مراجعة عامة

المحاضرة الاولى: المتغيرات

- **المتغير variable:** هو الخاصية أو الصفة المقاسة بمقياس معين والقابلة للملاحظة من طرف الباحث، ويرمز لها بالرمز **xi** وكمثال على ذلك متغير الطول، الوزن، الجنس، لون الشعر.
- **اقسام المتغيرات:** تقسم المتغيرات إلى قسمين رئيسيين:

1- المتغيرات الكيفية (النوعية)variables qualitatives: هذه المتغيرات وصفية و لا تأخذ فيها الأعداد معنى كمي مثل: اللغة، الجنس، الوظيفة. نلاحظ في هذا النوع من المتغيرات أن النصنيف يكون على أساس امتلاك الفرد للخاصية أو السمة أو عدم امتلاكه.

هذا النوع من المتغيرات يمكن أن يكون في مستوى قياس اسمي/ تصنيفي مثل: الديانة، اللون، الجنس، الحالة الاجتماعية (أعزب، متزوج، مطلق، أرمل). كما يمكن أن يكون له ترتيب أو تصنيف معين فيكون في مستوى ترتيبي مثل دخل الفرد (مرتفع، متوسط، منخفض) أودرجة مشاركة الطلبة في مقياس الاحصاء (كبيرة، متوسطة، ضعيفة)

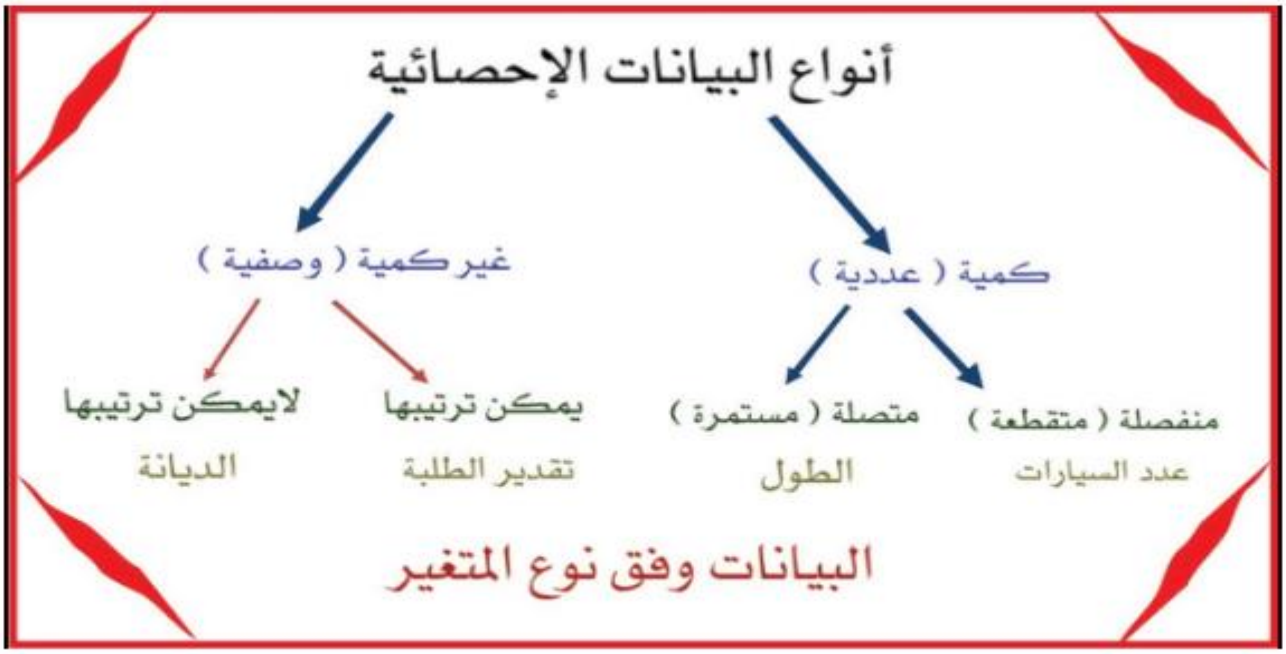
2- **المتغيرات الكمية variables qualitatives:** هي متغيرات نقاس بمقدار مثل: الوزن، الطول، السعة. ويقسم هذا النوع من المتغيرات إلى قسمين:

3- متغيرات كمية متصلة(مستمرة) variables continues: وهي تلك المتغيرات التي يمكن أن تأخذ قيمها ألرقاما صحيحة أوكسرية مثل: درجة الحرارة، الوزن، الطول، العمر، الأجر. ونلاحظ هنا في هذا النوع من المتغيرات أنه يمكن تقسيم وحدات قياسه إلى وحدات جزئية بحيث تكون هناك استمرارية في القياس.

2-2- متغيرات كمية منفصله(متقطعة) variables discrètes: وهي متغيرات نعبر عنها بأرقام عددية صحيحة مثل: عدد العمال، عدد المؤسسات الاقتصادية الخاصة، عدد الوفيات، عدد المساكن في حي من الأحياء...

إن هذه المتغيرات يمكن أن تندرج تحت مقياس المسافات المتساوية مثل: درجة الحرارة، درجة غليان الماء، عدد الوفيات، عدد العمال. كما يمكن أن تندرج تحت مقاييس النسبة مثل: الطول، الوزن...

والشكل التالي يوضح أنواع المتغيرات والمقاييس المناسبة لهذه المتغيرات



مستويات القياس

مدخل: إن النتائج التي يتوصل إليها أي باحث غالبا ما يترتب عليها اتخاذ قرارات، لذا يجب أن تكون الأساليب المتبعة مناسبة لطبيعة ونوعية البيانات التي يعطيها الاختبار المستخدم في البحث. ونرى أن فقدان قيمة أي بحث يعود إلى الأساليب الاحصائية غير المناسبة لمستوى القياس، ويعتبر **القياس** من أهم الامور التي يعتمد عليها **الاحصاء**، ويتم القياس للمتغيرات وفق المستويات التالية:

- **المقاييس/ البيانات الاسمية données nominales** : وهو أدنى مستوى للقياس ويستخدم في معظم الأحوال مع المتغيرات النوعية (الكيفية)، حيث يتم التصنيف طبقا لخصائص نوعية مثل: الجنس (ذكر، أنثى)، منطقة السكن (ريف، مدينة)، الجنسية (جزائرية، فرنسية).

فإذا حدد الباحث رقم 1 ليدل على أن المفحوص ذكر والرقم 2 ليدل على أن المفحوص أنثى، والأمر نفسه بالنسبة لمنطقة السكن والجنسية. فهذا لايعني أن 2 أكبر من 1 لأن رقم هنا ليس له معنى كمي، وإنما يؤدي وظيفة التصنيف فقط.

- **المقاييس/ البيانات الرتبيةdonnées ordinales** : وهو يلي في المستوى المتغير الاسمي، وهو بالاضافة إلى تصنيف الأفراد في مجموعات متمايزة يظهر ترتيبهم تصاعديا أو تنازليا في صفة أو خاصية.

وكمثال على ذلك ترتيب 5 طلبة حسب فعاليتهم في مادة الاحصاء واعطاء 5 لأكثرهم نشاطا و1 لأقلهم نشاطا. فإن الفرق بين الطالب الاول والثاني في درجة الفعالية لايشترط أن يكون مساويا للفرق بين الثالث والرابع....الخ. كما لايشنرط أن يكون الطالب الأول 5 أمثال الطالب رقم5

- **مقاييس/ بيانات المسافات المتساويةdonnées intervalles égaux** : هنا تكون الفروق بين المستويات المتتالية متساوية، حيث يسمح بتحديد الفرق بين كل مستويين، وهذا يعني أن لللمتغير هنا وحدة قياس إلا أن 0 نقطة البداية اختيارية (افتراضية) أي لاتعني غياب الظاهرة أو الخاصية المقاسة. في هذا المستوى يمكن اجراء العمليات الحسابية التقليدية مثل: الجمع أو الطرح.

مقاييس/ بيانات النسبةdonnées de ratio : يتميز بجميع الخصائص السابقة اضافة إلى وجود الصفر المطلق الذي يعني غياب الخاصية مثل: الطول، الوزن. تجدر الاشارة إلى أن المتغيرات في علم النفس لايتعدى قياسها مستوى المسافات المتساوية ولاتصل إلى مستوى النسبة.

