ففي المثال السابق نستخرج قيمة r الجدولية من جدول r عند درجة حرية 8 ومستوى احتمالية 0.05 فنجدها تساوي 0.63

الاستنتاج: بما أن قيمة r المحسوبة (0.95) هي اكبر من قيمة r الجدولية (0.63) نستنتج ان الارتباط معنوي.

جدول (r)

df\α	0.2	0.1	0.05	0.02	0.01	0.001
1	0.951057	0.987688	0.996917	0.999507	0.999877	0.999999
2	0.800000	0.900000	0.950000	0.980000	0.990000	0.999000
3	0.687049	0.805384	0.878339	0.934333	0.958735	0.991139
4	0.608400	0.729299	0.811401	0.882194	0.917200	0.974068
5	0.550863	0.669439	0.754492	0.832874	0.874526	0.950883
6	0.506727	0.621489	0.706734	0.788720	0.834342	0.924904
7	0.471589	0.582206	0.666384	0.749776	0.797681	0.898260
8	0.442796	0.549357	0.631897	0.715459	0.764592	0.872115
9	0.418662	0.521404	0.602069	0.685095	0.734786	0.847047
10	0.398062	0.497265	0.575983	0.658070	0.707888	0.823305
11	0.380216	0.476156	0.552943	0.633863	0.683528	0.800962
12	0.364562	0.457500	0.532413	0.612047	0.661376	0.779998
13	0.350688	0.440861	0.513977	0.592270	0.641145	0.760351
14	0.338282	0.425902	0.497309	0.574245	0.622591	0.741934
15	0.327101	0.412360	0.482146	0.557737	0.605506	0.724657
16	0.316958	0.400027	0.468277	0.542548	0.589714	0.708429
17	0.307702	0.388733	0.455531	0.528517	0.575067	0.693163
18	0.299210	0.378341	0.443763	0.515505	0.561435	0.678781
19	0.291384	0.368737	0.432858	0.503397	0.548711	0.665208
20	0.284140	0.359827	0.422714	0.492094	0.536800	0.652378
21	0.277411	0.351531	0.413247	0.481512	0.525620	0.640230
22	0.271137	0.343783	0.404386	0.471579	0.515101	0.628710
23	0.265270	0.336524	0.396070	0.462231	0.505182	0.617768
24	0.259768	0.329705	0.388244	0.453413	0.495808	0.607360
25	0.254594	0.323283	0.380863	0.445078	0.486932	0.597446
26	0.249717	0.317223	0.373886	0.437184	0.478511	0.587988
27	0.245110	0.311490	0.367278	0.429693	0.470509	0.578956
28	0.240749	0.306057	0.361007	0.422572	0.462892	0.570317
29	0.236612	0.300898	0.355046	0.415792	0.455631	0.562047
30	0.232681	0.295991	0.349370	0.409327	0.448699	0.554119
35	0.215598	0.274611	0.324573	0.380976	0.418211	0.518898
40	0.201796	0.257278	0.304396	0.357787	0.393174	0.489570
45	0.190345	0.242859	0.287563	0.338367	0.372142	0.464673
50	0.180644	0.230620	0.273243	0.321796	0.354153	0.443201
60	0.164997	0.210832	0.250035	0.294846	0.324818	0.407865
70	0.152818	0.195394	0.231883	0.273695	0.301734	0.379799
80	0.142990	0.182916	0.217185	0.256525	0.282958	0.356816
90	0.134844	0.172558	0.204968	0.242227	0.267298	0.337549
100	0.127947	0.163782	0.194604	0.230079	0.253979	0.321095
125	0.114477	0.146617	0.174308	0.206245	0.227807	0.288602
150	0.104525	0.133919	0.159273	0.188552	0.208349	0.264316
175	0.096787	0.124036	0.147558	0.174749	0.193153	0.245280
200	0.090546	0.116060	0.138098	0.163592	0.180860	0.229840
250	0.081000	0.103852	0.123607	0.146483	0.161994	0.206079
300	0.073951	0.094831	0.112891	0.133819	0.148019	0.188431
350	0.068470	0.087814	0.104552	0.123957	0.137131	0.174657
400	0.064052	0.082155	0.097824	0.115997	0.128339	0.163520
450	0.060391	0.077466	0.092248	0.109397	0.121046	0.154273
500	0.057294	0.073497	0.087528	0.103808	0.114870	0.146436
600	0.052305	0.067103	0.079920	0.094798	0.104911	0.133787
700	0.048427	0.062132	0.074004	0.087789	0.097161	0.123935
800	0.045301	0.058123	0.069234	0.082135	0.090909	0.115981
900	0.042711	0.054802	0.065281	0.077450	0.085727	0.109385
1000	0.040520	0.051993	0.061935	0.073484	0.081340	0.103800
1500	0.033086	0.042458	0.050582	0.060022	0.066445	0.084822
2000	0.028654	0.036772	0.043811	0.051990	0.057557	0.073488
3000	0.023397	0.030027	0.035775	0.042457	0.047006	0.060027
4000	0.020262	0.026005	0.030984	0.036773	0.040713	0.051996
5000	0.018123	0.023260	0.027714	0.032892	0.036417	0.046512

