التعليم عن بعد

علم اجتماع (المستوى السابع)

الإحصاء الاجتماعي

د/ علاء أيوب

تـنـسـيق :

أبو فيصل KFU

و حلم المشاعر

المحاضرة الأولى : مقدمة في علم الإحصاء

❖ المتغيرات : Variables

- يقصد بالمتغير : " أي خاصية يمكن قياسها وتتباين قيمها من فرد إلى أخر أو من مجموعة إلى أخرى "، والبيانات الإحصائية التي يتعامل معها الباحث النفسي أو يقوم بجمعها ما هي إلا درجات أو مؤشرات لمقدار الشيء أو الصفة أو الخاصية موضوع القياس لدى الفرد.

➢ **أمثلة :**

– متغير الجنس (ذكر ، أنثى) ، متغير الذكاء ، متغير القلق .

☒ المتغيرات المستقلة والمتغيرات التابعة :

- المتغير المستقل هو : المتغير الذي يخضع للتحكم والسيطرة وبتغير قيمه أو درجاته تتغير تبعاً لذلك قيم المتغير التابع .

– فإذا كان هناك متغيرين بينهما علاقة معينة فيمكن التنبؤ بقيمة أحدهما ويعرف في هذه الحالة بالمتغير التابع إذا علمت قيمة الأخر وهو المتغير المستقل .

➢ **أمثلة :**

– تأثير الذكاء على التحصيل الدراسي .

– أثر التدريس باستخدام الفصول الافتراضية على تحصيل الطلاب في مقرر الاحصاء الاجتماعي .

☒ متغيرات مستقلة ومتغيرات مترابطة :

- عندما يكون لدينا مجموعة من القياسات التي ترتبط أو تؤثر في بعضها البعض يقال للمتغيرات في هذه الحالة متغيرات مرتبطة أما إذا كانت القياسات غير مترابطة ولا تؤثر في بعضها البعض فإن المتغيرات في هذه الحالة تكون متغيرات مستقلة .

➢ **أمثلة :**

– إذا أردنا معرفة تأثير الذكاء على التحصيل فيمكن اعتبار الدرجات التي يحصل عليها الأفراد مستقلة ما دامت درجة الفرد لا ترتبط بدرجة غيره من الأفراد

– إذا أردنا معرفة الاختلاف بين تقدير الأم وتقدير الأب للعدوانية عن أطفالهم ، فهنا يكون لكل طفل درجتين في العدوانية إحداهما تقدير الأب والأخرى تقدير الأم وهنا يقال أن الدرجات مترابطة .

❖ طبيعة البيانات :

☒ البيانات الكيفية (النوعية) :

- هي المشاهدات أو الصفات التي لا يمكن قياسها مباشرة بوسائل القياس المألوفة كالعد والقياس (تكون في صورة غير عددية).

➢ **أمثلة :** لون العين (أسود ، أخضر ، عسلي ، أزرق) ، الجنس (ذكر ، أنثى) ، تقديرات الطلاب (ممتاز ، جيد جداً ، جيد ، مقبول) ، الجنسية (مصري ، سعودي ، ألماني) .

☒ البيانات الكمية (العددية) :

- هي المشاهدات أو الصفات التي يمكن قياسها مباشرة بوسائل القياس المألوفة (تكون في صورة عددية) .

➢ **أمثلة :** عدد طلاب التعليم الالكتروني ، الطول ، الوزن ، عدد أفراد الأسرة .

❖ أنواع البيانات الكمية :

☒ البيانات المنفصلة :

- هي المشاهدات أو الصفات التي تأخذ قيماً متمايزة عن بعضها ، مما يعني عدم اتصال البيانات ، ولا تتضمن كسوراً .

➢ **أمثلة :** عدد الطلاب الموزعين في كل تخصص أو شعبة أو فصل من فصول مدرسة .

☒ البيانات المتصلة :

- هي المشاهدات أو الصفات التي تأخذ مدى معين أو مجال معين من القيم ويمكن توزيعها على خط متصل بدون فواصل بينها لأنها تتضمن كسوراً .

➢ **أمثلة :** الطول ، والوزن .

❖ أساليب إجراء البحث :

☒ أسلوب الحصر الشامل :

- يتم فيه جمع البيانات عن الظاهرة موضوع الدراسة من جميع مفردات المجتمع الإحصائي المراد بحثه سواء أكان نطاقه أو مجاله واسعاً أو محدوداً.

مزايا أسلوب الحصر الشامل :
• خال من أخطاء الصدفة (الأخطاء العشوائية أو أخطاء المعاينة) .
• يعطي صورة مفصله عن مفردات الظاهرة موضوع الدراسة .

عيوب أسلوب الحصر الشامل :
• الزيادة الكبيرة في التكاليف المادية والبشرية والزمنية.
• طول الوقت اللازم لجمع البيانات يفقد نتائج البحث حداثتها وبالتالي قيمتها.
• وجود مجتمعات بطبيعتها غير محدودة وبالتالي يتعذر تحديد إطار مفرداتها.

☒ أسلوب العينات :

- يتم فيه جمع البيانات عن جزء فقط من مفردات المجتمع الإحصائي، ويتم سحب العينة بطريقة ما يساعد في تعميم نتائجها على مجتمع البحث.

مزايا أسلوب العينات :
• يوفر التكاليف المادية والبشرية والزمنية لإجراء الدراسة.
• زيادة الرقابة والضبط والتحكم في معظم الأسباب المؤدية إلى الأخطاء.
• يصلح للمجتمعات غير المحدودة.

عيوب أسلوب العينات :
• يتعرض أسلوب المعاينة إلى نوع آخر من الأخطاء ينفرد به هذا الأسلوب ويطلق عليه خطأ المعاينة أو خطأ الصدفة وخطأ التحيز.

❖ المجتمع والعينة :

☒ المجتمع:

• يعرف المجتمع بأنه مجموعة من العناصر، أو المفردات التي تخص ظاهرة معينة محل الدراسة . وهو مصطلح علمي يراد به كل من يمكن أن تعمم عليه نتائج البحث .

☒ العينة:

• تعرف العينة بأنها جزء من مفردات المجتمع الإحصائي يتم اختياره بطريقة علمية ، ثم دراسة خصائص هذا الجزء لغرض التعرف على خصائص المجتمع الذي اختير منه ذلك الجزء .

• ولكي تكون العينة مقبولة من الناحية الإحصائية ينبغي أن تكون عينة ممثلة للمجتمع ، أي أنها تحتوي على جميع الخصائص بنفس تواجدها في المجتمع الإحصائي الذي اختيرت منه.

❖ البارامترات (المعلمات) والإحصاءات :

• للمجتمع خصائص متعددة مثل المتوسط والوسيط والانحراف المعياري وكذلك لكل عينة تسحب من هذا المجتمع خصائصها أيضاً وما يتعلق بخصائص المجتمع يسمى مَعلَّم أو بارامتر Parameter بينما كل ما يتعلق بخصائص العينات يسمى إحصاءه Statistic ويمكن الاستفادة من إحصاءات العينة تقدير معلمات المجتمع.

❖ الإحصاء الوصفي والإحصاء الاستدلالي :

☒ الإحـــصــاء الوصفي :	• يقتصر على الوصف الكمي للظواهر وتصنيفها وتحليلها وعلاقتها بغيرها من الظواهر.
☒ الإحصاء الاستدلالي :	• يتعدى ذلك مستفيداً من نتائج الإحصاء الوصفي في الاستدلال على خصائص المجتمع العام للظاهرة فهو يهدف إلى تقدير خصائص المجتمع استنادا إلى نتائج دراسة عينة منتقاة من هذا المجتمع.

❖ الإحصاء البارامتري والإحصاء اللابارامتري :

☒ الأساليب البارامترية (المعلمية) :

• هي الأساليب التي تتطلب استيفاء افتراضات معينة حول المجتمع الذي تسحب منه عينة البحث ومن هذه الافتراضات أن يكون التوزيع طبيعياً وأن يكون هناك تجانس في التباين. والأساليب البارامترية تصلح للبيانات في المستوى الفتري والمستوى النسبي .

☒ الأساليب اللابارامترية (اللامعلمية) :

• هي الأساليب التي تستخدم في الحالات التي لا يكون فيها نوع التوزيع الاحتمالي للأصل الذي سحبت منه العينة معروفاً أو في حالة عدم استيفاء شرط التوزيع الاعتدالي للمجتمع. والأساليب الإحصائية اللابارامترية تصلح في حالة البيانات الرتبية والاسمية

طرق عرض البيانات

أولاً : العرض الجدولي للبيانات :

➢ **ويقصد بالعرض الجدولي للبيانات :**

– أن يتم تلخيص البيانات محل الدراسة وتصنيفها فى صورة جداول تعبر عن القيم التي أخذها المتغير من خلال البيانات التي جمعها و كرار كل قيمة من تلك القيم .

أهمية الجداول الاحصائية :
• تعبر عن الحقائق الكمية المعروضة بعدد كبير من الارقام في جداول بطريقة منظمة .
• تلخيص المعلومات الرقمية الكثيرة العدد، المتغيرة القيم ، مما يسهل التعرف عليها .
• الاستيعاب وبسهولة عدد كبير من الموضوعات .
• اظهار البيانات بأكبر وضوح ممكن وأصغر حيز مستطاع .

تتكون اجزاء الجدول مما يلي :	
رقـم الجـدول :	• يجب ان يرقم كل جدول حتى تسهل الاشارة اليه .
الـعـنـوان :	• يجب أن يعطي كل جدول عنوانا كاملا لتسهيل مهمة استخراج المعلومات منه، ويجب أن يكون هذا العنوان واضحا قصيرا بقدر الامكان، ويستخدم في بعض الاحيان عنوان توضيحي لبعض الجداول وذلك من أجل إعطاء معلومات إضافية عن بيانات الجدول .
الهيكل الرئيسي :	• ويتكون الجدول من أعمدة وصفوف، ويعتبر ترتيب المعلومات في الاعمدة والصفوف أهم خطوة في تكوين الجدول .
الـعـمـود :	• كل جدول يتكون من عمود أو اكثر ويوجد لكل عمود عنوان يوضح محتوياته .
الـحـواشي :	• قد يحتوي الجدول على مفردات بيانات لا ينطبق عليها عنوان الجدول أو عنوان العمود، ففي هذه الحالة تستعمل الحواشي لتوضيح ذلك وذلك اما بترقيم الملاحظات او باستعمال علامة (*) .. الخ .
المـصـدر :	• قد تؤخذ بيانات الجدول من مصادر جاهزة لذلك يجب إظهار المصدر في أسفل الجدول حتى يمكن الرجوع اليه عند الحاجة .

❖ أنواع الجداول الاحصائية :

تقسم الجداول تبعاً لدرجة تعقيدها إلى :	
جداول بسيطة :	• وفيها يتكون كل من موضوع الجدول ومادته من بضع أسطر وخانات تتعلق بالتقسيمات الزمانية (أي الأمور التي يتناولها الجدول أمور تتسلسل حسب السنوات) أو المكانية (أي توزيع الظاهرة حسب المكان) أو مؤشرات وصفية بسيطة وبأرقام بسيطة أيضا.
جداول التوزيع التكراري :	• وفيها تكون المعطيات مجمعة في فئات بمؤشر أو متغير واحد، ولكل فئة تكراراتها الخاصة عند ذلك المؤشر
جدول التوزيع التكراري المتجمع :	• وفيه تجمع التكرارات على التوالي من أحد طرفي الجدول الى طرفة الآخر فنحصل على التكرار الكلي (مجموعة التكرارات)، (فاذا بدأ من أعلى الى أسفل الجدول) سمي جدول تكراري متجمع صاعد، (واذا بدأ من أسفل الى أعلى الجدول) سمي جدول تكرار متجمع نازل أو هابط.
الجداول المزدوجة أو المركبة :	• وهي الجداول التي تتكون من متغيرين أو أكثر، وهذه المتغيرات قد توزع على أعمدة وحقول الجدول بصورة نظامية، تعبر عن الافكار العلمية التي يريد الباحث توضيحها توضيحا عدديا.

➢ **وهناك عدة ملاحظات يجب الانتباه إليها عند عمل جدول التوزيع التكراري لبيانات المتغير الكمي المتصل :**

١) إن تحديد عدد الفئات يتوقف على أمور عدة منها:
- عدد المفردات محل الدراسة
- انتظام وتوزيع تلك البيانات
- طبيعة بيانات المشكلة محل الدراسة

٢) طول الفئة لا بد أيضاً من تحديده بعناية حيث يمثل الوجه الآخر للعملة مع عدد الفئات، فمن الأفضل أن يكون تحديده بطريقة تجعل مركز الفئة قريباً من تركز البيانات بتلك الفئة بقدر الإمكان حيث يعبر مركز الفئة عن قيمة كل مفردة من المفردات التي تنتمي لتلك الفئة.

٣) أن تكون حدود الفئات واضحة بحيث لا يكون هناك أي تداخل فيما بينها.

➢ **ومن هنا يمكن إعداد جداول التوزيعات التكرارية للمتغيرات المتصلة بثلاث صور هي :**
- الجداول التكرارية المنتظمة .
- الجداول التكرارية غير المنتظمة .
- الجداول التكرارية المفتوحة .

لا بد من حضور المحاضرة المسجلة لفهم الشرح .

المحاضرة الثانية : تبويب وعرض البيانات الإحصائية

أولاً : العرض الجدولي للبيانات:

١) الجدول التكراري البسيط	٢) الجدول التكراري ذو الفئات	٣) الجدول التكراري المتجمع الصاعد
٤) الجدول التكراري المتجمع الهابط	٥) الجدول المزدوج	

١) تبويب البيانات في جدول تكراري بسيط:

مثـال :

البيانات التالية هى درجات حصل عليها عشرون طالباً فـى مـادة الإحصاء الاجتماعي بالفرقة الأولى قسم الاجتمـاع فـى امتحـان نهاية العام :

10	12	13	15	10	12	14	15	11	12
15	12	10	12	13	15	12	13	10	14

والمطلوب تبويب هذه البيانات فى جدول توزيع تكرارى بسيط ؟

الدرجة	العلامات	التكرار
١٠	////	٤
١١	/	١
١٢	///// /	٦
١٣	///	٣
١٤	//	٢
١٥	////	٤
المجموع		٢٠

التقدير	التكرار
مقبول	5
جيد	9
جيد جداً	3
ممتاز	3
المجموع	20

٢) تبويب البيانات في جدول تكراري ذو فئات:

– المقصود بالفئات :

78	60	74	61	54	73	73	68	64	67
65	77	66	72	69	60	63	65	74	80
55	70	63	78	66	68	69	76	50	74
77	65	57	72	72	61	76	64	67	67
67	73	66	74	63	58	78	79	71	59
55	84	61	64	81	73	68	69	71	61

➢ **أربعة طرق لكتابة الفئات :**

١

ف	ك
20-10	5
30-20	20
40-30	50
50-40	25

٢

ف	ك
19-10	5
29-20	20
39-30	50
49-40	25

٣

ف	ك
-10	5
-20	20
-30	50
-40	25

٤

ف	ك
20-	5
30-	20
40-	50
50-	25

مـثـال :

قام باحث بجمع بيانات تمثل درجات اختبار مادة الحاسب الآلي لخمسين طالباً من طلاب المرحلة الثانية من الثانوية العامة في الجدول التالي :

70	55	51	42	57
45	60	47	63	53
33	65	39	82	55
64	58	61	65	42
50	52	53	45	55
25	36	59	63	39
65	45	49	54	64
75	42	41	52	78
30	35	25	48	26
20	40	55	46	88

والمطلوب هو إعداد جدول توزيع تكراري ذو فئات للجدول السابق؟

✓ الحل :

➢ **حساب المدى** = أكبر قيمة – أصغر قيمة

= ٨٨ – ٢٠ = ٦٨

➢ **عدد الفئات** = ١ + ٣,٣ لو (ن)

= ١ + ٣,٣ × ١,٦٩٩ = ٦,٦١ .... ٧

➢ **طول الفئة** = المدى / عدد الفئات

= ٦٨/ ٧ = ٩,٧١ .... ١٠

➢ **بداية الفئة الأولى هو :** الحد الأدنى للدرجات (٢٠)

الفئات	العلامات	التكرار
-20	~~////~~	4
-30	~~////~~ /	6
-40	~~////~~ ~~////~~ //	12
-50	~~////~~ ~~////~~ ////	14
-60	~~////~~ ////	9
-70	///	3
90-80	//	2
	المجموع	50

٣) تبويب البيانات في الجدول التكراري المتجمع الصاعد :

- يقصد بالتكرار المتجمع الصاعد هو : تجميع تكرار كل فئة على جميع التكرارات السابقة لها بحيث يكون مجموع التكرار التصاعدي للفئة الأخيرة مساوي لمجموع التكرارات

الفئات	العلامات	التكرار
20-	////	4
30-	~~////~~ /	6
40-	~~////~~ ~~////~~ //	12
50-	~~////~~ ~~////~~ ////	14
60-	~~////~~ ////	9
70-	///	3
80-90	//	2
المجموع		50

حدود الفئات	التكرار المتجمع الصاعد
أقل من ٢٠	صفر
أقل من ٣٠	٤
أقل من ٤٠	١٠
أقل من ٥٠	٢٢
أقل من ٦٠	٣٦
أقل من ٧٠	٤٥
أقل من ٨٠	٤٨
أقل من ٩٠	٥٠

٤) تبويب البيانات في الجدول التكراري المتجمع الهابط :

- يقصد بالتكرار المتجمع الهابط هو : تجميع تكرار كل فئة على جميع التكرارات التالية لها بحيث يكون مجموع التكرار التنازلي للفئة الأولى مساوي لمجموع التكرارات

الفئات	العلامات	التكرار
20-	////	4
30-	~~////~~ /	6
40-	~~////~~ ~~////~~ //	12
50-	~~////~~ ~~////~~ ////	14
60-	~~////~~ ////	9
70-	///	3
80-90	//	2
المجموع		50

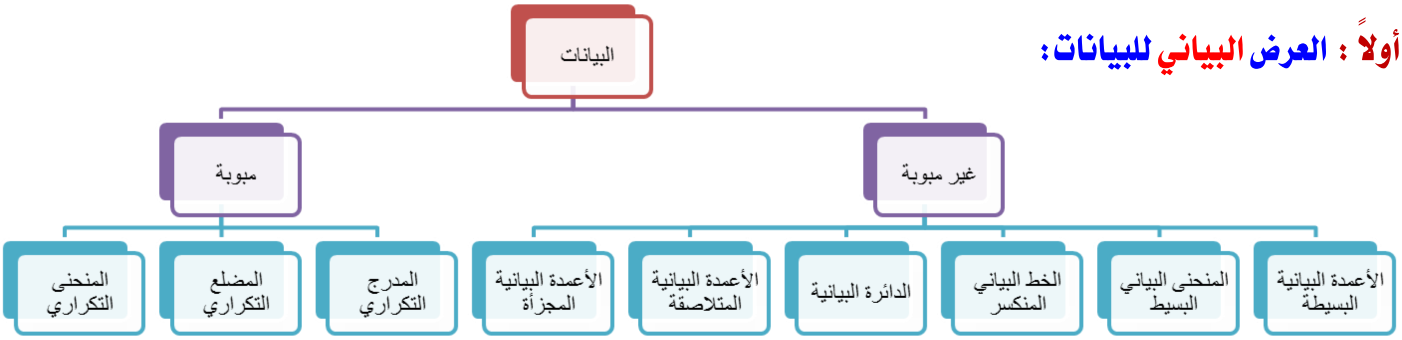
حدود الفئات	التكرار المتجمع الهابط
٢٠ فأكثر	٥٠
٣٠ فأكثر	٤٦
٤٠ فأكثر	٤٠
٥٠ فأكثر	٢٨
٦٠ فأكثر	١٤
٧٠ فأكثر	٥
٨٠ فأكثر	٢
٩٠ فأكثر	صفر

٥) تبويب البيانات في الجدول المزدوج :

التوزيع المشترك بين النوع وحضور المحاضرات :

	النوع		المجموع
	إناث	ذكور	
عدم الحضور	٥٤	١١٧	١٧١
الحضور	٣٤٨	٩٥٠	١٢٩٨
المجموع	٤٠٢	١٠٦٧	١٤٦٩

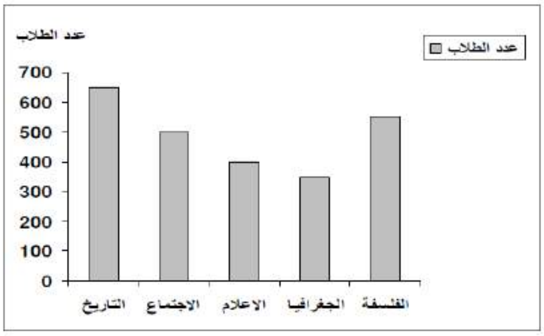
أولاً : العرض البياني للبيانات :



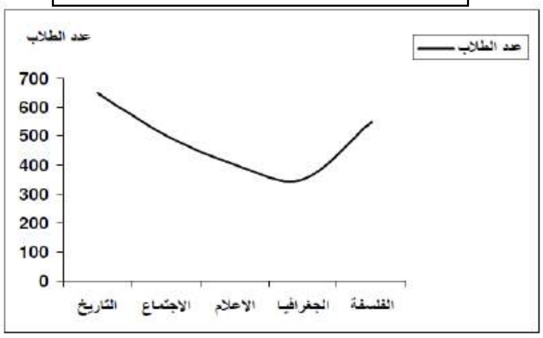
١) الأعمدة البيانية البسيطة :

الجدول التالى يوضح أعداد الطلاب ببعض أقسام كلية الآداب جامعة المنصورة والمطلوب عرض هذه البيانات باستخدام طريقة الأعمدة البيانية البسيطة ؟

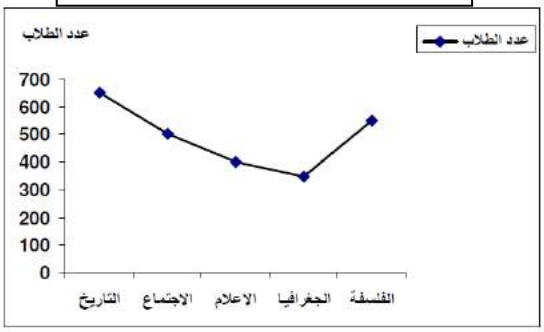
القسم	التاريخ	الاجتماع	الإعلام	الجغرافيا	الفلسفة
عدد الطلاب	650	500	400	350	550



٢) المنحنى البياني البسيط :



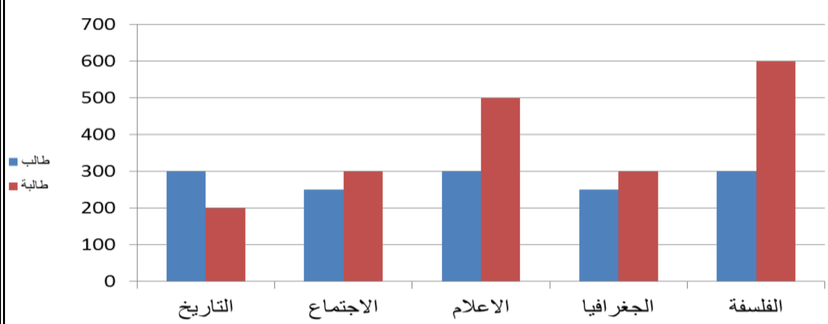
٣) الخط البياني المنكسر :



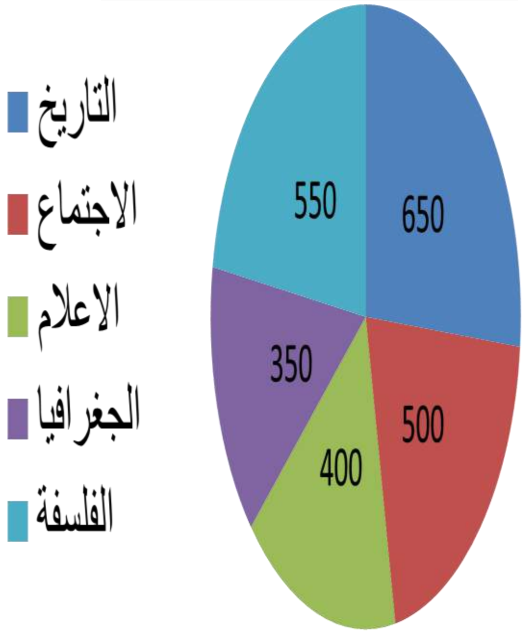
٤) الأعمدة البيانية المتلاصقة :

- الجدول التالي يوضح أعداد الطلاب ببعض أقسام كلية الآداب بجامعة الملك فيصل والمطلوب عرض هذه البيانات باستخدام طريقة الأعمدة المتلاصقة.