هذا وقد لخص Stevens هذه المستويات، طبيعة المتغيرات التابعة لها مع إعطاء أمثلة عن بعض استعمالاتها في الجدول الموالي. (1

المستوى	**إسمي**	**ترتيبي**	**مسافة**	**نسبة**
العمليات القاعدية	تحديد فقط المساواة من عدمها (x1=x2)	تحديد الرتبة	تحديد تساوي المسافات والفروق (x1−x2=x3−x4)	تحديد تساوي النسب (x1/x2=x3/x4)
التحويلات الممكنة	استبدال قيمة بقيمة أخرى (f(x)=y)	الترتيب التصاعدي أو التنازلي للقيم (f(x)=y)	وظيفة التقريب (y=ax+b)	وظيفة الضرب (y=ax)
طبيعة المتغيرات	منفصلة	منفصلة	متصلة	متصلة
مقاييس النزعة المركزية الممكنة	المنوال	الوسيط	المتوسط الحسابي	المتوسط الهندسي والتوافقي
مقاييس التشتت المتاحة	entropie	percentiles	التباين والانحراف المعياري	تحليل التباين
العلاقات بين	X^2	معاملات ارتباط	معاملات الارتباط	معاملات الارتباط

المتغيرات		الرتب	والانحدار	والانحدار
التمثيلات البيانية	المدرج التكراري وجداول العرض	التمثيلات البيانية الأخرى	المنحنيات الرياضية	المنحنيات الرياضية
أمثلة	الاستبيانات والسلالم الكيفية (Pougeon 1990)	سلالم Lickert	سلم Borg مثلا	مقاييس Stevensو Thurstone

جدول (1): ملخص عن مستويات القياس.

تجدر الاشارة إلى أنه يمكن تحويل البيانات المحصل عليها في القياس من المستوى الاكثر دقة (مستوى المسافات والنسبة) إلى المستوى الاقل دقة (المستوى الاسمي) والعكس غير ممكن.

- **الفرضية الإحصائية:** هي جملة علمية تعبر عن توقع أو احتمال أو اجابة مؤقتة لسؤال يضعة الباحث ويحاول التحقق منه احصائيا، وهي مرتبطة مباشرة بفرضيات البحث.

تصاغ الفروض الاحصائية في شكل صفري أو بديل.

***الفرض الصفري** H_0: يفترض الباحث أن العلاقة بين متغيرين أو الفرق بينهما يساوي صفر

مثال:

لا توجد علاقة ارتباطية ذات دلالة احصائية بين متغيري الدافعية والأداء لدى عمال شركة سونلغاز. $X1-x2=0$.

- لا توجد فروق ذات دلالة احصائية في الاداء بين العمال الذين التحقوا بتكوين والعمال الذين لم يلتحقوا بتكوين. $n1-n2=0$

- **الفرض البديل** H_1**:يفترض** الباحث أن هناك علاقة بين متغيرين أو فروق متوقعة بينهم.

مثال:

- توجد علاقة ارتباطية ذات دلالة احصائية بين متغيري الدافعية والأداء لدى عمال شركة سونلغاز. $X1-x2\neq0$
- توجد فروق ذات دلالة احصائية في الاداء بين العمال الذين التحقوا بتكوين والعمال الذين لم يلتحقوا بتكوين. $n1-n2\neq0$.
- **مستوى الدلالة:** هو المستوى الذي يطمئن عنده الباحث من صحة نتائجه وأنها لا تعود للصدفة. ويتم الكشف عنها من خلال جداول احصائية خاصة وذلك بعد تحديد القيمة المحسوبة. وتكون هذه الجداول غالبا في ملاحق كتب الاحصاء.

وهناك ثلاث مستويات دلالة مقبولة احصائيا:

1- مستوى دلالة 0.001، أي أن هناك ثقة في النتائج التي توصلت إليها كباحث بنسبة0.999 مقابل شك بنسبة 0.001 أي أن كل 1000مرة نقوم فيها بحساب معامل الارتباط مثلا، هناك 999 مرة صواب مقابل مرة واحدة محتملة للخطأ.

2- مستوى دلالة 0.01، أي أن هناك ثقة في النتائج التي توصلت إليها كباحث بنسبة0.99 مقابل شك بنسبة 0.01

أي أن كل 100مرة نقوم فيها بحساب معامل الارتباط ، هناك 99 مرة صواب مقابل مرة واحدة محتملة للخطأ.

3- مستوى دلالة 0.05، أي أن هناك ثقة في النتائج التي توصلت إليها كباحث بنسبة0.95 مقابل شك بنسبة0.05

أي أن كل 100مرة نقوم فيها بحساب معامل الارتباط ، هناك95مرة صواب مقابل 5 مرات محتملة للخطأ. ويعتبر هذا المستوى من الدلالة أقل مستوى نقبله كباحثين.

- **اختبار الفروض الاحصائية:** إن اختبار الفروض بأسلوب احصائي يؤدي إلى اتخاذ قرار اذا ماكان الفرض مقبولا أم مرفوضا. تجدر الاشارة إلى أن قبول الفرض لايعني بالضرورة أن يكون صحيحا، كما رفض الفرض لايعني بالضرورة أن يكون خاطئا. والجدول التالي يوضح ذلك:

القرار / الفرضية	(H0) صحيح	(H0) خاطيء
قبول(H0)	**صواب $(1-\alpha)$**	**خطأ2(خطا من النوع الثاني β)**
رفض(H0)	**خطأ1(خطا من النوع الاول α**	**صواب $(1-\beta)$**

فإذا كان H0 صحيحا ولم يتمكن الباحث من رفضه (قبله الباحث) فهو قرار صائب

أما إذا كان H0 خاطئا ورفضه الباحث (لم يقبله) فهو قرار صائب

أما عند رفض H0 وهو صحيح فالقرار خاطئ

اما عند رفض H0 وهو خطأ فالقرار خاطئ

يمكن توضيح نوعي الخطأ المبينين من خلال المثالين الأتيين**(صلاح الدين محمود علام 2005ص102)** :

مثال1: نفترض أن التغذية الرجعية ليس لها تأثير بالفعل على سلوك حل المشكلة، ولكننا لاحظنا عن طريق الصدفة أن سلوك حل المشكلة كان أفضل في وجود التغذية الرجعية، فإننا ربما نستنتج أن الرجعية تؤدي إلى تحسين سلوك حل المشكلة في حين أن الأمر ليس كذلك، وهنا نكون قد وقعنا في خطأ من النوع الأول.

مثال2: عند محاكمة متهم يمكن الوقوع في أي من نوعي الخطأ، فتجريم شخص بريء يعد خطأمن النوع الأول، وتبرئة شخص مذنب يعد خطأمن النوع الثاني

المحاضرة الثانية:

(معامل الارتباط بيرسونrp)

- **الارتباط بين متغيرين كميين**

يعتبر معامل الارتباط برافيس بيرسونBravis pearson والذي يرمز له بالرمز"rp" أحد الاختبارات الاحصائية البارامترية، ومن اكثر معاملات الارتباط استعمالا وهذا عندما تكون بيانات كلا المتغيرين كمية أي من مستوى قياس مسافات متساوية او نسبة مثل: العلاقة بين متغير الاقدمية في العمل ومتغير الدخل أو العلاقة بين الاجر والاداء في العمل.

شروط تطبيق rp

- بيانات كلا المتغيرين كمية (مستوى القياس مسافات متساوي أو نسبة)
- التوزيع الاعتدالي لبيانات كلا المتغيرين
- أن لا يقل حجم العينة عن 50 فردا (لضمان اقتراب توزيع البيانات من الاعتدالية)
- أن تكون العلاقة بين المتغيرين(x) و ()yخطية أي: كل زيادة في المتغيرx تصحبها زيادة في المتغير y أو أن كل تناقص في المتغير x يصاحبه تناقص في المتغير y ، أو أن الزيادة في المتغير x تصاحبه نقصا في المتغير y أو التناقص في المتغير xتصحبه زيادة في المتغير y

للتأكد من خطية العلاقة نقوم برسم لوحة الانتشار حيث يتم تمثيل احد المتغيرين على المحور الافقي (x) وقيم المتغير الاخر على المحور العمودي (y) حيث يتم تمثيل كل زوج من القيم المتناظرة بنقطة واحدة في المستوى.

<u>الارتباط :</u> هو تعيين طبيعة وقوة العلاقة بين متغيرين أو عدمها

<u>معامل الارتباط :</u>

يعرف معامل الارتباط والذي يرمز له بالرمز r بأنه عبارة عن مقياس رقمي يقيس قوة الارتباط بين متغيرين , حيث تتراوح قيمته بين (1+) و (1-) .

