مبادئ وتعاريف إحصائية

مقدمة:

يعتبر الإحصاء من العلوم الأكثر أهمية وفائدة في كل مجالات الحياة العملية فعن طريقه يمكن معالجة الكثير من المسائل المتعلقة بالحياة اليومية وأيضا لا يمكن الاستغناء عنه في الدراسات والأبحاث في كثير من فروع العلوم الأخرى. الإحصاء بالمعنى العلمي هو مجموعة من الطرائق العلمية والوسائل الحديثة التي تستخدم لجمع البيانات وتحليلها والوصول منها إلى نتائج معينة تخص ظاهرة معينة وتستخدم هذه النتائج عادة في عملية اتخاذ القرار.

يستخدم الإحصاء في مختلف الدراسات العلمية عن الظواهر سواء كانت طبيعية أو اجتماعية. فالظواهر الطبيعية مثلا في دوائر الأرصاد الجوي تجمع بيانات حول درجات الحرارة وسرعة الرياح والرطوبة الجوية وكمية الأمطار وغيرها في كل يوم وفي كل ساعة من ساعات اليوم وتستخدم هذه البيانات ليس لوصف حالة الطقس فقط بل للتنبؤ بما سيكون عليه الطقس بالمستقبل. أما الظواهر الاجتماعية مثلا في حالة إجراء تعداد سكاني ومن نتائجه تقدم فائدة كبيرة في مجالات متعددة مثل مجالات العمل والتشغيل والدفاع والتعليم والصحة....الخ.

تعريف علم الإحصاء: هو العلم الذي يبحث في :

- جمع المعلومات و البيانات و الحقائق الخاصة بمختلف الظواهر و تصميمها في صورة رقمية و تصنيفها في جداول منظمة و تمثيلها بيانياً.
- تحليل البيانات و استخلاص النتائج منها و اتخاذ القرارات المناسبة.
- مقارنة الظواهر ببعضها و دراسة العلاقة بينها.

يقسم الإحصاء إلى قسمين هما:

1) **الإحصاء الوصفي:** هو عملية جمع البيانات وتبويبها وتبسيطها وتلخيصها وعرضها بصورة رقمية عن طريق حساب بعض المقاييس الإحصائية عليها. أو بمعنى أخر يتناول الإحصاء الوصفي وصف ظاهرة معينة في زمان و مكان معينين عن طريق جمع بيانات حولها دون أي محاولة لتعميم النتائج على ظواهر أخرى في زمان آخر ومكان آخر. مثلا الظاهرة التي تدرس هي أعمار الطلاب اللذين يدرسون الإحصاء في كلية الزراعة /قسم البستنة/ المرحلة الأولى لعام معين ولنفرض إننا جمعنا البيانات عن عمر كل طالب وعددهم 150 طالب وقسمنا مجموع الأعمار على عدد الطلاب نحصل على مؤشر يسمى المتوسط الحسابي للعمر (المعدل) وليكن مثلا 19 سنة هذه النتيجة التي حصلنا عليها لا يمكن تعميمها على طلاب الكلية ككل ولا على طلاب الجامعة نظرا لخصوصية المجموعة الطلابية التي أخذنا منها المعلومات.

2) **الإحصاء الاستدلالي:** هو استخدام معلومات جزئية عن المجتمع لبيان مواصفات ذلك المجتمع ففيه نقوم بسحب عينة التي هي جزء من مجتمع كبير عادة ثم ندرس خصائص ظاهرة معينة في هذه العينة ومن ثم تعمم النتائج المتحصل عليها على المجتمع بدرجة معينة من الثقة فهو يعتمد بالدرجة الأساسية على نظرية الاحتمال. يضم هذا القسم فرعين هما:

- **التقدير:** يهتم بإيجاد قيم تقديرية للاستدلال منها على القيم الحقيقية لمصدر جمع البيانات.
- **اختبار الفرضيات:** يتضمن اختبار الفرضيات التي توضع كتفسير أولي للظاهرة المراد دراستها للوصول منها إلى قرار بقبولها أو رفضها.

تعاريف:

المجتمع Population: هو المجموعة التي تتكون من كل المفردات محل الدراسة. قد يكون المجتمع كبير جدا أو صغير حسب تعريف الباحث لمجتمع بحثه وقد يكون محدود الحجم مثلا عدد الأشجار في بستان معين أو غير محدود الحجم مثل عدد الأسماك في البحر.

العينة Sample: هي جزء من مفردات المجتمع يتم اختيارها بحيث تكون ممثلة للمجتمع ككل.

البيانات Data: تسمى المعلومات التي يتم جمعها و تنظيمها بواسطة الإحصائيين بيانات. و تسمى أيضاً **مشاهدات Observation** وهي تعد بمثابة المواد الاولية التي يتعامل معها الباحث وقد تكون أرقام كما في عدد التفرعات للنبات الواحد وليكن مثلا 6 فرع/ نبات فان هذا العدد يمثل مشاهدة واحدة وقد تكون وصف للظاهرة مثلا قصير, طويل او مصاب , سليم.

المتغير Varaible: هو أي ظاهرة تظهر اختلافات بين مفرداتها مثلا , ارتفاع النبات, الحاصل, عدد الأزهار, لون الأزهار.

