بسم الله الرحمن الرحيم

جامعة الملك سعود

كلية الآداب

قسم الجغرافيا

199 جغر

اساليب كمية في الجغرافيا

ا.د فهد بن محمد الكليبي

1438-1437

أهمية المنهج الكمي في الجغرافيا

يعتد المنهج الكمي في الجغرافيا على القياس و جمع المعلومات الجغرافية و تحليلها و اختبار النتائج للوصول إلى نتائج معينة تعيننا على حل كثير من المشاكل الجغرافية. و يعتمد هذا على تطبيق على كثير من الوسائل و الاساليب الاحصائية. لذالك المنهج الكمي:

1- يتطلب الاستعانة بالرياضيات و الاحصاءو إستخدام النماذج و النظم

2- يساعد على اعطاء اوصاف للظواهر الجغرافية على جانب كبير من الدقة

3- يمكن من إستخلاص نتائج عامة من نتائج جزئية (Deduction) و إعطاء تعميم حول ظاهرة معينة

4- يمكن من تحديد أثر متغير دون غيره ا متغيرات دون غيرها في الظاهرة الجغرافية

5- يمكن من فهم العلاقة التي قد تكون معقدة بين العديد من العوامل و المتغيرات.

6- يمكن من اختبار الفرضيات الجغرافية

7- يحدد امكانية الاعتماد على النتائج المتحصل عليها في البحث الجغرافي

المنهج الكمي و الاسلوب العلمي:

الاسلوب العلمي يمر بسبع خطوات ليعطي نتائج موثوق بها وهذي الخطوات هي:

1- الملاحظة المبنية على خبرة

2- الاحساس بمشكلة معينة و تحديدها

3- وضع فرضية (Hypothesis) او فرضيات او تساؤلات حول هذي المشكلة

4- جمع البيانات اوقياسها وترتيبها و تبويبها لغرض البحث

5- تحليل المعلومات و إختبار النتائج و تفسيرها

6- وضع التعميمات و القوانين

و المراحل الثلاث الأخيرة تتطلب معرفة في الاساليب الكمية. لذلك تعتبر المعرفة الكمية مهمة لتطبيق المنهج العلمي الصحيح.

<u>نبذة عن علم الاحصاء:</u>

<u>تعريف علم الحصاء:</u>

هو ذلك العلم الذي يهتم بالطرق الكمية لجمع البيانات وتنظيمها و تبويبها و تلخيصها و عرضها و تحليلها للوصول لنتائج علمية سليمة و من ثم اخذ قررات سليمة على ضوء هذه النتائج.

<u>قسمي علم الاحصاء:</u>

ا- **الاحصاء الوصفي**: هو طرق تنظيم المعلومات و تلخيصها للمساعدة على فهم المعلومات. و الطرق الوصفية تشمل التوزيعات التكرارية و الرسوم البيانية و قياس النزعة المركزية و مقاييس التشتت وغيرها من القياسات الاخرى التي تصف المعلومات.

ب- **الاحصاء الاستدلالى**: وهو الاحصاء الأستنتاجي الذي يهتم بتحليل البيانات للتوصل إلى التنبوء او الاستقراء و إتخاذ القرارت مثل تحليل الانحدار و تحليل التباين و غيرها.

<u>المجتمع الاحصائي (Population) و العينة (Sample)</u>

1- تعريف المجتمع الاحصائي: مجموعة ذات خصائص مشتركة من الافراد محل الدراسة و هي محدود مثل عدد الافراد في 119 جغر و غير محدود مثل عدد حبات القمح في مزرعة.

2- تعريف العينة: جزء من المجتمع يختار بحيث تمثل جميع صفات المجتمع و هي تكون صغيرة او كبيرة منتظمة و عشوائية او طبقية و غيرها. و كلما كبر حجم العينة كانت الاستنتاجات حول السكان ادق.

<u>تعريف البيانات:</u> هي مجموعة من المشاهدات تكون كمية (رقمية) مثل طول او وزن او درجات حرارة، و تكون وصفية مثل لون البشرة او الجنس. هي نوعين رئيسين: المقاسة و المجموعة.

<u>تعريف المعلمة</u>: شييء يميز المجتمع كله مثل متوسط الدخل الشهري للفرد في دولة معينة

تعرف الاحصاءة: شيء يميز العينة مثل متوسط الدخل الشهري لأفراد عددهم 1000 شخص.

تعريف المتغير: هو صفة او رقم او مقدار ينقص او يزيد مع الوقت او يأخذ قيم مختلفة في حالات مختلفة. و هو نوعين: المستقل الذي يأخذ مقادير مختلفة و يسسبب تغير تلك المقادير تغير في متغير ثاني يسمى المتغير التابع. أما النوع الثاني: فهو المتغير التابع الذي يأخذ قيم مختلفة فقط إستجابة للتغير في متغير أخر هو المتغير المستقل.

تفسير رمز الجمع Σ سيجما Sigma

تشير لعدد القراءات وان الجمع يشمل جميع القراءات

$$\sum_{i=1}^{n} X_i$$

العملية الرياضية وهنا جمع بسيط

تعني ان الجمع يبدأ من القراءة الاولى

ينتهي عند القراءة الثانية عشر

$$\sum_{i=3}^{12} X_i$$

يبدأ الجمع من القراءة الثالثة

مثال لرمز جمع $\sum$

لنفرض ان $n=5$ وعناصرها هي

$x_1 = 3$

$x_2 = 1$

$x_3 = 4$

$x_4 = 2$

$x_5 = 1$

يكون

$$\sum_{i=2}^{4} x_i = x_2 + x_3 + x_4 = 1 + 4 + 2 = 7$$

$$\sum_{i=1}^{n} x_i = x_1 + x_2 + x_3 + x_4 + x_5 = 3 + 1 + 4 + 2 + 1 = \underline{\underline{11}}$$

$$\sum_{i=1}^{2} 2x_i = 2x_1 + 2x_2 = 2(3) + 2(1) = \underline{\underline{8}}$$

$$\sum_{i=2}^{3} (x_i - 2) = (x_2 - 2) + (x_3 - 2) = (1-2) + (4-2) = -1 + 2 = \underline{\underline{1}}$$

$$\sum_{i=2}^{n} x_i^2 = x_2^2 + x_3^2 + x_4^2 + x_5^2 = 1^2 + 4^2 + 2^2 + 1^2 = 1 + 16 + 4 + 1 = \underline{\underline{22}}$$

مقاييس النزعة المركزية Measures of Central Tendency

مقدمة

نحتاج احيانا إلى معرفة أين تتركز المعلومات في ظاهرة معينة، لذلك نستخدم بعض المؤشرات التي يمكن الاعتماد عليها في وصف الظاهرة من حيث القيمة التي تتوسط القيم أو تنزع إليها القيم ، ومن حيث التعرف على مدى تجانس القيم التي يأخذها المتغير، وأيضا ما إذا كان هناك قيم شاذة أم لا .

مقاييس النزعة المركزية

تسمى مقاييس النزعة المركزية بمقاييس الموضع أو المتوسطات ، وهى القيم التى تتركز القيم حولها ، ومن هذه المقاييس ، الوسط الحسابي ، والمنوال ، والوسيط ، والوسط الهندسي ، والوسط التوافقي ، والرباعيات ، والمئينات ، وفيما يلي عرض لأهم هذه المقاييس. وبحكم أن هذا الجزء مراجعة لماسبق،سوف نقتصر فقط على الوسط الحسابي و الوسيط و المنوال للبيانات غير المبوبة.

الوسط الحسابي Arithmetic Mean

من أهم مقاييس النزعة المركزية ، وأكثرها استخداما في النواحي التطبيقية. يعرف الوسط الحسابي بشكل عام على أنه مجموع القيم مقسوما على عددها . فإذا كان لدينا n من القيم ، ويرمز لها بالرمز : $x_1, x_2, \dots, x_n$.

فإن الوسط الحسابي لهذه القيم ، ونرمز له بالرمز $\overline{x}$ يحسب بالمعادلة التالية :

$$\text{الوسط الحسابي} = \frac{\text{مجموع القيم}}{\text{عدد القيم}}$$

$$\overline{x} = \frac{x_1 + x_2 + \dots + x_n}{n} = \frac{\sum_{i=1}^{n} x_i}{n}$$

مثال (**3**–**1**)

فيما يلي الدرجات الفصلية لعدد 8 طلاب في مقرر **199** جغر .

34 32 42 37 35 40 36 40

والمطلوب إيجاد الوسط الحسابي لدرجة الطالب في الامتحان .

الحل

لإيجاد الوسط الحسابي للدرجات تطبق المعادلة رقم (**3**–**1**) كما يلي:

$$\bar{x} = \frac{x_1 + x_2 + \ldots + x_n}{n}$$
$$= \frac{34+32+42+37+35+40+36+40}{8} = \frac{296}{8} = 37$$

أي أن الوسط الحسابي لدرجة الطالب في اختبار مقرر **199** جغر يساوي 37 درجة

أهم خاصية للوسط الحسابي

مجموع انحرافات القيم عن وسطها الحسابي يساوى صفرا ، ويعبر عن هذه الخاصية بالمعادلة .

$$\boxed{\sum(x-\bar{x})=0} \qquad (٤-٣)$$

ويمكن التحقق من هذه الخاصية باستخدام بيانات المثال السابق ، نجد أن درجات الطلاب هي : ,34, 32, 42
37, 35, 40, 36, 40 ، والوسط الحسابي للدرجة هو $\bar{x}=37$ ، إذا :

x	34	32	42	37	35	40	36	40
$\sum(x-37)$	-3	-5	5	0	-2	3	-1	3

أي أن : $\sum(x-37)=0$

مزايا وعيوب الوسط الحسابي

يتميز الوسط الحسابي بالمزايا التالية :

- أنه سهل الحساب .
- يأخذ في الاعتبار كل القيم .
- أنه أكثر المقاييس استخداما وفهما .

ومن عيوبه .

- أنه يتأثر بالقيم الشاذة والمتطرفة .
- يصعب حسابه في حالة البيانات الوصفية .
- يصعب حسابه في حالة الجداول التكرارية المفتوحة .

الوسيط Median

هو أحد مقاييس النزعة المركزية، والذي يأخذ في الاعتبار رتب القيم ، ويعرف الوســـيط بأنه القيمة التي يقل عنها نصــف عدد القيم $(n/2)$ ، ويزيد عنها النصــف الآخر $(n/2)$ بعد رتيبها ترتيبا تصــاعديا، أي أن 50% من القيم أقل منه، 50% من القيم أعلى منه. وفيما يلي كيفية حساب الوسيط في حالة البيانات غير مبوبة:

الوسيط للبيانات غير المبوبة

لبيان كيف يمكن حساب الوسيط للبيانات غير المبوبة ، نتبع الخطوات التالية:

- ترتب القيم تصاعديا .
- تحديد رتبة الوسيط، وهي : رتبة الوسيط = $\left(\frac{n+1}{2}\right)$
- إذا كان عدد القيم *(n)* فردي فإن الوسيط هو:

$$\boxed{\text{الوسيط} = \text{القيمة رقم} \left(\frac{n+1}{2}\right)}$$

- إذا كان عدد القيم (n) زوجي، فإن الوســـيط يقع بين القيمة رقم $(n/2)$ ، والقيمة رقم $((n/2)+1)$ ، ومن ثم يحسب الوسيط بتطبيق المعادلة التالي:

$$\text{الوسيط} = \frac{\text{القيمة رقم}\left(\frac{n}{2}\right) + \text{القيمة رقم}\left(\frac{n}{2}+1\right)}{2} \qquad (٣-١٠)$$

مثال:

تم قياس كمية الامطار الشهر يناير في مدينتين a و b: الأول من عام 2001 حتى عام 2007 و الثانية من عام 2001 حتى عام 2010م كما هو موضح في الجدول اسفل.

وكانت على النحو التالي :

المدينة (a)	1.2	2.75	3.25	2	3	2.3	1.5			
المدينة (b)	4.5	1.8	3.5	3.75	2	2.5	1.5	4	2.5	3

والمطلوب حساب وسيط كمية الامطار في المدينيتين ثم قارن بينها.

الحل

أولا : حساب وسيط كمية الامطار للمدينة الاولى(a)

نرتب القيم تصاعديا فيكون:

كمية المطر

قيمة الوسيط

كمية المطر	1.2	1.5	2	2.3	2.75	3	3.25
الرتبة	1	2	2	4	5	6	7

رتبة الوسيط

الرتبة

- **عدد القيم فردى** $(n = 7)$
- **إذا رتبة الوسيط هي:** $\left((n+1)/2 = (7+1)/2 = 4\right)$.
- **ويكون الوسيط هو القيمة رقم 4 ، أي أن وسيط كمية الامطار للمدينة الاولى a هو:**

$$Med_a = 2.3 \text{ mm}$$

ثانيا : حساب وسيط كمية الامطار للمدينة الثانية(b) :

- **نرتب القيم تصاعديا فيكون:**

كمية المطر

قيمة الوسيط $= \frac{2.5 + 3}{2}$

2.75

1.5	1.8	2	2.5	2.5	3	3.5	3.75	4	4.5
1	2	3	4	5	6	7	8	9	10

5.5

رتبة الوسيط

الرتبة

- عدد القيم زوجي $(n = 10)$ إذا
- رتبة الوسيط هي : $((n+1)/2 = (10+1)/2 = 5.5)$.
- الوسيط = الوسط الحسابي للقيمتين الواقعتين في المنتصف (رقم 5 ، 6) .

$$Med_b = \frac{2.5+3}{2} = 2.75 \text{ mm}$$

وبمقارنة كمية المطر في المدينتين ، نجد أن وسيط كمية المطر في المدينة (a) أقل من وسيط كمية المطر في المدينة (b) ، أي أن : $Med_b > Med_a$.

● اهم مزايا وعيوب الوسيط

من مزايا الوسيط

1- لا يتأثر بالقيم الشاذة أو المتطرفة .

2- كما أنه سهل في الحساب .

ومن عيوب الوسيط

1- أنه لا يأخذ عند حسابه كل القيم في الاعتبار، فهو يعتمد على قيمة أو قيمتين فقط .

2- يصعب حسابه في حالة البيانات الوصفية المقاسة بمعيار اسمي nominal

المنوال Mode

يعرف المنوال بأنه القيمة الأكثر شيوعا أو تكرارا ، ويكثر استخدامه في حالة البيانات الوصفية ، لمعرفة النمط (المستوى) الشائع، ويمكن حسابة للبيانات المبوبة وغير المبوبة كما يلي:

●حساب المنوال في حالة البيانات غير المبوبة

المنوال (Mod) = القيمة (المستوى) الأكثر تكرارا

اختيرت عينات عشوائية من طلاب بعض أقسام الآداب ، وتم رصد درجات هؤلاء الطلاب في مقرر سلم **101** ، وكانت النتائج كالتالي:

قسم الجغرافيا	80	77	75	77	77	77	65	70	58	67
قسم اللغة العربية	88	68	60	75	93	65	77	85	95	90
قسم التاريخ	80	65	69	80	65	88	76	65	86	80
قسم الاعلام	85	73	69	85	73	69	69	73	72	85

والمطلوب حساب منوال الدرجات لكل قسم من الأقسام :

الحل

هذه البيانات غير مبوبة ، لذا فإن :

المنوال = القيمة الأكثر تكرارا

والجدول التالي يبين منوال الدرجة لكل قسم من الأقسام .

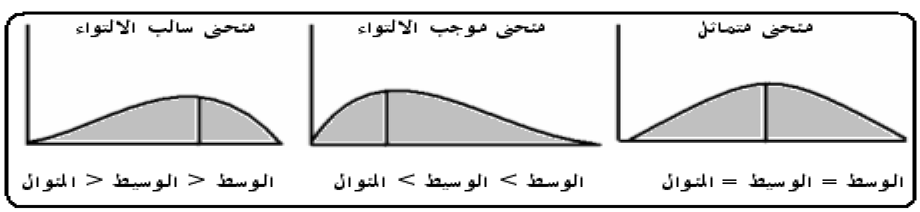
القسم	القيمة الأكثر تكرار	القيمة المنوالية
قسم الجغرافيا	الدرجة 77 تكررت 4 مرات	المنوال = 77 درجة
قسم اللغة العربية	جميع القيم ليس لها تكرار	لا يوجد منوال
قسم التاريخ	الدرجة 65 تكررت 3 مرات	يوجد منوالان هما :

المنوال الأول = 65 المنوال الثاني = 80	الدرجة 80 تكررت 3 مرات	
يوجد ثلاث منوال هي : المنوال الأول = 69 المنوال الثاني = 73 المنوال الثالث = 85	الدرجة 69 تكررت 3 مرات الدرجة 73 تكررت 3 مرات الدرجة 85 تكررت 3 مرات	قسم الاعلام

إستخدام مقاييس النزعة المركزية في تحديد شكل توزيع البيانات

يمكن استخدام الوسط الحسابي والوسيط والمنوال في وصف المنحنى التكراري، والذي يعبر عن شكل توزيع البيانات ، كما يلي :

شكل (**3-1**)



- يكون المنحنى متماثل إذا كان :

  الوسط = الوسيط = المنوال .

- يكون المنحنى موجب الالتواء (ملتوي جهة اليمين) إذا كان:

  الوسط > الوسيط > المنوال

- يكون المنحنى سالب الالتواء (ملتوي جهة اليسار) إذا كان :

  الوسط < الوسيط < المنوال

مثال عام (3-7)

قام مدير مراقبة الإنتاج بسحب عينة من 10 عبوات من المياه المعبأة للشرب ، ذات الحجم 5 لتر ، والمنتجة بواسطة إحدى شركات تعبئة المياه لفحص كمية الأملاح الذائبة، وكانت كالتالي :

115 123 119 123 124 119 123 121 123 121

<u>والمطلوب</u> : حساب الوسط الحسابي، والوسيط، والمنوال، ثم حدد شكل الالتواء لهذه البيانات .

الحل

حساب الوسط الحسابي :

$$\overline{x} = \frac{\sum x}{n} = \frac{1211}{10} = 121.1$$

- حساب الوسيط :

رتبة الوسيط : $(n+1)/2 = (10+1)/2 = 5.5$

ترتيب القيم تصاعديا

قيمة الوسيط

122

الطاقة	115	119	119	121	121	123	123	123	123	124
الرتبة	1	2	3	4	5	6	7	8	9	10

5.5

رتبة الوسيط

عدد القيم = 10 ، وهو عدد زوجي. الوسيط = الوسط الحسابي للقيمتين رقم (5 , 6)

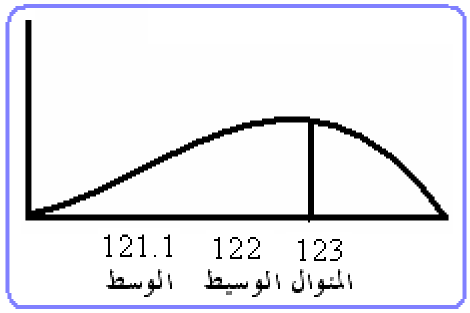
$$Med = \frac{121+123}{2} = \frac{244}{2} = 122$$

- حساب المنوال :

المنوال يساوى القيمة الأكثر تكرارا: القيمة 123 تكررت أكثر من غيرها ، إذا

$$Mod = 123$$

وبمقارنة الوسط والوسيط و المنوال نجد أن :



نجد أن : الوسط < الوسيط < المنوال ، إذا توزيع بيانات كمية الأملاح سالبة الالتواء.

● مقاييس التشتت Dispersion Measurements

مقدمة

الهدف من تلك المقييس هو معرفة تجانس مفردات البيانات و تقاربها أو تشتتها و تباعدها عن بعضها البعض و مدى مصداقية مقاييس النزعة المركزية في تمثيل البيانات. تخدم أيضا تلك المقاييس في المقارنة بين المجموعات.