خصائص معامل الارتباط

- معامل الارتباط مقياس وصفي
- تتراوح قيمة معامل الارتباط بين −1 و +1.
- معامل الارتباط يتأثر بالقيم الشاذة.
- إذا كانت قيمة معامل الارتباط قريبة من الصفر فهذا دليل على عدم وجود علاقة خطية بين المتغيرين. أما إذا كانت قيمة المعامل واحد صحيح فهذا دليل على أن العلاقة عكسية تامة ، أما إذا كانت قيمة معامل الارتباط عند الواحد الصحيح الموجب فهذا يدل على العلاقة الموجبة الطردية التامة، وفيما عدا ذلك فإن العلاقة توصف قوية أو متوسطة أو ضعيفة حسب الجدول التالي:
-

- ضعيفة جدا	- صفر – أقل من 0.20
- ضعيفة	- 0.20 – أقل من 0.40
- متوسطة	- 0.40 – أقل من 0.60
- قوية	- 0.60 – أقل من 0.80
- قوية جدا	- 0.80 – أقل من 1.00
- تام	- 1.00

-
- عيوب معامل الارتباط
- مقياس وصفي يقف عند حدود وصف العلاقة بين الظاهرتين ولا يسمح بالتنبؤ بقيمة أحد المتغيرين بمعلومية الآخر.
- لا يوضح العلاقة السببية بين المتغيرين أي أنه لا يميز بين المتغير المستقل والمتغير التابع
- لا يفرق بين العلاقة الحقيقة والعلاقة الناشئة من الصدفة.
- مثلا العلاقة بين درجة الاختبار ورقم السجل المدني تساوي 0.55
-

طرق حساب معاملrp

هناك ثلاثة طرق لحساب معامل الارتباط بيرسون:

- من خلال الدرجات الخام.
- من خلال الدرجات المعيارية.

من خلال الانحرافات المعيارية.

يستخدم معامل الارتباط في المنهج الوصفي للتعرف على طبيعة و قوة العلاقة بين متغيرين أو أكثر، فعندما يلاحظ تغير في المتغير (X) يتبعه تغير في المتغير (Y).

تتراوح قيمة معامل الارتباط بين −1 و +1

– معامل الارتباط بيرسون: لدرجات الخام

يتم استخدام هذا المعامل لدراسة الارتباط بين متغيرين من المستوى الكمي خلال المعادلة التالية:

$$r = \frac{n\sum(xy) - (\sum x)(\sum y)}{\sqrt{\left[n\sum x^2 - (\sum x)^2\right]\left[n\sum y^2 - (\sum y)^2\right]}}$$

حيث:

r معامل الارتباط

n حجم العينة

X Y متغيران

– **شكل الانتشار Scatter Diagram** أو مخطط الانتشار Scatter plot فهو أداة بيانية مفيدة ، بل أمر أساس قبل البدء بأي تحليل للارتباط وذلك لمعرفة المنحى العام للبيانات ، وشكل الانتشار مخطط يتم فيه رسم كل درجة من درجات الشخص للمتغير الأول مقابل درجته على المتغير الثاني – *كما يمكن رسم درجته على متغير ثالث*– ويستعمل لعرض العلاقة بين متغيرين ، إذ يعطي فكرة سريعة عن العلاقة واتجاهها دون حساب معامل الارتباط ، ويمكن رسم خط الملائمة الأفضل Best Fit Line او ما يعرف بخط الانحدار لأجراء المقارنة المنظورة بين هذا الخط وبين النقاط حوله والتي تمثل تقاطع قيم المتغيرين موضوع الدراسة ، فكلما كانت مجموعة النقاط قريبة من هذا الخط كلما

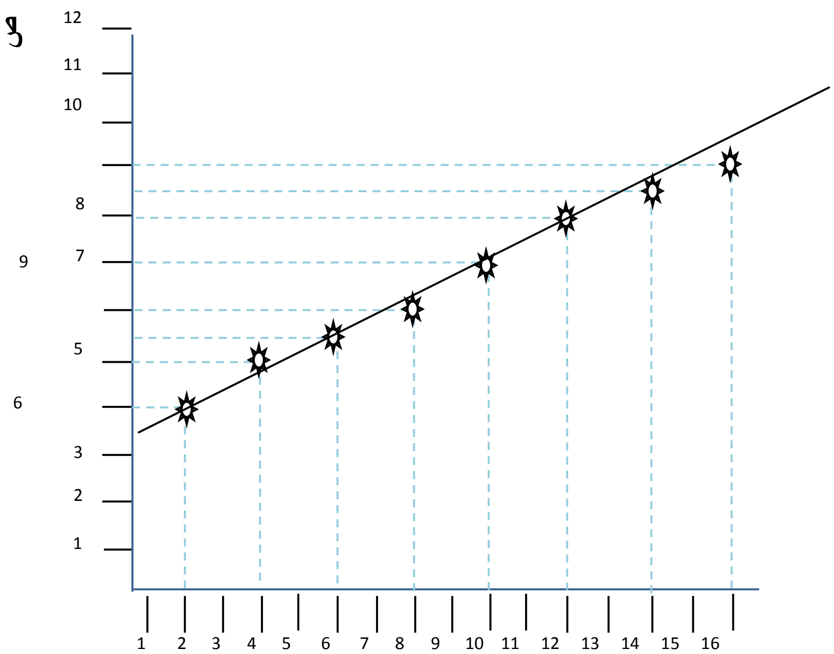
كانت العلاقة بين المتغيرين أقوى ، وكذلك كلما كانت هذه النقاط مبعثرة أكثر كانت العلاقة بين المتغيرين ضعيفة.

وسوف نستعرض مجموعة من الأمثلة توضح شكل الانتشار وتمثل العلاقة بيانياً وكيفية رسم مخطط الانتشار :

مثال : لمعرفة علاقة متغير العمر بمتغير الوزن لـ (8) أطفال كانت بياناتهم كالآتي :

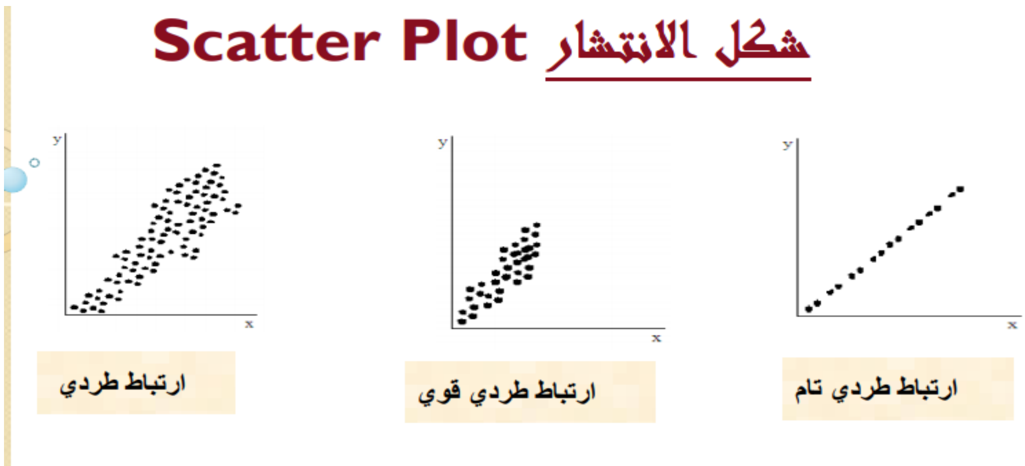
الاطفال	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
الوزن *بالكغم* **س**	4	5	5.5	6	7	8	8.5	9
العمر *بالشهر* **ص**	2	4	6	8	10	12	14	16

ولمعرفة شكل الانتشار واستنتاج نوع العلاقة من الرسم نقوم بتمثيل البيانات بيانياً ، ونقوم بتعيين النقطة الخاصة بكل طفل وذلك بمعرفة القيمتين (س) ، و(ص) على المحورين السيني (X–axes) والصادي (Y–axes) ، فالطفل الأول الذي حصل على قيمة مقدارها (4) في متغير الوزن (س) ، وقيمة مقدارها (2) في متغير العمر (ص) تتحدد النقطة الخاصة به بواسطة اقامة عمودين احدهما على المحور السيني عند القيمة (س) ، والآخر على المحور الصادي عند القيمة (ص) ، ويشكل التقاء العمودين النقطة المطلوبة لذلك الطفل وهكذا لجميع القيم ، وبعد الانتهاء من تعيين النقاط كافة يمكن ان تتضح لنا نوعية العلاقة من خلال اتجاه النقط وشكل انتشارها وكما موضح في الشكل الآتي:



4

ومن الشكل ومخطط الانتشار يتضح ان هناك علاقة موجبة بين متغيري الوزن والعمر للأطفال الثمانية



مثال:

البحث عن العلاقة بين متغيري عدد مرات التغيب و التحصيل الدراسي لمجموعة من التلاميذ. و هذا وفقا لمعطيات الجدول التالي:

n	التغيب X	التحصيلY	X^2	Y^2	X.Y
1	10	3	100	9	30
2	1	12	1	144	12
3	15	1	225	1	15
4	4	8	16	64	32

5	3	7	9	49	21
6	2	10	4	100	20
7	1	15	1	225	15
8	6	6	36	36	36
9	15	2	225	4	30
10	2	19	4	361	38
∑10	59	83	621	993	249

لقياس العلاقة بين متغير التحصيل و متغير التغيب نحتاج أولا إلى حساب الحدود التالية: $\sum X^2$ $\sum Y$ $\sum X$ $\sum Y^2$ $\sum(XY)$

ثم نطبق المعادلة كما يلي:

$$r = \frac{(10\sum(249)-(\sum 59)(\sum 83))/}{\sqrt{[10(621)-(59)^2][10(993)-(83)^2]}}$$

r= -0.83

إن العلاقة بين التحصيل الدراسي و الغياب علاقة عكسية سالبة قوية، ذلك أنه كلما زاد الغياب قل التحصيل الدراسي، و كلما قل الغياب زاد التحصيل الدراسي.

2- معامل ارتباط الرتب لسبيرمان:

يستخدم معامل الارتباط الرتبي لسبيرمان لإيجاد درجة العلاقة بين متغيرين كل منهما من المستوى الرتبي، أو يكون أحد المتغيرين رتبي و الثاني كمي يحول إلى رتبي، أو كلاهما من المستوى الكمي و يحولان إلى المستوى الرتبي. رلدراسة الارتباط في حالة متغيرات كيفية، حيث تستعمل رتبا تصاعدية أو تنازلية عوضا عن القيم العددية للمتغيرات المدروسة. و يعطى معامل سبيرمان من خلال المعادلة التالية:

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

rs : معامل الارتباط بيرسون

1 و 6 : ثوابت

D : الفرق بين رتب نفس الفرد على المتغير X و Y

D^2 : مربع الفرق بين رتب نفس الفرد على المتغيرين

n: حجم العينة

مثال 1:

يبين الجدول التالي الدرجة التقديرية لستة طلبة في إمتحاني الإحصاء و الإقتصاد

المطلوب: حساب معامل الارتباط بين الدرجات التقديرية

رقم الطالب	الدرجة التقديرية في الاحصاء	الدرجة التقديرية في الاقتصاد	رتب الاحصاء	رتب الاقتصاد	D_i	D^2
1	مقبول	ضعيف	4	5	−1	1
2	جيد جدا	ممتاز	2	1	1	1
3	جيد	جيد	3	3	0	0
4	ضعيف	ضعيف جدا	5	6	−1	1
5	ضعيف جدا	مقبول	6	4	2	4
6	ممتاز	جيد جدا	1	2	−1	1
Σ	/	/	/	/	/	8

$$rs = 1 - \frac{6\,(8)}{6(\,36 - 1)}$$

rs= 0.77

و عليه فالعلاقة بين المتغيرين طردية و قوية، يعني هذا بأن الطالب المتفوق في الإحصاء هو أيضا متفوق في علم الإقتصاد.

مثال2:

إنتاج مجموعة من العمال تابعوا فترات تدريبية مختلفة معبر عنها بالسنوات. نريد معرفة هل يوجد ارتباط بين مدة التكوين و الإنتاج باستخدام معامل الارتباط سبيرمان

مدة التكوينX	الإنتاجY	رتب X	رتبY	D	D^2
6	36	3	20	17−	289
8	23	5.5	6	0,5−	0.25
6	21	3	3	0	0
5	20	1	1	0	0
10	25	8	9.5	1,5 −	2.25
12	26	10.5	11	0.5 −	0.25
14	28	12	13	1 −	1
15	29	13	14.5	1.5 −	2.25
12	27	10.5	12	1.5 −	2.25
20	24	20	8	12	144
6	21	3	3	0	0
8	23	5.5	6	0.5 −	0.25

10	23	8	6	2	4
16	25	15.5	9.5	6	36
10	21	8	3	5	25
18	29	18.5	14.5	4	16
16	30	15.5	16.5	-1	1
18	32	18.5	19	-0.5	0.25
16	31	15.5	18	-2.5	6.25
16	30	15.5	16.5	-1	1
∑=242	∑=524	/	/	/	∑=531

ملاحظة: قمنا بالترتيب التصاعدي من أصغر قيمة إلى أكبر قيمة، و في حالة تكرار مجموعة من الرتب، فإننا نحسب المتوسط الحسابي لهذه الرتب، ثم نعطي كل الرتب المكررة نفس الرتبة التي هي المتوسط الحسابي لهذه الرتب.

مثال: تجدون رتبة القيمة 6 لقيم x هي 3 كما في الجدول. فإذا رتبنا قيم x تصاعديا فنجد القيم الأولى بتكراراتها كما يلي 5، 6، 6، 6 و بالتالي 6 تتكرر ثلاث مرات فهي بهذا الشكل تأخذ الرتبة الثانية، الثالثة و الرابعة و عليه نأخذ متوسط الرتب الثلاثة لرقم 6 كما يلي: 3 =3/(2+3+4) و على هذا الأساس أخذ الرقم 6 المتعلق ب x الرتبة 3.

$$rs = 1 - \frac{6\ \Sigma D^2}{n(n^2 - 1)}$$

و عليه $rs = 1 - \frac{6\,(531)}{20(20^2-1)}$

و بالتالي $rs = 0.59$

ومنه نستنتج أن العلاقة بين مدة التكوين و الإنتاج علاقة طردية متوسطة

اساليب ارتباطية بين متغيرين اسميين معامل الارتباط فاي

$(\emptyset\ \boldsymbol{Phi})$

يطبق معامل ارتباط فاي ϕ في حالة حساب العلاقة بين متغيرين اسميين منفصلين ويكون كلاهما لديه تقسيما ثنائيا. وكمثال على ذلك استجابة الشخص على استبيان حول رايه في التعليم المختلط في المرحلة الابتدائية (x) والتعليم المختلط في المرحلة المتوسطة(y)، وكانت بدائل الاجابة على السؤالين بنعم أو لا (سعيد التل 2006 ص 175).

او كانت لدينا اجابة ثنائية (نعم–لا) على سؤلين(x.y) من اختبار نفسي، وكان المطلوب التعرف على الارتباط بين هذين السؤالين.

يمكن تصنيف استجابة الافراد من خلال المثال الاول والمثال الثاني في جدول من 04 خلايا، كما يلي:

x / y	نعم	لا
نعم	A	B
لا	C	D

حيث : A ,B, C, D هي المشاهدات في صورة تكرارات والموزعة على الاقسام المختلفة لهذين المتغيرين أو السؤالين. والقانون الذي يستخدم لحساب معامل ارتباط فاي ...

$$\phi = \frac{AD-BC}{\sqrt{(A+B)(C+D)(A+C)(B+D)}}$$

A: عدد الافراد الذين أجابوا بنعم على x ونعم علىy

B: عدد الافراد الذين اجابوا بلا علىx ونعم على y

C: عدد الافراد الذين اجابوا بنعم على x ولا على y

D: عدد الافراد الذين اجابوا بلا على xولا على y

- **لمعرفة الدلالة الاحصائية لمعامل فاي.ϕ عند مستوى معين، علينا أن نحول قيمة فاي المحسوبة إلى Z كما يلي:**

$$Z = \phi\sqrt{n}$$

وبذلك تتحول قيمة معامل الارتباط إلى Z التي تكون قيمها الحرجة للرفض والقبول كما هو معروف عند مستوى 0.05 هي $\pm$ 1.96

وعند 0.01 هي $\pm$ 2.58

ملاحظة: للحصول على قيمة z في الجدول نأخذ مستوى الثقة ونقسمه على 2

مثال: $\frac{0.95}{2}=$ 0.4750 ونلاحظ القيمة المقابلة لها في جدول Z

تمارين:

1- جاءت بيانات الاستجابة على سؤالين من اسئلة ايزنك للشخصية كما هي موضحة في الجدول:

y \ x	نعم	لا
نعم	5	9
لا	13	4

هل العلاقة بين استجابات المفحوصين على السؤالين دالة احصائيا؟

المحاضرة الثالثة
التنبؤ

كيف نستطيع التنبؤ بظاهرة معينة من خلال معرفة ما له علاقة بهذه الظاهرة.

