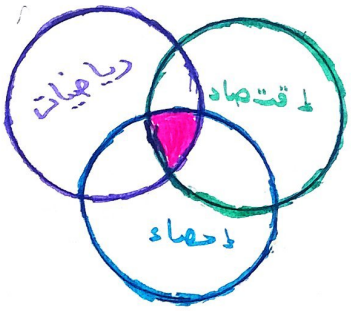
1- تعريف الإقتصاد القياسي :

- يعرف الإقتصاد القياسي أو ما يسمى بالقياس الإقتصادي بأنه العلم الذي يهتم بقياس العلاقات الإقتصادية من خلال بيانات واقعية بغرض إختبار مدى صحة هذه العلاقات كما تقدمها النظرية الإقتصادية أو تفسير بعض الظواهر أو رسم بعض السياسات أو التنبؤ بسلوك بعض المتغيرات الإقتصادية .



الإقتصاد القياسي

- نأخذ منه الجانب النظري للظاهرة المدروسة "النظرية الإقتصادية"
- نأخذ منه المعادلات أي الشكل الرياضي .
- من خلاله يتم جمع البيانات ، تقدير النموذج ومن ثم التنبؤ .

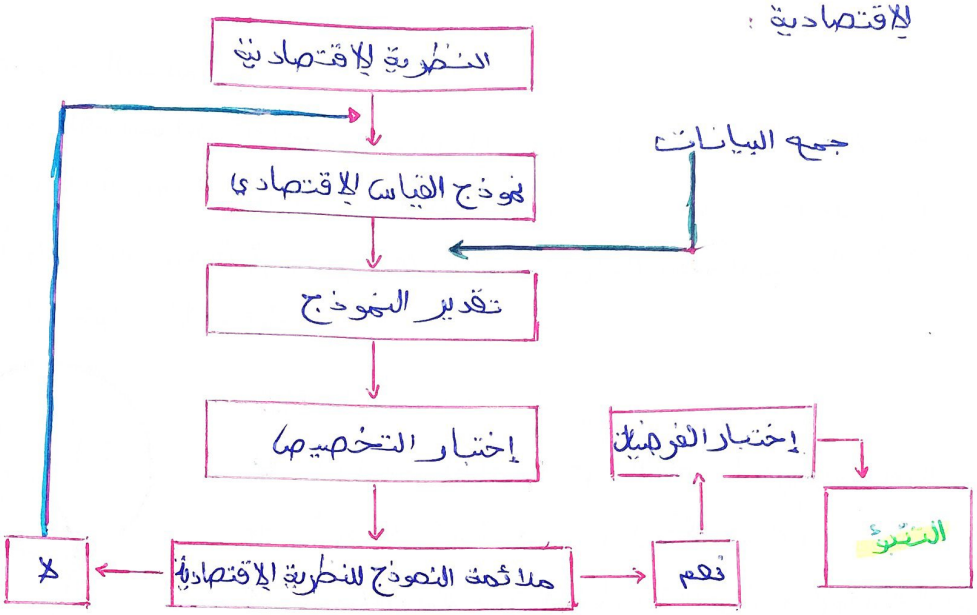
2- العلاقة بين الإقتصاد القياسي و الفروع الأخرى :

- يعتبر الإقتصاد القياسي محصلة لثلاث فروع من العلوم وهي : النظرية الإقتصادية و الإحصاء و الرياضيات ، أما عن الإحصاء فهو يمدنا بأساليب وطرق القياس مثل الإرتباط والإنحدار بالإضافة إلى البيانات الواقعية المبوبة ، أما بالنسبة للنظرية الإقتصادية فهي تحدد لنا العلاقات الإقتصادية المراد قياسها من خلال الكروف المفسرة التي تقدمها النظرية الإقتصادية ، بينما تيسر لنا الرياضيات هذه العلاقات الرياضية في صورة معادلات قابلة للقياس ولكن هذا لا يعني أن الإقتصاد القياسي ليس له صفة مستقلة عن هذه الفروع وإنما هو فرع متميز عن كل واحد منها .

3- أهداف الإقتصاد القياسي : هناك ثلاث أهداف رئيسية :

- بناء النماذج القياسية للإقتصادية في شكل قابل للإختبار الميداني وتمثل هذه المرحلة مشكلة تصور الصياغة الرياضية في منهجية القياس الإقتصادي .
- تقدير واختبار هذه النماذج مستعملين البيانات المتوفرة وتمثل هذه المرحلة القياسية في القياس الإقتصادي .
- إستعمال النماذج المقدرة لغرض التنبؤ التحليل الإقتصادي واتخاذ القرارات المناسبة .