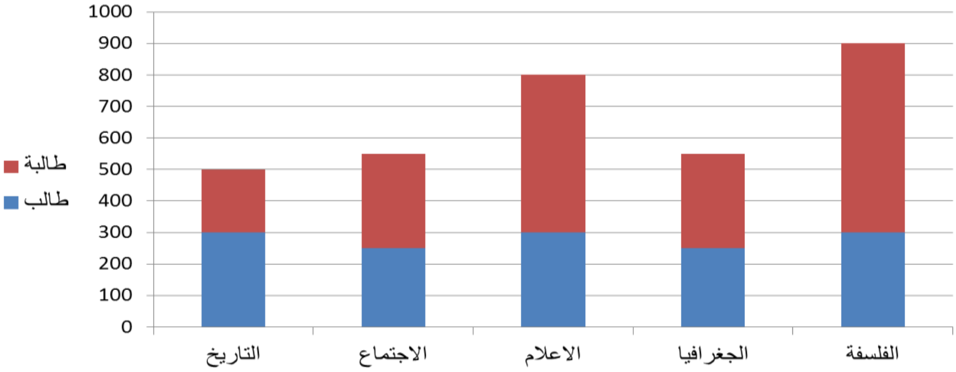
القسم	التاريخ	الاجتماع	الاعلام	الجغرافيا	الفلسفة
طالب	٣٠٠	٢٥٠	٣٠٠	٢٥٠	٣٠٠
طالبة	٢٠٠	٣٠٠	٥٠٠	٣٠٠	٦٠٠



٥) الدائرة البيانية :



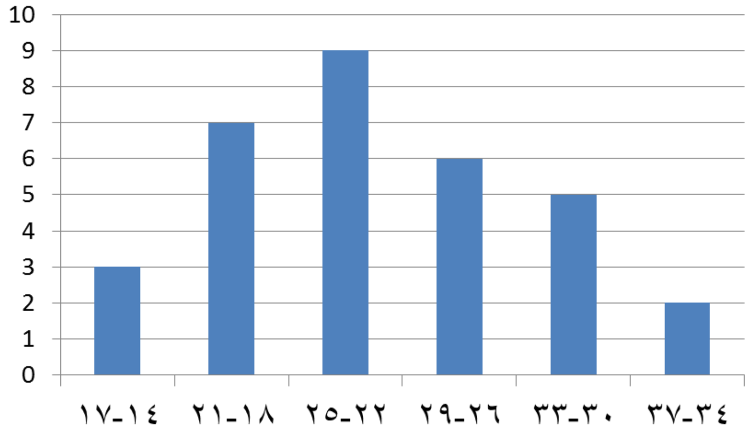
٦) الأعمدة البيانية المجزأة:



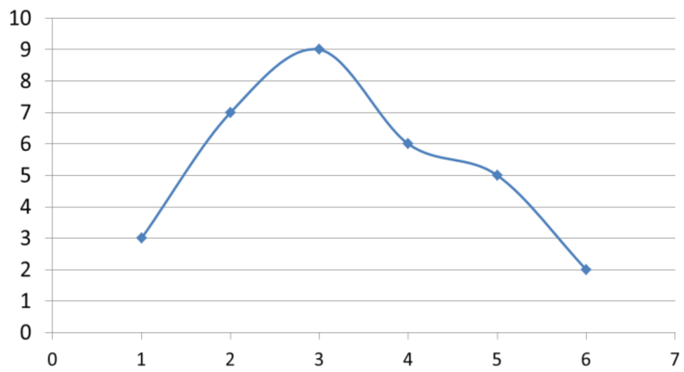
مبوبة

١) المدرج التكراري:

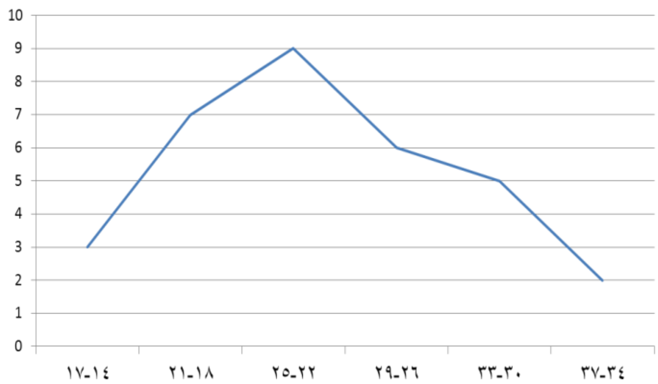
التكرارات	الفئات
٣	١٤-١٧
٧	١٨-٢١
٩	٢٢-٢٥
٦	٢٦-٢٩
٥	٣٠-٣٣
٢	٣٤-٣٧
٣٢	المجموع



٢) المنحنى التكراري:



٣) المضلع التكراري:



تمارين :

1- حصل عدد من الطلاب فى مادة الإحصاء على الدرجات التالية :

5	4	4	5	3	4	2	3	1	2
3	7	4	1	6	3	2	5	3	4
7	3	2	6	5	3	4	2	4	1

المطلوب : تكوين جدول تكرارى بسيط لهذه الدرجات.

2- تمثل البيانات التالية تقديرات عشرون طالبا فى مادة علم النفس والمطلوب وضعها فى جدول تكرارى بسيط لتلك التقديرات .

ممتاز	مقبول	جيد جدا	مقبول	جيد
جيد جدا	جيد	ضعيف	جيد	مقبول
جيد	ممتاز	مقبول	ضعيف	جيد
جيد جدا	جيد	مقبول	جيد	مقبول

المحاضرة الثالثة : مقاييس النزعة المركزية (المتوسط الحسابي، الوسيط، المنوال)

- كل ظاهرة في الحياة العامة لها ميل للتجمع حول نقطة معينة ؛ ومن ثم إذا استطعنا تحديد هذه النقطة فإننا سنصل إلى قيمة متوسطة تتجمع حولها القيم.
- **الميل إلى التجمع حول هذه يسمى القيمة بالنزعة المركزية وتسمى المقاييس المستخدمة مقاييس النزعة المركزية .**

مقاييس النزعة المركزية Measures Of Central Tendency

- المتوسط الحسابي Arithmetic Mean
- الوسيط Median
- المنوال Mode
- الوسط الهندسي GeometricMean
- الوسط التوافقي Harmonic Mean

يسمى الوسط الحسابي، والوسيط، والمنوال، مقاييس النزعة المركزية لأن كلا منها يحاول أن يصف نقطة تجمع مشاهدات التوزيع

أهمية مقاييس النزعة المركزية :

عند معرفتنا بتلك المتوسطات (مقاييس النزعة المركزية) يصبح أمامنا فرصة كبيرة لأن :

- ننظر لمتوسط مجموعة من البيانات لنعرف الكثير عن خصائص تلك المجموعة .
- نعقد مقارنة بين عدة مجموعات من البيانات في وقت واحد وذلك من خلال مقارنة متوسطات تلك المجموعات بعضها ببعض .

الوسط الحسابي Arithmetic Mean :

يعد من أكثر المقاييس المستخدمة في الاحصاء حيث انه بسيط وسهل الفهم و يصلح للمقارنة بين المجموعات.

إذا كانت قيم المتغير (x) هى $x_1, x_2, \dots, x_n$ حيث (n) يمثل حجم المجموعة ؛ فإن الوسط الحسابي يمكن التعبير عنه على النحو التالي :

$$\bar{x} = \frac{x_1 + x_2 + \dots + x_n}{n} = \frac{\sum x}{n}$$

الوسط الحسابي = مجموع قيم البيانات / عددها أي $\bar{x} = \frac{\sum x}{n}$

س ١ : درجات خمسة طلاب في مقرر ما [الدرجة العظمى 20] هي : 10 , 12 , 7 , 2 , 9 . أوجد الوسط الحسابي لدرجاتهم .

ج١ : $\bar{x} = \frac{\sum x}{n} = \frac{9+2+7+12+10}{5} = \frac{40}{5} = 8$

من هذا المثال البسيط يمكن ملاحظة الخصائص العامة التالية للوسط الحسابي :

١. يمكن تحديد قيمة الوسط الحسابي بالضبط، كما أن طريقة تحديده سهلة	٢. يأخذ في الاعتبار جميع البيانات
٣. لا يتأثر بترتيب البيانات	٤. لا يُشترط أن يكون الوسط الحسابي عدداً صحيحاً ولا يُشترط أن يكون إحدى قيم البيانات ولكنه قيمة تقع بين أقل قيمة في البيانات وأكبر قيمة فيها
٥. يتأثر بالقيم المتطرفة في البيانات [كما يتضح من السؤالين التاليين]	
س ٢ : احسب الوسط الحسابي للقيم : 40, 50, 45, 55, 35	**ج٢ :** $\frac{40+50+45+55+35}{5} = \frac{225}{5} = 45$
س ٣ : احسب الوسط الحسابي للقيم : 10 , 15 , 12 , 13 , 95	**ج٣ :** $\frac{40+50+45+55+95}{5} = \frac{285}{5} = 57$
لاختلاف هذا السؤال بالمحتوى والمسجلة تم مراسلة الدكتور و حل هذا السؤال بهذه الطريقة ..	

حاصل ضرب قيمة الوسط الحسابي في عدد البيانات = مجموع قيم البيانات

وهذا واضح من تعريف الوسط الحسابي : $(n \times \bar{x}) = \sum x$ تعني أن $\bar{x} = \frac{\sum x}{n}$

فمثلاً :

١٠	١٢	٧	٢	٩

خمس درجات وسطها الحسابي ٨

$\underbrace{٨ \times ٥}_{=٤٠} = \underbrace{١٠ + ١٢ + ٧ + ٢ + ٩}_{=٤٠}$

✓ **إذا أضفنا عدد ثابت c لكل قيمة من قيم البيانات فإن :**

الوسط الحسابي الجديد = الوسط الحسابي القديم + العدد الثابتc

بيانات جديدة: 11 ، 13 ، 8 ، 3 ، 10 — خمس درجات وسطها الحسابي 9 [8+1 =]

إضافة 1 لكل قيمة

بيانات قديمة: 10 ، 12 ، 7 ، 2 ، 9 — خمس درجات وسطها الحسابي 8

✓ **إذا ضربنا كل قيمة من قيم البيانات في عدد ثابت c ، فإن :**

الوسط الحسابي الجديد = الوسط الحسابي القديم x العدد الثابتc

بيانات جديدة: 11 ، 13.2 ، 7.7 ، 2.2 ، 9.9 — خمس درجات وسطها الحسابي 8.8 [8 x 1.1 =]

ضرب كل قيمة في 1.1

بيانات قديمة: 10 ، 12 ، 7 ، 2 ، 9 — خمس درجات وسطها الحسابي 8

اعتبر نفسك مدرساً للطلاب الخمسة المذكورين في س ١ [كانت درجاتهم (من 20) كالتالي : 10 , 12 , 7 , 2 , 9] وأردت أن تحسن من الوسط الحسابي لدرجاتهم ، أيهما أفضل : أن نزيد درجة كل طالب 5 درجات أم نزيد درجة كل طالب 50% من قيمتها ؟ علل إجابتك .

حساب الوسط الحسابي لبيانات غير مبوبة

س : أوجد الوسط الحسابي للأرقام : 8 , 8 , 8 , 2 , 2 , 4 , 4 , 4 , 4 , 4 , 6 , 6 , 3 , 3 , 5 , 5 , 5 , 5 , 5 , 5

ج : بتطبيق مباشر لتعريف . . .

$$\bar{x} = \frac{(5+5+5+5+5+5)+(3+3)+(6+6)+(4+4+4+4+4)+(2+2)+(8+8+8)}{20} = \frac{96}{20} = 4.8$$

لاحظ أن الرقم 5 متكرر 6 مرات ، الرقم 3 مرتان ، والرقم 6 مرتان ، والرقم 4 متكرر 5 مرات ، والرقم 2 مرتان ، والرقم 8 ثلاث مرات ، وبالتالي يمكن عمل العملية الحسابية السابقة كالآتي :

$$\bar{x} = \frac{(5\times 6)+(3\times 2)+(6\times 2)+(4\times 5)+(2\times 2)+(8\times 3)}{6+2+2+5+2+3} = \frac{30+6+12+20+4+24}{20} = \frac{96}{20} = 4.8$$

وهذا يمكن إنجازه بيسر من خلال الجدول التكراري للبيانات كالتالي :

$$\bar{x} = \frac{\sum fx}{\sum f} = \frac{96}{20} = 4.8$$

المتغير x	التكرار f	fx
5	6	30
3	2	6
6	2	12
4	5	20
2	2	4
8	3	24
	20	96
	$\sum f = 20$	$\sum fx = 96$

أي أنه في حالة البيانات الكمية المتقطعة ذات التكرارات يمكن حساب الوسط الحسابي من العلاقة :

$$\bar{x} = \frac{\sum fx}{\sum f}$$

حيث $\sum f$ **هو مجموع التكرارات**

$\sum fx$ **هو مجموع حاصل ضرب كل قيمة في تكرارها**

س : من مائة رقم يتكرر الرقم 4 عشرون مرة، والرقم 5 أربعون مرة، والرقم 6 ثلاثون مرة، والباقي كانوا الرقم 7 احسب الوسط الحسابي للمائة رقم .

ج : بتكوين الجدول التكراري للأرقام المذكورة ، ثم بضرب كل قيمة في تكرارها والتجميع [عمود fx] يكون الوسط الحسابي للأرقام المذكورة هو :

الجدول التكراري

المتغير x	التكرار f	fx
4	20	80
5	40	200
6	30	180
7	10	70
	100	530
	$\sum f=100$	$\sum fx=530$

$$\bar{x}=\frac{\sum fx}{\sum f}=\frac{530}{100}=5.3$$

حساب الوسط الحسابي لبيانات مبوبة

عندما نتعامل مع بيانات متصلة تُعطى فيها قيم المتغير على صورة فترات، فيمكن اعتبار أن جميع القيم داخل الفترة مطابقة لمركز الفئة ، وبالتالي يمكن استخدام الصيغة السابقة لحساب الوسط الحسابي :

$$\bar{x}=\frac{\sum f\,x_0}{\sum f}$$

حيث $\sum f$ هو مجموع التكرارات ، $\sum f x_0$ هو مجموع حاصل ضرب مركز كل فئة في تكرار الفئة

ففي المثال التالي والذي يوضح اطوال سيقان الزهار بالسنتيمتر، يكون الوسط الحسابي لأطوال سيقان الأزهار هو:

الجدول التكراري

الفئة	المتغير x (الطول)	التكرار f	مركز الفئة x_0	fx_0
الأولى	$0\le x<20$	4	10	40
الثانية	$20\le x<30$	16	25	400
الثالثة	$30\le x<35$	12	32.5	390
الرابعة	$35\le x<40$	10	37.5	375
الخامسة	$40\le x<50$	6	45	270
السادسة	$50\le x<60$	2	55	110
		$\sum f=50$		$\sum fx_0=1585$

$$\bar{x}=\frac{\sum f\,x_0}{\sum f}=\frac{1585}{50}=31.7$$

مزايا وعيوب الوسط الحسابي

من كل ما سبق يمكن استعراض مزايا وعيوب الوسط الحسابي كالتالي :

– يمكن تحديد قيمة الوسط الحسابي بالضبط، كما أن طريقة تحديده سهلة [ميزة] .	
– يأخذ في الاعتبار جميع البيانات [ميزة] .	– لا يتأثر بترتيب البيانات [ميزة] .
– يتأثر بالقيم المتطرفة في البيانات [عيب] .	– لا يمكن حسابه بالرسم ، أي بيانياً [عيب] .

الوسيط Median : استخدام الوسيط في حالة التعامل مع

– البيانات التي تكثر بها القيم الشاذة .

– الجداول التكرارية المفتوحة من أحد طرفيها أو من كليهما .

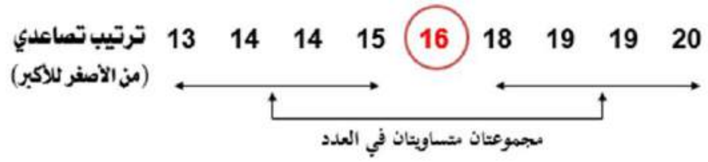
– التوزيعات التكرارية غير المتساوية في طول الفئات .

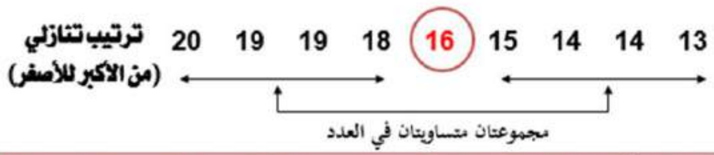
تعريف الوسيط :

(ببساطة) يُعرف الوسيط [ونرمز له بالرمز M] لمجموعة من القيم **(المرتبة تصاعدياً أو تنازلياً حسب قيمها)** على أنه القيمة التي تقسم مجموعة القيم إلى مجموعتين متساويتين في العدد ، أو بتعبير آخر هي القيمة التي في المنتصف .

فمثلاً لمجموعة القيم : 13 , 14 , 19 , 16 , 20 , 15 , 18 , 14 , 19 [عددها ٩ قيم **أي رقم فردي**] ، إذا قمنا بترتيبها تصاعدياً أو تنازلياً

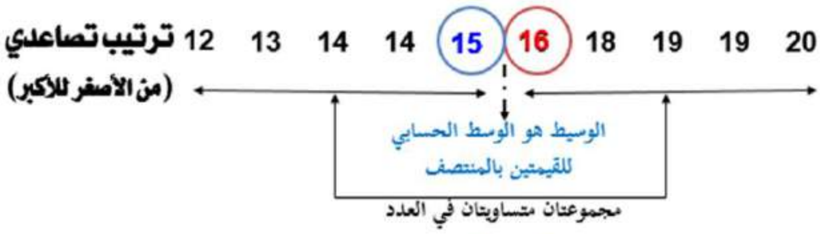
يكون الوسيط هو العدد الخامس [رتبة الوسيط أي ترتيبه بين القيم] وقيمته 16 **هام جداً : فرق بين رتبة الوسيط وقيمته**

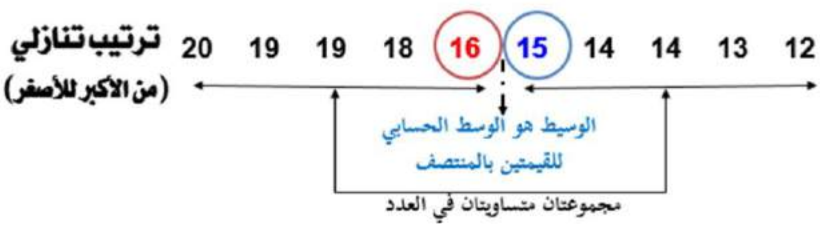




لاحظ هنا أن عدد القيم n [هنا = ٩] فردي وبالتالي هناك قيمة واحدة في منتصف المجموعة

أما لمجموعة القيم : 12 , 13 , 14 , 19 , 16 , 20 , 15 , 18 , 14 , 19 [عددها ١٠ قيم (**أي رقم زوجي**)] حيث أضفنا القيمة 12 للمجموعة السابقة]، إذا قمنا بترتيبها تصاعدياً أو تنازلياً ..





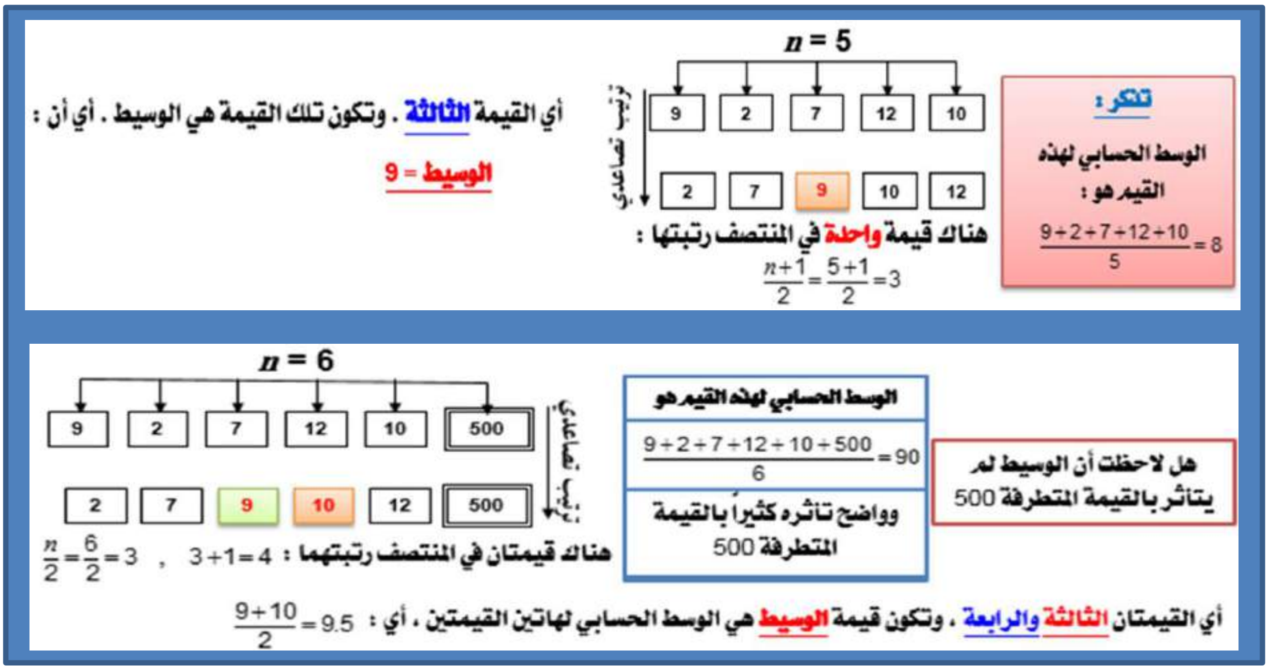
في هذه الحالة توجد قيمتان بالمنتصف وهما القيمة الخامسة والقيمة السادسة [وهما العددان 16 , 15] ، عندئذ يكون الوسيط هو الوسط الحسابي لهاتين القيمتين ، أي : $\frac{15+16}{2} = 15.5$

إذن من السابق يمكن استنتاج طريقة حساب الوسيط لمجموعة من القيم كالآتي :

- قم أولاً بترتيب البيانات **تصاعدياً** أو **تنازلياً** .
- حدد ما إذا كانت هناك قيمة واحدة بالمنتصف أم قيمتين، وهذا يتوقف على قيمة n .

فإذا كانت n فردية	وإذا كانت n زوجية
كانت هناك قيمة واحدة في المنتصف رتبتها	كانت هناك قيمتان في المنتصف رتبتهما
$\frac{n+1}{2}$	$\frac{n}{2}$, $\frac{n}{2}+1$
وتكون هذه القيمة هي الوسيط	ويكون الوسط الحسابي لهاتين القيمتين هو الوسيط

فمثلاً :



لاحظ من الأمثلة السابقة أن كلاً من المتوسطين **الوسط الحسابي و الوسيط** من السهل حسابهما ومن الممكن أن يمثل كل منهما مقياساً للنزعة المركزية للبيانات ، **لكن الأفضل (نسبياً هنا) أن نستخدم الوسط الحسابي** كمقياس للنزعة المركزية للبيانات حيث أنه يأخذ في الاعتبار جميع قيم البيانات، بينما يهتم الوسيط بقيم البيانات في المنتصف (وذلك بعد ترتيبها) .

مثال آخر : الأجر (بالريال) في الساعة لخمسة عاملين في مكتب هو : 37 , 92 , 32 , 39 , 25 . احسب الوسط الحسابي للأجور ووسيط هذه الأجور . أيهما تفضل كمقياس لمتوسط أجر الساعة ؟ ولماذا ؟

الوسط الحسابي للأجور هو : $\bar{x}=\frac{\sum x}{n}=\frac{25+39+32+92+37}{5}=\frac{225}{5}=45$

أما لتحديد الوسيط ، فلابد أولاً من ترتيب القيم (تصاعدياً مثلاً) : 92 , 39 , 37 , 32 , 25

وحيث أن عدد القيم فردي ، إذن هناك قيمة واحدة في المنتصف [هي 37] وهي الوسيط .

لاحظ في هذا السؤال أن الوسط الحسابي (بالرغم من عدم احتياجه لترتيب القيم وفي نفس الوقت يأخذ في الاعتبار جميع قيم البيانات) إلا أنه تأثر جداً بالقيمة المتطرفة 92 ، في حين لم يتأثر بها الوسيط لأنه يعتمد على البيانات في المنتصف . لذا يُفضل هنا استخدام الوسيط كمقياس للنزعة المركزية حيث يعطي دلالة أفضل لمتوسط الأجور من الوسط الحسابي .

الوسيط لبيانات كمية متصلة :

يمكن حساب الوسيط للبيانات الكمية المتصلة من خلال الرسم وكذلك من خلال المعادلات الاحصائية بسهولة

الوسيط من خلال الرسم البياني يتم كالتالي :

الجدول التالي يوضح العلاقة بين فئات الدخل بأحد المصانع وعدد العمال والمطلوب من واقع بيانات الجدول حساب الوسيط بالرسم من جدول التكرار المتجمع الصاعد والهابط معاً .

فئات الدخل	20-	30-	40-	50-	60-70
عدد العمال	20	40	100	30	10

طريقة تحديد الوسيط من :

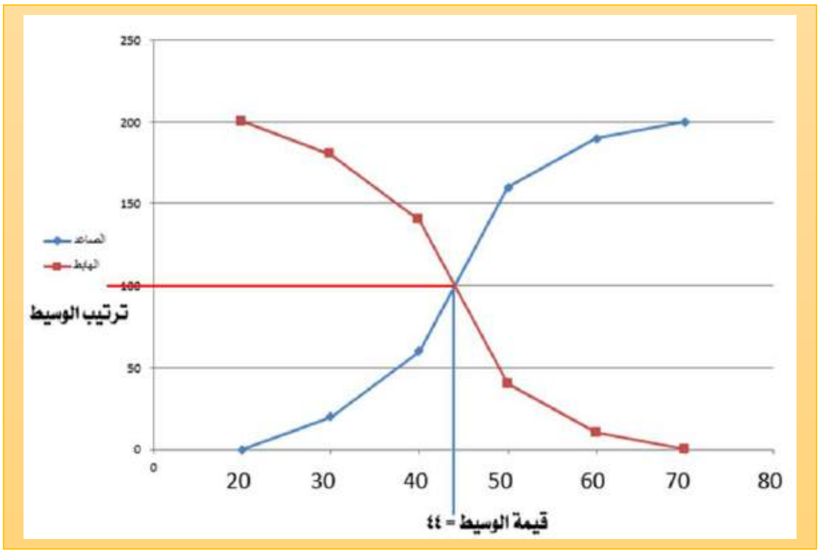
* من المنحنى المتجمع الصاعد فقط

* من المنحنى المتجمع الهابط فقط .

* من المنحنيين معاً .

ك . م . ص	الحدود العليا للفئات
صفر	أقل من ٢٠
٢٠	أقل من ٣٠
٦٠	أقل من ٤٠
١٦٠	أقل من ٥٠
١٩٠	أقل من ٦٠
٢٠٠	أقل من ٧٠

ك . م . هـ	الحدود العليا للفئات
٢٠٠	٢٠ فأكثر
١٨٠	٢٠ فأكثر
١٤٠	٢٠ فأكثر
٤٠	٢٠ فأكثر
١٠	٢٠ فأكثر
صفر	٢٠ فأكثر



الوسيط من خلال المعادلات الإحصائية :

مثال : في دراسة جغرافية لعدد من مساحات مجموعة من الأراضي لمنطقة سكنية بالرياض تبين أن التوزيع التكراري لها كما هو مبين .

المساحة (بالكيلومتر)	عدد قطع الأراضي
1 -	14
3 -	29
5 -	18
7 - 10	9

المطلوب : حساب الوسط الحسابي والوسيط لمساحة الأراضي .

المتغير x هنا هو مساحة الأرض (بالكيلومتر) ،

في حين يمثل عدد قطع الأراضي التكرار f .

أولاً : الوسط الحسابي : نستكمل الجدول التكراري كما هو مبين :

x

الجدول التكراري

الفئة	المتغير (المساحة)	التكرار f	المركز x_0	$f\,x_0$
الأولى	$1 \le x < 3$	14	2	28
الثانية	$3 \le x < 5$	29	4	116
الثالثة	$5 \le x < 7$	18	6	108
الرابعة	$7 \le x < 10$	9	8.5	76.5
		$\sum f = 70$		$\sum f x_0 = 328.5$

$$\therefore \quad \bar{x} = \frac{\sum f\,x_0}{\sum f} = \frac{328.5}{70} = 4.692857143 \quad \cong \underline{\underline{4.7}}$$

وهنا يتبادر إلى الذهن سؤالان هامان : أي أن الفئة الوسيطية هي تلك الفئة التي يقع داخلها الوسيط

السؤال الأول : هل من الممكن تحديد الفئة الوسيطية من الجدول التكراري مباشرةً أم لازم نعمل الجدول التكراري المتجمع الصاعد ونرسم المضلع التكراري المتجمع الصاعد ؟ .

