922034

الفصل	السنة	القسم
الثاني	الرابعة	معلم صف

رياضيات القياس

مقرر

الرياضيات (القياس)

أولاً : مقدمة :

نعرض بعض المفاهيم الأساسية في الاحتمالات

نرمي قطعة نقود معدنية مرة واحدة .

نسمي مجموعة النتائج الممكنة لهذه التجربة فضاء العينة و النتائج الممكنة $S =$ {عدد و شعار}

- نرمي قطعة النقود مرتين متتاليتين أو نرمي قطعتي نقود مرة واحدة فتكون النتائج الممكنة
- $S =$ {عدد عدد , شعار عدد , عدد شعار , شعار شعار}

نسمي هذه المجموعة بفضاء العينة

نرمي حجر نرد مرة واحدة فيكون فضاء العينة / الحالات الممكنة / $S = \{1,2,3,4,5,6\}$

نرمي حجري نرد معاً مرة واحدة فيكون فضاء العينة :

	١	٢	٣	٤	٥	٦
١	(1,1)	(1,2)	(1,3)	(1,4)	(1,5)	(1,6)
٢	(2,1)	(2,2)	(2,3)	(2,4)	(2,5)	(2,6)
٣	(3,1)	(3,2)	(3,3)	(3,4)	(3,5)	(3,6)
٤	(4,1)	(4,2)	(4,3)	(4,4)	(4,5)	(4,6)
٥	(5,1)	(5,2)	(5,3)	(5,4)	(5,5)	(5,6)
٦	(6,1)	(6,2)	(6,3)	(6,4)	(6,5)	(6,6)

إن عدد عناصر المجموعة S فضاء العينة يساوي $n(s) = 6^2 = 36$ نسمي كل مجموعة جزئية من المجموعة S حدثاً . فإذا كان الحدث A :ظهور الوجه ذي النقطتين على واحد من الوجهين الظاهرين فقط فإن

$$A = \{(2,1), (2,3), (2,4), (2,5), (2,6), (1,2), (3,2), (4,2), (5,2), (6,2)\}$$

أي : $n(A) = 10$ أما إذا كان الحدث B مجموع نقاط الوجهين الظاهرين أصغر تماماً من ٧ فإن : $n(B) = 15$

و يمكن الاستفادة من الجدول السابق لتحديد عناصر كل من المجموعات الآتية :

فإذا اعتبرنا أن الرمز X يدل على ظهور نقطتين على وجه واحد فقط من الوجهين الظاهرين و الرمز O يدل على مجموع نقط الوجهين الظاهرين أصغر تماماً من ٧ فإننا نحصل على الجدول التالي :

	١	٢	٣	٤	٥	٦
١	O	X	O	O	O	
٢	X	O	X	X	X	X
٣	O	X	O			
٤	O	X				
٥	O	X				
٦		X				

و نجد مباشرة من الجدول أن :

$$n(A) = 10$$

$$n(B) = 15$$

$$n(A \cap B) = 6$$

$$\mathrm{n}(\mathrm{A} \cup \mathrm{B}) = 19$$

ثانياً : المفاهيم الأساسية في الاحتمالات

الاختبار : هو تجربة مصممة يمكن إجراؤها بحيث نعلم مسبقاً مجموعة نتائجها الممكنة (لكننا لا نعلم النتيجة التي

فضاء العينة : هو مجموعة النتائج الممكنة للاختبار و سنرمز لها بالرمز S علماً أنه في بعض المراجع العلمية ترمز بالرمز Ω

الحدث : هو كل مجموعة جزئية من فضاء العينة S و سنرمز للحدث بأحد الحروف A,B,C……

مجموعة أجزاء S: هي مجموعة كل أحداث التجربة و نرمز لها بالرمز P(S) أي :

$$A \subseteq S \Leftrightarrow A \in P(S) \Leftrightarrow A \text{حدث}$$

تعريف : نقول عن الحدث A في اختبار ما ، أنه وقع إذا كانت نتيجة الاختبار ظهور أحد عناصره.

مثال : في اختبار رمي حجر النرد مرة واحدة ، نقول عن الحدث $A = \{1,2,5\}$ أنه قد وقع إذا كانت نتيجة الاختبار ظهور ١ أو ٢ أو ٥

أحداث مميزة :

- **الحدث الأكيد** S

- الحدث المستحيل $\emptyset$

- الحدثان المتنافيان (الحدثان A,B متنافيان) $\Leftrightarrow \{A \cap B = \emptyset\}$

- الحدثان المتضادان :

الحدثان B و A المتضادان $\Leftrightarrow A \cup B = S$ و $A \cap B = \emptyset \Leftrightarrow B = S/A = \grave{A}$

ونرمز للحدث المضاد A ب Ā

مثال : في تجربة رمي حجر النرد مرة واحدة يكون :

$$S = \{1,2,3,4,5,6\}$$

الحدثان $A = \{1,2,3\}$ و $B = \{2,6\}$ متنافيان

الحدثان $A = \{1,3,5\}$ و $B = \{2,4,6\}$ متضادان

الحدث $D = \{5\}$ حدث إبتدائي .

ثالثاً : الدالة الاحتمالية

تعريف : إذا كانت S فضاء العينة لاختبار معين P(S) مجموعة الأحداث فإن الدالة الاحتمالية هي كل دالة :

$P: P(S) \longrightarrow [0,1]$ يحقق الموضوعتين الآتيتين :

الموضوعة الأولى : $P(S) = 1$

الموضوعة الثانية : إذا كان A,B حدثين متنافيين فإن :

$$P(A \cup B) = P(A) + P(B)$$

ونسمي الثلاثية (S,P(S),P) فضاءً احتمالياً .

تعريف الفضاء الاحتمالي المنتظم (المتساوي الاحتمالات):

إذا كان الفضاء الاحتمالي (S,P(S),P) منتهياً وكانت احتمالات الأحداث الابتدائية متساوية قلنا إن الفضاء الاحتمالي هو فضاء منتظم (متساوي الاحتمالات)

ملاحظة : سنتعامل مع فضاء احتمالي منته .

نتائج :

إذا كان A و B حدثين من فضاء احتمالي (S,P(S),P) فإن

$$P(S) = 1$$

$$1)\ P(\emptyset) = 0$$

$$2)\ 0 \le P(A) \le 1$$

$$3)\ P\left(\grave{A}\right) = 1 - P(A)$$

$$4)\ P(A/B) = P(A) - P(A \cap B)$$

$$5)\ P(A \cup B) = P(A) + \ P(B) - P(A \cap B)$$

إذا كانت المجموعة S منتهية قلنا إن الفضاء الاحتمالي منته ، و إذا كان الحدث A

حدثاً في فضاء احتمالي منته منتظم

$$P(A) = \frac{n(A)}{n(S)} \text{فإن} (S, P(S), P)$$

تطبيقاً لهذه النتائج نورد الأمثلة التالية :

مثال: في تجربة رمي حجري نرد مرة واحدة :

الحدث A : ظهور ثلاث نقط على وجه واحد فقط من الوجهين الظاهرين .

الحدث B: ظهور وجهين مجموع نقاطهما أصغر تماماً من ٧

أحسب احتمال الأحداث :

$$A\ , B\ , A \cap B\ , A \cup B\ ,\ A/B\ , A \cap \grave{B}\ ,\ \grave{A} \cap \grave{B}\ , \qquad A \cup \grave{B}$$

الحل : نعتبر الرمز X الدال على وقوع الحدث A و الرمز O الدال على وقوع الحدث B و الرمز ⊗

الدال على وقوع الحدث $A \cap B$ نرتب جدولاً بذلك كما في الشكل و يكون :

	١	٢	٣	٤	٥	٦
١	○	○	⊗□	○	○	
٢	○	○	⊗□	○		
٣	⊗□	⊗□	○	X	X	X
٤	○	○	X			
٥	○		X			
٦			X			

من الجدول نجد :

$$n(S) = 36$$

$$n(A) = 10$$

$$n(B) = 15$$

$$n(A \cap B) = 4$$

$$n(A \cup B) = 21$$

و منه

$$P(A)=\frac{n(A)}{n(S)}=\frac{10}{36}=\frac{5}{18}$$

$$P(B)=\frac{n(B)}{n(S)}=\frac{15}{36}=\frac{5}{12}$$

$$P(A\cap B)=\frac{n(A\cap B)}{n(S)}=\frac{4}{36}=\frac{1}{9}$$

$$P(A\cup B)=\frac{n(A\cup B)}{n(S)}=\frac{21}{36}=\frac{7}{12}$$

$$P(A/B)=P(A)-P(A\cap B)=\frac{10}{36}-\frac{4}{36}=\frac{6}{36}=\frac{1}{6}$$

$$P\left(A\cap \grave{B}\right)=P(A/B)=\frac{6}{36}=\frac{1}{6}$$

$$P\left(\grave{A}\cap \grave{B}\right)=P(A\cup B)\grave{}=1-P(A\cup B)=1-\frac{21}{36}=1-\frac{7}{12}$$

$$=\frac{12}{12}-\frac{7}{12}=\frac{5}{12}$$

$$P\left(A\cup \grave{B}\right)= P(A)+P\left(\grave{B}\right)-P\left(A\cap \grave{B}\right)=\frac{10}{36}+\frac{21}{36}-\frac{6}{36}=\frac{25}{36}$$

مثال : يحوي صندوق ١٠ كرات متماثلة (٥ بيضاء و ٣ سوداء و ٢ حمراء) نسحب من الصندوق عشوائياً ثلاث كرات معاً . أحسب احتمال كل من الأحداث التالية :

١ - A : الكرات المسحوبة جميعها بيضاء .

٢ - B : الكرات المسحوبة من لون واحد.

٣ - C : الكرات المسحوبة من ألوان مختلفة (كرة من كل لون)

٤ - D : الكرات المسحوبة ليست من جميعها من لون واحد

٥ - E : كرة واحدة على الأقل من الكرات المسحوبة سوداء

الحل :

$$n(S) = C(10,3) = \frac{10 \times 9 \times 8}{3 \times 2 \times 1} = 120$$

١ - الحدث Aالكرات المسحوبة جميعها بيضاء

$$P(A) = \frac{n(A)}{n(S)} = \frac{C(5,3)}{\mathrm{C}(10,3)} = \frac{10}{120} = \frac{1}{12}$$

٢ - الحدث B هو (الكرات الثلاث بيضاء) أو (الكرات الثلاث سوداء)

$$n(B) = C(5,3) + C(3,3) = 10 + 1 = 11$$

و منه

$$P(B) = \frac{n(B)}{n(S)} = \frac{11}{120}$$

٣ - الحدث C هو الحصول (كرة بيضاء و كرة سوداء و كرة حمراء)

$$n(C) = C(5,1) \times C(3,1) \times C(2,1)$$

$$5 \times 3 \times 2 = 30$$

$$P(C) = \frac{n(C)}{n(S)} = \frac{30}{120} = \frac{1}{4}$$

٤ - الحدثان B و D متضادان ومنه

$$P(D) = P(\grave{B}) = 1 - P(B) = 1 - \frac{11}{120} = \frac{109}{120}$$

٥ - يتحقق الحدث E إذا حصلنا على (كرة سوداء و كرتين من اللونين الآخرين)

أو (كرتين سوداوان و كرة من اللونين الآخرين) أو (الكرات الثلاث سوداء) و

بالتالي :

$$n(E) = C(3,1) \times C(7,2) + C(3,2) \times C(7,1) + C(3,3)$$

$$3 \times 21 + 3 \times 7 + 1 = 8$$

$$n(E) = \frac{n(E)}{n(S)} = \frac{85}{120} = \frac{17}{24}$$

طريقة ثانية لحل الطلب (٥)

إن الحدث المتمم للحدث E هو :

E´ : ليس من الكرات المسحوبة أي كرة سوداء فيكون :

$$n(\grave{E}) = C(7,3) = \frac{7 \times 6 \times 5}{3 \times 2 \times 1} = 35$$

وبالتالي :

$$P(\grave{E}) = \frac{n(\grave{E})}{n(S)} = \frac{35}{120} = \frac{7}{24}$$

و منه

$$P(E) = 1 - P(\grave{E}) = 1 - \frac{7}{24} = \frac{17}{24}$$

تمارين

١ - إذا رمينا نرد و قطعة نقود معاً لمرة واحدة أكتب فضاء العينة ، ثم أحسب احتمالات الاحداث التالية:

١) A حدوث شعار من قطعة النقود و الرقم ٥ من النرد

٢) B حدوث شعار من قطعة النقود و رقم فردي من النرد

٣) C حدوث رقم من قطعة النقود مع (٢ أو ٣) من النرد

٢ - صندوق فيه ست بطاقات مرقمة بالأرقام ١,٢,٣,٤,٥,٦

نسحب من الصندوق بطاقتين معاً : احسب

١) عدد عناصر فضاء العينة n(S)

٢) اكتب فضاء العينة

٣) إذا كان A هو حدث ظهور العدد ٥ على إحدى البطاقتين فما هي مجموعة عناصر A

٤) إذا كان B هو حدث أن تكون مجموع رقمي البطاقتين يساوي ٦ فما هي مجموعة عناصر B

٣ - إذا رمينا قطعة نقد ثلاث مرات احسب

١) عدد عناصر فضاء العينة

٢) فضاء العينة

٣) إذا كانA هو حدث ظهور شعار واحد على الأقل فإن :

$$A = \cdots$$

٤) إذا كان B هو حدث ظهور شعارين فقط فإن:

$$B = \cdots.$$

٤ - يحتوي صندوق ثمان بطاقات تحمل الأرقام ١,٢,٣,٤,٥,٦,٧,٨ نسحب عشوائياً بطاقتين معاً من الصندوق ، احسب احتمالات الأحداث التالية :

١) A حدث الحصول على بطاقتين تحملان رقمين متساويين .

٢) B حدث أن يكون مجموع رقمي البطاقتين زوجياً.

٣) C حدث الحصول على بطاقة تحمل الرقم ٧

٤) D حدث أن يكون مجموع رقمي البطاقتين أكبر تماماً من ٨

٥) E حدث أن يكون مجموع رقمي البطاقتين أكبر تماماً من ٨ أو رقما البطاقتين أوليين.

٦) F حدث أن يكون رقما البطاقتين أوليين فيما بينهما

٥ - عدد طلاب السنة الرابعة معلم صف ٣٠٠ طالب و طالبة منهم ٥٠ طالب و٢٥٠ طالبة

يبدأ اسم ٦٠ طالب و طالبة بحرف ال ب منهم ١٠ طلاب ذكور و ٥٠ طالبة

إذا كان :

كتبنا أسماء جميع الطلاب على أوراق متماثلة و وضعناها في صندوق فإذا كانت الأحداث

A حدث سحب اسم طالبة من الصندوق

B حدث سحب اسم طالب من الصندوق

C حدث سحب اسم طالبة يبدأ بحرف ال ب

D حدث سحب اسم طالب يبدأ بحرف ال ب

E حدث سحب اسم طالب أو طالبة بحرف ال ب

إذا سحبنا ورقة واحدة من بين الأوراق عشوائياً احسب :

$$n(S), n(A), n(B), n(C), n(D), n(E)$$

ثم احسب

$$P(A), P(B), P(C), P(D), P(E)$$
$$P(\grave{A}), P(\grave{B}), P(\grave{C}), P(\grave{D}), P(\grave{E})$$

إن طرائق عرض البيانات الإحصائية عديدة و متنوعة و يمكن للباحث اختيار ما يعتقده الأفضل منها في دراسة ظاهرة ما وشرحها ، و من بين طرائق العرض الأكثر تداولاً و حاجة في تحليل البيانات هو التوزيع التكراري و التوزيعات التكرارية التراكمية .

و للجواب على سؤال كيف نبوب البيانات في جدول تكراري مناسب حتى نستطيع قراءة البيانات و معالجتها بشكل مختصر و واضح نورد المثال التالي:

مثال :

كوّن جدول توزيع تكراري لدرجات ٣٠ طالباً في أحد الامتحانات بمادة درجتها العظمى ١٠٠ درجة .

٥٠	٥٤	٥٨	٤٨	٤٩	٤٦
٧٥	٥٤	٤٨	٣٧	٦٢	٤٠
٥٨	٣٤	٤٥	٥٩	٤٨	٥٤
٣٩	٦٨	٤٤	٤٩	٦١	٤٧
٤٥	٤١	٥٧	٤٣	٥٦	٦٣

الحل: إن عرض البيانات بالطريقة المذكورة لا يعطي صورة واضحة عن مستوى التحصيل الدراسي لمجموعة الطلاب و لا يقودنا إلى أحكام دقيقة حوله . لذا سنقوم بتوزيع هذه الدرجات

الحل :

إن عرض البيانات بالطريقة المذكورة لا يعطي صورة واضحة عن مستوى التحصيل الدراسي لمجموعة الطلاب و لا يقودنا إلى أحكام دقيقة حوله . لذا سنقوم بتوزيع هذه الدرجات إلى فئات من أجل ذلك نتبع الخطوات التالية :

١ - نجد المدى المطلق (أو المدى) للبيانات المذكورة ، وهو الفرق بين أكبر مشاهدة في التوزيع و أصغر مشاهدة

$$R = X\,max - X\min = 75 - 34 = 41$$

٢ - نحدد عدد الفئات و يكون عادة اختياري يحدده الباحث حسب حجم العينة و المدى ويستعين بقانون ستيرجس في غالب الأحيان و هو

$$\text{عدد الفئات} = 1 + 3.321\log n$$

حيث n حجم العينة

$$\text{عدد الفئات} = 1 + 3.321\log 30 = 6$$

٣ - نحدد طول الفئة بتقسيم المدى المطلق على عدد الفئات

$$\text{طول الفئة } r = \frac{\text{المدى المطلق}}{\text{عدد الفئات}} = \frac{R}{\text{عدد الفئات}} = \frac{41}{6} \approx 7$$

و نعتبر طول الفئة دوماً عدداً صحيحاً لتسهيل العمل

٤ - نحدد الفئات :

نعين قبل كل شئ الحد الأدنى للفئة الأولى ، وهو أصغر قيمة في التوزيع ، و بعدها نجد الحد الأعلى للفئة الأولى و ذلك بإضافة طول الفئة إلى الحد الأدنى و نعبر عن الفئة بمجال نصف مفتوح [٣٤[أي الحد الأدنى للفئة يدخل مع الفئة نفسها و الحد الأعلى يدخل مع الفئة اللاحقة ، و هكذا نحدد الفئات الأخرى بحيث يكون الحد الأدنى للفئة هو الحد الاعلى للفئة التي تسبقها ، و هكذا تكون الفئة الأخيرة تضم أكبر قيمة بالتوزيع في هذه البيانات .