- لإيضاح أكثر نستطيع إدراج المخطط التالي الذي يبين وضعية الاقتصاد القياسي من النظرية الاقتصادية :



06- منهجية البحث في الاقتصاد القياسي:

- لا يختلف الاقتصاد القياسي عن غيره من العلوم الاقتصادية والرياضية في منهجية بحثه إلا أنه يركز في بحثه على إيجاد ثلاث قيم رئيسية وهي :

- **قيم المعاملات الاقتصادية القياسية :** وهي القيم التي تمثل حلقة الوصل والقياس بين المتغيرات المستقلة و المتغيرات التابعة ولهذا يتركز جل إهتمامنا في البحث عن أفضل الطرق الممكنة أن نجد بواسطتها هذه المعالم بأدق مستوى ممكن لتعبر بأصدق تعبير عن طبيعة العلاقة وقوة العلاقات الحقيقية الموجودة بين المتغيرات إن كان ذلك المجتمع أو العينة بحيث تصل إلى التطابق بين قيم العينة والقيم الخاصة بالمجتمع كما أنها تكون ممثلة لحقيقة الوضع الاقتصادي للظاهرة قيد الدراسة وسلوكها في الحاضر والماضي .

- **قيمة المتغير العشوائي :** ε أو u والذي يمثل حلقة الوصل بين القيم الحقيقية والقيم المقدرة والتي يقوم فيما بينها وصولا إلى هـ بقصد الوصول إلى أقرب دقة للقيم المقدرة .

- **قيم المتغير التابع المقدر أو المستقبلية وكذلك قيم المتغيرات المستقلة** من خلال الإستقراء والإستدلال والغرض الوصول إلى هذه البيانات يتطلب الأمر حسب منهج القياس إتباع الخطوات أو المراحل التالية :

- المرحلة 3 : مرحلة المعلوماتية .

- وهي المرحلة بالغة الأهمية في القياس والوصف والتحليل فرغم أنها تتعلق بالإحصاء وطبيعته إلا أنها القاعدة الوحيدة والمنطق الحاسم للقياس وتضم هذه المرحلة الإجراءات التالية :

- التحقق من البيانات المتوفرة والمجموعة عن تطور الظاهرة وسلوكها في الماضي وتدقيقها عبر دراسة المجتمع الإحصائي وأسلوب إختيار العينة وملائمتها للمجتمع ودقة المعالجة المعلوماتية والتقنية ودقة جدولة المعلومات ووضعه في رسوم بيانية مناسبة وعزل المعلومات المشوشة وتحديد التشويش ومصدره .

- المعالجة الإحصائية للبيانات لإيجاد المتوسطات المتحركة والثابتة وإزالة المعلومات الشاذة وتصنيفها لأغراض القياس ويقوم بهذه المهمة الإحصاء الإقتصادي وهو يعد بمثابة الوجلات التي يسير بها القياس الإقتصادي .

SHOT ON REDMI 9
AI QUAD CAMERA

• المرحلة 02: مرحلة تشخيص النموذج

- تعتبر الفروض الإقتصادية المعتمدة على النظرية الإقتصادية فروضا مؤكدة وذلك يعني إفتراض وجود علاقة بين المتغيرات التابعة و المستقلة ويتم تحديد ذلك كما يلي:

- إسقاط أزواج المتغيرات على إحداثيات معينة ورسم الشكل الإنتشاري الذي يمكن أن يأكد لنا وجود علاقة أو عدم وجود علاقة بين المتغيرات، و القوة التقديرية لهذه العلاقة وبناء النموذج الإقتصادي على ضوءها.

- صياغة العلاقة باستخدام الرموز الرياضية كتصوير العلاقة بين الإستثمار و أسعار الفائدة في النظرية الكلاسيكية بالشكل التالي: $I = I_0 - Ki$

- صياغة النموذج الرياضي وهو نموذج إفتراضي لخطية أو لاخطية العلاقة بين المتغيرات، وذلك إعتمادا على الشكل الإنتشاري وكذلك تحديد المعلمات التي تحل العلاقة بين المتغيرات.

- بعد إستكمال تحديد النموذج الرياضي يتم إدخال المتغير العشوائي ε_i لتقدير الأخطاء العشوائية للمعادلة ولصياغة النموذج القياسي. $\hat{I} = I_0 - Ki + \varepsilon_i$

• المرحلة الثانية: مرحلة تقدير المعلمات.

- في هذه المرحلة يتم معالجة المعلومات المتوفرة عن المجتمع و العينة رياضيا وإحصائيا لإستخراج قيم المعلمات و المتغير العشوائي والتي تتفق ضمنيا مع الفروض الإقتصادية ومنه نحصل على الصياغة الرقمية للنموذج باستخدام طريقة المربعات الصغرى (في الحالة العادية) أو أي تقديرات أخرى مناسبة للنموذج.

• المرحلة الثالثة: تقييم المقدرات.

- هي مطابقة مجموع القيم المقدرة مع القيم الحقيقية وفي حالة عدم التطابق يلجأ الحل حين الوصول إلى هذه النتيجة التي تعتبر عمليا أول إختبار لصحة التقدير وهي إختبارات من الدرجة الأولى. $\sum y_i = \sum \hat{y}_i$ وكذلك معلمات الإرتباط و الإنحدار و التحديد و الخطأ المعياري للتقدير بحيث يكون هذا الأخير في الحد الأدنى قياسيا لأي نموذج آخر.