فمثلا لدينا مدينتين X وY متوسط كمية الامطار السنوية فيهما متساوية و هي **216** ملم و كمية الامطار لكل شهر فيهما على النحو التالي:

X	8, 11, 10, 5, 7, 17, 35, 40, 45, 15, 14, 9
Y	17, 20, 21, 19, 20, 21, 25, 26, 13, 15, 19

بمتوسط كمية امطار شهرية متساوية و هي **18** ملم لكل منهما

نلاحظ أن كمية الامطار الشهرية في المجموعة Y متقاربة اكثر بكثير من المدينة X بالرغم ان متوسط كمية الامطار السنوية متساوية و متوسط الامطار الشهرية الكل منهما متساوي وهو **18**ملم. هنا تأتي مقاييس التشتت لتكشف الفرق في نمط كمية الامطار الشهرية بين هاتين المدينتين.

و من هذه أهم مقاييس التشتت: المدى، والتباين، والانحراف المعياري .

●المدى Rang

هو أبسط مقاييس التشتت ، ويحسب المدى في حالة البيانات غير المبوبة بتطبيق المعادلة التالية .

المدى في حالة البيانات غير المبوبة = أكبر قراءة – أقل قراءة

$$Rang = Max - Min$$

وأما المدى في حالة البيانات المبوبة له أكثر من صيغة، ومنها المعادلة التالية:

المدى في حالة البيانات المبوبة = مركز الفئة الأخيرة – مركز الفئة الأولى

مثـال :

وفيما يلي بيانات كمية الامطار لشهر مارس (ملم) لأحد المدن

48 62.1 54 51.8 52.9 51.8 50.8 46.3 50.3

والمطلوب حساب المدى .

الحل

المدى = أكبر قراءة – أقل قراءة

أكبر قراءة = 62.1 **أقل قراءة =** 46.3

إذا المدى هو :

Rang=Max-Min=62.1-46.3 =15.8

المدى يساوي 15.8ملم.

مثـال :

الجدول التكراري التالي يبين قيم درجة الحرارة المئوية الكبرى و الصغرى من شهر يناير إلى شهر يونيو في احد المدن .

درجة الحرارة بالمئوي	15-20	20-25	25-30	30-35	35-40	40-45

والمطلوب حساب المدى لدرجة الحرارة بين احر الشهور و ابردها

الحل:

المدى = مركز الفئة الأخيرة – مركز الفئة الأولى

مركز الفئة الأخيرة: (40+45)/2=85/2=42.5 **مركز الفئة الأولى:** (15+20)/2=35/2=17.5

إذا $Rang = 42.5 - 17.5 = 25$

أي أن المدى يساوي 25 درجة مئوية

مزايا وعيوب المدى

من مزايا المدى

1- أنه بسيط وسهل الحساب
2- يكثر استخدامه عند الإعلان عن حالات الطقس، و المناخ الجوي، مثل درجات الحرارة، والرطوبة، والضغط الجوي.
3- يستخدم في مراقبة الجودة .

2- ومن عيوبه

- أنه يعتمد على قيمتين فقط ، ولا يأخذ جميع القيم في الحسبان .
- يتأثر بالقيم الشاذة .

● الانحراف المتوسط Mean Deviation (MD)

هو أحد مقاييس التشتت، ويعبر عنه بمتوسط الانحرافات المطلقة للقيم عن وسطها الحسابي ، فإذا كانت $x_1, x_2, \ldots, x_n$ هي القراءات التي تم أخذها عن ظاهرة معينة ، وكان ($\bar{x} = \sum x/n$) عبارة عن الوسط الحسابي لهذه القراءات، فإن الانحراف المتوسط (MD) يحسب بتطبيق المعادلة التالية:

$$\boxed{MD = \frac{\sum |x - \bar{x}|}{n}} \qquad (٤-٤)$$

وهذه الصيغة تستخدم في حالة البيانات غير المبوبة .

مثال:

إذا كانت متوسط سرعة الرياح لخمس محطات مناخية (سم/الثانية):

4 5 2 10 7

أوجد قيمة الانحراف المتوسط سرعة الرياح

الحل

لحساب قيمة الانحراف المتوسط يتم اولا حساب

- الوسط الحسابي :

$$\bar{x} = \frac{\sum x}{n} = \frac{28}{5} = 5.6$$

ويتم تكوين الجدول التالي :

الطاقة التصديرية x	الانحرافات $(x-5.6)$ $(x-\bar{x})=$	الانحرافات المطلقة $\lvert x-5.6 \rvert$
4	4 - 5.6 = -1.6	1.6
5	5 - 5.6 = -0.6	0.6
2	2 - 5.6 = -3.6	3.6
10	10 - 5.6 = 4.4	4.4
7	7 - 5.6 = 1.4	1.4
Sum	0	11.6

- إذا الانحراف المتوسط قيمته هي :

$$MD = \frac{\sum |x-\bar{x}|}{n} = \frac{11.6}{5} = 2.32$$ (سم/الثانية)

مزايا وعيوب الانحراف المتوسط

من مزايا الانحراف المتوسط أنه يأخذ كل القيم في الاعتبار، ولكن يعاب عليه ما يلي:

- يتأثر بالقيم الشاذة .
- يصعب التعامل معه رياضيا.

●التباين Variance

هو أحد مقاييس التشتت ، وأكثرها استخداما في النواحي التطبيقية ، ويعبر عن متوسط مربعات انحرافات القيم عن وسطها الحسابي.

أولا: التباين في المجتمع (σ^2)

إذا توافر لدينا قراءات عن كل مفردات المجتمع ، ولتكن: $x_1, x_2, \dots, x_N$ ، فإن التباين في المجتمع ،

ويرمز له بالرمز σ^2 (سيجما) يحسب باستخدام المعادلة التالية :

حيث أن μ هو الوسط الحسابي في المجتمع ، أى أن : $\mu = \sum x/N$.

مثال:

مصنع لتعبئة المواد الغذائية ، يعمل به 15 عامل ، وكانت عدد سنوات الخبرة لهؤلاء العمال كما يلي :

5 13 7 14 12 9 6 8 10 13 14 6 11 12 10

بفرض أن هذه البيانات تم جمعها عن كل مفردات المجتمع ، فأوجد التباين لعدد سنوات الخبرة .

الحل

لحساب تباين سنوات الخبرة في المجتمع ، يتم استخدام المعادلة:

$$\sigma^2 = \frac{\sum (X - \mu)^2}{N}$$

- الوسط الحسابي في المجتمع μ

$$\mu = \frac{1}{N}\sum x$$
$$= \frac{1}{15}(5+13++7+\ldots+12+10) = \frac{1}{15}(150) = 10$$

- حساب مربعات الانحرافات $\sum (x-\mu)^2$

سنوات الخبرة x	$(x-\mu)$	$(x-\mu)^2$
5	5-10 = -5	25

13	3	9
7	-3	9
14	4	16
12	2	4
9	-1	1
6	-4	16
8	-2	4
10	0	0
13	3	9
14	4	16
6	-4	16
11	1	1
12	2	4
10	0	0
150	0	130

بما أن: $\sum(x-\mu)^2 = 130$

إذا التباين : $\sigma^2 = \frac{\sum(X-\mu)^2}{N}$

130/ 15 = 8.7

ثانيا: التباين في العينة (s^2)

في كثير من الحالات يكون تباين المجتمع σ^2 غير معلوم، وعندئذ يتم سحب عينة من هذا المجتمع ، ويحسب التباين من بيانات العينة كتقدير لتباين المجتمع ، فإذا كانت قراءات عينة عشوائية حجمها n هي $x_1, x_2, \ldots, x_n$ ، فإن تباين العينة ويرمز له بالرمز s^2 هو:

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

حيث أن $\bar{x}$ هو الوسط الحسابي لقراءات العينة ، أي أن : $\bar{x} = \sum x/n$ ، وتباين العينة هو التقدير غير المتحيز لتباين المجتمع .

مثال:

في المثال السابق ، إذا تم سحب عينة من عمال المصنع حجمها 5 عمال ، وسجل عدد سنوات الخبرة ، وكانت كالتالي .

8 13 10 5 9

احسب تباين سنوات الخبرة في العينة .

الحـل

لحساب التباين في العينة يتم تطبيق المعادلة:

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

ولكن قبل ذلك نحسب الوسط الحسابي في العينة :

$$\bar{x} = \frac{1}{n}\sum x = \frac{1}{5}(8+13+10+5+9) = \frac{1}{5}(45) = 9$$

- **حساب مربعات الانحرافات** $\sum(x-\bar{x})^2$

سنوات الخبرة x	8	13	10	5	9	45
$(x-\bar{x})$	-1	4	1	-4	0	0
$(x-\bar{x})^2$	1	16	1	16	0	34

أي أن : $\sum(x-\bar{x})^2 = 34$ ،

- **إذا تباين سنوات الخبرة في العينة قيمته هي :**

$$s^2 = \frac{\sum(x-\bar{x})^2}{n-1} = \frac{34}{(5-1)} = \frac{34}{4} = 8.5$$

- **في هذه الحالة يمكن القول بأن تباين العينة 8.5، وهو في نفس الوقت تقدير غير متحيز لتباين المجتمع .**

●الانحراف المعياري Standard Deviation

عند اســـتخدام التباين كمقياس من مقاييس التشـــتت، نجد أنه يعتمد علي مجموع مربعات الانحرافات، ومن ثم لا يتمشى هذا المقياس مع وحدات قياس المتغير محل الدراسة ، ففي المثال السابق ، نجد أن تباين سنوات الخبرة في العينة 8.5، فليس من المنطق عند تفسـير هذه النتيجة أن نقول ، " تباين سـنوات الخبرة هو 8.5 سـنة تربيع "، لأن وحدات قياس المتغير هو عدد الســـنوات، من أجل ذلك لجأ الإحصـــائيين إلى مقياس منطقي يأخذ في الاعتبار الجذر التربيعي للتباين ، لكي يناسب وحدات قياس المتغير، وهذا المقياس هو الانحراف المعياري.

إذا الانحراف المعياري ، هو الجذر التربيعي الموجب للتباين ، أي أن:

$$\boxed{\text{الانحراف المعيارى} = \sqrt{\text{التباين}}}$$

ومثال على ذلك :

- في مثال سابق وجدناأن الانحراف المعياري لسنوات الخبرة لعمال المصنع (المجتمع) ، ويرمز له بالرمز (σ) هو :

$$\sigma = \sqrt{\frac{1}{N}\sum x^2 - \mu^2}$$
$$= \sqrt{\frac{1}{15}1630 - 10^2} = \sqrt{8.67} = 2.94$$

في هذه الحالة ، يكون الانحراف المعياري لسنوات الخبرة في المجتمع هو 2.94 سنة .

- نجد أن الانحراف المعياري لسنوات الخبرة لعمال العينة ، ويرمز له بالرمز S ، هو :

$$s = \sqrt{\frac{1}{n-1}\left(\sum x^2 - \frac{(\sum x)^2}{n}\right)}$$
$$= \sqrt{\frac{1}{5-1}\left(439 - \frac{(45)^2}{5}\right)} = \sqrt{\frac{1}{4}(439-405)} = \sqrt{\frac{1}{4}(34)} = 2.92$$

أي أن الانحراف المعياري لسنوات الخبرة في العينة هو 2.92 سنة .

مزايا وعيوب الانحراف المعياري

من مزايا الانحراف المعياري

1–أنه أكثر مقاييس التشتت استخداما .

2– يسهل التعامل معه رياضيا .

3– يأخذ كل القيم في الاعتبار .

ومن عيوبه ، أنه يتأثر بالقيم الشاذة .

التحليل المكاني Locational Analysis

المتوسط المكاني location mean center

لنفرض ان N......1,2,3 (x,y) مجموعة نقاط تمثل مواقع في خريطة يمكن تحديد عن المتوسط المكاني على النحو التالي:

$\bar{X} = \sum x/N$ and

$\bar{y} = \sum y/N$

يعطينا النقطة $(\bar{X}, \bar{y})$ هي تمثل المتوسط المكاني µ

المتوسط المكاني الموزون Weighted Mean Lcation

$\bar{X}w = \sum xw/\sum w$ and

$\bar{y}w = \sum yw/\sum w$

يعطينا النقطة $(\bar{X}w, \bar{y}w)$ و هي تمثل المتوسط المكاني الموزون

مثال: لنفرض ان لدينا مواقع على الخريطة و هي: a.b.c.d,e,f,,g , h, و هي موضحة في الشكل اسفل و الجدول الذي يحوي المعلومات و هو يمثل الموقع وعدد السكان بالالف. اردت شركة غذائية فتح محل اغذية و ارادت تحديد المتوسط المكاني للمواقع و المتوسط المكاني الموزون لخدمة اكبر عدد من السكان:

الحل: اولا ننشيئ الجدول ثم نطبق المعادلات السابقة لإيجاد المتوسط و المتوسط المكاني الموزون

الموقع	X	Y	عدد السكان (1000)	xw	yw
A	1.5	.6	10	15	6
B	1	4	25	25	100
C	2	4	12	24	48
D	3	3	9	27	27
E	3.3	2	30	99	60
F	4	5	70	280	650
G	4.2	3	10	42	30
H	5	4	90	450	360
∑	**24**	**25.6**	**256**	**962**	**981**
µ	**3**	**3.2**		**3.7**	**3.83**

يحسب المتوسط المكاني:

$\mu = \bar{X} = \sum x/N = 24/8 = \underline{3}$

$\bar{y} = \sum y/N = 25.6/8 = 3.2$

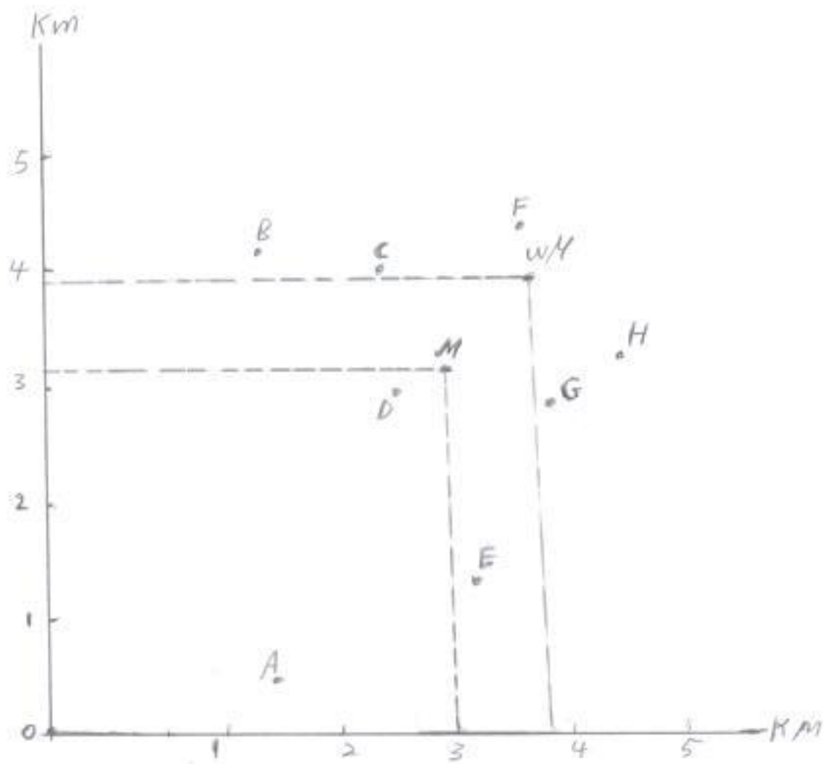
(3,3.2)

و يكون المتوسط المكاني الموزون:

$\bar{X}w = 962/256 = 3.75$ and

$\bar{y}w = 981/256 = 3.83$ (3.7,3.83)

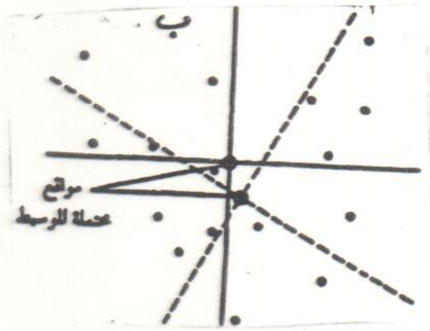
إذا المتوسط المكاني الموزون بعدد السكان يكون في الاحداثية كما هو موضح في الرسم wµ=x=3.75, y=3.83

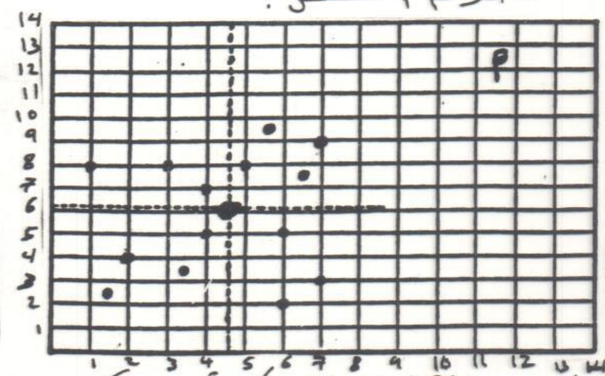

Km
5
4
3
2
1
0
1
2
3
4
5
KM
B
C
F
w/Y
M
H
G
D
E
A

locational

* الوسيط المكاني (الموقع الوسيط) - Locational Median

هنا لا نعتمد على تكرار النقاط وتكون ذلك الموقع الذي تقع نصف المواقع شماله والنصف الآخر جنوبه أو تقع نصف المواقع شرقه ونصفها غربه ويمكن تحديده بدون عمليات حسابية بل بواسطة رسم الرسم البياني كما في الرسم أ شكل:





يتضح من الشكل ب أنه يمكن تحديد أكثر من موقع وسيط. ومن أهم خواص الموقع المتوسط هو:

$$\Sigma |X - Med| = Min$$

أي أن مجموع الفروقات المطلقة للمشاهدات عن Med أقل منها عن أية نقطة أخرى. لذلك يستخدم لتحديد الموقع الذي تكون فيه المسافة مع المواقع الأخرى هي أقل مسافة (مثال الموقع على الطريق العام)

مثال: (مصنع الألبان وتحديد أفضل موقع له لخدمة البقالات)

الموقع	المسافة بين الموقع والوسط الحسابي	المسافة بين الموقع والوسيط
أ	٦	٤
ب	٥	٣
ج	٤	٢
د	(٢)	٠
ه	٠	٢
و	٨	١٠
ع	٩	١١
المجموع	٣٤	٣٢

* المنوال المكاني Locational Mode

يتم تحديده بواسطة الرسم البياني فالمنوال هو ذلك الموقع الذي يمثل أكبر عدد من المواقع وهو مفيد في حالات عديدة مثل شخص

يريد تأسيس مركز تجاري لخدمة اكثر الجهات تركزا للموقع

Measure of locational Dispersion * قياس التشتت المكاني

- مقدمة ......

- المسافة المعيارية Standard Distance (SD)

هي القيمة الثانية المساوية للانحراف المعياري ويتم حسابها بواسطة المعادلة التالية.

$$SD = \sqrt{\frac{\Sigma(X-\bar{X})^2}{N} + \frac{\Sigma(Y-\bar{Y})^2}{N}}$$

مع ملاحظة ان $(\bar{X}, \bar{Y})$ تمثل الاحداثية.