السؤال الثاني : هل من الممكن [بعد تحديد الفئة الوسيطية] تحديد الوسيط من الجدول التكراري مباشرة دون الحاجة للجدول التكراري المتجمع الصاعد أو المضلع التكراري المتجمع الصاعد ؟ .

والإجابة على السؤالين : نعم يمكن تحديد الفئة الوسيطية من الجدول التكراري مباشرةً ، ثم بعد ذلك يمكن أيضاً من هذا الجدول التكراري تحديد قيمة الوسيط دون أن نحتاج لعمل جدول تكراري متجمع صاعد ورسم المضلع التكراري المتجمع الصاعد ، وذلك كالتالي:

بالنسبة للسؤال الأول [تحديد الفئة الوسيطية]

١. احسب أولاً نصف مجموع التكرارات .

٢. ابدأ بالرقم صفر في ذهنك وزود تكرارات الفئات على التوالي وكل مرة قارن بنصف مجموع التكرارات السابق . أول ما يزيد الناتج عن نصف المجموع السابق أو يساويه تكون آخر فئة زودنا تكرارها تكون هي الفئة الوسيطية .

ويتم ذلك كالتالي :

- **احسب :** $\frac{1}{2}\sum f \longrightarrow \frac{1}{2}\sum f = \frac{70}{2} = 35$
- **نبدأ بالصفر** [في ذهننا]
- **نزود على الصفر السابق تكرار الفئة الأولى [14] ينتج 14**

14 أقل من 35 ، يبقى الفئة الأولى ليست الفئة الوسيطيه

- **نزود على الـ 14 الأخيرة تكرار الفئة الثانية [29] ينتج 43**

43 أكبر من 35 ، يبقى الفئة الثانية هي الفئة الوسيطيه

الجدول التكراري		
الفئة	المتغير (المساحة)	التكرار f
الأولى	$1 \le x < 3$	14
الثانية	$3 \le x < 5$	29
الثالثة	$5 \le x < 7$	18
الرابعة	$7 \le x < 10$	9
		$\sum f = 70$

وبالنسبة للسؤال الثاني [تحديد الوسيط (بعد ما حددنا الفئة الوسيطية)]

١. حدد الحد الأدنى للفئة الوسيطية وأيضاً طولها .

٢. احسب ما يُسمى بـ "التكرار المتجمع السابق" = مجموع تكرار الفئات السابقة للفئة الوسيطية .

٣. احسب الوسيط من العلاقة :

$$\text{الوسيط } M = \text{الحد الأدنى للفئة الوسيطية} + \left[\frac{\text{(نصف مجموع التكرارات} - \text{التكرار المتجمع السابق)}}{\text{تكرار الفئة الوسيطية}} \times \text{طول الفئة الوسيطية}\right]$$

الجدول التكراري		
الفئة	المتغير (المساحة)	التكرار f
الأولى	$1 \le x < 3$	14
الثانية	$3 \le x < 5$	29
الثالثة	$5 \le x < 7$	18
الرابعة	$7 \le x < 10$	9
		$\sum f = 70$

الفئة الوسيطة

- **الفئة الوسيطية هي الفئة الثانية :**

حدها الأدنى 3 وطولها 2 [3 − 5 =] وتكرارها 29

- **التكرار المتجمع السابق :**

يساوي مجموع تكرارات الفئات السابقة للفئة الوسيطية [أي تكرار الفئة الأولى فقط] − 14

بالتعويض في القانون السابق :

$$M = 3 + \left[\frac{35-14}{29} \times 2\right] = 3 + \left[\frac{21}{29} \times 2\right] = 3 + 1.44827 = 4.44827 \cong 4.4$$

تُسمى الطريقة الحسابية السابقة (لحساب الوسيط) بـ "طريقة الاستكمال"

المنوال :

تعريف المنوال [الشائع] :

يُعرف المنوال لمجموعة من القيم على أنه القيمة التي تتكرر أكثر من غيرها أو القيمة الأكثر شيوعاً [لذا يُسمى في بعض الأحيان بالـ "**الشائع**"] .

وأحياناً يُرمز للمنوال بالرمز .

فمثلاً :

مجموعة القيم 18 12 11 10 10 9 9 9 7 5 2 2 لها منوال 9

ومجموعة القيم 16 15 12 10 8 5 3 9 ليس لها منوال [أو عديمة المنوال]

ومجموعة القيم 9 7 7 7 5 5 4 4 4 3 2 لها منوالان 7 ، 4

> أي أن مجموعة القيم قد تكون وحيدة المنوال [لها منوال واحد] ، وقد تكون عديمة المنوال [منوالان أو أكثر] وقد تكون عديمة المنوال [لا يوجد لها منوال]

٧ ٧ ٦ ٦ ٥ ٥ ٤ ٤

فقد تتسرع وتقول أنها رباعية المنوال ومناولها : ٧ ، ٦ ، ٥ ، ٤

لكن [حيث أن جميع القيم لها نفس التكرار] هذه المجموعة الأخيرة عديمة المنوال

والمنوال [مقارنةً بالوسط الحسابي والوسيط] به العديد من العيوب منها :

– أنه لا يأخذ في الاعتبار جميع البيانات ولكنه يهتم فقط بالقيم الأكثر تكراراً .

– أنه قد لا يتواجد أو قد يكون هناك أكثر من منوال للبيانات .

إلا أنه أيضاً يتميز ببعض المزايا منها :

– أنه أسرع في تحديده من الوسط والوسيط

– من الممكن تحديده للتوزيعات التكرارية للبيانات المنفصلة سواء كانت تلك البيانات كمية متقطعة أو نوعية [والبيانات الأخيرة (النوعية) لا يمكن حساب الوسط الحسابي لها أو الوسيط]

بيانات منفصلة

درجات طلاب في مقرر الإحصاء

درجة الطالب	عدد الطلاب
12	28
14	24
16	39
18	9

بيانات كمية متقطعة

لها منوال وحيد وهو "الدرجة 16"

درجات طلاب في مقرر الإنجليزي

درجة الطالب	عدد الطلاب
12	23
14	30
16	30
18	17

بيانات كمية متقطعة

لها منوالان وهما "14 , 16"

درجات طلاب في مقرر الفقه

درجة الطالب	عدد الطلاب
12	25
14	25
16	25
18	25

بيانات كمية متقطعة

ليس لها منوال

سيارات في أحد المواقف

لون السيارة	عدد
أحمر R	10
أزرق B	23
أبيض W	12
أصفر Y	5

بيانات نوعية

لها منوال وهو "اللون الأزرق"

وماذا عن التوزيعات التكرارية للبيانات الكمية المتصلة ؟!

المنوال = أ + $\frac{\text{ف}_1}{\text{ف}_1 + \text{ف}_2}$ × ل

أ = الحد الأدنى للفئة المنوالية والمقصود بدايتها .

ف1 = ك – ك1

ف2 = ك – ك2

ك = تكرار الفئة المنوالية

ك1 = تكرار الفئة التى تسبق الفئة المنوالية

ك2 = تكرار الفئة التى تلى الفئة المنوالية

ل = طول الفئة

أوجد المنوال بطريقة بيرسون من الجدول التالى :

فئات الدخل	-10	-20	-30	-40	-50	-60	80-70
عدد العمال	5	12	22	38	22	12	5

	ف	ك	
	-10	5	
	-20	12	
	-30	22	ك1
أ	-40	38	ك
	-50	22	ك2
	-60	12	
	80-70	5	

نحسب ف1 = ك – ك1 = 38 – 22 = 16

نحسب ف2 = ك – ك2 = 38 – 22 = 16

نحسب ل = 10

ثم نعوض في القانون :

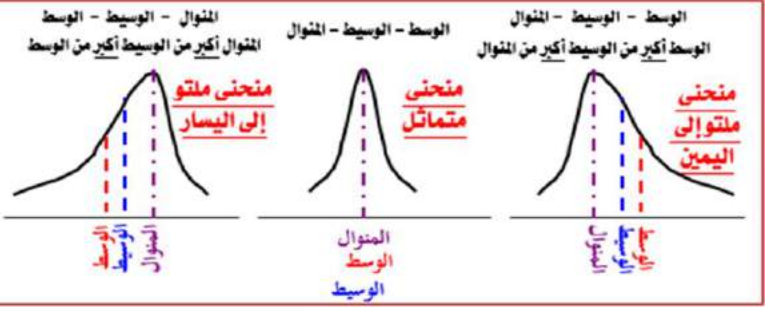
المنوال = 40 + $\frac{16}{16 + 16}$ × 10

المنوال = 40 + 5 = 45

مقارنة بين المتوسطات الثلاثة : الوسط ، الوسيط ، المنوال :

	الوسط الحسابي	الوسيط	المنوال
مزايا	١.سهولة حسابه .	١.سهولة حسابه حسابياً أو بيانياً .	١.سهولة حسابه .
	٢.يأخذ في الاعتبار جميع البيانات .	٢.لا يتأثر بالقيم المتطرفة .	٢.لا يتأثر كثيراً بالقيم المتطرفة .

	٣.لا يحتاج إلى ترتيب معين للبيانات .	٣.يمكن حسابه في حالة التوزيعات التكرارية المفتوحة .	٣.لا يحتاج لترتيب البيانات .
عيوبه	١.يتأثر بشدة بالقيم المتطرفة ٢.لا يمكن إيجاده بالرسم [بيانياً] ٣.لا يمكن حسابه في حالات التوزيعات التكرارية المفتوحة.	١.يحتاج إلى ترتيب للبيانات أولاً ٢.لا يأخذ في الاعتبار جميع البيانات	١. قد لا يتواجد وقد يكون له أكثر من قيمة



في جميع المنحنيات المبينة ، لاحظ الأتي

المنوال هو القيمة المناظرة لأعلى نقطة في المنحنى

الوسيط يقع دائماً بين الوسط و المنوال

$$\text{الوسط} - \text{المنوال} = ٣ \times (\text{الوسط} - \text{الوسيط})$$

وهذه العلاقة يمكن وضعها على أي صورة من الصور التالية

$$\text{الوسط} = \frac{(٣ \times \text{الوسيط}) - \text{المنوال}}{٢}$$

وهذه الصورة مفيدة عندما يكون الوسيط و المنوال معلومان ونريد معرفة الوسط الحسابي

$$\text{المنوال} = (٣ \times \text{الوسيط}) - (٢ \times \text{الوسط})$$

وهذه الصورة مفيدة عندما يكون الوسط الحسابي و الوسيط معلومان ونريد معرفة المنوال

$$\text{الوسيط} = \frac{(٢ \times \text{الوسط}) + \text{المنوال}}{٣}$$

وهذه الصورة مفيدة عندما يكون الوسط الحسابي و المنوال معلومان ونريد معرفة الوسيط

فمثلاُ إذا كان المنوال لمجموعة من القيم = 95 ، والوسيط لها = 85 ، فإن :

$$\text{الوسط} = \frac{(٣ \times \text{الوسيط}) - \text{المنوال}}{٢} = \frac{(٣ \times ٨٥) - ٩٥}{٢} = \frac{٢٢٥ - ٩٥}{٢} = \frac{١٦٠}{٢} = ٨٠$$

وإذا كان الوسط الحسابي لمجموعة من القيم = 80 ، والوسيط لها = 85 ، فإن :

$$\text{المنوال} = (٣ \times \text{الوسيط}) - (٢ \times \text{الوسط}) = (٣ \times ٨٥) - (٢ \times ٨٠) = ٢٥٥ - ١٦٠ = ٩٥$$

وإذا كان الوسط الحسابي لمجموعة من القيم = 80 ، والمنوال لها = 95 ، فإن :

$$\text{الوسيط} = \frac{(٢ \times \text{الوسط}) + \text{المنوال}}{٣} = \frac{(٢ \times ٨٠) + ٩٥}{٣} = \frac{١٦٠ + ٩٥}{٣} = \frac{٢٥٥}{٣} = ٨٥$$

سؤال : المنحنى التكراري للبيانات المذكورة في أي من الأمثلة السابقة :

( ) ملتو لليمين ( ) ملتو لليسار ( ) متماثل

تمرين: من واقع بيانات الجدول التالي :

الفئات	التكرار
٥ -	٢
١٠ -	٤
١٥ -	٦
٢٠ -	٨
٢٥ -	١٠
٣٠ -	١٦
٣٥ -	٤٠
٤٠ -	٢٤
٤٥ -	١٤
٥٠ -	١١
٥٥ - ٦٠	٥

- احسب الوسط الحسابي بطريقة مراكز الفئات ؟ ..
- احسب الوسيط بطريقتين مختلفتين ؟ ..
- احسب المنوال ؟ ..

المحاضرة الرابعة : مقاييس التشتت (المدى، الإنحراف المتوسط، التباين، الإنحراف المعياري)

تعريف التشتت :

درجة التباعد أو التقارب التي تتجه بها البيانات الكمية **للانتشار** حول قيمة متوسطة (أحد مقاييس النزعة المركزية) **تُسمى تشتت أو تغير البيانات**

. وتستخدم مقاييس التشتت في المقارنة بين مجموعات البيانات من حيث تشتتها.

كلما قل تشتت البيانات ← **كلما اقتربت من متوسطها** ← **كلما كانت اقرب للتجانس**

هل يمكن الاكتفاء بالوسط الحسابي في وصف البيانات؟

إذا كان لدينا ٣ مجموعات من الطلاب، كل مجموعة مكونة من خمسة طلاب، وكانت درجاتهم في أحد المقررات كالتالي:

المجموعة الأولى	المجموعة الثانية	المجموعة الثالثة
5,5,5,5,5	3,4,5,6,7	1,2,5,8,9
وسطها الحسابي 5	وسطها الحسابي 5	وسطها الحسابي 5

المجموعات الثلاثة لها وسط حسابي 5 ، لكن في المجموعة الأولى : جميع القيم متساوية وتساوي الوسط 5 ، في حين تنتشر البيانات في المجموعة الثانية حول هذا الوسط بقدرٍ ما ، وفي المجموعة الثالثة تنتشر البيانات حول الوسط بقدر آخر .

أي أن الوسط الحسابي وحده ليس كافياً وحده لوصف البيانات، ولكن لابد من وجود نوع آخر من المقاييس **لرصد مدى تشتت** البيانات عن تلك القيمة المتوسطة الممثلة للبيانات. هذا النوع من المقاييس هو **ما نسميه بـ مقاييس التشتت** .

أولاً : المدى :

البيانات الغير مبوبة ← الفرق بين أكبر قيمة و أصغر قيمة في البيانات و يرمز له بالرمز R

15 3 7 6 12 18 5 3 13 15 يكون المدى : $R=18-3=15$

البيانات المبوبة ← الفرق بين الحد الأعلى للفئة الأخيرة و الحد الأدنى للفئة الأولى

الفئة	العمر x
الأولى	$2 \le x < 6$
الثانية	$6 \le x < 12$
الثالثة	$12 \le x < 15$
الرابعة	$15 \le x < 18$

$R = 18 - 2 = 16$

الحد الأدنى للفئة الأولى الحد الأعلى للفئة الأخيرة

وبالرغم من بساطة تحديده إلا أن له بعض العيوب:

✓ تأثره بالقيم المتطرفة .

فمثلاً لمجموعة القيم : 15 3 7 6 12 18 5 3 13 15 يكون المدى $R=18-3=15$

ولمجموعة القيم : 16 3 14 15 17 18 17 13 14 16 يكون المدى $R=18-3=15$

16 15 14 15 17 18 17 13 14 16 يكون المدى $R=18-15=3$

أي أن المدى واحد للمجموعتين في حين يبدو للعين المجردة أن هناك تشتت للبيانات أكبر في المجموعة الأولى عنه في المجموعة الثانية ، مما يعني أن المدى هنا لا يظهر هذا الفارق.

لذا يُعد المدى مقياساً للتشتت لكنه غير جيد في كثير من الأحيان .

✓ لا يمكن تحديده في حالة التوزيعات التكرارية المفتوحة .

العمر x	الفئة
$x < 6$	الأولى
$6 \le x < 12$	الثانية
$12 \le x < 15$	الثالثة
$15 \le x < 18$	الرابعة

مفتوح من أسفل

العمر x	الفئة
$6 \le x < 12$	الأولى
$12 \le x < 15$	الثانية
$15 \le x < 18$	الثالثة
$x \ge 18$	الرابعة

مفتوح من أعلى

العمر x	الفئة
$x < 6$	الأولى
$6 \le x < 12$	الثانية
$12 \le x < 15$	الثالثة
$x \ge 15$	الرابعة

مفتوح من الطرفين

لا يمكن تحديد مدى البيانات

✓ لا يدخل في حسابه جميع البيانات

مثال :

البيانات التالية لدرجات ذكاء مجموعتين من الأطفال أوجد المدى وقارن بين المجموعتين:

القيمة الكبرى للذكاء	القيمة الصغرى للذكاء	متوسط الذكاء	
112	90	105	المجموعة A
140	75	120	المجموعة B

A المدى $= 112 - 90 = 22$ درجة

B المدى $= 140 - 75 = 65$ درجة

مثال :

البيانات التالية لدرجات الطلاب في مقرر الإحصاء الاجتماعي. احسب المدي لدرجات الطلاب؟

90 — 98	82 —	74 —	66 —	58 —	50 —	فئات الدرجات
8	15	40	24	10	3	عدد الطلاب

المدى $= 98 - 50 = 48$

ثانياً : الانحراف المتوسط [أو متوسط الانحرافات] $M.D$

يُعرف **الانحراف المتوسط** (أو **متوسط الانحرافات**) [**وسنرمز له بالرمز** $M.D$] على أنه متوسط القيم المطلقة للانحرافات عن قيمة متوسطة للبيانات [عادةً تكون **الوسط الحسابي** أو **الوسيط**] .

فإذا اعتبرنا أن القيمة المتوسطة للبيانات هي الوسط الحسابي ، فإن الانحراف المتوسط لمجموعة من البيانات عددها n يُعطى بـ :

$d = x - \bar{x}$ الانحراف عن الوسط

$$M.D = \frac{\sum |x - \bar{x}|}{n} \quad \text{أو} \quad M.D = \frac{\sum |d|}{n}$$

n عدد القيم

ملحوظة هامة : القيمة المطلقة لأي عدد y هي **القيمة العددية له دون إشارة** ، ونرمز له بنفس الرمز y لكن بين خطين رأسيين | | ، أي نكتب القيمة المطلقة لـ y على الصورة $|y|$. فمثلاً :

$|3| = 3$, $|-3| = 3$, $|2.5| = 2.5$, $|-3.25| = 3.25$

وهكذا .

حيث $d = x - \bar{x}$ هي انحراف القيمة x عن الوسط الحسابي ، $|d|$ هي القيمة المطلقة للانحراف d .

احسب الوسط الحسابي ← ثم احسب الانحرافات عن الوسط ← ثم احسب القيم المطلقة للانحرافات ← ثم احسب متوسط القيم المطلقة

وسطها الحسابي :

$$\bar{x} = \frac{15+13+3+5+18+12+6+7+3+15}{10} = 9.7$$

15	3	7	6	12	18	5	3	13	15
- 9.7	- 9.7	- 9.7	- 9.7	- 9.7	- 9.7	- 9.7	- 9.7	- 9.7	- 9.7
5.3	-6.7	-2.7	-3.7	2.3	8.3	-4.7	-6.7	3.3	5.3
5.3	6.7	2.7	3.7	2.3	8.3	4.7	6.7	3.3	5.3

لاحظ أن مجموع الانحرافات = صفر

إذن **الانحراف المتوسط** هو متوسط القيم المطلقة للانحرافات :

$$M.D = \frac{5.3+3.3+6.7+4.7+8.3+2.3+3.7+2.7+6.7+5.3}{10} = \underline{4.9}$$

وسطها الحسابي :

$$\frac{16+14+13+17+18+17+15+14+3+16}{10} = 14.3$$

16	14	13	17	18	17	15	14	3	16
-14.3	-14.3	-14.3	-14.3	-14.3	-14.3	-14.3	-14.3	-14.3	-14.3
1.7	-0.3	-1.3	2.7	3.7	2.7	0.7	-0.3	-11.3	1.7
1.7	0.3	1.3	2.7	3.7	2.7	0.7	0.3	11.3	1.7

لاحظ أن مجموع الانحرافات = صفر

$$M.D = \frac{1.7+0.3+1.3+2.7+3.7+2.7+0.7+0.3+11.3+1.7}{10} = 2.64$$

أي أن الانحراف المتوسط للمجموعة الثانية من القيم أقل من الانحراف المتوسط للمجموعة الأولى من القيم **مما يعني أن المجموعة الثانية أقل تشتتاً من المجموعة الأولى** وهذا الأمر لا يمكن ملاحظته عند استخدام المدى كمقياس للتشتت .

ويمكن أن يتم حل السؤال السابق وذلك بتنظيم خطواتنا من خلال جداول كالتالي :

المجموعة الأولى [$n = 10$]

x	$\bar{x}$	$d = x - \bar{x}$	$\|d\|$
15	9.7	15 – 9.7 = 5.3	5.3
13	9.7	13 – 9.7 = 3.3	3.3
3	9.7	3 – 9.7 = -6.7	6.7
5	9.7	5 – 9.7 = -4.7	4.7
18	9.7	18 – 9.7 = 8.3	8.3
12	9.7	12 – 9.7 = 2.3	2.3
6	9.7	6 – 9.7 = -3.7	3.7
7	9.7	7 – 9.7 = -2.7	2.7
3	9.7	3 – 9.7 = -6.7	6.7
15	9.7	15 – 9.7 = 5.3	5.3
97	97	0	49
$\sum x$		$\sum d$	$\sum \|d\|$

$$\bar{x} = \frac{\sum x}{n} = \frac{97}{10} = 9.7 \qquad M.D = \frac{\sum |d|}{n} = \frac{49}{10} = 4.9$$

المجموعة الثانية [$n = 10$]

x	$\bar{x}$	$d = x - \bar{x}$	$\|d\|$
16	14.3	16 – 14.3 = 1.7	1.7
14	14.3	14 – 14.3 = -0.3	0.3
13	14.3	13 – 14.3 = -1.3	1.3
17	14.3	17 – 14.3 = 2.7	2.7
18	14.3	18 – 14.3 = 3.7	3.7
17	14.3	17 – 14.3 = 2.7	2.7
15	14.3	15 – 14.3 = 0.7	0.7
14	14.3	14 – 14.3 = -0.3	0.3
3	14.3	3 – 14.3 = -11.3	11.3
16	14.3	16 – 14.3 = 1.7	1.7
143	143	0	26.4
$\sum x$		$\sum d$	$\sum \|d\|$

$$\bar{x} = \frac{\sum x}{n} = \frac{143}{10} = 14.3 \qquad M.D = \frac{\sum |d|}{n} = \frac{26.4}{10} = 2.64$$

– **وفي حالة البيانات الكمية المتقطعة ذات التكرارات :**

يمكن تحديد الانحراف المتوسط M.D من العلاقة : $M.D = \frac{\sum f\,|d|}{\sum f}$ **أي نضرب القيمة المطلقة لانحراف كل قيمة [عن الوسط] في تكرارها ، ثم نقسم الناتج على مجموع التكرارات**

فمثلاً : إذا كان المطلوب حساب الانحراف المتوسط للبيانات المبينة بالجدول التكراري :

المتغير x	التكرار f
4	20
5	40
6	30
7	10

الجدول التكراري

المتغير x	التكرار f	fx	$d = x - \bar{x}$	$\|d\|$	$f\|d\|$
4	20	80	4 – 5.3 = -1.3	1.3	20 x 1.3 = 26
5	40	200	5 – 5.3 = -0.3	0.3	40 x 0.3 = 12
6	30	180	6 – 5.3 = 0.7	0.7	30 x 0.7 = 21
7	10	70	7 – 5.3 = 1.7	1.7	10 x 1.7 = 17
	100	530			76
	$\sum f = 100$	$\sum fx = 530$			$\sum f\|d\|$

$$\bar{x} = \frac{\sum fx}{\sum f} = \frac{530}{100} = 5.3 \qquad M.D = \frac{\sum f|d|}{\sum f} = \frac{76}{100} = 0.76$$

انتبه :

مجموع الانحرافات هنا [والذي يجب أن يساوي صفراً] هو $\sum fd$ وليس $\sum d$

– **وفي حالة البيانات الكمية المتصلة :**

نستخدم نفس العلاقة السابقة لتحديد الانحراف المتوسط M.D ، أي يكون $M.D = \frac{\sum f\,|d|}{\sum f}$ حيث $d = x_0 - \bar{x}$ ، x_0 تمثل مراكز الفئات

.