SHOT ON REDMI 9
AI QUAD CAMERA

م - إختبار جودة الإستدلال :

- وذلك باختبار قدرة النموذج على التنبؤ و إجراء التنبؤ الفعلي بعد قبول النظرية إذا تطابقت البيانات

3- أنواع المتغيرات الإحصائية :

- المتغيرات هي الخاصية المدروسة أو الظاهرة الإحصائية المدروسة في مجتمع إحصائي

وتنقسم إلى قسمين :

• متغيرات كيفية (نوعية) : هي تلك المتغيرات التي لا يمكن قياسها أو غير قابلة للقياس

مثل : الجنسية، الحالة العائلية، الحالة المدنية ...

• متغيرات كمية : وهي تلك المتغيرات التي يمكن قياسها ، وهي أكثر المتغيرات إنتشارا

و إستعمالا لأن لغة الإحصاء هي لغة الأرقام مثل : الإنتاج، الإستهلاك، الإستثمار...

- وتنقسم المتغيرات الكمية بدورها إلى قسمين :

• متغيرات منفصلة (متقطعة) : هي تلك المتغيرات التي تأخذ قيما صحيحة لا يمكن تجزئتها مثل عدد الطلاب

عدد الأطفال، عدد قطع الغيار

• متغيرات المستمرة : هي تلك المتغيرات التي تأخذ كل القيم الممكنة لمجال الدراسة

ونظرا للعدد غير المتناهي لهذه القيم نقسم مجال الدراسة الى مجالات جزئية تسمى الفئات .

4- بنية المتغيرات الإقتصادية :

- تأخذ البيانات الإقتصادية عدة أشكال :

• البيانات المقطعية : تتكون مجموعة البيانات المقطعية من عينة الأفراد أو قطاع العائلات

أو الشركات أو الدول، المناطق أو المدن أو أي نوع من الوحدات في نقطة محددة من الزمن

وفي بعض الحالات لا تتماثل الفترة الزمنية للبيانات بالضبط ، مثل : مسح بيانات

العائلات المختلفة خلال أيام مختلفة من الشهر . وفي هذه الحالة يتم إهمال فروق

التوقيت في جمع البيانات وتسمى البيانات التي يتم جمعها بيانات مقطعية .

- : فإستخدام البيانات المقطعية واسع النطاق في الإقتصاد و العلوم الإجتماعية

ويساهم تحليلها بالإقتصاد الجزئي التطبيقي : إقتصاديات العمل ، المالية العامة

للدولة ، إقتصاديات الأعمال ، الإقتصاد الطبي وبعض الحقول في الإقتصاد الجزئي

في نقطة من الزمن

لإختبار فرضيات الإقتصاد الجزئي وتقييم السياسات الإقتصادية.

• بيانات السلاسل الزمنية - تتكون السلاسل الزمنية من مشاهدات متغير واحد أو أكثر عبر الزمن منظمة في ترتيب تسلسل زمني معين مثل: بيانات سنوية، شهرية، ومن أمثلة السلاسل الزمنية: أسعار الأسهم، الناتج المحلي الإجمالي، أسعار الفائدة.
وطبيعة البيانات الزمنية تجعل من تحليلها أكثر صعوبة من البيانات المقطعية. وتعتمد البيانات الاقتصادية على الزمن وهذا يعني أن أغلبها ترتبط بتاريخها الماضي وتطبق أغلب الإجراءات القياسية على البيانات المقطعية.

- لأننا نحتاج في هذه الحالة إلى إجراءات لتحديد النموذج القياسي المناسب كما أن البيانات الاقتصادية الزمنية تتضمن اتجاهًا زمنيًا يقودنا إلى أساليب قياسية جديدة.

• بيانات السلاسل الزمنية المقطعية (Panel Data) : تتكون بيانات Panel من سلاسل زمنية لكل طرف مقطعي في مجموعة البيانات مثل: المبيعات وعدد العاملين لخمس شركات خلال فترة زمنية 10 سنوات، يمكن جمع بيانات Panel على أساس الجغرافي مثل: الناتج المحلي الإجمالي وعرض النقود لدول خلال فترة زمنية 20 سنة

• المحور 3 : نظرية الإرتباط

3 : معامل الإرتباط :

• هو نوعين :

• معامل الإرتباط البسيط : مؤشر إحصائي يقيس العلاقة بين متغيرين أحدهما مستقل والآخر تابع

خصائصه :

✓ قيمة معامل الإرتباط البسيط محصورة بين $[-1, 1]$

✓ إذا كان $ر = 0$ فإنه لا يوجد إرتباط أي أن المتغيرين X و Y مستقلين

$ع = 0$ أو $ع = 1$

✓ إذا كان $r = 1$ فإنه توجد ارتباط قوي موجب.

✓ إذا كان $r = -1$ فإنه توجد علاقة ارتباط قوية جدا سالبة.

✓ معامل الإرتباط تناظري.

✓ معامل الإرتباط يوضح طبيعة العلاقة: $r_{xy} = r_{yx}$.

✓ إذا تم طرح أو ضرب أو قسمة الـ x و y على نفس العدد فإن الإرتباط لا يتغير.