قسم المحاصيل الحقلية

مبادئ الإحصاء

المحاضرة الخامسة

α : مستوى الاحتمالية

df: درجات الحرية

المحاضرة الخامسة

مثال //

جد قيمة معامل الارتباط وبين نوع الارتباط مع اختبار معنوية الارتباط للمتغيرين x وy

2	10	8	1	6	X
9	3	4	7	5	y

الحل//

لإيجاد معامل الارتباط (r) نرتب البيانات في جدول كما يلي:

x_i	y_i	x_iy_i	x_i^2	y_i^2
6	5	30	36	25
1	7	7	1	49
8	4	32	64	16
10	3	30	100	9
2	9	18	4	81
$\sum x_i=27$	$\sum y_i=28$	$\sum x_iy_i=117$	$\sum x_i^2=205$	$\sum y_i^2=180$

$$r = \frac{\sum x_i y_i - \frac{(\sum x_i)(\sum y_i)}{n}}{\sqrt{\left(\sum x_i^2 - \frac{(\sum x_i)^2}{n}\right)\left(\sum y_i^2 - \frac{(\sum y_i)^2}{n}\right)}}$$

$$r = \frac{117 - \frac{(27)(28)}{5}}{\sqrt{\left(205 - \frac{(27)^2}{5}\right)\left(180 - \frac{(28)^2}{5}\right)}}$$

$$r = \frac{-34.2}{\sqrt{(59.2)(23.2)}}$$

$$r = \frac{-34.2}{37.05}$$

$$r = -0.92$$

العلاقة عكسية (سالبة)

لاختبار معنوية الارتباط نستخرج قيمة r الجدولية من جدول r عند درجة حرية 3 (n-2) ومستوى احتمال 0.05 فنجدها تساوي 0.87

الإحصاء مبادئ

المحاضرة الخامسة

بما أن قيمة r المحسوبة (-0.92) قيمتها المطلقة (0.92)أكبر من قيمة r الجدولية (0.87) نستنتج ان الارتباط معنوي على مستوى احتمال 0.05.

اما على مستوى احتمال 0.01 نجد ان قيمة r الجدولية تساوي 0.95

بما أن قيمة r المحسوبة (-0.92) قيمتها المطلقة (0.92) اقل من قيمة r الجدولية (0.95) نستنتج ان الارتباط غير معنوي على مستوى احتمال 0.01

المصادر//

1. الراوي، خاشع محمود. 1989. المدخل الى الإحصاء. وزارة التعليم العالي والبحث العلمي. جامعة الموصل.
2. طبيه، احمد عبدالسميع. 2007. مبادئ الإحصاء، عمان. دار البداية. ر.أ: (2007/6/1709)

www.daralbedayah.com

3. David, M. Lane. Introduction to Statistics. Online Edition.

وزارة التعليم العالي والبحث العلمي
جامعة الموصل - كلية الزراعة والغابات
قسم المحاصيل الحقلية - المرحلة الاولى
أحصاء زراعي