أو هي

$$SD = \sqrt{\sigma_x^2 + \sigma_y^2}$$

يلاحظ استخدم N بدل $n-1$ في المقام. والمسافة المعيارية تمثل نصف قطر الدائرة المحيطة بالوسط الحسابي للموقع. (الموقع المتوسط) وكلما ازدادت قيمة SD كلما أشار ذلك الى تفرقه وتبعثر الاماكن والعكس صحيح

تركز المواقع

تبعثر المواقع

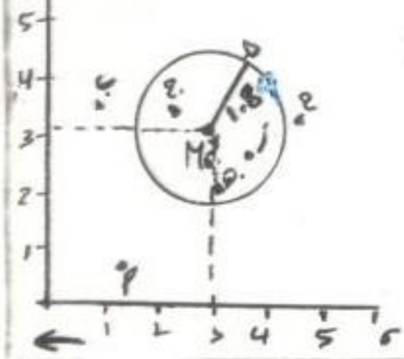
مثال: اذا اردنا ان نحسب المسافة المعيارية للمواقع في المثال السابق فاننا نقوم بتطبيق المعادلة السابقة الذكر والتي تحتاج لانشاء الجدول التالي:

$(Y-\bar{Y})^2$	$(X-\bar{X})^2$	$Y-\bar{Y}$	$X-\bar{X}$	Y	X	الموقع
6.76	2.25	-2.6	-1.5	0.6	1.5	أ
.64	4	.8	-2	4	1	ب
.64	1	.8	-1	4	2	ج
0.04	0	-.2	0	3	3	د
1.44	0.09	-1.2	.3	2	3.3	هـ
3.24	1	1.8	1	5	4	و
0.04	1.44	-.2	1.2	3	4.2	ز
.64	4	.8	2	4	5	ح
Σ 13.44	13.78	0	0	25.6	24	
M				3.2	3	

وتطبيقا للقانون أعلاه

$$SD = \sqrt{\frac{13.78}{8} + \frac{13.44}{8}}$$

$$= \sqrt{1.72 + 1.68} = \boxed{1.84}$$

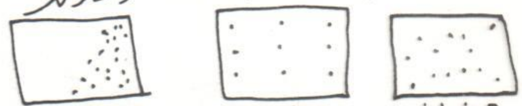


4 locational

Neurst Neighbor Index

* قرينة الجار الأقرب

يعتبر هذا من مقاييس التشتت التي تقيس التشتت المكاني بالإضافة إلى كونها تقيس انتظام التوزيع المكاني من عدمه. فهو تقيس فيما إذا كان التوزيع المكاني عشوائي Random Distribution أو منتظم Uniform أو متركز clustered



متركز منتظم متماثل عشوائي

و يمكن حساب قرينة الجار الأقرب على النحو التالي

$$R = \frac{\Sigma d / N}{\bar{d}e} \quad \text{أو} \quad \frac{\bar{d}}{\bar{d}e}$$

حيث أن $\bar{d}$ تمثل متوسط المسافة التي تفصل النقاط عن بعضها و $\bar{d}e$ تمثل متوسط المسافة المتوقعة التي تفصل بين النقاط و جاراتها الأقرب في التوزيع العشوائي. وعند وضع المواقع في خريطة شبكية إحداثيات X, Y فإن المسافة بين النقاط يمكن حسابها بواسطة قانون فيثاغورس المعروف على النحو التالي Pythagorus Theroum

$$d_{ij} \sqrt{(X-X_m)^2+(Y-Y_m)^2}$$

- أما القيمة de فيمكن حسابها على النحو التالي

$$\bar{d}e = \frac{1}{2\sqrt{N/A}}$$

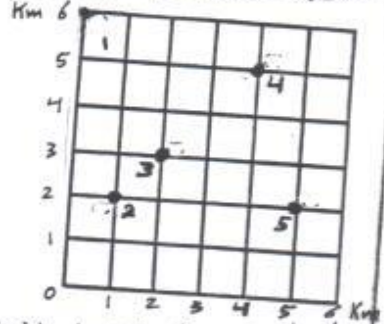
و N/A تمثل الكثافة الحسابية للنقاط في المنطقة حيث N عدد النقاط و A المساحة الكلية للمنطقة ويرمز أحيانا لـ N/A بالرمز λ.

- والقيم التي يحصل عليها لـ R دائما ما بين صفر و 2.149 . وعندما يكون قيمة R تساوي 1 أي أن $\frac{\bar{d}}{\bar{d}e} = 1$ فإن ذلك يشير إلى أن توزيع النقاط توزيع عشوائي تام وعندما تكون $R < 1$ فإن ذلك يشير إلى تركز فإذا أصبحت $R=0$ معنى ذلك أن جميع النقاط تحتل نفس الموقع. أما إذا كانت قيمة $R = 2.149$ فإن ذلك يشير إلى أن التوزيع توزيع منتظم تماما. لذلك التوزيع ذو الصفة العشوائية دائما يكون $.1 < R < 2.149$

5
Locational

مثال لحساب قرينة الجار الأقرب.

الشكل: شكل يمثل مواقع خمس مدن موزعة في مساحة مقدارها $36 Km^2$ $(6 \times 6Km)$.



المطلوب

1- حساب المسافة بين كل مدينتين.
2- حساب الجار الأقرب لكل مدينة.
3- حساب قرينة الجار الأقرب R.

الحل: أولاً: نوجد المسافة بين النقاط (المدن) باستخدام قانون فيثاغورس
مثلاً الموقع 1 مسافته من الموقع 2 تحسب بالنحو التالي
بعد احتساب إحداثية الموقع 1 هي (X, Y) والموقع 2 هو (X_m, Y_m)

$$d_{12} = \sqrt{(X-X_m)^2+(Y-Y_m)^2} = \sqrt{(0-1)^2+(6-2)^2} = \sqrt{17} = 4.12$$

وبنفس الطريقة نحسب المسافات بين جميع النقاط.
ثانياً: نحسب الجار الأقرب لكل نقطة (مدينة) كما هو موضح في الجدول: شكل على النحو التالي

$$\bar{d} = \Sigma d/N = 12.49/5 = 2.49$$

$$\bar{d}_e = \frac{1}{2\sqrt{N/A}} = \frac{1}{2\sqrt{5/36}} = \frac{1}{2\sqrt{.138}} = \frac{1}{2 * .372} = \frac{1}{.745} = 1.34$$

$$R = \frac{\bar{d}}{\bar{d}_e} = \frac{2.49}{1.34} = 1.85$$

المسافة للجار الأقرب	5	4	3	2	1	المدينة
3.61	5.39	4.12	3.61	4.12	0	1
1.41	4	4.24	1.41	0	4.12	2
1.41	3.17	2.83	0	1.41	3.61	3
2.83	3.17	0	2.83	4.24	4.12	4
3.17	0	3.17	3.17	4	5.39	5
$\Sigma d = 12.43$						
$\bar{d} = 2.49$						

- حيث إن القيمة $1 < R \leq 2.149$ فإن التوزيع توزيع متساوٍ
غير منتظم وغير متركز لهذه المدن.

* مربع كاي X^2 chi square للأهمية الإحصائية للنتائج ولكنه
عادة يستخدم مربع كاي لاختبار
ويمكن استخدامه أيضاً لقياس مدى التركز لها في خاصة عندما
تكون النقاط (المواقع) المدروسة كثيرة ونحسب مربع كاي
في هذه الحالة على النحو التالي

$$X^2 = \Sigma \frac{(f-F)^2}{F}$$

6 Locational

Loranz Index

* قرينة لورنز (موشر لورنز)

يستخدم لأغراض جغرافية عديدة تهتم بالتركز مثل تركز الدخل بين السكان وتركز النشاط البشري ... الخ، ومن ضمن استخدامات ~~التركز~~ قياس التركز السكاني لظاهرة ما مثل تركز نشاط معين في مكان معين أو تركز السكان في إقليم معين من الدولة ... الخ. ويمكن صياغة على النحو التالي

$$I = (A-R) / (M-R)$$

حيث تمثل I = قرينة لورنز

A = مجموع التكرارات التراكمية للتوزيع الفعلي للظاهرة

R = " " " " المنتظم الافتراضي " "

M = " " " " " المركز للظاهرة

وتتراوح قيمة I ما بين صفر إلى 1 حيث يشير 1 إلى تركز تام وصفر إلى عدم تركز تام

مثال: إذا كان لدينا دولة فيها سبع مناطق إستيطانية وحصة كل إقليم منها من السكان على النحو التالي:

الإقليم	أ	ب	ج	د	هـ	و	ز
عدد السكان بالألف	72.6	58.9	33.9	10.0	8.2	2.9	2.5

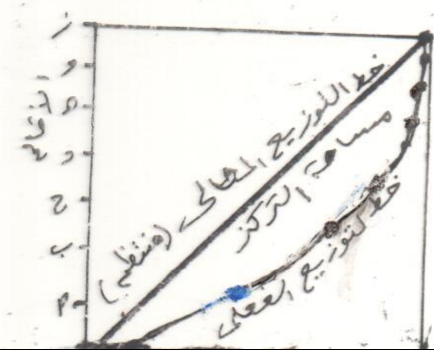
المطلوب: إحسب قرينة لورنز للتركز السكاني في أقاليم تلك الدولة.

الحل: ننشئ الجدول أسفل ثم نطبق قانون قرينة لورنز وتطبيقاً للقانون

الإقليم	التكرار للسكان تنازلياً	التكرار النسبي التراكمي A	توزيع منتظم نسبي	توزيع منتظم تراكمي R	توزيع مركز	توزيع مركز تراكمي M
أ	72.6	39	14.3	14.3	100	100
ب	58.9	70	14.3	28.6	0	100
ج	33.9	88	14.3	43	0	100
د	10.0	93	14.3	57.3	0	100
هـ	8.2	97	14.3	71.6	0	100
و	2.9	99	14.3	85.9	0	100
ز	2.5	100	14.3	100	0	100
مج	189	586	100	400.7	100	700

$$I = (586 - 400.7) / (700 - 400.7)$$
$$= 0.63$$

وهذا يشير إلى وجود تركز إلى حد ما حيث يوجد 88% من السكان في 3 أقاليم فقط، ويوضح منحنى لورنز ذلك



الارتباط والانحدار الخطي البسيط

1/6 مقـدمة

في الفصول الثلاث السابقة تم عرض بعض المقاييس الوصفية، مثل مقاييس النزعة المركزية، والتشتت، وغيرها من المقاييس الأخرى والتي يمكن من خلالها وصف شكل توزيع البيانات التي تم جمعها عن متغير واحد، وننتقل من التعامل مع متغير واحد إلى التعامل مع متغيرين أو أكثر، ويتناول هذا الفصل دراسة وتحليل العلاقة بين متغيرين، وذلك باستخدام بعض طرق التحليل الإحصائي مثل تحليل الارتباط، والانحدار الخطي البسيط، فإذا كان اهتمام الباحث هو دراسة العلاقة بين متغيرين استخدم لذلك أسلوب تحليل الارتباط، وإذا كان اهتمامه بدراسة أثر أحد المتغيرين على الآخر استخدم لذلك أسلوب تحليل الانحدار، ومن الأمثلة على ذلك:

1- الإنفاق، والدخل العائلي.

2- سعر السلعة، والكمية المطلوبة منها.

3- تقديرات الطلاب في مقرر الإحصاء، وتقديراتهم في مقرر الرياضيات.

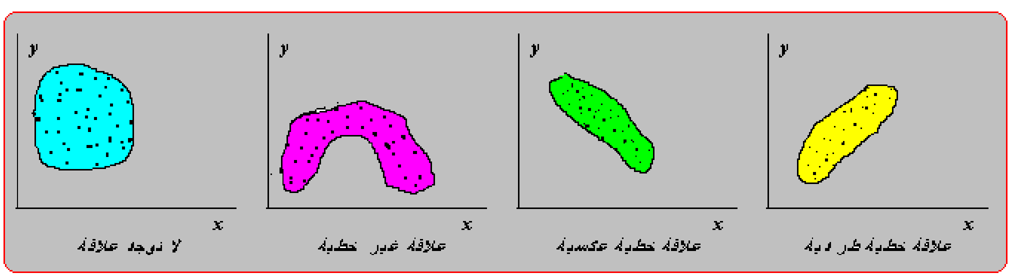
4- كميات السماد المستخدمة، وكمية الإنتاج من محصول معين تم تسميده بهذا النوع من السماد.

5- وزن الجسم، وضغط الدم.

والأمثلة على ذلك في المجال التطبيقي كثيرة، فإذا كان لدينا المتغيرين $(x \ , \ y)$ ، وتم جمع بيانات عن أزواج قيم هذين المتغيرين، وتم تمثيلها بيانيا فيما يسمى بشكل الانتشار، فإن العلاقة بينها تأخذ أشكالا مختلفة على النحو التالي :

شكل(6-1)

شكل الانتشار لبيان نوع العلاقة بين x , y



2/6 الارتباط الخطى البسيط Simple Correlation

إذا كان الغرض من التحليل هو تحديد نوع وقوة العلاقة بين متغيرين ، يستخدم تحليل الارتباط ، وأما إذا كان الغرض هو دراسة وتحليل أثر أحد المتغيرين على الآخر ، يستخدم تحليل الانحدار، وفي هذا الفصل يتم عرض أسلوب تحليل الارتباط الخطي البسيط، أي في حالة افتراض أن العلاقة بين المتغيرين تأخذ الشكل الخطي ، وسوف يجرى حسابه في حالة البيانات الكمية ، والبيانات الوصفية المقاسة بمعيار ترتيبي.

الغرض من تحليل الارتباط الخطى البسيط

الغرض من تحليل الارتباط الخطي البسيط هو تحديد نوع وقوة العلاقة بين متغيرين، ويرمز له في حالة المجتمع بالرمز ρ (رو)، وفي حالة العينة بالرمز r ، وحيث أننا في كثير من النواحي التطبيقية نتعامل مع بيانات عينة مسحوبة من المجتمع، سوف نهتم بحساب معامل الارتباط في العينة r كتقدير لمعامل الارتباط في المجتمع، ومن التحديد السابق للغرض من معامل الارتباط، نجد أنه يركز على نقطتين هما:

- **نوع العلاقة:.** وتأخذ ثلاث أنواع حسب إشارة معامل الارتباط كما يلي:

1– إذا كانت إشارة معامل الارتباط سالبة $(r < 0)$ توجد علاقة عكسية بين المتغيرين، بمعنى أن زيادة أحد المتغيرين يصاحبه انخفاض في المتغير الثاني، والعكس.

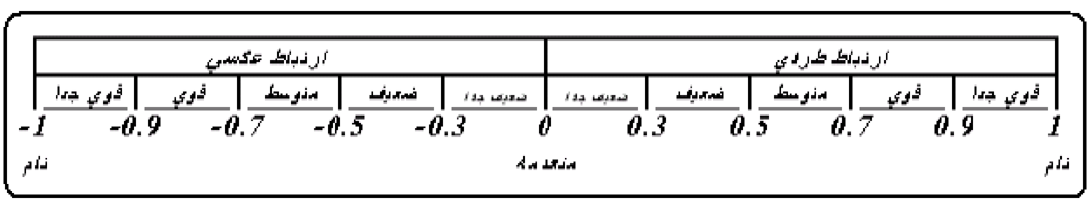
2– إذا كانت إشارة معامل الارتباط موجبة $(r > 0)$توجد علاقة طردية بين المتغيرين، بمعنى أن زيادة أحد المتغيرين يصاحبه زيادة في المتغير الثاني، والعكس .

3– إذا كان معامل الارتباط قيمته صفرا $(r = 0)$ دل ذلك على انعدام العلاقة بين المتغيرين.

- **قوة العلاقة:.** ويمكن الحكم على قوة العلاقة من حيث درجة قربها أو بعدها عن (± 1) ، حيث أن قيمة معامل الارتباط تقع في المدى$(-1 < r < 1)$، وقد صنف بعض الإحصائيين درجات لقوة العلاقة يمكن تمثيلها على الشكل التالي:

شكل (**6–2**)

درجات قوة معامل الارتباط



2/2/6 معامل الارتباط الخطى البسيط " لبيرسون" Pearson

في حالة جمع بيانات عن متغيرين كميين $(y \ , \ x)$ ، يمكن قياس الارتباط بينهما، باستخدام طريقة "بيرسون" Pearson، ومن الأمثلة على ذلك: قياس العلاقة بين الوزن والطول، والعلاقة بين الإنتاج والتكلفة، والعلاقة بين الإنفاق الاستهلاكي والدخل، والعلاقة بن الدرجة التي حصل عليها الطالب وعدد ساعات الاستذكار، وهكذا الأمثلة على ذلك كثيرة.

ولحساب معامل الارتباط في العينة ، تستخدم صيغة " بيرسون" التالية :

$$r = \frac{S_{xy}}{S_x\ S_y} = \frac{\dfrac{\sum(x-\bar{x})(y-\bar{y})}{(n-1)}}{\sqrt{\dfrac{\sum(x-\bar{x})^2}{(n-1)}}\sqrt{\dfrac{\sum(y-\bar{y})^2}{(n-1)}}} \qquad (٦-١)$$

حيث أن :

$S_{xy} = \sum(x-\bar{x})(y-\bar{y})/(n-1)$: هو التغاير بين $(x\ ,\ y)$ ،

$S_x = \sqrt{\sum(x-\bar{x})^2/(n-1)}$: هو الانحراف المعياري لقيم (x) ،

$S_y = \sqrt{\sum(y-\bar{y})^2/(n-1)}$: هو الانحراف المعياري لقيم (y) .

ويمكن اختصار الصيغة السابقة على النحو التالي:

$$r = \frac{\sum(x-\bar{x})(y-\bar{y})}{\sqrt{\sum(x-\bar{x})^2}\sqrt{\sum(y-\bar{y})^2}} \qquad (٦-٢)$$

مثال

فيما يلي كمية الامطار السنوية في دولة ما وإجمالي إنتاج القمح بالطن، خلال الفترة من 1995حتى عام 2002.

السنة	1995	1996	1997	1998	1999	2000	2001	2002
كمية الامطار سم	305 سم	313	297	289	233	214	240	217
إنتاج القمح بالطن	592 طن	603	662	607	635	699	719	747

والمطلوب: حساب معامل الارتباط بين االمطر والكمية، وما هو مدلوله ؟

الحل

بفرض أن (x) هي كمية المطر، (y) هي كمية الانتاج، ولحساب معامل الارتباط بين $(x\ ,\ y)$ يتم تطبيق المعادلة ، وذلك على النحو التالي:

- حساب الوسط الحسابي لكل من المساحة، والكمية $(\bar{x}\ ,\ \bar{y})$.

$$\bar{x}=\frac{\sum x}{n}=\frac{2108}{8}=263.5\ ,\ \ \bar{y}=\frac{\sum y}{n}=\frac{5264}{8}=658$$

- حساب المجاميع

x	y	$x-\bar{x}$	$(x-\bar{x})^2$	$y-\bar{y}$	$(y-\bar{y})^2$	$(x-\bar{x})(y-\bar{y})$
305	592	41.5	1722.25	-66	4356	-2739
313	603	49.5	2450.25	-55	3025	-2722.5
297	662	33.5	1122.25	4	16	134
289	607	25.5	650.25	-51	2601	-1300.5
233	635	-30.5	930.25	-23	529	701.5
214	699	-49.5	2450.25	41	1681	-2029.5
240	719	-23.5	552.25	61	3721	-1433.5
217	747	-46.5	2162.25	89	7921	-4138.5
2108	5264	0	12040	0	23850	-13528

$$\sum(x-\bar{x})^2=12040\ ,\ \sum(y-\bar{y})^2=23850\ ,$$
$$\sum(x-\bar{x})(y-\bar{y})=-13528$$

إذا معامل الارتباط قيمته هي:

$$r=\frac{\sum(x-\bar{x})(y-\bar{y})}{\sqrt{\sum(x-\bar{x})^2}\sqrt{\sum(y-\bar{y})^2}}=\frac{-13528}{\sqrt{12040}\sqrt{23850}}$$
$$=\frac{-13528}{(109.727)(154.434)}=\frac{-13528}{16945.619}=-0.798$$

- يوجد ارتباط عكسي قوي بين المساحة المنزرعة، وكمية إنتاج اللحوم.