القانون:

$$r_{xy} = \frac{cov(x,y)}{\sigma_x \cdot \sigma_y}$$

$$cov(x,y) = \frac{\Sigma(x_i - \bar{x})(y - \bar{y})}{n} \quad \text{و} \quad \sigma_x = \sqrt{\frac{\Sigma(x_i - \bar{x})^2}{n}}$$

$$\sigma_y = \sqrt{\frac{\Sigma(y_i - \bar{y})^2}{n}}$$

$$r_{xy} = r_{yx} = \frac{\Sigma(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\Sigma(x_i - \bar{x})^2 \cdot \Sigma(y_i - \bar{y})^2}}$$

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

- الدلالة الإحصائية لمعامل الإرتباط:

- من أجل معرفة الدلالة الإحصائية لمعامل الإرتباط نقوم بحساب إحصائية ستيودنت ونقارنها بالقيمة المجدولة، فإذا كانت قيمة ستيودنت المحسوبة أكبر من المجدولة فإن معامل الإرتباط له معنوية أي: أن هناك علاقة خطية بين الـ x و y وإذا كانت إحصائية (T) المحسوبة أقل من المجدولة فإنه لا توجد دلالة إحصائية لمعامل الإرتباط أي لا توجد علاقة خطية بين المتغيرين. (يوجد إرتباط لكن لا توجد علاقة)

- وإحصائية ستيودنت تحسب بالقانون التالي:

$$T_{cal} = \frac{r}{\sqrt{\frac{1-r^2}{n-2}}} \quad , \quad T_{tab}\left(n-2 \; , \; \frac{\alpha}{2}\right)$$

(درجة المعنوية)

- معامل الإرتباط المتعدد : مؤشر إحصائي يقيس قوة العلاقة بين أكثر من متغيرين ، إشارته لا تعني طبيعة العلاقة .

خصائصه :

✓ لا نستطيع فصله عن معامل الإرتباط البسيط ، لأنه يعتمد في حسابه الإرتباط البسيط .

✓ يجب تحديد المتغير التابع والمستقل .

✓ عند تحديد العلاقة بين X_1 و X_2 و X_3 تصاغ في المعادلة التالية :

X_1 هو المتغير المستقل . ← X_1 تابع

$$r_{1,23} = \sqrt{\frac{r_{12}^2 + r_{13}^2 - 2 r_{23} \cdot r_{12} \cdot r_{13}}{1 - r_{23}^2}}$$

X_2 هو التابع ←

$$r_{2,13} = \sqrt{\frac{r_{21}^2 + r_{23}^2 - 2 r_{13} \cdot r_{21} \cdot r_{23}}{1 - r_{13}^2}}$$

- الدلالة الإحصائية لمعامل الإرتباط :

- لمعرفة الدلالة الإحصائية نقوم بحساب إحصائية فيشر

↳ فإذا كانت قيمة فيشر المحسوبة أكبر من المجدولة هذا يعني أن معامل الإرتباط له دلالة إحصائية ، أي هناك علاقة خطية بين المتغيرات المستقلة والمتغير التابع .

↳ إذا كانت أقل هذا يعني أنه لا توجد علاقة خطية بين المتغيرات .

- وتحسب إحصائية فيشر بالعلاقة التالية :

حجم العينة ← n ، عدد المتغيرات المستقلة ← k

$$F_{cal} = \frac{r^2_{y,12\ldots}}{1 - r^2_{y,12\ldots}} \times \frac{n - k - 1}{k}$$

تابع لمعامل الإرتباط البسيط.

- معامل سبيرمان:

- إذا كان لدينا متغيرين كميين أو كيفيين فإن العلاقة بينهما تحسب من خلال رتبتهما ومعامل الإرتباط يحسب بالعلاقة التالية:

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

حيث d هي فرق رتبتي x و y
و n هي عدد المشاهدات

ملاحظة:

معامل سبيرمان نستطيع أن نحسب قوة العلاقة بين متغيرين كميين وكيفيين

- معامل الإقتران:

- هو وصف درجة العلاقة بين الظواهر غير الكمية لا يمكن التعبير عنها بالأرقام. بل هو بالوصف فقط مثل إرتباط بين الجنسين أو بين الحالة التعليمية للأفراد أو دراسة التطعيم بمصل واقي والإصابة بمرض معين... الخ.

- ولدراسة هذا النوع من العلاقات نستعمل معامل الإقتران ويعرض وفقا للجدول التالي لمتغيرين x و y.

معامل الإقتران

$$T = \frac{AD - BC}{AD + BC}$$

	x	□	
y	A	B	
□	C	D	-

الإجابتي هو الصحيح.

- تتراوح قيمة هذا المعامل بين $[-1, 1]$

- معامل إرتباط فاي

- له نفس خصائص معامل الإقتران حيث يتعامل مع متغيرات منفصلة ثنائية التصنيف تم تطويره من قبل أوندي ويل عام 1912، وهو معطى بالعلاقة التالية:

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

SHOT ON REDMI
AI QUAD CAMERA

مثال :

- ليكن لدينا الجدول التالي يبين لنا نتائج أجريت على 400 طالب بحيث حاولنا تتبع تأثير الدروس الخصوصية على النجاح في شهادة الباكالوريا .