(إذاكان كذا ، سيكون كذا)

التنبؤ علاقة بين متغيرين $[\ y = f(x)\ ]$ $x\ ,\ y$ من خلال f وهي دالة تربط بين متغيرين يفترض الباحث أن أحدها تابع والآخر مستقل

والتنبؤ هو معرفة أحد المتغيرين بمعلومية الآخر ، والمتغير المطلوب معرفته يسمى (المتغير التابع) ، والمتغير المعلوم يسمى (المتغير المستقل)

وهذا هو أبسط وأسهل أنواع التنبؤ $[\ y = f(x)\ ]$، فالبساطة جاءت من عدد المتغيرات فهو أقل ما يكون. وقد تكون أكثر من متغير $[\ y = f(x_1, x_2, x_3 \ldots\ldots\ldots\ldots x_n\ ]$ وهذا أكثر واقعية لأن الظاهرة تتأثر بعد متغيرات وليس متغيرا واحدا. فالتحصيل لا يتأثر بالذكاء فقط بل أيضا بالبيئة وبالمثابرة وبتعلم الوالدين ....الخ

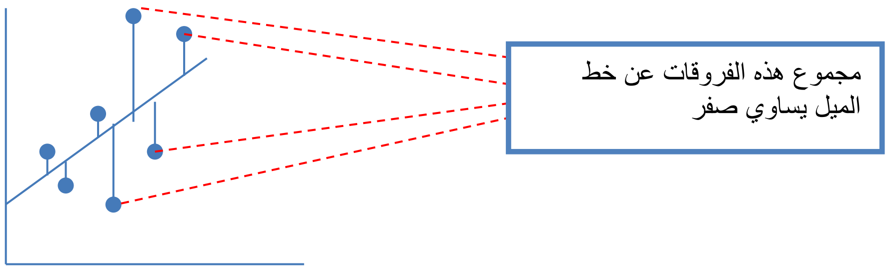
فالسهولة هنا جاءت من المعادلة الرياضية فهذه الدالة فقط عن العلاقة الخطية ، ولكن هناك صور للعلاقات غير خطية ، فهناك العلاقات المنحنية وعلاقات تأخذ شكل حرف S وعلاقات عكسية وعلاقات لوغارتمية وعلاقات أسية.

فيجب معرفة طبيعة العلاقة بين الظاهر هل هي خطية أم لا ، ويكون ذلك من خلال الرسم للبيانات.

وهناك طرق أخرى بالحسابات لكنها أصعب.

أسلوب تحليل الانحدار

الأساس العلمي الذي بني عليه أسلوب تحليل الانحدار يعتمد على طريقة المربعات الصغرى **(OLS)** Ordinary Least Squares وتتلخص هذه الطريقة في أن تجعل مجموع مربعات الأخطاء أقل ما يمكن.
الفرق بين التقدير والمعلمة يسمى الخطأ والمجموع يساوي صفر $(\mu - \bar{x}) = 0$



تفترض أن معادلة الانحدار الخطي البسيط بين متغيرين أحدهما تابع والآخر مستقل على الصورة التالية:

$$Y = a + bx + \epsilon$$

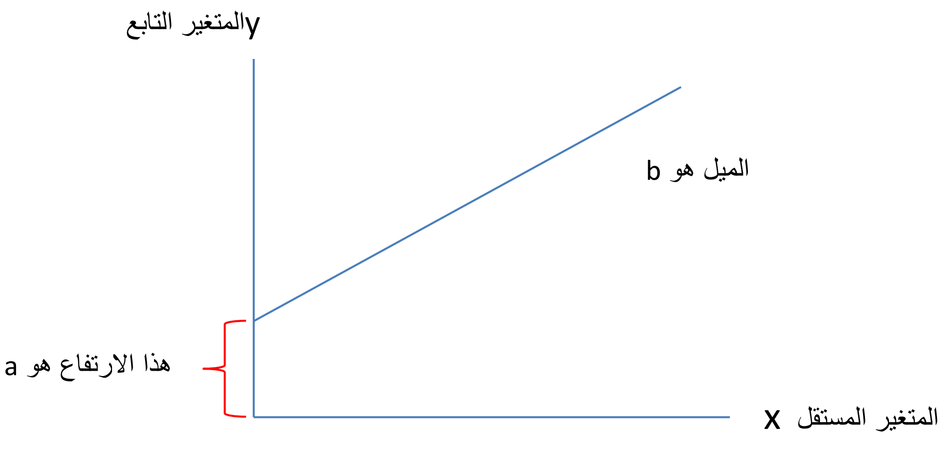
المتغير التابع (Y)

الجزء المقطوع من المحور الرأسي (ثابت) (a)

ميل الخط المستقيم على المحور الأفقي (b)

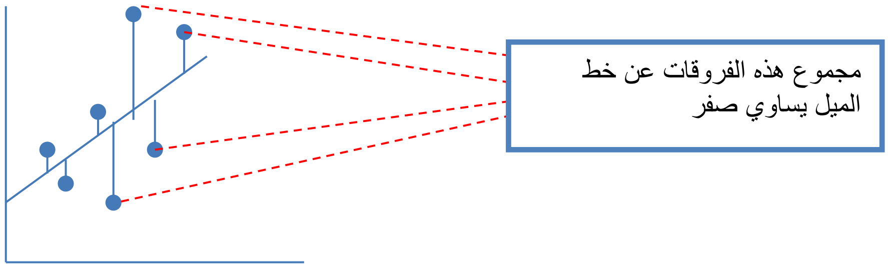
المتغير المستقل (x)

الخطأ العشوائي (ϵ)



$$y = a + bx + \epsilon$$

يبقى أن نعرف كيف نقدر (a) و (b) لذلك يجب أن يختفي الخطأ العشوائي(ϵ) من المعادلة ، وذلك من خلال طريقة المربعات الصغرى التي تجعل مجموع مربعات الأخطاء أقل ما يمكن (أي صفر) كما في الشكل التالي:



وتصبح المعادلة هكذا: $\hat{Y} = \hat{a} + \hat{b}x$

و نلاحظ أن x ليس عليها علامة ̂ لأنها معلومة في الأصل

حيث أن ̂ عبارة عن : (معامل الارتباط × الانحراف المعياري y / الانحراف المعياري x) وتحسب بالمعادلة التالية:

$$\hat{b} = r \times \frac{S_y}{S_x}$$

وتدل (b) على كمية الزيادة في (y) إذا زادت (x) بوحدة واحدة.

كما أن $\hat{a}$ عبارة عن الثابت في المعادلة ، تحسب عن طريق المعادلة التالية:

$$\hat{a} = \bar{y} - \hat{b}\bar{x}$$

حيث $(\bar{x})$ متوسط المتغير المستقل و $(\bar{y})$ متوسط المتغير التابع

=====

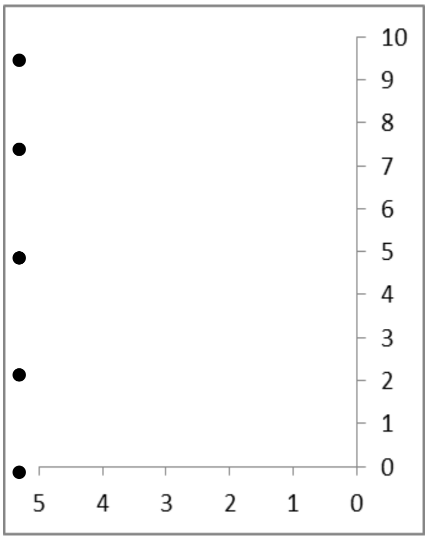
معامل التحديد r^2(التفسير) : يتراوح بين (صفر و 1) ، دائما موجب فهو مربع معامل الارتباط.

انه نسبة مساهمة المتغير المستقيل في تفسير التغيرات التي تحدث في المتغير التابع

تمرين1 :

- أوجد معامل الارتباط بين المتغيرين (x) و(Y).
- احسب معامل التحديد واشرح معناه
- قدّر معادلة انحدار (y) على (x) [معادلة التنبؤ] $\hat{\ } = \hat{a} + \hat{b}(x)$
- ارسم الشكل الانتشاري للعلاقة بين المتغيرين، مع رسم خط ميل الانحدار.

