تمهيد: يعتبر الانحدار أحد الأساليب الإحصائية التي تستخدم في قياس العلاقات الاقتصادية بين متغير ما يسمى بالمتغير التابع ومتغير آخر أو مجموعة من المتغيرات تسمى بالمتغيرات المفسرة أو المستقلة، كما تنقسم نماذج الانحدار إلى عدة أنواع فهنالك الانحدار الخطي والانحدار غير الخطي، وهناك الانحدار البسيط والانحدار المتعدد وتحدد درجة الخطية على أساس العلاقة المراد قياسها وصفتي البسيط والمتعدد على حسب عدد المتغيرات التفسيرية .

1- الشكل العام لنموذج الانحدار الخطي البسيط: يستخدم النموذج الخطي البسيط لتكوين علاقة بين متغير تابع y ومتغير مستقل مفسر x بواسطة عينة n من الملاحظات، هذه العلاقة تسمح بشرح قيم y بواسطة قيم مأخوذة من طرف x، كعلاقة الكمية المطلوبة من سلعة ما مع سعرها أو علاقة الإنفاق الاستهلاكي بالدخل المتاح....إلخ، أو بعبارة أخرى أثر متغير مستقل واحد على متغير آخر يسمى بالتابع، وتعرف علاقة الانحدار بالشكل التالي:

$$Y_i = \alpha + \beta X_i + \varepsilon_i \qquad i = 1.....n$$

حيث: Y_i المتغير التابع (الداخلي) ، X_i : المتغير المستقل (الخارجي)، ε_i : حد الخطأ (المتغير العشوائي)، أما α يمثل الجزء الثابت وهو الجزء المقطوع من المحور الرأسي، وهو عبارة عن قيمة متوسط المتغير التابع لما تنعدم قيمة المتغير المفسر ، بينما β فيمثل ميل المستقيم وهو يعبر عن مقدار التغير في المتغير التابع إذا حدث تغير في المتغير المستقل بوحدة واحدة ويطلق عليه أيضا بمعامل الانحدار، وإشارته تبين ما إذا كانت هنالك علاقة طردية أو عكسية بين المتغير التابع والمتغير المستقل، n : عدد المشاهدات أو الملاحظات .

ويرجع وجود حد الخطأ ε_i إلى عدة أسباب منها :

- إهمال بعض المتغيرات المستقلة التي يمكن أن تؤثر على المتغير التابع في النموذج.
- الصياغة الرياضية غير السليمة للنموذج.
- حدوث خطأ في كل من تجميع البيانات وقياس المتغيرات الاقتصادية .

2 - فرضيات النموذج: يعتبر ε_i الخطأ متغير عشوائي حيث يخضع للفرضيات الأساسية:

الفرضية الأولى : الأمل الرياضي للأخطاء معدوم أي $E(\varepsilon_i) = 0, \forall i = 1...n$

وتعني هذه الفرضية أن الأخطاء لا تدخل في تفسير y، إذ أنها تعبر عن حدود عشوائية تأخذ قيما سالبة

أو موجبة أو معدومة وبالتالي لا يمكن تحديدها أو قياسها بدقة، وهي تخضع لقوانين الاحتمال بوسط أو توقع معدوم .

الفرضية الثانية: ثبات (تجانس) تباين الأخطاء Homoscedasticity

وهو ما يعني أن تبعثرها أو تشتتها حول المتوسط ثابت، ونعبر عن ذلك رياضيا بـ:

$$V(\varepsilon_i) = E\left(\varepsilon_i^2\right) = \sigma^2, \forall i = 1...n$$

الفرضية الثالثة: الأخطاء تتوزع طبيعيا

ε_i موزع توزيعا طبيعيا ونكتب : $\varepsilon_i \to N(0,\sigma^2)$.

الفرضية الرابعة: لا يوجد إرتباط ذاتي بين الأخطاء

تعني بأن التباينات المشتركة لأخطاء الملاحظات المختلفة تكون معدومة أي $COV(\varepsilon_i, \varepsilon_j) = 0, \forall i \neq j$

الفرضية الخامسة: لا يوجد إرتباط بين المتغير X_i والخطأ ε_i

أي أن المعطيات التي جمعت بالنسبة لهذا المتغير قادرة على إظهار تأثيرها على مستوى التابع، بحيث تكون قيمة واحدة على الأقل تختلف عن بقية القيم، لأجلها تكون العبارة الآتية تختلف عن الصفر $\frac{1}{n}\sum\left(X_i - \bar{X}\right)^2 \neq 0$ ونعبر عن ذلك بالعبارة الآتية $COV(X_i, \varepsilon_i) = 0$

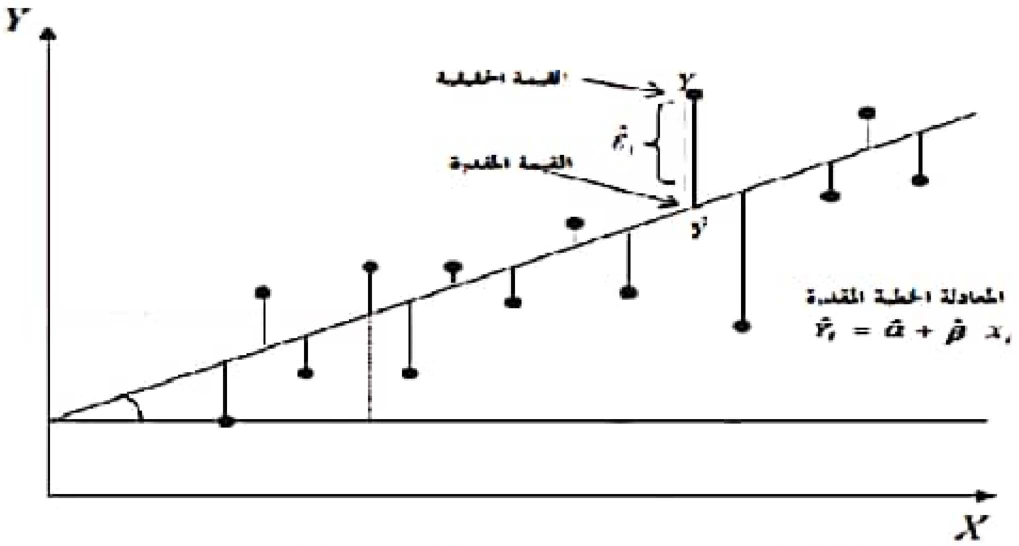
3 -تقدير معاملات النموذج بطريقة المربعات الصغرى OLS:

تعتمد هذه الطريقة في إيجادها لقيم تقديرية لمعلمات النموذج على التقليل قدر الإمكان من الفوارق بين القيم المشاهدة والقيم النظرية أو المقدرة التي يمنحها لنا مستقيم الانحدار أي $e_i = \hat{\varepsilon}_i = (Y_i - \hat{Y}_i)$ وتكون هذه الفوارق موجبة أحيانا وسالبة أحيانا أخرى ولتخطي عقبة الإشارة نلجأ إلى استعمال مربعات الفوارق مما يعطينا مجموعا دائما موجبا وبالتالي فإن مبدأ هذه الطريقة هو تصغير مجموعة مربعات الأخطاء ونكتب : $MIN\sum_{i=1}^{n} e_i^2 = MIN\sum\left(Y_i - \hat{\beta}X_i - \hat{\alpha}\right)^2$

حيث $\hat{\alpha}$: القيمة المقدرة لـ α ، $\hat{\beta}$: القيمة المقدرة لـ β

$e_i = \hat{\varepsilon}_i = (Y_i - \hat{Y}_i)$ قيم البواقي ، $\hat{Y}_i$: النموذج المقدر ، Y_i : النموذج الاقتصادي.

والشكل الموالي يوضح المعادلة الخطية المقدرة



ولإيجاد القيم المقدرة نشتق $\sum_{i=1}^{n} e_i^2$ بالنسبة لكل من $\hat{\alpha}$ و $\hat{\beta}$ بالطريقة الآتية :

$$\partial \sum_{i=1}^{n} e_i^2 \Big/ \partial \hat{\alpha} = -2\sum_{i=1}^{n} (Y_i - \hat{\alpha} X_i - \hat{\beta}) = 0$$

$$\partial \sum_{i=1}^{n} e_i^2 \Big/ \partial \hat{\beta} = -2\sum_{i=1}^{n} (Y_i - \hat{\alpha} X_i - \hat{\beta}) = 0$$

وبالتبسيط نجد:

$$\hat{\beta} = \frac{\sum_{i=1}^{n} (X_i - \bar{X})(Y_i - \bar{Y})}{\sum_{i=1}^{n} (X_i - \bar{X})^2} = \frac{\sum_{i=1}^{n} X_i Y_i - n\bar{X}\bar{Y}}{\sum_{i=1}^{n} X_i^2 - n\bar{X}^2} = \frac{COV(X_i, Y_i)}{V(X_i)}$$

$$\hat{\alpha} = \bar{Y} - b\bar{X}$$

4 – معامل التحديد (R^2):

هذا المعامل يقيس جودة النموذج، أي يوضح نسبة انحرافات قيم (y) الموضحة في النموذج بالنسبة للإنحرافات الكلية، وهو عدد موجب محصور بين [1;0] ويرمز له بالرمز (R^2) حيث هو مربع معامل الارتباط الخطي (r)، ويتم استخراج قيمته الجبرية كالتالي:

$$\sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n} (Y_i - \hat{Y}_i)^2$$

$$Y_i - \hat{Y}_i = e_i \Rightarrow \sum Y_i - \sum \hat{Y}_i = \sum e_i = 0 \Rightarrow \bar{Y} = \hat{Y}$$

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

$$\sum_{i=1}^{n} (y_i - \bar{y})^2 = \sum_{i=1}^{n} (\hat{y} - \bar{\hat{y}})^2 + \sum_{i=1}^{n} e_i^2 \ldots\ldots\ldots.(1)$$

بقسمة طرفي المعادلة الأخيرة على $\sum_{i=1}^{n}(Y_i-\bar{Y})^2$ نحصل على:

$$1=\frac{\sum_{i=1}^{n}(\hat{Y}-\bar{Y})^2}{\sum_{i=1}^{n}(Y_i-\bar{Y})^2}+\frac{\sum_{i=1}^{n}e_i^2}{\sum_{i=1}^{n}(Y_i-\bar{Y})^2}$$

ومنه قيمة R^2 هي :

$$R^2=1-\left(\sum_{i=1}^{n}e_i \Big/ \sum_{i=1}^{n}(Y_i-\hat{Y})^2\right)$$

وتعد المعادلة (1) مفيدة جدا لخدمة أغراضنا فيما يتعلق بقياس القدرة التفسيرية للنموذج، لذا من المهم أن نفحص بعناية معنى كل حد من حدودها:

$\sum_{i=1}^{n}(Y_i-\bar{Y})^2$: تعبر عن مجموع مربعات الانحرافات الكلية (TSS) TOTAL SUM OF SQUARES

$\sum_{i=1}^{n}(\hat{Y}_i-\bar{Y})^2$: تمثل مجموع مربعات الانحرافات المشروحة EXPLAINED SUM OF SQUARES (ESS)

$\sum_{i=1}^{n}(e_i)^2$: تمثل مجموع مربعات البواقي RESIDUAL SUM OF SQUARES (RSS)

ومنه نعيد صياغة المعادلة (1) على الشكل الآتي : TSS=ESS+RSS

وبتقسيم كل الأطراف على TSS نجد : 1=(ESS/TSS)+(RSS/TSS) ومنه نعرف معامل التحديد كمايلي : $R^2=r^2=\frac{ESS}{TSS}=1-\frac{RSS}{TSS}$ وهو معامل يقيس ويشرح نسبة الانحرافات الكلية أو التغيرات التي تحدث في التابع Y والمشروحة بواسطة تغيرات المتغير المستقل X .