أحصاء زراعي

المحاضرة الثامنة

(اختبار الفرضيات)

م.م. ظفر عبدالرزاق فرحان النجماوي

مبادئ الإحصاء

المحاضرة الثامنة (اختبار الفرضيات)

اختبار الفرضيات Tests of Hypotheses

يعد اختبار الفرضيات الإحصائية من أهم المواضيع في مجال اتخاذ القرارات

مصطلحات ضرورية:

1. **الفرضية الاحصائية:** هي عبارة عن أدعاء أو تصريح قد يكون صائبا أو خاطئا حول معلمة أو أكثر لمجتمع أو لمجموعة من المجتمعات.

تؤخذ العينة من المجتمع وتدرس وتستخدم جميع المعلومات المتحصل عليها للوصول الى قرار بقبول أو رفض الفرضية الاحصائية. في حالة كون بيانات العينة تساند النظرية فأن الفرضية تقبل، أما أذا كانت البينات تناقض النظرية ففي هذه الحالة ترفض الفرضية. تجدر الاشارة هنا الى ان قبولنا الفرضية الاحصائية هو ناتج عن عدم وجود أدلة كافيه لرفضها من بيانات العينة ولذلك فأن قبولنا لهذه الفرضية لا يعني بالضرورة كونها صحيحه أما أذا

رفضنا الفرضية بن "اء على المعلومات الموجودة في بيانات العينة فأن ذلك يعني بأن الفرضية خاطئة. لذلك فأن

الباحث أو الاحصائي يحاول دائما أن يضع الفرضية بشكل يأمل أن يرفضها، فمثلا أذا أراد الباحث أن يقارن صنفا جديدا من الحنطة مع الصنف المحلي فأنه يضع فرضية مفادها بأنه لا يوجد فرق معنوي بين الصنفين.

أن الفرضية التي يضعها الباحث على امل أن يرفضها تدعى بفرضية العدم Null Hypothesis ويرمز لها بـH_0، ورفضنا لفرضية العدم يقودنا الى قبول فرضية بديلة عنها هذه الفرضية تدعى الفرضية البديلة ويرمز لها بـ H_1 **2. الاخطاء التي يقع فيها الباحث**

أن طريقة اتخاذنا القرارات قد يقودنا الى الوقوع في نوعين من الخطأ هما

- خطأ من النوع الاول: يقع الباحث في الخطأ من النوع الاول أذا رفض فرضية العدم عندما تكون هي الفرضية الصحيحة.
- خطأ من النوع الثاني: يقع الباحث في الخطأ من النوع الثاني أذا قبل فرضية العدم عندما تكون هي الفرضية الخاطئة

كما هو موضح في الجدول التالي:

الحالة الحقيقية / القرار	H_0 صحيحه	H_1 خاطئة
قبول H_0	قرار صائب	خطاء من النوع الثاني
رفض H_0	خطاء من النوع الاول	قرار صائب

3. مستوى المعنوية Level of Significant

أو مستوى الاحتمال Probability level Size

أو حجم الاختبار of the test

يعرف **مستوى المعنوية** بأنه درجة الاحتمال الذي نرفض به فرضية العدم H_0 عندما تكون هي الصحيحة أو بعبارة أخرى هو احتمال الوقوع في الخطاء من النوع الاول ويرمز لها بـ α أي:

$$\alpha = P\ (\text{Type I error})$$
$$= P(\text{Reject } H_0 \mid H_0 \text{ is true})$$

فاذا أستخدمنا مستوى معنوية 5% ورفضنا فرضية العدم H_0 نقول بأن هناك فرق معنوي، أما أذا رفضت فرضية العدم بمستوى معنوية 1% فيقال بأن هناك فرق معنوي جدا وفي الحالة الاولى فهذا يعني بأنه لو تكررت التجربة 100 مره فأن 5 منها تؤيد صحة فرضية العدم و 95 لا تطابق فرضية العدم.