تبسيط العمليات الحسابية:

في بعض الأحيان، يكون استخدام صيغة هذه المعادلة في غاية الصعوبة، خاصة إذا لازم العمليات الحسابية قيما كسرية، من أجل ذلك يمكن تبسيط هذه الصيغة إلى صيغة أسهل تعتمد على مجموع القيم وليس على انحرافات القيم عن وسطها الحسابي، وهذه الصيغة هي:

$$r = \frac{\sum xy - \frac{\sum x \sum y}{n}}{\sqrt{\left(\sum x^2 - \frac{(\sum x)^2}{n}\right)\left(\sum y^2 - \frac{(\sum y)^2}{n}\right)}} \qquad (٣-٦)$$

وبالتطبيق على بيانات المثال السابق ، يتبع الآتي :

- حساب المجاميع:

x	y	xy	x^2	y^2
305	592	180560	93025	350464
313	603	188739	97969	363609
297	662	196614	88209	438244
289	607	175423	83521	368449
233	635	147955	54289	403225
214	699	149586	45796	488601
240	719	172560	57600	516961
217	747	162099	47089	558009
2108	5264	1373536	567498	3487562

المجاميع المطلوبة
$\sum x = 2108 \ , \ \sum y = 5264$
$\sum xy = 1373536$
$\sum x^2 = 567498$
$\sum y^2 = 3487562$

- حساب معامل الارتباط:

باستخدام المجاميع السابقة، وبالتطبيق على المعادلة أعلاه، نجد أن معامل الارتباط قيمته هي:

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

$$= \frac{1373536 - \frac{(2108)(5264)}{8}}{\sqrt{\left(567498 - \frac{(2108)^2}{8}\right)\left(3487562 - \frac{(5264)^2}{8}\right)}}$$

$$= \frac{-13528}{\sqrt{(12040)(23850)}} = \frac{-13528}{16945.619} = -0.798$$

وهي نفس النتيجة السابقة:

3/2/6 معامل ارتباط الرتب (اسبيرمان) Spearman

إذا كانت الظاهرة محل الدراسة تحتوي على متغيرين وصفيين ترتيبين، ومثال على ذلك قياس العلاقة بين تقديرات الطلبة في مادتين ، أو العلاقة بين درجة تفضيل المستهلك لسلعة معينة ، ومستوى الدخل، فإنه يمكن استخدام طريقة "بيرسون" السابقة في حساب معامل ارتباط يعتمد على رتب مستويات المتغيرين كبديل للقيم الأصلية ، ويطلق على هذا المعامل " معامل ارتباط اسبيرمان " Spearman ، ويعبر عنه بالمعادلة التالية :

$$\boxed{r_s = 1 - \frac{6\sum d^2}{n(n^2-1)}} \qquad (٤-٦)$$

حيث أن d هي الفرق بين رتب مســـــتويات المتغير الأول x ، ورتب مســـــتويات المتغير الثاني y ، أي أن : $d = R_x - R_y$.

مثـال

فيما يلي تقديرات 10 طلاب في مادتي الإحصاء، والاقتصاد:

تقديرات إحصاء	ب+	ب+	ب	أ+	ج+	ب+	د+	د	ج+	أ
تقديرات اقتصاد	ب	ج	ب	ب+	ب	أ	ج	ج	د	أ+

والمطلوب:

1- احسب معامل الارتباط بين تقديرات الطلبة في المقررين.

2- وما هو مدلوله ؟

الحـل

1- بفرض أن x هي تقديرات الإحصــاء، y هي تقديرات الاقتصــاد، يمكن حســاب معامل الارتباط بينهما باســتخدام المعادلة (**6-4**)، وذلك بإتباع الآتي:

الرتب	1	2	3	4	5	6	7	8	9	10
تقديرات إحصاء	أ+	أ	ب+	ب+	ب+	ب	ج+	ج+	د+	د
رتب x	1	2	(3+4+5)/3=4			6	(7+8)/2=7.5		9	10
تقديرات اقتصاد	أ+	أ	ب+	ب	ب	ب	ج	ج	ج	د
رتب y	1	2	3	(4+5+6)/3=5			(7+8+9)/3=8			10

- إذا يمكن حساب المجموع: $\sum d^2$ كما يلي:

x	y	رتب x	رتب y	d	d^2
أ	أ+	2	1	1	1
ج+	د	7.5	10	-2.5	6.25
د	ج	10	8	2	4
د+	ج	9	8	1	1
ب+	أ	4	2	2	1
ج+	ب	7.5	5	2.5	6.25
أ+	ب+	1	3	-2	4
ب	ب	6	5	1	1
ب+	ج	4	8	-4	16
ب+	ب	4	5	-1	1
					44.5

$\sum d^2 = 44.5$

- **معامل الارتباط هو:**

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

$$= 1 - \frac{6(44.5)}{10(10^2-1)} = 1 - \frac{267}{990}$$

$$= 1 - 0.2697 = 0.7303$$

2- مدلول معامل الارتباط :

بما أن $r = 0.703$ ، ويدل ذلك على وجود ارتباط طردي قوي بين تقديرات الطالب في مادة الإحصاء ، ومادة الاقتصاد .

ملحوظة:- يمكن استخدام صيغة معامل ارتباط "اسبيرمان" في حساب الارتباط بين متغيرين كميين، حيث يتم استخدام رتب القيم التي يأخذها المتغير، ونترك للطالب القيام بحساب معامل ارتباط الرتب بين المساحة والكمية في مثال (5-1) السابق، وعليه أن يقوم بتفسير النتيجة: (معاونة : $\sum d^2 = 148$)

3/6 الانحدار الخطى البسيط Simple Regression

إن الغرض من استخدام أسلوب تحليل الانحدار الخطي البسيط، هو دراسة وتحليل أثر متغير كمي على متغير كمي آخر، ومن الأمثلة على ذلك ما يلي:

- **دراسة أثر كمية السماد على إنتاجية الدونم.**
- **دراسة أثر الإنتاج على التكلفة.**
- **دراسة أثر كمية البروتين التي يتناولها الأبقار على الزيادة في الوزن.**
- **أثر الدخل على الإنفاق الاستهلاكي.**

وهكذا هناك أمثلة في كثير من النواحي الاقتصادية، والزراعية، والتجارية، والعلوم السلوكية، وغيرها من المجالات الأخرى.

نموذج الانحدار الخطي

في تحليل الانحدار البسيط، نجد أن الباحث يهتم بدراسة أثر أحد المتغيرين ويسمى بالمتغير المستقل أو المتنبأ منه، على المتغير الثاني ويسمى بالمتغير التابع أو المتنبأ به، ومن ثم يمكن عرض نموذج الانحدار الخطي في شكل معادلة خطية من الدرجة الأولى، تعكس المتغير التابع كدالة في المتغير المستقل كما يلي:

$$\boxed{y = \beta_0 + \beta_1 x + e} \qquad (٥-٦)$$

.

حيث أن:

y : هو المتغير التابع (الذي يتأثر)

x : هو المتغير المستقل (الذي يؤثر)

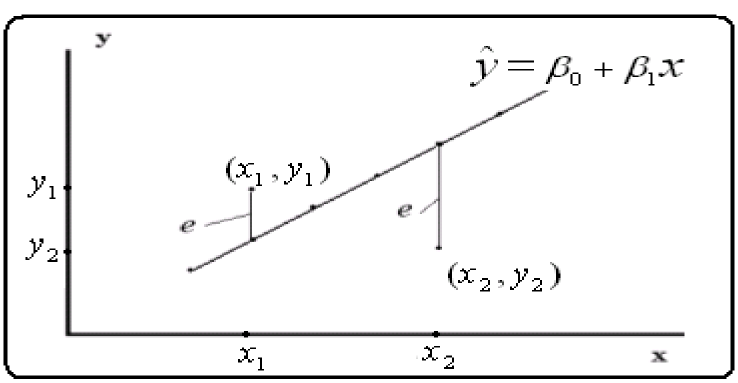
β_0 : هو الجزء المقطوع من المحور الرأسي y ، وهو يعكس قيمة المتغير التابع في حالة انعدام قيمة المتغير المستقل x ، أي في حالة $x = 0$

و يسمى بثابت الانحدار او)(a)

β_1 : ميل الخط المستقيم $(\beta_0 + \beta_1 x)$ ، ويعكس مقدار التغير في y إذا تغيرت x بوحدة واحدة.

و يسمى معامل الانحدار او (b)

e : هو الخطأ العشـــوائي، والـذي يعبر عن الفرق بين القيمـة الفعليـة y ، والقيمـة المقـدرة $\hat{y} = \beta_0 + \beta_1 x$ ، أي أن : $e = y - (\beta_0 + \beta_1 x)$ ، ويمكن توضـــيح هذا الخطأ على الشكل التالي لنقط الانتشار.



و يمكن كتابة المعادلة

$$\boxed{y = a + b(x) + e}$$

و هي تكتب هكذا في بعض المراجع

تقدير نموذج الانحدار الخطي البسيط

يمكن تقدير معاملات الانحدار $(\beta_1 \ , \ \beta_0)$ في النموذج باستخدام طريقة المربعات الصغرى، وهذا التقدير هو الذي يجعل مجموع مربعات الأخطاء العشوائية $\sum e^2 = \sum\left(y - (\beta_0 + \beta_1 x)\right)^2$ أقل ما يمكن، ويحسب هذا التقدير بالمعادلة التالية:

$$\boxed{\begin{array}{l} \hat{\beta}_1 = \dfrac{n\sum xy - \sum x \ \sum y}{n\sum x^2 - (\sum x)^2} \ , \\ \hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x} \end{array}} \qquad (٦-٦)$$

حيث أن $\bar{x}$ هو الوسط الحسابي لقيم x ، $\bar{y}$ هو الوسط الحسابي لقيم y ، وتكون القيمة المقدرة للمتغير التابع هو: $\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x$ ، ويطلق على هذا التقدير " تقدير معادلة انحدار y على x .

مثال

فيما يلي بيانات عن كمية البروتين اليومي بالجرام التي يحتاجها العجل الرضيع، ومقدار الزيادة في وزن العجل بالكجم، وذلك لعينة من العجول الرضيعة حجمها 10.

كمية البروتين	10	11	14	15	20	25	46	50	59	70
الزيادة في الوزن	10	10	12	12	13	13	19	15	16	20

والمطلوب :

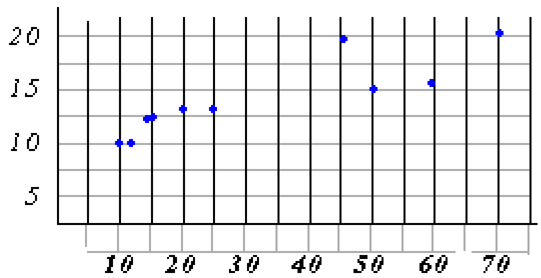
1- ارسم نقط الانتشار، وما هو توقعاتك لشكل العلاقة ؟
2- قدر معادلة انحدار الوزن على كمية البروتين.
3- فسر معادلة الانحدار.
4- ما هو مقدار الزيادة في الوزن عند إعطاء العجل 50 جرام من البروتين ؟ وما هو مقدار الخطأ العشوائي؟

5- ارسم معادلة الانحدار على نقط الانتشار في المطلوب (**1**) .

الحـل

1- رسم نقط الانتشار:

مقدار الزيادة y



كمية البروتين x

من المتوقع أن يكون لكمية البروتين أثر طردي (إيجابي) على مقدار الزيادة في الوزن.

2- تقدير معادلة الانحدار.

بفرض أن x هي كميـة البروتين، y هي مقـدار الزيـادة في الوزن، يمكن تطبيق المعـادلتين ومن ثم يتم حســـاب المجاميع التالية:

كمية البروتين x	الزيادة في الوزن y	$x\ y$	x^2
10	10	100	100
11	10	110	121
14	12	168	196
15	12	180	225
20	13	260	400
25	13	325	625
46	19	874	2116
50	15	750	2500
59	16	944	3481
70	20	1400	4900
320	140	5111	14664

المجاميع المطلوبة
$\sum x = 320$ $\sum y = 140$ $\sum xy = 5111$ $\sum x^2 = 14664$ **إذا الوسط الحسابي:** $\bar{x} = \frac{\sum x}{n} = \frac{320}{10} = 32$ $\bar{y} = \frac{\sum x}{n} = \frac{140}{10} = 14$

- بتطبيق المعادلة الأولى يمكن حساب $\hat{\beta}_1$ كما يلي:

$$\hat{\beta}_1=\frac{n\sum xy-\sum x\sum y}{n\sum x^2-(\sum x)^2}=\frac{(10)(5111)-(320)(140)}{(10)(14664)-(320)^2}$$
$$=\frac{6310}{44240}=0.1426$$

- بتطبيق المعادلة الثانية يمكن حساب $\hat{\beta}_0$ كما يلي:

$$\hat{\beta}_0=\bar{y}-\hat{\beta}_1\bar{x}=14-(0.1426)(32)=9.4368$$

- إذا معادلة الانحدار المقدرة، هي:

$$\hat{y}=9.44+0.143x$$

3- تفسير المعادلة:

- الثابت $\hat{\beta}_0=9.44$: يدل على أنه في حالة عدم اســـتخدام البروتين قي التغذية، فإن الوزن يزيد 9.44 كجم.
- معامل الانحدار $\hat{\beta}_1=0.143$: يدل على أنه كلما زادت كمية البروتين جرام واحد، حدث زيادة في وزن العجل بمقدار 0.143 كجم، أى زيادة مقدارها 143 جرام.

4- مقدار الزيادة في الوزن عند $x=50$ هو:

$$\hat{y}=9.44+0.143(50)=16.59$$

وأما ومقدار الخطأ العشوائي هو:

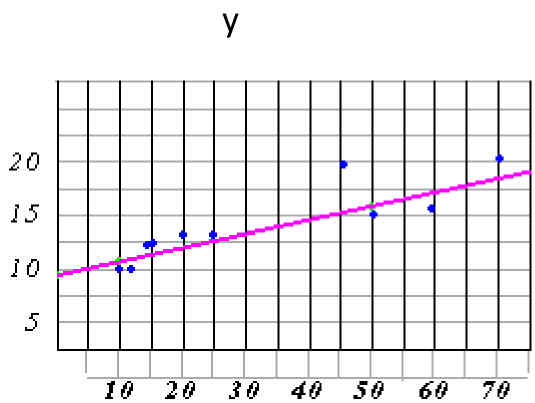
$$\hat{e}_{x=50}=y_{x=50}-\hat{y}_{x=50}=15-16.59=-1.59$$

5- رسم معادلة الانحدار على نقط الانتشار.

يمكن رسم معادلة خط مستقيم إذا علم نقطتين على الخط المستقيم.

x	50	10
$\hat{y}$	16.59	10.87

إذا معادلة الانحدار هي:



x

معامل الارتباط الخطي لبيرسون
Pearson's Coefficent of Correlation

يستخدم هذا المعامل لحساب الارتباط الخطي بين متغيرين يرمز للاول X وهو المتغير المستقل والثاني y وهو المتغير التابع ويمكن حسابه بالمعادلة التالية:

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

مثال:
تم حساب عدد الحوادث في أربعة أيام وكمية المطر cm في قرية على النحو التالي:

كمية المطر X	2	1	5	0
الحوادث y	5	3	6	2

أوجد هل هناك ارتباط بين عدد الحوادث وكمية المطر وما هي قوة ونوع هذا الارتباط؟

الحل: ننشئ الجدول اللازم لإيجاد عناصر المعادلة على النحو التالي:

X	y	xy	X^2	y^2
2	5	10	4	25
1	3	3	1	9
5	6	30	25	36
0	2	0	0	4
$\sum x$ 8	$\sum y$ 16	$\sum xy$ 43	$\sum x^2$ 30	$\sum y^2$ 74

ثم نطبق المعادلة السابقة:

$$r = \frac{43 - \frac{(8)(16)}{4}}{\sqrt{30 - \frac{(8)^2}{4}}\ \sqrt{74 - \frac{(16)^2}{4}}}$$

$$r = \frac{43 - 32}{\sqrt{30 - 16}\ \sqrt{74 - 64}}$$

$$r = \frac{11}{\sqrt{14}\ \sqrt{10}} = \frac{11}{(3.74)(3.16)} = \frac{11}{11.8}$$

$$r = \boxed{0.93}$$

إذا هناك علاقة خطية طردية موجبة بين عدد الحوادث وكمية المطر.

ويمكن كتابة معادلة بيرسون للانحدار على النحو التالي أيضاً

$$r = \frac{n(\Sigma xy) - (\Sigma x)(\Sigma y)}{\sqrt{n(\Sigma x^2) - (\Sigma x)^2}\sqrt{n(\Sigma y^2) - (\Sigma y)^2}}$$

ويمكن حل المثال السابق على النحو التالي:

$$r = \frac{4(43) - (8)(16)}{\sqrt{4(30) - (64)}\sqrt{4(74)(258)}}$$

$$r = \frac{(172) - (128)}{\sqrt{56}\sqrt{38}}$$

$$r = \frac{(44)}{(7.5)(6.2)}$$

$$r = \frac{44}{47} = \boxed{.93}$$

وهذا كما أشرنا سابقاً بين أن هناك علاقة خطية طردية موجبة بين كمية المطر الساقطة وعدد الحوادث في تلك القرية.

Spearman Coefficient
Rank Correlation

c- معامل الارتباط الخطي للرتب لسبيرمان

هذا المعامل يقيس الارتباط لمتغيرين ممثلين ببيانات وصفية ~~ونحو~~ مثل مقياس العلاقة بين مذاقه ماله معينة وغيرها. ويمكن تطبيقه للعلاقات كمية بعد ترتيبها ووضعها في رتب معينة. ولكن أيضاً هذا المعامل المعامل يصلح بشكل خاص إذا كان عدد الأزواج للعلاقات ما بين 25-30 أو أقل ويكون حساب معامل سبيرمان على النحو التالي

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

حيث أن n = عدد أزواج البيانات (X,Y) و d تساوي الفرق بين رتب (Y,X) أي أن

$$d^2 = (a-b)^2$$

مثال لعمل الرتب: أوجد رتب التقديرات الآتية B, C, B, D, D, A, E

الحل

X - التقديرات	A	D	D	E	B	C	B
X رتب	7	2.5	2.5	1	5.5	4	5.5

مثال لإيجاد معامل سبيرمان للعلاقات وصفية وضعت في رتب

أوجد معامل ارتباط الرتب لتقديرات الطلاب في مادتي الإحصاء والرياضيات كما هو موضح بالجدول التالي:

الرياضيات X	A	C	C	C	B	D
الإحصاء Y	B	B	D	C	A	E

الحل

الرياضيات X	الإحصاء Y	رتبة X a	رتبة Y b	d = a − b	$d^2 = (a-b)^2$
A	B	6	4.5	1.5	2.25
C	B	3	4.5	− 1.5	2.25
C	D	3	2	1	1
C	C	3	3	0	0
B	A	5	6	− 1	1
D	E	1	1	0	0
				0	6.5

$\sum d^2$

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

$$\therefore r_s = 1 - \frac{6(6.5)}{6(36-1)}$$

$$= 1 - 0.186 = 0.814$$

أي أنه يوجد ارتباط طردي (موجب) قوي بين تقديرات الطلاب في مادة الإحصاء والرياضيات.