٥ - نبدأ بعملية التفريغ وفق الجدول التالي:

الفئات	التفريغ	التكرارات
[٣٤,٤١[	||||	٤
[٤١,٤٨[	||||| ||	٧
[٤٨,٥٥[	||||| |||||	١٠
[٥٥,٦٢[	|||||	٥
[٦٢,٦٩[	|||	٣
[٦٩,٧٦[	|	١
المجموع		٣٠

التكرار التراكمي: (التكرار التجميعي الصاعد و التكرار الهابط)

التكرار الهابط	التكرار التجميعي الصاعد	التكرارات	الفئات
٣٠	٤	٤	[٣٤,٤١[
٣٠- ٤=٢٦	٤+٧=١١	٧	[٤١,٤٨[
٢٦- ٧=١٩	١١+١٠=٢١	١٠	[٤٨,٥٥[
١٩- ١٠=٩	٢١+٥=٢٦	٥	[٥٥,٦٢[
٩- ٥=٤	٢٦+٣=٢٩	٣	[٦٢,٦٩[
٤- ٣=١	٢٩+١=٣٠	١	[٦٩,٧٦[
		٣٠	المجموع

المدرج التكراري و المدرج التكراري التراكمي

المدرج التكراري :

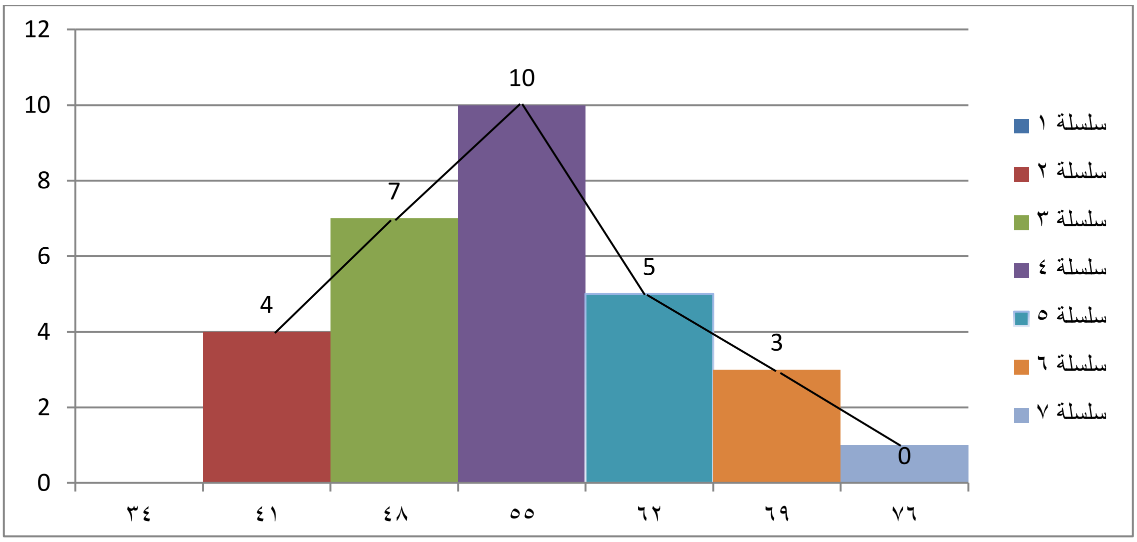
يحتاج الإحصائي لتمثيل البيانات الإحصائية على شكل رسوم بيانية فيلجأ لرسم المدرج التكراري الذي يتكون من مجموعة من المستطيلات المتلاصقة بحيث أن قاعدة كل مستطيل تمثل الحدود الفعلية للفئة و ارتفاعه يتناسب مع تكرار كل فئة ، و بالعودة للمثال السابق نجد :

تكرار الهابط	Fi	fi	الفئات
٣٠	٤	٤	[٣٤,٤١[
٢١	١١	٧	[٤١,٤٨[
١٩	٢١	١٠	[٤٨,٥٥[
٩	٢٦	٥	[٥٥,٦٢[
٤	٢٩	٣	[٦٢,٦٩[
١	٣٠	١	[٦٩,٧٦[

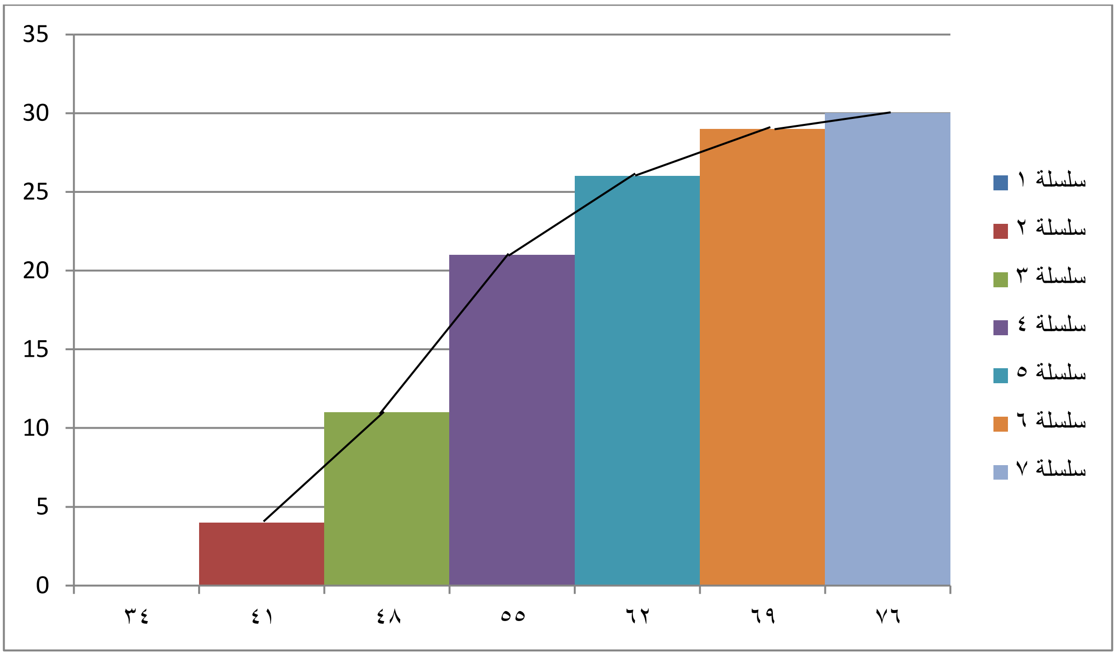
المدرج التكراري

المضلع التكراري : يتكون من الخط المنكسر الذي يصل بين منتصفات رؤوس المستطيلات المشكلة للمدرج التكراري

المدرج التكراري التراكمي و المنحنى التكراري التراكمي :



المدرج التكراري التراكمي و المنحنى التكراري التراكمي لعلامات ٣٠ طالب



الفصل الثالث

مقاييس النزعة المركزية

أولاً : المتوسط الحسابي The Arithetic Mean

كما درسنا في السنة الأولى بمادة الإحصاء فإن الوسط الحسابي لبيانات غير مبوبة يحسب من القانون :

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

حيث X_i مشاهدات العينة (القيم) وn حجم العينة و في مثالنا السابق و قبل تبويب البيانات نجد

$$\bar{X} = \frac{x_1 + x_2 + x_3 \ldots\ldots + x_{30}}{30}$$

أما في حالة حساب الوسط الحسابي من خلال جداول التوزيع التكراري فيصبح القانون

$$\bar{X} = \frac{\sum_{i=1}^{n} fi\ xi}{\sum fi}$$

حيث X_i وسطي كل فئة i

fi تكرار الفئة i

$\sum fi = n$ حجم العينة

وفي مثالنا السابق نجد

الفئات	Fi	Xi	fi Xi
[٣٤,٤١[	٤	٣٧.٥	١٥٠
[٤١,٤٨[	٧	٤٤.٥	٣١١.٥
[٤٨,٥٥[	١٠	٥١.٥	٥١٥
[٥٥,٦٢[	٥	٥٨.٥	٢٩٢.٥
[٦٢,٦٩[	٣	٦٥.٥	١٩٦.٥
[٦٩,٧٦[	١	٧٢٥	٧٢.٥
المجموع			١٥٣٨

حيث Xi = الحد الأدنى للفئة i +الحد الأعلى للفئة نفسها / ٢

$$\bar{X} = \frac{\sum_{i=1}^{n} xi\ fi}{\sum_{i=1}^{n} fi} = \frac{1538}{30} = 51.2667$$

ثانياً : الوسيط The Median:

لإيجاد الوسيط في حالة المشاهدات (المفردات)

المفردة أي حال البيانات الغير مبوبة ، نرتب هذه البيانات تصاعدياً و يكون :

الوسيط = المفردة التي رقمها $\frac{n+1}{2}$ إذا كان n فردياً

فالوسيط هو مقياس ترتيبي يقع في منتصف القيم عند ترتيبها تصاعدياً أو تنازلياً

مثال : إذا كان لدينا ١١ مفردة ولتكن علامات ١١طالب في مادة الرياضيات :

٨٠ ، ٧٢ ، ٧٠ ، ٦٠ ، ٥٩ ، ٥٧ ، ٥٣ ، ٥٠ ، ٤٦ ، ٤٥ ، ٤٠

١١ ١٠ ٩ ٨ ٧ ٦ ٥ ٤ ٣ ٢ ١

نوجد أولاً ترتيب الوسيط$= 6 = \frac{11+1}{2} = \frac{n+1}{2}$

فالمفردة رقم ٦ هي الوسيط = ٥٧ علامة

أما إذا كانت n عدد زوجي فترتيب الوسيط $= \frac{n+1}{2} = T$

مثال : إذا كانت ١٠= n بدلاً من ١١

٧٢ ، ٧٠ ، ٦٠ ، ٥٩ ، ٥٧ ، ٥٣ ، ٥٠ ، ٤٦ ، ٤٥ ، ٤٠

١٠ ٩ ٨ ٧ ٦ ٥ ٤ ٣ ٢ ١

$$T = \frac{n+1}{2} = \frac{10+1}{2} = \frac{11}{2} = 5.5$$

فالوسيط يقع بين المفردة رقم ٥ والمفردة رقم ٦ فالوسط الحسابي لهذه المفردتين هو

الوسيط $Me = \frac{53+57}{2} = 55$

حيث Me الوسيط

الوسيط لبيانات مبوبة في فئات نتبع الخطوات التالية لحسابه :

١ - نحسب ترتيب الوسيط من خلال العلاقة $T = \frac{\sum fi}{2}$

٢ - نحسب التكرارات التجميعية الصاعدة F_i (التكرار التراكمي)

وذلك بهدف تحديد الفئة التي يقع فيها الوسيط

٣ - نبحث عن الفئة التي تقع فيها قيمة ترتيب الوسيط بين التكرارت التجميعية الصاعدة بعد حساب T ترتيب الوسيط و أول قيمة في عمود التكرارات التجميعية الصاعدة تساوي أو أكبر من قيمة T تكون الفئة التي تمثلها هي فئة الوسيط

٤ - نطبق العلاقة التالية لحساب الوسيط :

$$Me = Lm + \frac{\frac{\sum fi}{2} - F_{m-1}}{fm} . r$$

حيث Lm: الحد الأدنى للفئة الوسيطة .

F_{m-1} : التكرار التجميعي الصاعد للفئة السابقة للفئة الوسيطية

f_m : تكرار الفئة الوسيطية

r : طول الفئة الوسيطية (مدى الفئة)

Me: الوسيط

بالعودة لمثالنا السابق

رقم الفئة	الفئات	fi	Fi
١	٣٤,٤١	٤	٤
٢	٤١,٤٨	٧	١١
٣	٤٨,٥٥	١٠	٢١
٤	٥٥,٦٢	٥	٢٦
٥	٦٢,٦٩	٣	٢٩
٦	٦٩,٧٦	١	١

٢١ هو $F3$ ، ١٠ هو $f3$

ترتيب الوسيط $T = \frac{\sum fi}{2} = \frac{30}{2} = 15$

الفئة الوسيطية هي الفئة رقم ٣ أي الوسيط يقع بين]48,55[لأن $F_3 = 21$ و

$T = 15$

فيصبح :

$$Me = 48 + \frac{15 - 11}{10} \times 7 = 48 + 2.8 = 50.8$$

ثالثاً: المنوال The Mode:

المنوال لمجموعة من البيانات: هو القيمة الأكثر تكراراً و يحسب في البيانات (القيم) المبوبة بشكل مباشر

مثال :

إذا كانت لدينا مجموعة المفردات التالية :

٥,٢,٤,٦,٤,٧,٨,٤

نلاحظ أن القيمة الأكثر تكراراً هي ٤

يمكن أن تتساوى تكرارات إحدى القيم مع أخرى فنقول أن هناك منوالين

مثال : لدينا االبيانات التالية :

٣,٩,٧,٣,٥,٧,٨,١٠,٢

نلاحظ أن لهذه البيانات منوالين هما ٣ و ٧ لأنهما أكثر القيم تكراراً و نسميها ثنائية المنوال ، و ممكن أن يكون أكثر من منوال وأن لا يكون منوال إطلاقاً

- إيجاد المنوال في حالة البيانات المبوبة (التكرارية)

لإيجاد المنوال في حال البيانات المبوبة نتبع مايلي :

١ - نحدد الفئة المنوالية و هي الفئة الأكثر تكراراً

٢ - نقول بحساب المنوال من العلاقة :

$$Mo = L_{mo} + \frac{\Delta_1}{\Delta_1 + \Delta_2} . r$$

Mo : المنوال

L_{mo} : الحد الأدنى للفئة المنوالية

r : مدى الفئة

$$\Delta_1 = f_{mo} - f_{mo-1}$$

$$\Delta_1 = f_{mo} - f_{mo+1}$$

بالعودة إلى مثالنا السابق نجد :

الرقم	الفئات	f_i
١	[٣٤,٤١[	٤
٢	[٤١,٤٨[	٧
٣	[٤٨,٥٥[	١٠
٤	[٥٥,٦٢[	٥
٥	[٦٢,٦٩[	٣
٦	[٦٩,٧٦[	١

f_{mo-1} هي 7

١٠ هي f_{mo} الفئة المنوالية

f_{mo+1} هي ٥

$$\Delta_1 = 10 - 7 = 3$$

$$\Delta_2 = 10 - 5 = 5$$

$$M_o = 48 + \frac{3}{3+5} \times 7$$

$$48 + 2.625 = 50.625$$

الربيعات Quartiles :

و هي القياسات التي تقسم البيانات إلى أربعة أقسام متساوية فإذا كانت $x_1, x_2 ...... x_n$ مجموعة من البيانات المرتبة تصاعدياً يعرف الربيع الأدنى Q_1 بأنه القيمة التي يسبقها ٢٥٪ أو ربع البيانات ، كما يعرف الربيع الأعلى Q_3 بأنه القيمة التي يسبقها ٧٥٪ أي ثلاثة أرباع البيانات ، كما يعرف الربيع الأوسط Q_2 بأنه القيمة التي تتوسط قيم التوزيع أي يسبقها ٥٠٪ من البيانات و هو بهذا التعريف ينطبق على الوسيط الذي حسبناه سابقاً

- إيجاد الربيعات في حال البيانات المفردة (الغير مبوبة)

١ - نوجد ترتيب الربيع الأدنى من العلاقة :

$$T_{Q_1} = \frac{n+1}{4}$$

بينما ترتيب الربيع الأعلى :

$$T_{Q_3} = \frac{3(n+1)}{4}$$

و الربيع الأوسط و هو الوسيط :

$$T_{Q_2} = \frac{n+1}{2}$$

مثال :

تمثل البيانات التالية أوزان أعضاء فريق كرة اليد

٩ ٨ ٧ ٦ ٥ ٤ ٣ ٢ ١

٧٤ ، ٧٢ ، ٧٠ ، ٦٩ ، ٦٧ ، ٦٥ ، ٥٩ ، ٥٨ ، ٥٢

أوجد التربيعات الثلاثة

١ - نرتب البيانات تصاعدياً من الأصغر إلى الأكبر

٢ - نحسب الربيع الأدنى

$$T_{Q_1} = \frac{n+1}{4} = \frac{9+1}{4} = 2.5$$

الربيع الأدنى يقع بين القيمة رقم ٢ و القيمة رقم ٣

و الوسط الحسابي لهاتين القيمتين هو الربيع الأدنى

$$Q_1 = \frac{58+59}{2} = 58.5$$

- الربيع الأوسط و هو الوسيط ترتيبه:

$$T_{Q_2} = \frac{n+1}{2} = \frac{10}{2} = 5$$

القيمة رقم ٥ هي الربيع الأوسط $Q_2 = 67$

- الربيع الأعلى ترتيبه :

$$T_{Q_3} = \frac{3(n+1)}{4} = \frac{30}{4} = 7.5$$

أي أنه يساوي الوسط الحسابي للقيمتين رقم ٧ و رقم ٨

$$Q_3 = \frac{70+72}{2} = 71$$

- إيجاد الربيعات في حالة البيانات المبوبة:

١ - نشكل جدول التوزيع التكراري التراكمي F_i

٢ - نحسب الربيع (إن ترتيب الربيع الأدنى $\frac{\sum f_i}{4}$ بينما ترتيب الربيع الأعلى $\frac{3\sum f_i}{4}$

٣ - نحدد فئة الربيع من جدول التوزيع التكراري الصاعد (التراكمي) تماماً كما فعلنا عند حساب الوسيط .