منطقة الرفض: هي تلك المنطقة التي أذا وقعت قيمة المختبر الإحصائي داخلها تسبب في رفض فرضية العدم H_0. وتحدد منطقة الرفض بعد تعيين مستوى المعنوية

تسمى المنطقة تحت المنحني غير منطقة الرفض بمنطقة القبول التي أذا وقعت قيمة المختبر الاحصائي فيها يسبب في قبول فرضية العدم H_0.

المحاضرة الثامنة (اختبار الفرضيات)

المختبر الإحصائي: هو عبارة عن متغير عشوائي له توزيع احتمالي معلوم ويصف المختبر الاحصائي العلاقة بين القيم النظرية في المجتمع والقيم المعلومة أو القيم المحسوبة في المجتمع وعادة تقارن قيمة المختبر الاحصائي المحسوب من العينة مع قيمته المستخرجة من توزيعه الاحتمالي (الموجود في جداول خاصة) ومنها تتخذ القرارات برفض أو قبول فرضية العدم H_0.

خطوات اختبار الفرضية:

1- تحديد نوع توزيع المجتمع أي هل هو توزيع طبيعي أو نوع من التوزيعات الاخرى أو توزيع معتدل. أن البحث يوضح التوزيع حول المتغير العشوائي هل هو توزيع طبيعي أو توزيع ذي حدين علما بأن معظم الظواهر يكون توزيعها مشابه للتوزيع الطبيعي.

2- صياغة فرضية العدم والبديلة، تصاغ الفرضيات بلغة معالم المجتمع فمثلا عندما يكون الاختبار عن المتوسط الحسابي بأنه يساوي قيمة معينة، فرضا " أن معدل درجات الطلبة 70 إذن فرضية العدم H_0: $\mu=70$ والفرضية البديلة $\mu \neq 70$

3- اختبار مستوى المعنوية، يكون اختيار مستوى المعنوية مسبقا" من قبل الباحث وفي البحوث العلمية يأخذ مستوى المعنوية 5% (معنوي) أو 1% (معنوي جدا") وبعد ذلك تحدد منطقة القبول

4- المختبر الإحصائي، حيث يتم اختيار المختبر الإحصائي الذي سيكون قاعدة لاختبار الفرضيات، والمختبرات الإحصائية منها Z او T أو X^2 أو اختبار F.

5- جمع البيانات من العينة وحساب المختبر الاحصائي بحسب صيغة حساب معينه نذكرها لاحقا".

6- اتخاذ القرارات أذا وقعت قيمة المختبر الإحصائي في منطقة الرفض ترفض فرضية العدم وتقبل البديلة ويقال بأن هنالك فرق معنوي بين القيم النظرية للمجتمع والقيم المحسوبة من العينة وإذا وقعت قيمة المختبر في منطقة القبول تقبل فرضية العدم ويقال بأنه لا يوجد فرق معنوي بين نتائج العينة والقيم النظرية للمجتمع.

اختبار الفرضيات الخاصة بالمتوسط الحسابي الحقيقي للمجتمع μ: 1- اختبار Z:

شروط استخدامه

- حجم العينة العشوائية يجب أن يكون اكبر من 30 ($n \geq 30$)
- توزيع المجتمع يجب أن يكون طبيعيا أو مقارب منه
- تباين المجتمع σ^2 معلوم

وهو من الاختبارات التي تستخدم في حالة متوسط حسابي واحد ويحسب المختبر الاحصائي Z من المعادلة:

$$Z = \frac{\bar{Y}-\mu_0}{\frac{\sigma}{\sqrt{n}}}$$

حيث:

$\bar{Y}$: الوسط الحسابي للعينة μ_0: الوسط الحسابي للمجتمع

σ: الانحراف القياسي للمجتمع، ويمكن أن يستعاض عنه بالانحراف القياسي للعينة (S) n: حجم العينة