٢- معامل الاقتران ومعامل التوافق (Coefficient of Contingency)

أ- معامل الاقتران يستخدم عندما يكون لكل ظاهرتين تحتمل وجود تداخل بينهما صفتين فقط مثل تقسيم العلاقة بين الحالة الزوجية (متزوج ، غير متزوج) والحالة الوظيفية (موظف ، غير موظف) .

أما معامل التوافق فيستخدم عندما يكون لكل ظاهرة أو واحدة من الظواهر أكثر من صفتين مثل الحالة الزوجية (متزوج ، مطلق ، غير متزوج) . في هذا الفصل لن نتطرق لمعامل التوافق ونكتفي بمعامل الاقتران الذي يمكن حسابه على النحو التالي

$$C.C = \frac{AD - BC}{AD + BC}$$

حيث أن AB تمثل الصفتين للظاهرة الأولى و CD تمثل الصفتين للظاهرة الثانية

مثال : عند دراسة علاقة الترفيه بالتعليم في مؤسستين مختلفتين أخذت عينة مكونة من 17 شخص وحددت الصفات على النحو التالي

التعليم \ الترفيه	يرفض	لا يرفض
متعلم	A	B
غير متعلم	C	D

وبعد فرز المعلومات كانت النتائج على النحو التالي في المؤسسة الأولى

التعليم \ الترفيه	يرفض	لا يرفض
متعلم	5	5
غير متعلم	3	4

في هذه المؤسسة

$$CC = \frac{(5\times4)-(3\times5)}{(5\times4)+(3\times5)} = \frac{5}{35} = .14$$

وهو ارتباط ضعيف

أما المؤسسة الثانية

التعليم \ الترفيه	يرفض	لا يرفض
متعلم	2	8
غير متعلم	5	2

في الحالة الثانية

$$CC = \frac{(2\times2)-(8\times5)}{(2\times2)+(8\times5)} = \frac{-36}{44}$$

$$= -0.9$$

أي كلما قل التعليم زاد الترفيه

*الارتباط ومعنوية الارتباط.

تعلمنا فيما مضى طريقة حساب معامل بيرسون للارتباط والذي يمكن حسابه بالنحو التالي

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

والتي يمكن كتابتها بصيغ أخرى.

مثال: أحسب معامل الارتباط بين معدل الرطوبة النسبية في المدينة A و B حيث A هي المتغير التابع و B المتغير المستقل (X و Y).

الحل: ننشئ الجدول أسفل للحصول على قيم عناصر المعادلة أعلاه:

X	Y	XY	X^2	Y^2
75	69	5175	5625	4761
82	85	6970	6724	7225
65	55	3575	4225	3025
90	90	8100	8100	8100
77	80	6160	5929	6400
60	50	3000	3600	2500
55	50	2750	3025	2500
87	90	7830	7569	8100
91	85	7735	8281	7225
682	654	51295	53078	49836

$$r=\frac{9\times 51295-682\times 654}{\sqrt{9\times 53078-(682)^2}\sqrt{9\times 49836-(654)^2}}$$

$$r=\frac{461655-446028}{\sqrt{477702-465124}\sqrt{448524-427716}}$$

$$=\frac{15627}{\sqrt{12578}\sqrt{20808}}$$

$$=\frac{15627}{16177.856}=0.966$$

$n=9$

$\bar{X}=75.78$

$\bar{Y}=72.67$

واضح جداً أن هناك ارتباط طردي قوي جداً ولكن هل هذا الارتباط معنوي؟ للتحقق من ذلك نقيس معنوية الارتباط

*معنوية الارتباط تقاس على النحو التالي

$$t=\frac{r\sqrt{n-2}}{\sqrt{1-r^2}} \quad \text{أو} \quad t=\frac{r}{\sqrt{(1-r^2)/(n-2)}}$$

ففي المثال السابق

$$t=\frac{.966}{\sqrt{(1-.966^2)/(9-2)}}$$

$$t=\frac{.966}{\sqrt{(1-.93)/(7)}}=\frac{.966}{\sqrt{.07/7}}=9.7$$

بعد ذلك نقارن بين قيمة t المحسوبة والجدولية عند $t(\alpha/2, n-2)$ والتي مقدارها 2.365 وبما أن المحسوبة أكبر نستطيع أن نقول أن الارتباط $r=.966$ ذو معنوية ويمكن صياغة ذلك على هيئة اختبار فرضية حيث

$H_0: R=0 \qquad H_1: R\neq 0$

حيث أن R تمثل معامل الارتباط للمجتمع الإحصائي.

* الانحدار والانحدار المتعدد Regression and Multiple regression

في ٣-٢ تحدثنا أن العلاقة الخطية بين متغيرين أحدهما X المستقل والآخر Y التابع يمكن أن يعبر عنها بالمعادلة الخطية التالية ←

$$\hat{Y} = a + b(x)$$

وذكرنا أن من أهم أغراض دراسة الانحدار هو التنبؤ بقيمة المتغير التابع بمعرفة قيمة المتغير المستقل ونود أن نشير هنا أن المعادلة السابقة هي تقدير لمعادلة الانحدار للمجتمع الإحصائي وهي

$$\hat{Y} = A + B(X)$$

* الفرق بين القيم الفعلية لـ Y والقيم المقدرة $\hat{Y}$ يسمى الخطأ أو الانحراف من القيمة الفعلية وهو e

$$e = Y - \hat{Y}$$

وقيم a و b يمكن إيجادها بطريقة المربعات الصغرى least square والتي تجعل مجموع مربعات الانحرافات $Y - \hat{Y}$ أصغر ما يمكن.

* وقيم b والتي تسمى معامل الانحدار regression coefficient هي مقدار التغير في المتغير التابع لكل وحدة تغير في المتغير المستقل ويمكن حسابها على النحو التالي

$$b = \frac{\sum x_i y_i - n\bar{x}\bar{y}}{\sum x_i^2 - n\bar{x}^2}$$

أما a وهي المتقاطع أو هي قيمة Y عندما تكون قيمة X صفر يمكن حسابها على النحو التالي

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

مثال: أحسب معادلة خط الانحدار للعلاقة بين الرطوبة النسبية في المدينة B (المتغير المستقل) والمدينة A (المتغير التابع) ثم قدر الرطوبة النسبية في A عندما تكون الرطوبة النسبية في B = 60%.

الحل بتطبيق المعادلة أعلاه

$$b = \frac{51245 - 9 * 75.8 * 72.7}{53078 - 9 * (75.8)^2}$$

$$= \frac{51295 - 49558.8}{53078 - 51680.4} = \frac{1736.2}{1397.6} = 1.24$$

وهذا الانحدار ←

أما المتقاطع فهو

$$a = 72.7 - 1.24 * 75.8 = -21.45$$

وبذلك تكون معادلة الانحدار هي ←

$$\boxed{\hat{Y} = -21.45 + 1.24X}$$

وقيمة Y عندما تكون قيمة X = 60% هي

$$\hat{Y} = -21.45 + 1.24 * 60 = 52.9\%$$

3 Grr
page 9

* الخطاء المعياري لخط الانحدار والدلالة الإحصائية لنموذج الانحدار

- الفرق بين القيم الحقيقة Y والمتوقعة $\hat{Y}$ يسمى بالخطاء أو البواقي Residual. الخطاء المعياري هو تقدير للانحراف المعياري للانحرافات القيم $\hat{Y}$ عن Y وهو يقيس مدى يقيس الخطاء في تقدير Y.

ويمكن حسابه على النحو التالي

$$S_{Y\cdot X} = \sqrt{\frac{\Sigma (Y-\hat{Y})^2}{n-2}}$$

ودرجة الحرية تعتمد على عدد المتغيرات وهي هنا = 2 لذلك $n-2$

مثال احسب الخطاء المعياري لنموذج خط الانحدار الذي حصلنا عليه في المثال السابق وهو $\hat{Y} = -21.45 + 1.24X$

الحل: ننشئ الجدول أسفل ثم نطبق المعادلة أعلاه على النحو التالي

	X	Y	$\hat{Y} = -21.45 + 1.24X$	$Y-\hat{Y}$	$(Y-\hat{Y})^2$
1	75	69	71.67	-2.67	7.13
2	82	85	80.36	4.64	21.49
3	65	55	59.25	-4.25	18.06
4	90	90	90.30	-.30	.09
5	77	80	74.15	5.85	34.18
6	60	50	53.04	-3.04	9.24
7	55	50	46.83	3.17	10.05
8	87	90	86.57	3.43	11.74
9	91	85	91.54	-6.54	42.80
Σ	682	654	653.7	0	154.8

$$S_{X\cdot Y} = \sqrt{154.8/7} = \sqrt{22.11} = 4.70$$

وهذا يعني أنه عند تطبيق المعادلة للتوقع بقيم Y وهي $\hat{Y}$ فإن النتيجة

$\hat{Y} \pm 4.7$

فمثلاً في الحالة الأولى كانت النتيجة المتوقعة هي 71.67 وعند القياس 90% فإن القيمة المتوقعة هنا تساوي 71.67 ± 4.7

فالحد الأعلى = 76.37

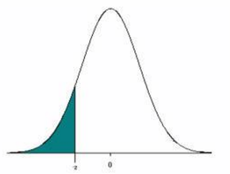
والحد الأدنى = 66.97

* معامل التحديد r^2 أو Coefficient of Determination

هو النسبة بين التباين المفسر في المتغير التابع Y إلى التباين الكلي للمتغير التابع أي أنه التباين في Y والذي يعزى للمتغير X وهو المتغير المستقل. وهو مربع معامل الارتباط بيرسون في مثالنا: حيث أن $r = .966$ إذاً $r^2 = .933$

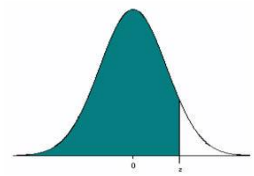
وهذا يعني ....

Table of Standard Normal Probabilities for Negative Z-scores



z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
-3.4	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0002
-3.3	0.0005	0.0005	0.0005	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004	0.0003
-3.2	0.0007	0.0007	0.0006	0.0006	0.0006	0.0006	0.0006	0.0005	0.0005	0.0005
-3.1	0.0010	0.0009	0.0009	0.0009	0.0008	0.0008	0.0008	0.0008	0.0007	0.0007
-3.0	0.0013	0.0013	0.0013	0.0012	0.0012	0.0011	0.0011	0.0011	0.0010	0.0010
-2.9	0.0019	0.0018	0.0018	0.0017	0.0016	0.0016	0.0015	0.0015	0.0014	0.0014
-2.8	0.0026	0.0025	0.0024	0.0023	0.0023	0.0022	0.0021	0.0021	0.0020	0.0019
-2.7	0.0035	0.0034	0.0033	0.0032	0.0031	0.0030	0.0029	0.0028	0.0027	0.0026
-2.6	0.0047	0.0045	0.0044	0.0043	0.0041	0.0040	0.0039	0.0038	0.0037	0.0036
-2.5	0.0062	0.0060	0.0059	0.0057	0.0055	0.0054	0.0052	0.0051	0.0049	0.0048
-2.4	0.0082	0.0080	0.0078	0.0075	0.0073	0.0071	0.0069	0.0068	0.0066	0.0064
-2.3	0.0107	0.0104	0.0102	0.0099	0.0096	0.0094	0.0091	0.0089	0.0087	0.0084
-2.2	0.0139	0.0136	0.0132	0.0129	0.0125	0.0122	0.0119	0.0116	0.0113	0.0110
-2.1	0.0179	0.0174	0.0170	0.0166	0.0162	0.0158	0.0154	0.0150	0.0146	0.0143
-2.0	0.0228	0.0222	0.0217	0.0212	0.0207	0.0202	0.0197	0.0192	0.0188	0.0183
-1.9	0.0287	0.0281	0.0274	0.0268	0.0262	0.0256	0.0250	0.0244	0.0239	0.0233
-1.8	0.0359	0.0351	0.0344	0.0336	0.0329	0.0322	0.0314	0.0307	0.0301	0.0294
-1.7	0.0446	0.0436	0.0427	0.0418	0.0409	0.0401	0.0392	0.0384	0.0375	0.0367
-1.6	0.0548	0.0537	0.0526	0.0516	0.0505	0.0495	0.0485	0.0475	0.0465	0.0455
-1.5	0.0668	0.0655	0.0643	0.0630	0.0618	0.0606	0.0594	0.0582	0.0571	0.0559
-1.4	0.0808	0.0793	0.0778	0.0764	0.0749	0.0735	0.0721	0.0708	0.0694	0.0681
-1.3	0.0968	0.0951	0.0934	0.0918	0.0901	0.0885	0.0869	0.0853	0.0838	0.0823
-1.2	0.1151	0.1131	0.1112	0.1093	0.1075	0.1056	0.1038	0.1020	0.1003	0.0985
-1.1	0.1357	0.1335	0.1314	0.1292	0.1271	0.1251	0.1230	0.1210	0.1190	0.1170
-1.0	0.1587	0.1562	0.1539	0.1515	0.1492	0.1469	0.1446	0.1423	0.1401	0.1379
-0.9	0.1841	0.1814	0.1788	0.1762	0.1736	0.1711	0.1685	0.1660	0.1635	0.1611
-0.8	0.2119	0.2090	0.2061	0.2033	0.2005	0.1977	0.1949	0.1922	0.1894	0.1867
-0.7	0.2420	0.2389	0.2358	0.2327	0.2296	0.2266	0.2236	0.2206	0.2177	0.2148
-0.6	0.2743	0.2709	0.2676	0.2643	0.2611	0.2578	0.2546	0.2514	0.2483	0.2451
-0.5	0.3085	0.3050	0.3015	0.2981	0.2946	0.2912	0.2877	0.2843	0.2810	0.2776
-0.4	0.3446	0.3409	0.3372	0.3336	0.3300	0.3264	0.3228	0.3192	0.3156	0.3121
-0.3	0.3821	0.3783	0.3745	0.3707	0.3669	0.3632	0.3594	0.3557	0.3520	0.3483
-0.2	0.4207	0.4168	0.4129	0.4090	0.4052	0.4013	0.3974	0.3936	0.3897	0.3859
-0.1	0.4602	0.4562	0.4522	0.4483	0.4443	0.4404	0.4364	0.4325	0.4286	0.4247
-0.0	0.5000	0.4960	0.4920	0.4880	0.4840	0.4801	0.4761	0.4721	0.4681	0.4641

Table of Standard Normal Probabilities for Positive Z-scores



z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998

Note that the probabilities given in this table represent the area to the LEFT of the z-score.
The area to the RIGHT of a z-score = 1 – the area to the LEFT of the z-score

Student t-Table

Alpha	0.250	0.200	0.150	0.100	0.050	0.025	0.010	0.005	0.0005
df									
1	1.000	1.376	1.963	3.078	6.314	12.706	31.821	63.656	636.578
2	0.816	1.061	1.386	1.886	2.920	4.303	6.965	9.925	31.600
3	0.765	0.978	1.250	1.638	2.353	3.182	4.541	5.841	12.924
4	0.741	0.941	1.190	1.533	2.132	2.776	3.747	4.604	8.610
5	0.727	0.920	1.156	1.476	2.015	2.571	3.365	4.032	6.869
6	0.718	0.906	1.134	1.440	1.943	2.447	3.143	3.707	5.959
7	0.711	0.896	1.119	1.415	1.895	2.365	2.998	3.499	5.408
8	0.706	0.889	1.108	1.397	1.860	2.306	2.896	3.355	5.041
9	0.703	0.883	1.100	1.383	1.833	2.262	2.821	3.250	4.781
10	0.700	0.879	1.093	1.372	1.812	2.228	2.764	3.169	4.587
11	0.697	0.876	1.088	1.363	1.796	2.201	2.718	3.106	4.437
12	0.695	0.873	1.083	1.356	1.782	2.179	2.681	3.055	4.318
13	0.694	0.870	1.079	1.350	1.771	2.160	2.650	3.012	4.221
14	0.692	0.868	1.076	1.345	1.761	2.145	2.624	2.977	4.140
15	0.691	0.866	1.074	1.341	1.753	2.131	2.602	2.947	4.073
16	0.690	0.865	1.071	1.337	1.746	2.120	2.583	2.921	4.015
17	0.689	0.863	1.069	1.333	1.740	2.110	2.567	2.898	3.965
18	0.688	0.862	1.067	1.330	1.734	2.101	2.552	2.878	3.922
19	0.688	0.861	1.066	1.328	1.729	2.093	2.539	2.861	3.883
20	0.687	0.860	1.064	1.325	1.725	2.086	2.528	2.845	3.850

Ch Reg

Cont

* الدلالة الإحصائية لنموذج الانحدار

الأساسية: اختبار الدلالة هو الاعتماد على تحليل التباين ANOVA حيث نحتاج حساب مجموع المربعات SST للمتغير Y ونقوم بتجزئته إلى جزئين هما الجزء المفسر SSR والآخر غير مفسر SSE

Sum of Squares for Regression and Sum of squares for errors

ومن ثم بواسطة هذين الجزئين نحسب قيمة F ونقارنها بقيمة F الجدولية عند مستوى معنوية معين

وهناك طرق عديدة لحساب SST, SSR and SSE

ولكن من أسهل الطرق المعادلات الثلاث الآتية

1 – $SST = \sum Y^2 - n\bar{Y}$ or $SSR + SSE$

2 – $SSR = b[\sum xy - n\bar{x}\bar{y}]$ or $SST - SSE$

3 – $SSE = \sum(Y - \hat{Y})^2$ or $SST - SSR$

وعند الانتهاء من حساب تلك القيم نتوجه لإنشاء جدول تحليل التباين وبعد ذلك نقارن بين قيم F الجدولية و القيم المحسوبة ثم نحدد معنوية أو عدم معنوية بناء على مستوى المعنوية المطلوب وجدول تحليل التباين في تحليل الانحدار البسيط (متغير مستقل واحد فقط) يكون على النحو التالي:

هنا MS_x هو معدل المربعات لـ SSR وحسبت بقسمة SSR على درجة حرية واحدة في تحليل التباين للانحدار البسيط. أما S_{YX}^2 فهو معدل المربعات لـ SSE وهو تقدير تباين خطأ انحدار Y على X ويحسب

Source of Variation مصدر التغير	Sum of squars مجموع المربعات SS	درجات الحرية DF	معدل مجموع المربعات SSM	المحسوبة f
بسبب الانحدار (SSR)	$\Sigma(\hat{Y}_i - \bar{Y})^2$	$k-1$	$MS_x = \frac{SSR}{1}$	$\frac{MS_x}{S_{YX}^2} = f$
الانحرافات عن الانحدار (SSE)	$\Sigma(Y_i - \hat{Y}_i)^2$	$n-k$	$S_{YX}^2 = \frac{SSE}{n-2}$	
الكلي SST	$\Sigma(Y_i - \bar{Y})^2$	$n-1$		

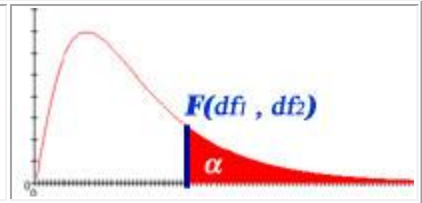
بالمعادلة $S_{YX}^2 = \frac{SSE}{n-2}$ والجذر التربيعي له يعطي ما يعرف الخطأ المعياري لتقدير خطأ انحدار Y على X وهو $S_{YX} = \sqrt{\frac{\sum(Y-\hat{Y})^2}{n-2}}$ والذي سيتم الحديث عنه