٤ - نحدد الربيع الأدنى من العلاقة :

$$Q_1 = L_{Q_1} + \frac{\frac{\sum f_i}{4} - F_{Q1-1}}{f_{Q_1}}.r$$

و الربيع الأعلى من العلاقة :

$$Q_3 = L_{Q_3} + \frac{\frac{3\sum f_i}{4} - F_{Q3-1}}{f_{Q_3}}.r$$

حيث Q_1 وQ_3 الربيع الأدنى و الأعلى على التوالي

L_{Q1} و L_{Q3} :الحد الأدنى للفئة الربيعية على التوالي

F_{Q1}-1 و F_{Q3}-1 :التكرار التجميعي الصاعد للفئة التي تسبق فئة الربيع

f_{Q1} و f_{Q3}: تكرارات الفئات الربيعية على التوالي

r :طول الفئة (مدى الفئة)

مثال :

بالعودة إلى مثالنا السابق نجد :

F_i	f_i	الفئات	رقم الفئة
٤	٤	٣٤,٤١	١
١١	٧	٤١,٤٨	٢
٢١	١٠	٤٨,٥٥	٣
٢٦	٥	٥٥,٦٢	٤
٢٩	٣	٦٢,٦٩	٥
٣٠	١	٦٩,٧٦	٦

لحساب الربيع الأول (الأدنى) نوجد ترتيبه من العلاقة :

$$T = \frac{\sum f_i}{4} = \frac{30}{4} = 7.5$$

ننظر إلى عمود التكرارات التجميعية الصاعدة و أول عدد يتضمن ٧.٥ تكون هو العدد المقابل لفئة الربيع الأدنى إذاً الربيع الأدنى يقع في الفئة رقم ٢ لأن تكراراته التجميعية ١١ تتضمن ترتيب الربيع ٧.٥ نطبق العلاقة:

$$Q_1 = L_{Q_1} + \frac{\frac{\sum f_i}{4} - F_{Q1-1}}{f_{Q_1}} . r$$

$$Q_1 = 41 + \frac{7.5 - 4}{7} \times 7 = 41 + 3.5 = 44.5$$

- الربيع الأعلى :

$$Q_3 = L_{Q_3} + \frac{\frac{3\sum f_i}{4} - F_{Q3-1}}{f_{Q_3}} . r$$

ترتيبه :

$$T = \frac{3\sum f_i}{4} = \frac{3 \times 30}{4} = 22.5$$

أي أن فئة الربيع الأعلى هي الفئة الرابعة ٤

$$Q_3 = 55 + \frac{22.5 - 21}{5} \times 7 = 55 + \frac{1.5}{5} \times 7 = 55 + \frac{10.5}{5}$$
$$= 55 + 2.1 = 57.1$$

و بنفس الطريقة نحسب الربيع الأوسط و هو الوسيط و قد قمنا بحسابه سابقاً.

الفصل الرابع

بعض مقاييس التشتت:

تستخدم مقاييس التشتت لإعطاء صورة عن مدى تقارب المشاهدات (البيانات) أو تباعدها (تشتتها)بعضها عن بعض فكلما زادت قيمة مقياس التشتت ازداد تشتت المشاهدات و لكلّما قلت قيمته زاد التجانس بين المشاهدات ، و تجدر الإشارة إلى أن جميع مقاييس التشتت موجبة و مقاييس التشتت عديدة سندرس منها المدى ، الانحراف عن المتوسط، التباين، و الانحراف المعياري .

١. المدى *The Range*:

إن المدى R لمجموعة من البيانات $x_1,x_2,.....x_n$ هو الفرق بين أكبر قيمة و أصغر قيمة من هذه البيانات أي :

$$R = X_{max} - \mathrm{X}_{\min}$$

أما المدى في حال البيانات المبوبة فيعرف بأنه الفرق بين مركزي الفئتين الأخيرة (الكبرى)و الأولى (الصغرى)

ملاحظة :إن المدى مقياس غير مُرضٍ للتشتت ، لأن المدى يعطي الفرق بين اكبر قيمة و أصغر قيمة أي يعتمد في قياسه على مشاهدتين فقط من البيانات الإحصائية .

٢. الانحراف المتوسط The Average Deviation :

من الممكن معرفة تشتت القيم باستخدام انحراف كل قيمة من قيم التوزيع عن الوسط الحسابي لقيم التوزيع ، أي $X_i - \bar{X}$ فإذا اتضح لنا بأن هذه الانحرافات كبيرة ، يعني ذلك أن تشتت البيانات هو اكبر من تلك الحالة التي تكون فيها الانحرافات عن الوسط الحسابي تشكل قيمة صغيرة و بالتالي يمكن استخدام انحرافات القيم عن وسطها الحسابي لمعرفة مقدار تبعثر أو تشتت البيانات المدروسة ، و قد يتوقع البعض أنه يمكن اسنخدام الوسط الحسابي لهذه الانحرافات كمقياس للتشتت ، و لكن ذلك سوف يكون مستحيلاً بسبب أن بعض هذه الانحرافات سيكون سالباً و البعض الآخر موجباً.

و إن مجموعها الجبري سيكون معدوماً (كما شاهدنا ذلك في مادة الإحصاء في السنة الأولى) بأنه من خواص الوسط الحسابي أن مجموع انحرافات قيم التوزيع عن الوسط الحسابي تساوي الصفر أي :

$$\sum (X_i - \bar{X}) = 0$$

و للتخلص من هذه الخاصة بحساب الانحراف عن المتوسط نهمل إشارات الانحرافات و نأخذ المجموع بالقيمة المطلقة فيكون قانون حساب الانحراف المتوسط حسب ما سبق :

$$D_m = \frac{\sum_{i=1}^{n} |X_i - \bar{X}|}{n}$$

حيث $|X_i - \bar{X}|$ هي القيمة المطلقة لإنحراف X_i عن الوسط الحسابي

D_m : الانحراف المتوسط

مثال :

أحسب الانحراف المتوسط لعلامات إحدى الطالبات في ٦ مواد

المادة	علامة الطالبة X_i	الانحرافات $\lvert X_i - \overline{X} \rvert$
١	٦٨	+١٥
٢	٤٦	- ٧
٣	٤٧	- ٦
٤	٤٩	- ٤
٥	٤١	- ١٢
٦	٦٧	+١٤
المجموع	٣١٨	٥٨

نحسب أولاً الوسط الحسابي لعلامات الطالبة :

$$\overline{X} = \frac{\sum X_i}{n} = \frac{318}{6} = 53 \text{علامة}$$

ثم نحسب انحرافات القيم عن الوسط الحسابي فنجد من الجدول أن مجموعهما

الجبري يساوي الصفر لذلك نهمل الإشارة و نجمع فيكون المجموع ٥٨ و يكون :

$$D_m = \frac{\sum \lvert X_i - \overline{X} \rvert}{n} = \frac{58}{2} = 9.67 \text{علامة}$$

- حساب الانحراف المتوسط لبيانات مبوبة :

الانحراف المتوسط لبيانات مبوبة في فئات هو عبارة عن مجموع انحرافات قيم مراكز الفئات عن الوسط الحسابي بالقيمة المطلقة مضروبة بتكراراتها ، و تقسيم الناتج على مجموع التكرارات التي هي تساوي حجم العينة أي :

$$D_{\mathrm{m}} = \frac{\sum|X_i - \bar{X}|f_i}{\sum \mathrm{f_i}}$$

مثال :

الجدول التالي يمثل توزيع النفقات الأسبوعية لعينة من ٢٠٠ أسرة (بآلاف الليرات)

فئات النفقات (آلاف الليرات)	عدد الأسر	وسطي الفئة X_i	$X_i f_i$	$\lvert X_i - \bar{X}\rvert$	$\lvert X_i - \bar{X}\rvert f_i$
١.٥- ٢.٥	١٨	٢	٣٦	- ٢	٣٦
٢.٥- ٣.٥	٤٢	٣	١٢٦	- ١	٤٢
٣.٥- ٤.٥	٨٠	٤	٣٢٠	٠	٠
٤.٥- ٥.٥	٤٦	٥	٢٣٠	+١	٤٦
٥.٥- ٦.٥	١٠	٦	٦٠	+٢	٢٠
٦.٥- ٧.٥	٤	٧	٢٨	+٢	١٢
المجموع	٢٠٠		٨٠٠		١٥٦

$$\bar{X} = \frac{\sum \mathrm{X_i f_i}}{\sum \mathrm{f_i}} = \frac{800}{200} = 4$$

منه :

$$D_m = \frac{\sum|X_i - \bar{X}|f_i}{\sum f_i} = \frac{156}{200} = 0.78 \text{ ألف ليرة}$$

أي الانحراف المتوسط = ليرة $0.78 \times 1000 = 780$

- الانحراف المتوسط النسبي :

لكي نزيل وحدات القياس و يصبح مقياس التشتت و قابل للمقارنة بين مثيلاته لبيانات مختلفة عنه بالواحدات يفضل حساب الانحراف المتوسط النسبي و هو :

$$\text{الانحراف المتوسط النسبي} = \frac{\text{الانحراف المتوسط}}{\text{الوسط الحسابي}} \times 100$$

أي

$$D_m\% = \frac{D_m}{\bar{X}} \times 100$$

وفي مثالنا :

$$D_m\% = \frac{0.78}{4} \times 100 = 19.5\%$$

٣. الانحراف المعياري Standard Deviation :

في سبيل التغلب على مشكلة الإشارات الجبرية عند جمع الانحرافات ، فبدلاً من أخذ القيم المطلقة للانحرافات و إهمال الإشارات الجبرية ، فإننا نستطيع أن نتغلب

على هذه الخاصة التي يتمتع بها الوسط الحسابي بتربيع قيم الانحرافات ، وبالتالي تصبح جميعها موجبة و بالتالي نحصل على مجموع مربعات الانحرافات بدلاً من مجموع الانحرافات فمجموع مربعات الانحرافات $\sum(X_i - \bar{X})^2$ يتم تقسيمها على حجم العينة n في البيانات الغير مبوبة بذلك نحصل على ما يدعى بالتباين و نرمز له في العينة S^2

فإذا كان لدينا عينة حجمها nمشاهدة $x_i=x_1,x_2....x_n$

و وسطها الحسابي $\bar{X}$ يكون التباين :

$$S^2 = \frac{\sum(X_i - \bar{X})^2}{n}$$

و حتى نرد القيم المربعة إلى أصلها نقوم بجذر قيمة التباين و هذا يعطينا الانحراف المعياري أي :

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

حساب الانحراف المعياري لبيانات غير مبوبة :

هناك عدة طرق لحساب الانحراف المعياري نكتفي منها بالطريقة المباشرة و تتلخص العمليات الحسابية الواجب إجرائها بما يلي :

١ - نقوم بحساب الوسط الحسابي للبيانات .

٢ - نوجد انحرافات القيم x_i عن وسطها الحسابي

٣ - نربع كل انحراف من الانحرافات

٤ - نحسب مجموع مربع الانحرافات .

٥ - نقسم مجموع مربع الانحرافات على حجم العينة n فنحصل على التباين S^2

٦ - نأخذ الجذر التربيعي الموجب للتباين فنحصل على الانحراف المعياري S

مثال :

لدينا المعلومات الآتية و التي تمثل الساعات الأسبوعية لدراسة أحد الطلاب خلال سبعة أيام :

الأسبوع	عدد الساعات X	$X_i - \bar{X}$ $X_i - 17$	$(X_i - \bar{X})^2$
١	١٥	- ٢	٤
٢	٩	- ٨	٦٤
٣	١٤	- ٣	٩
٤	٦	- ١١	١٢١
٥	٢٣	٦	٣٦
٦	١٩	٢	٤
٧	٣٣	١٦	٢٥٦
المجموع	١١٩	٠	٤٩٤

متوسط الدراسة الأسبوعية للطالب:

$$\bar{X} = \frac{\sum X}{n} = \frac{119}{7} = 17$$

الخطوات المتبقية موضحة على الجدول

الانحراف المعياري:

$$S = \sqrt{\frac{\sum (X_i - \bar{X})^2}{n}} = \sqrt{\frac{494}{7}} = 8.4$$

حساب الانحراف المعياري للبيانات المبوبة في فئات :

إذا كانت البيانات مبوبة و x_i تمثل مراكز الفئات في الجدول التكراري و إن تكراراتها المقابلة f_i فالانحراف المعياري بالطريقة المباشرة هو :

$$S = \sqrt{\frac{\sum (X_i - \bar{X})^2 . f_i}{\sum f_i}}$$

مثال :

الجدول الآتي تمثل التوزيع التكراري لمدة الخدمة بالسنوات لعينة مؤلفة من ١٠٠ عامل من عمال إحدى الشركات .

مدة الخدمة بالسنوات	عدد العمال f_i	نحسب x_i	نحسب $x_i f_i$	نحسب $(X_i - \bar{X})$	نحسب $(X_i - \bar{X})^2$	نحسب $(X_i - \bar{X})^2 . f_i$
٠.٥,٥.٥	٦	٣	١٨	- ١٢	١٤٤	٨٦٤
٥.٥,١٠.٥	٢٥	٨	٢٠٠	- ٧	٤٩	١٢٢٥
١٠.٥,١٥.٥	٢٩	١٣	٣٧٧	- ٢	٤	١١٦
١٥.٥,٢٠.٥	٢٠	١٨	٣٦٠	+٣	٩	١٨٠
٢٠.٥,٢٥.٥	١٠	٢٣	٢٣٠	+٨	٦٤	٦٤٠
٢٥.٥,٣٠.٥	٥	٢٨	١٤٠	+١٣	١٦٩	٨٤٥
٣٠.٥,٣٥.٥	٣	٣٣	٩٩	+١٨	٣٢٤	٩٧٢
٣٥.٥,٤٠.٥	٢	٣٨	٧٦	+٢٣	٥٢٩	١٠٥٨
المجموع	١٠٠		١٥٠٠			٥٩٠٠

الوسط الحسابي : $\bar{X} = \frac{\sum X_i f_i}{\sum f_i} = \frac{1500}{100} = 15$

بعد إجراء الخطوات الموضحة على الجدول :

$$S = \sqrt{\frac{\sum (X_i - \bar{X})^2 . f_i}{\sum f_i}} = \sqrt{\frac{5900}{100}} = 7.68$$

- معامل الإختلاف (الانحراف المعياري النسبي) Coefficient of variation :

يعتبر معامل الاختلاف (C.V) من أهم مقاييس التشتت النسبية ، و يستعمل بشكل واسع في التحليل الإحصائي لمقارنة مقدار التشتت أو الاختلاف لمجموعتين أو أكثر من البيانات في حالة اختلاف الوحدات القياسية المستخدمة مع وحدات كل مجموعة، و كذلك عند اختلاف قيم الوسط الحسابي ، و لو كانت الوحدات القياسية غير مختلفة ، فمعامل الاختلاف يعبر عن الانحراف المعياري كنسبة مئوية من الوسط الحسابي و صيغته للعينة :

$$C.V = \frac{S}{\overline{X}}.100$$

و حسابه في المثال السابق

$$C.V = \frac{7.68}{15}.100 = 51.2\%$$

- الدرجة المعيارية Standardized Scores :

في بعض الأحيان نضطر لمقارنات من نوع آخر ، فمثلاً يمكن أن نقارن مشاهدتين من مجموعتين مختلفتين ، في هذه الحالة يتم تحويل وحدات كل مفردة إلى وحدات معيارية حتى نتمكن من المقارنة ، و ذلك باستخدام كل من الوسط الحسابي و الانحراف المعياري و بذلك تصبح الدرجات المعيارية خالية من وحدات القياس و عملية التحويل إلى درجات معيارية تتم وفق الصيغة التالية :

$$Z = \frac{X_i - \bar{X}}{S}$$

حيث Z هي الدرجة المعيارية

- مثال يوضح كيفية استخدام الدرجات المعيارية و أهميتها في المقارنة في امتحان مادة القياس حصل أحد الطلاب على ٧٥ علامة علماً بأن متوسط علامات جميع الطلاب في مبادئ الاحصاء يساوي ٦٧ علامة بانحراف معياري يساوي ١٠ علامات و أن هذا الطالب حصل في مادة العلوم الاجتماعية على ٨٥ علامة علماً أن متوسط علامات جميع الطلاب في مقرر العلوم الاجتماعية يساوي ٧٧علامة بانحراف معياري ١٦ علامة

السؤال :

في أي المقررين كانت علامة الطالب أعلى بالمقارنة مع زملائه

الحل :

عند مقارنة الامتحانين لهذا الطالب نجد أن علامته في مادة العلوم الاجتماعية أعلى من علامته في مقرر القياس لكن عند تحويل علامات الطالب إلى درجات معيارية نجد أن :

$$\text{للقياس}\ Z_1 = \frac{X_1 - \overline{X_1}}{S_1} = \frac{75 - 67}{10} = 0.8\ \text{درجة معيارية}$$

$$\text{العلوم الاجتماعية}\ Z_2 = \frac{\mathrm{X}_2 - \overline{\mathrm{X}_2}}{\mathrm{S}_2} = \frac{85 - 77}{16} = 0.5\ \text{درجة معيارية}$$

يتضح من مقارنات الدرجات المعيارية لهذا الطالب بأن تفوقه في مقرر القياس بين زملائه هو أعلى من تفوقه في مقرر العلوم الاجتماعية بالمقارنة مع جميع الطلاب الممتحنين في تلك المقررات .

تمارين

١ - تمثل البيانات التالية عدد المرضى في أحد الأقسام خلال ٢٥ يوم

٩ ، ٨ ، ٨ ، ٦ ، ٦ ، ٦ ، ٥ ، ٥ ، ٥ ، ٤ ، ٤ ، ٢ ، ٢ ، ١

٢٥ ، ٢٤ ، ٢٢ ، ٢١ ، ٢٠ ، ١٩ ، ١٨ ، ١٧ ، ١٥ ، ١٣ ، ١٠

و المطلوب :

١. إيجاد الوسيط و المنوال

٢. تفريغ هذه البيانات في جدول توزيع تكراري مناسب

٣. إيجاد المتوسط الحسابي – الوسيط – المنوال

٤. ارسم مدرج التوزيع التكراري .

٢ - البيانات الآتية تبين توزع عدد من الشركات التجارية بحسب المبيعات الشهرية بآلاف الليرات .