بعد اختيار مستوى المعنوية فأن منطقة الرفض ومنطقة القبول ستحدد

لنفرض بأن Y يتوزع توزيعا طبيعيا بمتوسط μ وتباين σ^2 أي: $Y \sim N(\mu_1\ \sigma^2)$ وكانت الرغبة باختبار فرضية العدم ضد الفرضية البديلة

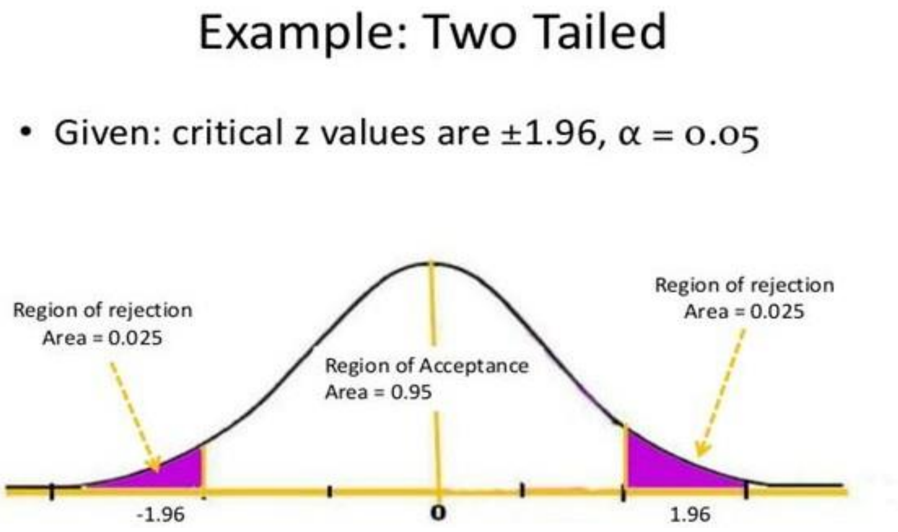
$$H_0 \quad \mu = \mu_0$$
$$H_1 \quad \mu \neq \mu_0$$

فاذا أخترنا $\alpha= .05$ كمستوى معنوية فمنطقة الرفض ستكون في هذه الحالة عن يمين ويسار القيمة μ_0، أي ذات طرفين لأن الفرضية البديلة $\mu \neq \mu_0$ من μ_0 $H_1:$ يعني بأنه في حالة رفض فرضية العدم فأن μ قد تكون أقل أو أكثر ومنطقة الرفض هي

$$\mu_0 \pm Z_{\alpha/2}\, \sigma_{\bar{Y}}$$

حيث أن : $Z_{\alpha/2}$ هو المتغير القياسي الذي يقطع المنحني بـ $\alpha/2$ في كلا الطرفين من التوزيع الطبيعي وبما أن $\alpha = 0.05$ أذاً

$$Z_{\alpha/2} = Z_{.025} = 1.96$$



جدول مناطق الرفض عند اختبار $H_0 : \mu=\mu_0$ ضد ثلاث حالات من الفرضية البديلة H_1 وان: $Y \sim N(\mu_1\ \sigma^2_y)$ عندما يكون حجم العينة كبير و σ^2 معلوم:

حالات الاختبار	اتخاذ القرارات		
	$\alpha=\alpha$ ترفض H_0 إذا كانت قيمة	$\alpha=0.05$ ترفض H_0 اذا كانت قيمة	$\alpha=0.01$ ترفض H_0 اذا كانت قيمة
$H_0 : \mu=\mu_0$ $H_1: \mu\neq\mu_0$	$Z \geq Z_{\alpha/2}$ and $Z \leq -Z_{\alpha/2}$ or $\lvert Z \rvert \geq Z_{\alpha/2}$	$Z \geq 1.96$ and $Z \leq -1.96$ or $\lvert Z \rvert \geq 1.96$	$Z \geq 2.58$ and $Z \leq -2.58$ or $\lvert Z \rvert \geq 2.58$
$H_0 : \mu = \mu_0$ $H_1 : \mu > \mu_0$	$Z \geq Z_\alpha$	$Z \geq 1.65$	$Z \geq 2.33$
$H_0 : \mu = \mu_0$ $H_1 : \mu < \mu_0$	$Z \leq Z_\alpha$	$Z \leq -1.65$	$Z \leq -2.33$