لما يأخذ R^2 أكبر قيمة وهي 1 أي عندما تقع كل نقاط الملاحظات (Yi , Xi) على الخط المقدر $\hat{Y}_i=\hat{\alpha}+\hat{\beta}X_i$ فالقدرة التفسيرية للنموذج عالية جدا (تامة)، أي هناك جودة في التوفيق والارتباط بين المتغير التابع والمستقل.

أما إذا كان R^2 يأخذ أصغر (أسوء) قيمة له وهي الصفر، فنفسر ذلك بأنه ليس هناك أي جودة في التوفيق والارتباط بين المتغير التابع والمستقل أي ليس للنموذج قدرة تفسيرية على الإطلاق وقد يرجع السبب في ذلك إلى أن العلاقة الموجودة بين المتغيرين هي غير خطية أو إلى غياب السببية بينهما.

نذكر أن الفرق الجوهري بين معامل التحديد ومعامل الارتباط يكمن في السببية حيث يقيس معامل

الارتباط العلاقة بين متغيرين بغض النظر عن الدور الذي يلعبه كل متغير، أما معامل التحديد فيقيس أيضا الارتباط ولكن يأخذ بعين الاعتبار السببية حيث أن المتغير X، هو الذي يشرح الظاهرة .

7- حساب تباين المقدرات :

7-1- حساب تباين مقدرتي معلمتي النموذج:

- تباين الثابت $\hat{\alpha}$: نقوم أولا بحساب المقدار $\hat{\alpha}-\alpha$ أين وجدنا سابقا بأن $\hat{\alpha}=\bar{Y}-\hat{\beta}\bar{X}$ ولدينا كذلك $\bar{Y}=\alpha+\beta\bar{X}+\bar{\varepsilon}$ وبالتعويض بقيمة $\bar{Y}$ في قيمة $\hat{\alpha}$ نجد :

$$\hat{\alpha}-\alpha=\beta\bar{X}-\hat{\beta}\bar{X}+\bar{\varepsilon}=\bar{\varepsilon}-\left(\hat{\beta}-\beta\right)\bar{X}=\sum_{i=1}^{n}\left(\frac{1}{n}-w_i\bar{X}\right)\varepsilon_i=\sum_{i=1}^{n}w_i^*\varepsilon_i$$

وعليه : $(\hat{\alpha}-\alpha)^2=\sum_{i=1}^{n}w_i^{*2}\varepsilon_i^2+2\sum_i\sum_j w_i^*w_j^*\ \varepsilon_i\varepsilon_j$

وبعد إدخال التوقع الرياضي على الطرفين نجد : $E(\hat{\alpha}-\alpha)^2=\sum_{i=1}^{n}w_i^{*2}E(\varepsilon_i^2)+2\sum_i\sum_j w_i^*w_j^*\ E(\varepsilon_i\varepsilon_j)$

مع العلم بأنه لدينا: $\sum_{i=1}^{n}w_i^{*2}=\sum_i\left(\frac{1}{n^2}+w_i^2\bar{X}^2-\frac{2}{n}w_i\bar{X}\right)=\frac{1}{n}+\sum_i w_i^2\bar{X}^2$

$$E\left(\varepsilon_i\varepsilon_j\right)=0 \qquad , \qquad E\left(\varepsilon_i^2\right)=\sigma_\varepsilon^2$$

وبالتالي يصبح لدينا تباين الثابت هو : $\mathrm{var}\left(\hat{\alpha}\right)=\left(\frac{1}{n}+\frac{\bar{X}^2}{\sum_i\left(X_i-\bar{X}^2\right)}\right)\sigma_\varepsilon^2=\left(\frac{\sum_i X_i^2}{n\sum_i\left(X_i-\bar{X}\right)^2}\right)\sigma_\varepsilon^2$

-حساب تباين $\hat{\beta}$:

نعلم بأن :

$$\hat{\beta}=\frac{\sum_{i=1}^{n}(X_i-\bar{X})(Y_i-\bar{Y})}{\sum_{i=1}^{n}(X_i-\bar{X})^2}=\frac{\sum_{i=1}^{n}x_iy_i}{\sum_{i=1}^{n}x_i^2}$$

بحيث:

$$y_i=Y_i-\bar{Y} \qquad ,x_i=X_i-\bar{X}$$

نضع : $w_i=\frac{x_i}{\sum_i x_i^2}$ وبالتالي يمكن كتابة $\hat{\beta}$ كمايلي: $\hat{\beta}=\sum_{i=1}^{n}w_iy_i$

يجب أن نذكر بعض الخصائص الخاصة ب w_i :

$$\sum_i w_i^2=\frac{1}{\sum_i x_i^2} \qquad \sum_i w_ix_i=1 \qquad \sum_i w_i=0$$

لدينا : $\hat{\beta}=\sum_{i=1}^{n} w_i y_i=\sum_i w_i\left(Y_i-\bar{Y}\right)$ كما لدينا $Y_i=\alpha+\beta X_i+\varepsilon_i$ و $\bar{Y}=\alpha+\beta\bar{X}+\bar{\varepsilon}$

بالتعويض نحصل على :

$$\hat{\beta}=\sum_i w_i\left(\alpha+\beta X_i+\varepsilon_i-\alpha-\beta\bar{X}-\bar{\varepsilon}\right)$$
$$=\sum_i w_i\left(\beta X_i+\varepsilon_i-\beta\bar{X}-\bar{\varepsilon}\right)=\sum_i w_i\left[\beta\left(X_i-\bar{X}\right)+\left(\varepsilon_i-\bar{\varepsilon}\right)\right]$$

المقدار $\varepsilon_i-\bar{\varepsilon}$ يؤول إلى ε_i لأن الأمل الرياضي ل $\bar{\varepsilon}$ يساوي الصفر كما لدينا $\sum_i w_i x_i=1$ ومنه

يمكن كتابة $\hat{\beta}$ كمايلي : $\hat{\beta}-\beta=\sum_{i=1}^{n} w_i\varepsilon_i$ أي $\hat{\beta}=\beta+\sum_{i=1}^{n} w_i\varepsilon_i$