ونود أن نشير إلى أن R و R^2 يمكن

$R^2 = SSR / SST$

$R = \sqrt{SSR/SST}$

F Table for α = 0.10

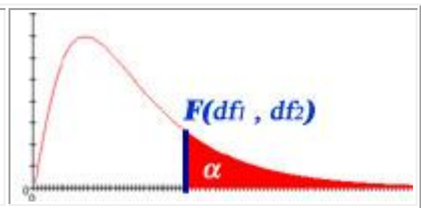


\	df_1=1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
df_2=1	39.86346	49.50000	53.59324	55.83296	57.24008	58.20442	58.90595	59.43898	59.85759	60.19498	60.70521	61.22034	61.74029	62.00205	62.26497	62.52905	62.79428	63.06064	63.32812
2	8.52632	9.00000	9.16179	9.24342	9.29263	9.32553	9.34908	9.36677	9.38054	9.39157	9.40813	9.42471	9.44131	9.44962	9.45793	9.46624	9.47456	9.48289	9.49122
3	5.53832	5.46238	5.39077	5.34264	5.30916	5.28473	5.26619	5.25167	5.24000	5.23041	5.21562	5.20031	5.18448	5.17636	5.16811	5.15972	5.15119	5.14251	5.13370
4	4.54477	4.32456	4.19086	4.10725	4.05058	4.00975	3.97897	3.95494	3.93567	3.91988	3.89553	3.87036	3.84434	3.83099	3.81742	3.80361	3.78957	3.77527	3.76073
5	4.06042	3.77972	3.61948	3.52020	3.45298	3.40451	3.36790	3.33928	3.31628	3.29740	3.26824	3.23801	3.20665	3.19052	3.17408	3.15732	3.14023	3.12279	3.10500
6	3.77595	3.46330	3.28876	3.18076	3.10751	3.05455	3.01446	2.98304	2.95774	2.93693	2.90472	2.87122	2.83634	2.81834	2.79996	2.78117	2.76195	2.74229	2.72216
7	3.58943	3.25744	3.07407	2.96053	2.88334	2.82739	2.78493	2.75158	2.72468	2.70251	2.66811	2.63223	2.59473	2.57533	2.55546	2.53510	2.51422	2.49279	2.47079
8	3.45792	3.11312	2.92380	2.80643	2.72645	2.66833	2.62413	2.58935	2.56124	2.53804	2.50196	2.46422	2.42464	2.40410	2.38302	2.36136	2.33910	2.31618	2.29257
9	3.36030	3.00645	2.81286	2.69268	2.61061	2.55086	2.50531	2.46941	2.44034	2.41632	2.37888	2.33962	2.29832	2.27683	2.25472	2.23196	2.20849	2.18427	2.15923
10	3.28502	2.92447	2.72767	2.60534	2.52164	2.46058	2.41397	2.37715	2.34731	2.32260	2.28405	2.24351	2.20074	2.17843	2.15543	2.13169	2.10716	2.08176	2.05542
11	3.22520	2.85951	2.66023	2.53619	2.45118	2.38907	2.34157	2.30400	2.27350	2.24823	2.20873	2.16709	2.12305	2.10001	2.07621	2.05161	2.02612	1.99965	1.97211
12	3.17655	2.80680	2.60552	2.48010	2.39402	2.33102	2.28278	2.24457	2.21352	2.18776	2.14744	2.10485	2.05968	2.03599	2.01149	1.98610	1.95973	1.93228	1.90361
13	3.13621	2.76317	2.56027	2.43371	2.34672	2.28298	2.23410	2.19535	2.16382	2.13763	2.09659	2.05316	2.00698	1.98272	1.95757	1.93147	1.90429	1.87591	1.84620

14	3.10221	2.72647	2.52222	2.39469	2.30694	2.24256	2.19313	2.15390	2.12195	2.09540	2.05371	2.00953	1.96245	1.93766	1.91193	1.88516	1.85723	1.82800	1.79728
15	3.07319	2.69517	2.48979	2.36143	2.27302	2.20808	2.15818	2.11853	2.08621	2.05932	2.01707	1.97222	1.92431	1.89904	1.87277	1.84539	1.81676	1.78672	1.75505
16	3.04811	2.66817	2.46181	2.33274	2.24376	2.17833	2.12800	2.08798	2.05533	2.02815	1.98539	1.93992	1.89127	1.86556	1.83879	1.81084	1.78156	1.75075	1.71817
17	3.02623	2.64464	2.43743	2.30775	2.21825	2.15239	2.10169	2.06134	2.02839	2.00094	1.95772	1.91169	1.86236	1.83624	1.80901	1.78053	1.75063	1.71909	1.68564
18	3.00698	2.62395	2.41601	2.28577	2.19583	2.12958	2.07854	2.03789	2.00467	1.97698	1.93334	1.88681	1.83685	1.81035	1.78269	1.75371	1.72322	1.69099	1.65671
19	2.98990	2.60561	2.39702	2.26630	2.17596	2.10936	2.05802	2.01710	1.98364	1.95573	1.91170	1.86471	1.81416	1.78731	1.75924	1.72979	1.69876	1.66587	1.63077
20	2.97465	2.58925	2.38009	2.24893	2.15823	2.09132	2.03970	1.99853	1.96485	1.93674	1.89236	1.84494	1.79384	1.76667	1.73822	1.70833	1.67678	1.64326	1.60738
21	2.96096	2.57457	2.36489	2.23334	2.14231	2.07512	2.02325	1.98186	1.94797	1.91967	1.87497	1.82715	1.77555	1.74807	1.71927	1.68896	1.65691	1.62278	1.58615
22	2.94858	2.56131	2.35117	2.21927	2.12794	2.06050	2.00840	1.96680	1.93273	1.90425	1.85925	1.81106	1.75899	1.73122	1.70208	1.67138	1.63885	1.60415	1.56678
23	2.93736	2.54929	2.33873	2.20651	2.11491	2.04723	1.99492	1.95312	1.91888	1.89025	1.84497	1.79643	1.74392	1.71588	1.68643	1.65535	1.62237	1.58711	1.54903
24	2.92712	2.53833	2.32739	2.19488	2.10303	2.03513	1.98263	1.94066	1.90625	1.87748	1.83194	1.78308	1.73015	1.70185	1.67210	1.64067	1.60726	1.57146	1.53270
25	2.91774	2.52831	2.31702	2.18424	2.09216	2.02406	1.97138	1.92925	1.89469	1.86578	1.82000	1.77083	1.71752	1.68898	1.65895	1.62718	1.59335	1.55703	1.51760
26	2.90913	2.51910	2.30749	2.17447	2.08218	2.01389	1.96104	1.91876	1.88407	1.85503	1.80902	1.75957	1.70589	1.67712	1.64682	1.61472	1.58050	1.54368	1.50360
27	2.90119	2.51061	2.29871	2.16546	2.07298	2.00452	1.95151	1.90909	1.87427	1.84511	1.79889	1.74917	1.69514	1.66616	1.63560	1.60320	1.56859	1.53129	1.49057
28	2.89385	2.50276	2.29060	2.15714	2.06447	1.99585	1.94270	1.90014	1.86520	1.83593	1.78951	1.73954	1.68519	1.65600	1.62519	1.59250	1.55753	1.51976	1.47841

29	2.88703	2.49548	2.28307	2.14941	2.05658	1.98781	1.93452	1.89184	1.85679	1.82741	1.78081	1.73060	1.67593	1.64655	1.61551	1.58253	1.54721	1.50899	1.46704
30	2.88069	2.48872	2.27607	2.14223	2.04925	1.98033	1.92692	1.88412	1.84896	1.81949	1.77270	1.72227	1.66731	1.63774	1.60648	1.57323	1.53757	1.49891	1.45636
40	2.83535	2.44037	2.22609	2.09095	1.99682	1.92688	1.87252	1.82886	1.79290	1.76269	1.71456	1.66241	1.60515	1.57411	1.54108	1.50562	1.46716	1.42476	1.37691
60	2.79107	2.39325	2.17741	2.04099	1.94571	1.87472	1.81939	1.77483	1.73802	1.70701	1.65743	1.60337	1.54349	1.51072	1.47554	1.43734	1.39520	1.34757	1.29146
120	2.74781	2.34734	2.12999	1.99230	1.89587	1.82381	1.76748	1.72196	1.68425	1.65238	1.60120	1.54500	1.48207	1.44723	1.40938	1.36760	1.32034	1.26457	1.19256
∞	2.70554	2.30259	2.08380	1.94486	1.84727	1.77411	1.71672	1.67020	1.63152	1.59872	1.54578	1.48714	1.42060	1.38318	1.34187	1.29513	1.23995	1.16860	1.00000

F Table for α = 0.05

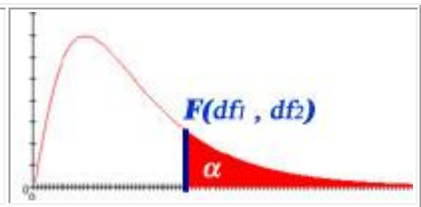


/	**df_1 =1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**12**	**15**	**20**	**24**	**30**	**40**	**60**	**120**	∞
df_2= 1	161.4476	199.5000	215.7073	224.5832	230.1619	233.9860	236.7684	238.8827	240.5433	241.8817	243.9060	245.9499	248.0131	249.0518	250.0951	251.1432	252.1957	253.2529	254.3144
2	18.5128	19.0000	19.1643	19.2468	19.2964	19.3295	19.3532	19.3710	19.3848	19.3959	19.4125	19.4291	19.4458	19.4541	19.4624	19.4707	19.4791	19.4874	19.4957
3	10.1280	9.5521	9.2766	9.1172	9.0135	8.9406	8.8867	8.8452	8.8123	8.7855	8.7446	8.7029	8.6602	8.6385	8.6166	8.5944	8.5720	8.5494	8.5264
4	7.7086	6.9443	6.5914	6.3882	6.2561	6.1631	6.0942	6.0410	5.9988	5.9644	5.9117	5.8578	5.8025	5.7744	5.7459	5.7170	5.6877	5.6581	5.6281
5	6.6079	5.7861	5.4095	5.1922	5.0503	4.9503	4.8759	4.8183	4.7725	4.7351	4.6777	4.6188	4.5581	4.5272	4.4957	4.4638	4.4314	4.3985	4.3650
6	5.9874	5.1433	4.7571	4.5337	4.3874	4.2839	4.2067	4.1468	4.0990	4.0600	3.9999	3.9381	3.8742	3.8415	3.8082	3.7743	3.7398	3.7047	3.6689
7	5.5914	4.7374	4.3468	4.1203	3.9715	3.8660	3.7870	3.7257	3.6767	3.6365	3.5747	3.5107	3.4445	3.4105	3.3758	3.3404	3.3043	3.2674	3.2298

8	5.3177	4.4590	4.0662	3.8379	3.6875	3.5806	3.5005	3.4381	3.3881	3.3472	3.2839	3.2184	3.1503	3.1152	3.0794	3.0428	3.0053	2.9669	2.9276
9	5.1174	4.2565	3.8625	3.6331	3.4817	3.3738	3.2927	3.2296	3.1789	3.1373	3.0729	3.0061	2.9365	2.9005	2.8637	2.8259	2.7872	2.7475	2.7067
10	4.9646	4.1028	3.7083	3.4780	3.3258	3.2172	3.1355	3.0717	3.0204	2.9782	2.9130	2.8450	2.7740	2.7372	2.6996	2.6609	2.6211	2.5801	2.5379
11	4.8443	3.9823	3.5874	3.3567	3.2039	3.0946	3.0123	2.9480	2.8962	2.8536	2.7876	2.7186	2.6464	2.6090	2.5705	2.5309	2.4901	2.4480	2.4045
12	4.7472	3.8853	3.4903	3.2592	3.1059	2.9961	2.9134	2.8486	2.7964	2.7534	2.6866	2.6169	2.5436	2.5055	2.4663	2.4259	2.3842	2.3410	2.2962
13	4.6672	3.8056	3.4105	3.1791	3.0254	2.9153	2.8321	2.7669	2.7144	2.6710	2.6037	2.5331	2.4589	2.4202	2.3803	2.3392	2.2966	2.2524	2.2064
14	4.6001	3.7389	3.3439	3.1122	2.9582	2.8477	2.7642	2.6987	2.6458	2.6022	2.5342	2.4630	2.3879	2.3487	2.3082	2.2664	2.2229	2.1778	2.1307
15	4.5431	3.6823	3.2874	3.0556	2.9013	2.7905	2.7066	2.6408	2.5876	2.5437	2.4753	2.4034	2.3275	2.2878	2.2468	2.2043	2.1601	2.1141	2.0658
16	4.4940	3.6337	3.2389	3.0069	2.8524	2.7413	2.6572	2.5911	2.5377	2.4935	2.4247	2.3522	2.2756	2.2354	2.1938	2.1507	2.1058	2.0589	2.0096
17	4.4513	3.5915	3.1968	2.9647	2.8100	2.6987	2.6143	2.5480	2.4943	2.4499	2.3807	2.3077	2.2304	2.1898	2.1477	2.1040	2.0584	2.0107	1.9604
18	4.4139	3.5546	3.1599	2.9277	2.7729	2.6613	2.5767	2.5102	2.4563	2.4117	2.3421	2.2686	2.1906	2.1497	2.1071	2.0629	2.0166	1.9681	1.9168
19	4.3807	3.5219	3.1274	2.8951	2.7401	2.6283	2.5435	2.4768	2.4227	2.3779	2.3080	2.2341	2.1555	2.1141	2.0712	2.0264	1.9795	1.9302	1.8780
20	4.3512	3.4928	3.0984	2.8661	2.7109	2.5990	2.5140	2.4471	2.3928	2.3479	2.2776	2.2033	2.1242	2.0825	2.0391	1.9938	1.9464	1.8963	1.8432
21	4.3248	3.4668	3.0725	2.8401	2.6848	2.5727	2.4876	2.4205	2.3660	2.3210	2.2504	2.1757	2.0960	2.0540	2.0102	1.9645	1.9165	1.8657	1.8117
22	4.3009	3.4434	3.0491	2.8167	2.6613	2.5491	2.4638	2.3965	2.3419	2.2967	2.2258	2.1508	2.0707	2.0283	1.9842	1.9380	1.8894	1.8380	1.7831
23	4.2793	3.4221	3.0280	2.7955	2.6400	2.5277	2.4422	2.3748	2.3201	2.2747	2.2036	2.1282	2.0476	2.0050	1.9605	1.9139	1.8648	1.8128	1.7570
24	4.2597	3.4028	3.0088	2.7763	2.6207	2.5082	2.4226	2.3551	2.3002	2.2547	2.1834	2.1077	2.0267	1.9838	1.9390	1.8920	1.8424	1.7896	1.7330
25	4.2417	3.3852	2.9912	2.7587	2.6030	2.4904	2.4047	2.3371	2.2821	2.2365	2.1649	2.0889	2.0075	1.9643	1.9192	1.8718	1.8217	1.7684	1.7110
26	4.2252	3.3690	2.9752	2.7426	2.5868	2.4741	2.3883	2.3205	2.2655	2.2197	2.1479	2.0716	1.9898	1.9464	1.9010	1.8533	1.8027	1.7488	1.6906
27	4.2100	3.3541	2.9604	2.7278	2.5719	2.4591	2.3732	2.3053	2.2501	2.2043	2.1323	2.0558	1.9736	1.9299	1.8842	1.8361	1.7851	1.7306	1.6717

28	4.1960	3.3404	2.9467	2.7141	2.5581	2.4453	2.3593	2.2913	2.2360	2.1900	2.1179	2.0411	1.9586	1.9147	1.8687	1.8203	1.7689	1.7138	1.6541
29	4.1830	3.3277	2.9340	2.7014	2.5454	2.4324	2.3463	2.2783	2.2229	2.1768	2.1045	2.0275	1.9446	1.9005	1.8543	1.8055	1.7537	1.6981	1.6376
30	4.1709	3.3158	2.9223	2.6896	2.5336	2.4205	2.3343	2.2662	2.2107	2.1646	2.0921	2.0148	1.9317	1.8874	1.8409	1.7918	1.7396	1.6835	1.6223
40	4.0847	3.2317	2.8387	2.6060	2.4495	2.3359	2.2490	2.1802	2.1240	2.0772	2.0035	1.9245	1.8389	1.7929	1.7444	1.6928	1.6373	1.5766	1.5089
60	4.0012	3.1504	2.7581	2.5252	2.3683	2.2541	2.1665	2.0970	2.0401	1.9926	1.9174	1.8364	1.7480	1.7001	1.6491	1.5943	1.5343	1.4673	1.3893
120	3.9201	3.0718	2.6802	2.4472	2.2899	2.1750	2.0868	2.0164	1.9588	1.9105	1.8337	1.7505	1.6587	1.6084	1.5543	1.4952	1.4290	1.3519	1.2539
∞	3.8415	2.9957	2.6049	2.3719	2.2141	2.0986	2.0096	1.9384	1.8799	1.8307	1.7522	1.6664	1.5705	1.5173	1.4591	1.3940	1.3180	1.2214	1.0000

F Table for α = 0.025

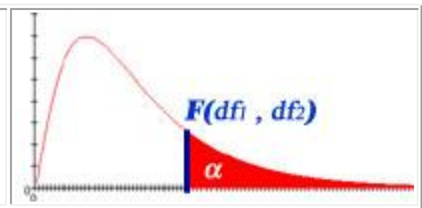


/	df_1=1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
df_2=1	647.7890	799.5000	864.1630	899.5833	921.8479	937.1111	948.2169	956.6562	963.2846	968.6274	976.7079	984.8668	993.1028	997.2492	1001.414	1005.598	1009.800	1014.020	1018.258
2	38.5063	39.0000	39.1655	39.2484	39.2982	39.3315	39.3552	39.3730	39.3869	39.3980	39.4146	39.4313	39.4479	39.4562	39.465	39.473	39.481	39.490	39.498
3	17.4434	16.0441	15.4392	15.1010	14.8848	14.7347	14.6244	14.5399	14.4731	14.4189	14.3366	14.2527	14.1674	14.1241	14.081	14.037	13.992	13.947	13.902
4	12.2179	10.6491	9.9792	9.6045	9.3645	9.1973	9.0741	8.9796	8.9047	8.8439	8.7512	8.6565	8.5599	8.5109	8.461	8.411	8.360	8.309	8.257
5	10.0070	8.4336	7.7636	7.3879	7.1464	6.9777	6.8531	6.7572	6.6811	6.6192	6.5245	6.4277	6.3286	6.2780	6.227	6.175	6.123	6.069	6.015
6	8.8131	7.2599	6.5988	6.2272	5.9876	5.8198	5.6955	5.5996	5.5234	5.4613	5.3662	5.2687	5.1684	5.1172	5.065	5.012	4.959	4.904	4.849
7	8.0727	6.5415	5.8898	5.5226	5.2852	5.1186	4.9949	4.8993	4.8232	4.7611	4.6658	4.5678	4.4667	4.4150	4.362	4.309	4.254	4.199	4.142
8	7.5709	6.0595	5.4160	5.0526	4.8173	4.6517	4.5286	4.4333	4.3572	4.2951	4.1997	4.1012	3.9995	3.9472	3.894	3.840	3.784	3.728	3.670