	طلبة قاموا بدروس خصوصية	طلبة لم يقوموا بدروس خصوصية	
طلبة نجحوا	220	80	300
طلبة لم ينجحوا	60	40	100
	280	120	400

1- أحسب معامل الإقتران ؟

2- حساب معامل الإرتباط لفاي ؟

$$T = \frac{AD - BC}{AD + BC} \qquad T = \frac{(220)(40) - (80)(60)}{(220)(40) + (80)(60)} \Rightarrow T = 0,29$$

$$V_\varphi = \frac{AD - BC}{\sqrt{(A+B)(A+C)(A+D)(C+D)}} \qquad V_\varphi = \frac{(220)(40) - (80)(60)}{\sqrt{(220+80)(220+60)(220+40)(60+40)}} = V_{\varphi = 0,07}$$

* معامل التوافق لبيرسون :

- يدرس هذا المعامل العلاقة بين الظواهر الوصفية التي تنقسم إلى أكثر من نوعين حيث لا يساعد معامل الإقتران السابق الذكر ، بل يستخدم معامل التوافق لبيرسون الذي يصلح أيضا لقياس العلاقة بين الظواهر الكمية القابلة للقياس وأخرى وصفية لا يمكن قياسها ويحسب بموجب العلاقة التالية :

$$T_P = \sqrt{\frac{G-1}{G}}$$

حيث : G هي حاصل قسمة مربع كل تكرار في الجدول التكراري على حاصل ضرب مجموع بيانات الصف في مجموع بيانات العمود التي يقع فيها التكرار حيث نقوم بترتيب كل تكرار وارد في الجدول التكراري ثم نقوم بقسمته على حاصل ضرب التكرار الكلي الوحيد

▫ المحور الثالث : نظرية الإنحدار

1- تعريف تحليل الإنحدار : هو أسلوب لتحليل البيانات يتوقع قيمة بيانات غير معروفة باستخدام بيانات أخرى ذات صلة ومعروفة ، يقوم هذا النموذج على النموذج الرياضي لمتغير غير معروف (المتغير التابع) ومتغير معروف (المستقل) كمعادلات رياضية قد تكون خطية (من الدرجة الأولى) أو غير خطية (انحدار لوجستي)

أولا : تحليل الإنحدار الخطي : هو عبارة عن أسلوب يكون فيه النموذج الرياضي الممثل للعلاقة بين المتغيرين أو أكثر في شكل معادلة خطية ، وينقسم الإنحدار الخطي إلى : إنحدار خطي بسيط و إنحدار خطي متعدد .

▪ تحليل الإنحدار الخطي البسيط :

1- تعريف : هو علاقة خطية بين متغيرين و يهتم بدراسة أثر أحد المتغيرين يسمى بالمتغير المستقل على المتغير الثاني يسمى المتغير التابع . ومن أمثلة ذلك الإنفاق الإستهلاكي وعلاقته بالدخل التصرفي ، كميات الطلب على سلعة وسعرها ... الخ

ومن ثم يمكن عرض نموذج الإنحدار الخطي في شكل معادلة خطية من الدرجة الأولى تعكس المتغير التابع كدالة في المتغير المستقل $y_i = \alpha + \beta x_i + \varepsilon_i$

حيث y_i : تمثل المتغير التابع ، x_i : المتغير المستقل ، α تمثل ثابت يعكس قيمة المتغير التابع إذا انعدم المتغير المستقل

β : يمثل الميل : وهو يعكس التغيرات الحاصلة في المتغير تابع إذا تغير المستقل بوحدة واحدة .

ε_i : هو الخطأ العشوائي والذي يمثل الفرق بين القيمة الحقيقية والقيمة المقدرة من المعادلة

لدقة القياس ← وذلك لتصحيح بيانات لانتهاء عوامل أخرى تأثر في القيمة

لمعالجة بعض المشاكل كالأخطاء في القياس أو بعض العوامل الغير قابلة للقياس .

ب/ المتغير العشوائي : • أسباب إضافة المتغير العشوائي ε : يمكن إرجاع أسباب إضافة المتغير العشوائي له عدة أسباب أهمها مايلي :

- ✓ حذف بعض المتغيرات من الدالة المدروسة .
- ✓ السلوك العشوائي للجنس البشري .
- ✓ الصياغة الناقصة للنموذج القياسي .
- ✓ أخطاء التجميع .
- ✓ أخطاء القياس .

التوزيع التكراري الاحتمالي .

مثال : ليكن لدينا الجدول التالي :

المعايير / البدائل	ضعيف	مقبول	حسن	قريب من الجيد	جيد	جيد جدا	ممتاز	w_i
ضعيف	α_{11}	α_{12}	α_{13}	α_{14}	α_{15}	α_{16}	α_{17}	B_1
مقبول	α_{21}							B_2
حسن	α_{31}							B_3
قاصر الجيد	α_{41}							B_4
جيد	α_{51}							B_5
جيد جدا	α_{61}							B_6
ممتاز	α_{71}						α_{77}	B_7
$\sum$	γ_1	γ_2	γ_3	γ_4	γ_5	γ_6	γ_7	

$$G = \frac{(\alpha_{11})^2}{B_1 \cdot \gamma_1} + \frac{(\alpha_{12})^2}{B_1 \cdot \gamma_2} + \text{———} + \cdots + \frac{(\alpha_{77})^2}{B_7\ \gamma_7}$$

$$T = \sqrt{\frac{G-1}{G}}$$

- ثم نقوم بتطبيق هذا القانون

- الفرضيات الخاصة بالمتغير العشوائي:

1- القيمة المتوقعة للمتغير العشوائي معدومة $E(\varepsilon_i) = 0$ *

2- تباين المتغير العشوائي ثابت عند كل قيمة من قيم x_i.