وعليه : $(\hat{\beta}-\beta)^2=\left(\sum_{i=1}^{n} w_i\varepsilon_i\right)^2=\sum w_i^2\varepsilon_i^2+2\sum_i\sum_j w_i w_j\varepsilon_i\varepsilon_j$

وبعد إدخال التوقع الرياضي نحصل على :

$$E(\hat{\beta}-\beta)^2=\sum w_i^2 E(\varepsilon_i^2)+2\sum_i\sum_j w_i w_j E(\varepsilon_i\varepsilon_j)$$

بما أن : $E\left(\varepsilon_i^2\right)=\sigma_\varepsilon^2$ و $E\left(\varepsilon_i\varepsilon_j\right)=0$ فإن :

$$\operatorname{var}\left(\hat{\beta}\right)=\sigma_\varepsilon^2\sum w_i^2=\left(\frac{\sigma_\varepsilon^2}{\sum_i\left(X_i-\bar{X}\right)^2}\right)$$

نلاحظ أن تباين كل مقدر غير معروف لأنه يرتبط بتباين الأخطاء النظري σ_ε^2 فينبغي في هده الحالة تقدير تباين الأخطاء للحصول على تباين البواقي .

2-7- حساب تباين الأخطاء (البواقي) σ_ε^2 :

إن تقدير تباين الأخطاء العشوائية هو :

$$\hat{\sigma}_\varepsilon^2=\frac{\sum_{i=1}^{n}\hat{\varepsilon}_i^2}{n-2}$$

القيمة التي في البسط تعبر عن مجموع مربعات البواقي أما n-2 فهي درجة الحرية التي تعبر عن حجم العينة n ناقص 2 وذلك لوجود معلمتين مقدرتين في النموذج الخطي البسيط .

8 -بناء مجال الثقة للمعالم:

بمعرفة توزيع $\hat{\alpha}$ و $\hat{\beta}$ يمكن تكوين مجالات ثقة وإجراء اختبار الفرضيات الموضوعة حول معالم الانحدار α و β على التوالي، نعطي مجالا للقيم التي يمكن أن تحتوي عليها معالم الانحدار الحقيقية، مع كل مجال ثقة نضع مستوى إحصائيا للمعنوية، حيث أن احتمال احتواء المجال المذكور على معلمة الانحدار الحقيقية يكون واحد مطروحا منه مستوى المعنوية، أي ($1-\alpha$) ولتكوين مجال الثقة من التوزيع t بالنسبة للمعلمين α و β نكتب القانون الخاص لكل معلمة:

- **في حالة $n \leq 30$ و σ^2 غير معروف :**

$$\frac{\hat{\alpha}-\alpha}{\hat{\sigma}_{\hat{\alpha}}} \rightarrow t_{n-2}$$

$$\frac{\hat{\beta}-\beta}{\hat{\sigma}_{\hat{\beta}}} \rightarrow t_{n-2}$$

عند مستوى معنوية (α%) يكون مجال الثقة لكلا المعلمتين :

$$pr\left[-t_{n-2,\frac{\alpha}{2}} \leq \frac{\hat{\alpha}-\alpha}{\hat{\sigma}_{\hat{\alpha}}} \leq t_{n-2,\frac{\alpha}{2}}\right]=1-\alpha$$

$$pr\left[-t_{n-2,\frac{\alpha}{2}} \leq \frac{\hat{\beta}-\beta}{\hat{\sigma}_{\hat{\beta}}} \leq t_{n-2,\frac{\alpha}{2}}\right]=1-\alpha$$

إذا ضربنا جميع أطراف الإحتمال الأول ب $\hat{\sigma}_{\hat{\alpha}}$ وأضفنا المقدار α، أما الاحتمال الثاني فنضرب أطرافه ب $\hat{\sigma}_{\hat{\beta}}$ ونضيف β فستحصل في الأخير على فترة الثقة الخاصة بكل معلمة :

$$\alpha \in \left[\hat{\alpha}-\hat{\sigma}_{\hat{\alpha}} t_{(n-2,\frac{\alpha}{2})} ; \hat{\alpha}+\hat{\sigma}_{\hat{\alpha}} t_{(n-2,\frac{\alpha}{2})}\right]$$

$$\beta \in \left[\hat{\beta}-\hat{\sigma}_{\hat{\beta}} t_{(n-2,\frac{\alpha}{2})} , \hat{\beta}+\hat{\sigma}_{\hat{\beta}} t_{(n-2,\frac{\alpha}{2})}\right]$$

- **في حالة حجم العينة $n > 30$ و التباين σ^2 معروف :**

$$\hat{\alpha} \rightarrow N\left(\alpha, \frac{\hat{\sigma}_{\varepsilon}^2}{\sum_i x_i^2}\right) \Rightarrow \frac{\hat{\alpha}-\alpha}{\hat{\sigma}_{\hat{\alpha}}} \rightarrow N(0,1)$$

$$\hat{\beta} \rightarrow N\left(\beta, \frac{\hat{\sigma}_{\varepsilon}^{2}\sum_{i} X_{i}^{2}}{n\sum_{i} x_{i}^{2}}\right) \Rightarrow \frac{\hat{\beta}-\beta}{\hat{\sigma}_{\hat{\beta}}} \rightarrow N(0,1)$$

عند مستوى المعنوية ($\alpha\%$) يكون مجال الثقة لكلا المعلمتين كمايلي:

$$\alpha \in \left[\hat{\alpha}-\hat{\sigma}_{\hat{\alpha}} z_{\frac{\alpha}{2}} ; \hat{\alpha}+\hat{\sigma}_{\hat{\alpha}} z_{\frac{\alpha}{2}}\right]$$

$$\beta \in \left[\hat{\beta}-\hat{\sigma}_{\hat{\beta}} z_{\frac{\alpha}{2}} , \hat{\beta}+\hat{\sigma}_{\hat{\beta}} z_{\frac{\alpha}{2}}\right]$$

مع العلم بأن القيمة $z_{\frac{\alpha}{2}}$ تمثل القيمة الحرجة للتوزيع الطبيعي المعياري وهي القيمة المحسوبة نستخرجها من الجدول مباشرة .