علما ان Z في هذه الحالة هو $\mathbf{Z}= \dfrac{\bar{Y}-\mu_0}{\frac{\sigma}{\sqrt{n}}}$

مثال: ينتج معمل لتعليب الزيت علبا متوسط وزنها 15 كغم بانحراف قياسي 0.5 كغم، أجري اختبار للتأكد من أن المعمل لازال ينتج عند المستوى المطلوب وبذلك أخذت عينة من 50 علبة ووجد أن متوسط وزنها 14.8 كغم فاذا

كان وزن العلب متغير عشوائي يتوزع توزيعا طبيعيا فهل تدل العينة على أن أنتاج المعمل لازال 15 كغم، على مستوى معنوية $\alpha = 0.01$.

$H_0: \mu=15$

$H_1: \mu\neq15$

منطقة الرفض: $Z > 2.58$

$Z < -2.58$

$\mu_0=15$

$\bar{Y}= 14.8$

$\sigma= 0.5$

$n= 50$

$$Z= \frac{\bar{Y}-\mu_0}{\frac{\sigma}{\sqrt{n}}}$$

$$Z= \frac{14.8-15}{\frac{0.5}{\sqrt{50}}}$$

$$Z = \mathbf{-2.83}$$

القرار: بما أن قيمة Z المحسوبة 2.83 – هي أقل من 2.58 – أي أن قيمة Z المحسوبة تقع في منطقة الرفض لذا نرفض فرضية العدم وتقبل الفرضية البديلة أي أن المعمل لا ينتج علبا أوزانها 15 كغم.

او بما أن قيمة Z المحسوبة المطلقة 2.83 هي أكبر من 2.58 أي أن قيمة Z المحسوبة تقع في منطقة الرفض لذا نرفض فرضية العدم وتقبل الفرضية البديلة أي أن المعمل لا ينتج علبا أوزانها 15 كغم.

مثال// اذا كان معدل انتاج احد أصناف الحنطة هو 1600 كغم هـ$^{-1}$ بانحراف قياسي =15 كغم. وادعى أحد الباحثين انه استنبط سلالة من هذا الصنف تعطي انتاجا أكبر. ولاختبار صحة ادعاء الباحث اخذت عينة عشوائية مؤلفة من 81 نباتا ووجد ان متوسط انتاجها 1630 كغم هـ$^{-1}$.

هل نتائج العينة تؤيد ادعاء الباحث عند مستوى معنوية = 0.01؟ الحل:

فرضية العدم $H0: \mu=1600$

الفرضية البديلة $H1: \mu>1600$

مستوى المعنوية $\alpha= 0.01$

منطقة الرفض $Z \geq 2.33$

المختبر الاحصائي

$\bar{Y} =1630$ $n=81$ $\sigma=S=15$

$$Z= \frac{\bar{Y}-\mu_0}{\frac{\sigma}{\sqrt{n}}}$$

$$Z= \frac{1630-1600}{\frac{15}{\sqrt{81}}}$$

$$= 18$$

القرار: بما ان قيمة Z المحسوبة (18) هي اكبر من قيمة Z الجدولية (2.33) لذا نرفض فرضية العدم H_0 ونقبل الفرضية البديلة H_1 أي ان ادعاء الباحث كان صحيحا.