الفئة	المتغير x	التكرار f	المركز x_0	fx_0	$d = x_0 - \bar{x}$	$\|d\|$	$f\,\|d\|$
الأولى	$50 \le x < 60$	6	55	330	$55 - 83.75 = -28.75$	28.75	172.5
الثانية	$60 \le x < 70$	9	65	585	$65 - 83.75 = -18.75$	18.75	168.75
الثالثة	$70 \le x < 80$	15	75	1125	$75 - 83.75 = -8.75$	8.75	131.25
الرابعة	$80 \le x < 90$	12	85	1020	$85 - 83.75 = 1.25$	1.25	15
الخامسة	$90 \le x < 100$	9	95	855	$95 - 83.75 = 11.25$	11.25	101.25
السادسة	$100 \le x < 120$	6	110	660	$110 - 83.75 = 26.25$	26.25	157.5
السابعة	$120 \le x < 180$	3	150	450	$150 - 83.75 = 66.25$	66.25	198.75
		60		5025			945
		$\sum f$		$\sum fx_0$			$\sum f\,\|d\|$

$$\bar{x} = \frac{\sum f x_0}{\sum f} = \frac{5025}{60} = \underline{83.75}$$

$$M.D = \frac{\sum f\,|d|}{\sum f} = \frac{945}{60} = 15.75$$

ثالثاً : التباين s^2 والانحراف المعياري s :

يُعرف متوسط مربعات الانحرافات عن الوسط الحسابي على أنه **تباين** مجموعة البيانات [ويُرمز له بالرمز s^2] ، ويُعرف الجذر التربيعي للتباين على أنه **الانحراف المعياري** للبيانات [ويُرمز له بالرمز s] ، أي أن :

التباين $s^2 = \frac{\sum d^2}{n} =$

$s = \sqrt{s^2} = \sqrt{\frac{\sum d^2}{n}} =$ الانحراف المعياري

المجموعة الأولى [$n = 10$]

x	$d = x - \bar{x}$	d^2
15	$15 - 9.7 = 5.3$	28.09
13	$13 - 9.7 = 3.3$	10.89
3	$3 - 9.7 = -6.7$	44.89
5	$5 - 9.7 = -4.7$	22.09
18	$18 - 9.7 = 8.3$	68.89
12	$12 - 9.7 = 2.3$	5.29
6	$6 - 9.7 = -3.7$	13.69
7	$7 - 9.7 = -2.7$	7.29
3	$3 - 9.7 = -6.7$	44.89
15	$15 - 9.7 = 5.3$	28.09

$$\bar{x} = \frac{\sum x}{n} = 9.7$$

$$s^2 = \frac{\sum d^2}{n} = \frac{274.1}{10} = 27.41$$

$$s = \sqrt{s^2} = \sqrt{27.41} \cong \underline{5.24}$$

المجموعة الثانية [$n = 10$]

x	$d = x - \bar{x}$	d^2
16	$16 - 14.3 = 1.7$	2.89
14	$14 - 14.3 = -0.3$	0.09
13	$13 - 14.3 = -1.3$	1.69
17	$17 - 14.3 = 2.7$	7.29
18	$18 - 14.3 = 3.7$	13.69
17	$17 - 14.3 = 2.7$	7.29
15	$15 - 14.3 = 0.7$	0.49
14	$14 - 14.3 = -0.3$	0.09
3	$3 - 14.3 = -11.3$	127.69
16	$16 - 14.3 = 1.7$	2.89

$$\bar{x} = \frac{\sum x}{n} = 14.3$$

$$s^2 = \frac{\sum d^2}{n} = \frac{164.1}{10} = 16.41$$

$$s = \sqrt{s^2} = \sqrt{16.41} \cong \underline{4.05}$$

– **وفي حالة البيانات الكمية المتقطعة ذات التكرارات :**

يمكن تحديد التباين s^2 والانحراف المعياري s من :

التباين $s^2 = \frac{\sum fd^2}{\sum f}$ ومنه يكون $s = \sqrt{s^2} = \sqrt{\frac{\sum fd^2}{\sum f}} =$ الانحراف المعياري

فمثلاً : إذا كان المطلوب حساب الانحراف المعياري للبيانات المبينة بالجدول التكراري :

الجدول التكراري

المتغير x	التكرار f
4	20
5	40
6	30
7	10

المتغير x	التكرار f	fx	$d = x - \bar{x}$	d^2	fd^2
4	20	80	$4 - 5.3 = -1.3$	1.69	$20 \times 1.69 = 33.8$
5	40	200	$5 - 5.3 = -0.3$	0.09	$40 \times 0.09 = 3.6$
6	30	180	$6 - 5.3 = 0.7$	0.49	$30 \times 0.49 = 14.7$
7	10	70	$7 - 5.3 = 1.7$	2.89	$10 \times 2.89 = 28.9$
	100	530			81

$\sum f = 100$ $\sum fx = 530$ $\sum fd^2$

التباين $= s^2 = \dfrac{\sum fd^2}{\sum f}$ ومنه يكون الانحراف المعياري $= \sqrt{\dfrac{\sum fd^2}{\sum f}} = \sqrt{s^2} = s$ حيث $d = x_0 - \bar{x}$

$\bar{x} = \dfrac{\quad}{\sum f} = \dfrac{\quad}{100} = 5.3$

– وفي حالة ال

أي أنه عند حساب الانحرافات نعتبر أن مركز أي فئة يمثل جميع القيم الموجودة في تلك الفئة :

الفئة	المتغير x	التكرار f	المركز x_0	fx_0	$d = x_0 - \bar{x}$	d^2	fd^2
الأولى	$0 \le x < 20$	4	10	40	$10 - 31.7 = -21.7$	470.89	$4 \times 470.89 = 1883.56$
الثانية	$20 \le x < 30$	16	25	400	$25 - 31.7 = -6.7$	44.89	$16 \times 44.89 = 718.24$
الثالثة	$30 \le x < 35$	12	32.5	390	$32.5 - 31.7 = 0.8$	0.64	$12 \times 0.64 = 7.68$
الرابعة	$35 \le x < 40$	10	37.5	375	$37.5 - 31.7 = 5.8$	33.64	$10 \times 33.64 = 336.4$
الخامسة	$40 \le x < 50$	6	45	270	$45 - 31.7 = 13.3$	176.89	$6 \times 176.89 = 1061.34$
السادسة	$50 \le x < 60$	2	55	110	$55 - 31.7 = 23.3$	542.89	$2 \times 542.89 = 1085.78$
		50		1585			5093
		$\sum f$		$\sum fx_0$			$\sum fd^2$

$\therefore\ \bar{x} = \dfrac{\sum f x_0}{\sum f} = \dfrac{1585}{50} = \underline{\underline{31.7}}$ $s^2 = \dfrac{\sum fd^2}{\sum f} = \dfrac{5093}{50} = 101.86$ $s = \sqrt{s^2} = \sqrt{101.86} \cong \underline{\underline{10.09}}$

<u>وبنفس الأسلوب يمكن التعامل مع المثال السابق لحساب التباين والانحراف المعياري:</u>

الفئة	المتغير x	التكرار f	المركز x_0	fx_0	$d = x_0 - \bar{x}$	d^2	$f \times d^2$
الأولى	$50 \le x < 60$	6	55	330	$55 - 83.75 = -28.75$	826.56	4959.38
الثانية	$60 \le x < 70$	9	65	585	$65 - 83.75 = -18.75$	351.56	3164.04
الثالثة	$70 \le x < 80$	15	75	1125	$75 - 83.75 = -8.75$	76.56	1148.4
الرابعة	$80 \le x < 90$	12	85	1020	$85 - 83.75 = 1.25$	1.56	18.72
الخامسة	$90 \le x < 100$	9	95	855	$95 - 83.75 = 11.25$	126.56	1139.04
السادسة	$100 \le x < 120$	6	110	660	$110 - 83.75 = 26.25$	689.06	4134.36
السابعة	$120 \le x < 180$	3	150	450	$150 - 83.75 = 66.25$	4389.06	13167.18
		60		5025			27731.12
		$\sum f$		$\sum fx_0$			$\sum fd^2$

$\bar{x} = \dfrac{\sum f x_0}{\sum f} = \dfrac{5025}{60} = \underline{\underline{83.75}}$ $s^2 = \dfrac{\sum fd^2}{\sum f} = \dfrac{27731.12}{60} \cong 462.19$ $s = \sqrt{s^2} = \sqrt{46219} \cong \underline{\underline{215}}$

من السابق يتضح أن كلاً من الانحراف المتوسط والانحراف المعياري يعتمدا تماماً في حساباتهما على الوسط الحسابي ، وبالتالي فلهما نفس مزايا وعيوب الوسط الحسابي . أي :

المزايا	العيوب
• من السهل حسابهما	• يتأثرا بشدة بالقيم المتطرفة
• يأخذ في الاعتبار جميع البيانات	• لا يمكن إيجادهما بالرسم (بيانياً)
• لا يحتاجا لترتيب معين للبيانات	• لا يمكن حسابهما للتوزيعات التكرارية المفتوحة

ويمكن تلخيص كل ما يخص الوسط الحسابي والانحراف المتوسط والانحراف المعياري في الآتي :

للقيم المفردة :

$\bar{x} = \frac{\sum x}{n}$ = الوسط الحسابي

$M.D = \frac{\sum |d|}{n}$ = الانحراف المتوسط

$s^2 = \frac{\sum d^2}{n} \longrightarrow s = \sqrt{s^2}$

قيم عددها n	الانحرافات عن الوسط	القيم المطلقة للانحرافات	مربع الانحرافات
x	$d = x - \bar{x}$	$\|d\|$	d^2
...	...	...	...
...	...	...	...
$\sum x$		$\sum \|d\|$	$\sum d^2$

• ولتوزيع تكراري :

$\bar{x} = \frac{\sum fx}{\sum f}$

$M.D = \frac{\sum f|d|}{\sum f}$

$s^2 = \frac{\sum fd^2}{\sum f} \longrightarrow s = \sqrt{s^2}$

القيم	التكرار		الانحرافات عن الوسط	القيم المطلقة للانحرافات	مربع الانحرافات		
x	f	fx	$d = x - \bar{x}$	$\|d\|$	d^2	$f\|d\|$	fd^2
...	...	...	...	...	...	...	...
...	...	...	...	...	...	...	...
	$\sum f$	$\sum fx$				$\sum f\|d\|$	$\sum fd^2$

وللبيانات المتصلة :

الفئات	التكرار			الانحرافات عن الوسط	القيم المطلقة للانحرافات	مربع الانحرافات		
x	f	مراكز الفئات	fx_0	$d = x_0 - \bar{x}$	$\|d\|$	d^2	$f\|d\|$	fd^2
...	...	x_0	...	...	...	...	...	...
...	...	...	...	...	...	...	...	...
	$\sum f$	...	$\sum fx$				$\sum f\|d\|$	$\sum fd^2$

خاصيتان هامتان للانحراف المتوسط والانحراف المعياري :

الخاصية الأولى : إضافة عدد ثابت c لكل قيمة من قيم البيانات لا يؤثر على قيمة الانحرافين المتوسط والمعياري .

الانحراف المتوسط (أو المعياري) الجديد = الانحراف المتوسط (أو المعياري) القديم

الخاصية الثانية : ضرب كل قيمة من قيم البيانات في عدد ثابت c يجعل :

الانحراف المتوسط (أو المعياري) الجديد = الانحراف المتوسط (أو المعياري) القديم × القيمة المطلقة للثابت c

فمثلاً : لو كانت لدينا البيانات التالية والتي توضح درجات مجموعة من الطلاب كالتالي :

الدرجات الأصلية			
x	d	$\|d\|$	d^2
9	1	1	1
2	-6	6	36
7	-1	1	1
12	4	4	16
10	2	2	4
40		14	58

$\bar{x} = \frac{\sum x}{n} = \frac{40}{5} = 8$

$M.D = \frac{\sum |d|}{n} = \frac{14}{5} = 2.8$

$s^2 = \frac{\sum d^2}{n} = \frac{58}{5} = 11.6$

$s = \sqrt{11.6} \approx 3.4$

بعد إضافة 5 لكل درجة			
x	d	$\|d\|$	d^2
14	1	1	1
7	-6	6	36
12	-1	1	1
17	4	4	16
15	2	2	4
65		14	58

$\bar{x} = \frac{\sum x}{n} = \frac{65}{5} = 13$

$M.D = \frac{\sum |d|}{n} = \frac{14}{5} = 2.8$

$s^2 = \frac{\sum d^2}{n} = \frac{58}{5} = 11.6$

$s = \sqrt{11.6} \approx 3.4$

المحاضرة الخامسة : الفروض الإحصائية

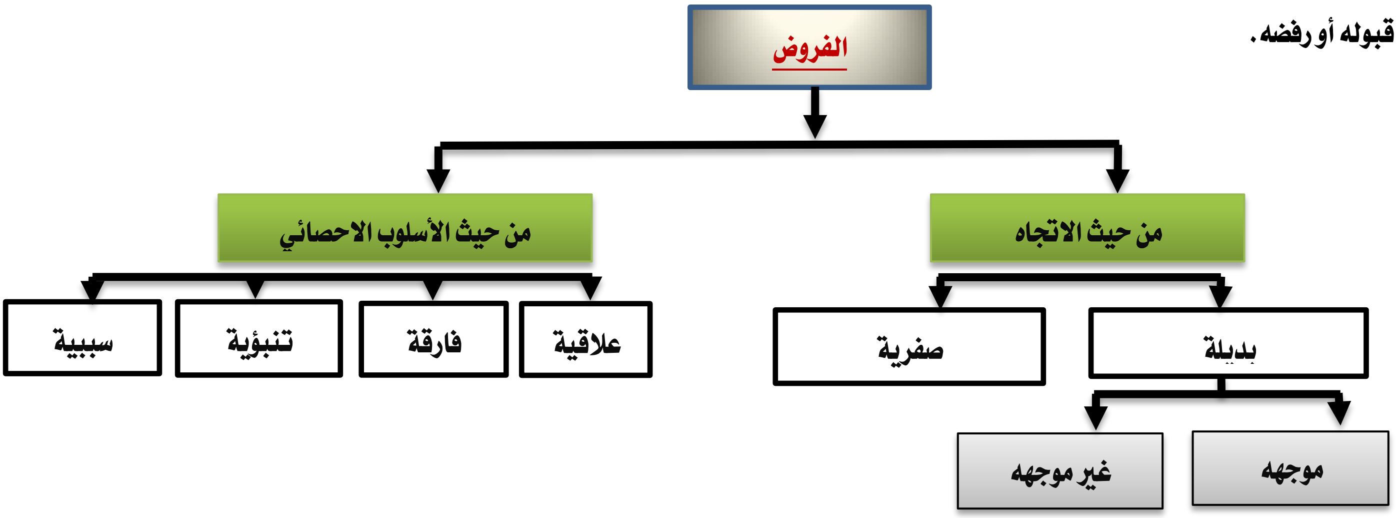
يعرف الفرض بأنه إجابة متوقعة لسؤال من الأسئلة التي تراود ذهن الباحث أو المهتم، وهذه الإجابة لا تكون نهائية وإنما خاضعة للدراسة والتحقق من مدى صحتها فإما أن تكون الإجابة صحيحة وإما أن تكون الإجابة خاطئة.

وتوقع الإجابة من جانب الباحث لا يتم من فراغ وإنما بناءً على خلفية نظرية متعلقة بهذا السؤال ونتائج دراسات سابقة حوله.

فمثلاً: يراود ذهن الباحث سؤال مضمونه: **ما طبيعة العلاقة بين حب الاستطلاع والقدرة الابتكارية لدى طلاب قسم علم الاجتماع ؟** وبناء على الخلفية النظرية ونتائج الدراسات السابقة المرتبطة بطبيعة العلاقة بين المتغيرين يصيغ الباحث إجابة متوقعة لهذا السؤال وهي تمثل إحدى فروض بحثه **وتكون صياغة الفرض كالتالي:**

- توجد علاقة بين حب الاستطلاع والابتكارية
- لا توجد علاقة بين حب الاستطلاع والابتكارية

الفرض هو اقتراح لقضية معينة وبالتالي فإن قرار قبولنا هذا الاقتراح كاقتراح صحيح أو رفضنا إياه كاقتراح خاطئ لا بد أن يؤجل حتى نجمع دليل يؤكد قبوله أو رفضه.



الفرضية الصفرية (فرضية العدم) (H_0) (Null Hypothesis) :

هي الفرضية حول معلمة المجتمع التي نجري اختبار عليها باستخدام بيانات من عينة والتي تشير أن الفرق بين معلمة المجتمع والإحصائي من العينة ناتج عن الصدفة ولا فرق حقيقي بينهما. وهي الفرضية التي ننطلق منها ونرفضها عندما تتوفر دلائل على عدم صحتها، وخلاف ذلك نقبلها وتعني كلمة Nul انه لا يوجد فرق بين معلمة المجتمع والقيمة المدعاة (إحصائية العينة).

الفرضية البديلة (H_a) Alternative Hypothesis :

هي الفرضية التي يضعها الباحث كبديل عن فرضية العدم و نقبلها عندما نرفض فرضية العدم باعتبارها ليست صحيحة بناء على المعلومات المستقاة من العينة.

وفي اختبار الفروض يمكن أن نرتكب نوعين من الخطأ :

الخطأ من النوع الأول Type I error: الخطأ من النوع الأول هو "رفض الفرض العدمي بينما هو صحيح ". أي أنه على الرغم من أن الفرض العدمي في الواقع صحيح وكان من الواجب قبوله فقد تم أخذ قرار خاطئ برفضه. وباختصار شديد فإن الخطأ من النوع الأول هو : " رفض فرض صحيح". **ويرمز له بالرمز α .**

الخطأ من النوع الثاني Type II error: وفي المقابل فإن الخطأ من النوع الثاني يعني " قبول الفرض العدمي بينما هو خاطئ " أي أنه على الرغم من أن الفرض العدمي خاطئ وكان من الواجب رفضه فقد تم أخذ قرار خاطئ بقبوله وباختصار شديد فإن الخطأ من النوع الثاني هو " قبول فرض خاطئ ". **ويرمز له بالرمز β .**

ويمكن أن نمثل ذلك في الجدول التالي:

الفرضية / القرار	صحيحة (H_0)	خاطئة (H_a)
قبول (H_0)	صواب	خطأ ٢ بيتا (B)
رفض (H_0)	خطأ ١ ألفا (α)	صواب

١. فرضية صحيحة نتائج العينة تؤيد صحتها. (قبول صواب)

٢. فرضية صحيحة نتائج العينة غير مؤيدة لصحتها. (رفض صواب) وهذا يعطينا خطأ من النوع الأول ألفا (α)

٣. فرضية خاطئة نتائج تؤيد صحتها (قبول خطأ) وهذا يعطينا خطأ من النوع الثاني بيتا (B) ويمكن أن يقلل بزيادة حجم العينة

٤. فرضية خاطئة نتائج غير مؤيدة صحتها (رفض خطأ)

الفروض البحثية:

هي الفروض التي يصيغها الباحث في بحثه بناءً على خلفيته النظرية ونتائج الدراسات السابقة.

١) الفروض العلاقية	
أ. الفرض البديل العلاقي غير الموجه	توجد علاقة دالة إحصائياً بين الاتجاه نحو الدراسة والبيئة الدراسية
ب. الفرض البديل العلاقي الموجه	توجد علاقة ايجابية دالة إحصائياً بين الاتجاه نحو الدراسة والبيئة الدراسية
ج. الفرض الصفري العلاقي	لا توجد علاقة دالة إحصائياً بين نحو الدراسة والبيئة الدراسية
٢) الفروض الفارقة	
أ. الفرض البديل الفارق غير الموجه	توجد فروق ذات دلالة إحصائية بين متوسطي درجات الذكور والإناث في الذكاء الوجداني
ب. الفرض البديل الفارق الموجه	توجد فروق ذات دلالة إحصائية بين متوسطي درجات الذكور والإناث في الذكاء الوجداني لصالح الذكور
ج. الفرض الصفري الفارق	لا توجد فروق ذات دلالة إحصائية بين متوسطي درجات الذكور والإناث في الذكاء الوجداني
٣) الفروض التنبؤية	
أ. الفرض البديل التنبؤي غير الموجه	يمكن التنبؤ من المتغيرات المستقلة (الدافعية، وحب الاستطلاع، والقلق) بالمتغير التابع (التحصيل الدراسي) لدى طلاب جامعة الملك فيصل
ب. الفرض البديل التنبؤي الموجه	يمكن التنبؤ من المتغيرات المستقلة (الدافعية كمنبئ موجب، وحب الاستطلاع كمنبئ موجب، والقلق كمنبئ سالب) بالمتغير التابع (التحصيل الدراسي) لدى طلاب جامعة الملك فيصل
ج. الفرض الصفري التنبؤي	لا يمكن التنبؤ من المتغيرات المستقلة (الدافعية، وحب الاستطلاع، والقلق) بالمتغير التابع (التحصيل الدراسي) لدى طلاب جامعة الملك فيصل
٤) الفروض السببية	
أ. الفرض البديل السببي غير الموجه	يمكن التوصل إلى نموذج سببي يفسر العلاقة بين المتغيرات المستقلة (المعاملة الوالدية، والذكاء، والضغوط النفسية، والاتجاه نحو الدراسة) والمتغير التابع (مستوى الطموح) لدى طلاب جامعة الملك فيصل .

ب. الفرض البديل السببي الموجه	يمكن التوصل إلى نموذج سببي يفسر العلاقة بين المتغيرات المستقلة (المعاملة الوالدية «تأثير موجب»، والذكاء «تأثير موجب»، والضغوط النفسية تأثير سالب»، والاتجاه نحو الدراسة «تأثير موجب») والمتغير التابع (مستوى الطموح) لدى طلاب جامعة الملك فيصل
ج. الفرض الصفري السببي	لا يمكن التوصل إلى نموذج سببي يفسر العلاقة بين المتغيرات المستقلة (المعاملة الوالدية، والذكاء، والضغوط النفسية، والاتجاه نحو الدراسة) والمتغير التابع (مستوى الطموح) لدى طلاب جامعة الملك فيصل

الفروض الإحصائية :

ما الفرق بين الفروض البحثية والفروض الإحصائية ؟!

الفروض البحثية هي الفروض التي يصيغها الباحث بنفسه في ضوء اطلاعه على الخلفية النظرية ونتائج الدراسات السابقة، وبناء على اطلاعه يحدد اتجاه الفرض هل هو فرض بديل موجه أم فرض بديل غير موجه أم فرض صفري.

أما **الفروض الإحصائية** فتهدف إلى تفسير نتيجة معالجة الأسلوب الإحصائي للفرض البحثي، والذي بناء عليه نتقبل الفرض البحثي أو نرفضه، وبالتالي فالذي يجعلنا نقبل الفرض البحثي ليس الأسلوب الإحصائي فقط ولكن الفرض الإحصائي المرتبط به.

المحاضرة السادسة : مربع كاي كا^2

كا^2 ← التكرارات المشاهدة أو الملاحظة + التكرار النظري أو المتوقع

التكرارات المشاهدة أو الملاحظة: التكرارات الناتجة من التجربة الفعلية، أي التكرارات التي حصل عليها الباحث باستخدام منهج البحث الملائم سواء عن طريق الملاحظة أو التجريب

التكرار النظري أو المتوقع: افتراض من الباحث قائم على أساس معين يحدده الباحث أو تأمل نظري مستقل عن البيانات التي حصل عليها الباحث

هل تحب الإحصاء؟

من المتوقع أن يجيب ٥٠ منهم بـ (نعم) ويجيب ٥٠ الآخرين بـ (لا) وهذا ما يسمى بالتكرار المتوقع Expected Frequency حيث إن:

$$\text{التكرار المتوقع} = \frac{\text{عدد أفراد العينة}}{\text{عدد الاستجابات}}$$

لكن ما حدث أن أجاب ٢٠ منهم بـ (نعم) ، وأجاب ٨٠ بـ (لا) وهذا ما يسمى بالتكرار المشاهد أو الملاحظ Observed Frequency

يتعامل اختبار مربع كاي مع تكرارات البيانات الاسمية

هل يتعامل اختبار مربع كاي مع تكرارات البيانات الفترية أو الرتبية ؟

السؤال : هل توجد فروق بين من قالوا نعم وبين من قالوا لا؟

اختبار كا^2 *Chi-squared Test* هو أحد اختبارات الدلالة الإحصائية اللابارامترية .

يتعامل مع تكرارات الدرجات وليس الدرجات نفسها، ويستخدم في دراسة الفروق بين تكرارات استجابات أفراد عينة ما على سؤال أو عدة أسئلة.

ويتم حساب اختبار (كا^2) من المعادلة التالية : $x^2 = \sum \frac{(O-E)^2}{E}$

حيث: O : التكرار المشاهد *Observed Frequency*

E : التكرار المتوقع *Expected Frequency*

$$\text{كا}^2 = \text{مجـ} \frac{(\text{تو} - \text{تم})^2}{\text{تو}}$$

حيث :

تو : هو التكرار الواقعي الذي يحدث بالفعل والموجود بالجدول .

تم : هو التكرار المتوقع حدوثه ويختلف حسابه بـاختلاف نـوع الجدول المطلوب حساب كا^2 منه .

مثال :

الرأي	موافق	لا أدري	معارض	مج
التكرار	12	2	16	30

حساب التكرار المتوقع (تم) :

$$\text{تم} = \frac{16 + 2 + 12}{3} = 10$$

تو	تم	تو - تم	$(\text{تو} - \text{تم})^2$	$\frac{(\text{تو} - \text{تم})^2}{\text{تم}}$
12	10	2	4	0.4
2	10	8−	64	6.4
16	10	6	36	3.6
−	−	−	مجموع	10.4

درجة الحرية = عدد الأعمدة − 1 = 3 − 1 = 2

بالبحث في جداول كا^2 عند درجة حرية = 2 ومستوى دلالة 0.05 نجد قيمة كا^2 الجدولية = 5.991 .

تحديد مدى دلالة كا^2 :

نقارن قيمة كا^2 المحسوبة بقيمة كا^2 الجدولية نجد أن

قيمة كا^2 المحسوبة = 10.4 > قيمة كا^2 الجدولية = 5.991

لذا فان كا^2 دالة إحصائية عند مستوى دلالة 0.05 .

القرار :

نقارن كا٢ المحسوبة بالجدولية، فعندما تكون **قيمة كا٢ المحسوبة أكبر من قيمة كا٢ المجدولة** فإننا **نرفض** الفرضية الصفرية أو فرض العدم والتي تنص على أنه لا توجد أي علاقة بين المتغيرين ونقبل الفرض البديل والتي تثبت وجود علاقة بين المتغيرين تحت الدراسة .

أما إذا كانت **قيمة كا٢ المحسوبة أقل من قيمة كا٢ المجدولة** فإننا **نقبل** الفرضية الصفرية أو فرض العدم .

طريقة آخرى :

الرأي	موافق	لا أدري	معارض	مج
التكرار	12	2	16	30

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

$$x^2 = \frac{(12-10)^2}{10} + \frac{(2-10)^2}{10} + \frac{(16-10)^2}{10}$$

$$x^2 = \frac{4}{10} + \frac{64}{10} + \frac{36}{10}$$

$$x^2 = 10.4$$

تمرين :

قام باحث بتطبيق استبيان على مجموعة من الأفراد لأخذ أراءهم في قضية الدروس الخصوصية وذلك بتوجيه سؤال واحد إليهم: هل توافق على الدروس الخصوصية (نعم – لا ولكن بشرط – لا)، فحصل على التكرارات التالية:

الاستجابة	نعم	لا ولكن بشروط	لا
التكرار	٢١	٥٤	١٤

المطلوب اختبار الفرض البحثي : لا يختلف التكرار التجريبي الذي حصلنا عليه من استجابات الأفراد على قضية الدروس الخصوصية عن التكرار النظري.

مثال :

أراد معلم معرفة علاقة نجاح تلاميذه في المقرر الذي يقوم بتدريسه بأماكنهم في الفصل، فحسب عدد الناجحين في الامتحان وعدد الراسبين وحدد منهم عدد الجالسين في المقاعد الأمامية وعدد الجالسين في المقاعد الخلفية فتوصل إلى الجدول التالي:

	مقاعد أمامية	مقاعد خلفية	المجموع
ناجح	٢٧	٩	٣٦
راسب	٤	٢٠	٢٤
المجموع	٣١	٢٩	٦٠

المطلوب اختبار الفرض البحثي : توجد علاقة بين نجاح التلاميذ في الامتحان وبين أماكنهم في الفصل.