عدد الشركات	المبيعات الشهرية
٢٥	[٥,١٠[
٣٠	[١٠,١٥[
٤١	[١٥,٢٠[
٤٩	[٢٠,٢٥[
٣٥	[٢٥,٣٠[
٣٠	[٣٠,٣٥[
٢٠	[٣٥,٤٠[
١٠	[٤٠,٤٥[

المطلوب :

١ - أحسب معامل الاختلاف

٢ - أحسب الانحراف المتوسط

٣ - الجدول الآتي يمثل توزع مجموعة من الطلبة الناجحين في مقرر القياس تبعاً لعلاماتهم في ذلك المقرر:

العلامات	عدد الطلاب
أقل من ٥٠ علامة	٠
أقل من ٥١ علامة	٥
أقل من ٥٢ علامة	٢٦
أقل من ٥٣ علامة	٧٥
أقل من ٥٤ علامة	١٥٢
أقل من ٥٥ علامة	٢٢٤
أقل من ٥٦ علامة	٢٨٩
أقل من ٥٧ علامة	٣٣٣
أقل من ٥٨ علامة	٣٦١
أقل من ٥٩ علامة	٣٨٣
أقل من ٦٠ علامة	٤٠٠

المطلوب :

١ - إجراء جدول توزيع تكراري على أساس طول الفئة يساوي علامتين

٢ - أحسب معامل الاختلاف

٣ - أحسب العلامة الأكثر شيوعاً

قياس و تحليل الارتباط و الانحدار:

رأينا في الفصل السابق كيف يمكن حساب مقاييس النزعة المركزية و التشتت و احتمال أن يأخذ المتغير قيمة ما أو تكون قيمته في مجال ما .

سوف نتعرف في هذا الفصل على أساليب قياس و تحليل جديدة نستطيع من خلالها تحليل متغير ما من خلال علاقته بمتغير آخر أو بعدة متغيرات ، حيث نستطيع أن نذكر عدداً كبيراً من الأمثلة على الظواهر أو المتغيرات التي يوجد علاقة بينها مثل دخل الأسرة و إنفاقها ، دخل الأسرة و مدخراتها ، علامة طالب مقرر الرياضيات و علامته في مقرر الإحصاء .

إن قياس و تحليل الارتباط و الانحدار يتم من خلال إجراء عدد من الحسابات ، يسبقها التأكد من وجود علاقة بين المتغيرات عن طريق التحليل المنطقي و بالاعتماد على معارفنا العامة

بعد التأكد من وجود علاقة بين الظواهر (المتغيرات) يجب تحديد أي المتغيرات يؤثر بالآخر و هذا ندعوه بالمتغير المستقل و نرمز له ب X

أما المتغير الذي يتأثر بالآخر ندعوه بالمتغير التابع و نرمز له بالرمز Y . و يتم تحديد طبيعة المتغيرات في أكثر الأحيان من خلال التحليل المنطقي ، فالدخل يؤثر باستهلاك الأسرة فهو المستقل ، و الاستهلاك تابع ، أما إذا تعذر تحديد طبيعة المتغيرات بالطريقة

السابقة فعندها من المنطق القول إن من يحدث أولاً هو الذي يؤثر بالآخر ، و بالتالي هو المستقل

١ - تحليل الارتباط :

يمكن تقسيم العلاقات بين الظواهر إلى قسمين : علاقة تابعية تامة عندما يكون لكل قيمة للمتغير المستقل قيمة بالذات للمتغير المستقل ، مثل علاقة مساحة الدائرة بنصف القطر ، و مساحة المربع بطول الضلع ، و استهلاك معمل ما من أحد المواد الأولية بحجم الإنتاج ، وعلاقات ارتباطية عندما يكون مقابل كل قيمة للمتغير المستقل قيمة تقريبية أو احتمالية للمتغير التابع و هذا ما سوف ندرسه و هذا ما ندعوه الارتباط البسيط

- فالارتباط البسيط : هو الذي يبين العلاقة بين متغير تابع واحد و متغير مستقل واحد مع افتراض عدم وجود أي متغيرات مستقلة أخرى مؤثرة. و ندعوه بالارتباط الخطي لأنه يمثل بمعادلة مستقيم خطية من الدرجة الأولى .

- وللتمييز بين المتغير المستقل و المتغير التابع نلخص ما أسلفناه

من حيث الزمن : المتغير الأسبق زمنياً هو المستقل و اللاحق هو التابع

من ناحية التأثير : المتغير المؤثر هو السابق و بالتالي هو المستقل

من ناحية العلاقة المنطقية : كلما ازداد الطول يزداد الوزن فالطول هو المستقل و الوزن هو التابع

و سوف ندرس الارتباط البسيط لنوعين من البيانات

١ - الارتباط البسيط للمتغيرات الكمية

٢ - الارتباط البسيط للمتغيرات النوعية .

المتغيرات المقاسة بالكم مثل (الطول ، الوزن ، العلامة)

- معامل ارتباط بيرسون :

يساعد معامل ارتباط بيرسون على تحديد متانة و طبيعة العلاقة بين المتغيرين الكميين Xو Y و تعطى قيمته بالعلاقة الآتية :

$$r_{xy} = \frac{\sum(X-\bar{X})(Y-\bar{Y})}{nS_X S_Y}$$

حيث $\bar{X}$و$\bar{Y}$ الاوساط الحسابية للمتغير المستقل و المتغير التابع على التوالي

S_X و S_y الانحرافات المعيارية للمتغير المستقل و المتغير التابع

n : عدد المشاهدات (حجم العينة)

r_{xy} : معامل ارتباط بيرسون

يمكن إصلاح العلاقة السابقة و كتابتها بأشكال أخرى أكثر سهولة لحساب معامل الارتباط و أسهلها العلاقة :

$$r_{XY} = \frac{\sum XY - n\bar{X}\bar{Y}}{\sqrt{[\sum X^2 - n\bar{X}^2][\sum Y^2 - n\bar{Y}^2]}}$$

مثال :

الجدول الآتي يمثل بيانات الدخل الشهري ، و الإنفاق الشهري لعينة مؤلفة من ١٠ أسر (بآلاف الليرات)

أحسب معامل ارتباط بيرسون .

رقم الأسرة	X الدخل	Y الإنفاق	نحسب XY	نحسب X^2	نحسب Y^2
١	٤	٣	١٢	١٦	٩
٢	٥	٤	٢٠	٢٥	١٦
٣	٣	٢	٦	٩	٤
٤	٦	٥	٣٠	٣٦	٢٥
٥	٧	٧	٤٩	٤٩	٤٩
٦	٨	٧	٥٦	٦٤	٤٩
٧	٩	٦	٥٤	٨١	٣٦
٨	٤	٤	١٦	١٦	١٦
٩	٥	٣	١٥	٢٥	٩
١٠	١٢	٩	١١٨	١٤٤	٨١
المجموع	٦٣	٥٠	٣٦٦	٤٦٥	٢٩٤

نحسب $\bar{Y}$و $\bar{X}$

$$\bar{X} = \frac{\sum \mathrm{X}}{\mathrm{n}} = \frac{63}{10} = 6.3$$

$$\bar{Y} = \frac{\sum \mathrm{Y}}{\mathrm{n}} = \frac{50}{10} = 5$$

الحل : بعد حساب المؤشرات المصاغة في القانون على الجدول نكتب :

$$\mathrm{r_{xy}} = \frac{\sum XY - n\bar{X}\bar{Y}}{\sqrt{[\sum \mathrm{X}^2 - \mathrm{n}\bar{\mathrm{X}}^2][\sum \mathrm{Y}^2 - \mathrm{n}\bar{\mathrm{Y}}^2]}}$$

$$r_{\mathrm{xy}} = \frac{366 - 10(6.3)(5)}{\sqrt{[465 - 10(6.3)^2][294 - 10(5)^2]}} = +0.9317$$

إذاً معامل ارتباط بيرسون $r_{\mathrm{xy}} = +0.9317$

ماذا تعني هذه القيمة

التفسير : الارتباط بين المتغيرين الدخل و الإنفاق طردي (لأن إشارته موجبة) أي كلما ازداد الدخل يزداد الإنفاق و قوي جداً لأن قيمته تقترب من الواحد

و نستطيع أن نقول دائماً بأنه إذا كان :

r =0 : ارتباط معدوم (لا يوجد ارتباط بين المتغيرين)

$r < 0.5$: ارتباط ضعيف جداً و طردي

$r \geq 0.5$: ارتباط شديد و طردي

r = +1 : ارتباط تام و طردي

أما إذا كانت النتيجة سالبة فيكون الارتباط عكسي و تغيّر شدته حسب اقتراب قيمته من الواحد الصحيح أو الصفر تماماً كما سبق أي أن :

$$-1 \leq r \leq 1$$

- معامل ارتباط سبيرمان (الرتب)

إن معامل ارتباط بيرسون لا يقيس الارتباط بين المتغيرات إذا كانت هذه المتغيرات نوعية (غير كمية)

المتغيرات النوعية (الوصفية) : هي البيانات التي تدل على الصفة أو النوعية الموجودة فيها الظاهرة أي أنها معطيات وصفية ملازمة للموصوف و لا يمكن قياسها كمياً مثل لون البشرة أو لون العيون أو لون الشعر أو لون الزهور ..... الخ

في هذه الحالة نلجأ إلى استخدام معامل ارتباط الرتب لقياس متانة العلاقة بين متغيرين نوعيين ، لكل منهما عدد كبير من الصفات أما إذا كان أحد المتغيرين نوعياً و الآخر كمياً ، و كذلك الأمر يمكن استخدامه في خالة المتغيرات الكمية أحياناً لاختصار الحسابات .

يتم حساب ارتباط الرتب باستبدال الصفات النوعية بقيم عددية تعبر عنها و تسمى الرتب Rank ، حيث يتم وضعها عن طريق إيجاد العلاقة الداخلية للصفات و

لكل متغير على حدى ، فنبدأ بأضعف الصفات مثلاً و نعطيها الرتبة ١ و الأقوى منها رقم ٢ و هكذا ، بحيث يتولد لدينا متوالية عددية لكل متغير يكون مجموعها متساوياً ، لأن عدد الصفات لكل متغير متساوية ، و بعد ذلك نحسب معامل ارتباط الرتب بالعلاقة :

$$r_s = 1 - \frac{6\sum D^2}{n(n^2-1)}$$

حيث :

r_s : معامل ارتباط سبيرمان (الرتب)

D : فرق رتبة X من رتبة Y لكل ثنائية

n : عدد المشاهدات (حجم العينة).

مثال :

ليكن لدينا الجدول الآتي الذي يمثل البيانات الخاصة بشكل علبة الحلوى و الطلب على هذا النوع من الحلوى

شكل علبة الحلوى X	نوع الطلب Y	رتب X	رتب Y	رتبX -رتبY D	D^2
علبة بسيطة دون رسوم	ضعيف	١	١	٠	٠
علبة بسيطة برسوم	وسط	٢	٢	٠	٠
علبة بسيطة برسوم و غلاف	جيد جداً	٣	٤	١ -	١
علبة مزخرفة	جيد	٤	٣	+١	١
علبة حديدية تقليدية	عالي الجودة	٥	٥	٠	٠
المجموع		١٥	١٥		٢

$$r_s = 1 - \frac{6\sum D^2}{n(n^2-1)} = 1 - \frac{6(2)}{5(5^2-1)} = 1 - \frac{12}{120} = 0.9$$

أي أن هناك علاقة قوية جداً بين شكل العلبة و نوع الطلب عليها . و تجب الإشارة هنا إلى أن معامل ارتباط الرتب يتمتع بالمزايا نفسها التي يتمتع بها معامل ارتباط بيرسون .

- إيجاد معادلة مستقيم الانحدار

إن دراسة الانحدار بين متغيرين X,Y تبدأ بتمثيل العلاقة بينهما بخط مستقيم يدعى

مستقيم الانحدار ، المعادلة الممثلة لهذا المستقيم تسمى بمعادلة الانحدار و هي من الشكل

$$\widehat{\mathrm{Y}}_{\mathrm{i}} = a + bX_i$$

حيث : $\widehat{Y}_{\mathrm{i}}$: القيمة المقدرة أو المحسوبة للمتغير التابع .

X_i : قيمة المتحول المستقل

a : ثابت الانحدار و يعني قيمة Y عندما تنعدم قيمة X أي نقطة تقاطع مستقيم الانحدار مع محور Y

b :معامل الانحدار (ميل مستقيم الانحدار) و يعني قيمة الزيادة في المتغير التابع *Y* عندما يزداد المتغير المستقل وحدة واحدة و يدعى ***a,b*** بثوابت معادلة مستقيم الانحدار.

و هذه المعادلة تدرس العلاقة الوسطية بين المتغيرين و شروطها

١. أن تكون العلاقة بين متغيرين فقط أحدهما مستقل و الآخر تابع

٢. أن يمكن تمثيل العلاقة بخط مستقيم أي معادلة من الدرجة الأولى .

و لإيجاد هذه المعادلة $\widehat{\mathrm{Y}}_{\mathrm{i}} = a + bX_i$ ما علينا إلا إيجاد قيمة الثوابت ***a,b*** اللذان يمكن إيجاد قيمهم من القوانين التالية :

$$b = \frac{\sum X_i Y_i - \bar{X} \sum Y_i}{\sum X_i^2 - \bar{X} \sum X_i}$$

$$a = \bar{Y} - b\bar{X}$$

مثال :

يبين الجدول الآتي كمية السماد المستخدم و غلة القطعة من القمح ل ١٢ قطعة أرض متساوية المساحة استخدمت في تجربة قام بها أحد الباحثين لمعرفة مدى تأثير السماد على الغلة فكانت النتائج :

رقم قطعة الأرض	كمية السماد ١ كغ X	الغلة ١٠ كغ Y	$X.Y$	X^2
١	٠	١٩	٠	٠
٢	٠	٢٠	٠	٠
٣	٠	٢١	٠	٠
٤	١	٢١	٢١	١
٥	١	٢٢	٢٢	١
٦	١	٢٣	٢٣	١
٧	٢	٢٣	٤٦	٤
٨	٢	٢٤	٤٨	٤
٩	٢	٢٥	٥٠	٤
١٠	٣	٢٥	٧٥	٩
١١	٣	٢٦	٧٨	٩
١٢	٣	٢٧	٨١	٩
المجموع	١٨	٢٧٦	٤٤٤	٤٢

$$b = \frac{\sum X_i Y_i - \bar{X} \sum Y_i}{\sum X_i^2 - \bar{X} \sum X_i} = \frac{444 - 1.5(276)}{42 - 1.5(18)} = \frac{30}{15} = 2$$

$$a = \bar{Y} - b\bar{X} = 23 - 2(1.5) = 20$$

و تصبح معادلة مستقيم الانحدار

$$\widehat{\mathrm{Y}}_{\mathrm{i}} = 20 + 2(X_i)$$

تفسير الثوابت a=20 و b=2

a=20 : أي نقطة تقاطع المستقيم مع محور Y أي عندما ينعدم السماد فإن الغلة بالمتوسط تكون ٢٠ كغ .

b=2 أي إذا زاد السماد بمقدار ١ كغ فإن الغلة تزداد بمقدار ٢ كغ

عندما $X_i = 0 \Leftarrow \widehat{\mathrm{Y}}_{\mathrm{i}} = 20 + 2(0) = 20$

و عندما $X_i = 1 \Leftarrow \overline{\mathrm{Y}}_{\mathrm{i}} = 20 + 2(1) = 22$

و عندما $\mathrm{X_i} = 2$ أي تزداد(١) فإن :

$$\widehat{\mathrm{Y}}_{\mathrm{i}} = 20 + 2(2) = 24$$

ازدادت الغلة بمقدار٢ بمقدار قيمة b

تستخدم معادلة مستقيم الانحدار للتنبؤ بقيم المتغير التابع عندما يأخذ المتغير المستقل قيماً محددة ، و لكن هنا يجب الانتباه إلى أي مدى يمكن أن تحافظ هذه العلاقة على نوعها و طبيعتها خارج حدود هذه البيانات ، و بشكل عام يمكننا القول إنه يمكن استخدام المعادلة في حدود معينة من الثقة على طرفي مجال قيم المتغير المستقل بمعنى أننا لا نستطيع أن نعطي قيم للمتحول المستقل X كبيرة في هذا المثال السماد بكميات كبيرة يصبح ضار على الغلة و الأرض

تمارين

١ - أحسب متانة العلاقة بين درجة نضج العنب ، و المسافة بين الأشجار في البيانات التي تم الحصول عليها من إحدى التجارب

٧	٦	٥	٤.٥	٤	٣	المسافة بالمتر
أحمر	محمر قليلاً	منقط بالأحمر	منقط بالأحمر	حصرم طري	حصرم	درجة النضج

٢ - الجدول الآتي يبين النفقات الشهرية و الدخل الشهري ل ١٠ مؤسسات متشابهة :

١٠	٩	٨	٧	٦	٥	٤	٣	٢	١	المؤسسة
٢٥	٢٠	١٩	١٧	١٦	١٥	١٢	١٠	١٠	٨	الدخل بالآلاف
٣٦	٣٥	٣٢	٣٠	٣٠	٢٦	٢٥	٢٤	٢٢	٢٠	الإنفاق بالآلاف

المطلوب :

١ - إيجاد العلاقة بين مستقيم الانحدار الممثلة للعلاقة و تفسير ثوابتها .

٢ - حساب معامل الارتباط المناسب و تفسيره .

٣ - بهدف دراسة العلاقة بين عدد سنوات الخبرة و الأجر الشهري للعامل في ورشات صيانة السيارات ثم جمع البيانات التالية عن ١٢ عاملاً:

١٢	١١	١٠	٩	٨	٧	٦	٥	٤	٣	٢	١	رقم العامل
٢٥	٢٣	٢٠	١٥	١٢	١٠	٩	٧	٥	٤	٢	١	عدد سنوات الخبرة
٦٠	٥٥	٥٠	٤٠	٣٠	٢٥	٢٠	١٦	١٢	١٠	٧	٥	الأجر الشهري

١٠٠٠ل.س												

المطلوب :

١. أوجد متانة العلاقة بين عدد سنوات الخبرة و الأجر الشهري بعد تحديد المتحول المستقل و المتحول التابع

٢. أوجد معادلة مستقيم الانحدار و فسّر ثوابتها

٤ -بهدف دراسة تأثير نوع من السماد على إنتاج القمح ثم استخدام كميات مختلفة من ذلك السماد في ٧ قطع متشابهة الظروف:

القطعة	١	٢	٣	٤	٥	٦	٧
كمية السماد كغ	٠	١	٢	٤	٥	٧	٩
الغلة كغ	٢٠٠	٢١٠	٢٣٥	٢٦٥	٢٦٠	٣٢٩	٣٢٥

المطلوب :

١. حساب معامل الارتباط المناسب بين كمية السماد و الغلة و تفسير قيمته.