9	7.2093	5.7147	5.0781	4.7181	4.4844	4.3197	4.1970	4.1020	4.0260	3.9639	3.8682	3.7694	3.6669	3.6142	3.560	3.505	3.449	3.392	3.333
10	6.9367	5.4564	4.8256	4.4683	4.2361	4.0721	3.9498	3.8549	3.7790	3.7168	3.6209	3.5217	3.4185	3.3654	3.311	3.255	3.198	3.140	3.080
11	6.7241	5.2559	4.6300	4.2751	4.0440	3.8807	3.7586	3.6638	3.5879	3.5257	3.4296	3.3299	3.2261	3.1725	3.118	3.061	3.004	2.944	2.883
12	6.5538	5.0959	4.4742	4.1212	3.8911	3.7283	3.6065	3.5118	3.4358	3.3736	3.2773	3.1772	3.0728	3.0187	2.963	2.906	2.848	2.787	2.725
13	6.4143	4.9653	4.3472	3.9959	3.7667	3.6043	3.4827	3.3880	3.3120	3.2497	3.1532	3.0527	2.9477	2.8932	2.837	2.780	2.720	2.659	2.595
14	6.2979	4.8567	4.2417	3.8919	3.6634	3.5014	3.3799	3.2853	3.2093	3.1469	3.0502	2.9493	2.8437	2.7888	2.732	2.674	2.614	2.552	2.487
15	6.1995	4.7650	4.1528	3.8043	3.5764	3.4147	3.2934	3.1987	3.1227	3.0602	2.9633	2.8621	2.7559	2.7006	2.644	2.585	2.524	2.461	2.395
16	6.1151	4.6867	4.0768	3.7294	3.5021	3.3406	3.2194	3.1248	3.0488	2.9862	2.8890	2.7875	2.6808	2.6252	2.568	2.509	2.447	2.383	2.316
17	6.0420	4.6189	4.0112	3.6648	3.4379	3.2767	3.1556	3.0610	2.9849	2.9222	2.8249	2.7230	2.6158	2.5598	2.502	2.442	2.380	2.315	2.247
18	5.9781	4.5597	3.9539	3.6083	3.3820	3.2209	3.0999	3.0053	2.9291	2.8664	2.7689	2.6667	2.5590	2.5027	2.445	2.384	2.321	2.256	2.187
19	5.9216	4.5075	3.9034	3.5587	3.3327	3.1718	3.0509	2.9563	2.8801	2.8172	2.7196	2.6171	2.5089	2.4523	2.394	2.333	2.270	2.203	2.133
20	5.8715	4.4613	3.8587	3.5147	3.2891	3.1283	3.0074	2.9128	2.8365	2.7737	2.6758	2.5731	2.4645	2.4076	2.349	2.287	2.223	2.156	2.085
21	5.8266	4.4199	3.8188	3.4754	3.2501	3.0895	2.9686	2.8740	2.7977	2.7348	2.6368	2.5338	2.4247	2.3675	2.308	2.246	2.182	2.114	2.042
22	5.7863	4.3828	3.7829	3.4401	3.2151	3.0546	2.9338	2.8392	2.7628	2.6998	2.6017	2.4984	2.3890	2.3315	2.272	2.210	2.145	2.076	2.003
23	5.7498	4.3492	3.7505	3.4083	3.1835	3.0232	2.9023	2.8077	2.7313	2.6682	2.5699	2.4665	2.3567	2.2989	2.239	2.176	2.111	2.041	1.968
24	5.7166	4.3187	3.7211	3.3794	3.1548	2.9946	2.8738	2.7791	2.7027	2.6396	2.5411	2.4374	2.3273	2.2693	2.209	2.146	2.080	2.010	1.935
25	5.6864	4.2909	3.6943	3.3530	3.1287	2.9685	2.8478	2.7531	2.6766	2.6135	2.5149	2.4110	2.3005	2.2422	2.182	2.118	2.052	1.981	1.906
26	5.6586	4.2655	3.6697	3.3289	3.1048	2.9447	2.8240	2.7293	2.6528	2.5896	2.4908	2.3867	2.2759	2.2174	2.157	2.093	2.026	1.954	1.878
27	5.6331	4.2421	3.6472	3.3067	3.0828	2.9228	2.8021	2.7074	2.6309	2.5676	2.4688	2.3644	2.2533	2.1946	2.133	2.069	2.002	1.930	1.853
28	5.6096	4.2205	3.6264	3.2863	3.0626	2.9027	2.7820	2.6872	2.6106	2.5473	2.4484	2.3438	2.2324	2.1735	2.112	2.048	1.980	1.907	1.829

29	5.5878	4.2006	3.6072	3.2674	3.0438	2.8840	2.7633	2.6686	2.5919	2.5286	2.4295	2.3248	2.2131	2.1540	2.092	2.028	1.959	1.886	1.807
30	5.5675	4.1821	3.5894	3.2499	3.0265	2.8667	2.7460	2.6513	2.5746	2.5112	2.4120	2.3072	2.1952	2.1359	2.074	2.009	1.940	1.866	1.787
40	5.4239	4.0510	3.4633	3.1261	2.9037	2.7444	2.6238	2.5289	2.4519	2.3882	2.2882	2.1819	2.0677	2.0069	1.943	1.875	1.803	1.724	1.637
60	5.2856	3.9253	3.3425	3.0077	2.7863	2.6274	2.5068	2.4117	2.3344	2.2702	2.1692	2.0613	1.9445	1.8817	1.815	1.744	1.667	1.581	1.482
120	5.1523	3.8046	3.2269	2.8943	2.6740	2.5154	2.3948	2.2994	2.2217	2.1570	2.0548	1.9450	1.8249	1.7597	1.690	1.614	1.530	1.433	1.310
∞	5.0239	3.6889	3.1161	2.7858	2.5665	2.4082	2.2875	2.1918	2.1136	2.0483	1.9447	1.8326	1.7085	1.6402	1.566	1.484	1.388	1.268	1.000

F Table for α = 0.01



/	**df1 =1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**12**	**15**	**20**	**24**	**30**	**40**	**60**	**120**	∞
df2= 1	4052.181	4999.500	5403.352	5624.583	5763.650	5858.986	5928.356	5981.070	6022.473	6055.847	6106.321	6157.285	6208.730	6234.631	6260.649	6286.782	6313.030	6339.391	6365.864
2	98.503	99.000	99.166	99.249	99.299	99.333	99.356	99.374	99.388	99.399	99.416	99.433	99.449	99.458	99.466	99.474	99.482	99.491	99.499
3	34.116	30.817	29.457	28.710	28.237	27.911	27.672	27.489	27.345	27.229	27.052	26.872	26.690	26.598	26.505	26.411	26.316	26.221	26.125
4	21.198	18.000	16.694	15.977	15.522	15.207	14.976	14.799	14.659	14.546	14.374	14.198	14.020	13.929	13.838	13.745	13.652	13.558	13.463
5	16.258	13.274	12.060	11.392	10.967	10.672	10.456	10.289	10.158	10.051	9.888	9.722	9.553	9.466	9.379	9.291	9.202	9.112	9.020
6	13.745	10.925	9.780	9.148	8.746	8.466	8.260	8.102	7.976	7.874	7.718	7.559	7.396	7.313	7.229	7.143	7.057	6.969	6.880
7	12.246	9.547	8.451	7.847	7.460	7.191	6.993	6.840	6.719	6.620	6.469	6.314	6.155	6.074	5.992	5.908	5.824	5.737	5.650
8	11.259	8.649	7.591	7.006	6.632	6.371	6.178	6.029	5.911	5.814	5.667	5.515	5.359	5.279	5.198	5.116	5.032	4.946	4.859
9	10.561	8.022	6.992	6.422	6.057	5.802	5.613	5.467	5.351	5.257	5.111	4.962	4.808	4.729	4.649	4.567	4.483	4.398	4.311
10	10.044	7.559	6.552	5.994	5.636	5.386	5.200	5.057	4.942	4.849	4.706	4.558	4.405	4.327	4.247	4.165	4.082	3.996	3.909

11	9.646	7.206	6.217	5.668	5.316	5.069	4.886	4.744	4.632	4.539	4.397	4.251	4.099	4.021	3.941	3.860	3.776	3.690	3.602
12	9.330	6.927	5.953	5.412	5.064	4.821	4.640	4.499	4.388	4.296	4.155	4.010	3.858	3.780	3.701	3.619	3.535	3.449	3.361
13	9.074	6.701	5.739	5.205	4.862	4.620	4.441	4.302	4.191	4.100	3.960	3.815	3.665	3.587	3.507	3.425	3.341	3.255	3.165
14	8.862	6.515	5.564	5.035	4.695	4.456	4.278	4.140	4.030	3.939	3.800	3.656	3.505	3.427	3.348	3.266	3.181	3.094	3.004
15	8.683	6.359	5.417	4.893	4.556	4.318	4.142	4.004	3.895	3.805	3.666	3.522	3.372	3.294	3.214	3.132	3.047	2.959	2.868
16	8.531	6.226	5.292	4.773	4.437	4.202	4.026	3.890	3.780	3.691	3.553	3.409	3.259	3.181	3.101	3.018	2.933	2.845	2.753
17	8.400	6.112	5.185	4.669	4.336	4.102	3.927	3.791	3.682	3.593	3.455	3.312	3.162	3.084	3.003	2.920	2.835	2.746	2.653
18	8.285	6.013	5.092	4.579	4.248	4.015	3.841	3.705	3.597	3.508	3.371	3.227	3.077	2.999	2.919	2.835	2.749	2.660	2.566
19	8.185	5.926	5.010	4.500	4.171	3.939	3.765	3.631	3.523	3.434	3.297	3.153	3.003	2.925	2.844	2.761	2.674	2.584	2.489
20	8.096	5.849	4.938	4.431	4.103	3.871	3.699	3.564	3.457	3.368	3.231	3.088	2.938	2.859	2.778	2.695	2.608	2.517	2.421
21	8.017	5.780	4.874	4.369	4.042	3.812	3.640	3.506	3.398	3.310	3.173	3.030	2.880	2.801	2.720	2.636	2.548	2.457	2.360
22	7.945	5.719	4.817	4.313	3.988	3.758	3.587	3.453	3.346	3.258	3.121	2.978	2.827	2.749	2.667	2.583	2.495	2.403	2.305
23	7.881	5.664	4.765	4.264	3.939	3.710	3.539	3.406	3.299	3.211	3.074	2.931	2.781	2.702	2.620	2.535	2.447	2.354	2.256
24	7.823	5.614	4.718	4.218	3.895	3.667	3.496	3.363	3.256	3.168	3.032	2.889	2.738	2.659	2.577	2.492	2.403	2.310	2.211
25	7.770	5.568	4.675	4.177	3.855	3.627	3.457	3.324	3.217	3.129	2.993	2.850	2.699	2.620	2.538	2.453	2.364	2.270	2.169
26	7.721	5.526	4.637	4.140	3.818	3.591	3.421	3.288	3.182	3.094	2.958	2.815	2.664	2.585	2.503	2.417	2.327	2.233	2.131
27	7.677	5.488	4.601	4.106	3.785	3.558	3.388	3.256	3.149	3.062	2.926	2.783	2.632	2.552	2.470	2.384	2.294	2.198	2.097
28	7.636	5.453	4.568	4.074	3.754	3.528	3.358	3.226	3.120	3.032	2.896	2.753	2.602	2.522	2.440	2.354	2.263	2.167	2.064
29	7.598	5.420	4.538	4.045	3.725	3.499	3.330	3.198	3.092	3.005	2.868	2.726	2.574	2.495	2.412	2.325	2.234	2.138	2.034
30	7.562	5.390	4.510	4.018	3.699	3.473	3.304	3.173	3.067	2.979	2.843	2.700	2.549	2.469	2.386	2.299	2.208	2.111	2.006

40	7.314	5.179	4.313	3.828	3.514	3.291	3.124	2.993	2.888	2.801	2.665	2.522	2.369	2.288	2.203	2.114	2.019	1.917	1.805
60	7.077	4.977	4.126	3.649	3.339	3.119	2.953	2.823	2.718	2.632	2.496	2.352	2.198	2.115	2.028	1.936	1.836	1.726	1.601
120	6.851	4.787	3.949	3.480	3.174	2.956	2.792	2.663	2.559	2.472	2.336	2.192	2.035	1.950	1.860	1.763	1.656	1.533	1.381
∞	6.635	4.605	3.782	3.319	3.017	2.802	2.639	2.511	2.407	2.321	2.185	2.039	1.878	1.791	1.696	1.592	1.473	1.325	1.000

5 GW

مثال: لحساب معنوية نموذج الانحدار بالإشارة للمثال السابق والذي يحدد علاقة الرطوبة النسبية في المدينة A و B: فاختبر معنوية نموذج الانحدار مع كتابة جدول تحليل التباين للانحدار ومستوى الدلالة $\alpha = .01$

الحل

$$SST = \sum Y_i^2 - n\bar{Y}^2 \rightarrow \sum (Y - \bar{Y})^2$$
$$= 49836 - 9 \times \left(\frac{654}{9}\right)^2$$
$$= 49836 - 47524$$
$$= 2312$$
$$SSR = b[\sum X_i Y_i - n\bar{X}\bar{Y}] \longrightarrow \sum (\hat{Y} - \bar{Y})^2$$
$$= 1.242 \times 1736.189$$
$$= 2156.35$$
$$SSE = SST - SSR \longrightarrow \sum (Y - \hat{Y})^2$$
$$= 2312 - 2156.35$$
$$= 155.65$$
$$S_{YX}^2 = \frac{SSE}{n-2} = \frac{155.65}{7}$$
$$= 22.236$$
$$MS_x = \frac{SSR}{1} = 2156.35$$
$$f = \frac{MS_x}{S_{YX}^2} = \frac{2156.35}{22.236}$$
$$= 96.98$$

∴ القيمة الجدولية لاختبار F عند

$F(1-\alpha, 1, n-2)$

تساوي $= F(.99, 1, 7) = 12.2$

وحيث أن القيمة المحسوبة أكبر من الجدولية

$96.98 > 12.2$

إذاً نستطيع القول وبدرجة ثقة 99% أن نموذج خط الانحدار ذو معنوية ويمكن الاعتماد عليه في تقدير قيم Y عند معرفة قيم X.

* جميع القياسات السابقة وهي معامل التحديد وحساب الخطأ في التقدير وحساب تباين الانحدار تعتبر من مقاييس جودة خط الانحدار ونموذج الانحدار وللتأكد أكثر من جودة النموذج Goodness of fit يمكن حساب معامل

جدول تحليل التباين للانحدار:

مصدر التغير	مجموع المربعات	درجات الحرية	معدل المربعات	ƒ المحسوبة
بسبب الانحدار SSR	2156.35	1	2156.35	96.98
الانحرافات عن الانحدار SSE	155.65	7	22.236	
الكلي SST	2312	8		

التباين C.V لنموذج الانحدار على النحو التالي

$C.V = S_{x.y} / \bar{Y} * 100 = \frac{4.7}{72.67} \times 100 = 6.48\%$

وهنا كلما صغرت القيمة كلما أشار ذلك إلى جودة خط الانحدار لتمثيل العلاقة بين X و Y أو بعبارة أخرى جودة النموذج.

* الدلالة الإحصائية لمعامل الانحدار b هنا نستخدم t-test على النحو التالي

$t = b / SE_b$

مع العلم أن SE_b تحسب بالمعادلة التالية

$SE_b = \sqrt{(SSE / n-k) / \sum (X - \bar{X})^2}$

وبعد الحصول على النتيجة نقارنها مع t الجدولية عند $t(\alpha/2, V)$ وهي هنا حسب درجة الثقة.

b corr
Key

* مثال لحساب الدلالة الإحصائية لمعامل الانحدار b.
في مثالنا السابق أحسب معنوية معامل الانحدار b عند مستوى معنوية $\alpha = .05$

الحل: نحسب t ثم نقارنها بالقيمة الجدولية وذلك باتباع الخطوات التالية:

أولاً: نحسب Sxx (وهو في الواقع $= \Sigma(x-\bar{x})^2$)

ويمكن حسابه على النحو التالي $Sxx = \Sigma x^2 - \frac{(\Sigma x)^2}{n}$

وبالرجوع إلى جدول مثال الارتباط نجد أن

$$\Sigma x^2 = 53078 \quad , \quad (\Sigma x)^2 = 465124 \quad , \quad n = 9$$

وبتطبيق المعادلة أعلاه نجد أن

$$Sxx = 53078 - \frac{465124}{9} = 53078 - 51680.4 = \underline{\underline{1397.6}}$$

وبذلك نستطيع حساب SEb على النحو التالي:

$$SEb = \sqrt{\frac{SSE}{n-2} / Sxx} = \sqrt{\frac{154.8}{7} / 1397.6} = \sqrt{22.1 / 1397.6}$$
$$= \sqrt{.0158} = \underline{\underline{.125}}$$

وأخيراً نستطيع حساب قيمة t على النحو التالي

$$t = b/SEb = \frac{1.24}{.125} = \underline{\underline{9.8}}$$

و... قيمة t الجدولية عند $t(\alpha/2, v)$ هي

$$t = (.025, 7) = 2.36$$

وحيث أن القيمة المحسوبة 9.8 أكبر من الجدولية فإن معامل الانحدار b معنوي ويمكن الاعتماد عليه واستخدامه في المعادلة.

7 corr

Reg

Multiple Regression

*** الانحدار المتعدد**

معادلة الانحدار المتعدد: $\hat{Y} = a + b_1X_1 + b_2X_2 \cdots\cdots b_nX_n$

حيث: a تمثل التقاطع و $b_1 \dots b_n$ تمثل معاملات الانحدار الجزئية (Partial Regression Coefficients)

و $X_1 \dots X_n$ هي المتغيرات المستقلة. وحساب التقاطع ومعاملات الانحدار يعتمد أيضا على طريقة المربعات الصغرى Least Squares حيث تكون $\sum(Y-\hat{Y})^2$ صغرى. حساب المعاملات $b_1 \dots b_n$ والتقاطع في الانحدار البسيط حيث لدينا متغير مستقل واحد سهل وسبقه شرحه. أما الانحدار المتعدد وهنالك ثلاث متغيرات أو أكثر تتفاعل مع بعضها لذلك نحتاج لحل العديد من المعادلات لايجاد المجاهيل a و $b_1 \dots b_n$ ويعتمد عدد هذه المعادلات على عدد المتغيرات المستقلة $X_1 \dots X_n$.