	مقاعد أمامية	مقاعد خلفية	المجموع
ناجح	٢٧	٩	٣٦
راسب	٤	٢٠	٢٤
المجموع	٣١	٢٩	٦٠

ك م (لأي خلية) = حاصل ضرب مجموعي تكرارات الصف والعمود المنتميان إليهما الخلية / المجموع الكلي للتكرارات

	ك و	ك م	(ك و - ك م)	(ك و - ك م)٢	(ك و - ك م)٢ / ك م
ناجح – مقاعد أمامية	٢٧	١٨,٦	٨,٤	٧٠,٥٦	٣,٧٩
ناجح – مقاعد خلفية	٩	١٧,٤	- ٨,٤	٧٠,٥٦	٤,٠٦
راسب – مقاعد أمامية	٤	١٢,٤	- ٨,٤	٧٠,٥٦	٥,٦٩
راسب في مقاعد خلفية	٢٠	١١,٦	٨,٤	٧٠,٥٦	٦,٠٨
المجموع	٦٠	٦٠	صفر		كا٢ = ١٩,٦٢

الطريقة المختصرة لحساب مربع كاي من الجدول التكراري ٢ × ٢

	مقاعد أمامية	مقاعد خلفية	المجموع
ناجح	٢٧ أ	٩ ب	٣٦ ح
راسب	٤ ج	٢٠ د	٢٤ ز
المجموع	٣١ هـ	٢٩ و	٦٠ ن

كا^2 = فاى^2 × ن

حيث :

فاى : هو معامل ارتباط فاى والذى يحسب من العلاقة :

$$\text{فاى} = \frac{\text{أ} \times \text{د} - \text{ب} \times \text{جـ}}{\sqrt{\text{هـ} \times \text{و} \times \text{ز} \times \text{ح}}}$$

$$\text{فاي} = \frac{(٢٧ \times ٢٠) - (٩ \times ٤)}{\sqrt{٣١ \times ٢٩ \times ٣٦ \times ٢٤}}$$

فاي = ٠,٥٧ مربع فاي = ٠,٣٣

كا^2 = ٠,٣٣ × ٦٠
= ١٩,٦٢

المحاضرة السابعة : معامل الارتباط

وعندما نقول **مقاييس العلاقة نعني بذلك تلك المقاييس التي تبين درجة العلاقة والارتباط بين متغيرين أو أكثر** مثلا، كأن يكون الهدف معرفة هل هناك علاقة بين مستوى الإنتاجية وجودة المنتج في مصنع ما؟، أي هل كلما زادت الإنتاجية تقل جودة المنتج أو العكس

معامل الارتباط: هو تعبير يشير إلى المقياس الإحصائي الذي يدل على مقدار العلاقة بين المتغيرات سلبية كانت أم إيجابية، **وتتراوح قيمته بين الارتباط الموجب التام (+١) وبين الارتباط السالب التام (-١)** .

العلاقة الطردية بين المتغيرات : هو تعبير يشير إلى تزايد المتغيرين المستقل والتابع معا، فإذا كانت الإنتاجية مرتفعة، ومستوى الجودة مرتفع، يقال حينئذ أن بينهما ارتباط موجب، **وأعلى درجة تمثله هي** (+١) .

العلاقة العكسية بين المتغيرات : هو تعبير يشير إلى تزايد في متغير يقابله تناقص في المتغير الآخر، فإذا كانت الإنتاجية منخفضة ومستوى الجودة مرتفع، يقال حينئذ أن بينهما ارتباط سالب، **وأعلى درجة تمثله هي** (-١) .

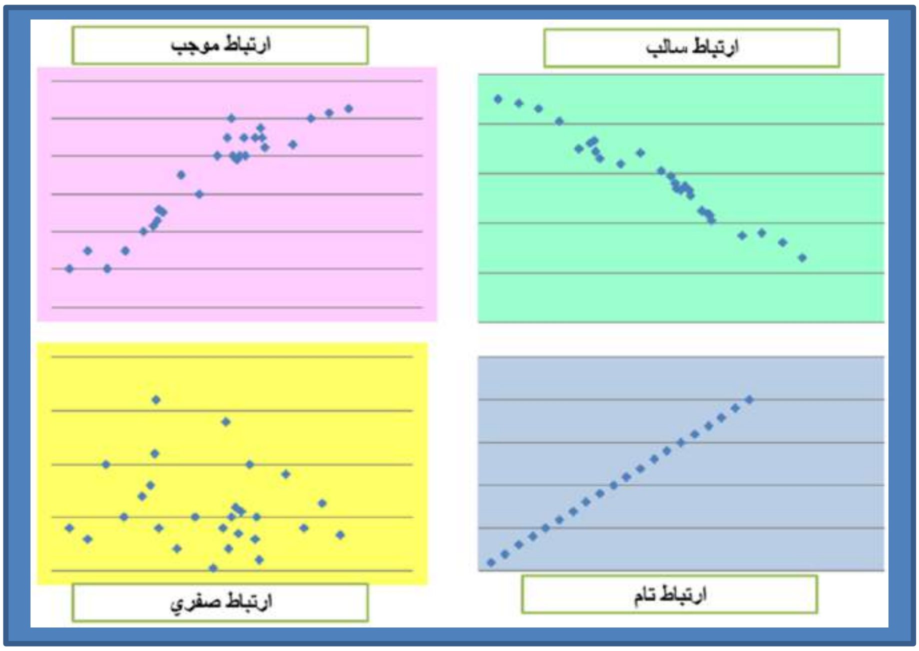
إن معامل الارتباط التام الموجب (+١) يعنى التغير في اتجاه واحد في كلا الظاهرتين مع بقاء الأوضاع النسبية لوحدات الظاهرة ثابتة، سواء كان هذا التغير في اتجاه الزيادة (أي زيادة قيم الظاهرة الأولى تتبعها زيادة في قيم الظاهرة الأخرى)، أو في اتجاه النقص (أي نقص قيم الظاهرة الأولى يتبعها نقص في قيم الظاهرة الأخرى) .

طرق التعرف على العلاقة بين متغيرين وحسابها

أولاً : طريقة شكل الانتشار Scatter Diagram :

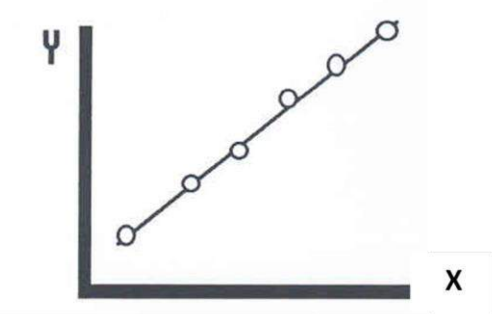
هناك وسيلة مبدئية يعرف الباحث من خلالها نوع الارتباط بين المتغيرين وما إذا كان الارتباط قوياً وضعيفاً أو منعدماً، وما إذا كانت العلاقة خطية أو غير خطية، موجبة أو سالبة. هذه الوسيلة هي " شكل الانتشار " والتي تصلح إذا كان المتغيران كميين. وجدير بالذكر أن هذه وسيلة مبدئية تساعد فقط في معرفة نوع الارتباط ولا تعتبر بديلاً عن الطرق الإحصائية التي سوف نتناولها بالتفصيل في هذه المحاضرة.

والمقصود بشكل الانتشار هو تمثيل قيم الظاهرتين بيانياً على المحورين، المتغير الأول X على المحور الأفقي، والمتغير الثاني Y على المحور الرأسي، حيث يتم تمثيل كل زوج Pair من القيم بنقطة، فنحصل على شكل يمثل كيفية انتشار القيم على المستوى، وهو الذي يسمى شكل الانتشار. وطريقة انتشار القيم تدل على وجود أو عدم وجود علاقة بين المتغيرين ومدى قوتها ونوعها. فإذا كانت تتوزع بشكل منتظم دل ذلك على وجود علاقة (يمكن استنتاجها)، أما إذا كانت النقط مبعثرة ولا تنتشر حسب نظام معين دل ذلك على عدم وجود علاقة بين المتغيرين أو أن العلاقة بينهما ضعيفة. **والأشكال التالية تظهر بعض أشكال الانتشار المعروفة :**

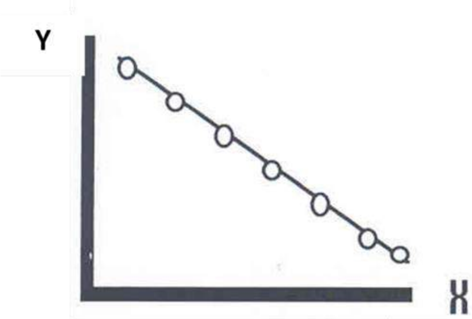


الشكل الأول :

إذا وقعت جميع النقاط على خط مستقيم، دل ذلك على أن العلاقة بينهما خطية وأنها ثابتة أو تامة. وهذه تمثل أقوى أنواع الارتباط بين المتغيرين " ارتباط تام ". فإذا كانت العلاقة طردية فإن " الارتباط طردي تام " كما في الشكل الأول (أ). ومثاله العلاقة بين الكمية المشتراة من سلعة والمبلغ المدفوع لشراء هذه الكمية. أما إذا كانت العلاقة عكسية (وجميع النقاط تقع على خط مستقيم واحد فإن "الارتباط عكسي تام" كما في الشكل الأول (ب). ومثال على ذلك العلاقة بين السرعة والزمن.



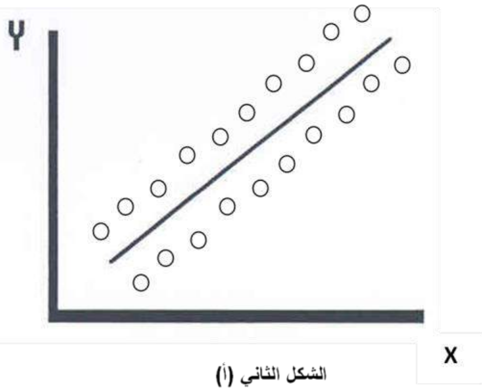
الشكل الأول (أ) ارتباط طردي تام (موجب)



الشكل الأول (ب) ارتباط عكسي تام (سالب)

الشكل الثاني :

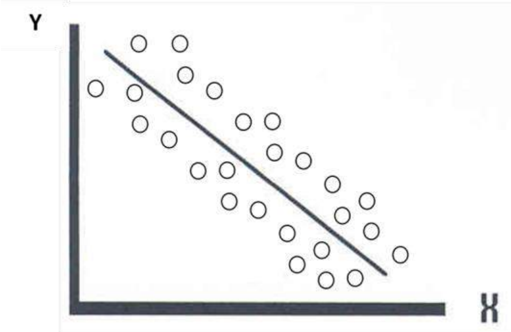
أما إذا كانت النقاط تأخذ شكل خط مستقيم ولكن لا تقع جميعها على الخط قيل أن العلاقة خطية (موجبة أو سالبة) كما في الشكل الثاني أ، ب.



الشكل الثاني (أ)

ارتباط موجب قوي

(ارتباط خطي طردي)



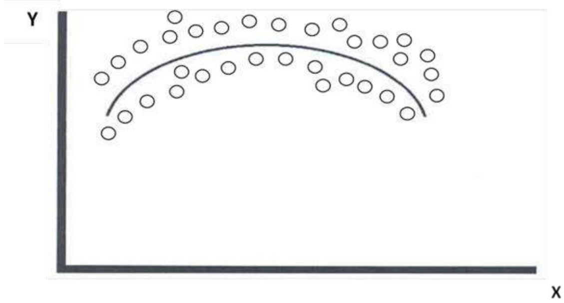
الشكل الثاني (ب)

ارتباط سالب قوي

(ارتباط خطي عكسي)

الشكل الثالث :

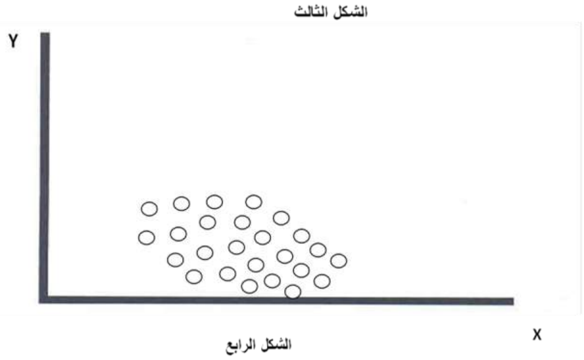
وإذا كانت العلاقة تأخذ شكل منحنى فإن الارتباط لا يكون خطياً "ارتباط غير خطي " Non Linear Correlation كما في الشكل الثالث :



الشكل الثالث

الشكل الرابع :

أما إذا كانت النقاط تتبعثر بدون نظام معين فإن ذلك يدل على عدم وجود علاقة بين المتغيرين (أو أن العلاقة بينهما ضعيفة جداً) كالعلاقة مثلاً بين دخل الشخص وطوله كما في الشكل الرابع :



الشكل الرابع

(لا توجد علاقة)

ثانياً : معامل الارتباط Correlation Coefficient :

يقاس الارتباط بين متغرين بمقياس إحصائي يسمى " معامل الارتباط " ويعكس هذا المقياس درجة أو قوة العلاقة بين المتغيرين واتجاه هذه العلاقة. وتنحصر قيمة معامل الارتباط بين + ١، - ١.

– فإذا كانت قيمة معامل الارتباط تساوي + ١ فمعنى ذلك أن الارتباط بين المتغيرين طردي تام، وهو أقوى أنواع الارتباط الطردي بين متغيرين.

– وإذا كانت قيمة معامل الارتباط تساوي – ١ فمعنى ذلك أن الارتباط بين المتغيرين عكسي تام، وهو أقوى أنواع الارتباط العكسي بين متغيرين.

– وإذا كانت قيمة معامل الارتباط تساوي صفر، فمعنى ذلك أنه لا يوجد ارتباط بين المتغيرين.

– وكلما اقتربت قيمة معامل الارتباط من + ١ أو – ١ كلما كان الارتباط قوياً، وكلما اقترب من الصفر كلما كان الارتباط ضعيفاً.

والخلاصة :

أنه كلما كانت العلاقة قوية بين المتغيرين كلما اقترب معامل الارتباط من + ١ أو – ١ فإذا وصلت قيمة المعامل إلى + ١ أو – ١ كان الارتباط تاماً بين المتغيرين. وأنه كلما كانت العلاقة ضعيفة بين المتغيرين كلما اقترب معامل الارتباط من الصفر، فإذا وصلت قيمة المعامل إلى الصفر كان الارتباط منعدماً بين المتغيرين. ومعنى ذلك أيضاً أنه لا يوجد ارتباط بين متغيرين تكون قيمة المعامل فيه أكبر من + ١ ولا أصغر من – ١. ويمكن تمثيل قوة العلاقة بالشكل التالي:

قيمة معامل الارتباط	المعنى
+1	ارتباط طردي تام
من ٠,٧٠ إلى ٠,٩٩	ارتباط طردي قوي
من ٠,٥٠ إلى ٠,٦٩	ارتباط طردي متوسط
من ٠,٠١ إلى ٠,٤٩	ارتباط طردي ضعيف
0	لا يوجد ارتباط

يمكن تفسير الارتباط العكسي بنفس الطريقة مع المعاملات السالبة

معامل بيرسون للارتباط الخطى البسيط Simple Correlation :

يفترض بيرسون Pearson أن المتغيرين كميان، وأن العلاقة بينهما خطية (أي تأخذ شكل خط مستقيم، ويرى بيرسون أن أفضل مقياس للارتباط بين متغيرين قد يختلفان في وحدات القياس و / أو في مستواهما العام (مثل الارتباط بين العمر والدخل) حيث يقاس العمر بالسنوات ويقاس الدخل بالعملة، بالريال أو الدولار.. كما أن المستوى العام للعمر – أي متوسط العمر – قد يساوي أربعين عاماً. فبينما المستوى العام – أي متوسط – الدخل الشهري قد يكون خمسة آلاف ريال مثلاً.

وبالتالي فإن أفضل مقياس للارتباط بين مثل هذين المتغيرين – حسب رأي بيرسون – هو عن طريق حساب انحرافات كل من المتغيرين عن وسطه الحسابي وقسمة هذه الانحرافات على الانحراف المعياري لكل منهما، فنحصل على ما يسمى بالوحدات المعيارية لكل متغير. ويكون معامل ارتباط بيرسون هو " متوسط حاصل ضرب هذه الوحدات المعيارية ". ومعامل الارتباط يكون بدون تمييز.

ويتم حساب معامل ارتباط بيرسون من خلال العلاقة التالية:

$$r = \frac{\sum XY - \frac{(\sum X)(\sum Y)}{n}}{\sqrt{\left(\sum X^2 - \frac{(\sum X)^2}{n}\right)\left(\sum Y^2 - \frac{(\sum Y)^2}{n}\right)}}$$

حيث :

$\sum XY$ تعني مجموع حاصل ضرب كل قيمة من X في Y.

$(\sum X)$ تعني مجموع قيم المتغير X.

$(\sum Y)$ تعني مجموع قيم المتغير Y.

$\sum X^2$ تعني مجموع مربع قيم المتغير X.

$(\sum X)^2$ تعني مربع مجموع قيم المتغير X.

$\sum Y^2$ تعني مجموع مربع قيم المتغير Y.

$(\sum Y)^2$ تعني مربع مجموع قيم المتغير Y.

n عدد قيم الدراسة (عدد الأزواج المطلوب حساب الارتباط بينها).

مثال :

رغبة إحدى الشركات معرفة العلاقة بين عدد ساعات العمل لموظفيها ومستوى الإنتاجية لهم ، فقاموا بجمع معلومات عن هذا الموضوع وحصلوا على النتائج التالية :

الموظفين	أ	ب	ج	د	هـ	و	ز	ح	ط	ي
ساعات العمل X	٨	٢	٨	٥	١٥	١١	١٣	٦	٤	٦
مستوى الإنتاجية Y	٣	١	٦	٣	١٤	١٢	٩	٤	٤	٥

المطلوب : حساب معامل ارتباط بيرسون للبيانات السابقة .

الموظفين	ساعات العمل X	مستوى الإنتاجية Y	X^2	Y^2	XY
أ	٨	٣	٦٤	٩	٢٤
ب	٢	١	٤	١	٢
ج	٨	٦	٦٤	٣٦	٤٨
د	٥	٣	٢٥	٩	١٥
هـ	١٥	١٤	٢٢٥	١٩٦	٢١٠
و	١١	١٢	١٢١	١٤٤	١٣٢
ز	١٣	٩	١٦٩	٨١	١١٧
ح	٦	٤	٣٦	١٦	٢٤
ط	٤	٤	١٦	١٦	١٦
ي	٦	٥	٣٦	٢٥	٣٠
المجموع	٧٨	٦١	٧٦٠	٥٣٣	٦١٨

نطبق معادلة معامل ارتباط بيرسون كالتالي :

$$r = \frac{\sum XY - \frac{(\sum X)(\sum Y)}{n}}{\sqrt{\left(\sum X^2 - \frac{(\sum X)^2}{n}\right)\left(\sum Y^2 - \frac{(\sum Y)^2}{n}\right)}}$$

$$r = \frac{618 - \frac{(78)(61)}{10}}{\sqrt{\left(760 - \frac{(78)^2}{10}\right)\left((533) - \frac{(61)^2}{10}\right)}} = \frac{618 - 475.8}{\sqrt{(760 - 608.4)(533 - 372.1)}}$$

$$= \frac{142.2}{\sqrt{(151.6)(160.9)}} = \frac{142.2}{\sqrt{24392.44}} = \frac{142.2}{156.2} = 0.91$$

وهذه النتيجة توضح أن درجة الارتباط = ٠.٩١ وهذه النتيجة تعتبر مؤشر على علاقة إيجابية قوية بين ساعات العمل ومستوى الإنتاجية.

ملاحظة مهمة :

من خواص معامل بيرسون للارتباط الخطي أنه لا يتأثر بالعمليات الحسابية التي تجري على المتغيرين y ، x. بمعنى أنه لا يتأثر بالطرح (أو الجمع)، ولا بالقسمة (أو الضرب). أي إذا طرحنا (أو جمعنا) **قيمة معينة** من كل قيم x وقيمة أخرى من كل قيم y، أو قسمنا (أو ضربنا) قيم x على قيمة معينة وكل قيم y على قيمة أخرى فإن قيمة معامل الارتباط لا تتغير أي نحصل على القيمة نفسها .

معامل ارتباط الرتب *Rank Correlation* :

قد يرغب الباحث في حساب معامل الارتباط بين رتب المتغيرين وليس بين القيم ذاتها، فقد يكون المتغيران **وصفيين ترتيبين** Ordinal **أو أن يكون أحد المتغيرين كمياً بينما الأخر وصفياً ترتيبياً**، أو أن يكون **المتغيران كميين**، ويكون اهتمام الباحث منصباً على الرتب أكثر من القيم. ففي انتخابات مجلس الشيوخ أو النواب الأمريكي مثلاً، يعتبر المرشح الأول هو من حصل على أعلى الأصوات بغض النظر عن عددها، والذي يحصل على عدد أصوات أقل منه مباشرة هو الثاني.. وهكذا.

فإذا كانت رتب المتغيرين تسير في الاتجاه نفسه: بمعنى أن الرتب الأعلى للمتغير الأول تناظرها رتب أعلى للمتغير الثاني كانت العلاقة طردية بينهما. وإذا كانت الرتب الأعلى للمتغير الأول تناظرها رتب أدنى للمتغير الثاني كانت العلاقة بينهما عكسية. ففي مثالنا السابق عن العلاقة بين دخل الناخب وعمره، كان الناخب الأكبر عمراً (بصفة عامة) هو الأعلى دخلاً، فمن الواضح أن العلاقة بينهما طردية، أما إذا كان الناخب الأكبر عمراً (بصفة عامة) هو الأقل مشاركة في العمل السياسي فإننا في هذه الحالة نكون أمام علاقة عكسية.

ولحساب معامل ارتباط الرتب هناك طرق مختلفة أهمها معاملي سبيرمان وكيندال.

حساب معامل سبيرمان لارتباط الرتب *Spearman rank Correlation Coefficient* :

لحساب معامل سبيرمان لارتباط الرتب نقوم بترتيب كل من المتغيرين ترتيباً تصاعدياً أو تنازلياً (أما تصاعدياً لكلا المتغيرين أو تنازلياً لكليهما). وفي حالة الترتيب التصاعدي تأخذ أقل قيمة من قيم المتغير الرتبة رقم ١، والقيمة الأعلى منها مباشرة الرتبة رقم ٢ وهكذا (بالنسبة لكل من المتغيرين). أما في حالة الترتيب التنازلي تأخذ أكبر قيمة من قيم المتغير الرتبة رقم ١، والقيمة الأقل منها مباشرة الرتبة رقم ٢ وهكذا (بالنسبة لكل من المتغيرين). وعند تساوي قيمتين (أو أكثر) من قيم المتغير نعطي كل قيمة رتبة مختلفة (كما لو كانت القيم غير متساوية) ثم نحسب متوسط هذه الرتب، ويعطى هذا المتوسط لكل من هذه القيم المتساوية.

وبعد ترتيب المتغيرين نحسب الفروق بين رتب كل من المتغيرين (ونرمز للفروق بالرمز d) ثم نقوم بتربيع هذه الفروق ونحصل على مجموعها أي نحصل على $\sum d^2$ ثم نعوض في معامل سبيرمان لارتباط الرتب والذي يأخذ الشكل التالي :

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

حيث : $\sum d^2$ هو مجموع مربعات الفروق بين رتب المتغيرين n هي عدد أزواج القيم.

مما سبق نستطيع إجمال بعضاً من الملاحظات فيما يلي :

١ – مجموع الفروق بين الرتب يساوي صفر.

٢ – أن قيمة معامل ارتباط الرتب تنحصر بين – ١، + ١ فإذا كانت الرتبة رقم ١ للمتغير الأول تناظرها الرتبة ١ للمتغير الثاني، والرتبة ٢ للمتغير الأول تناظرها الرتبة رقم ٢ للمتغير الثاني، وهكذا.. فإن معامل ارتباط الرتب يساوي + ١ (ارتباط طردي تام بين الرتب).

وإذا كانت الرتبة رقم ١ (أقل رتبة) للمتغير الأول تناظرها أعلى رتبة للمتغير الثاني وهكذا.. فإن معامل ارتباط الرتب يساوي – ١ (ارتباط عكسي تام بين الرتب).

٣ – كذلك نلاحظ أن مجموع الرتب لكل من المتغيرين تساوي

مثال :

البيانات التالية تمثل إجابات عينة من سبعة أشخاص حول برامج الضمان الاجتماعي، ومدى ملاءمتها لحاجات الناس.

السؤال الأول	جيدة	مقبولة	ممتازة	جيدة	جيدة جداً	مقبولة	جيدة
السؤال الثاني	جيدة جداً	مقبولة	جيدة جداً	جيدة	جيدة	جيدة	ممتازة

والمطلوب : حساب معامل سبيرمان لارتباط الرتب بين هذين السؤالين ؟

السؤال الأول X	السؤال الثاني Y	رتب X	رتب Y	d الفرق بين الرتب	d^2 مربعات الفرق
جيدة	جيدة جداً	4	2.5	1.5	2.25
مقبولة	مقبولة	6.5	7	- 0.5	0.25
ممتازة	جيدة جداً	1	2.5	- 1.5	2.25
جيدة	جيدة	4	5	- 1.0	1.00
جيدة جداً	جيدة	2	5	- 3.0	9.00
مقبولة	جيدة	6.5	5	1.5	2.25
جيدة	ممتازة	4	1	3.0	9.00
المجموع				Zero	26.0

$$r_s = 1 - \frac{6\left(\sum d^2\right)}{n\left(n^2-1\right)} = 1 - \frac{6\,(26)}{7\,(49-1)}$$

$$= 1 - \frac{156}{336} = 1 - 0.46$$

$$r_s = 0.54$$

وهذا يعني أن الارتباط بين إجابات المبحوثين بالنسبة للسؤالين هو ارتباط طردي متوسط. وبالتالي فليس بالضرورة أن يكون رأي المجيبين في برامج الضمان الاجتماعي تعني ملاءمتها لحاجات الناس.

لدراسة العلاقة بين تقدير الطالبة في الإحصاء و تقديرها في الرياضيات ، اخترنا خمس طالبات و كانت تقديراتهم كالتالي:

تقدير الإحصاء X	A	C	D	F	A
تقدير الرياضيات y	B	C	B	D	A

x	y	رتب x	رتب y	d	d^2
A	B	5	4	1	1
C	C	3	3	0	0
D	B	2	4	-2	4
F	D	1	2	-1	1
A	A	5	5	0	0
$\sum$					6

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

$$r_s = 1 - \frac{6\times 6}{5(5^2-1)} = 1 - \frac{36}{5\times 24} = 1 - \frac{36}{120}$$

$$= 1 - 0.3 = 0.7$$

معامل بوينت بايسيريال Point Biserial للارتباط :

يستخدم لقياس الارتباط بين متغير كمي X و متغير اسمي ٢مستويين (نعم –لا) أو (ذكر – أنثى) و غيرها.

اشارة معامل الارتباط ليس لها معنى في حالة المتغيرات الوصفية فتقاس قوة العلاقة و ليس اتجاهها.