اسم الطالب	x	y	$(x-\bar{x})$	$(y-\bar{y})$	$(x-\bar{x})^2$	$(y-\bar{y})^2$	$(x-\bar{x})(y-\bar{y})$
محمد	1	2					
فهد	2	4					
سعد	3	6					
خالد	4	8					
سعود	5	10					
المجموع∑	15	30					



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

$$\bar{y} = \frac{\sum y}{n} = \frac{\ }{\ } =$$

$$S_x = \sqrt{\frac{\sum(x-\bar{x})^2}{n}} = \sqrt{\frac{\ }{\ }} = \sqrt{\ } =$$

$$S_y = \sqrt{\frac{\sum(y-\bar{y})^2}{n}} = \sqrt{\frac{\ }{\ }} = \sqrt{\ } =$$

$$r = \frac{\sum(x-\bar{x})(y-\bar{y})}{\sqrt{\sum(x-\bar{x})^2 \sum(y-\bar{y})^2}} = \frac{\ }{\sqrt{\ \times\ }} = \frac{\ }{\sqrt{\ }} = \frac{\ }{\ } =$$

- **التعليق :** توجد علاقة ارتباط .... بين عدد ساعات المذاكرة والدرجة التي يحصل عليها الطالب
- **معامل التحديد r^2يساوي** .... أي أن المتغير (x) يفسر التغيرات التي يحدث في المتغير(y) بنسبة ....%
- $\hat{b} = r \times \frac{S_y}{S_x} = \quad \times \frac{\ }{\ } = \quad \times \quad =$
- $\hat{a} = \bar{y} - (\hat{b} \times \bar{x}) = \quad -(\quad \times \quad) =$

إذن معادلة انحدار(y)على(x) في هذا المثال هي: $\hat{y} = \hat{a} + \hat{\ }\ (x)$

$\hat{y} = \quad + \quad (\ )$

تطبيق

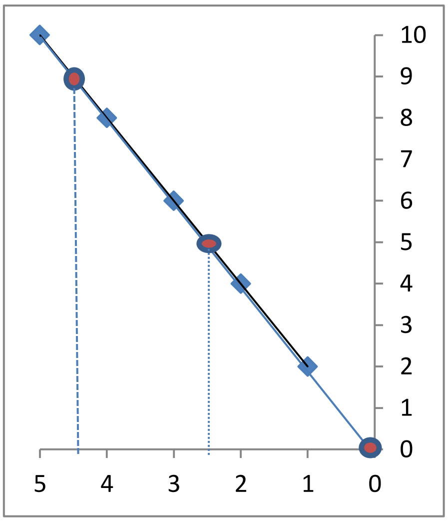
- **تنبأبالقيم التي ستكون للمتغير (y) في حالات قيم (X) التالية ، ثم حدد مكانها على الشكل الانتشاري:**

x=4.5	x=2.5	x= 0
$\hat{y} = \quad + \ (\quad) =$	$\hat{y} = \quad + \ (\quad) =$	$\hat{y} = \quad + \ (\quad) =$

حل تمرين 1 :

- أوجد معامل الارتباط بين المتغيرين (x) و(Y).
- احسب معامل التحديد واشرح معناه
- قدّر معادلة انحدار (y) على (x) [معادلة التنبؤ] $\hat{\ } = \hat{a} + \hat{b}(x)$
- ارسم الشكل الانتشاري للعلاقة بين المتغيرين، مع رسم خط ميل الانحدار.

اسم الطالب	x	y	$(x-\bar{x})$	$(y-\bar{y})$	$(x-\bar{x})^2$	$(y-\bar{y})^2$	$(x-\bar{x})(y-\bar{y})$
محمد	1	2	-2	-4	4	16	8
فهد	2	4	-1	-2	1	4	2
سعد	3	6	0	0	0	0	0
خالد	4	8	1	2	1	4	2
سعود	5	10	2	4	4	16	8
المجموع∑	15	30			10	40	20



$$\bar{x} = \frac{\sum x}{n} = \frac{15}{5} = 3$$

$$\bar{y} = \frac{\sum y}{n} = \frac{30}{5} = 6$$

$$S_x = \sqrt{\frac{\sum(x-\bar{x})^2}{n}} = \sqrt{\frac{10}{5}} = \sqrt{2} = 1.41$$

$$S_y = \sqrt{\frac{\sum(y-\bar{y})^2}{n}} = \sqrt{\frac{40}{5}} = \sqrt{8} = 2.82$$

$$r = \frac{\sum(x-\bar{x})(y-\bar{y})}{\sqrt{\sum(x-\bar{x})^2 \sum(y-\bar{y})^2}} = \frac{20}{\sqrt{10\times 40}} = \frac{20}{\sqrt{40}} = \frac{20}{20} = 1.00$$

- **التعليق** : توجد علاقة ارتباط تامة بين عدد ساعات المذاكرة والدرجة التي يحصل عليها الطالب
- **معامل التحديد r^2يساوي 1.00**أي أن المتغير (x) يفسر التغيرات التي يحدث في المتغير(y) بنسبة 100%
- $\hat{b} = r \times \frac{S_y}{S_x} = 1.00 \times \frac{2.82}{1.41} = 1.00 \times 2.00 = 2.00$
- $\hat{a} = \bar{y} - \hat{b}\bar{x} = 6 - 2(3) = .0$

- إذن معادلة انحدار(y)على(x) في هذا المثال هي: $\hat{y} = 0 + 2(x)$

تطبيق

- **قدّر القيم التي ستكون للمتغير (y) في حالات قيم (X) التالية ، ثم حدد مكانها على الشكل الانتشاري:**

x=4.5	x=2.5	x= 0
$\hat{y} = 0 + 2(4.5) = 9$	$\hat{y} = 0 + 2(2.5) = 5$	$\hat{y} = 0 + 2\ (0) = 0$

التوزيعات الاحتمالية و تطبيقاتها

يجب التمييز هنا بين نوعين من التوزيعات الاحتمالية، فهناك التوزيعات الاحتمالية المنفصلة والتوزيعات الاحتمالية المتصلة .

والأمثلة الاحتمالية التي تحتوي على توزيعات احتمالية منفصلة كثيرة، ولكل تجربة مجموعة منفصلة من النتائج الممكنة ولكل نتيجة احتمال مرتبط بها. وكمثال بسيط على فكرة التوزيع الاحتمالي المنفصل هو قائمة التسجيل التي نحصل عليها عند رمي حجر الفرد. فالمتغير يمكن أن يأخذ ستة قيم منفصلة1 ، 2، 3، 4، 5، 6واحتمال كل نتيجة هو. $\frac{1}{6}$

أما التوزيعات الاحتمالية المتصلة فإننا نتعامل فيها مع متغيرات غالباً ما تأخذ أي قيم على الأقل في مدى معين. وعند التعامل مع متغيرات التوزيعات المتصلة لا يمكن الحديث عن احتمال أن يأخذ المتغير قيمة واحدة بعينها كما هو الحال في التوزيعات المنفصلة. ومثال ذلك أننا لا نستطيع توقع احتمال أختيار شخص عشوائياً وزنه 17.325 كغم.

التوزيعات الاحتمالية :

توزيع ذي الحدين

التوزيع الطبيعي

العلاقة بين توزيع ذي الحدين والتوزيع الطبيعي المعتاد

توزيعt

الفرق بين توزيع t والتوزيع الطبيعي

توزيع(X^2 كاي تربيع (

توزيع ذي الحدين

يسمى غالباً التوزيع الاحتمالي المنفصل بتوزيع ذي الحدين. إذا كانت P احتمال وقوع حدث ما (احتمال النجاح) و q = 1 - p احتمال عدم وقوع الحدث (احتمال الفشل. (

فإذا كان احتمال وقوع الحدث X من N من المحاولات، فإن (X) نجاح و (N-X) فشل، وبالتالي فإن احتمال يعطى من خلال :

$$P(X) = {}^{N}C_{X} P^{X} q^{N-X}$$
$$= \frac{N!}{X!(N-X)!} P^{X} q^{N-X}$$

مثال (1 : (

احتمال الحصول على صورتين بالضبط من ستة رميات لعملة غير متحيزة هو :

$${}^{6}C_{2}\left(\frac{1}{2}\right)^{2}\left(\frac{1}{2}\right)^{6-2} = \frac{6!}{2!4!}\left(\frac{1}{2}\right)^{6}$$
$$= \frac{15}{64}$$

مثال (2 : (

احتمال الحصول على 4 صورة من ستة رميات لعملة غير متحيزة

$${}^{6}C_{4}\left(\frac{1}{2}\right)^{4}\left(\frac{1}{2}\right)^{6-4}$$

التوزيع الطبيعي

هو أفضل وأكثر التوزيعات الاحتمالية المتصلة استخداماً .