- البيانات تتكون من نوعين من المتغيرات.

المتغيرات الوصفية أو النوعية **Qualitative Variables**	**المتغيرات الكمية** **Quantitative Variables**
هي المتغيرات التي تكون في صورة غير عددية وغير مقاسه بالأرقام مثل لون الأزهار, لون العيون, الجنس (ذكر أو أنثى) مصاب أو سليم أو متوسط الإصابة	هي المتغيرات التي تكون في صورة عددية ويمن قياسها مباشرة بالأرقام مثل طول النبات. الوزن , العمر....الخ المتغيرات الكميـــة تنقسم إلى قسمين : أ - المتغيرات الكمية المتصلة(المستمرة):هي المتغيرات التي يمكن أن تأخذ جميع القيم بين حدي التغير. فيمكنها أن تأخذ أي قيمة صحيحة أو كسور. مثال أوزان مجموعة من الأفراد و كذلك الأطوال و العمر, ارتفاع النبات ب- المتغيرات الكمية المنفصلة(المتقطعة):هي المتغيرات التي لا تمكنها طبيعتها من أن تأخذ جميع القيم بين حدي التغير. مثال عدد أفراد الأسرة و عدد الحوادث . فهي تأخذ أرقام صحيحة, عدد الأزهار

معالم Parameters : النتائج التي تؤخذ من دراسة المجتمع

إحصائيةVarieties : النتائج التي تؤخذ من دراسة العينة.

إن الاختلاف بين القيمة الإحصائية وقيمة معالم المجتمع يسمى بالخطأ العيني أو الخطأ المعياري. إن أهم ما يميز المجتمع هو الدقة في القياس واتساع استخدام القياس.

طرائق جمع البيانات

1. **طريقة المسوحات:** وفيها تجمع البيانات كما هي في الميدان. البيانات تكون متوفرة سواء كانت منشورة آو غير منشورة نتيجة جمعها من قبل الجهات الرسمية وغير الرسمية ولإغراض مختلفة. مثل الجهاز المركزي للإحصاء أو مراكز البحوث أو محطات التجارب أو المنظمات الإقليمية والدولية أو الشركات أو الجمعيات التعاونية... الخ.

2 - طريقة التجربة العلمية: يتم فيها جمع البيانات من خلال التحكم أو السيطرة على واحد أو أكثر من العوامل لتحديد تأثير هذه العوامل على المتغير المراد دراسته كما هو الحال عند دراسة تأثير نوع معين من السماد في إنتاج محصول ما مع مراعاة تثبيت العوامل الأخرى اللازمة لإنتاج المحصول وعادة تتطلب طريقة التجربة العلمية مهارة فنية عالية وتخصص في مجال الدراسة.

بعد إن تجمع البيانات يتم عرضها كالآتي:

1. الاشكال وتشمل:

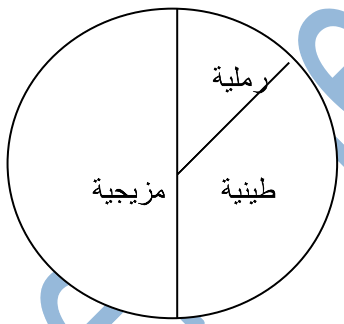
أ) دوائر مجزأة: تستخدم الدائرة اذا كانت بيانات الظاهرة موضوع الدراسة عبارة عن مجموع عام مقسم إلى أجزائه المختلفة و تمثل المساحة الكلية للدائرة المجموع الكلي ثم نقسم الدائرة إلى قطاعات و نميز بينها بالتضليل أو بالألوان على الرسم .

تحسب الزاوية المقابلة لكل قطاع بالقانون التالي :

$$\text{زاوية القطاع} = \frac{\text{قيمة القطاع}}{\text{المجموع العام}} \times 360^\circ$$

مثال: البيانات الآتية : تمثل توزيع خصائص تربة معينة في احد االمناطق ولسنة معينة وهي كما يلي:
تربة مزيجية 50%, تربة طينية 30% , تربة رملية 20% .

$$\text{التربة المزيجية} = \frac{50}{100} \times 360^\circ = 180^\circ$$



$$\text{الطينية} = \frac{30}{100} \times 360^\circ = 108^\circ$$

$$\text{الرملية} = \frac{20}{100} \times 360^\circ = 72^\circ$$

ب - الأعمدة البيانية:

هي عبارة عن مجموعة من الأعمدة الرأسية أو المستطيلات المتساوية القاعدة و التي يتناسب ارتفاعها مع البيانات التي تمثلها. و تستخدم ، لإظهار التطور الذي يطرأ على ظاهرة ما على مدار عدة سنوات ، و عادة ما يؤخذ المحور الرأسي لتمثيل قيم الظاهرة و المحور الأفقي لتمثيل الزمن ، و يجب مراعاة أن يقسم المحور الرأسي

بحيث يسمح مقياس الرسم بإظهار جميع الظاهرة و كذلك يراعى أن تكون المسافات بين الأعمدة متساوية و قاعدة الأعمدة متساوية .