- **اختبار Z لمتوسطين حسابيين**

1- وضع فرضية العدم H_0

$$H_0 : \mu_1 - \mu_2 = d$$

علما أن d تمثل الفرق الافتراضي بين المتوسطين

2- تحديد الفرضية البديلة وتكون الفرضية البديلة أحد الاشكال الثلاثة الاتية

$H_1: \mu_1 - \mu_2 \neq d$
$H_1: \mu_1 - \mu_2 > d$
$H_1: \mu_1 - \mu_2 < d$

3- تحديد مستوى المعنوية α أما ان يكون 5% او 1% 4- تحديد منطقة الرفض:

ترفض فرضية العدم وتقبل الفرضية البديلة كما يلي

- في حالة كون الفرضية البديلة

$H_1: \mu_1 - \mu_2 \neq d$

تكون منطقة الرفض $|Z| \geq Z_{\alpha/2}$ (القيمة المطلقة لـ Z المحسوبة) ترفض فرضية العدم وتقبل البديلة

- في حالة كون الفرضية البديلة

$H_1: \mu_1 - \mu_2 > d$

ترفض فرضية العدم وتقبل الفرضية البديلة عندما تكون قيمه Z المحسوبة أكبر من الجدولية

- في حالة كون الفرضية البديلة

$H_1: \mu_1 - \mu_2 < d$

ترفض فرضية العدم وتقبل الفرضية البديلة عندما تكون قيمه Z المحسوبة أقل من الجدولية علما أن قيمة المختبر الاحصائي يحسب بالصيغة الرياضية

$$Z = \frac{\bar{Y}1 - \bar{Y}2 - d}{\sqrt{\frac{S1^2}{n1} + \frac{S2^2}{n2}}}$$

مثال// في امتحان لمادة الاحصاء لمجموعتين من الطلبة، المجموعة الاولى من 50 طالبة متوسط درجاتهن في الامتحان 82 بانحراف قياسي (S) 8، والمجموعة الثانية من 75 طالب كان متوسط درجاتهم 76 بانحراف قياسي (S) 6، فهل يوجد فرق معنوي بين مستوى الطالبات والطلاب في مستوى معنوية 5%

الحل:

فرضية العدم الفرضية $H_0 : \mu_1 - \mu_2 = 0$

البديلة $H_1 : \mu_1 - \mu_2 \neq 0$

مستوى المعنوية 5% اذاً منطقة الرفض هي $|Z| \geq Z_{\alpha/2}$

$|Z| \geq 1.96$

$$Z = \frac{82 - 76 - 0}{\sqrt{\frac{64}{50} + \frac{36}{75}}}$$

$$Z = \frac{6}{\sqrt{1.28 + 0.48}}$$

$$= 4.52$$

القرار: بما أن قيمة Z المطلقة المحسوبة (4.52) أكبر من Z الجدولية (1.96) لذا ترفض فرضية العدم (H0) وتقبل البديلة (H1) أي يوجد فرق معنوي بين درجات الطلاب والطالبات تحت مستوى احتمال 5%.

2- اختبار توزيع t المعتدل t-Distribution

وهو من التوزيعات الاحتمالية المستمرة المهمة والمستخدمة مع العينات صغيرة الحجم (أقل من 30 وأكبر من 4) وتحدد قيمة t المحسوبة وفق المعادلة التالية ونقارنها مع t الجدولية بدرجات حرية n-1

$$t= \frac{\bar{Y}-\mu}{\frac{S}{\sqrt{n}}}$$

مثال // لو أريد تقدير المتوسط الحسابي الحقيقي لوزن الدجاجة لمجتمع معين من الدجاج على أساس اختيار عينة عشوائية ممثلة لهذا المجتمع مكون من عشر دجاجات وكانت أوزان الدجاجات هي: 1.2 ,1.1 ,0.9 ,0.8 ,1.2 ,1.1 ,0.8 ,1.2 ,1.0 ,0.9 كغم، **بانحراف قياسي** (S) = 0.16، فان اختبار فرضية العدم القائلة بأن المتوسط الحسابي الحقيقي لوزن الدجاجة لمجتمع معين من الدجاج هو 1.25 كغم يتم كما يلي:

$$H_0 : \mu_x = 1.25$$
$$H_1 : \mu_x \neq 1.25$$

المتوسط الحسابي للعينة = 1.02 كغم

$$t= \frac{\bar{Y}-\mu}{\frac{S}{\sqrt{n}}}$$

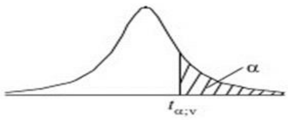
$$t= \frac{1.02-1.25}{\frac{0.16}{\sqrt{10}}}$$

$$= - 4.510$$

القرار: بما أن قيمة t المحسوبة المطلقة (4.510) تزيد على قيمة t الجدولية (2.821) بمستوى معنوية 1% ودرجات حرية 9 لذا فأننا نرفض فرضية العدم بمستوى معنوية 1%.

Table of the Student's *t*-distribution

The table gives the values of $t_{\alpha;\nu}$ where $\Pr(T_\nu > t_{\alpha;\nu}) = \alpha$, with ν degrees of freedom



ν \ α	0.1	0.05	0.025	0.01	0.005	0.001	0.0005
1	3.078	6.314	12.076	31.821	63.657	318.310	636.620
2	1.886	2.920	4.303	6.965	9.925	22.326	31.598
3	1.638	2.353	3.182	4.541	5.841	10.213	12.924
4	1.533	2.132	2.776	3.747	4.604	7.173	8.610
5	1.476	2.015	2.571	3.365	4.032	5.893	6.869
6	1.440	1.943	2.447	3.143	3.707	5.208	5.959
7	1.415	1.895	2.365	2.998	3.499	4.785	5.408
8	1.397	1.860	2.306	2.896	3.355	4.501	5.041
9	1.383	1.833	2.262	2.821	3.250	4.297	4.781
10	1.372	1.812	2.228	2.764	3.169	4.144	4.587
11	1.363	1.796	2.201	2.718	3.106	4.025	4.437
12	1.356	1.782	2.179	2.681	3.055	3.930	4.318
13	1.350	1.771	2.160	2.650	3.012	3.852	4.221
14	1.345	1.761	2.145	2.624	2.977	3.787	4.140
15	1.341	1.753	2.131	2.602	2.947	3.733	4.073
16	1.337	1.746	2.120	2.583	2.921	3.686	4.015
17	1.333	1.740	2.110	2.567	2.898	3.646	3.965
18	1.330	1.734	2.101	2.552	2.878	3.610	3.922
19	1.328	1.729	2.093	2.539	2.861	3.579	3.883
20	1.325	1.725	2.086	2.528	2.845	3.552	3.850
21	1.323	1.721	2.080	2.518	2.831	3.527	3.819
22	1.321	1.717	2.074	2.508	2.819	3.505	3.792
23	1.319	1.714	2.069	2.500	2.807	3.485	3.767
24	1.318	1.711	2.064	2.492	2.797	3.467	3.745
25	1.316	1.708	2.060	2.485	2.787	3.450	3.725
26	1.315	1.706	2.056	2.479	2.779	3.435	3.707
27	1.314	1.703	2.052	2.473	2.771	3.421	3.690
28	1.313	1.701	2.048	2.467	2.763	3.408	3.674
29	1.311	1.699	2.045	2.462	2.756	3.396	3.659
30	1.310	1.697	2.042	2.457	2.750	3.385	3.646
40	1.303	1.684	2.021	2.423	2.704	3.307	3.551
60	1.296	1.671	2.000	2.390	2.660	3.232	3.460
120	1.289	1.658	1.980	2.358	2.617	3.160	3.373
∞	1.282	1.645	1.960	2.326	2.576	3.090	3.291

المصادر//

1. الراوي، خاشع محمود. 1989. المدخل الى الإحصاء. وزارة التعليم العالي والبحث العلمي. جامعة الموصل.
2. طبيه، احمد عبدالسميع. 2007. مبادئ الإحصاء، عمان. دار البداية. ر.أ: (2007/6/1709) www.daralbedayah.com
3. David, M. Lane. Introduction to Statistics. Online Edition