* $Var(\varepsilon_i) = E(\varepsilon_i - E(\varepsilon_i))^2 = E(\varepsilon_i)^2 = \sigma^2$

- وفي حال عدم ثبات التباين تظهر لدينا مشكلة عدم تجانس التباين.

3- المتغير العشوائي يتوزع توزيعا طبيعيا بتوقع معدوم وتباين ثابت يساوي σ^2 (وسط حسابي)

* $\varepsilon_i \sim N(0, \sigma^2)$

4- قيم المتغير العشوائي غير مرتبطة بالمتغيرات المستقلة أي انعدام التباين المشترك بين المتغير العشوائي والمتغيرات المستقلة $Cov(\varepsilon_i, x_i) = 0$ *

5- قيم المتغير العشوائي غير مترابطة فيما بينها $cov(\varepsilon_i, \varepsilon_j) = 0$ *

II تقدير النموذج الخطي البسيط:

- يمكن الحصول على مقدرات المعالمات بعدة طرق ومنها طريقة المربعات الصغرى هذه الطريقة تعتبر الأحسن والأدق مقارنة بمثيلاتها حيث تعتمد في الحصول على مقدرات خاصة للانحدار التالي ← $y_i = \alpha + \beta x_i + \varepsilon_i$

- وتهدف هذه الطريقة إلى تصغير مجموع مربعات البواقي إلى أدنى قيمة لها بمعنى آخر أن مجموع مربعات الأخطاء يكون أقل ما يمكن، حيث يجري تعريف مكون يطلق عليه مجموع مربعات البواقي $\sum \varepsilon_i^2$ والتي تمثل انحرافات القيم المقدرة عن القيم الفعلية ويعبر عنه رياضيا بالكتابة على الشكل التالي

$y_i = \alpha + \beta x_i + \varepsilon_i$ ... (1)

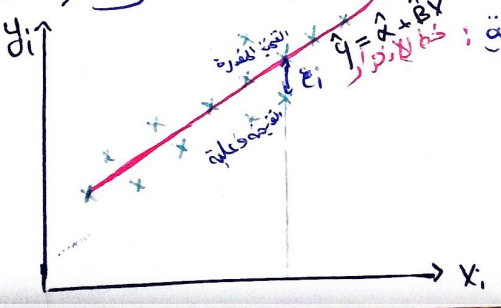
النموذج المقدر يساوي: $\hat{y}_i = \hat{\alpha} + \hat{\beta} x_i$ ... (2)

معادلة البواقي هي الفرق بين (1) و (2) ←

$\varepsilon_i = y_i - \hat{y}_i$

$\varepsilon_i = y_i - (\hat{\alpha} + \hat{\beta} x_i)$

حيث يعبر الفرق بين القيمة الفعلية y والقيمة المقدرة $\hat{y}$ ، ويمكن توضيح هذا الخطأ على شكل لوحة الانتشار التالية:



نرمز إلى مجموع بمربعات البواقي بـ $\sum \varepsilon_i^2$

$$\varepsilon_i = y_i - (\hat{\alpha} + \hat{\beta} x_i)$$

مربعات البواقي ← $\varepsilon_i^2 = (y_i - (\hat{\alpha} + \hat{\beta} x_i))^2$

مجموع مربعات البواقي ← $\sum \varepsilon_i^2 = \sum (y_i - (\hat{\alpha} + \hat{\beta} x_i))^2$

$$\text{Min} \sum \varepsilon_i^2 = \text{Min} \sum (y_i - (\hat{\alpha} + \hat{\beta} x_i))^2$$

↗

ا إشتقاق المعلمات المقدرة:

من أجل تقدير $\hat{\alpha}$ نقوم

$\frac{\delta \sum \varepsilon_i^2}{\delta \hat{\alpha}}$ - نشتق مجموع مربعات البواقي بالنسبة لـ α

$$\frac{\delta \sum \varepsilon_i^2}{\delta \hat{\alpha}} = \sum (2)(-1)(y_i - \hat{\alpha} - \hat{\beta} x_i)^{2-1}$$

$$= (-2) \sum (y_i - \hat{\alpha} - \hat{\beta} x_i) = 0$$

نضع المشتقة الأولى = 0

$$\sum (y_i - \hat{\alpha} - \hat{\beta} x_i) = 0$$

$$\sum y_i - \sum \hat{\alpha} - \sum \hat{\beta} x = 0$$

$$\sum y_i - n\hat{\alpha} - \hat{\beta} \sum x_i = 0$$

$$n\hat{\alpha} = \sum y_i - \hat{\beta} \sum x_i$$

$$\hat{\alpha} = \frac{\sum y_i}{n} - \hat{\beta} \frac{\sum x_i}{n}$$

$$\boxed{\hat{\alpha} = \bar{y} - \hat{\beta} \bar{x}}$$

• تذكير:

$$f(x) = s(x)^n$$

$$f'(x) = n(s'(x)) s(x)^{n-1}$$

$$\sum k.x = k \sum x$$

$$\sum A+B = \sum A + \sum B$$

$$\sum A \times B \neq \sum A \sum B$$

$$\sum k = n.k$$

من أجل تقدير $\hat{\beta}$ نقوم بإشتقاق مجموع مربعات البواقي بالنسبة لـ β.