9 – اختبارات المعنوية أو الدلالة:

قد يكون النموذج المبني من طرفنا صحيحا أو غير صحيح ونثبت صحته من خلال اختباره، ويتم ذلك بواسطة فرض معلمة من معالم النموذج تساوي الصفر أو أي عدد آخر، وتسمى فرضية العدم H0 وما دامت العلاقة بين Y و X قائمة على أساس النموذج الخطي، فإن انعدام هذه العلاقة يعني بأن خط انحدار المجتمع هو عبارة عن خط أفقي أي $H_0: \beta = 0$ ، وبما أن الافتراض H0 خاضع للاختبار فإنه لا يكون بالضرورة صحيحا الأمر الذي يتطلب منا وضع فرض بديل $H_1: \beta \neq 0$ وفي حالة معرفة إشارة β مسبقا من النظرية الاقتصادية فإن الافتراض البديل يكون $H_1: \beta \succ 0$ (أو $H_1: \beta \prec 0$)

فإذا أردنا أن نختبر العلاقة بين المتغير المستقل (X) والتغير التابع (Y) وذلك بوضع الفرضية H_0 التي تنص على عدم وجود علاقة بينهما فتكون الفرضية $H1$ عكس H_0 ويكون شكل الاختبار كمايلي :

$$\begin{cases} H_0: \beta = 0 \\ H_1: \beta \neq 0 \end{cases}$$

ولاختبار صحة إحدى الفرضيتين السابقتين نستعمل إختبار ستودنت (T) أو إختبار فيشر (F).

1- **اختبار ستودنت** (T): ويتم هذا الاختبار بحساب الإحصائية التالية: $T_c = \left|\frac{\hat{\beta}-B}{\delta_{\hat{\beta}}}\right|$

حيث $\delta_{\hat{\beta}}$: الانحراف المعياري للمقدرة $\hat{\beta}$ ، وبما أن الفرضية H_0 تنص على إنعدام B فإن قيمة (T) تصبح:

$$T_c = \left|\frac{\hat{\beta}}{\delta_{\hat{\beta}}}\right|$$

ويتم قبول أو رفض H_0 بمستوى معنوية معين ($\alpha\%$) بمقارنة قيمة (T) المحصل عليها مع القيمة المجدولة عند درجة الحرية $(N-2)$، حيث: 2 هو عدد المقدرات في هذه الحالة، و N هو عدد المشاهدات، وقرار هذا الاختبار يكون كالآتي:

$T_c \rangle T_t$: فإن نرفض H_0 : إذن $\hat{\beta} \neq 0$ ومنه المتغير له معنى (تأثير) في النموذج أي أنه معنوي.

$T_c \langle T_t$: فإننا نقبل H_0 : إذن $\hat{\beta} = 0$ ومنه $\hat{\beta}$ ليس معنوي أي أن المتغير المفسر له دور في النموذج.

حيث T_t المجدولة تمثل القيم الحرجة وتحدد المنطقة الحرجة للاختبار ذو الطرفين عند درجة الحرية $(N-2)$ وبمستوى معنوية محدد ونكتب $t_{n-2;\frac{\alpha}{2}}$.

ملاحظة: عندما يكون حجم العينة كبيرا $(n>30)$ فينبغي استعمال التوزيع الطبيعي ويمكن أخذ القيمة الحرجة $z_{\alpha/2}$ وذلك بحساب المساحة المظلة للتوزيع الطبيعي .

2- اختبار فيشر (F): يوضح لنا هذا الاختبار دلالة النموذج بصورة عامة، وكذلك حساب نسبة الانحرافات الموضحة إلى الانحرافات غير الموضحة بواسطة النموذج:

$$\begin{cases} H_0 : \alpha = \beta = 0 \\ H_1 : \alpha \neq 0 \text{ أو } \beta \neq 0 \end{cases}$$

- شكل الاختبار: ويتم الاختبار بحساب الإحصائية:

$$F_c = \frac{\sum_{i=1}^{n} (\hat{Y}_i - \bar{Y}) \Big/ 1}{\sum_{i=1}^{n} e_i^2 \Big/ n-2}$$

حيث n: هو عدد المشاهدات

نقوم بمقارنة القيمة F_c مع القيمة (F_t) عند درجة الحرية $(1,n-2)$ بمستوى معنوية α .

- قرار الاختبار:

إذا كان $F_t \langle F_c$ فإننا نرفض H_0 : أي أن المتغيرات x تؤثر (أي تفسر) y.

إذا كان $F_t \rangle F_c$ فإننا نقبل H_0 : أي أن المتغيرات x لا يؤثر (أي تفسر) y.

10 -استعمال برنامج Eviews9.0 في تطبيقات لمعادلة الانحدار الخطي البسيط :

سنعتمد على المثال الآتي من أجل معرفة تقدير النماذج الخطية البسيطة وذلك بإختبارها يدويا أولا، ثم باستعمال برنامج Eviews9.0

مثال 01 : لدراسة أثر سعر الفائدة (X) كمتغير مستقل على إدخار القطاع العائلي (Y) كمتغير تابع في دولة ما تم تجميع البيانات الخاصة بالمتغيرين خلال الفترة 2008 – 2017 كما هو موضح بالجدول الموالي:

السنة	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017
سعر الفائدة	2	3	5	4	3	5	7	6	7	8
الإدخار	20	28	40	45	37	52	54	43	65	56

المطلوب :

1 -رسم بيانات الجدول في شكل إنتشاري .

2 -تقدير معادلة الانحدار الخطي البسيط بين المتغيرين باعتبار العلاقة بينهما خطية باستخدام طريقة المربعات الصغرى OLS يدويا ثم عن طريق برنامج Eviews.

3 -قدم تفسيرا إقتصاديا لمعلمة النموذج المتوصل إليها .

4 -أحسب مختلف قيم حد الخطأ $e_i = \hat{\varepsilon}_i$ مقدر الخطأ العشوائي ε_i

5 -إيجاد تباين البواقي والانحراف المعياري للمقدرات .