٢. إيجاد معادلة مستقيم الانحدار و تفسيره .

اختبار الفرضيات

إن اختبار الفرضيات مسألة هامة جداً في التحليل الاحصائى ، لأنها تساعد على تفسير المعطيات والنتائج التجريبية و اتخاذ القرار المناسب لتلك المعطيات و تستند في اتخاذ القرار على الفرضية الصفرية (العدم) H_0 .

و هذه الفرضية لا تعترف بوجود الفروقات الحقيقية بين العينات أو بين متوسطات العينات المدروسة .

و للتحقق من صحة أي فرضية نقوم بسحب عينة عشوائية من المجتمع المدروس و نقوم بإجراء الحسابات اللازمة ثم نقارن القيمة التي حصلنا عليها من العينة مع القيمة النظرية التي نستخرجها من الجداول الخاصة باختبارات المعنوية ، و هذه الاختبارات متعددة و سنذكر منها فقط الاختبارات التالية :

1- اختبار t ستودنت .

2- اختبار كاي مربع X^2 .

3- اختبار فيشر F

وهذه الاختبارات تسمى اختبارات المعنوية لأنها تستخدم لاختبار صحة الفرضية الصفرية (العدم) أو عدم صحتها و ذلك بمقارنة القيم الفعلية (المحسوبة) لهذه الاختبارات مع القيم النظرية التي نستخرجها من الجداول الخاصة بهذه الاختبارات بعد معرفة درجات الحرية و مستوى المعنوية المطلوب

وقبل الحديث عن اختبارات المعنوية لا بد من توضيح بعض المفاهيم و العناوين و هي :

أولاً- الفرضية الصفرية (فرضية العدم) : و نرمز لها بالرمز H_0 .

عندما نريد اتخاذ قرار في أي موضوع من المواضيع قيد البحث من المفيد أن نضع فروض أو تخمينات عن المجتمع المدروس و هذه الفروض أو التخمينات قد تكون صحيحة أو غير صحيحة ، فمثلاً عندما نريد مقارنة عدة أسماء تجارية لمعالجة حالة صحية محددة بهدف اختبار الأفضل فالفرض الإحصائي يقول أن كل الأسماء التجارية المدروسة متساوية في العلاج أو لا يوجد فروقات بين تأثيرها أي الفرق بينها يساوي الصفر .

لنفترض أن هذه الأسماء التجارية عددها 3 فتكون الفرضية الصفرية كالتالي :

$$H_0 : \overline{X}_1 = \overline{X}_2 = \overline{X}_3$$

و كما هو واضح من العلاقة السابقة أو من الفرضية السابقة بأنه لا يوجد فروقات معنوية (حقيقية) بين تلك الأسماء التجارية و إن وجدت الفروقات فهي ليست حقيقية و إنما تعود للصدفة و العشوائية .

و بذلك يمكننا أن نقدم النص التالي للفرضية الصفرية التي تستخدم في جميع مجالات حياتنا مهما تكن .

بمعنى أن كل الباحثين في مختلف المجالات يضعون الفرضية الصفرية كأساس نظري لمقارنة نتائج أبحاثهم .

<u>الفرضية الصفرية تنص</u> : على عدم وجود فروقات معنوية (جوهرية أو حقيقية) بين متوسطات العينات المدروسة و إن وجدت بعض هذه الفروقات الظاهرية فهي ليست حقيقية و إنما تعود للصدفة و العشوائية أي لطريقة أخذ العينة و تسجيل القراءات أو بسبب حجم العينة أو لأي أسباب أخرى .

ثانياً- الفرضية البديلة (العكسية) : و نرمز لها بالرمز H_1 .

هذه الفرضية كما هو العنوان فرضية عكسية أي أننا نلجأ إليها عندما لا تتحقق الفرضية الصفرية ، أي عندما يكون المتوسط أكبر أو أصغر من الفرض و للتوضيح نأخذ المثال التالي : الفرضية الصفرية المتعلقة بمتوسط طول مجموعة من الطلاب تقول : $H_0: \overline{X} = 167cm$.

و الفرضية البديلة تقول أن متوسط طول هذه المجموعة هو أقل أو أكثر من 167 سم أو لا يساوي 167 سم

أي : $H_1: \overline{X} \neq 167cm$.

ثالثاً- مستوى المعنوية :

مستوى المعنوية في العرف الإحصائي يعبر عن درجة الاحتمال التي تقبل أو ترفض عنده الفرضية الصفرية و بتعبير آخر مستوى المعنوية يعبر عن درجة الدقة المطلوبة للنتائج أو درجة الخطأ المسموح فمثلاً عندما نقول إننا نريد دقة النتائج عند المستوى 5% فهذا يعني أن الخطأ المسموح هو 5% و دقة النتائج هي 95% و نفس الشيء يقال بالنسبة للمستوى 1% فالدقة المطلوبة هنا هي 99% و هكذا بالنسبة لبقية المستويات فكل باحث يختار مستوى المعنوية المناسب للظاهرة التي يدرسها فيمكن أن يختار مثلا واحد بالألف لأن النتائج المطلوبة تتطلب دقة عالية جداً و هذا يحدث مع الظواهر التي تتطلب مثل هذه الدقة لا سيما تلك الأشياء التي تتعلق بحياة الإنسان .

اختبارات المعنوية :

أولاً: اختبار t ستودنت:

هذا الاختبار يستخدم بكفاءة عالية عند مقارنة متوسطي عينتين أو متوسطي مجتمعين أو أي متوسطين ، و تحدد مجالات استخدامه في التالي :

أ- تقدير مجال الثقة - تقدير حدود المجتمع ..

ب- مقارنة متوسطي عينتين مستقلتين .

جـ- مقارنة متوسطي عينتين غير مستقلتين (مرتبطتين) .

أولاً- تقدير مجال الثقة :

إن العلاقة الرياضية المناسبة لإنشاء مجال الثقة يحوي القيمة الحقيقية لمتوسط المجتمع M تكون :

$$\overline{X} - \frac{\delta}{\sqrt{n}}(t_\alpha) < M < \overline{X} + \frac{\delta}{\sqrt{n}}(t_\alpha)$$

حيث أن الاحتمال $1-\alpha$ يساعدنا في حساب قيمة t_α النظرية (الجدولية) حيث α تمثل مستوى المعنوية الذي نختاره لدقة النتائج .

δ - الانحراف المعياري

$\overline{X}$ - المتوسط الحسابي للعينة

M - متوسط المجتمع المدروس .

t_α - هي القيمة النظرية التي نستخرجها من الجدول الخاص بتوزيع t عند مستوى المعنوية المحدد و أمام درجات الحرية n-1 .

و العلاقة الرياضية السابقة تفيدنا بمعرفة حدود العينة المدروسة أي حدها الأعلى و الأدنى أو حدود المجتمع المدروس بمعرفة حديه الأدنى و الأعلى .

فالحد الأدنى يساوي

$$\overline{X} - \frac{\delta}{\sqrt{n}}(t_{\alpha})$$

هذين الحدين كما يلاحظ من العلاقة السابقة

و الحد الأعلى يساوي

$$\overline{X} + \frac{\delta}{\sqrt{n}}(t_{\alpha})$$

و الحدين الأعلى و الأدنى تم حسابهما من العلاقة الخاصة بتوزيع ستودنت أي :

$$t = \frac{\overline{X} - M}{\frac{\delta}{\sqrt{n}}}$$

$\overline{X}$ متوسط العينة .

M متوسط المجتمع

δ الانحراف المعياري

و لتوضيح ذلك نأخذ المثال التالي :

مثال 1 :

سحبنا عينة مكونة من 25 طالب من طلاب كلية التربية و قمنا بحساب متوسط الوزن لهؤلاء الأشخاص فوجدناه 69 كغ بانحراف معياري قدره 3 كغ .

و المطلوب : حدد مجال الثقة لهذه العينة بدرجة ثقة تساوي 95% إذا علمت أن قيمة t النظرية عند المستوى 5% تساوي 2,064 .

الحل :

الحد الأدنى لوزن العينة =

$$\overline{X} - \frac{\delta}{\sqrt{n}}(t_{0,05}) = 69 - \frac{3}{\sqrt{25}}(2,064) = 67,76$$

الحد الأعلى لوزن العينة =

$$\overline{X} + \frac{\delta}{\sqrt{n}}(t_{0,05}) = 69 + \frac{3}{\sqrt{25}}(2,064) = 70,24$$

كما نلاحظ أن الحد الأدنى لأوزان هذه العينة هي 67,76 كغ و الحد الاعلى 70,24 كغ .

إذن $(67,76 < M < 70,24)$

أي أقل من 70,24 و أكثر من 67,76

مثال 2 :

إذا كان متوسط سكر الدم لجميع العاملين في إحدى كليات الجامعة يساوي 89 ، سحبنا عينة من 50 شخصاً فوجدنا أن متوسط سكر الدم لهؤلاء الأشخاص هو 94 بانحراف معياري قدره 2,5 .

و **المطلوب** :

هل هذه العينة تنتمي لمجتمع العاملين المذكورين أعلاه عند مستوى المعنوية 1% (درجة الثقة 99%) إذا علمت أن قيمة t عند المستوى 1% تساوى 2,58 .

الحل : الحد الأدنى للعينة السابقة :

$$\overline{X} - \frac{\delta}{\sqrt{n}}(t_{0,01}) = 94 - \frac{2,5}{\sqrt{50}}(2,58) = 93,08$$

الحد الأعلى لنفس العينة :

$$\overline{X}+\frac{\delta}{\sqrt{n}}(t_{0,01})=94+\frac{2,5}{\sqrt{50}}(2,58)=94,91$$

من هنا يتضح أن العينة ذات المتوسط 94 لا تتبع المتوسط 89 لأن حدي الثقة لمتوسط المجتمع التابعة له هذه العينة تقع ما بين 93,08 و 94,91 و لا تشمل القيمة 89 .

و يمكن أن نثبت هذا الكلام بطريقة أخرى و هي باستخدام اختبار t أي باستخدام العلاقة الرياضية الخاصة بهذا الاختبار في حال وجود عينة و مجتمع كما هو وارد في مثالنا السابق .

$$t=\frac{\overline{X}-M}{\delta_{\overline{X}}}$$

حيث : $\delta_{\overline{X}}$ تمثل الخطأ القياسي للعينة و تساوي $\frac{\delta}{\sqrt{n}}$

إذن :

$$t=\frac{94-89}{\frac{2,5}{\sqrt{50}}}=14,28$$

و هي قيمة t الفعلية (المحسوبة) .

و هذه القيمة نقارنها مع القيمة النظرية (الجدولية) لهذا الاختبار و التي تساوي 2,58 .

نلاحظ أن القيمة الفعلية (المحسوبة) لـــ t أكبر من القيمة النظرية لذلك نرفض نص الفرضية الصفرية ($H_0:M=\overline{X}$) الذي يقول أن متوسط العينة يساوي متوسط مجموعة العاملين في إحدى كليات الجامعة أو ينتمي له و بالتالي نقبل نص الفرضية البديلة ($H_1:M\neq\overline{X}$) الذي يقول أن متوسط العينة لا يساوي متوسط

مجموعة العاملين إذن العينة لا تنتمي للمجتمع المذكور أعلاه أي لمجتمع مجموعة العاملين في إحدى كليات الجامعة .

ثانياً- مقارنة متوسطي عينتين مستقلتين :

نفترض أنه لدينا عينتين عشوائيتين غير مرتبطتين أي أن أفراد العينة الأولى يختلفون تماماً عن أفراد العينة الثانية

و عدد أفراد العينتين n_1، n_2 تم اختيارهما من مجتمعين طبيعيين فالفرضية الصفرية لهاتين العينتين هي :

$H_0: \overline{X}_1 = \overline{X}_2$ و الفرضية البديلة $H_1: \overline{X}_1 \neq \overline{X}_2$ و العلاقة الرياضية لاختبار t في هذه الحالة تساوي :

$$t = \frac{\overline{X}_1 - \overline{X}_2}{S_D}$$

S_D

حيث تمثل δ_D الخطأ القياسي و يتوقف حسابها على حجم العينة .

فإذا كان حجم العينة كبيراً (أكبر من 30) فتحسب δ_D من العلاقة :

$$S_D = \sqrt{\frac{\delta_1^2}{n_1} + \frac{\delta_2^2}{n_2}}$$

S_1^2

حيث : $\delta_1^2 = \frac{\sum\left(X_1 - \overline{X}_1\right)^2}{n_1}$ التباين للعينة الأولى .

S_2^2

$\delta_2^2 = \frac{\sum\left(X_2 - \overline{X}_2\right)^2}{n_2}$ التباين للعينة الثانية .

أما إذا كان حجم العينة صغيراً (30 و أقل) فإن قيمة الخطأ القياسي تحسب من العلاقة التالية :

$$S_D = \sqrt{\frac{\sum\left(X_1 - \overline{X}_1\right)^2 + \sum\left(X_2 - \overline{X}_2\right)^2}{n_1 + n_2 - 2}} \times \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

و يمكن أن تحسب $\sum\left(X_1 - \overline{X}_1\right)^2$ و $\sum\left(X_2 - \overline{X}_2\right)^2$ بطريقة تربيع القيم التالية :

$$\sum\left(X_1 - \overline{X}_1\right)^2 = \sum X_1^2 - \frac{\left(\sum X_1\right)^2}{n_1}$$

$$\sum\left(X_2 - \overline{X}_2\right)^2 = \sum X_2^2 - \frac{\left(\sum X_2\right)^2}{n_2}$$

فهنا نختار الطريقة التي نراها أسهل .

<u>مثال للعينات الصغيرة :</u>

إحدى مستشفيات التوليد الخاصة

قمنا بزيارة قسم التوليد في [redacted] و سجلنا أوزان ستة أطفال حديثي الولادة و كانت كما يلي :

X_1 : 3,8 3,7 2,9 3,5 2,6 3,3

و ذهبنا إلى المشفى الوطني و بقسم التوليد أيضاً سجلنا أوزان ثمانية أطفال :

X_2 : 3,7 4,6 5,4 6,2 4,2 3,5 5,3 5,5

و المطلوب :

هل يوجد اختلاف معنوي (حقيقي) بين متوسطي العينتين إذا علمت أن قيمة t عند المستوى 1% تساوي 3,055 .

الحل :

$$t = \frac{\overline{X}_1 - \overline{X}_2}{S_D}$$

بما أن العينة صغيرة فإن قيمة الخطأ القياسي S_D و تحسب من العلاقة :

$$S_D = \sqrt{\frac{\sum\left(X_1 - \overline{X}_1\right)^2 + \sum\left(X_2 - \overline{X}_2\right)^2}{n_1 + n_2 - 2}} \times \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

و لحساب هذه الانحرافات ننشئ الجدول التالي :

X_1	$\left(X_1 - \overline{X}_1\right)$	$\left(X_1 - \overline{X}_1\right)^2$	X_1^2	X_2	$\left(X_2 - \overline{X}_2\right)$	$\left(X_2 - \overline{X}_2\right)^2$	X_2^2
3,8	0,5	0,25	14,44	3,7	−1,1	1,21	13,69
3,7	0,4	0,16	13,69	4,6	0,2	0,04	21,16
2,9	−0,4	0,16	8,41	5,4	0,6	0,36	29,16
3,5	0,2	0,04	12,25	6,2	1,4	1,96	38,44
2,6	−0,7	0,49	6,76	4,2	−0,6	0,36	17,64
3,3	0	0	10,89	3,5	1,3	1,69	12,25
				5,3	0,5	0,25	28,09
				5,5	0,7	0,49	30,25
19,8	0	1,1	66,44	38,4		6,36	190,68

إذن يمكن أن نطبق إحدى الطريقتين الانحرافات أو تربيع القيم بعد أن قمنا بحساب متوسط العينة الأولى .

$$\overline{X}_1 = \frac{\sum X_1}{n_1} = \frac{19,8}{6} = 3,3$$

$$\overline{X}_2 = \frac{\sum X_2}{n_2} = \frac{38,4}{8} = 4,8$$

$$\sum\left(X_1 - \overline{X}_1\right)^2 = \sum X_1^2 - \frac{\left(\sum X_1\right)^2}{n_1} = 66,44 - \frac{(19,8)^2}{6} = 1,1$$

$$\sum\left(X_2 - \overline{X}_2\right)^2 = \sum X_2^2 - \frac{\left(\sum X_2\right)^2}{n_2} = 190,68 - \frac{(38,4)^2}{8} = 6,36$$

$$S_D = \sqrt{\frac{1,1+6,36}{6+8-2}} \times \sqrt{\frac{1}{6}+\frac{1}{8}} = 0,426$$

$$t = \frac{3,3-4,8}{0,426} = -3,$$[redacted]$$52$$

و الآن نقارن قيمة t الفعلية $\left|-3,52\right|$ بالقيمة المطلقة مع قيمة t النظرية 3,055 نلاحظ أن قيمة t المحسوبة أكبر من قيمة t النظرية لذلك نرفض نص الفرضية الصفرية ($H_0: \overline{X}_1 = \overline{X}_2$) الذي يقول بعدم وجود فروقات معنوية (حقيقية) بين متوسطي العينتين و بالتالي نقبل نص الفرضية البديلة ($H_1: \overline{X}_1 \neq \overline{X}_2$)

(إذن يوجد فرق معنوي (حقيقي) أو جوهري بين متوسط أوزان الأطفال في المستشفى الخاص ~~مستشفى الجامعي~~ و أوزان الأطفال في المشفى الوطني .