1- الأعمدة البيانية المزدوجة:
تستخدم إذا كان الهدف من الرسم مقارنة ظاهرتين أو أكثر لعدة سنوات . نحصل عليها برسم عمودين متلاصقين يمثلان قيمة ظاهرتين محل الدراسة في كل سنة بحيث يتناسب طول العامود مع العدد الذي يمثله و نفرق بين الأعمدة بالتضليل أو بالألوان على الرسم .. مع ضرورة أن تكون المستطيلات متساوية و المسافات بينها متساوية

2- الأعمدة البيانية المجزأة :
تستخدم في نفس الحالة التي تستخدم فيها الأعمدة البيانية المزدوجة و يتم الحصول عليها برسم عامود واحد يمثل جملة الظواهر محل الدراسة في كل سنة كما في حالة الأعمدة البيانية البسيطة ، ثم تقسم كل عامود الى مكوناته بحيث يتناسب كل جزء مع العدد الذي يمثله و نميز بين هذه الأجزاء بالتضليل أو بالألوان على الرسم .

2- المنحنى:
يستخدم المنحى أساساً لتوضيح الإتجاه العام للظاهرة خلال فترة من الزمن ، و يتم الحصول على المنحنى بتوقيع مجموعة من النقاط على مستوى المحاور . يمثل المحور الأفقي الزمن و المحور الرأسي الظاهرة أو قيم المتغير ، ثم نوصل هذه النقاط بخط متصل فنحصل على المنحنى .

3- الرسم الدائري (الدائرة)

الرسوم البيانية في حالة البيانات الكمية المتصلة :

1- المدرج التكراري:
هي أعمدة متلاصقة تبين العلاقة بين الفئات و التكرارات.
2- المضلع التكراري:
تبين العلاقة بين مراكز الفئات و التكرارات .... و نقوم بوضع نقاط تساوي مركز الفئة و التكرار (X , Y) و يجب إغلاق المضلع من الأمام وذلك بافتراض فئة سابقة للفئة الأولى و من الخلف و ذلك بافتراض فئة لاحقة للفئة الأخيرة ، و تكرار كل منها (صفر) ويمكن رسم المضلع بطريقتين : يرسم على رسم مستقل أو يرسم على المدرج التكراري.
3- المنحنى التكراري:
يرسم بنفس طريقة المضلع التكراري و لكن باليد . ويبين العلاقة بين مراكز الفئات و التكرارات .

المقاييس الاحصائية

1 – مقاييس التوسط او النزعة المركزية
تعريف النزعه المركزية:
ميل المفردات الى التراكم او الاقتراب حول قيمة معينة ، هي القيمة المتوسطة.

الوسط الحسابي :
هو مجموع القيم أو القراءات مقسومة على عددها .... يرمز له µ (للمجتمع) و $\bar{Y}$, $\bar{Z}$, $\bar{X}$ (للعينة)
القاعدة الرياضيه لحساب الوسط الحسابي ..

الوسط الحسابي = $\frac{\text{مجموع القيم}}{\text{عددها}}$ $\frac{\sum Xi}{n}$

المنوال : (Mo)
هي القيمة الأكثر شيوعاً... أو القيمة التي تكرر أكثر من غيرها

الوسيط :
هي القيمة التي تتوسط مجموعة من القسم بعد ترتيبها تصاعدياً أو تنازلياً ... أي أنه القيمة التي يكون عدد القيم الأصغر منها مساوياً لعدد القيم الأكبر منها و يرمز إليها بـ M

مزايا و عيوب مقاييس النزعة المركزية :
1- الوسط الحسابي :

مزايا الوسط الحسابي	عيوب الوسط الحسابي
السهولة في الحساب ، لذلك هو أكثر المتوسطات استخداماً.	لا يمكن إيجاده بالرسم
تدخل جميع القيم في حسابه .	يتأثر بالقيم الشاذة

2- المنوال:
المزايا :
1- سهل الحساب سواء بالحساب أو بالرسم
2- لا يتأثر بالقيم الشاذة .
3- المقياس الوحيد المناسب للبيانات الوصفية .

العيـــوب : قد لا يوجد في بعض الحالات و قد يوجد أكثر من منوال إذا كان اكبر تكرار يتحقق لقيمين أو أكثر

3- الوســـيط :
المزايــا:
1- من مزايا الوسيط أنه لا يتأثر بالقيم الشاذة .
2- يمكن الحصول علية بالرسم .

العيـــوب:
لا يدخل في حسابه جميع القيم ، سوى قراءة واحدة أو قراءتين من المجوعة كلها .

2 - مقايـيس التشـــتت

تعريف التشتت:
هو مدى قرب أو بعد مفردات الظاهرة بعضها عن بعض .....

و هذا التشتت يكون صغيراً إذا كان الاختلاف بين قيم المفردات قليلاً و يكون كبيراً إذا كان الاختلاف بين قيم المفردات كبيرا ... و لذلك يمكننا اتخاذ مقدار تشتت القيم مقاييس لمعرفة قرب القيم أو تباعدها بعضها عن بعض .

المدى :

هو أبسط مقاييس التشتت و يعرف بأنه الفرق بين أكبر و أصغر قراءة في المجموعة و يرمز له R .