6 -اختبار عند مستوى معنوية 5% معنوية المقدرات.

7 -إختبر المعنوية الإحصائية الكلية للنموذج عند مستوى معنوية 5%

8 -إيجاد معامل التحديد للنموذج المقدر ثم اشرح النتيجة .

-الحل باستخدام الطريقة الحسابية اليدوية :

1 **-تقدير النموذج الخطي البسيط:** نقوم أولا بحساب المجاميع التي يوضحها الجدول الموالي :

السنوات	X	Y	XY	X^2	Y^2
2008	2	20	40	4	400
2009	3	28	84	9	784
2010	5	40	200	25	1600
2011	4	45	180	16	2025
2012	3	37	111	9	1369
2013	5	52	260	25	2704
2014	7	54	378	49	2916
2015	6	43	258	36	1849
2016	7	65	455	49	4225

2017	8	56	448	64	3136
المجموع	50	440	2414	286	21008

يمكن حساب المتوسط الحسابي للمتغيرين :

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{10} X_i = \frac{50}{10} = 5 \qquad \bar{y} = \frac{1}{n}\sum_{i=1}^{10} Y_i = \frac{440}{10} = 44$$

لدينا العلاقة بين المتغيرين هي من الشكل: $Y_i = \alpha + \beta X_i + \varepsilon_i$ وتقديرها يكون من الشكل الموالي:

$\hat{Y}_i = \hat{\alpha} + \hat{\beta} X_i$ ، ومنه يمكن إيجاد قيمة المعلمة B كمايلي :

$$\hat{\beta} = \frac{\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})}{\sum_{i=1}^{n}(X_i - \bar{X})^2} = \frac{\sum_{i=1}^{n} X_i Y_i - n\bar{X}\bar{Y}}{\sum_{i=1}^{n} X_i^2 - n\bar{X}^2} = \frac{2414 - (10)(44)(5)}{286 - 10.(5)^2}$$

$$\Rightarrow \hat{\beta} = 5.94$$

$$\hat{\alpha} = \bar{Y} - \hat{\beta}\bar{X} \Rightarrow \hat{\alpha} = 44 - 5.94(5) \Rightarrow \hat{\alpha} = 14.3$$

إذن المعادلة المقدرة تكتب كمايلي : $\hat{Y}_i = 14.3 + 5.94 X_i$

2 –التفسير الإقتصادي لمعلمة النموذج:

النموذج المتوصل إليه هو $\hat{Y}_i = 14.3 + 5.94 X_i$ وبغض النظر عن معنوية النموذج من عدمها فإن القيمة 5.94 تعني بأنه إذا تغير المتغير المستقل (إدخار العائلات) بوحدة واحدة فإن المتغير التابع (سعر الفائدة) سيتغير بمقدار 5.94 وحدة وفي نفس الاتجاه أي العلاقة بين المتغيرين هي علاقة طردية.

3 –حساب مختلف قيم حد الخطأ $e_i = \hat{\varepsilon}_i$:

لدينا البواقي تكتب من الشكل : $e_i = (Y_i - \hat{Y}_i)$

السنوات	X	Y	$\hat{Y}_i = 14.3 + 5.94 X_i$	$e_i = (Y_i - \hat{Y}_i)$	e_i^2
2008	2	20	26.18	-6.18	38.19
2009	3	28	32.12	-4.12	16.97
2010	5	40	44	-4	16
2011	4	45	38.06	6.94	48.16
2012	3	37	32.12	4.88	23.81

2013	5	52	44	8	64
2014	7	54	55.88	-1.88	3.53
2015	6	43	49.94	-6.94	48.16
2016	7	65	55.88	9.12	83.17
2017	8	56	61.82	-5.82	33.87
المجموع	50	440	/	/	375.86

3- تقدير تباين حد الخطأ العشوائي :

لدينا:

$$\hat{\sigma}_{\varepsilon}^{2}=\frac{\sum_{i=1}^{n}\hat{\varepsilon}_{i}^{2}}{(n-2)}=\frac{\sum_{i=1}^{n}\left(Y_{i}-\hat{Y}_{i}\right)}{(n-2)}$$

ومنه فإن مقدر تباين حد الخطأ يساوي :

$$\hat{\sigma}_{\varepsilon}^{2}=\frac{\sum_{i=1}^{n}\hat{\varepsilon}_{i}^{2}}{(n-2)}=\frac{375.86}{8}=46.98$$

4- إختبار معنوية معلمتي النموذج المقدر :

نقوم باختبار الفرضيتين الآتيتين:

فرض العدم :سعر الفائدة ليس له أثر معنوي على إدخار العائلات $H_{0}:\beta=0$

فرض البديل: سعر الفائدة له أثر معنوي على إدخار العائلات $H_{1}:\beta\neq 0$

وللقيام بهذا الاختبار نقوم بحساب إحصائية ستيودنت t بحيث :

$$T_{c}=\left|\frac{\hat{\beta}}{\hat{\sigma}_{\hat{\beta}}}\right|$$

بحيث أن :

$$\hat{\sigma}_{\hat{\beta}}^{2}=\frac{\hat{\sigma}_{\varepsilon}^{2}}{\sum_{i=1}^{n}(X_{i}-\bar{X})^{2}}$$

السنوات	X	$X_i-\bar{X}$	$(X_i-\bar{X})^2$
2008	2	-3	9
2009	3	-2	4

2010	5	0	0
2011	4	-1	1
2012	3	-2	4
2013	5	0	0
2014	7	2	4
2015	6	1	1
2016	7	2	4
2017	8	3	9
المجموع	50	/	36

لدينا : $\hat{\sigma}_{\hat{\beta}}^2 = \dfrac{\hat{\sigma}_e^2}{\sum_{i=1}^{n}(x_i-\bar{x})^2} = \dfrac{46.98}{36} = 1.3$ ، $\hat{\sigma}_{\hat{\beta}} = \sqrt{1.3} = 1.14$