$$\frac{\delta \sum \varepsilon_i^2}{\delta \hat{\beta}} = \sum (2)(-x_i)(y_i - \hat{\alpha} - \hat{\beta} x_i) = 0$$

$$= -2 \sum [x_i y_i - \hat{\alpha} x_i - \hat{\beta} x_i x_i] = 0$$

$$= \sum x_i y_i - \hat{\alpha} \sum x_i - \hat{\beta} \sum x_i^2 = 0$$

$$= \sum x_i y_i - \left(\frac{\sum y_i}{n} - \hat{\beta} \frac{\sum x_i}{n}\right) \sum x_i - \hat{\beta} \sum x_i^2 = 0$$

$$= \sum x_i y_i - \frac{\sum y_i \cdot \sum x_i}{n} = \hat{\beta} \sum x_i^2 - \hat{\beta} \frac{\sum x_i \sum x_i}{n}$$

$$\hat{B}\left(\sum x_i^2 - \frac{\sum x_i \sum x_i}{n}\right) = \sum x.y_i - \frac{\sum y_i \sum x_i}{n}$$

$$\hat{B}\left(\frac{n\sum x^2 - (\sum x_i)^2}{n}\right) = \sum x_i y_i - \frac{\sum y_i \sum x_i}{n}$$

$$\boxed{\hat{B} = \frac{n\sum x_i y_i - \sum x_i \sum y_i}{n\sum x_i^2 - (\sum x_i)^2}}$$

← القانون العام كعامل د الانحدار

نضرب البسط و المقام $\times n$.

$$\hat{B} = \frac{\sum x.y - n\bar{x}\bar{y}}{\sum x^2 - n(\bar{x})^2}$$

القانون الخاص:

- باستخدام الانحرافات يمكن الحصول على المقدرات كما يلي:

$$\sum (y_i - \bar{y})^2 = \sum y_i^2 - n\bar{y}^2$$

$$\sum (x_i - \bar{x})(y_i - \bar{y}) = \sum x_i y_i - n\bar{x}\bar{y}$$

$$\sum (x_i - \bar{x})^2 = \sum x_i^2 - n\bar{x}^2$$

$$\sum x_i y_i = \hat{\alpha}\sum x_i + \hat{B}\sum x_i^2$$

$$\sum x_i y_i = \sum x_i (\bar{y} - \hat{B}\bar{x}) + \hat{B}\sum x^2$$

$$= n\bar{x}(\bar{y} - \hat{B}\bar{x}) + \hat{B}\sum x_i^2$$

$$\sum x_i y_i - n\bar{x}\bar{y} = -\hat{B}\bar{x}^2 + \hat{B}\sum x^2$$

$$= \hat{B}\left(\sum x_i^2 - n\bar{x}^2\right)$$

$$\hat{B} = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sum (x_i - \bar{x})^2}$$

$$\hat{B} = \frac{\sum xy}{\sum x^2}$$

← القانون الخاص (المختصر)

- إذا كانت $\hat{B} = 0$ لا يوجد معادلة لانحدار y قيمة ثابتة وهي $\hat{\alpha}$.

- فرضيات النموذج :

- لطريقة المربعات الصغرى 6 فرضيات أساسية فالإخلال بإحدى منها يؤدي إلى نتائج مضللة خاصة عند دراسة صلاحية النموذج وتكمن هذه الفرضيات في النقاط التالية:

- أن تكون العلاقة خطية بين المتغير التابع والمستقل وبالتالي فإن الشكل الوطيني للعلاقة بين y و x يكون بالشكل التالي $y_i = \alpha + \beta x_i + \varepsilon_i$.
- أن التوقع أو الأمل الرياضي أو متوسط البواقي يكون معدوم $E(\varepsilon_i) = 0$.
- تجانس تباين الأخطاء أي أن تباين أو تشتت البواقي (المتغير العشوائي) متجانس ومتابت، من أجل كل القيم i مهما كانت $Var(\varepsilon_i) = \sigma^2_i = \sigma^2$
- عدم وجود إرتباط ذاتي للأخطاء بين البواقي معناه أن التباين المشترك بين البواقي يساوي الصفر مهما كانت i تختلف عن j $\forall i \neq j$ و $cov(\varepsilon_i, \varepsilon_j) = 0$
- عدم وجود إرتباط بين المتغير المستقل والمتغير العشوائي معناه أن التباين المشترك بينهما معدوم $cov(x_i, \varepsilon_i) = 0 \quad \forall i$
- التوزيع الإحتمالي للمتغير العشوائي هي توزيع طبيعي $\varepsilon_i \sim N(0, \sigma^2)$

■ الخصائص العددية والخصائص الإحصائية للمعلمات المقدرة :

- الخصائص العددية : تتناول التوقع والتباين .