إذا كان المدى صغير تكون المجموعة متقاربة أي متجانسة ..

و إذا كان المدى كبير فإنه يدل على أن مفردات المجموعة مبعثرة و مشتتة و متباعدة عن بعضها .

مزايا و عيوب المدى :

المزايا: سهل الحساب .

العيوب:

1- لا يدخل في حسابه إلا قراءتين ، هما أكبر وأصغر قيمة .

2- شديد التأثر بالقيم الشاذة .

الانحراف المعياري : S (للعينة) او σ (للمجتمع) هو أكثر مقايــيس التشتت استخداماً . ويساوي الجذر التربيعي للتباين

شروط استخدام الانحراف المعياري للمقارنة بين ظاهرتين:

1-أن تكون وحدة القياس للظاهرتين متساوية .

2- تساوي حجم العينات

3- تساوي الأوساط الحسابية للظاهرتين .

*إذا فقد أو اختلف احد هذه الشروط نستخدم مقاييس التشتت النسبي ...

التباين : S^2 (للعينة) او σ^2 (للمجتمع) : او يسمى متوسط المربعات يحسب من حاصل قسمة مجموع مربع الفروقات بين اي قيمة عن وسطها الحسابي مقسومة على درجات الحرية (n -1) هذا يمثل تباين العينة

تباين المجتمع يحسب من حاصل قسمة مجموع مربع الفروقات بين اي قيمة عن وسطها الحسابي مقسومة على حجم المجتمع (N).

مقايــيس التشتت النسبي :

معامل الاختلاف : (C.V) و يستخدم إذا اختلف احد شروط الانحراف المعياري و للمقارنة بين تشتت مجموعتين .

يحسب من حاصل قسمة الانحراف المعياري على المتوسط الحسابي كنسبة مئوية

سؤال : أي المقايـيس يفضل استخدامها في حساب القيم التالية :

1- إذا كانت القيم شاذة : يفضل استخدام المنوال (Mo) أو الوسيط (M)

2- إذا كانت القيم غير شاذة : نستخدم أي من المقاييس الأخرى

التصميم التام التعشية Completely Randomized Design (CRD)

مميزاته:

-أسهل التصاميم تنفيذا وتحليلاً.

- يمكن استخدام هذا التصميم عندما تكون الوحدات التجريبية متجانسة الى درجة كبيرة.

- بسبب متطلبات التجانس ، قد يكون من الصعب استخدام هذا التصميم في الحقل.

-يكون مناسب في التجارب ذات عدد قليل من المعاملات.

العشوائية:

- يتم توزيع المعاملات على الوحدات التجريبية المتجانسة بشكل عشوائي تمامًا.

- كل وحدة تجريبية لديها نفس الاحتمال في تلقي أي معاملة

- يتم تنفيذ العشوائية باستخدام جدول الأرقام العشوائية ، الكمبيوتر ، البرنامج ، إلخ.

مثال على التوزيع العشوائي

- لديك 4 معاملات يرمز لها (A و B و C و D) واريد تطبيقها في 5 مكررات ، سيكون عدد الوحدات التجريبية الكلية تساوي 20 وحدة تجريبية. ترقم الوحدات التجريبية من 1 – 20

1 D	2 D	3 B	4 C	5 D	6 C	7 A	8 A	8 B	10 D
11 C	12 B	13 A	14 B	15 C	16 B	17 C	18 D	19 A	20 A

ملاحظة : توزيع المعاملات بشكل عشوائي تام على كل الوحدات التجريبية بحيث ان كل وحدة تجريبية لها نفس الاحتمال في تلقي اي معاملة

محاسن CRD

1. التصميم مرن جدًا (على سبيل المثال عدد المعاملات والتكرارات يتحدد فقط بعدد الوحدات التجريبية المتجانسة المتوفره).

2. التحليل الإحصائي بسيط مقارنة بالتصاميم الأخرى.

3. فقدان المعلومات عن احد المعاملات يكون تاثيره صغير مقارنة بالتصميمات الأخرى وذلك لان درجات الحرية للخطأ التجريبي تكون اكبر من اي تصميم اخر يضم نفس عدد المعاملات والمكررات.

مساوئ CRD

1. إذا لم تكن الوحدات التجريبية متجانسة يؤدي التوزيع العشوائي الى الفشل في تقليل الاختلاف وبالتالي يكون هناك فقدان بالدقة في النتائج.

2. عادة ما يكون التصميم الأقل كفاءة ما لم تكن الوحدات التجريبية متجانسة.

3. غير مناسب في حالة عدد كبير من المعاملات في التجربة

مثال: لديك البيانات التالية لتجربة ضمن تصميم CRD لمقارنة ثلاث درجات حرارة وباربع تكرارات في احد الصفات المقاسة (الوزن الجاف غم للنبات)

	المشاهدات Y_{ij}			
التكرارات	T1	T2	T3	
1	23	42	47	
2	36	26	43	
3	31	47	43	
4	33	34	39	Y..
Yi.(المجموع)	123	149	172	444
$\bar{Y}_i$(المتوسطات)	30.75	37.25	43	37

الحل:

Step 1.