مثال آخر للعينات الكبيرة :

لدينا العينتين التاليتين : الأولى و هي الشاهد و عدد أفرادها 50 و متوسطها الحسابي 10,51 و تباينها 7,99 و عدد أفراد العينة الثانية 64 و متوسطها الحسابي يساوي 9,23 و التباين 7,23

السؤال :

هل يوجد اختلاف معنوي بين متوسطي العينتين إذا علمت أن قيمة t النظرية (الجدولية) تساوي 1,96 عند المستوى 5% .

الحل :

$$S_D = \sqrt{\frac{\delta_1^2}{n_1} + \frac{\delta_2^2}{n_2}}$$

بما أن n_1+n_2 كبيراً لذلك S_D تحسب من العلاقة التالية :

و في نص المسألة :

$$\left[\delta_1^2 = 7,99\right] \; و \; \left[\overline{X}_1 = 10,51\right] \; و \; \left[n_1 = 50\right]$$

$$\left[\delta_2^2 = 7,23\right] \; و \; \left[\overline{X}_2 = 9,23\right] \; و \; \left[n_2 = 64\right]$$

$$S_D = \sqrt{\frac{7,99}{50} + \frac{7,23}{64}} = 0,52$$

$$t = \frac{10,51 - 9,23}{0,52} = 2,46$$

و هي قيمة t الفعلية

و للإجابة عن السؤال السابق نقارن قيمة t المحسوبة 2,46 مع قيمة t النظرية (الجدولية) 1,96 نلاحظ أن قيمة t المحسوبة أكبر من قيمة t النظرية لذلك

نرفض نص الفرضية الصفرية ($H_0: \overline{X}_1 = \overline{X}_2$) الذي يقول بعدم وجود فروقات معنوية بين متوسطي العينتين أو بتساوي متوسطي العينتين و بالتالي نقبل نص الفرضية البديلة ($H_1: \overline{X}_1 \neq \overline{X}_2$)

إذن يوجد فرق جوهري (معنوي) بين متوسطي العينتين المدروستين أي يوجد اختلاف بين المتوسطين .

مثال 1 :

أجريت معايرة كمية الخضاب في الدم لعينة مؤلفة من 48 طفلاً فكان متوسط هذه العينة 11,3 مختارة من مجتمع انحرافه المعياري 2,5 غ و المطلوب : أوجد مجال الثقة لمجتمع الأطفال الذي أخذت منه هذه العينة إذا علمت أن قيمة t النظرية بدرجة ثقة 95% تساوي 1,96 أو يمكن أن نطرح السؤال بشكل آخر وهو أوجد حدود المجتمع الذي أخذت منه العينة أو حدود العينة بنفس الشروط السابقة .

الحل :

مجال الثقة المطلوب هو :

$$L_1 = \overline{X} - \frac{\delta}{\sqrt{n}}(t_{0,05})$$

$$L_2 = \overline{X} + \frac{\delta}{\sqrt{n}}(t_{0,05})$$

و منه نجد أن الحد الأدنى يساوي :

$$L_1 = 11,3 - \frac{2,5}{\sqrt{49}}(1,96) = 10,6$$

$$L_2 = 11,3 + \frac{2.5}{\sqrt{49}}(1,96) = 12$$

إذن مجال الثقة في مجتمع الأطفال هو [10,6 و 12] أو بمعنى آخر أن حدود مجتمع العينة يقع بين 10,6 و 12 .

مثال 2 :

لدى قياس معدل البولة في الدم لـــ 25 شخصاً مختارين من مجتمع إحصائي له التوزيع الطبيعي و وجدنا أن المتوسط الحسابي لمعدل البولة في دم العينة يساوي $\overline{X} = 0,29g/l$ و الانحراف المعياري $\delta = 0,05g/l$ و المطلوب : إيجاد مجال الثقة لمتوسط مجتمع الأشخاص الذي أخذت منه هذه العينة إذا علمت أن قيمة t النظرية بدرجة ثقة 90% تساوي 1,711 .

الحل :

مجال الثقة هو :

$$L_1 = \overline{X} - \frac{\delta}{\sqrt{n}}(t_{0.01})$$

$$L_2 = \overline{X} + \frac{\delta}{\sqrt{n}}(t_{0.01})$$

و الآن نجد الحد الأدنى لمجتمع العينة :

$$L_1 = 0,29 - \frac{0,05}{\sqrt{25}}(1,711) = 0,288$$

$$L_2 = 0,29 + \frac{0,05}{\sqrt{25}}(1,711) = 0,291$$

إذن مجال الثقة في مجتمع العينة هو [0,288 و 0,291]

أو بمعنى آخر أن حدود مجتمع العينة يقع بين 0,288 و 0,291 .

مثال :

إذا كان متوسط عمر صمام في جهاز طبي من نوع ألفا هو 6,5 سنة بانحراف معياري قدره 0,9 سنة بينما متوسط عمر صمام آخر في جهاز طبي من نوع بيتا هو 6 سنة بانحراف معياري 0,8 سنة .

و المطلوب :

اختبار معنوية الفروق بين متوسط عمر عينة الصمامات من نوع ألفا التي حجمها 36 و متوسط عمر عينة الصمامات من نوع بيتا و حجمها 49 بدرجة ثقة 95%.

الحل :

لدينا المعطيات التالية :

$n_1 = 36$, $\overline{X_1} = 6.5$, $\delta_1 = 0.9$

$n_2 = 49$, $\overline{X_2} = 6$, $\delta_2 = 0,8$

نلاحظ أن عدد أفراد العينة الأولى و الثاني كبير لذلك سنقوم بحساب الخطأ القياسي بما يتناسب و العينات الكبيرة إذن :

$$t = \frac{\overline{X}_1 - \overline{X}_2}{S_D}, S_D = \sqrt{\frac{\delta_1^2}{n_1} + \frac{\delta_2^2}{n_2}} = \sqrt{\frac{(0,9)^2}{36} + \frac{(0,8)^2}{49}} = 0,189$$

ملاحظة : نلاحظ من العلاقة السابقة أننا قمنا بتربيع قيمة الانحراف المعياري لكلا العينتين و هذا منبه اننا نريد التباين و ليس الانحراف المعياري كما هو واضح في العلاقة السابقة حيث يرمز للتباين بـ δ^2

$$t = \frac{6,5-6}{0,189} = 2,64$$

ونعود لحساب t الفعلية من العلاقة السابقة :

و نستخرج قيمة t النظرية من الجدول بدرجة ثقة 95% و أمام درجات الحرية :

$$n_1 + n_2 - 2 = 36 + 49 - 2 = 83$$

فنجد أن قيمة t الفعلية (المحسوبة) 2,64 مع قيمة t النظرية 1,96 نلاحظ أن قيمة t الفعلية أكبر من قيمة t النظرية لذلك نرفض نص الفرضية الصفرية (H_0: $\overline{X_1} = \overline{X_2}$) الذي ينص بعدم وجود فروقات معنوية (جوهرية) بين متوسط عمر الصمام من نوع ألفا و عمر الصمام من نوع بيتا و بالتالي نقبل نص الفرضية البديلة (H_1: $\overline{X_1} \neq \overline{X_2}$) إذن يوجد فرق معنوي (حقيقي و جوهري) بين متوسط عمر الصمامين .

مثال :

أجريت دراسة على عينتين من الأطفال المولودين حديثاً إحداهما من الذكور و الأخرى من الإناث حجمها على التوالي $n_1=n_2=10$ و وجد أن متوسط وزن الطفل في العينة الأولى (ذكور) = 3,5 و متوسط وزن العينة الثانية (إناث) = 3 و تباين العينة الأولى 25 و الثانية 22 .

و المطلوب :اختبار معنوية الفروق بين متوسطي العينتين إذا علمت أن t تحسب عند المستوى 0,05% (بدرجة ثقة 95%) .

الحل :

المعطيات المتوفرة بنص المسألة : $n_1=10$ ، $\overline{X_1} = 3{,}5$ ، $\delta_1^2 = 25$

$n_2 = 10$ ، $\overline{X_2} = 3$ ، $\delta_2^2 = 22$

و كما نلاحظ أن عدد أفراد العينتين صغير لذلك سنقوم بحساب الخطأ القياسي S_D بما يتوافق و العينات الصغيرة :

$$t=\frac{\overline{X_1}-\overline{X_2}}{S_D}, S_D=\sqrt{\frac{\sum\left(X_1-\overline{X_1}\right)^2+\sum\left(X_2-\overline{X_2}\right)^2}{n_1+n_2-2}}\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}=\sqrt{\frac{\delta_1^2}{n_1}+\frac{\delta_2^2}{n_2}}=2,168$$

لعدم توفر البيانات المطلوبة طبقنا قانون العينات الكبيرة

إذن :

$$t=\frac{3,5-3}{2,168}=0,23$$

و هي قيمة t الفعلية .

نقارن قيمة t الفعلية 0,23 مع قيمة t النظرية بدرجة ثقة 95% و درجات الحرية 18 فنجدها تساوي 1,734 نلاحظ أن قيمة t الفعلية أصغر من القيمة النظرية لذلك نقبل نص الفرضية الصفرية الذي يقول بعدم وجود فروقات بين متوسطي الوزن في كلتا العينتين $\left(H_0:\overline{X_1}=\overline{X_2}\right)$ و بالتالي نرفض نص الفرضية البديلة .

إذن لا يوجد فرق معنوي (جوهري) حقيقي بين متوسط وزن عينة الذكور و متوسط وزن عينة الإناث و الفرق الموجود (الظاهري) بين العينتين غير حقيقي و إنما يعود للصدفة و العشوائية .

تطبيقات غير محلولة :

مثال 1:

قام أحد الباحثين بأخذ عينة حجمها n=33 من مياه بحيرة سد 16 تشرين و قام بقياس كمية الصوديوم في هذه العينة فوجد أن متوسط وزن الصوديوم في هذه العينة 24,11 ملغ/ليتر و الانحراف المعياري 2,44 ملغ/ليتر

و المطلوب :

إيجاد مجال الثقة لهذه العينة (حدي مجتمع العينة) بدرجة ثقة 95% (قيمة t النظرية عند المستوى 5% = 1,96) .

مثال 2:

أجري اختبار في دار التوليد بدمشق لمعايرة التيروكسين لدى 49 وليداً فوجدنا أن المتوسط الحسابي لهذه العينة $\overline{X}=9,8$ بانحراف معياري قدره 3,1

والمطلوب :

تقدير حدود الثقة لمجتمع العينة (مجال للثقة) إذا علمت أن قيمة t تحسب عند المستوى 0,1% (درجة الثقة 90%) علماً أن قيمة t النظرية بدرجة ثقة 90 % = 1,296

مثال 3:

مجموعتين من المرضى ، تتكون الأولى من 50 شخصاً و الثانية من 100 شخص ، المجموعة الأولى أعطيت نوعاً جديداًمن الحبوب المنومة فكان وسط ساعات النوم 7,82 بانحراف 0,25 ساعة و أعطيت المجموعة الثانية نوعا

معروفاً من الحبوب فكان متوسط ساعات النوم 6,75 ساعة بانحراف معياري 0,3 ساعة .

و المطلوب :

اختبار معنوية الفرق بين نوعي الحبوب المستخدمة بمستوى 0,05 (قيمة t عند المستوى 0,05 = 1,96) .

مثال 4:

لدينا العينتين التاليتين اللتين تمثلان مستوى التحصيل لدى الطلبة :

X_1 : 15 14 10 9 11 16

X_2 : 20 19 14 15 16 17

و المطلوب :

اختبار معنوية الفرق بين مستوى التحصيل لكلا المجموعتين إذا علمت أن :

قيمة t تحسب بدرجة ثقة 95% و تساوي 2,179 .

ثالثاً– مقارنة متوسطي عينتين غير مستقلتين (مرتبطتين) :

في كثير من الأحيان نأخذ عينة واحدة و نسجل المشاهدات لمفرداتها ، ثم تخضع هذه العينة تحت تأثير معين و نعود لتسجيل الجديدة لمفردات العينة نفسها و بعملية المقارنة لهاتين القراءتين لمفردات العينة يمكن أن نستنتج مدى تأثير هذا المؤثر الذي وضعت العينة تحته .

و هذا الاختبار يستخدم كثيراً في المجالات المختلفة لحياة الإنسان ، فمثلاً يستخدم في مجال تجارب العقاقير الطبية لبيان تأثير عقار معين قبل استخدامه و بعده و

كذلك تأثير برنامج غذائي معين على زيادة الوزن و أشياء أخرى كثيرة و يتم حساب قيمة الاختبار الفعلية كما يلي :

$$t = \frac{\overline{d}}{\delta_d}$$

حيث : $\overline{d}$ تمثل المتوسط الحسابي للفرق بين قيم أفراد العينة المدروسة قبل و بعد الاختبار (التأثير)

δ_d الخطأ القياسي للفرق و يتم حساب المتوسط السابق و الخطأ القياسي للفرق كما يلي :

1- نحسب الفرق بين كل زوج مقابل من أزواج القيم (أفراد العينة قبل و بعد التأثير)

$$\sum (X_1 - \overline{X}_2) = \sum d$$

2- نحسب متوسط الفرق :

$$\overline{d} = \frac{\sum d}{n}$$

3- نحسب تباين الفرق :

$$\delta_d^2 = \frac{\sum d^2}{n} - \left(\frac{\sum d}{n}\right)^2 \text{ أو } \delta_d^2 = \frac{\sum \left(d - \overline{d}\right)^2}{n} = \frac{\sum d^2 - \frac{\left(\sum d\right)^2}{n}}{n}$$

4- الخطأ القياسي للفرق :

$$\delta_{\overline{d}} = \sqrt{\frac{\delta_d^2}{n}}$$

5- نقارن قيمة t المحسوبة مع t النظرية (الجدولية) عند مستوى المعنوية المطلوب و أمام درجات الحرية $n-1$ ثم نأخذ القرار الإحصائي بقبول الفرضية الصفرية أو برفضها .

مثال :

أخذت عينة من عشرة أطفال و سجلت أوزانها الأولية ثم أعطيت وجبة غذائية محددة و لمدة خمسة عشر يوما ، ثم سجلت أوزانها من جديد و كانت النتائج على الشكل التالي :

الوزن قبل التغذية X_1	17	17,5	18	18,5	19	19,5	20	21	21	21,5	المجموع
الوزن بعد التغذية X_2	17,7	18,5	18,9	19,3	20,1	19,9	20,9	21,7	21,7	22,3	-
$d=X_1-X_2$	-0,7	-1	-0,9	-0,8	-1,1	-0,4	-0,9	-0,7	-0,7	-0,8	-8
d^2	0,49	1	0,81	0,64	1,21	0,16	0,81	0,49	0,49	0,64	6,74

و المطلوب : اختبار معنوية الفروق بين الوزنين بدرجة ثقة 95% .

الحل : الفروض

$$H_0 : \bar{d} = 0$$
$$H_1 : \bar{d} \neq 0$$

$$\bar{d} = \frac{\sum d}{n} = \frac{X_1 - X_2}{n} = \frac{-8}{10} = -0,8$$

نحسب تباين الفرق :

$$\delta_d^2 = \frac{\sum d^2 - \frac{\left(\sum d\right)^2}{n}}{n} = \frac{\sum d^2}{n} - \left(\frac{\sum d}{n}\right)^2 = \frac{6,74 - \frac{(-8)^2}{10}}{10} = \frac{6,74}{10} - \left(\frac{-8}{10}\right)^2 = 0,034$$

إذن الخطأ يساوي :

$$\delta_{\bar{d}} = \sqrt{\frac{\delta_d^2}{n}} = \sqrt{\frac{0,034}{10}} = 0,058$$

نحسب قيمة t الفعلية :

$$t = \frac{\bar{d}}{\delta_{\bar{d}}} = \frac{-0,8}{0,058} = -13,79 = |13,79|$$

نضع الفروض

$$H_0 : \bar{d} = 0$$
$$H_1 : \bar{d} \neq 0$$

القرار الإحصائي : نقارن قيمة t المحسوبة (الفعلية) 13,79− بالقيمة المطلقة مع قيمة t الجدولية (النظرية) عند المستوى 0,05% و درجات الحرية n-1 أي : 9=1−10 فتساوي 2,26 فنلاحظ أن t المحسوبة أكبر من t الجدولية لذلك نص الفرضية الصفرية الذي يقول بعدم وجود فروقات معنوية أي عدم وجود تأثير للبرنامج الغذائي على زيادة الوزن و بالتالي نقبل الفرضية البديلة إذن يوجد فرق معنوي أي يوجد تأثير للبرنامج الغذائي (الوجبة الغذائية) على زيادة أوزان الأطفال .

توزيع كاي مربع X^2

كثيراً ما يحتاج الباحث لمعرفة مدى مطابقة نتائج حصل عليها بطريقة الملاحظة أو التجربة و نتائج أخرى يمكن الحصول عليها بالطرق النظرية لنفس الظاهرة .

فاختبار كا مربع أو اختبار المطابقة يجري بين مؤشر تجريبي و الآخر نظري و عادة النتائج التجريبية ليس من الضروري أن تتفق مع النتائج المتوقعة وفقاً لقواعد الاحتمال أو وفقاً للتوزيع النظري ، فعلى سبيل المثال إن الاحتمالات النظرية تؤدي بنا إلى توقع 50 أنثى و 50 ذكر في 100 ولادة و الحقيقة أنه من النادر الحصول على هذه النتيجة بالضبط .

و هذا الاختبار يستعمل بشكل واسع في التحليل الوراثي و الطبي و مجالات كثيرة و متعددة و نرمز للتكرارات الفعلية (التجريبية) بالرمز [O]و التكرارات النظرية أو المتوقعة [E] و بذلك تكون معادلة الاختبار :

$$\chi^2 = \sum \frac{(O-E)^2}{E}$$

حيث : χ^2 – كاي مربع

O – المشاهدات الفعلية (الحقيقية)

E – المشاهدات النظرية (المتوقعة)

- القيمة الناتجة لـ χ^2 نقارنها مع القيمة النظرية التي نستخرجها من الجدول الخاص بهذا الاختبار (توزيع كا مربع) و الموجود في كتب الاحصاء ، على أساس درجات الحرية k-1 (حيث k تمثل عدد المجموعات المدروسة) ومستوى المعنوية المطلوب أي درجة الثقة المطلوبة (درجة الدقة التي نختارها لنتائج البحث) .