إذن يصبح $T_c = \left|\dfrac{\hat{\beta}}{\hat{\sigma}_{\hat{\beta}}}\right| = \dfrac{5.94}{1.14} = 5.21$

11 - إختبار المعنوية الكلية للنموذج :

نقوم باختبار الفرضيتين الآتيتين:

فرض العدم :النموذج الحالي غير مناسب لتمثيل العلاقة بين سعر الفائدة كمتغير تابع وإدخار العائلات كمتغير مستقل ونعبر عنه بالصيغة الآتية: $H_0 : \alpha = \beta = 0$

فرض البديل: النموذج الحالي مناسب لتمثيل العلاقة بين سعر الفائدة كمتغير تابع وإدخار العائلات كمتغير مستقل ونعبر عنه بالصيغة الآتية: $H_1 : \alpha \neq 0$ أو/ $\beta \neq 0$

ولإجراء هذا الاختبار نقوم بحساب إحصائية فيشر F وفق العلاقة الآتية:

$$F_c = \frac{\sum_{i=1}^{n}(\hat{Y}_i-\bar{Y})^2 / 1}{\sum_{i=1}^{n} e_i^2 / n-2}$$

نقوم بحساب المقدار $\sum_i(\hat{Y}_i-\bar{Y})^2$ أولا كما في الجدول :

السنوات	Y	$Y_i-\bar{Y}$	$(Y_i-\bar{Y})^2$	$\hat{Y}_i$	$(\hat{Y}_i-\bar{Y})$	$(\hat{Y}_i-\bar{Y})^2$	e_i^2
2008	20	-24	576	26.18	-17.82	317.55	38.19
2009	28	-16	256	32.12	-11.88	141.13	16.97

2010	40	-4	16	44	0	0	16
2011	45	1	1	38.06	-5.94	35.28	48.16
2012	37	-7	49	32.12	-11.88	141.13	23.81
2013	52	8	64	44	0	0	64
2014	54	10	100	55.88	11.88	141.13	3.53
2015	43	-1	1	49.94	5.94	35.28	48.16
2016	65	21	441	55.88	11.88	141.13	83.17
2017	56	12	144	61.82	17.82	317.55	33.87
المجموع	440	/	1648	/	/	1270.18	375.86

$$F_c = \frac{\sum_{i=1}^{n}(\hat{Y}_i - \bar{Y})^2 / 1}{\sum_{i=1}^{n} e_i^2 / n-2} = \frac{1270.18/1}{375.86/8} = 27.03$$

لدينا قيمة فيشر المجدولة عند درجة حرية F(1،8) وبمستوى معنوية 5% تساوي : 5.32

بمأن قيمة فيشر المحسوبة أكبر من المجدولة فإننا نقبل بالفرضية البديلة التي تنص على وجود العلاقة مابين كل من سعر الفائدة وإدخار العائلات ويستدل من هذا أن النموذج المقترح مناسب لتمثيل العلاقة بين المتغيرين .

12 - حساب معامل التحديد R^2 :

$$R^2 = 1 - \left(\sum_{i=1}^{n} e_i^2 \Big/ \sum_{i=1}^{n}(Y_i - \bar{Y})^2 \right)$$

$$R^2 = 1 - \frac{RSS}{TSS} = \frac{ESS}{TSS}$$

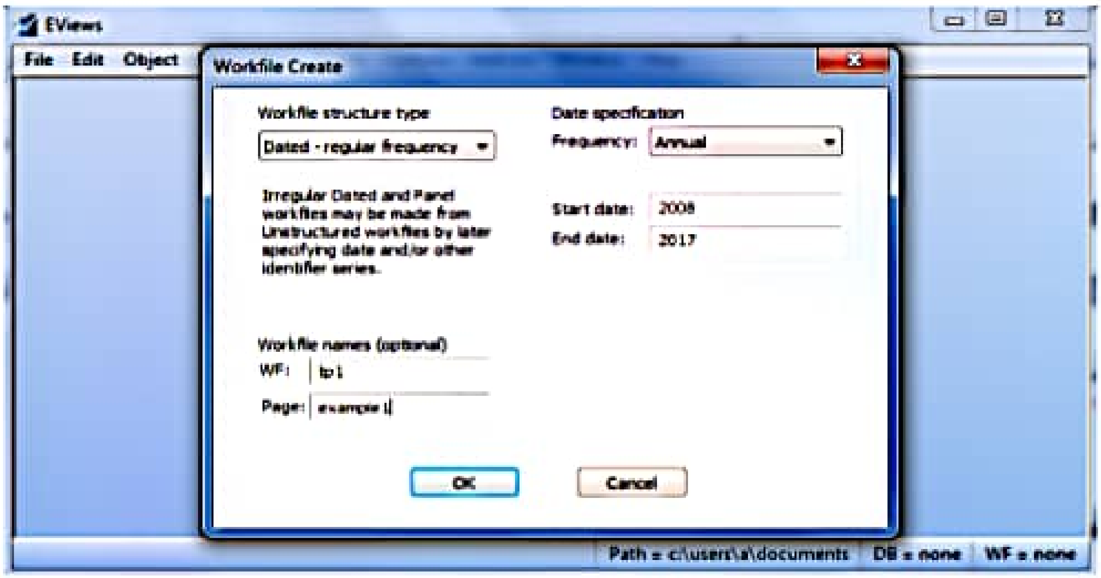
$$R^2 = \frac{\sum(\hat{Y}_i - \bar{Y})^2}{\sum(Y_i - \bar{Y})^2} = \frac{1270.18}{1648} = 0.7707$$

يمكن تفسير هذه النتيجة بأن 77.07 % من التغيرات التي تحدث في المتغير التابع المتمثل في سعر الفائدة ناتجة من مجموع التغيرات التي تحدث في معدلات إدخار العائلات ، أما النسبة المتبقية (22.93%) ترجع لعوامل أخرى (أخطاء عشوائية) أي أن للنموذج ذو قوة تفسيرية عالية.