نفرض فرضية العدم H0 التي تنص على ان المتوسطات الحسابية للمعاملات الثلاثة متساوية أي ان متوسط الوزن الجاف للنبات بتأثير درجات الحرارة متساوي ولا يوجد فرق مهم من الناحية الاحصائية بينهم

HO: T1 = T2 =T3

Hi: T1 ≠ T2 ≠ T3 الفرضية البديلة على الاقل يوجد متوسط واحد غير متساوي

Step 2. Calculate the Correction Factor.

$$CF = \frac{Y^2}{rt} = \frac{444^2}{4*3} = 16,428.0$$

Step 3. Calculate the Total SS

$$TotalSS = \sum Y_{ij}^2 - CF$$

$$= (23^2 + 36^2 + 31^2 + ... + 39^2) - CF$$

$$= 17,108 - 16,428$$

$$= 680.0$$

Step 4. Calculate the Treatment SS (TRT SS)

$$TRTSS = \sum \frac{Y_i^2}{r} - CF$$

$$= \left(\frac{123^2}{4} + \frac{149^2}{4} + \frac{172^2}{4}\right) - 16,428$$

$$= 16728.5 - 16428.0$$

$$= 300.5$$

Step 5. Calculate the Error SS

$$\text{Error SS} = \text{Total SS} - \text{Treatment SS}$$

$$= 680 - 300.5$$

$$= 379.5$$

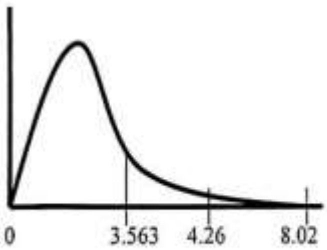
Step 6. Complete the ANOVA table

Sources of variation	Df	SS	MS	F
Treatment	t-1 = 2	300.5	150.25	3.563^{NS}
Error	t(r-1) = 9	379.5	42.167	
Total	rt-1 = 11	680.0		

Step 7. Look up Table F-values.

$F_{0.05;2,9} = 4.26$
$F_{0.01;2,9} = 8.02$

Step 8. Make conclusions.



-Since F-calc (3.563) < 4.26 we fail to reject Ho: $\mu_1 = \mu_2 = \mu_3$ at the 95% level of confidence.
-Since F-calc (3.563) < 8.02 we fail to reject Ho: $\mu_1 = \mu_2 = \mu_3$ at the 99% level of confidence.

Step 9. Calculate Coefficient of Variation (CV).

$$\%CV = \frac{s}{\bar{Y}} * 100$$

Remember that the Error MS = s^2.

$$\%CV = \frac{\sqrt{42.167}}{\left(\frac{444}{4*3}\right)} * 100$$

$$= (6.494/37) * 100$$

$$= 17.6\%$$

ANOVA for Any Number of Treatments with Unequal Replication

Given the following data:

	Treatment				
Replicate	A	B	C	D	
1	2.0	1.7	2.0	2.1	
2	2.2	1.9	2.4	2.2	
3	1.8	1.5	2.7	2.2	
4	2.3		2.5	1.9	
5	1.7		2.4		
$Y_{i.}$	10	5.1	12	8.4	Y_=35.5
$\sum Y_{ij}^2$	20.26	8.75	29.06	17.7	

Step 1. Write the hypotheses to be tested.

$H_o: \mu_1 = \mu_2 = \mu_3 = \mu_4$
H_A: At least one of the means is different from one of the other means.

Step 2. Calculate the Correction Factor.

$$CF = \frac{Y^2}{\sum r_i} = \frac{35.5^2}{17} = 74.132$$

Step 3. Calculate the Total SS

$$TotalSS = \sum Y_{ij}^2 - CF$$

$$= (2.0^2 + 2.2^2 + 1.8^2 + ... + 1.9^2) - CF$$

$$= 75.77 - 74.132$$

$$= 1.638$$

Step 4. Calculate the Treatment SS (TRT SS)

$$TRTSS = \sum \frac{Y_i^2}{r_i} - CF$$

$$= \left(\frac{10^2}{5} + \frac{5.1^2}{3} + \frac{12^2}{5} + \frac{8.4^2}{4}\right) - 74.132$$

$$= 75.110 - 74.132$$

$$= 0.978$$

Step 5. Calculate the Error SS

Error SS = Total SS – Treatment SS

$= 1.638 - 0.978$

$= 0.660$

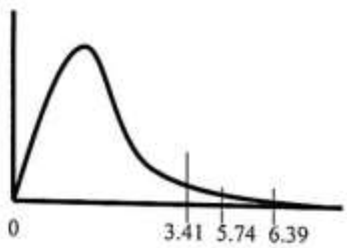
Step 6. Complete the ANOVA table

Sources of variation	Df	SS	MS	F
Treatment	t-1 = 3	0.978	0.326	6.392**
Error	By subtraction = 13	0.660	0.051	
Total	Total number of observations -1 = 16	1.638		

Step 7. Look up Table F-values.

$F_{0.05;3,13} = 3.41$
$F_{0.01;3,13} = 5.74$

Step 8. Make conclusions.