أوجد قيمة معامل الارتباط بين مشاركة الطالبة فى المحاضرة و درجتها فى الاختبار للبيانات التالية :

y المشاركة	نعم	نعم	نعم	لا	لا
X درجة الاختبار	15	19	20	15	11

$$n_1 = 3$$

$$\overline{x_1} = \frac{15+19+20}{3} = \frac{54}{3} = 18$$

$$n_2 = 2$$

$$\overline{x_2} = \frac{15+11}{2} = \frac{26}{2} = 13$$

$$r_{pb} = \frac{\overline{x}_1 - \overline{x}_2}{s_x}\sqrt{\frac{n_1 n_2}{n(n-1)}}$$

$$r_{pb} = \frac{18-13}{13}\sqrt{\frac{3\times 2}{5\times 4}} = \frac{5}{13}\sqrt{\frac{6}{20}} \approx 0.21$$

أوجد قيمة معامل الارتباط بين الإجابة على السؤال الإجباري y و الدرجة الإجمالية y لستة من الطلاب، حيث ١ تعني الإجابة على السؤال و ٠ تعني عدم الإجابة.

Y	1	1	1	0	0	0
X	14	16	19	11	7	8

$s_x = 4,68$

$$n_1 = 3 \qquad \overline{x_1} = \frac{14+16+19}{3} = 16.3$$

$$n_2 = 3 \qquad \overline{x_2} = \frac{11+7+8}{3} = 8.67$$

$$r_{pb} = \frac{\overline{x}_1 - \overline{x}_2}{s_x}\sqrt{\frac{n_1 n_2}{n(n-1)}}$$

$$r_{pb} = \frac{16.3-8.67}{4.68}\sqrt{\frac{3\times3}{6\times5}} = 1.63 \times \sqrt{\frac{9}{30}} \approx 0.893$$

معامل الاقتران (معامل فاي) Phi :

- يستخدم للعلاقة بين متغيرين اسميين كل منهما ثنائي التقسيم.
- اشارة معامل فاي ليس لها معنى فهو يقيس قوة العلاقة دون اتجاهها.

	X1	X2	Sum
Y1	a	b	a+b
Y2	c	d	C+d
Sum	a+c	b+d	a+b+c+d

$$r_{\emptyset} = \frac{a\times d \; - b\times C}{\sqrt{(a+b)(c+d)(a+c)(b+d)}}$$

أوجدي قيمة معامل الاقتران بين النوع (ذكر/ أنثى) و بين الاصابة بمرض الاكتئاب (مصاب/ غير مصاب) للبيانات التالية :

	lwhf	غير مصاب	المجموع
ذكر	12	7	19
أنثى	10	5	15
المجموع	22	12	36

$$r_{\emptyset} = \frac{12 \times 5 \; - 7 \times 10}{\sqrt{22 \times 12 \times 19 \times 15}} =$$

$$= \frac{60 - 70}{\sqrt{75240}} = \frac{-10}{274.299} = -0.037$$

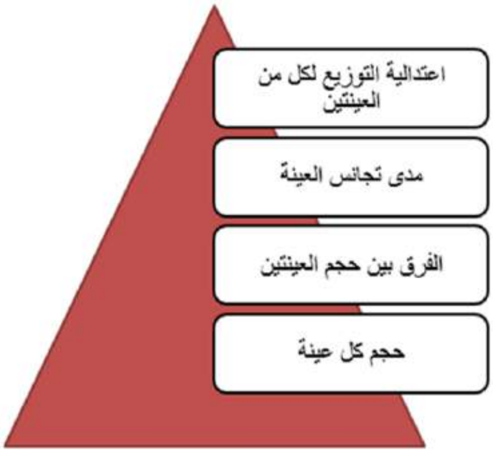
معاملات الارتباط

- معامل بيرسون — كلا المتغيرين كمي
- معامل بوينت بايسيريال — متغير كمي+ متغير اسمي في مستوبين
- معامل سبيرمان — متغيرين كميين أو متغيرين وصفيين ترتيبين
- معامل فاي — متغيرين اسميين في مستوين

المحاضرة الثامنة : اختبار «ت» *t. test*

يعد اختبار (ت) من أكثر اختبارات الدلالة شيوعاً في الأبحاث النفسية والتربوية. ويستخدم لقياس دلالة فروق المتوسطات غير المرتبطة والمرتبطة، للعينات المتساوية وغير المتساوية.

شروط استخدام اختبار (ت) لدلالة فوق المتوسطات :



١. حجم كل عينة :

الأصل في اختبار (ت) أنه من مقاييس دلالة العينات الصغيرة ولكن هذا لا يحول دون استخدام (ت) للعينات الكبيرة.

- العينة الصغيرة هي التي يقل حجمها عن ٣٠
- العينة الكبيرة هي التي يزيد حجمها عن ٣٠
- في حالة العينات الصغيرة جداً يتم استخدام البدائل اللابارامترية للدلالة التي تصلح للتوزيعات الحرة غير المقيدة باعتدالية التوزيع.

٢. الفرق بين حجم العينتين :

من الأفضل أن يكون حجم العينتين متقارباً فلا يكون حجم أحد العينتين ٤٠٠ وحجم الآخر ٥٠ لأن للحجم أثره على مستوى دلالة (ت).

٣. مدى تجانس العينتين :

يقاس مدى التجانس بالفرق بين تباين العينتين ولا يقاس هذا الفرق بطرح التباين الأصغر من التباين الأكبر، وإنما يقاس بقسمة التباين الأكبر على التباين الأصغر.

$$\text{ف} = \frac{\text{التباين الاكبر}}{\text{التباين الأصغر}} \qquad \text{ف} = \frac{\text{ع}_1^2}{\text{ع}_2^2}$$

يتحقق التجانس بين العينتين عندما تصبح ف مساوية للواحد الصحيح، أي عندما يصبح التباين الكبير مساوياً للتباين الصغير.

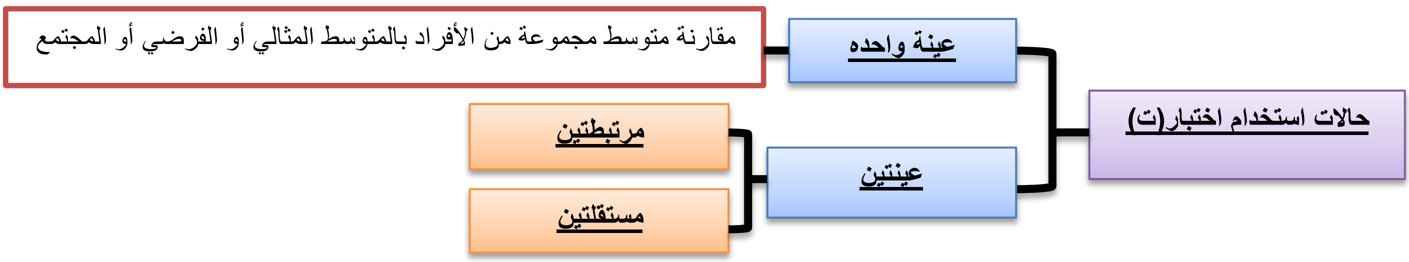
٤. مدى اعتدالية التوزيع التكراري للعينتين :

نعني بمدى الاعتدالية تحرر التوزيع التكراري من الالتواء، والالتواء اما أن يكون سالباً أو موجباً.

التوزيع الاعتدالي لا التواء له، ويمتد من – ٣ إلى + ٣ مقياس الالتواء التالي:

$$\text{الإلتواء} = \frac{3\,(\text{المتوسط} - \text{الوسيط})}{\text{الانحراف المعياري}}$$

كلما اقترب الالتواء من الصفر كان التوزيع اعتدالياً، لأن المتوسط في التوزيع الاعتدالي يساوي الوسيط.



استخدام اختبار ت للتعرف على دلالة الفرق بين متوسط عينة ما ومحك ثابت :

$$ت = \frac{م - س}{خ_م}$$

حيث أن ت تمثل النسبة التائية، م متوسط العينة،

س متوسط المجتمع أو المحك، خ م الخطأ المعياري للمتوسط .

درجات الحرية = ن – ١

مثال :

طبق باحث اختبار في اللغة الانجليزية على مجموعة من المفحوصين عددهم (٢٠) مفحوصاً، فحصل على البيانات التالية:

٦٢	٤٨	٣٠	٣٨	٣٩	٤٠	٤٦	٢٢	٤٠	٣٨
٥٠	١٩	٤١	٤٢	٧٢	١٧	٦٦	٢٤	٣٥	٤٥

المطلوب اختبار الفرض البحثي : يختلف متوسط درجات المجموعة في اللغة الانجليزية عن الدرجة ٣٩.

٦٢	٤٨	٣٠	٣٨	٣٩	٤٠	٤٦	٢٢	٤٠	٣٨
٥٠	١٩	٤١	٤٢	٧٢	١٧	٦٦	٢٤	٣٥	٤٥

$$ت = \frac{م - س}{خ_م} \qquad م = \frac{٨١٤}{٢٠} = ٤٠,٧$$

$$\text{الخطأ المعياري للمتوسط} = \frac{\text{الانحراف المعياري}}{\text{الجذر التربيعي لحجم العينة}} \qquad \hat{\sigma}_{\bar{X}} = \frac{S}{\sqrt{n}}$$

x	$d = x - \bar{x}$	d^2
62	62 – 40.7 = 21.3	453.69
48	48 – 40.7 = 7.3	53.29
30	30 – 40.7 = -10.7	114.49
38	38 – 40.7 = -2.7	7.29
39	39 – 40.7 = -1.7	2.89
40	40 – 40.7 = -0.7	0.49
46	46 – 40.7 = 5.3	28.09
22	22 – 40.7 = -18.7	349.69
40	40 – 40.7 = -0.7	0.49
38	38 – 40.7 = -2.7	7.29
50	50 – 40.7 = 9.3	86.49
19	19 – 40.7 = -21.7	470.89
41	41 – 40.7 = 0.3	0.09
42	42 – 40.7 = 1.3	1.69
72	72 – 40.7 = 31.3	979.69
17	17 – 40.7 = -23.7	561.69
66	66 – 40.7 = 25.3	640.09
24	24 – 40.7 = -16.7	278.89
35	35 – 40.7 = -5.7	32.49
45	45 – 40.7 = 4.3	18.49
		4088.2

$$s^2 = \frac{\sum d^2}{n} = \frac{4088.2}{20} = 204.41$$

$$s = \sqrt{s^2} = \sqrt{204.41} \cong 14.30$$

$$\text{الخطأ المعياري للمتوسط} = \frac{١٤,٣٠}{\sqrt{٢٠}} = ٣,٢٠$$

$$ت = \frac{م - س}{خ_م} = \frac{٤٠,٧ - ٣٩}{٣,٢٠} = ٠,٥٣$$

يعتمد تطبيق اختبار ت لحساب دلالة الفروق بين متوسطات درجات العينات على حساب درجتين لاختبار ت :

الأولى : تسمى القيمة **المحسوبة** لاختبار ت يتم حسابها من خلال معادلة خاصة.

الثانية : تسمى القيمة **الجدولية** لاختبار ت، ويتم حسابها من جدول يسمى جدول ت.

ويعتمد الكشف في هذه الجداول على ما يسمى بـ"درجات الحرية" .

درجات الحرية = **عدد الأفراد – عدد المجموعات**

= **ن – ١**

يتم مقارنة قيمة ت المحسوبة بقيمة ت الجدولية فإذا كانت :

- إذا كانت قيمة ت المحسوبة أكبر من قيمة ت الجدولية فذلك يعنى أن (ت) دالة إحصائيا وذلك يعنى أن الفروق بين المتوسطات فروق حقيقية وجوهرية ولها معنى وليست فروقا ظاهرية .
- أما إذا كانت قيمة ت المحسوبة أقل من الجدولية فذلك يعنى أن (ت) غير دالة إحصائيا وذلك يعنى أن الفروق بين المتوسطات غير جوهرية بل فروق ظاهرية ليس لها أي تأثير .

الحالات التي يستخدم فيها اختبار (ت) لدى عينة واحدة :

يمكن استخدام اختبار (ت) لدى عينة واحدة في حالات كثيرة منها الحالات التالية :

- دراسة الفرق بين متوسط مجموعة من الأفراد في متغير ما والمتوسط المثالي لهذا المتغير.
- دراسة الفرق بين متوسط التحصيل الدراسي لطلاب فصل دراسي معين في مقرر دراسي أو مقررات دراسية معينة والمتوسط العام للتحصيل الدراسي لطلاب المدرسة أو الإدارة التعليمية أو المحافظة في نفس المقرر أو المقررات الدراسية.
- دراسة الفرق بين متوسط ذكاء مجموعة من الطلاب بمدرسة معينة ومتوسط الذكاء العام لدى طلاب المنطقة أو المحافظة التي تقع بها المدرسة .
- المقارنة بين متوسط أداء مجموعة من الأفراد في شيء ما، ومستوى معين لأداء هذا الشيء .

البيانات المطلوب توافرها لاستخدام اختبار (ت) لدى عينة واحدة :

يحتاج استخدام اختبار (ت) لدى عينة واحدة إلى توافر البيانات التالية :

- **البيانات الخام** (أو الدرجات الخام) لدى عينة الأفراد موضع الدراسة، أو (متوسط العينة + الخطأ المعياري لمتوسط العينة)، أو (متوسط العينة + الانحراف المعياري لدرجات العينة + عدد أفراد العينة) .
- **المتوسط المثالي** أو الفرضي لدى المجتمع الذي سنقارن به متوسط العينة .

صياغة الفروض عند استخدام اختبار (ت) لدى عينة واحدة :

عند استخدام اختبار (ت) لدى عينة واحدة يمكن صياغة الفروض التالية :

- H_0 : لا يوجد فرق دال إحصائيًا بين متوسط عينة البحث والمتوسط العام (أو المثالى أو الفرضى) لدى مجتمع البحث فى المتغير (**فرض صفري**).
- H_1 : يوجد فرق دال إحصائيًا بين متوسط عينة البحث والمتوسط العام (أو المثالي أو الفرضي) لدى مجتمع البحث في المتغير (**فرض بديل غير موجه**).

يوجد فرق دال إحصائيًا بين متوسط عينة البحث والمتوسط العام (أو المثالي أو الفرضي) لدى مجتمع البحث في المتغير (س)، لصالح متوسط عينة البحث أو لصالح مجتمع البحث. (فرض بديل موجه) .

حالات استخدام اختبار ت :

عينتان مرتبطتان :

عبارة عن مجموعتين من الدرجات لكنهما ناتجتان عن مجموعة واحدة من الأفراد لكل فرد درجتين على الأقل مثل :

- إجراء قياس قبلي وقياس بعدي لمتغير ما لدى عينة واحدة من الأفراد.
- أو تطبيق اختبارين على مجموعة واحدة أو تطبيق اختبار واحد مرتين على العينة.

اختبار ت لمجموعتين مرتبطتين :

$$ت = \frac{م\,ف}{\sqrt{\frac{مجـ\ ح^2\,ف}{ن\,(ن-1)}}}$$

حيث :

م ف = متوسط الفروق ويحسب من العلاقة : $م\,ف = \frac{مجـ\ ف}{ن}$ ح ف = ف – م ف

ف = س١ – س٢ / س١ درجات الاختبار الأول / س ٢ درجات الاختبار الثاني / ن = عدد الأفراد في أي من الاختبارين

مثال ١ :

الاحصاء الاجتماعي	٢٦	١٨	٢٠	٢٤	٢٢	١٤	٢٣	١٦	٢٢	١١
مشروع التخرج	٢٣	١٦	١٩	٢١	١٨	١٢	٢٤	١١	٢٣	٩

س١	س٢	ف	ح ف	(ح ف)٢
٢٦	٢٣	٣	١	١
١٨	١٦	٢	٠	٠
٢٠	١٩	١	١-	١
٢٤	٢١	٣	١	١
٢٢	١٨	٤	٢	٤
١٤	١٢	٢	٠	٠
٢٣	٢٤	١-	٣-	٩
١٦	١١	٥	٣	٩
٢٢	٢٣	١-	٣-	٩
١١	٩	٢	٠	٠
		٢٠		٣٤

$$ت = \frac{م\,ف}{\sqrt{\frac{مجـ\ ح^2\,ف}{ن\,(ن-1)}}} \qquad م\,ف = \frac{مجـ\ ف}{ن}$$

$$ح\,ف = ف - م\,ف$$

$$ت = \frac{٢}{\sqrt{\frac{٣٤}{١٠\,(١٠-١)}}}$$

$$ت = ٣,٢٥$$

مثال ٢ :

الاحصاء الاجتماعي	١٠	٥	٦	٧	١٠	٦	٧	٨	٦	٥
مشروع التخرج	٧	٣	٧	٥	٨	٤	٥	٢	٣	٦

س١	س٢	ف	ح ف	(ح ف)٢
١٠	٧	٣	١	١
٥	٣	٢	٠	٠
٦	٧	١-	٣-	٩
٧	٥	٢	٠	٠
١٠	٨	٢	٠	٠
٦	٤	٢	٠	٠
٧	٥	٢	٠	٠
٨	٢	٦	٤	١٦
٦	٣	٣	١	١
٥	٦	١-	٣-	٩
		٢٠		٣٦

$$ت = \frac{م_ف}{\sqrt{\frac{مجـ\ ح^2_ف}{ن\,(ن-1)}}}$$

$$ت = \frac{٢}{\sqrt{\frac{٣٦}{١٠\,(١٠-١)}}}$$

$$ت = ٣,١٦$$

عينتين غير مرتبطتين (مستقلتين) : حيث ن١= ن٢

$$ت = \frac{م_1 - م_2}{\sqrt{\frac{ع_1^2 + ع_2^2}{ن-1}}}$$

حيث :

$م_1$ = المتوسط الحسابي للمجموعة الأولى .

$م_2$ = المتوسط الحسابي للمجموعة الثانية .

$ع_1^2$ = تباين المجموعة الأولى .

$ع_2^2$ = تباين المجموعة الثانية .

ن = عدد أفراد العينة الأولى أو الثانية حيث أنهما متساويتان .

عينتين غير مرتبطتين (مستقلتين) :

عبارة عن مجموعتين من الدرجات ناتجة عن مجموعتين مستقلتين من الأفراد مثل (المجموعة التجريبية والمجموعة الضابطة؛ أو الذكور والإناث؛ أو القسم العلمي والقسم الأدبي).

ذكور	٧	٤	٥	٣	٨	٦	٢
إناث	٣	٥	١٥	٢	١٠	١٣	١

المجموعة الأولى [n = 7]

x	d = x − x̄	d²
7	2	4
4	-1	1
5	0	0
3	-2	4
8	3	9
6	1	1
2	-3	9
35		28

$\bar{x} = \frac{\sum x}{n} = 5$

$s^2 = \frac{\sum d^2}{n} = \frac{28}{7} = 4$

المجموعة الثانية [n = 7]

x	d = x − x̄	d²
3	-4	16
5	-2	4
15	8	64
2	-5	25
10	3	9
13	6	36
1	-6	36
49		

$\bar{x} = \frac{\sum x}{n} = 7$

$s^2 = \frac{\sum d^2}{n} = \frac{190}{7} = 27.14$

$$ت = \frac{م_1 - م_2}{\sqrt{\frac{ع_1^2 + ع_2^2}{ن-1}}}$$

$$ت = \frac{5 - 7}{\sqrt{\frac{4 + 27.14}{7 - 1}}}$$

ت المحسوبة = -0.88

عينتين غير مرتبطتين (مستقلتين) : حيث ن١≠ ن٢

$$\text{ت} = \frac{\text{م}_1 - \text{م}_2}{\sqrt{\frac{\text{ع}_1^2}{\text{ن}_1} + \frac{\text{ع}_2^2}{\text{ن}_2}}}$$

حيث :

م_1 = المتوسط الحسابي للمجموعة الأولى .

م_2 = المتوسط الحسابي للمجموعة الثانية .

ع_1^2 = تباين المجموعة الأولى .

ع_2^2 = تباين المجموعة الثانية .

ن_1 = عدد أفراد المجموعة الأولى .

ن_2 = عدد أفراد المجموعة الثانية .

العينة الأولى	٣٥	١٧	٢٢	٣٢	١٩	٤٨	١٣	١٩	٢٠
العينة الثانية	١١	٣	٩	١٠	١٤	٢	٧		

المجموعة الأولى [n = 9]

x	$d = x - \bar{x}$	d^2
35	10	100
17	-8	64
22	-3	9
32	7	49
19	-6	36
48	23	529
13	-12	144
19	-6	36
20	-5	25
225		992

$\bar{x} = \frac{\sum x}{n} = 25$

$s^2 = \frac{\sum d^2}{n} = \frac{992}{9} = 110.2$

المجموعة الثانية [n = 7]

x	$d = x - \bar{x}$	d^2
11	3	9
3	-5	25
9	1	1
10	2	4
14	6	36
2	-6	36
7	-1	1
56		112

$\bar{x} = \frac{\sum x}{n} = 8$

$s^2 = \frac{\sum d^2}{n} = \frac{112}{7} = 16$

$$\text{ت} = \frac{\text{م}_1 - \text{م}_2}{\sqrt{\frac{\text{ع}_1^2}{\text{ن}_1} + \frac{\text{ع}_2^2}{\text{ن}_2}}}$$

$$\text{ت} = \frac{٢٥ - ٨}{\sqrt{\frac{١١٠{,}٢}{٩} + \frac{١٦}{٧}}}$$

ت = ٤,٤٦

يعتمد تطبيق اختبارت لحساب دلالة الفروق بين متوسطات درجات العينات على حساب درجتين لاختبارت :

الأولى : تسمى القيمة المحسوبة لاختبار ت يتم حسابها من خلال معادلة خاصة .

الثانية : تسمى القيمة الجدولية لاختبار ت، ويتم حسابها من جدول يسمى جدول ت .

ويعتمد الكشف في هذه الجداول على ما يسمى بـ"درجات الحرية" .

درجات الحرية = **عدد الأفراد – عدد المجموعات**

= ن – ١

يتم مقارنة قيمة ت المحسوبة بقيمة ت الجدولية فإذا كانت :

- إذا كانت قيمة ت المحسوبة أكبر من قيمة ت الجدولية فذلك يعني أن (ت) دالة إحصائيا وذلك يعني أن الفروق بين المتوسطات فروق حقيقية وجوهرية ولها معنى وليست فروقا ظاهرية .
- أما إذا كانت قيمة ت المحسوبة أقل من الجدولية فذلك يعني أن (ت) غير دالة إحصائيا وذلك يعني أن الفروق بين المتوسطات غير جوهرية بل فروق ظاهرية ليس لها أي تأثير .

صياغة الفروض عند استخدام اختبار (ت) لمجموعتين :

مجموعتين مرتبطتين :

Ho : لا توجد فروق دالة إحصائيًا بين متوسطي درجات طلاب قسم الاجتماع في مادتي الإحصاء الاجتماعي ومناهج البحث (**فرض صفري**) .

H1 : توجد فروق دالة إحصائيًا بين متوسطي درجات طلاب قسم الاجتماع في مادتي الإحصاء الاجتماعي ومناهج البحث (**فرض بديل غير موجه**).

مجموعتين مستقلتين :

H_0 : لا توجد فروق دالة إحصائيًا بين متوسطي درجات الذكور والإناث في مقرر الإحصاء الاجتماعي (**فرض صفري**).

H_1 : توجد فروق دالة إحصائيًا بين متوسطي درجات الذكور والإناث في مقرر الإحصاء الاجتماعي (**فرض بديل غير موجه**).

المحاضرة العاشرة : تحليل التباين *ANOVA*

« لا توجد فروق دالة إحصائياً بين متوسطات درجات طلاب كليات العلوم والآداب والتربية في الذكاء الاجتماعي »

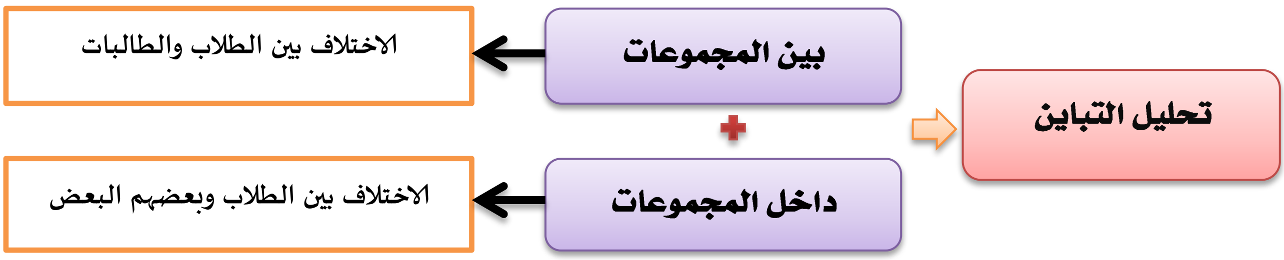
- طلاب كلية العلوم .
- طلاب كلية الآداب .
- طلاب كلية التربية .

الصعوبات :

- عدد أزواج المقارنات بين المتوسطات كبيراً جداً .
- العمليات الحسابية ستكون كبيرة جداً .
- المقارنات الزوجية بين المتوسطات سوف لا تعطينا القرار المطلوب بخصوص مقارنة جميع المتوسطات (جميع المجموعات) في آن واحد .
- عملية المقارنات الزوجية للمعالجات تؤدي إلى زيادة الاختلاف بين تأثيرات المعالجات لأسباب غير الأسباب محل الدراسة، وهذا بالطبع سيزيد من عدم التجانس بين مجموعات المعالجات وبالتالي سيزيد من مقدار الخطأ التجريبي بين المشاهدات .

تحليل التباين (*Analysis of Variance (ANOVA)*)

المعنى العام للتباين	اختلاف الأشياء عن بعضها البعض، هذا الاختلاف هو الذي يجعلنا نميز بين هذه الأشياء. أي أن أي مجموعة من الأشياء مختلفة عن بعضها معناها متباينة.
المعنى النفسي للتباين	يتشابه مع معنى الفروق الفردية، أي اختلاف الأفراد عن بعضهم البعض، وأحياناً يكون الاختلاف داخل الأفراد، أي اختلاف مجموعة من الظواهر الاجتماعية أو النفسية .
المعنى الإحصائي للتباين	هو مربع الانحراف المعياري ع٢.
معنى تحليل التباين	تحليل التباين هو البحث عن مكونات هذا الاختلاف (أو التباين). دراسة مكونات الاختلاف بين مجموعة من الأفراد في ظاهرة معينة وحساب نصيب كل مكون بواسطة معادلات إحصائية معينة.



شروط استخدام أسلوب تحليل التباين :

١. وجود مجموعتين من البيانات أو أكثر.
٢. أن تكون البيانات الخاصة بالمجموعات من النوع الفتري.
٣. اعتدالية توزيع بيانات المتغير التابع.
٤. وجود تجانس بين المجموعات الداخلة في التحليل.

أسس تحليل التباين :

١. البحث عن مقدار الاختلاف بين المجموعات.
٢. الأساس الذي تختلف فيه المجموعات وهو ما يسمى (المتغير التابع).
٣. الأساس الذي تقسم على أساسه المجموعات يسمى (المتغير المستقل).