ويكمن السبب في ذلك:

أن توزيعات كثيرة لمتغيرات مثل الطول والوزن تتبع توزيعات طبيعية .

النتيجة الرياضية التي تسمى بنظرية النهاية المركزية .

نظرية النهاية المركزية تقول بأنه إذا أضفنا عدداً كبيراً كبراً كافياً من المتغيرات العشوائية المستقلة والمتماثلة في التوزيع إلى بعضها بأي طريقة فإن توزيع المجموع سيكون تقريباً هو التوزيع الطبيعي.

منحنى دالة الاحتمال للتوزيع الطبيعي له خاصية شكل الجرس. ويتحدد شكل الجرس تماماً لأي توزيع طبيعي خاصة إذا علمنا الوسط الحسابي μ والانحراف المعياري σ لهذا التوزيع .

تدل قيمة μ على مكان مركز الجرس، كما تدل σ على كيفية الانتشار .

القيمـــة الصغـــيرة لــ σ تعني أنـــه لدينا جرس طويل مدبب، القيمة الكبيرة لــ σ تعني أن الجرس قصير ومفرطح .

المساحة الكلية تحت كل جرس يجب أن تكون واحد صحيح .

إذا كان لدينا توزيع طبيعي ذو وسط حسابي μ وانحراف معياري σ فإن معادلة منحنى دالة كثافة الاحتمال تكون على الصورة التالية :

$$Y = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(x-\mu)^2/2\sigma^2}$$

ملاحظات متعلقة بالتوزيع الطبيعي

إذا كانت X لها توزيع طبيعي، ووسطها الحسابي μ ، وانحرافها المعياري σ فإنه يمكن التعبير عنها :

$$X \sim N(\mu, \sigma)$$

للحصول على قيمة احتمال أن تقع X بين نقطتين X1 و X2 . فإن الاجراء العادي هو الحصول على تكامل الدالة بينX1 , X2 ، وهذا يتم باستعمال طريقة التكامل العددي .

هناك جداول توزيع طبيعي تعطي احتمال وقوع Z بين الصفر وأي قيمة موجبة Z) هي أكبر من صفر وأقل من 0.5 (

القيمة المعيارية لـــ $$Z = \frac{X - \mu}{\sigma} = X$$

معالم التوزيع الطبيعي المعياري . $\mu = 0$ $\sigma = 1$

لحساب احتمال وقوع X بين الوسط الحسابي وأي قيمة موجبة مثل X1 يجب الحصول على

$$Z_1 = \frac{X_1 - \mu}{\sigma} = ؟؟$$

ثم من جداول التوزيع الطبيعي نعرف أن

$$P(\mu \le X \le X_1) = P(X \langle Z \langle Z_1)$$

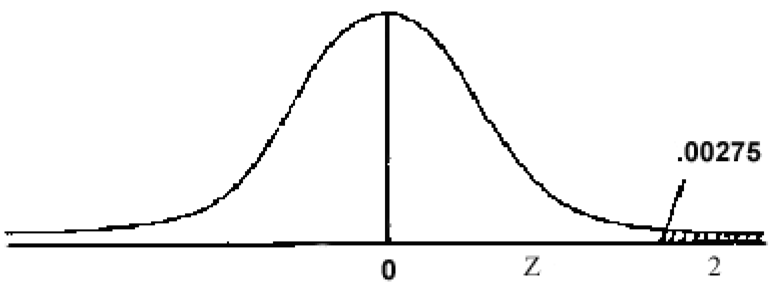
لحساب احتمال وقوع X بين قيمتين X1 و X2 نحسب القيم المعيارية للقيمتين، ثم نستعمل جداول التوزيع الطبيعي .

أمثلة

أمثلة

مثال:(1)

احتمال أن تكون قيمة Z أكبر من 2 :



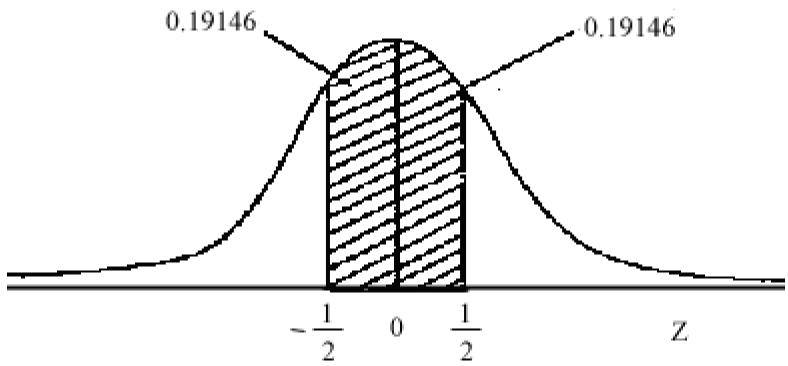
الحل: حيث أن احتمال أن تكون Z أقل من صفر = 0.5000 ومن الجدول احتمال Z في (0،2) = 0.47725 اذن احتمال أن تكون قيمة Z أكبر من 2 هي :

0.02275 = 0.47725 - 0.5000

مثال:(2)

احتمال أن تقع Z بين صفر و 0.5 .

احتمال أن تقع Z بين 0.5 و 0.5 - .



الحل :من الرسم إعادة : المساحة المظللة بين 0 و 0.5 تمثل احتمال أن تقع Z بين 0 و 0.5، والمساحة المظللة إلى شمال 0 ويمين 0.5 - = Z هي احتمال أن تقع Z في الفترة (0.5،0 -) .
واحتمال أن تقع Z في الفترة (0.5،0 = (المساحة المقابلة لقيمة 0.5 Z = هي 0.19146 ، كذلك احتمال أن تقع Z في الفترة0) ، 0.19146 = (- 0.5

اذن احتمال أن تقع Z في الفتـ(-0.5, 0.5) = 2 * 0.19146 = 3292وهي تتمثل

بالمساحة المظللة في الرسم أعلاه .

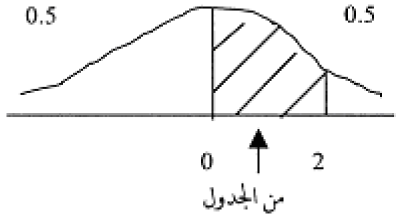
مثال تطبيقي

مثال تطبيقي

مثال : استخدام التوزيع الطبيعي:

يود أحد البنوك دراسة خدماته تجاه عملائه. فإذا كان الايداع اليومي للعملاء يتبع التوزيع الطبيعي بمتوسط 500 د. ك وانحراف معياري 100 د. ك فما هو الاحتمال أن :

-1يقل الإيداع النقدي عن 700 د.ك .

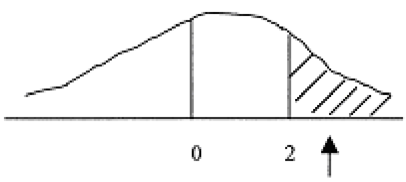


$P(x \leq 700)$

$Z = \frac{700-500}{100} = 2$

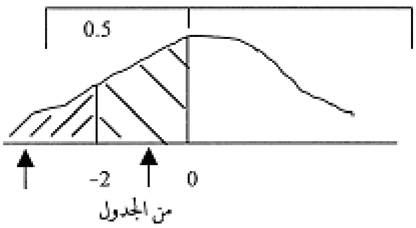
$P(Z \leq Z) = 0.5 + 0.4772 = \underline{0.9772}$

-2أن يزيد الايداع النقدي عن 700 د. ك .