- Since F-calc (6.392) > 3.41 we reject Ho: $\mu_1 = \mu_2 = \mu_3 = \mu_4$ at the 95% level of confidence.
- Since F-calc (6.392) > 5.74 we reject Ho: $\mu_1 = \mu_2 = \mu_3 = \mu_4$ at the 99% level of confidence.

Step 9. Calculate Coefficient of Variation (CV).

$$\%CV = \frac{s}{\bar{Y}} * 100$$

Remember that the Error MS = s^2_e

$$\%CV = \frac{\sqrt{0.051}}{\left(\frac{35.5}{17}\right)} * 100$$

$$= (0.2259 / 2.088) * 100$$

$$= 10.82\%$$

(المصدر: الأسدي, ماهر حميد سلمان. 2019. GenStat لتحليل التجارب الزراعية. دار الوارث للطباعة والنشر. العراق. ع ص: 304.)

2) تجربة عاملية بتصميم القطاعات العشوائية الكاملة

Factorial Experiment Designing with RCBD

يتضمن جدول تحليل التباين لمثل هذه التجارب ما يلي:

S.O.V	df	SS	MS	F cal.
Blocks	r - 1	$SS_r = R - CF \quad R = \frac{\sum yi^2}{ab}$		
treats	ab -1	$SS_t = AB - CF \quad AB = \frac{\sum y^2}{r}$		
A	a - 1	$SS_a = A - CF \quad A = \frac{\sum ai^2}{br}$	$= \frac{SS_a}{df_a}$	$= \frac{MS_a}{MS_e}$
B	b - 1	$SS_b = B - CF \quad B = \frac{\sum bi^2}{ar}$	$= \frac{SS_b}{df_b}$	$= \frac{MS_b}{MS_e}$
AB	(a -1)(b - 1)	$SS_{ab} = AB - B - A + CF$	$= \frac{SS_{ab}}{df_{ab}}$	$= \frac{MS_{ab}}{MS_e}$
Error	(ab – 1)(r - 1)	$SS_e = RAB - AB - R + CF$	$= \frac{SS_e}{df_e}$	
Total	abr – 1	$SS_T = RAB - CF \quad RAB = \sum x^2$		

مثال:

أجريت تجربة لمعرفة تأثير الكثافة النباتية 44000 و 41066 و 38133 للذرة العلفية صنف فجر 1 ورش حامض الهيومك بتركيز 3% و 6% فضلاً عن معاملة المقارنة (رش ماء مقطر فقط) في عدد العرانيص.نبات$^{-1}$ بثلاث مكررات، وكانت نتائج كما في الجدول التالي:

Treats		R1	R2	R3	∑xi	Means
Verities (A)	Extracts (B)					
44000	0.0	1.6	1.5	1.5	4.6	1.5
	3%	1.9	2.0	2.1	6	2.0
	6%	2.3	2.5	2.3	7.1	2.3
41066	0.0	1.7	1.6	1.7	5	1.6
	3%	2.4	2.3	2.2	6.9	2.3
	6%	2.5	2.3	2.4	7.2	2.4
38133	0.0	1.9	1.8	1.8	5.5	1.8
	3%	2.4	2.3	2.5	7.2	2.4
	6%	2.7	2.6	2.5	7.8	2.6
∑ri		19.0	18.9	19.4	57.3	

المطلوب:

1. حل البيانات وفق إختبار F عند مستوى معنوية 0.05؟

2. حدد الأفضل بين الكثافات النباتية وتراكيز الرش والتداخل عند مستوى إحتمال 0.05؟

3. أعرض النتائج وفق التحليل الإحصائي؟

الحل:

1) إستخراج مجموع المكررات ومتوسطاتها في الجدول أعلاه، ثم إعداد جدول التداخل AB كما يلي:

AB	b1	b2	b3	∑ai	Mean A
a1	4.6	6.0	7.1	17.7	1.9
a2	5.0	6.9	7.2	19.1	2.1
a3	5.5	7.2	7.8	20.5	2.2
∑bi	15.1	20.1	22.1	57.3	
Mean B	1.7	2.2	2.5		

2) حساب معامل التصحيح Correction Factor

$$C.F. = \frac{(\sum xi)^2}{n} \qquad n = a \times b \times r \qquad = \frac{(57.3)^2}{3\times3\times3} = 121.6$$

3) حساب مجموع مربعات القطاعات Sum Square of Blocks R

$$SS_r = R - CF \qquad R = \frac{\sum yi^2}{br}$$

$$R = \frac{(19.0)^2 + \ldots + (19.4)^2}{3 \times 3} = 121.61 \qquad SS_r = 121.61 - 121.60 = 0.01$$

4) حساب مجموع مربعات العامل الأول (الكثافات النباتية) Sum Square of Factor A

$$SS_a = A - CF \qquad A = \frac{\sum ai^2}{br}$$

$$A = \frac{(17.7)^2 + \ldots + (20.5)^2}{3 \times 3} = 122 \qquad SS_a = 122 - 121.6 = 0.43$$

5) حساب مجموع مربعات العامل الثاني (رش حامض الهيومك) Sum Square of Factor B

$$SS_b = B - CF \qquad B = \frac{\sum bi^2}{ar} - C.F$$

$$B = \frac{(15.1)^2 + \ldots + (22.1)^2}{3 \times 3} = 124.4 \qquad SS_b = 124.4 - 121.6 = 2.8$$