س : هل الاختلافات التي نبحثها في المتغير المستقل أم في المتغير التابع ؟

ج : إن اختلافات الدرجات التي نبحثها تكون في درجات المتغير التابع طبقاً لاختلافات المتغير المستقل، أي أن تحليل التباين (أو الاختلاف) يكون في درجات المتغير التابع وفقاً لطبيعة المتغير المستقل.

أنواع تحليل التباين :

- تحليل التباين أحادي الاتجاه
- تحليل التباين المتعدد
- تحليل التباين ذي القياسات المتعددة

حساب التباين الداخلي (داخل المجموعات) وذلك بحساب مجموع المربعات داخل المجموعات :

مجموع المربعات داخل المجموعات = ن^{١} (ع^{١})^{٢} + ن^{٢} (ع^{٢})^{٢} + ... + ن$_{ن}$ (ع$_{ن}$)^{٢}

حيث :

ن^{١} ، ن^{٢} ، .... ن$_{ن}$ عدد افراد المجموعات ١ ، ٢ ، ٣ ، ... ، ن

(ع^{١})^{٢} ، (ع^{٢})^{٢} ، .... ، (ع$_{ن}$)^{٢} تباين درجات المجموعات ١ ، ٢ ، ٣ ، ... ، ن على الترتيب

حساب التباين الخارجي (بين المجموعات) وذلك بحساب مجموع المربعات بين المجموعات :

مجموع المربعات بين المجموعات = ن^{١} (م$_{١}$ – م)^{٢} + ن^{٢} (م$_{٢}$ – م)^{٢} + ... + ن$_{ن}$ (م$_{ن}$ – م)^{٢}

حيث :

م^{١} ، م^{٢} ، ..... ، م$_{ن}$ متوسطات درجات المجموعات ١ ، ٢ ، ٣ ، ... ن على الترتيب

م = المتوسط الوزني .

المتوسط الوزني = $\dfrac{\text{ن}^{١}\text{م}^{١} + \text{ن}^{٢}\text{م}^{٢} + \ldots \text{ن}_{\text{ن}}\text{م}_{\text{ن}}}{\text{ن}^{١} + \text{ن}^{٢} + \ldots + \text{ن}_{\text{ن}}}$

صياغة الفروض عند استخدام تحليل التباين أحادي الاتجاه :

الفرض الصفري :

«لا توجد فروق دالة إحصائياً بين متوسطات درجات طلاب كليات العلوم والآداب والتربية في الذكاء الاجتماعي»

الفرض بديل :

«توجد فروق دالة إحصائياً بين متوسطات درجات طلاب كليات العلوم والآداب والتربية في الذكاء الاجتماعي»

المحاضرة الحادية عشر : تحليل الانحدار

يعتبر تحليل الانحدار أكثر طرق التحليل الإحصائي استخداماً، حيث يتم من خلاله التنبؤ بقيمة احد المتغيرات (المتغير التابع) عند قيمة محددة لمتغير أو متغيرات أخرى (المتغيرات المستقلة).

وتسمى العلاقة الرياضية التي تصف سلوك المتغيرات محل الدراسة والتي من خلالها يتم التنبؤ بسلوك احد المتغيرين عند معرفة الاخر بمعادلة خط الانحدار.

وهناك صورتان أساسيتان لمعادلة الانحدار وهما :

الصورة الأولى : معادلة انحدار x|y (التى يطلق عليها معادلة انحدار y على x)

الصورة الثانية : معادلة انحدار x|y (التى يطلق عليها معادلة انحدار x على y)

معادلة انحدار x على y :

وهى التي يطلق عليها معادلة انحدار x | y. أي تتحدد قيمة المتغير x تبعا لقيمة المتغير y لذلك يمكن التعبير عن تلك العلاقة الخطية بالمعادلة التالية :

$$\hat{x} = c_0 + c_1 y$$

حيث يسمى C_0 ثابت الانحدار او الجزء الثابت بينما C_1 يطلق عليها معامل الانحدار أو معدل التغير في الدالة .

مثال :

أراد باحث دراسة العلاقة بين عدد مشاركات الطلاب في المحاضرات (متغير مستقل) ودرجاتهم في الاختبار (متغير تابع) ، وكانت الدرجات كما هو موضح بالجدول :

التحصيل	9	7	10	5	3	7	8	10	4	6
عدد المشاركات	12	9	14	6	4	7	10	10	5	8

أوجد :

- قيمة معامل الانحدار
- معادلة انحدار x على y

الحل :

x	y	xy	x^2	y^2
9	12	108	81	144
7	9	63	49	81
10	14	140	100	196
5	6	30	25	36
3	4	12	9	16
7	7	49	49	49
8	10	80	64	100
10	10	100	100	100
4	5	20	16	25
6	8	48	36	64
69	85	650	529	811

من خلال الجدول السابق يمكن تقدير معادلة انحدار x على y كما يلي :

اولاً- يتم تقدير قيمة معامل الانحدار C_1 :

$$c_1 = \frac{n\sum xy - (\sum x)(\sum y)}{n\sum y^2 - (\sum y)^2} = \frac{10(650)-(69)(85)}{10(811)-(85)^2} \cong 0.718$$

ثانياً – تقدير قيمة C_0 :

$$c_0 = \frac{\sum x}{n} - c_1 \frac{\sum y}{n}$$

$$c_0 = \bar{x} - c_1\bar{y} = 6.9 - 0.718(8.5) \cong 0.797$$

معادلة انحدار x على y وهي :

وبالتالي تكون معادلة انحدار x على y

$$\hat{x} = c_0 + c_1 y$$

$$\hat{x} = c_0 + c_1 y \Rightarrow \hat{x} = 0.797 + 0.718y$$

إذا كان عدد المشاركات يساوي ٢٥

فإن التحصيل المتوقع هو:

يتم التعويض في معادلة الانحدار التي سبق إيجادها عندما تكون y = 25 كما يلي :

$$\hat{x} = 0.797 + 0.718(25) = 18.747 \cong 19$$

صياغة الفروض :

الفرض الصفري : «لا يمكن التنبؤ بالتحصيل الأكاديمي بمعلومية الدافعية وحب الاستطلاع والقلق لدى طلاب جامعة الملك فيصل»

الفرض البديل : «يمكن التنبؤ بالتحصيل الأكاديمي بمعلومية الدافعية وحب الاستطلاع والقلق لدى طلاب جامعة الملك فيصل»

حجم التأثير :

مثال :

أثر طريقة التدريس على التحصيل الدراسي لدى طلاب المرحلة الابتدائية

المتغير المستقل : طريقة التدريس .

المتغير التابع : التحصيل الدراسي .

$$\boxed{\text{حجم التأثير} = \frac{\text{ت}^{٢}}{\text{ت}٢ + \text{درجات الحرية}}}$$

حجم التأثير الذي يفسر ١% (٠,١) .... حجم تأثير **ضعيف**

حجم التأثير الذي يفسر ٦% (٠,٠٦) .... حجم تأثير **متوسط**

حجم التأثير الذي يفسر ١٥% (٠,١٥) .... حجم تأثير **كبير**

مثال :

عند دراسة أثر برنامج لتنمية التفكير القائم على الحكمة على اتخاذ القرار لدى طلاب جامعة الملك فيصل، أشارت النتائج إلى أن قيمة "ت" تساوي (٢.٧)، ودرجات الحرية (٣٠). وفق هذه النتائج فإن قيمة حجم التأثير تساوي ....

$$\text{حجم التأثير} = \frac{\text{ت}^{٢}}{\text{ت}^{٢} + \text{درجات الحرية}}$$

$$\text{حجم التأثير} = \frac{(٢,٧)^{٢}}{(٢,٧)^{٢} + ٣٠}$$

= ٠,١٩ ⟵ حجم تأثير كبير

المحاضرة الثانية عشر : العينات

❖ **المجتمع:**

❖ يعرف المجتمع بأنه : « مجموعة من العناصر ، أو المفردات التي تخص ظاهرة معينة محل الدراسة. وهو مصطلح علمي يراد به كل من يمكن أن تعمم عليه نتائج البحث » .

❖ **العينة:**

- تعرف العينة بأنها : « جزء من مفردات المجتمع الإحصائي يتم اختياره بطريقة علمية، ثم دراسة خصائص هذا الجزء لغرض التعرف على خصائص المجتمع الذي اختير منه ذلك الجزء » .
- ولكي تكون العينة مقبولة من الناحية الإحصائية ينبغي أن تكون عينة ممثلة للمجتمع، أي أنها تحتوي على جميع الخصائص بنفس تواجدها في المجتمع الإحصائي الذي اختيرت منه.

أساليب جمع المعلومات :	
أسلوب الحصر الشامل :	• يتم فيه جمع البيانات عن الظاهرة موضوع الدراسة من جميع مفردات المجتمع الإحصائي المراد بحثه سواء أكان نطاقه أو مجاله واسعاً أو محدوداً.
أسـلـوب العيـنـات :	• يتم فيه جمع البيانات عن جزء فقط من مفردات المجتمع الإحصائي، ويتم سحب العينة بطريقة ما يساعد في تعميم نتائجها على مجتمع البحث.

مزايا « أسلوب الحصر الشامل » :
١) خال من أخطاء الصدفة (الأخطاء العشوائية أو أخطاء المعاينة)
٢) يعطي صورة مفصله عن مفردات الظاهرة موضوع الدراسة.

عيوب « أسلوب الحصر الشامل » :
١) الزيادة الكبيرة في التكاليف المادية والبشرية والزمنية.
٢) طول الوقت اللازم لجمع البيانات يفقد نتائج البحث حداثتها وبالتالي قيمتها.
٣) وجود مجتمعات بطبيعتها غير محدودة وبالتالي يتعذر تحديد إطار مفرداتها.

مزايا « أسلوب العينات » :
١) يوفر التكاليف المادية والبشرية والزمنية لإجراء الدراسة.
٢) زيادة الرقابة والضبط والتحكم في معظم الأسباب المؤدية إلى الأخطاء.
٣) يصلح للمجتمعات غير المحدودة.

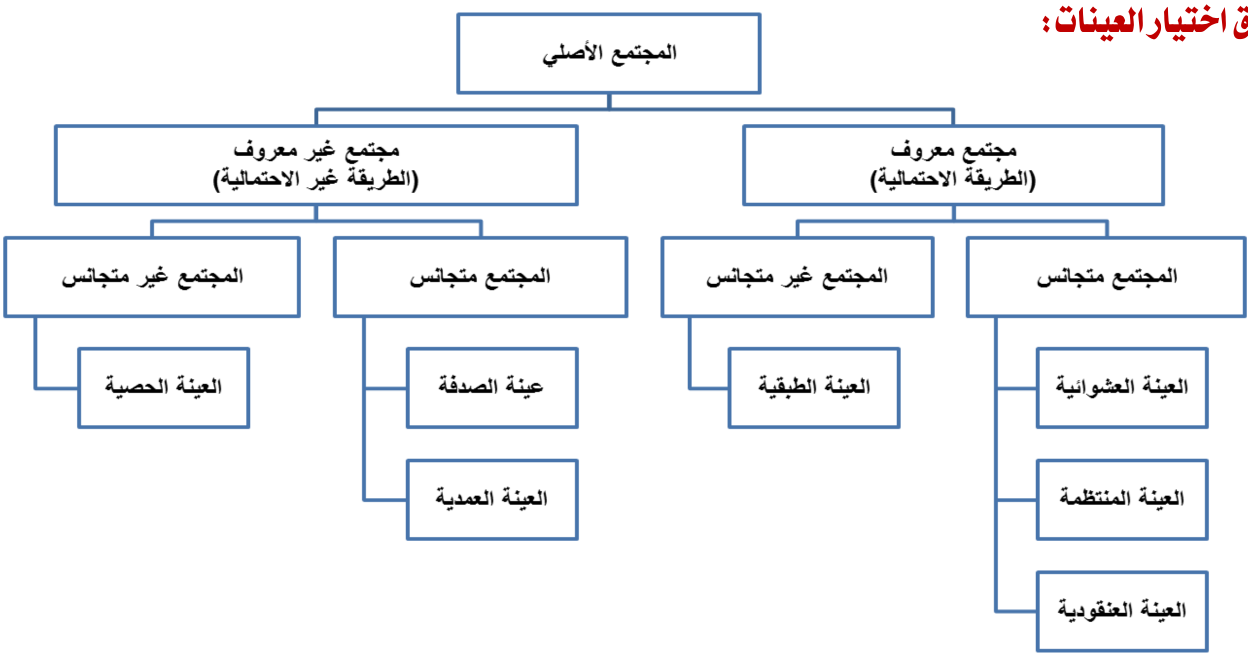
عيوب « أسلوب العينات » :
– يتعرض أسلوب المعاينة إلى نوع آخر من الأخطاء ينفرد به هذا الأسلوب ويطلق عليه خطأ المعاينة أو خطأ الصدفة وخطأ التحيز.

❖ اختيار العينة :

تمر عملية اختيار العينة بالخطوات التالية :	
١) تحديد المجتمع الأصلي للدراسة.	٢) تحديد افراد المجتمع الأصلي للدراسة.
٣) اختيار عينة ممثلة.	٤) اختيار عدد كاف من الأفراد في العينة.

ويتحدد الحجم المناسب للعينة من خلال العوامل التالية :		
١) تجانس أو تباين المجتمع الأصلي.	٢) اسلوب البحث المستخدم.	٣) درجة الدقة المطلوبة.

❖ طرق اختيار العينات :



أولاً : العينات الاحتمالية :

- يختار الباحث افراد المجتمع الأصلي للبحث معروفين ومحددين. فالتمثيل هنا يكون دقيقا ويتم الاختيار العشوائي وفق شرط محدد لا وفق الصدفة وهذا الشرط هو: ان يتوفر لدى كل فرد من افراد المجتمع الاصلي الفرصة المكافئة لكل فرد اخر في اختياره للعينة دون أي تحيز من قبل الباحث.

ثانياً : العينات اللااحتمالية :

➢ مثال: إذا كان المجتمع الأصلي طالبات كلية التربية – قسم اجتماع بجامعة الملك فيصل وعددهن (٥٠٠) طالبة، ونريد اختيار عينة من هذا المجتمع عددها (٥٠) طالبة ٠٠٠٠ ماذا نفعل وفقاً لهذه الطريقة؟

☒ العينة العشوائية البسيطة :

- تؤدى هذه الطريقة إلى احتمال اختيار أي فرد من أفراد المجتمع كعنصر من عناصر العينة .
- لكل فرد فرصة متساوية لاختياره ضمن العينة.
- اختيار فرد في العينة لا يؤثر على اختيار أي فرد آخر.

طريقة القرعة :

➢ **مثال:**

– إذا كان المجتمع الأصلي طالبات كلية التربية – قسم اجتماع بجامعة الملك فيصل وعددهن (١٠٠٠) طالبة ، ونريد اختيار عينة من هذا المجتمع عددها (١٠٠) طالبة ..... ماذا نفعل وفقاً لهذه الطريقة؟

☒ العينة العشوائية المنتظمة :

- يتم اختيار الحالة الأولى من العينة بطريقة عشوائية ثم يمضى الباحث في اختيار بقية الحالات على أبعاد رقمية منتظمة أو متساوية بين الحالات، بحيث تكون المسافة بين أي وحدتين متتاليتين ثابتة في جميع الحالات.
  - تحديد المجتمع الأصلي (n)
  - تحديد حجم العينة المرغوب فيه ((N)
  - تحديد المسافة بين أفراد العينة S=N/n
  - اختر عشوائياً عدداً ينحصر بين (١ وقيمة S).
  - أضف إلى العدد المختار قيمة S بشكل منتظم، لتحصل على العينة التي تريدها.

➢ **مثال:**

– إذا كان المجتمع الأصلي طالبات كلية التربية – قسم اجتماع بجامعة الملك فيصل وعددهن (٥٠٠) طالبة، ونريد اختيار عينة من هذا المجتمع عددها (٥٠) طالبة ..... ماذا نفعل وفقاً لهذه الطريقة؟

☒ العينة العنقودية :

- يختار الباحث النوع من العينات اذا كان مجتمع الدراسة على مستوى دولة كبيرة. حيث يصعب علية استخدام العينة البسيطة او العينة المنتظمة او العينة الطبقية. ويتبع الباحث في هذه الحالة تقسيم الدولة الي مناطق ثم الي محافظات ثم الي اجزاء صغيرة حتى يصل الي الافراد المطلوبين للعينة، والصالحين لتمثيل مجتمع الدراسة.

➢ **مثال:**

– اراد الباحث ان يتعرف على مدى استخدام اعضاء هيئة التدريس بكليات الآداب في المملكة للتقنيات الحديثة في التدريس.

✓ يكتفي بعدد ممثل من هذه الكليات.

☒ العينة الطبقية :

- نستخدم هذا النوع من العينات عندما يكون هناك تباين (عدم تجانس) واضح في مجتمع الدراسة، بحيث يمكن تقسيم مجتمع الدراسة إلى مجموعات أو طبقات بناءً على هذا التباين.

➢ **مثال:**

– أراد باحث إجراء دراسة على عينة عددها (٢٠٠) من طلاب كليات العلوم والتربية والآداب، إذا علمت أن عدد الطلاب (٢٥٠ العلوم ، و٣٥٠ التربية ، ٤٠٠ الآداب) . كيف يتم اختيار العينة ؟

$$\text{عينة طلاب كلية العلوم} = \frac{\text{عدد طلاب كلية العلوم}}{\text{العدد الكلي}} \times \text{عدد العينة}$$

$$\text{عينة طلاب كلية التربية} = \frac{\text{عدد طلاب كلية التربية}}{\text{العدد الكلي}} \times \text{عدد العينة}$$

$$\text{عينة طلاب كلية الآداب} = \frac{\text{عدد طلاب كلية الآداب}}{\text{العدد الكلي}} \times \text{عدد العينة}$$

☒ العينة الصدفة (العرضية) :

- هذا النوع من العينة يتم اختياره بالصدفة مثلما تستطلع صحيفة معينة الرأي العام حول قضية معينة أو مرشح ما، وغالبا ما يكون هذا النوع من العينات غير ممثلا لمجتمع الدراسة، وتستخدم هذه العينة في الدراسات الاستطلاعية المسحية المبدئية.

➢ **مثال:**

– اختيار الباحث لعدد من المصلين عند خروجهم من المساجد، أو الطلاب عند خروجهم من مدارسهم ويسالهم عن موقفهم حيال تأثير الفضائيات على التحصيل الدراسي للطلاب.

☒ العينة القصدية :

- ينتقي الباحث أفراد عينته بما يخدم أهداف دراسته وبناءً على معرفته دون أن يكون هناك قيود أو شروط غير التي يراها هو مناسبة من حيث الكفاءة أو المؤهل العلمي أو الاختصاص أو غيرها، وهذه عينة غير ممثلة لكافة وجهات النظر ولكنها تعتبر أساس متين للتحليل العلمي ومصدر ثري للمعلومات التي تشكل قاعدة مناسبة للباحث حول موضوع الدراسة.

➢ **مثال:**

– تحليل محتوى مجلة محددة، الخصائص النفسية لدى مدمني المخدرات، دراسة متعمقة لبعض حالات التخلف العقلي.

☒ العينة الحصصية :

يقوم الباحث اذا اراد الاخذ بالعينة الحصصية بتقسيم مجتمع الدراسة الي فئات، ثم يختار عددا من الافراد من كل فئة بما يتناسب وحجم الفئة في مجتمع الدراسة. وتشبه العينة الحصصية العينة الطبقية في هذا المعنى، لكن تختلف عنها في ان العينة الحصصية يتدخل الباحث في اختيار افراد العينة. ويعاب على هذا النوع من العينات، هو انه لا يمثل مجتمع الدراسة بصوره دقيقة.

المحاضرة الثالثة عشر : أدوات جمع البيانات

أدوات جمع البيانات Data Collection Instruments :

يقصد بأداة جمع البيانات الوسيلة التي تتم بواسطتها عملية جمع البيانات بهدف اختبار فرضيات البحث أو الإجابة عن تساؤلاته. ويتوقف اختيار الأداة المناسبة لجمع البيانات اللازمة والتي ستستخدم في إجراء بحث معين على نوعية البحث نفسه وطبيعته ، وعلى الهدف من تطبيقه ، وعلى نوعية المفحوصين وخصائصهم ...الخ ، وقد يستخدم الباحث أداة واحدة فقط لجمع البيانات التي يحتاج إليها في بحثه ، وقد يستخدم أكثر من أداة إذا وجد مبررا لذلك. وتجدر الإشارة إلى أن خطوة جمع البيانات في البحث تعتبر من الخطوات الأساسية التي يبدأ منها عمل الباحث ، لذا فالهدف النهائي من إعداد وسائل وأدوات جمع البيانات هو الحصول على تلك المعلومات التي تخدم في تحقيق أغراض البحث ودراسة مشكلته ، وإيجاد الحلول المناسبة له .

أولاً : الاختبارات والمقاييس Tests & Scales :

من المثيرات التي تقدم للفرد لاستثارة استجابات تكون أساساً لإعطاء الفرد درجة رقمية، وهذه الدرجة القائمة على عينة ممثلة لسلوك الفرد ، تعتبر مؤشراً للقدر الذى يمتلكه الفرد من الخاصية التي يقيسها الاختبار، ومنها :

١. **الاختبارات التحصيلية**	هي الاختبارات التي يراد بها قياس مستوى التحصيل الدراسي للطلاب، وهى واسعة الاستخدام في البحوث النفسية والتربوية، وقد تكون تحريرية، عملية أدائية، شفوية. ويجب أن تتمتع بخصائص سيكومترية جيدة في بنائها.
٢. **اختبارات الاستعدادات العقلية**	وتدخل جميع اختبارات الذكاء فى نطاق هذا النوع من الاختبارات ، وهى تختلف عن الاختبارات التحصيلية التي تقيس النواتج النهائية للتعلم المدرسي في الجانب المعرفي ، حيث تركز اختبارات الذكاء والقدرات العقلية عموماً على تحديد مدى استعداد الفرد للتعلم والدراسة من خلال نسبة ذكائه ومستوى قدراته العقلية ، ويمكن تصنيفها إلى اختبارات الاستعداد العقلي العام ، واختبارات الاستعداد العقلي الخاص " القدرات الخاصة"
٣. **اختبارات الشخصية**	١. **استبيانات الشخصية :** – مجموعة من العبارات تصف السلوك موجهة للمفحوص ، وعليه أن يجيب على كل عبارة أو سؤال بالاختيار الذى يناسبه. – لا يوجد عبارات صحيحة وأخرى خاطئة. – الاستجابات قد تكون ثنائية (نعم - لا) أو ثلاثية (موافق - غير متأكد - غير موافق). – بعض الاختبارات تقيس بعد واحد وبعضها الآخر متعدد الأبعاد. – تتميز بأنها اقتصادية ، بسيطة ، موضوعية. – **من عيوبها :** المرغوبية الاجتماعية أو التزييف، الاستجابات تعتمد على معرفة الفرد لنفسه وتصرفاته في المواقف المختلفة ، الاختيار من بين الاستجابات الموجودة وعدم إضافة شيء، ضرورة معرفة القراءة، لا توضح الأسس والدوافع التي تجعل المستجيب يختار إجابة دون غيرها. ٢. **الأساليب الاسقاطية :** – مثير غامض يستجيب له الفرد استجابة حرة بالطريقة الحرة التي يريدها. – يستخدمها الأخصائيون النفسيون الإكلينيكيون لدراسة وتشخيص المشكلات الانفعالية للفرد. – من أشهرها: اختبار روشاخ – اختبار تفهم الموضوع .

٤. **مقاييس الاتجاهات**	– الاتجاه هو استجابة موجبة أو سالبة للفرد نحو موضوع، أو مؤسسة، أو مفهوم أو قضية ذات صبغة اجتماعية غالباً. – **يتضمن الاتجاه ثلاثة جوانب :** هدف، حالة انفعالية، توجيه السلوك.
٥. **مقاييس التقدير**	– تستخدم عندما نريد تحديد درجة حدوث السلوك. – تتكون من مجموعة الخصائص أو الصفات للحكم عليها، ومقياس مدرج لتحديد درجة تواجد الخاصية أو الصفة. – الاستمارة المستخدمة هي مجرد أداة لتسجيل الملاحظات، وتتوقف قيمتها في جمع البيانات على الدقة في البناء والتنفيذ.

ثانياً : الاستبيانات *Questionnaire* :

– عبارة عن وثائق توجه نفس الأسئلة إلى جميع الأفراد في العينة.

– يسجل المستجيبون إجابات مكتوبة لكل مفردة من المفردات ، فهم يتحكمون في جمع البيانات حيث يملؤون الاستبيان بالطريقة التي تناسبهم وبالترتيب الذى يرونه.

– **يمكن تصنيف أسئلة الاستبيان إلى : الأسئلة المفتوحة ، والأسئلة المقيدة. ويمكن إجراء مقارنة بين مزايا وعيوب النوعين :**

	الاستبيانات المقيدة	**الاستبيانات المفتوحة**
مميزاتها	١. أسهل للمستجيبين وأسرع في الإجابة ٢. يسهل مقارنة إجابات المستجيبين ٣. يسهل ترميز الإجابات وتحليلها إحصائياً ٤. يزيد احتمال استجابة أفراد العينة للأسئلة ٥. يقل عدد الأسئلة الغامضة والمحيرة	١. للمستجيب حرية إعطاء أى عدد من الإجابات ٢. يمكن الحصول على نتائج غير متوقعة واستجابات كافية لقضايا معقدة ٣. تسمح بحرية الابتكار والتعبير عن الذات وتكشف عن طريقة التفكير ٤. يستطيع المستجيب إعطاء مبررات لإجاباته
عيوبها	١. تعطى الفرد فرصة إعطاء إجابات لم يفكر فيها ٢. يصعب التمييز بين الإجابات المختلفة ٣. يصاب الفرد بالإحباط لعدم توفر إجابة تناسبه ٤. من ليس لديه فكرة عن الموضوع يستطيع الإجابة ٥. عند زيادة عدد الإجابات عن عشرة يقع المفحوص في حيرة وقلق	١. يختلف المستجيبون فيما بينهم في درجة التفصيلات التى يعطونها ٢. يصعب مقارنة الإجابات وترميزها وتحليلها إحصائياً ٣. تتسم الأسئلة بالعمومية ، وتحتاج إلى وقت كبير ، ومساحة للكتابة ٤. المستوى التعليمى يؤثر على الإجابة

ثالثاً : المقابلة *Interview* :

– مجموعة أسئلة شفوية يسألها المقابل ويحصل على استجابات شفوية من المشاركين.

– أكثر استخداماً في البحوث الكيفية، لأنها تسمح بالاستكشافات ذات الطبيعة المفتوحة ، كما أنها تسمح للمستجيبين بحرية غير محدودة في الإدلاء بما يريدون من استجابات.

– استبيان منطوق، والفرق الأساسي بينهما أن المقابلة تتضمن التفاعل المباشر بين الباحث والمستجيب.

– تفضل المقابلة في الموضوعات الشخصية بينما يفضل الاستبيان في الموضوعات العامة.