ثم نجري المقارنة بين قيمة χ^2 الفعلية (المحسوبة) من العلاقة الخاصة بهذا الاختبار و الموجودة أعلاه مع القيمة النظرية (الجدولية) لهذا الاختبار .

فإذا كانت قيمة كاي مربع χ^2 الفعلية أكبر أو تساوي قيمة كاي مربع النظرية نرفض الفرضية الصفرية $(H_0:O=E)$ و التي تقول بعدم وجود فروقات معنوية (حقيقية) بين التكرارات التجريبية (الفعلية) و التكرارات النظرية و بالتالي نكون قد قبلنا الفرضية البديلة $(H_1:O\neq E)$.

أما إذا كانت قيمة كاي مربع الفعلية (المحسوبة) أصغر من قيمة كاي مربع النظرية (الجدولية) نقبل الفرضية الصفرية $(H_0:O=E)$ القائلة بعدم وجود فروقات معنوية (جوهرية) و نرفض الفرضية البديلة (العكسية) $(H_1:O\neq E)$ إذن لا يوجد فرق معنوي بين التكرارات الفعلية و التكرارات النظرية .

مثال 1 :

بالعودة إلى سجلات أحد مشافي التوليد وجدنا توزيع 1000 طفل مولود تبعاً للجنس كالآتي : 560 إناث و 440 ذكور .

و السؤال :

هل هده النتيجة تؤيد أو تؤكد الفرضية القائلة بتساوي الذكور و الإناث عند الولادة بدرجة ثقة 95% .

الحل :

الفرضية الصفرية و التي تقول بتساوي ولادة الذكور و الإناث $H_0:O=E$

O – تمثل المشاهدات الفعلية كما هو وارد في نص المسألة .

E – المشاهدات النظرية حسب احتمال ولادة الذكور و الإناث .

و الفرضية البديلة (العكسية) و التي تقول بعدم تساوي ولادة الذكور و الإناث $H_1: O \neq E$

إذن التوزيع الفعلي هو كالتالي : O = 560 إناث + 440 ذكور

و التوزيع النظري هو : E = 500 إناث + 500 ذكور (حسب احتمال ولادة الذكور و الإناث) .

و الآن نطبق العلاقة كاي مربع :

$$\chi^2 = \sum \frac{(O-E)^2}{E} = \frac{(560-500)^2}{500} + \frac{(440-500)^2}{500} = 14,4$$

و هي القيمة الفعلية .

و الآن نستخرج من جدول كاي مربع القيمة النظرية عند المستوى 5% و درجات الحرية k-1=2-1=1 فنجدها 3,841 بينما قيمة كاي مربع الفعلية 14,4 و هي أكبر من القيمة النظرية (الجدولية) و هذا يؤدي إلى رفض الفرضية الصفرية السابقة و القائلة بعدم وجود فروقات معنوية (حقيقية) بين ولادة الذكور و الإناث و بالتالي نقبل نص الفرضية القائلة بعدم تساوي ولادة الذكور و الإناث أي وجود اختلاف جوهري (معنوي) بين ولادة الذكور و الإناث و الاختلاف الموجود هو اختلاف حقيقي و لا يعود للصدفة و العشوائية كما تنص الفرضية الصفرية .

مثال 2 :

أثبتت الدراسات في العام الماضي أن 65% من السكان يفضلون الملابس التي تصنع محلياً و بعد ذلك أدخلت الكثير من التعديلات لتحسين جودة الإنتاج و صار هناك شعور بأن هذه النسبة قد زادت للتأكد من صحة الشعور أو عدمه أخذت

عينة حجمها 300 شخص و اتضح أن هناك 188 شخص يستعملون الملابس المنتجة محلياً .

السؤال :

هل هناك ما يؤكد تصور المنتجين المحليين أو هل يتطابق ما يتصوره المنتجين المحليين مع نتائج العينة باحتمال قدره 5% .

الحل :

الفرضية الصفرية $H_0:O=E$ و الفرضية البديلة $H_1:O\neq E$

إذن التوزيع الفعلي هو كالتالي : O = 188 يستعملون الملابس المحلية + 122 يستعملون الملابس المستوردة و التوزيع النظري هو :

E = الملابس المحلية (195=0,65×300) + الملابس المستوردة (105) .

و الآن نطبق العلاقة كاي مربع :

$$\chi^2=\sum\frac{(O-E)^2}{E}=\frac{(188-195)^2}{195}+\frac{(122-105)^2}{105}=0,718$$

و هي القيمة الفعلية

و الآن نستخرج من جدول كاي مربع القيمة النظرية $\chi^2=3,841$

و الآن نقارن قيمة χ^2 الفعلية 0,718 مع القيمة النظرية 3,841 فنجد أن القيمة الفعلية (المحسوبة) أصغر من القيمة النظرية لذلك نقبل نص الفرضية الصفرية الذي يقول بأن 65% من السكان يفضلون الملابس المحلية و نرفض الفرضية البديلة التي تقول بأن النسبة أكثر من ذلك.

إذن عند إجراء التطابق بين الدراسات التي أجريت في الماضي و التي أثبتت أن 65% من السكان يفضلون الملابس المصنوعة محلياً وبين الشعور بزيادة النسبة تبين أن النسبة 65% ما زالت قائمة و لم تتغير نسبة استعمال الناس للملابس المحلية .

اختبار الاستقلالية أو اختبار χ^2 لبيانات جدول التوافق :

هو نفس اختبار التطابق أو اختبار كاي مربع إلا أنه يرتب النتائج (البيانات الإحصائية) في صفوف و أسطر أي ضمن جداول فهنا ندرس مدى استقلالية صفتين أو خاصتين فالخاصة الأولى نضعها ضمن الأسطر و الخاصية الثانية ضمن الأعمدة و نطبق العلاقة المعروفة لاختبار كاي مربع :

$$\chi^2 = \sum \frac{(O-E)^2}{E}$$

مثال :

قام فريق من الأطباء بدراسة لمعرفة آثار التدخين على الصحة العامة و تم اختبار عينتين تضم الأولى 425 من المدخنين و الثانية 425 من غير المدخنين و بعد إجراء الفحوصات اللازمة توصل فريق الأطباء إلى النتائج التالية :

الحالة الصحية / العينة	جيدة	غير جيدة	المجموع
يدخن	311	114	425
لا يدخن	375	50	425
المجموع	686	164	850

و **المطلوب :**

اختبار مدى تأثير التدخين على الصحة العامة بدرجة ثقة 99%

الحل :

لدينا توزيعين التوزيع الأول يتعلق بالنتائج التجريبية أي النتائج التي توصل إليها فريق الأطباء وهي موجودة في الجدول أعلاه .

أما النتائج النظرية فنحن من سيقوم بحسابها ثم نضعها في جدول كالجدول السابق و بذلك سيصبح لدينا جدولين الأول يحتوي النتائج الفعلية و الجدول الثاني النتائج النظرية .

– الأشخاص اللذين يدخنون و حالتهم الصحية جيدة = $\frac{686 \times 425}{850} = 343$ شخص

كما لاحظنا قمنا بضرب مجموع السطور الذي يحتوي على المدخنين و هو 425 بمجموع الأعمدة الذي يحوي المدخنين أيضاً و هو 686 و قسمنا على المجموع العام .

– الأشخاص اللذين يدخنون و حالتهم الصحية غير جيدة = $\frac{164 \times 425}{850} = 82$ شخصاً

قمنا بضرب مجموع السطور الذي يحوي على غير المدخنين 425 و مجموع العمود الذي يحوي غير المدخنين أيضا 164 و قسمنا على المجموع العام .

– الأشخاص اللذين لا يدخنون و حالتهم الصحية جيدة = $\frac{686 \times 425}{850} = 343$

– الأشخاص اللذين يدخنون و حالتهم الصحية غير جيدة = $\frac{164 \times 425}{850} = 82$

و بعدها ننشئ الجدول التالي :

الحالة الصحية / العينة	جيدة	غير جيدة	المجموع
يدخن	343	82	425
لا يدخن	343	82	425
المجموع	686	164	850

نتائج الفحوصات نظرياً

ثم نطبق العلاقة :

$$\chi^2 = \sum\frac{(O-E)^2}{E} = \frac{(311-343)^2}{343} + \frac{(114-82)^2}{82} + \frac{(375-343)^2}{343} + \frac{(50-82)^2}{82} = 30,96$$

و هي القيمة الفعلية

و الآن نستخرج القيمة النظرية لكاي مربع من الجدول الخاص بهذا الاختبار بعد معرفة عدد درجات الحرية من العلاقة

(عدد الأسطر - 1) × (عدد الأعمدة - 1) = (2-1) (2-1) = 1

و مستوى المعنوية 0,01% تكون قيمة كاي مربع الجدولية

$\chi^2 = 6,635$

الفروض الإحصائية : الصفرية H_0: $O=E$

البديلة H_1: $O \neq E$

القرار الإحصائي : إن قيمة χ^2 الفعلية 30,96 أكبر من قيمة χ^2 الجدولية (النظرية) 6,635 لذلك نرفض الفرضية الصفرية القائلة بعدم وجود تأثير

للتدخين على الصحة العامة و بالتالي نقبل بالفرضية البديلة القائلة بوجود تأثير للتدخين على الصحة العامة ، إذن يوجد تأثير معنوي (جوهري) حقيقي للتدخين على الصحة العامة .

مثال : في دراسة لمعرفة مدى الارتباط بين الحالة الزوجية و القدرة على الإدخار فقام فريق من الباحثين بأخذ عينة مؤلفة من 880 أعزب و عينة أخرى مؤلفة من 2447 متزوج و سجلت النتائج التالية :

الحالة الزوجية	عدد المدخرين	غير المدخرين	المجموع
أعزب	490	390	880
متزوج	1552	895	2447
المجموع	2042	1285	3327

و المطلوب : اختبار الفرض القائل بأن سلوك المتزوجين في الادخار لا يختلف اختلافاً جوهرياً عن سلوك غير المتزوجين بدرجة ثقة 95% .

الحل : ننشئ جدول مماثل للجدول السابق يتعلق بالقيم النظرية (النتائج النظرية) كما يلي :

عدد المدخرين من غير المتزوجين =

$$\frac{\text{(مجموع السطر الذي يحوي غير المتزوجين × مجموع العمود الذي يحوي المدخرين)}}{\text{المجموع العام}}$$

عدد المدخرين غير المتزوجين = $\frac{2042 \times 880}{3327} = 540,1$

عدد غير المدخرين من غير المتزوجين = $\frac{1285 \times 880}{3327} = 339,3$

عدد المدخرين من المتزوجين = $\frac{2042 \times 2447}{3327} = 1501,9$

عدد غير المدخرين من المتزوجين = $\frac{1285 \times 2447}{3327} = 945,1$

و الآن نضع النتائج السابقة في جدول مشابه للجدول السابق :

الحالة الزوجية	عدد المدخرين	غير المدخرين	المجموع
أعزب	540,1	339,9	880
متزوج	1501,9	945,1	2447
المجموع	2042	1285	3327

و نطبق العلاقة كاي مربع :

$$\chi^2 = \sum \frac{(O-E)^2}{E} = \frac{(490-540,1)^2}{540,1} + \frac{(390-339,9)^2}{339,9} + \frac{(1552-1501,9)^2}{1501,9} + \frac{(895-945,1)^2}{945,1} = 16,36$$

و هي قيمة χ^2 الفعلية

نستخرج من الجدول قيمة χ^2 النظرية أمام درجات الحرية

(عدد الأسطر -1) × (عدد الأعمدة -1) = (2-1) × (2-1) = 1

و مستوى المعنوية 0,05% فنجدها تساوي 3,841 .

الفروض الاحصائية : الفرضية الصفرية $H_0: O=E$

الفرضية البديلة $H_1: O \neq E$

القرار الاحصائي : نلاحظ أن قيمة χ^2 الفعلية 16,34 أكبر من قيمة χ^2 النظرية لذلك نرفض نص النظرية الصفرية القائلة بأن سلوك غير المتزوجين في الادخار لا يختلف اختلافا معنويا (جوهريا) عن سلوك المتزوجين و بالتالي نقبل نص الفرضية البديلة (العكسية) القائلة بأن سلوك غير المتزوجين في الادخار يختلف اختلافاً معنوياً (جوهرياً) عن سلوك المتزوجين .

مثال محلول :

أردنا اختبار فعالية أحد الأدوية الجديدة في شفاء الأمراض فأخذنا 200 مريض و قسمناهم إلى مجموعتين الأولى 100 مريض و أعطيناهم الدواء و الثانية 100 مريض و لم نعطهم الدواء (مجموعة المراقبة) و بعد فترة محددة من الزمن سجلنا في جدول خاص عدد الأشخاص اللذين تعافوا من المرض و اللذين لم يتعافوا من المرض من كلتا المجموعتين

العينة	شفوا	لم يشفوا	المجموع
المجموعة الأولى	75	25	100
المجموعة الثانية	65	35	100
المجموع	140	60	200

و **المطلوب** : اختبار الفرض القائل بعجم فعالية الدواء بدرجة ثقة 99% .

الحل :

يجب أن ننشئ جدول للنتائج النظرية للاختبار :

– الأشخاص اللذين شفوا من المجموعة الأولى = $\frac{140\times100}{200}=70$ شخص

– الأشخاص اللذين لم يشفوا من المجموعة الأولى = $\frac{60\times100}{200}=30$ شخص

– الأشخاص اللذين شفوا من المجموعة الثانية = $\frac{140\times100}{200}=70$ شخص

– الأشخاص اللذين لم يشفوا من المجموعة الثانية = $\frac{60\times100}{200}=30$ شخص

ننشئ الجدول :

العينة	شفوا	لم يشفوا	المجموع
المجموعة الأولى	70	30	100
المجموعة الثانية	70	30	100
المجموع	140	60	200

نطبق العلاقة :

$$\chi^2=\sum\frac{(O-E)^2}{E}=\frac{(75-70)^2}{70}+\frac{(25-30)^2}{30}+\frac{(65-70)^2}{70}+\frac{(35-30)^2}{30}=2,38$$

و هي القيمة الفعلية لــ χ^2 .

نستخرج من الجدول الخاص بتوزيع كاي مربع χ^2 القيمة النظرية عند درجات الحرية المساوية 1 و مستوى

المعنوية 1% فنجدها تساوي 6,635

الفروض الإحصائية : الفرضية الصفرية $H_0: O=E$

الفرضية البديلـــة $H_1: O\neq E$

القرار الإحصائي : إن قيمة χ^2 الفعلية أصغر من قيمة χ^2 النظرية لذلك نقبل نص الفرضية الصفرية بأنه لا يوجد تأثير للدواء الجديد على المرض (غير فعال) و بالتالي نرفض الفرضية البديلة القائلة بوجود تأثير للدواء الجديد بفعاليته ، إذن : الدواء الجديد غير فعال في معالجة المرض لأن الأشخاص اللذين شفوا لوحدهم يساوي عدد الأشخاص اللذين شفوا باستخدامه .

بعض الأمثلة غير المحلولة الخاصة باختبار كاي مربع χ^2 :

مثال 1 :

في دراسة قام بها فريق من الباحثين لمعرفة نسبة ولادة الذكور للإناث و سجلت نتائج 400 حالة ولادة في أحد مشافي التوليد فكانت النتيجة على الشكل التالي :

185 ذكور + 215 إناث

و المطلوب :

اختبار الفرض القائل بتساوي ولادة الذكور و الإناث بدرجة ثقة 95% (χ^2 النظرية 3,841) .

مثال 2 :

في دراسة لمعرفة العلاقة بين لون البشرة للإنسان ولون الشعر أخذت عينة مؤلفة من 927 شخص و بعد المعاينة سجلت النتائج التالية :

لون الشعر / لون الجلد	أسود	أشقر	خرنوبي	المجموع
أبيض	63	45	140	248
أسود	32	12	60	104
حنطي	210	170	195	575
المجموع	305	227	395	927

و المطلوب :

اختبار الفرض القائل بعدم وجود ارتباط بين لون البشرة و لون الشعر بدرجة ثقة %95

(قيمة χ^2 النظرية 9,488) .

مثال 3 :

في دراسة لمعرفة الارتباط بين نوعية المقررات التي يدرسها الطالب و النتيجة في الامتحان ، سجلت نتائج الطلاب اللذين نجحوا و اللذين رسبوا في ثلاثة مقررات فحصلنا على النتائج الفعلية التالية :

النتيجة \ اسم المقرر	احصاء حيوي	فسيولوجيا	تدريب عملي	المجموع
نجاح	240	220	345	805
رسوب	60	70	56	186
المجموع	600	290	401	991

و المطلوب :

اختبار الفرض القائل بعدم وجود علاقة بين النتيجة و نوعية المقررات العلمية بدرجة ثقة 99% (χ^2 النظرية = 9,21) .

العمود الأول = الأرقام المسلسلة.

العمود الثاني = التطبيق الأول.

العمود الثالث = التطبيق الثاني.

العمود الرابع = الفرق بين التطبيق الأول والثاني.

العمود الخامس = انحراف الفرق عن المتوسط في العمود الرابع.

العمود السادس = مربع الانحراف للفرق.

الحل:

١-إيجاد الفرق بين التطبيقين.

٢-إيجاد المتوسط الحسابي للفرق وهو $\frac{١٠}{١٠} = ١$

٣-إيجاد الانحراف للفرق عن المتوسط.

٤-إيجاد مربع الانحراف للفرق.

٥-تطبيق المعادلة على النحو التالي .

$$ت = \frac{|م\ ف|}{\sqrt{\frac{ع^2 ف}{ن - ١}}}$$

ولإيجاد $ع^2 ف = \frac{مج\ (ح^2 ف)}{ن}$

صورة أخرى :

$$ت = \frac{|م\ ف|}{\sqrt{\frac{مج\ (ح^2 ف)}{ن\ (ن - ١)}}}$$

وبالرجوع إلى (ت) الجدولية عند درجة حرية (١٥٠) لا نجد قيمتها في الجدول ، ولذا تكشف عن قيمتها مقابل درجة الحرية الأدنى = (١٠٠) ومستوى دلالة (٠٥,٠) ، فنجد أنّ قيمة (ت) =١,٦٦(دلالة الطرف الواحد)، و ١,٩٨(دلالة الطرفين).