- ليكن النموذج الهيكلي الخطي البسيط التالي : $y_i = \alpha + \beta x_i + \varepsilon_i$ ، حيث النموذج المقدر لهذا النموذج هو $\hat{y}_i = \hat{\alpha} + \hat{\beta} x_i$ ، علما أن $\hat{\alpha}$ و $\hat{\beta}$ هي مقدرات α و β على التوالي :

- التوقع الرياضي للمعلمات المقدرة :

$E(\hat{\beta}) = \beta$ ، $E(\hat{\alpha}) = \alpha$

التباين للمعلمات المقدرة:

بالنسبة لـ β .

$$Var(\hat{\beta}) = \frac{\frac{\sum(y_i - \hat{y})^2}{n-2}}{\sum(x_i - \bar{x})^2} = \frac{\sigma_{\varepsilon_i}^2}{\sum(x_i - \bar{x})^2}$$

$$\text{انحراف } \sigma_{\hat{\beta}} = \sqrt{Var(\hat{\beta})} = \sqrt{\frac{\sigma_{\varepsilon_i}^2}{\sum(x_i - \bar{x})^2}}$$

بالنسبة لـ α .

$$Var(\hat{\alpha}) = \frac{\frac{\sum(y_i - \hat{y})^2}{n-2}}{n} + \bar{X}^2 Var(\hat{\beta})$$

$$\sigma_{\hat{\alpha}} = \sqrt{Var(\hat{\alpha})} = \sqrt{\frac{\sum(x_i - \hat{y})^2/n-2}{n} + \bar{X}^2 Var(\hat{\beta})}$$

- **الخصائص الإحصائية:**

- المعلمات المقدرة بطريقة المربعات الصغرى 3 خصائص إحصائية وهي:
  - خاصية عدم التحيز
  - خاصية الكفاءة
  - خاصية الإتساق

- خاصية عدم التحيز = نقول أن $\hat{\beta}$ هو أفضل مقدر غير متحيز إذا كان توقع المعلمة المقدرة يساوي إلى المعلمة $E(\hat{\beta}) = \beta$
- خاصية الكفاءة = نقول أن المعلمة المقدرة كفأة إذا كانت تتحيز بأصغر تباين.
- خاصية الإتساق = نقول أن المعلمات المقدرة أنها متسقة إذا تحقق ما يلي: كلما كبر حجم العينة فإن قيمة توقع المعلمة المقدرة تقترب من قيمة المعلمة الحقيقية و قيمة المعلمة المقدرة يقترب من الصفر، ويمكن التعبير عنها رياضيا

$$\begin{cases} \lim_{n\to\infty} E(\hat{\beta}) = \beta \;,\; \lim_{n\to\infty} E(\hat{\alpha}) = \alpha \\ \lim_{n\to\infty} Var(\hat{\beta}) = 0 \;,\; \lim_{n\to\infty} Var(\hat{\alpha}) = 0 \end{cases}$$

- دراسة صلاحية النموذج :

* معامل التحديد R^2 : عبارة عن مربع معامل الإرتباط .

نفرض النموذج $\hat{y} = \hat{\alpha} + \hat{\beta}x$ معامل التحديد يحدد كم النموذج مفسر للظاهرة $\hat{y}$ (القوة التفسيرية للمتغير الحاصل في المتغير التابع) .

و القيمة المتبقية يفسرها ε (عوامل لم تأخذ بالإعتبار) .

مثلا : $R^2 = 0,98$ أي أن النموذج $\hat{\alpha} + \hat{\beta}x$ مفسر للظاهرة $\hat{y}$ بـ 98%

و ε المتبقية يفسرها

- كلما زاد معامل التحديد يكون النموذج أفضل

- إذا كان $F_{tab} > R^2$ ← رفض النموذج (الجزء المفسر الأكبر هي عوامل لم تأخذ) .

- إذا كان $R^2 > F_{tab}$ ← قبول النموذج (النموذج مفسر للظاهرة $\hat{y}$)

حساب R^2 :

$$R^2 = \frac{SSR}{SST} = \frac{\sum(\hat{y}-\bar{y})^2}{\sum(y_i-\bar{y})^2}$$

مجموع مربعات الإنحدار / مجموع مربعات الإنحرافات الكلية

مجموع إنحرافات الأخطاء

$$SST = SSR + SSE$$

$$\sum(y_i-\bar{y})^2 = \sum(\hat{y}-\bar{y})^2 + \sum(y-\hat{y})^2$$

$$R^2 = 1 - \frac{SSE}{SST} = 1 - \frac{\sum(y_i-\hat{y})^2}{\sum(y_i-\bar{y})^2}$$

$$R^2 = \frac{\hat{\beta}\sum(x_i-\bar{x})^2}{\sum(y_i-\bar{y})^2}$$ ← دائما موجب

* المعنوية الكلية - : $\hat{y} = \hat{\alpha} + \hat{\beta}x$ ⇐ نحسب F_{cal} ⇐ $\begin{cases} H_0 & \text{غير معنوي احصائيا} \\ H_1 & \text{معنوي} \end{cases}$

$$F_{cal} = \frac{SSR}{SSE/n-k}$$

k : عدد المتغيرات

المعنوية الفردية : t

$$t_{cal} = \frac{|\hat{\beta}|}{S_{\hat{\beta}}}$$