– مرنة ويمكن تعديلها حسب الموقف، ويمكن استخدامها مع انواع مختلفة من المشكلات والأشخاص.

أنواع المقابلة :

١. **مقننة** ، وفيها تكون الأسئلة محددة ، ويتبع كل سؤال مجموعة من الاختيارات أو الإجابات يختار من بينها المستجيب الإجابة التى تتفق مع رأيه .وتتميز بالثبات والصدق والموضوعية المرتفعة.

٢. **شبه المقننة** ، وفيها لا يتبع الأسئلة اختيارات محددة ولكن تصاغ بحيث تسمح بالإجابات الفردية ، فالسؤال مفتوح ولكنه محدد للغاية في محتواه.

٣. **غير المقننة** ، وفيها يقوم الباحث بتوجيه أسئلة واسعة في أي ترتيب يراه مناسباً ، والتركيز هنا على المستجيب ، ودرجة ثباتها وصدقها محدودة.

- يفضل استخدام مزيج من المقابلة المقننة وغير المقننة.
- يفضل تسجيل الإجابات حرفياً كما أعطاها المستجيب.
- وجود متغيرات شخصية تتعلق بالباحث تؤثر في المقابلة منها: عمر الباحث ، التخصص ، المستوى التعليمي ، الخبرة ، الجنس.

رابعاً : الملاحظة *Observation* :

- طريقة لجمع المعلومات عن سلوك فى سياقه الطبيعى ، وتوصف الملاحظة بأنها أفضل طرق جمع المعلومات عن السلوك ، لأنها لا تتطلب وسيطاً كالاختبارات أو الاستبيانات ، ومع أنها تمدنا بمعلومات ثرية إلا أنها معقدة وتحتاج لجهد وترتيب مكثفين.
- أدوات الملاحظة هى الأدوات التى نستخدمها أثناء الملاحظة لتسجيل الملاحظات مثل قوائم المراجعة ، مقاييس التقدير ، السجلات القصصية.
- أسلوب الملاحظة هو عملية ملاحظة السلوك ذاتها تمهيداً لتسجيلها.

لكي تكون الملاحظة دقيقة وصادقة يجب :

١. التخطيط مسبقاً لما نلاحظه، وذلك بناء على أهداف المشكلة التي ندرسها.
٢. التركيز على نوع أو نوعين من السلوك فقط .
٣. استخدام صفات واضحة غير غامضة حتى تكون الملاحظة محددة تصف السلوك وصفاً سليماً.
٤. أن يكون كل سلوك ملاحظ مختلفاً عما عداه من أنواع السلوك الأخرى.
٥. أن يكون الباحث واعياً بما يحدث من أخطاء الملاحظة التي تحدث نتيجة لاختيار أوقات معينة نلاحظ فيها السلوك.
٦. تسجيل وتلخيص الملاحظات عقب حدوثها مباشرة.
٧. أن يختار الباحث من يلاحظه في كل مرة .
٨. تأجيل تفسير السلوك إلى ما بعد جمع البيانات.
٩. ألا يظهر الباحث أنه يلاحظ سلوكاً ما أو فرداً ما.

خامساً : استطلاعات الرأي :

تشكل استطلاعات الرأي مصدرا مهما للمعلومات حول الرأي العام ، وهي من أهم الأدواتَ التي تساعد على كتابة تقارير معلوماتية دقيقة وموضوعية.

هدف استطلاع الرأي :

توقع النتائج جملة شديدة البساطة. وبالتالي، فالاستطلاع الجيد هو من يقدم نتائج هي الأقرب للنتيجة النهائية.

ما يجب مراعاته عند قراءة نتائج استطلاع للرأي :

- من الذي أجرى الاستطلاع؟
- ما الجهة الممولة للاستطلاع ، ولماذا ؟
- كم عدد عينة الاستطلاع ؟

- كيف تم اختيار العينة؟ أو هل هي عينة ممثلة؟
- هل يوجد تطابق بين نتائج الاستطلاع وأجوبة العينة؟
- من كان مفترضا ان يتم استطلاع رأيه ولم يحدث؟
- متى اجري الاستطلاع؟
- كيف أجريت المقابلات ، ومن أجراها ؟
- كيف تم ترتيب الاسئلة؟ أو ما طريقة تقديم الأسئلة ؟
- ماذا عن الاستطلاعات الاخرى التي اجريت على نفس الموضوع؟ هل وصلت الى نفس النتائج؟ ولماذا يوجد اختلاف أن كان موجود ؟
- ما الاشياء الاخرى التي كان من الضروري ان يوردها تقرير الاستطلاع وغير موجودة ؟

وسنركز في هذه المحاضرة على الاستبانة كأحد وسائل جمع البيانات المهمة في البحوث والدراسات الاجتماعية

تعتبر الاستبانة من الأدوات البحثية شائعة الاستخدام في أغلب البحوث والدراسات النفسية والاجتماعية ، وخاصة تلك التي تركز على جمع معلومات وبيانات متعلقة بمعتقدات ورغبات المستجيبين ، وكذلك الحقائق التي هم على علم بها.

خطوات بناء الاستبانة :

أولاً : الاطلاع على الدراسات والبحوث السابقة :

إن الهدف الأساسي من الرجوع إلى الدراسات السابقة هو تكوين فكرة عامة عن الظاهرة موضع الدراسة ، ومحاولة تحديد مشكلة البحث ، والتعرف على ما تم التوصل إليه في هذا الموضوع ، والعمل على حصر الموضوعات التي ستتضمنها الاستبانة ، والمساعدة على تحديد الكثير من فقرات الاستبانة بشكلها النهائي (توماس Thomas 1999) .

ثانياً : تحديد الأسئلة الرئيسة للبحث موضع الدراسة :

بعد الاطلاع على الدراسات السابقة وقبل الشروع في بناء الاستبانة لابد من تحديد الأسئلة الرئيسة التي يرغب الباحث الإجابة عليها (ديرشوسكي Dereshiwsky 1993) . ومن الخصائص الأساسية لهذه الأسئلة أن تكون محددة ، واضحة ، دقيقة ..الخ ، كذلك لابد أن تكون هذه الأسئلة محددة لنوع المعلومات التي من أجلها تبنى الاستبانة ، وقد تصاغ هذه الأسئلة بصورة عامة ، وقد تصاغ بصورة محددة . وإن ما يجب التنويه إليه هنا أن هذه الخطوة لا تتعلق ببناء فقرات الاستبانة ، بل هي تعتبر بمثابة الإطار العام للاستبانة ، والتركيز على هذه الأسئلة الأساسية يساعد الباحث على عدم إضافة فقرات ليس لها علاقة بالبحث .

وهناك عدد من الخطوات الأساسية التي تساعد الباحث على كتابة الأسئلة الرئيسة للبحث وتحديدها وهي :

- ✓ الرجوع إلى الدراسات السابقة من كتب وبحوث ورسائل علمية .
- ✓ مناقشة الموضوع مع المتخصصين .
- ✓ مناقشة الموضوع مع صناع القرار .
- ✓ النزول إلى الميدان للاطلاع على الواقع الفعلي للظاهرة موضع الدراسة .

ثالثاً : تحديد الأسئلة الفرعية المبنية على الأسئلة الرئيسية :

عادة الأسئلة الرئيسة للدراسة تحوي بعض الكلمات العامة والتي يمكن أن تحمل أكثر من معنى ، مثلا " التدريس ، التعليم الجماعي ، المهارات الإدارية، تحقيق الذات ، إدارة الصف ، الاتجاهات الخ ، هذه الكلمات تحتاج إلى نوع من التحديد والتعريف ، وعادة يتم هذا من خلال التعريف الإجرائي لهذه الكلمات إضافة إلى وضع الأسئلة الفرعية التي تتناول بالتفصيل الموضوعات المندرجة تحت هذه الكلمات .

وقد أشار (كوكس Cox، 1997) أن الأسئلة الفرعية لابد أن تتصف بالآتي :

- ✓ أن تكون قابلة للقياس
- ✓ أن تكون دقيقة و تعالج موضوعا محددا .
- ✓ أن تكون على مستوى واحد من الصياغة .

و يمكن استنباط هذه الأسئلة الفرعية من خلال :

- ✓ الرجوع إلى الكتب ، البحوث ، الدراسات العلمية ذات العلاقة بالموضوع .
- ✓ الحوار و المناقشة مع المتخصصين .

رابعاً : الدراسة الاستطلاعية :

بعد الاطلاع على الدراسات السابقة وتحديد الاسئلة الرئيسة والفرعية ذات العلاقة بموضوع البحث يأتي دور الدراسة الاستطلاعية وذلك لوضع المشكلة المدروسة ضمن الاطار الحضاري لها ، لأن الدراسات السابقة كما هو معروف قد اجريت في مجتمع غير مجتمع البحث . فالدراسة الاستطلاعية تساعد على التأكد من أن التصميم الذي انتهى إليه الباحث واقعي ويمكن تنفيذه

زد على ذلك أن الدراسة الاستطلاعية تمكن الباحث من التالي :

- ✓ التعرف على أفضل الأساليب لمخاطبة أفراد العينة .
- ✓ تحديد الموضوعات التي سوف تدور حولها اسئلة الاستبانة .
- ✓ تحديد شكل الأسئلة التي تدور حولها تلك الموضوعات .
- ✓ تحديد الصورة الاجمالية للاستبانة .
- ✓ تحديد أفضل ترتيب للأسئلة التي تدور حول موضوعات الاستبانة .
- ✓ تحديد الصياغة اللفظية للغة الاستبانة .

خامساً : كتابة فقرات الاستبانة :

من أجل ضمان دقة المعلومات التي تحصل عليها من خلال وسيلة جمع البيانات ، فإنه لابد لفقرات الإستبانة أن تكون دقيقة ومحدده وغير قابلة لأكثر من تفسير (Fink, 1995). إن الأسئلة الدقيقة تضمن لنا إجابة المستجيب على فقرات الإستبانة بطريقة تتوافق مع الهدف الذي وضعت من أجله. إن وضوح فقرات الإستبانة لا يتيح للمستجيب قراءة الفقرة مرة واحدة لفهمها فقط بل يساعد ايضا على تقليص الوقت المتطلب لإكمال هذه الإستبانة وبالتالي ضمان إعادتها . يجب على الباحث اثناء كتابة فقرات الإستبانة أن يضع نفسه موضع المستجيب ، فالبعد عن استخدام الكلمات غير المفهومة مطلب اساسي لضمان دقة الإجابة.

ولقد اشار (Cox, 1996 ،Fink, 1995) إلى بعض الإرشادات التي تساعد الباحث على كتابة فقرات الإستبانة بطريقة جيدة :

- ✓ أن تكون الفقرة واضحة وبسيطة .
- ✓ تجنب استخدام المصطلحات العامة ، والكلمات الغامضة (اكتب باللغة التي يفهمها المستجيب) .
- ✓ استخدام الاسئلة القصيرة المحددة المعنى .
- ✓ صياغة العبارات بصورة لا توحي بالتحيز إلى أحد الاتجاهات .
- ✓ مراعاة عدم وضع اسئلة تمس شعور المفحوص أو عقائده .
- ✓ صياغة الاسئلة والعبارات بصورة تسمح بمعرفة شدة الاستجابة .
- ✓ تجنب صياغة الاسئلة بالنفي .
- ✓ تجنب الاسئلة التي تحوي على فكرتين .

سادساً : الشكل العام للاستبانة :

تعرف الاستبانة بأنها عبارة عن استمارة تضم مجموعة من الأسئلة توجه للأفراد بغية الحصول على بيانات معينة ، وهذه الأسئلة التي تتضمنها الاستمارة **يمكن تقسيمها إلى نوعين تبعاً لأسلوب الحصول على البيانات :**

- **الأسئلة المباشرة :** وهي التي تهدف إلى الحصول على المعلومات بطريقة واضحة وصريحة .
- **الأسئلة غير المباشرة :** وهي التي يمكن من خلال الإجابة عنها استنتاج البيانات المطلوبة .

وكذلك يمكن تقسيمها إلى نوعين وفقاً لأسلوب تقنينها :

- **أسئلة مغلقة :** وهي التي تحدد إجابة الفرد في إطار المتغيرات المحددة كأن تكون نعم ولا ، أو موافق ، وغير موافق .. الخ .
- **أسئلة مفتوحة :** وهي التي تسمح للمستجيب بالإجابة الحرة دون التقيد بإجابات معينة .

ولا بد أن يراعي الباحث عند اعداده الاستبانة جملة من الأمور تتعلق بالشكل العام للاستبانة منها :

١. **طول الاستبانة :** يجب أن يكون طول الاستبانة معقول ، فعندما تكون الاستبانة طويلة فهذا يؤدي إلى عدم الإجابة الكاملة على بنودها . لذا ينصح بأن تكون المدة المحددة للإجابة على الاستبانة من ١٠ إلى ١٢ (cox, 1996) .
٢. **تصنيف الفقرات :** العبارات ذات الاستجابة الموحدة من الأولى أن تكون مع بعض ولذا يجب مراعاة عدم تشتيت المستجيب في الانتقال من شكل استجابة إلى شكل آخر .
٣. **استغلالية الصفحات :** يجب أن لا توزع المعلومة المراد الإجابة عليها على أكثر من صفحة حتى لا يؤدي ذلك إلى ازعاج المستجيب في الرجوع إلى معلومات في صفحات سابقة .
٤. **المسافات :** يجب عدم ضغط المعلومات والفقرات في صفحات محددة مما يجعلها مزدحمة وغير واضحة للمستجيب .
٥. **وضوح الخط المستخدم :** ينبغي أن يكون الخط المستخدم في كتابة فقرات الاستبانة واضح ومقروء للجميع من حيث الخط ومقاسه .
٦. **المراجعة اللغوية لمحتويات الاستبانة :** ينبغي على الباحث المراجعة اللغوية لجميع محتويات الاستبانة لأن الخطأ الإملائي قد يؤدي إلى خطأ في الاستجابة مما قد يؤثر على النتائج المتحصلة .

سابعاً : اختبار الاستبانة :

اختبار الاستبانة يعني التأكد من أنها أصبحت صالحة للاستخدام من حيث المدلول والمحتوى لجمع المعلومات حول المشكلة قيد البحث . **وبهذا المفهوم لاختبار الاستبانة يمكن التفريق بين :**

- ✓ الاختبار الذي يهدف إلى تصحيح المدلول اللفظي لكل بند من بنود الاستبانة وإزالة ما يمكن أن يؤدي إلى غموض أو عدم معرفة المراد منه . ويتم ذلك من خلال عرض الاستبانة على من لهم خبرة علمية في مجال البحث .
- ✓ الاختبار الذي يهدف إلى التأكد من مدى صدق الاستبانة وثباتها ويتم ذلك باختيار عدة أشخاص من مجتمع البحث ثم يطلب منهم إجابة الاستبانة ويقاس في ضوء استجابتهم مدى صدق الاستبانة وثباتها .
- ✓ الاختبار الذي يهدف إلى التأكد من مدى جِّدية المجيب في إجابته للاستبانة ، وذلك من خلال تنويع صياغة سؤال أو أكثر ذي مدلول واحد ليتبين له من خلال مقارنة الاجابة مدى جديتها .

ثامناً : كتابة تعليمات الإجابة :

بالإضافة إلى الاعتناء بمحتوى وشكل الاستبانة لا بد من تزويد المجيب بتعليمات واضحة للإجابة على بنود هذه الاستبانة، تكون على شكل رسالة مصاحبة يوضح فيها المشكلة قيد الدراسة باختصار، والهدف من بحثها . ومدى أهمية مشاركة المجيب في تحقيق ذلك الهدف .

تاسعاً : توزيع الاستبانة ومتابعتها :

بعد أن يقوم الباحث ببناء الاستبانة واختبارها من حيث سلامة المدلول اللفظي لبنودها وصحة معناها ، وحساب معامل صدقها وثباتها ، **يتعين عليه اختيار الطريقة المناسبة التي سيستخدمها للتوزيع ومنها :**

- ✓ **التوزيع المباشر:** وهو أن يقوم الباحث بنفسه أو من يمثله بتسليم الاستبانة لأفراد العينة .
- ✓ **التوزيع غير المباشر:** وفيها يقوم الباحث بإرسال الاستبانة عبر البريد .

إن التوزيع بكلا الطريقتين يحتاج إلى متابعة حثيثة من قبل الباحث، وذلك لتدني نسبة المستجيبين والتي تعتبر من أهم العقبات التي تقف في طريق الباحثين عند استخدامهم للاستبانة كوسيلة لجمع البيانات

عاشراً : تبويب وترميز بيانات الاستبانة بالطريقة المناسبة :

بعد أن يتم جمع المعلومات من خلال الاستبانة يقوم بمراجعتها وذلك بهدف استبعاد الاستمارات التي لم يجب عليها أولاً، ثم التأكد من مدى جدية المجيب في إجابته من خلال مراجعة البنود التي وضعت لقياس هذا الأمر، ومن ثم **يبدأ عملية التبويب من خلال الآتي:**

- ✓ وضع رقم لكل استبانة .
- ✓ وضع رقم لكل عبارة أو سؤال .
- ✓ وضع رقم لكل إجابة من إجابات العبارة أو السؤال .

حادي عشر : تفريغ معلومات الاستبانة وادخالها بالطريقة المناسبة في الحاسب الآلي :

بعد أن يتم ترميز بيانات الاستبانة وإعطاء رقم لكل استمارة وبند، واجابة ، يتم بعد ذلك تفريغ هذه المعلومات وإدخالها بالطريقة المناسبة في الحاسب الآلي (باستخدام احد البرامج الإحصائية المناسبة) ويجب هنا مراعاة المتغيرات موضع الدراسة وطريقة تحليلها . لأن طريقة إدخال هذه البيانات في الحاسب تؤثر بطريقة مباشرة على النتائج المتحصلة .

ثاني عشر : تحليل بيانات الاستبانة :

بعد أن يتم ادخال بيانات الاستبانة في الحاسب الآلي يأتي دور معالجة هذه البيانات معالجة رقمية وذلك من خلال تطبيق أساليب الإحصاء بنوعيه الوصفي والاستنتاجي ، وهنا يحتاج الباحث إلى التأني في الاختيار المناسب للأسلوب الإحصائي لأن ذلك قد يؤثر بطريقة مباشرة على النتائج المتحصلة .

المحاضرة الرابعة عشر : الثبات والصدق للاختبار والمقاييس

الشروط العلمية للاختبار :

١. **موضوعية الاختبار :** ويقصد بموضوعية الاختبار عدم تأثر المصحح بالعوامل الذاتية عند تصميمه لأوراق الإجابة .
٢. **صدق الاختبار :** يقصد بصدق الاختبار مدى قدرته على قياس المجال الذي وضع من أجله أو بمعنى أكثر تحديدا مدى صلاحية درجاته للقيام بتفسيرات مرتبطة بالمجال المقاس .
٣. **ثبات الاختبار :** يقصد بصدق الاختبار دقته واتساقه وبمعنى أدق أن يعطي الاختبار نفس النتائج إذا ما تم استخدامه أكثر من مره تحت ظروف مماثلة .

معنى الثبات :

إذا أجري اختبار ما على مجموعة من الأفراد ورصدت درجات كل فرد في هذا الاختبار ثم أعيد إجراء نفس هذا الاختبار على نفس هذه المجموعة ورصدت أيضا درجات كل فرد ودلت النتائج على أن الدرجات التي حصل عليها الطلاب في المرة الأولى لتطبيق الاختبار هي نفس الدرجات التي حصل عليها هؤلاء الطلاب في المرة الثانية ، نستنتج من ذلك أن النتائج الاختبار ثابتة تماما لأن نتائج القياس لم تتغير في المرة الثانية بل ظلت كما كانت قائمة في المرة الأولى .

- ✓ درجة الاتساق في قياس السمة موضوع القياس من مرة لأخرى فيما لو أعدنا تطبيق الأداة عددا من المرات (**يسمى دقة القياس**) .
- ✓ يعبر عن الثبات بصورة كمية يطلق عليها **معامل الثبات تتراوح بين صفر والواحد الصحيح (٠_١)** .
- ✓ كلما **زادت** قيمة المعامل دلت على (**أن الأداة تتمتع بثبات مرتفع والعكس صحيح**) .

أخطاء تؤثر على الثبات بشكل أساسي :

- أخطاء القياس المنتظمة والتي تعود الى أداة القياس كأن تكون صعبة جدا أو سهلة جدا .
- أخطاء القياس العشوائية والتي تعود للمفحوص نفسه كأن يكون مريض أو غير مهتم .
- الاختبار الصادق هو اختبار ثابت وليس كل اختبار ثابت هو اختبار صادق .

أنواع الثبات :

١. ثبات الإعادة .
٢. ثبات الصورة المتكافئة .
٣. الثبات بالطريقة النصفية .
٤. ثبات المصححين .

١. **ثبات التطبيق وإعادة التطبيق**	– يطبق الاختبار على عينة ما . – يعطي الباحث مهلة . – يعيد الباحث تطبيق نفس الاختبار على نفس العينة . – يقارن الباحث نتائج التطبيق الأول مع نتائج إعادة التطبيق – إذا كانت متطابقة أو متقاربة فإن الأداة تتمتع بمعامل ثابت مرتفع .
٢. **ثبات الصورة المتكافئة**	– إعداد صورتين متكافئتين لأداه ما – يتم تطبيق الصورتين على عينة ما . – يتم حساب معامل الارتباط بين نتائج صورتي الأداة . – إذا كانت معامل الارتباط عالي فإن الأداة تتمتع بمعامل ثابت مرتفع .

٣. **ثبات الطريقة النصفية (التجزئة النصفية)**	– يطبق الاختبار أو الأداة مره واحدة فقط . – تقسم فقرات الاختبار أو أسئلته إلى نصفين (الفقرات الفردية معا والزوجية معا) – مثال :الفقرات ١١,٩,٧,٥,٣,١ معا ١٠,٨,٦,٤,٢ معا – يقوم الباحث بحساب معامل الثبات باستخدام طريقة سيبرمان - براون Spear man-Brown . – إذا كانت معامل الثبات عالي فإن الأداة تتمتع بمعامل ثابت مرتفع .
٤. **ثبات المصححين**	– حساب ثابت الأداة إذا كانت هناك أكثر من مصحح أو ملاحظ اشتركوا في التصحيح أو جمع البيانات . – تحسب من خلال إعداد قائمة بدرجات كل مصحح على حده . – ثم يحسب معامل الارتباط بين قوائم المصححين هذه . – إذا كانت معامل الارتباط عالي فإن الأداة تتمتع بمعامل ثبات مرتفع .

العوامل المؤثرة في الثبات :

١. **طول الاختبار أو كثرة عدد فقراته :** كلما زادت الفقرات زاد معامل الثبات (أن لا يزيد طول الأداة عن ٣٥ إلى ٤٥ فقرة)
٢. **زمن الاختبار :** كلما زاد زمن الاختبار زاد معامل الثبات (مع ملاحظة أن هذا الأمر قد يكون مناسبا للاختبارات التحصيلية لكن أدوات القياس فالأمر يختلف).
٣. **تباين مجموعة الثبات (العينة) :** كلما كان أفراد العينة متباينين كلما زاد معامل الثبات .
٤. **صعوبة الاختبار :** يرتفع معامل الثبات إذا كانت متوسط الصعوبة (الاختبار الصعب أو السهل يؤدي إلى معاملات ثبات منخفضة).

حساب معامل الثبات :

يحسب الثبات من خلال حساب معامل الارتباط وهو خير طريقة لمقارنة هذه الدرجات التي حصل عليها الطلاب في الاختبارين ويحسب معامل التباين من العلاقة التالية : $Re\ liability = \frac{2(r)}{1+(r)}$

وقيمة r لبيرسون يتم حسابها من العلاقه التالية :

$$r = \frac{\sum XY - \frac{(\sum X)(\sum Y)}{n}}{\sqrt{\left(\sum X^2 - \frac{(\sum X)^2}{n}\right)\left(\sum Y^2 - \frac{(\sum Y)^2}{n}\right)}}$$

الصدق :

معنى الصدق :

الاختبار الصادق يقيس ما وضع لقياسه . فاختبار الذكاء الذي يقيس الذكاء فعلا اختبار صادق مثله في ذلك كمثل المتر في قياسه للأطوال والكيلو في قياسه للأوزان والساعة في قياسها للزمن . وتختلف الاختبارات في مستويات صدقها تبعا لاقترابها أو ابتعادها من تقدير تلك الصفة التي تهدف إلى قياسها . فاختبار الذكاء الذي يصل في قياسه لتلك القدرة إلى مستوى ٠.٨ أصدق في هذا القياس من أي اختبار آخر للذكاء لا يصل إلى هذا المستوى أي إنه أصدق مثلا من الاختبار الذي يصل في قياسه للذكاء إلى مستوى ٠.٥ .

أنواع الصدق :

١- صدق المحتوى .
٢- صدق المفهوم أو صدق البناء .
٣- الصدق التلازمي .
٤- الصدق التنبؤي .

١. صدق المحتوى	– إعداد وتحليل محتوى الظاهرة محور الدراسة . – صياغة الفقرات . – عرض الفقرات ونتائج تحليلها على مجموعة من الخبراء في ميدان البحث لمعرفة مدى مناسبة الفقرات وسلامتها وانتمائها للظاهرة المقاسة – أحيانا يقوم الباحث بإعداد كشف يتكون من درجات للخبراء لوضع تقيمهم عليه . **مثال :** الفقرة مناسبة (١,٢,٣,٤,٥,٦,٧,٨,٩,١٠) . **اللغة سليمة :** (١,٢,٣,٤,٥,٦,٧,٨,٩,١٠) .
٢. صدق المفهوم أو صدق البناء	– قياس مفهوم افتراضي غير قابل للملاحظة مثل الذكاء أو الدافعية . – يبين هذا النوع من الصدق مدى العلاقة بين الأساس النظري للاختبار وبين فقرات الاختبار ، وبمعنى آخر إلى أي مدى يقيس الاختبار الفرضيات النظرية التي يبنى عليها الاختبار .
٣. الصدق التلازمي : مهم	مدى ارتباط الدرجات المحققة على الأداة بالدرجات المحققة على أداة آخري تقيس نفس السمة . **مثال :** قام باحث بإعداد اختبار ذكاء ويريد حساب دلالات صدق هذا الاختبار . – يقوم بتطبيق اختباره . – يقوم بتطبيق اختبار آخر من اختبارات الذكاء المعروفة . – يقوم بحساب معامل الارتباط بيرسون بين الإختباريين – إذا كان معامل الارتباط قوي بين الإختباريين وذو دلالة عندها نقول أنه يوجد صدق تلازمي للاختبار .
٤. الصدق التنبؤي	هو الدرجة التي يمكن من خلالها للمقياس أن يكون قادرا على التنبؤ بأداء معين (محك) في المستقبل . **مثال :** قدرة اختبارات الذكاء على التنبؤ بالتحصيل الأكاديمي المستقبلي للطلاب .

تم بحمد الله