6) حساب مجموع مربعات المعاملات Sum Square of treatments

$SS_t = AB - CF$ $\qquad AB = \frac{\sum y^2}{r}$

$AB = \frac{(4.6)^2 + \dots + (8.7)^2}{3} = 124.9 \qquad SS_t = 124.9 - 121.6 = 3.3$

7) حساب مجموع مربعات معاملات التداخل بين العاملين Sum Square of Interaction AB

$SS_{ab} = AB - B - A + CF \qquad = 124.9 - 124.4 - 122 + 121.6 = 0.1$

8) حساب مجموع المربعات الكلية Sum Square of Total

$SS_T = RAB - CF \qquad RAB = \sum x^2$

$= (1.6)^2 + \dots + (2.5)^2 = 125.13 \qquad SS_T = 125.13 - 121.6 = 3.53$

9) حساب مجموع مربعات الخطأ التجريبي Sum Square of Error

$SS_e = RAB - AB - R + CF \qquad = 125.13 - 124.9 - 121.61 + 121.6 = 0.22$

10) حساب درجات الحرية Degree of Freedom للمعاملات والكلي والخطأ التجريبي

$df_t = ab - 1 \qquad = 3\times3 - 1 = 8 \qquad df_r = r - 1 \qquad = 3 - 1 = 2$
$df_a = a - 1 \qquad = 3 - 1 = 2 \qquad df_b = b - 1 \qquad = 3 - 1 = 2$
$df_{ab} = (a-1)(b-1) = 2\times2 = 4 \qquad df_T = abr - 1 \qquad = 27 - 1 = 26$
$df_e = (ab-1)(r-1) = (3\times3-1)(3-1) = 16$

11) حساب متوسط المربعات Mean Square للمعاملات والقطاعات والخطأ التجريبي

$MS_a = \frac{SS_a}{df_a} = \frac{0.43}{2} = 0.215 \qquad MS_b = \frac{SS_b}{df_b} = \frac{2.8}{2} = 1.4$

$MS_{ab} = \frac{SS_{ab}}{df_{ab}} = \frac{0.1}{4} = 0.02 \qquad MS_e = \frac{SS_e}{df_e} = \frac{0.22}{16} = 0.013$

12) يعد جدول تحليل التباين Analysis of Variance (ANOVA Table)

S.O.V	df	SS	MS	F cal.	F tab.
Blocks	2	0.01			
treats	8	3.3			
A	2	0.43	0.21	16.15	3.36*
B	2	2.8	1.4	107.6	3.36*
AB	4	0.1	0.02	1.53	3.01^{NS}
Error	16	0.22	0.013		
Total	26	3.53			

13) إستخراج القيمة المحسوبة لـ F (F. calculated)

$$F_{cal} = \frac{MS_a}{MS_e} = \frac{0.21}{0.013} = 16.15 \qquad F_{cal} = \frac{MS_b}{MS_e} = \frac{1.4}{0.013} = 107.6$$

$$F_{cal} = \frac{MS_{ab}}{MS_e} = \frac{0.02}{0.013} = 1.53$$

14) إستخراج قيمة F الجدولية F table من جدول F-values بتقاطع df_e (16) في المحور العمودي وفق مستوى الإحتمالية 0.05 ودرجات الحرية للمعاملات df_a و df_b (2) و df_{ab} (4) في المحور الأفقي.

بما أن قيمة F المحسوبة لكل من A و B هي أكبر من قيمة F الجدولية

∴ توجد فروق معنوية أي ترفض نظرية العدم H_0 (القائلة بعدم وجود فروق معنوية) وتقبل النظرية البديلة H_a (القائلة بوجود فروق معنوية) لكل من تأثير الكثافة النباتية والرش بحامض الهيومك، لكن التداخل بين العاملين غير معنوي ولإجراء الاختبار في هذه الحالة يجب إجراء إختبار دانكن متعدد الحدود بدل من اختبار LSD لعدم معنوية التداخل ولتوحيد طريقة الإختبار بين العاملين والتداخل بينها، ولإجراء إختبار دانكن للعوامل المفردة والتداخل بينها تتبع الخطوات الآتية:

1) مقارنة متوسطات العامل A

يتم إستخراج قيمة SSR من جدول SSR-Duncan بدلالة قيمة df_e (16) ومستوى الإحتمالية 0.05 في المحور العمودي وعدد المعاملات 3 في المحور الأفقي لمعاملات العامل A كما في المعادلة والجدول أدناه:

$$LSR = S_{\bar{x}} \times SSR$$

$$S_{\bar{x}}A = \sqrt{\frac{MSe}{br}} = \sqrt{\frac{0.013}{3\times 3}} = 0.038$$

ثم إعداد جدول لإستخراج LSR كمايلي:

SSR	3.00	3.15
$S_{\bar{x}}$	0.038	
LSR	0.11	0.119

ثم إعداد جدول للمقارنة بين متوسطات A كمايلي:

LSR		0.11	0.119
Treats	Means	a1 1.9	a2 2.1
a3	2.2	0.3*	0.1*
a2	2.1	0.2*	0.0

وتكون نتيجة الإختبار كمايلي:

Treats A	Means
a3	2.2 a
a2	2.1 b
a1	1.9 c

المتوسطات التي تشترك بالحرف نفسه لا يوجد بينها فرق معنوي حسب إختبار دانكن متعدد الحدود (DMR) عند مستوى إحتمال 0.05

∴ تفوق معنوياً الكثافة النباتية 38133 نبات.هكتار $^{-1}$ (a3) على بقية الكثافات إذ أعطى أعلى متوسط لعدد العرانيص بلغ 2.2 عرنوص.نبات $^{-1}$ بالمقارنة مع أقل متوسط للكثافة النباتية 44000 نبات.هكتار $^{-1}$ (a1) بلغ 1.9 عرنوص.نبات $^{-1}$.

2) مقارنة متوسطات العامل B

بالنظر لتساوي الخطأ التجريبي (بسط معادلة $S_{\bar{x}}$) ومستويات العاملين A و B (مقام معادلة $S_{\bar{x}}$ وأعداد وقيم SSR) لذلك تكون قيمة LSR مشتركة بينهما وتجرى المقارنة بين متوسطات العامل B (تراكيز الرش) باستعمال قيم الجدول السابق كمايلي:

LSR		0.11	0.119
Treat B	Mean B	b1 1.7	b2 2.2
b3	2.5	0.8*	0.3*
b2	2.2	0.5*	0.0

وتكون نتيجة الإختبار كمايلي:

Treats	Means
b3	2.5 a
b2	2.2 b
b1	1.7 c

المتوسطات التي تشترك بالحرف نفسه لا يوجد بينها فرق معنوي حسب إختبار DMR عند مستوى إحتمال 0.05

∴ تفوق معنوياً تركيز رش حامض الهيومك 6% (b3) على بقية التراكيز إذ أعطى أعلى متوسط لعدد العرانيص بلغ 2.5 عرنوص.نبات$^{-1}$ بالمقارنة مع أقل متوسط لتركيز الرش بالماء المقطر (b3) الذي بلغ 1.7 عرنوص.نبات$^{-1}$.

3) مقارنة متوسطات التداخل AB

يتم إستخراج قيمة SSR من جدول SSR-Duncan بدلالة قيمة df_e (16) ومستوى الإحتمالية 0.05 في المحور العمودي وعدد المعاملات 9 في المحور الأفقي لمعاملات التداخل AB كما في المعادلة والجدول أدناه:

$$LSR = S_{\bar{x}} \times SSR$$

$$S_{\bar{x}}A = \sqrt{\frac{MSe}{r}} = \sqrt{\frac{0.013}{3}} = 0.06$$

ثم إعداد جدول لإستخراج LSR كمايلي:

SSR	3.00	3.15	3.23	3.3	3.34	3.37	3.39	3.41
$S_{\bar{x}}$	0.06							
LSR	0.18	0.189	0.19	0.198	0.2	0.2	0.2	0.2

ثم إعداد جدول

للمقارنة بين متوسطات التداخل AB (الصنف × تراكيز الرش) كمايلي:

Treat AB	Mean AB	0.18 a1b1 1.5	0.189 a2b1 1.7	0.19 a3b1 1.8	0.198 a1b2 2.0	0.2 a2b2 2.3	0.2 a3b2 2.4	0.2 a2b3 2.4	0.2 a1b3 2.4
a3b3	2.6	1.1*	0.9*	0.8*	0.6*	0.3*	0.2*	0.2*	0.2*
a1b3	2.4	0.9*	0.7*	0.6*	0.4*	0.1	0.0	0.0	0.0
a2b3	2.4	0.9*	0.7*	0.6*	0.4*	0.1	0.0	0.0	
a3b2	2.4	0.9*	0.7*	0.6*	0.4*	0.1	0.0		
a2b2	2.3	0.8*	0.6*	0.5*	0.3*	0.0			
a1b2	2.0	0.5*	0.3*	0.2*	0.0				
a3b1	1.8	0.3*	0.1	0.0					
a2b1	1.7	0.2*	0.0						

وتكون نتيجة الإختبار كمايلي:

Treats AB	Means AB
a3b3	2.6 a
a1b3	2.4 b
a2b3	2.4 b
a3b2	2.4 b
a2b2	2.3 b
a1b2	2.0 c
a3b1	1.8 d
a2b1	1.7 d
a1b1	1.5 e

المتوسطات التي تشترك بالحرف نفسه لا يوجد بينها فرق معنوي حسب إختبار DMR عند مستوى إحتمال 0.05

∴ تفوق معنوياً معاملة التداخل بين الكثافة النباتية 38133 مع رش حامض الهيومك بتركيز 6% (a3b3) على بقية المعاملات إذ أعطى أعلى متوسط بلغ 2.6 عرنوص.نبات$^{-1}$ بالمقارنة مع أقل متوسط لمعاملات تداخل الكثافة النباتية 44000 مع تركيز الرش بالماء المقطر فقط (a1b1) التي أعطت 1.5 عرنوص.نبات$^{-1}$.