أمّا عند مستوى دلالة (٠١,٠) فإنّ قيمة (ت) الجدولية = ٢,٣٦(دلالة الطرف الواحد) و ٦٣ر٢ (دلالة الطرفين).

وبما أن (ت) المحسوبة = ٨,٤٦ ، وهي أكبر من قيمة (ت) الجدولية في جميع الحالات ، فإنّ ثمّة فروقاً دالة إحصائياً بين المجموعتين لصالح المتوسط الأفضل.

٣/٢-دلالة فرق متوسطين مرتبطين لعينة واحدة:

يطبّق هذا القانون في حال جرى اختبار على مجموعة من الأفراد ، ثم أعيد الاختبار على المجموعة نفسها في وقت آخر ، أي أنّ الاختبارين أجريا على العينة نفسها بفاصل زمني .

مثال: أوجد قيمة (ت) لتطبيق اختبارين في مادة (الرياضيات) على عينة مؤلّفة من عشرة طلاّب ، كما هو مبيّن في بيانات الجدول التالي:

م	التطبيق١	التطبيق٢	ف	ح ف	ح^2 ف
١	٨	٦	٢	١	١
٢	٦	٢	٤	٣	٩
٣	٤	٥	-١	-٢	٤
٤	٧	٦	١	٠	٠
٥	٩	٧	٢	١	١
٦	٥	٤	١	٠	٠
٧	٧	٦	١	٠	٠
٨	٧	٧	٠	-١	١
٩	٥	٤	١	٠	٠
١٠	٦	٧	-١	-٢	٤
مج	٦٤	٥٤	١٠		٢٠

$$ت = \frac{١}{\sqrt{\frac{٢}{٩}}} = \frac{١}{٠,٤٧} = ٢,١٣$$

وبالرجوع إلى (ت) الجدولية عند درجة حرية ١٠-١=٩، وعند مستوى دلالة (٠.٥ر.) نجد أن قيمة ت = ١,٨٣دلالة الاتجاه الواحد، و٢,٢٦ دلالـة الاتجـاهين .وعند مستوى دلالة (١.٠ر.) قيمة ت =٢,٨٢، دلالة الطـرف الواحـد ، و٣,٢٥ دلالة الطرفين.

وبما أنّ (ت) المحسوبة = ٢,١٣، فهي دالة فقط عند مستوى (٠,٠٥) ، لدلالة الطرف الواحد فقط.

ملاحظة: يحدّد الاتجاه الواحد أو الاتجاهين عند مستوى دلالة (٠,٠٥) أو (٠,٠١) تبعاً لتحديد فرضيات البحث من البداية.

٤/٢- دلالة فرق متوسطين لعينتين غير متجانستين:

عندما يختلف حجم العينة في المجموعتين ، يختلف أيضاً الانحراف المعياري لكل عينة عن الأخرى اختلافاً كبيراً ، وفي هذه الحال لا يمكن استخدام (ت) كما سـبق أن أوضحنا ولكن يستخدم بدلاً منها صور أخرى.

مثال:

أوجد دلالة فرق متوسطين لعينتين غير متجانسين ، بالنظر لاختلاف حجم العينة ، من خلال البيانات في الجدول التالي :

البيـــــانات	المجمـــــوعة الأولى	المجمـــــوعة الثانية
الوسط الحسابي	١١,٣٥	١٨,٠٧
الوسيط	١١,٠٠	٢٠,٢٣
التباين	١٩,٢٢	٥,٣٢
عدد الأفراد	١٤	٢٥

الحل:

١.حساب التجانس عن طريق النسبة الفائية (ف):

$ف = \frac{١٩,٢٢}{٥,٣٢} = ٣,٦١$ ، وبالرجوع لقيمة(ف)الجدولية عند درجة حرية ٢٥,١٤

ومستوى دلالة (٠٥ ,.) نجد أنها =٢,١١، وعند مستوى دلالة (٠١ ,.)=٢,٨٩

وبما أنّ قيمة (ف) المحسوبة أكبر من الجدولية، فالعينتان غير متجانستين لأن الفرق بين ع٢١،ع٢٢ ، هو فرق معنوي سواء عند مستوى (٠٥ ,.) أوعند مستوى (٠١ ,.)

٢-تطبيق صورة المعادلة التالية .

$$ت = \frac{م١ - م٢}{\sqrt{\frac{ع^{2}١}{ن١} + \frac{ع^{2}٢}{ن٢}}}$$

$$ت = \frac{١٨,٠٧ - ١١,٣٥}{} = \frac{٦,٧٢}{} = \frac{٦,٧٢٢}{} = \frac{٦,٧٢}{}$$

$= ٥,٣٣$

$$\sqrt{\frac{١٩,٢٢}{١٤} + \frac{٥,٣٢}{٢٥}} = \sqrt{١,٣٧ + ,٢١} = \sqrt{١,٥} = ١,٢٦$$

٣-استخراج قيمة (ت) الجدولية لكل من العينة الأولى ، العينة الثانية.

قيمة(ت) الجدولية = ١٦ر٢ العينة الأولى ، عند مستوى دلالة (٠٥ ,.) ودرجة حرية(١٣)

وقيمة(ت) الجدولية = ٢.٠٦ للعينة الثانية عند مستوى دلالة (٠,٠٥) ودرجة حرية(٢٤)

ملحوظة : تطبق الخطوات السابقة نفسيها بدلالة٠١, إذا أراد الباحث ذلك.

٢/٥– مثال توضيحي:

((فاعلية طريقة المناقشة في تدريس مادة الجغرافية))

دراسة تجريبية لطلاب الصف الثاني الثانوي الأدبي في مدارس مدينة دمشق.

أدوات البحث:

استبانة لتعرف آراء الطلبة (ذكور +إناث) نحو طريقة المناقشة في التدريس.

فرضية البحث:

" لا يوجد فرق ذو دلالة إحصائية بين متوسط درجات آراء الذكور ومتوسط آراء الإناث نحو استخدام طريقة المناقشة في التدريس."

التحقق من الفرضية:

لحساب دلالة الفروق بين المتوسطات الحسابية بين المجموعات ولمقارنة فيما بينها، بحسب الجنس، استخدام اختبار ستيودنت (ت) عند مستوى دلالة (٠.٠٥)

والهدف من ذلك هو تحديد ما إذا كان هناك فروق دلالة إحصائية بين المتوسطات الحسابية، فإذا كانت قيمة (ت) المحسوبة أصغر من قيمة (ت) الجدولية، فالفروق ظاهرية يمكن إرجاعها إلى عاملي الحظ والصدفة، وإذا كانت قيمة (ت) المحسوبة أكبر أو تساوي قيمة (ت) الجدولية فالفروق دالة إحصائية "فروق جوهرية حقيقية".

بعد تفرغ إجابات الاستبانة بحسب الجنس (الذكور والإناث) تبيّن ما يلي :

مجموع تكرارات إجابات الذكور=٢٩١٩ ، مجموع ن = ٤٠ ، متوسط تكرار الذكور=٧٢,٤٧٥

مجموع تكرارات إجابات الإناث = ٢٧٢٨، مجموع ن = ٣٥ ، متوسّط تكرار الإناث =٧٧٫٩٤

وحتى نستخدم الصورة المناسبة لاختبار ((ت)) لابد لنا من معرفة تجانس العينتين، ويقاس مدى التجانس باستخدام اختبار ((ف)) بقسمة التباين الأكبر على التباين الأصغر ،أي بالنسبة الفائية ، حيث أنّ :

$$ف = \frac{\text{التباين الأكبر}}{\text{التباين الأصغر}}$$

فإذا كانت قيمة (ف) الجدولية أكبر من قيمة(ف) المحسوبة ،عند درجة حرية ما ، ومستوى دلالة ما ، فعندها نقول إنّ العينتين متجانستين، ونستخدم صورة اختبار (ت) المناسبة لذلك، وهي:

صورة اختبار (ت) لـ:(متوسطين غير مرتبطين لعينتين غير متساويتين و متجانستين).

$$ت = \frac{م١ - م٢}{\sqrt{\frac{ع^2١ \times ن١ + ع^2٢ \times ن٢}{ن١ + ن٢ - ٢}\left[\frac{١}{ن١} + \frac{١}{ن٢}\right]}}$$

حيث أن : ت = قيمة (ت) المحسوبة - (م١) أو (سَ١) = المتوسط الحسابي للمجموعة الأولى.

- (م٢) أو (سَ٢) = المتوسط الحسابي للمجموعة الثانية.

-ع^{2}١= تباين المجموعة الأولى. -ع^{2}٢= تباين المجموعة الثانية.

-ن١= عدد أفراد المجموعة الثانية . -ن٢= عدد أفراد المجموعة الثانية.

وإذا كانت قيمة (ت) المحسوبة أكبر من (ت) الجدولية ، فعندها نقول إنّ العينتين غير متجانستين ،ونستخدم صورة اختبار(ت) المناسبة لذلك، وهي:

صورة اختبار(ت) لـ:((متوسطين غير مرتبطين لعينتين غير متساويتين وغير متجانستين))

-تطبيق صورة المعادلة التالية:

$$\text{ت} = \frac{\text{م}_{١} - \text{م}_{٢}}{\sqrt{\dfrac{\text{ع}^{2}_{١}}{\text{ن}_{١}} + \dfrac{\text{ع}^{2}_{٢}}{\text{ن}_{٢}}}}$$

-ثم تطبيق صورة المعادلة التالية:

$$\text{ت} = \frac{(\text{ت}_{١})\text{الجدولية للعينة الأولى} \times \dfrac{\text{ع}^{2}_{١}}{\text{ن}_{١}} + (\text{ت}_{٢})\text{الجدولية للعينة الثانية} \times \dfrac{\text{ع}^{2}_{٢}}{\text{ن}_{٢}}}{\dfrac{\text{ع}^{2}_{١}}{\text{ن}_{١}} + \dfrac{\text{ع}^{2}_{٢}}{\text{ن}_{٢}}}$$

وحتى نتأكد من تجانس عينتي البحث نطبق اختبار (ف) السابق ذكره ، وهو:

ف = التباين الكبير= ١٤,٤٩ = ٢,٧

التباين الصغير= ٥,٥٩

وبالرجوع لقيمة((ف)) الجدولية عند درجة حريـة(٣٩) كبير،ودرجـة حريـة صغير(٣٤) ، وعند مستوى دلالة(٠,٠٥) نجدها = ١,٨٠

وبما أنّ قيمة (ف) الجدولية أصغر من قيمة (ف) المحسوبة نقول إنّ العينتين غير متجانستين، لذلك نستخدم صورة اختبار((ت)) لـ:

((متوسطين غير مرتبطين لعينتين غير متساويتين وغير متجانستين))

١- تطبيق صورة المعادلة التالية:

$$ت = \frac{|م_١ - م_٢|}{\sqrt{\frac{ع^2_١}{ن_١} + \frac{ع^2_٢}{ن_٢}}}$$

$$ت = \frac{٧٥,٤٧ - ٧٧,٩٤}{\sqrt{\frac{١٤,٤٩}{٤٠} + \frac{٥,٠٩}{٣٥}}}$$

$$= \frac{٢,٤٧}{\sqrt{٠,٥٢}} = ٣,٤٣$$

٢- نستخرج قيمة (ت) الجدولية للعينة الأولى = ٢,٠٢١، عند درجة حرية (٣٩) ومستوى دلالة(٠,٠٥). وقيمة(ت) الجدولية للعينة الثانية=٢,٠٢١ عند مستوى دلالة (٠,٠٥) ودرجة حرية (٣٤)

٣- تطبيق صورة المعادلة التالية:

$$ت = \frac{(ت_١)\text{الجدولية للعينة الأولى} \times \frac{ع^2_١}{ن_١} + (ت_٢)\text{الجدولية للعينة الثانية} \times \frac{ع^2_٢}{ن_٢}}{\frac{ع^2_١}{ن_١} + \frac{ع^2_٢}{ن_٢}}$$

$$\text{فتكون } ت = \frac{٢,٠٢١ \times \frac{١٤,٤٩}{٤٠} + ٢,٠٢١ \times \frac{٥,٠٩}{٣٥}}{\frac{١٤,٤٩}{٤٠} + \frac{٥,٠٩}{٣٥}}$$

$$= \frac{٢,٠٢١ \times ٠,٣٦ + ٢,٠٢١ \times ٠,١٦}{٠,٣٦ + ٠,١٦}$$

$$= \frac{١,٠٥}{٠,٥٢} = ٢,٠١٩$$

والجدول التالي يوضّح معطيات تلك المعادلة :

	حجم العينة	المتوسط الحسابي	التباين	(ت) المحسوبة	(ت) الجدولة
ذكور	ن١=٤٠	م١=٧٥,٤٧	ع١²=١٤,٤٩	٢ر٠١٩	٣,٤٣
إناث	ن٢=٣٥	م٢=٧٧,٩٤	ع٢²=٥,٥٩		

يبين الجدول السابق أنّ قيمة (ت) المحسوبة (٢ر٠١٩) أصغر من قيمة (ت) المجدولة(٣ر٤٣)،عند مستوى دلالة (٠,٠٥)، مما يشير إلى عدم وجود فروق دالة إحصائياً بين آراء الذكور وآراء الإناث بشأن طريقة المناقشة في التدريس ، وهذا يسمح لنا بقبول الفرضية.

ثالثاً– اختبار كاي مربّع (كا٢)

١ – استخدام كاي مربّع :

إنّ الفكرة الأساسية التي يقوم عليها هذا الأسلوب الإحصائي (كاي مربّع) مصوغة على أساس الفرض الصفري ، هي أن التكرار الملاحظ في الفئة أو الفئات موضع الدراسة ، يختلف عن التكرار المتوقع أو الفرضي اختلافاً يرجع إلى الصدفة.

ويتمّ حساب (كاي مربّع)بوساطة المعادلة التالية:

$$\text{كا}^2 = \text{مجموع} \frac{(\text{التكرار الملاحظ} - \text{التكرار المتوقع})^2}{\text{التكرار المتوقع}}$$

٢-نماذج تطبيقيّة لاستخدام كاي مربّع :

٢/١- التوزيع العادل بين القيم :

مثال : لنفرض أنّنا قمنا بتحليل مضمون قصّة للأطفال ، وحصلنا على تكرارت عدد من القيم ، ونريد معرفة مدى عدالة توزيع هذه القيم ، على النحو التالي :

القيمة	الصدق	الأمانة	العدالة	التعاون	محبّةالوطن	الصداقة	المجموع
التكرار	٩	٧	٦	٥	٥	٤	٣٦

لدينا التكرار الملاحظ ، ونريد التكرار المتوقّع ، ونفرضه هنا متوسط التكرارات (٦)

فتكون قيمة كا٢ المحسوبة =

$$\frac{(٩-٦)^{٢}}{٦} + \frac{(٧-٦)^{٢}}{٦} + \frac{(٦-٦)^{٢}}{٦} + \frac{(٥-٦)^{٢}}{٦} + \frac{(٥-٦)^{٢}}{٦} + \frac{(٤-٦)^{٢}}{٦} = ٣ر٦١$$

وبالعودة إلى جداول (كا٢) نجد أنّ قيمة $كا^2$ الجدولية =١١ر٠٧عند درجة حرية (٥) ومستوى دلالة (٠ر٠٥) وهي أكبر من قيمة $كا^2$ المحسوبة ، وبناء" عليه يمكن القول بأنّ الفرق ظاهري بين القيم وهناك توزيع عادل لها.

٢/٢- معرفة مدى التطابق بين آراء عينتين أو متغيرين حول موضوع معيّن :

مثال توضيحي- لو كان لدينا الفرضية الصفرية التالية:

(لا توجد فروق بين آراء مدرسي ومدرسات الجغرافية بشأن التدريس بطريقة المناقشة) وكانت نتائج السؤال كما في الجدول التالي :

جنس المستجيب	الاستجابة للسؤال			
	موافق	لا رأي لي	أرفض	المجموع
ذكر	٤٢	١٣	٣٣	٨٨
أنثى	٢٠	٨	٢٥	٥٣
	٦٢	٢١	٥٨	١٤١

ومن أجل حساب قيمة (كاي مربّع) نقوم أولاً بحساب التكرار المتوقّع لكل تكرار ظاهري على النحو التالي :

- حساب التكرار المتوقع لذكور موافق = ٦٢×٨٨ / ١٤١ = ٣٨,٧٠

~~- حساب التكرار المتوقع لذكور لا أدري = ٢١×٨٨ / ١٤١ = ١٣,١١~~

- حساب التكرار المتوقع لذكور أرفض = ٥٨×٨٨ / ١٤١ = ٣٦,٢٠

- حساب التكرار المتوقع لإناث موافق = ٦٢×٥٣ / ١٤١ = ٢٣,٣٠

- حساب التكرار المتوقع لإناث أدري = ٢١×٥٣ / ١٤١ = ٧,٨٩

- حساب التكرار المتوقع لإناث أرفض = ٥٨×٥٣ / ١٤١ = ٢١,٨

وبعد ذلك نقوم بحساب قيمة كاي مربّع وفق القانون المعتمد ، كما يلي :

كا^2 = (٣٨,٧٠−٤٢)^2 / ٣٨,٧ + (١٣,١١−١٣)^2 / ١٣,١١ + (٢٣,٣٠−٢٠)^2 / ٢٣,٣٠

+ (٣٦,٢٠−٣٣)^2 / ٣٦ر٢ + (٧,٨٩−٨)^2 / ٧,٨٩ + (٢١,٨−٢٥)^2 / ٢١,٨ = ١,٥

وبالعودة إلى جداول (كاي مربّع) نجد أنّ قيمة (كا^2) الجدولية = (١١ر٠٧)
عند درجة حرية (٥) ومستوى دلالة (٠,٠٥)

وبما أنّ قيمة (كا^2) الجدولية أكبر من قيمة (كا^2) المحسوبة نرفض الفرض الصفري، ونقول لا توجد فروق جوهرية بين آراء المدرسين والمدرسات حول طريقة المناقشة في تدريس الجغرافية .

انضموا إلى مجموعة زرقاء اليمامة

على الفيس بوك لمعرفة كل

ماهو جديد

مركز زرقاء اليمامة للخدمات الجامعية