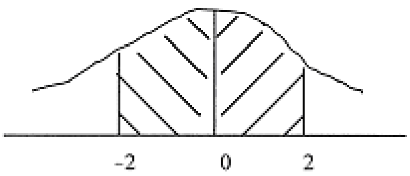
$$P(x > 700) = P(Z > 2)$$
$$= 1 - P(Z \le 2) = 1 - 0.9772$$
$$= \underline{0.0228}$$

-3أن يزيد الايداع النقدي عن 700 د. ك .



$$P(x \le 300)$$
$$Z = \frac{300 - 500}{100} = -2$$
$$\therefore P(Z \le -2) = 1 - P(Z > 2)$$
$$= 1 - 0.9772$$
$$= 0.0228$$

-4أن يتراوح الايداع اليومي بين 300 و 700 د. ك



$$P(300 \le x \le 700) =$$
$$P(-2 \le Z \le 3)$$
$$= 0.4772 + 0.4772 = \underline{\underline{0.9554}}$$

العلاقة بين توزيع ذي الحدين والتوزيع الطبيعي المعتاد

خصائص التوزيع الطبيعي		خصائص توزيع ذي الحدين
μ	الوسط	$\mu = NP$
σ^2	التباين	$\sigma^2 = NPq$
σ	الانحراف المعياري	$\sigma = \sqrt{NPq}$
$\frac{X-\mu}{\sigma}$	Z	$\frac{X-NP}{\sqrt{NPq}}$
$f(X) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2}$	الدالة	$P(X) = \binom{N}{X} P^X q^{N-X}$
• الوسط الحسابي • الانحراف المعياري	← يتحدد بـ →	• نسبة النجاح (p) • نسبة الفشل (1-p)=(q)

إذا كانت N كبيرة (في توزيع ذي الحدين (

وإذا كان كل من q و p غير قريبين من الصفر فإنه يمكن تقريب توزيع ذي الحدين بصورة جيدة إلى التوزيع الطبيعي المعياري.

$$Z = \frac{X - NP}{\sqrt{NPq}}$$

يصبح التقريب أكثر جودة كلما زادت (N) حتى تصبح متطابقة في المالا نهاية∞

يعد التقريب جيداً إذا كان كل من Nq و NP ‏0.5(

توزيع t

إذا كان متوسط العينة $\overline{X}$ مأخوذاً من توزيع طبيعي، فإن هذه العينة متوزعة توزيعاً طبيعياً .

كما تكون العينة موزعة توزيعاً طبيعياً إذا كان حجم العينة كبيراً. وذلك في حالة كون الانحراف المعياري σ معروفاً .

إذا استخدمت S (الانحراف المعياري المحسوب من العينة) لتقدير الانحراف المعياري، فإن التوزيع الاحتمالي في هذه الحالة سيكون توزيع t.

الفرق بين توزيع t والتوزيع الطبيعي

التوزيع الطبيعي	توزيع t
(1) الانحراف المعياري σ معروف $Z = \frac{\bar{X}-\mu}{\sigma/\sqrt{n}}$	(1) الانحراف المعياري (σ) غير معروف، وبالتالي يؤخذ تقدير الانحراف المعياري S وهو $\sigma_{\bar{x}}$.
(2) إذا كان حجم العينة كبيراً (أكبر من 30 أو أكبر من 5% من حجم المجتمع، فإنه يتم تعديل $\sigma_{\bar{x}}$ لتصبح $\frac{\sigma}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}}$ إذا أخذت العينة بدون ارجاع. أما إذا أخذت بارجاع فإن $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$	(2) حجم العينة صغير أو أقل من 5% . وهو قيمته تقديرية. $t = \frac{\bar{X}-\mu}{S_{\bar{x}}} = \frac{\bar{X}-\mu}{S/\sqrt{n}}$
(3) $Z \simeq N(0,1)$	(3) هناك معامل يحدد شكل التوزيع وهو درجات الحرية (n-1)

يقترب توزيع t من الطبيعي عند زيادة حجم العينة فعندما يكون $n \geq 30$ فإنهما يتساويان .

التوزيعات الاحتمالية و تطبيقاتها > توزيع χ^2 (كاي تربيع)

إذا كان $Z_1, Z_2, \ldots, Z_n$ عينة عشوائية حجمها n من توزيع طبيعي معياري أي أز

Z1 ~ N (0,1)فإن مجموع مربعات هذه القيم لها توزيع كاي تربيع أو χ^2 بدرجات حرية n.

$$X^2 = \sum Z_i^2$$

حيث

$$Z_i^2 = \frac{(X_i - \mu)^2}{\sigma^2}$$

شكل المنحنى الاحتمالي يعتمد اعتماداً كلياً على درجات الحرية، فكلما زادت درجات الحرية اتجه المنحنى للتماثل، حيث عندما تكبر قيمة n فإننا يمكن أن نستخدم جداول التوزيع الطبيعي وذلك بأخذ قيمته . $Z = \sqrt{2X_i^2} - \sqrt{2n-1}$

قيم كاي تربيع موجبة دائماً لأنها عبارة عن مجموع مربعات . Z_i

تستعمل جداول القيم الاحتمالية لتوزيع كاي، كما هو الحال في توزيع t.

إذا كانت القيمة X_1^2 لها توزيع كاي تربيع بدرجات حرية n_2 فإن نسـ X_1^2/n_1 إلى X_2^2/n_2

لها توزيع F بدرجات حرية.n_1 ، n_2

$$F_{n_1, m_1} = \frac{X_1^2/n_1}{X_2^2/n_2}$$

استعمالات توزيع كاي تربيع :

يعتبر إختبار مربع كاي من الاختبارات اللابارامترية الأكثر استخداما في البحوث النفسية والتربوية والاجتماعية، وهو يستعمل في حالة البيانات الإسمية(مستوى قياس إسمي/تصنيفي)،حيث يصنف الأفراد أو الأشياء أو الإستجابات في فئات أو أقسام مختلفة. ومثال ذلك الإستجابة على بنود استبيان يحتوي فقرات تتطلب الاجابة عن كل فقرة بديلا من 3(موافق، محايد، معارض).أو :

-من بين عدة تخصصات في علم النفس، ماهو التخصص الذي تختاره؟

علم النفس العيادي علم النفس العمل علم النفس التربوي علم النفس الاجتماعي

إن هدف اختبار مربع كاي هو المقارنة بين التوزيع التكراري المشاهد والتوزيع التكراري المتوقع لإحدى العينات إذا افترضنا وجود توزيع احتمالي للمجتمع أو التحقق ما إذا كانت هناك فروق دالة إحصائيا بين التكرارات المشاهدة والتكرارات المتوقعة لاستجابات مجموعة من الأفراد.

يطبق اختبار مربع كاي في حالة وجود متغير نوعي واحد ويسمى اختبار حسن المطابقة، أو في حالة وجود متغيرين أو أكثر ويسمى كا2 للاستقلالية.

<u>شروط تطبيق كا2</u>

- البيانات نوعية
- مستوى القياس إسمي
- أن تكون العينة منتقاة عشوائيا من المجتمع المستهدف.
- أن يكون حجم العينة أكبر من 30وهذا لأن اختبار كا2 يتأثر تأثرا واضحا بالعينات الصغيرة .

- أن لا يقل أي تكرار متوقع في كل خلية من خلايا الجدول عن5 وذلك لأنه إذاكانت التكرارات المكتوقعة صغيرة جدا فإن توزيع معاينات كا2 لن تكون ممثلة تمثيلا كافيا لتوزيع المتغير موضع البحث.

ويعطى كا^2 بالمعادلة التالية:

- **معادلة اختبار X^2:**

$$x^2 = \Sigma \frac{(f_o - f_t)^2}{f_t}$$

حيث:

f_o يمثل التكرار المشاهد (fréquences observées)

f_t يمثل التكرار النظري (fréquences théoriques)

1-اختبار X^2 لعينة واحدة (حسن المطابقة): بالإضافة إلى الشروط السابقة الخاصة بتطبيقه فإن x^2 لحسن المطابقة يستخدم:

- للتعرف على وجود أو عدم وجود تطابق بين التوزيع المشاهد والتوزيع النظري في خاصية ما.
- وكذا الكشف عن الدلالة الإحصائية للفروق بين التكرارات الملاحظة والتكرارات النظرية.

2- **اختبار X^2 للاستقلالية:** في الكثير من الاحيان يتعرض الباحث بالدراسة إلى ظاهرة تتطلب استخدام أكثر من عينتين يقوم بتصنيف أفرادها في أقسام مستقلة تبعا لمتغيرين أو أكثر من مستوى القياس الاسمي، في هذه الحالة يستوجب عليه- الباحث- تطبيق بعض الأساليب الاحصائية الاستدلالية االتي تناسب هاته البيانات. ويعتبر كا2 للاستقلالية أكثر المقاييس استخداما في التحقق من استقلالية المتغيرات.