مثلا في حالة وجود متغيرين مستقلين هي X_1, X_2 نحتاج حل ثلاث معادلات هي

$$1- \sum Y = an + b_1\sum X_1 + b_2\sum X_2$$

$$2- \sum X_1Y = a\sum X_1 + b_1\sum X_1^2 + b_2\sum X_2X_1$$

$$3- \sum X_2Y = a\sum X_2 + b_1\sum X_1X_2 + b_2\sum X_2^2$$

وهذه ثلاث معادلات فيها ثلاث مجاهيل هي a و b_1 و b_2 ويمكن حلها بعدة طرق نجدها في كتب الجبر مثل طريقة الحذف والتعويض وطريقة المصفوفات البسيطة. وليس من الضروري بالنسبة لنا أن نطبق أحد تلك الطرق ونتعلمها بل نستخدم الحاسب الآلي لعمل ذلك. إلا أنه من الضروري جدا أن نفهم طريقة إيجاد المجاميع الضرورية لبناء تلك المعادلات والتي تمكننا في النهاية من حل المعادلات بالطرق السابقة الذكر ومن ثم إيجاد المجاهيل a و b_1 و b_2 وبالتالي بناء معادلة هي معادلة الانحدار المتعدد التي تمكننا من توقع قيم Y استنادا على معرفة قيم المتغيرات المستقلة. والمجاميع التي نحتاجها لبناء المعادلات الثلاث هي:

$$\sum X_1, \sum X_2, \sum Y, \sum X_1^2, \sum X_2^2, \sum X_1Y, \sum X_2Y, \sum X_1X_2$$

مثال للانحدار المتعدد

أراد باحث جغرافي أن يحدد العلاقة بين سعر المنزل وعمره ومساحته في أحد المدن فقام بأخذ عينة لعشرين منزل وحصل على المعلومات التالية في الجدول:

Y	X_2	X_1
32	30	1.8
24	33	1.0
27	25	1.7
47	12	2.8
35	26	2.2
17	25	0.8
52	28	3.6
20	29	1.1
38	25	2.0
45	2	2.6
44	30	2.3
19	23	0.9
25	12	1.2
50	33	3.4
30	1	1.7
43	12	2.5
27	17	1.4
50	16	3.3
37	22	2.2
28	29	1.5
690	430	40

وفي هذا الجدول:

Y = سعر المنزل بآلاف الريالات

X_1 = مساحة المنزل بالآف الأمتار

X_2 = عمر المنزل بالسنة

الحل من أجل إيجاد a, b_1, b_2 نحتاج إلى حل المعادلات الثلاث السابقة الذكر ولكن هذا يتطلب إيجاد:

$\Sigma X_1, \Sigma X_2, \Sigma Y, \Sigma X_1^2, \Sigma X_2^2, \Sigma X_1Y, \Sigma X_2Y, \Sigma X_1X_2$

لذلك ننشئ الجدول التالي:

X_1^2	X_2^2	X_1X_2	X_1Y	X_2Y	
3.24	900	54.0	57.6	960	
1.00	1089	33.0	24.0	792	
2.89	625	42.5	45.9	675	
7.84	144	33.6	131.6	564	
4.84	676	57.2	77.0	910	
0.64	625	20.0	13.6	425	
12.96	784	100.8	187.3	1456	
1.21	841	31.9	22.0	580	
4.00	625	50.0	76.0	950	
6.76	4	5.2	117.0	90	
5.29	900	69.0	101.2	1320	
0.81	529	20.7	17.1	437	
1.44	144	14.4	30.3	300	
11.56	1089	112.2	170.0	1650	
2.89	1	1.7	51.0	30	
6.25	144	30.0	107.5	516	
1.96	289	23.8	37.8	459	
10.89	256	52.8	165.0	800	
4.84	484	48.4	81.4	814	
2.25	841	43.5	42.0	812	
93.56	10990	844.7	1554.9	14540	المجموع

نتائج هذا الجدول تمكننا من حل المعادلات الثلاث السابقة الذكر والحصول على a, b_1, b_2 على النحو التالي:

$$690 = 20a + 40b_1 + 430b_2$$
$$1554.9 = 40a + 93.56b_1 + 844.7b_2$$
$$14540 = 430a + 844.7b_1 + 10990b_2$$

وحل تلك المعادلات يعطي

$a = 10.05$

$b_1 = 12.83$

$b_2 = -0.06$

9 Reg

• هذا أخيراً يمكننا كتابة معادلة خط الانحدار المتعدد على النحو التالي

$$\hat{Y} = 10.05 + 12.83X_1 + (-0.06X_2)$$

• هذا يعني أن سعر المنزل في تلك المدينة يساوي 10.05 ألف ريال زائداً مقدارين هما أ- مساحة المنزل مضروبة في 12.83 ألف ريال، ب- عمر المنزل مضروب في -0.06 ألف ريال

فمثلاً إذا أراد شخص شراء منزل في المنطقة مساحته الفين متر مربع وعمره 8 سنوات فإن سعره يقدر بحل المعادلة ←

$$\hat{Y} = 10.05 + 12.83(2) + (-0.06(8))$$
$$= 10.05 + 25.66 + (-.48) = 10.05 + 25.18 = \underline{35.23}$$ الف ريال

• وبهذا يستطيع هذا الشخص إعطاء سعر مناسب للمنزل الذي يريد أن يشتريه.

⊗ تباين انحدار Y على X_1 و X_2 هو $S^2_{Y.X} = \frac{\sum(Y-\hat{Y})^2}{n-k}$

و k هنا = 3 والجذر التربيعي لهذه القيمة هو الخطأ المعياري للتقدير Standard Error of the Estimate

وحيث أن $\sum(Y-\hat{Y})^2$ هنا تساوي 111.4 ،

درجة الحرية = $n-k = 17 = 20-3$ إذن

$$S_{Y.X_1X_2} = \sqrt{\frac{111.4}{17}} = \sqrt{6.55} = 2.56$$

لذلك يكون تقديرنا للسعر في المثال السابق:

35.23 ± 2.56 ألف ريال

فيكون الحد الأدنى للتقدير هو 35.23 − 2.56 = 32.7 ألف ريال

والحد الأعلى هو 35.23 + 2.56 = 37.8 ألف ريال

ملاحظات 1- يمكن حساب معنوية النموذج عن طريق تحليل التباين وهذا يحتاج حساب P ← $\sum(Y_i-\bar{Y})^2$ وهو SST

ب ← $\sum(\hat{Y}-\bar{Y})^2$ وهو SSR

ج ← $\sum(Y-\hat{Y})^2$ وهو SSE

2- يمكن حساب معنوية معامل الانحدار عن طريق اختبار t وهي $t = b/S_{Eb}$

3- يمكن حساب معامل التحديد بنفس الطريقة السابقة الذكر في النقاط المثال السابق توضح جودة ومعنوية النموذج.

15 Reg Corr

⊛ تحديد المتغيرات المهمة في الانحدار المتعدد.
في المثال السابق هل عمر المنزل ومساحته لهما نفس الأهمية في تحديد سعر المنزل؟ للتأكد من ذلك دعونا نسرد نتائج تحليل الانحدار المتعدد التالية:

(1)

$r_{x_1y} = .975$, $r^2_{x_1y} = .950$ و $r_{x_2y} = -.145$, $r^2_{x_2y} = .02$

هذا يشير إلى أن مساحة المنزل تفسر 95٪ من التغير في السعر وفقط 2٪ من التغير في السعر يعزى لعمر المنزل

(2)

$b_1 = 12.8$, $b_2 = -.06$

وهذا يعني زيادة مساحة المنزل ألف متر تعني زيادة سعره 12.8 ألف ريال وأن زيادة عمر المنزل سنة واحدة تعني نقص سعر المنزل فقط 0.06 ألف ريال = 60 ريال

(3) حساب معنوية معاملات الانحدار ومعاملات الارتباط يوضح أن b_2 ليس معنوي و أن r_{x_2y} ليس معنوي.

⊛ النتائج الواردة في النقاط الثلاث السابقة توضح أن العامل الأهم في تحديد سعر المنزل هو مساحته أما عمره فلا يؤثر كثيراً في سعره.

⊛ تحليل التباين في الانحدار المتعدد.
لا يختلف عن الانحدار البسيط إلا في درجات الحرية لأننا نتعامل مع أكثر من متغير مستقل. لذلك يكون تحليل التباين في مثالنا السابق على النحو التالي: (عند مستوى معنوية $\alpha = .05$)

جدول تحليل التباين للانحدار المتعدد مع نتائج المثال السابق

مصدر التغير Source of Variation	مجموع المربعات SS	درجات الحرية Df	معدل المربعات SSM	f المحسوبة
بسبب الانحدار SSR	$\Sigma(\hat{Y}-\bar{Y})^2 =$ 2261.4	$K-1$ $3-1=2$	$MS_x = \frac{SSR}{2} = 1130.7$	$\frac{MS_x}{S^2_{Y\cdot X}}$
بسبب الخطأ SSE	$\Sigma(Y-\hat{Y})^2 =$ 111.6	$n-R$ $20-3=17$	$S^2_{Y\cdot X} = \frac{SSE}{17} = 6.6$	172.4
SST	$\Sigma(Y-\bar{Y})^2 =$ 12373	$n-1 =$ 19		

وحيث أن القيمة الجدولية = $F(.95, 2, 17) = 3.29$ ← $F(1-\alpha, V_1, V_2)$
فإن ذلك يشير إلى أن القيمة المحسوبة أكبر بشكل واضح من القيمة الجدولية فإن النموذج يكون ذو معنوية عالية ويمكن الاعتماد عليه في تقدير سعر المنزل.

1

ANOVA

* تحليل التباين يوضح درجة التشابه والاختلاف في ~~[illegible]~~ معالم مجتمعات إحصائية عديدة لظاهرة معينة مثل إنتاج 5 مزارع نخيل نستخدم 5 أنواع من الأسمدة، متوسط شفاء 4 مجموعات من المرضى يستخدمون 4 أنواع مختلفة من العقاقير، تفاوت متوسطات درجة الحرارة في 8 محطات مناخية في إقليم واحد ---- إلخ.
* الاختبار المستخدم في قياس درجة التشابه والاختلاف في معالم تلك المجتمعات الإحصائية المختلفة هو إختبار F.
* تحليل التباين أنواع عديدة حسب طبيعة المدروسة وعدد المعاملات Treatments وسوف نقتصر هنا على تحليل التباين في المتوسطات ودراسة الاختلاف والتشابه في خاصية واحدة المسمى One-Way ANOVA والتشابه والاختلاف في خاصيتين المسمى Tow-Way ANOVA.

One-Way ANOVA

1. تحليل التباين الأحادي

وهذا أبسط أنواع تحليل التباين، وهو عبارة عن دراسة خاصية واحدة (One Treatment) لعدد من العينات أو المجتمعات إحصائية تمثل ظاهرة معينة وسوف نسمي كل مجموعة أو كل مجتمع إحصائي معاملة k ونفترض أن تكون مستقلة عن بعضها البعض وذات توزيع طبيعي Indipendent and Normal D. ولها متوسطات مختلفة $M_1, M_2, M_3 \dots M_k$ والانحراف التباين. وللتحقق من تشابه واختلاف المتوسطات الحسابية لـ W المجتمعات الإحصائية نأخذ عينات كل عينة تمثل مجموعة، حجمها $= n$ ونضع الفرضيتين

2

ANOVA

الأولى وهي فرضية للباحث والتي مفادها

H_1: يوجد على الأقل معدلين M غير متساويين من تلك المجموعات

حيث تقول هذه الفرضية أن هنالك اختلافات معنوية بين تلك المجموعات لذلك هنالك اختلافات ذات معنوية بين متوسطاتها

الثانية هي فرضية العدم وهي $H_0: M_1 = M_2 = M_3 = \dots M_n$

ومفادها أنه ليس هنالك اختلاف بين متوسطات المجموعات.

* تحليل التباين يفسر مصدر التباين في مجموعات ذلك المتغير أي مصدرين

1- التباين بين المتوسطات لكل مجموعة وهو يسمى التباين بين المجموعات Between groups Variation

وهو يحسب بواسطة حساب مجموع مربعات المعاملات والذي يسمى Treatment Sum of Squares SST

وهو $\sum^{k} n(\bar{Y}_i - \bar{Y})^2$ حيث تمثل n عدد المشاهدات في كل مجموعة و $\bar{Y}_i$ متوسط المجموعة و $\bar{Y}$ متوسط كل المشاهدات في كل المجموعات ويسمى $\bar{Y}$ المتوسط الأعظم

ونشرح $\sum n(\bar{Y}_i - \bar{Y})^2$ هو التالي

ا- نطرح الوسط الحسابي الأعظم من متوسط كل مجموعة ونربعه

ب- نضرب الناتج في عدد n لكل مجموعة

ج- نجمع الناتج ونحصل على SST

2 المصدر الثاني للتباين هو التباين في n لكل المجموعات مجتمعة وهو. $\sum^{k}\sum^{n_i}(Y_{ij} - \bar{Y}_i)^2$

وهو يسمى Within groups Variation ويتم حسابه بواسطة حساب كل مربعات الأخطاء والذي يسمى Error Sum of Squares أو ESS ونقصد هنا بالخطأ هو الفرق بين كل مشاهدة والمتوسط الذي يمثلها و ESS يسمى أيضا مجموع المربعات داخل المجموعات أو Within groups Sum of Squares

وطريقة حسابه والتي تشرح $\sum^{k}\sum^{n_i}(Y_{ij} - \bar{Y}_i)^2$ هي

أ- نطرح الوسط الحسابي للمجموعة من كل مشاهدة في تلك المجموعة بعد ذلك نربع الناتج

ب- نطبق النقطة السابقة لكل المجموعات

ج نجمع الناتج في الخطوة أ و ب فيعطينا ESS.

3 ANOVA

ونشرح الفقرة السابقة نأخذ جدول بيانات المعلومات لتحليل التباين مع مثال:

المجموعات groups. k	المشاهدات Observations	متوسط المجموعات Mean	مجموع المربعات Sum of Squares
Treatment 1	$y_{11}, y_{12}, \dots, y_{1n_1}$	$\bar{y}_1$	$\sum_{j=1}^{n_1}(y_{1j}-\bar{y}_1)^2$
Treatment 2	$y_{21}, y_{22}, \dots, y_{2n_2}$	$\bar{y}_2$	$\sum_{j=1}^{n_2}(y_{2j}-\bar{y}_2)^2$
$\vdots$	$\vdots$	$\vdots$	$\vdots$
Treatment k	$y_{k1}, y_{k2}, \dots, y_{kn_k}$	$\bar{y}_k$	$\sum_{j=1}^{n_k}(y_{kj}-\bar{y}_k)^2$

المتوسط العام وهذا الجدول هو متوسط جميع n لكل المجموعات على عدد n لكل المجموعات ويسمى grand Mean

مثال: شخص أراد أن يقيس الاختلاف والتشابه بين متوسط درجة الحرارة لشهر يناير في منطقة شمالية من المملكة، أخذ عينات تمثل أيام في شهر يناير لأربع محطات مناخية هي A, B, C, D حيث أن n لكل مجموعة هي كالآتي:

$A=5,\ B=4,\ C=7,\ D=6$ يوم

المطلوب هو اختبار الفرضية

$H_0: M_1=M_2=M_3=M_4$

وهي فرضية العدم. مقابل: يوجد على الأقل معدلين غير متساويين H_1:

عند مستوى معنوية $\alpha = .05$ أي أن 0.05 أن الاختلافات بين المتوسطات معنوية

الحل نضع المعلومات التي حصلنا عليها في جدول مماثل للجدول أعلاه للحصول على الجامع المطلوب. ثم بعد ذلك نحسب كل من SST و SSE و TSS

المحطة	Observations	Mean	Sum of Squares
A	10, 15, 8, 12, 15	$\bar{y}_1 = 12$	$\sum_{j=1}^{5}(y_{1j}-\bar{y}_1)^2 = 38$
B	14, 18, 21, 15	$\bar{y}_2 = 17$	$\sum_{j=1}^{4}(y_{2j}-\bar{y}_2)^2 = 30$
C	17, 16, 14, 15, 17, 15, 18	$\bar{y}_3 = 16$	$\sum_{j=1}^{7}(y_{3j}-\bar{y}_3)^2 = 12$
D	12, 15, 17, 15, 16, 15	$\bar{y}_4 = 15$	$\sum_{j=1}^{6}(y_{4j}-\bar{y}_4)^2 = 14$
		Grand mean $\bar{y} = 15$	

أولاً نحسب SST وهو

$$\sum_{i=1}^{k} n(\bar{Y}_i-\bar{Y})^2$$

ويمكن إيجاده كما يلي:

مثلاً $12-15=-3$ على التوالي: $n(-3)^2+n(2)^2+n(1)^2+n(0)^2$

$5(-3)^2+4(2)^2+7(1)^2+6(0)^2=(45)+(16)+(7)+0=\underline{\underline{68}}$

ثانياً نحسب SSE وهو $\sum\sum(Y_{ij}-\bar{Y}_i)^2$ لذلك نحسبه على التوالي

$(10-12)^2+(15-12)^2+(8-12)^2+\dots+(14-17)^2+\dots+(16-15)^2+(15-15)^2$

$(-2)^2+(3)^2+(-4)^2\dots+(1)^2+(0)^2=\underline{94}$

4 ANOVA

أما مجموع المربعات وهو TSS فيمكن حسابه على النحو
التالي: $\sum^{k}\sum^{n_i}(Y_{ii}-\bar{Y})^2$ وهو في الواقع مجموع مربعات
الفرق بين كل مشاهدة في كل المجموعات والوسط الحسابي الأعظم
الممثل لكل المجموعات لذلك نجد أن TSS في مثالنا يحسب كالتالي
على النحو التالي:

$$(10-15)^2+(15-15)^2+\dots+(16-15)^2+(15-15)^2$$
$$(-5)^2+(0)^2+\dots+(1)^2+(0)^2=162$$

وإذا تذكرنا أن $TSS=SST+ESS$ لذلك يمكن حساب TSS
بجمع SST مع ESS دون الحاجة لإجراء حسابات مطولة
وفي مثالنا $TSS=68+94=162$

* ولكي نختبر الفرضيتين Ho و H1 نضع النتائج السابقة في جدول تحليل
التباين؛ مع التذكير أن فرضية الباحث تقول أن هناك
فروقات معنوية والفرضية البديلة Ho تقول أن المتوسطات
للمجموعات متساوية

(جدول تحليل التباين للمتوسطات للتباين الأحادي) وللتذكير الحاصبة بالدروس هنا هي متوسطات درجة الحرارة لشهور يناير:

Source of Variation مصدر التغير	مجموع المربعات SS	درجة الحرية D.F.	متوسط المربعات Mean square	F
Treatment SST	$\sum^{k} n(\bar{Y}_i-\bar{Y})^2$	$k-1$	$MS_T=\frac{SS_T}{k-1}$	$\frac{MST}{MSE}$
Error SSE	$\sum^{k}\sum^{n_i}(Y_{ii}-\bar{Y}_i)^2$	$\sum^{k} n-k$	$MSE=\frac{SSE}{\sum^{k} n-k}$	
Total TSS	$\sum^{k}\sum^{n_i}(Y_{ii}-\bar{Y})^2$			

وفي مثالنا نتائج الجدول على النحو التالي:

Source of Variation	مجموع المربعات Sum of Squares	درجة الحرية DF	المتوسط MSS	F
SS_T	68	3	22.7	4.4
SSE	94	18	5.2	
SST	162	21		

وعند النظر في جدول F عند
$F(1-\alpha, V_1, V_2)$
$F(95,3,18)=$
$=3.10$
وهذه القيمة أقل من القيمة المحسوبة وهذا يشير إلى
أن هناك فروقات ذات معنوية بين متوسطات المجموعات الأربع
أي نقبل الفرضية البديلة (الباحث) ونرفض فرضية العدم Ho.

5 ANOVA

طريقة أخرى لحساب مجموع المربعات في تحليل التباين:

في المحاضرات السابقة شرحنا طريقة حساب مجموع المربعات بشكل مفصل ولكن هناك طريقة أسهل لإيجاد مجموع تلك المربعات وهذه تتمثل في المعادلات التالية:

$$① \; TSS = \sum^{k}\sum^{n_i} y_{ii}^2 - \frac{T_g^2}{n_g}$$

T هنا هي مجموع كل n أو المجموع الأعظم grand Total

لذلك في مثالنا السابق $\sum^{n}\sum^{n_i} y_{ii}^2 = (10)^2 + (15)^2 - - - - - + (16)^2 + (15)^2 = \underline{5112}$

المجموع الأعظم $T_g = 10 + 15 + 8 + - - - + 14 + 18 + - - - - + 16 + 15 = \underline{\underline{330}}$

ثم نطبق المعادلة $TSS = 5112 - \frac{(330)^2}{22} = 5112 - 4950 = \underline{\underline{162}}$

$$② \; SS_T = \sum^{k} \frac{T_i^2}{n} - \frac{T_g^2}{n_g}$$

وهنا T_i مجموع كل مجموعة. لذلك في مثالنا السابق

$$\sum^{k} \frac{T_i^2}{n} = \frac{(60)^2}{5} + \frac{(68)^2}{4} + \frac{(112)^2}{7} + \frac{(90)^2}{6} = 5018$$

ونطبق القانون $SST = 5018 - 4950 = \underline{\underline{68}}$

وأخيرا $③ \; SSE = TSS - SS_T = 162 - 68 = \underline{94}$

ثم نستكمل الحل كما ورد في المحاضرة السابقة للتأكد من الفرضيات H1 و H0.