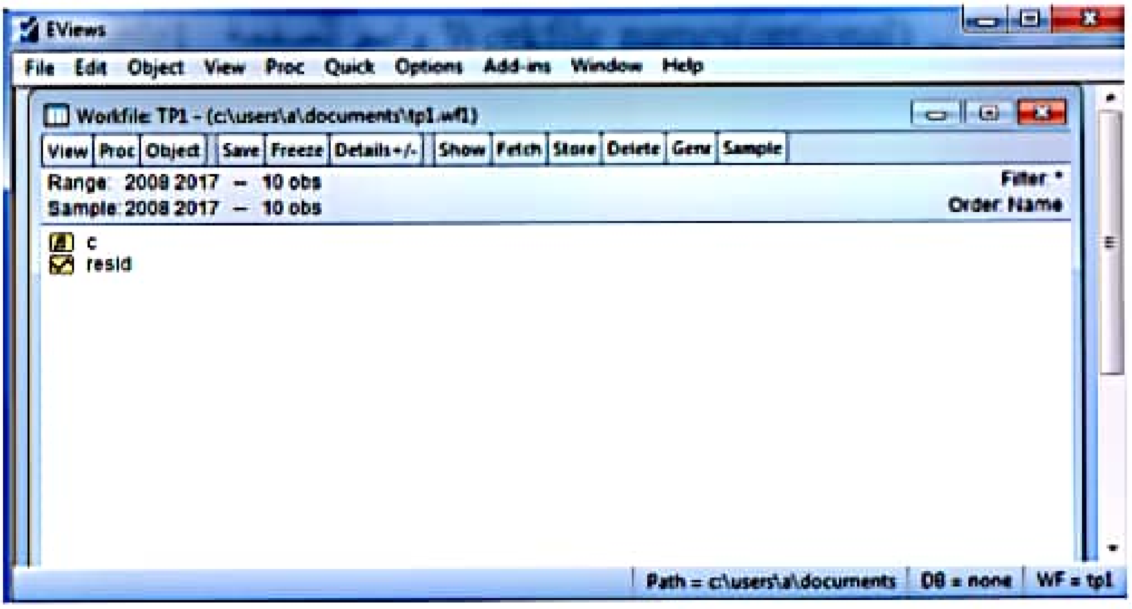
-الحل باستخدام برنامج Eviews :

قبل تقدير العلاقة بين متغيرتي النموذج نقوم أولا وكما تطرقنا في الفصل الأول بإنشاء ملف جديد بالبرنامج أين ندخل بيانات متغيرتي الدراسة وذلك بعد فتح البرنامج على الجهاز ومن خلال القائمة الرئيسية ننفذ الأوامر وفق الخطوات الآتية:

File ⟵ new ⟵ workfile أو بالضغط على الزر CTRL+N تظهر لنا شاشة من أجل تحديد نوعية البيانات، وهي عبارة عن سلاسل زمنية سنوية Annual لنكتب تاريخ بداية السلسلة 2008 مقابل Start date وتاريخ نهاية السلسلة 2017 مقابل End date في مثالنا كما نقوم بتسمية الملف مثلا TP1 من خلال Workfile names(optional) واسم للصفحة example1 كما في الشكل الموالي :

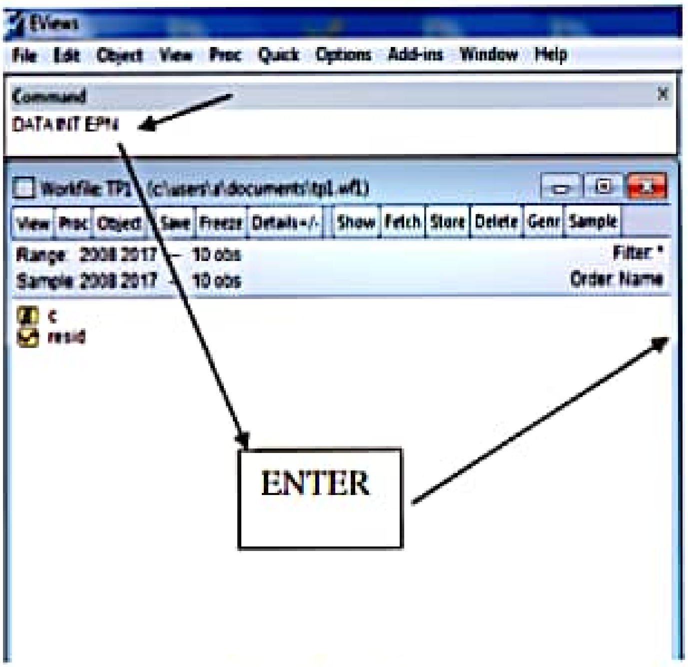


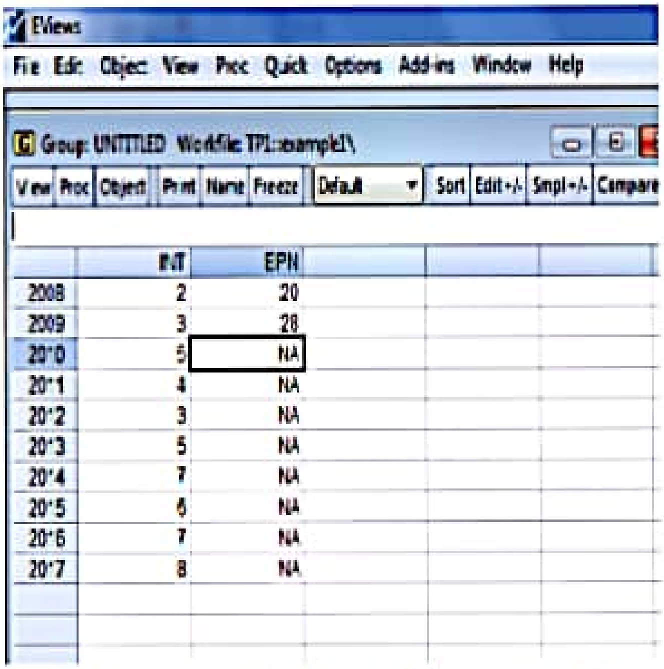
ثم نضغط على الأمر OK لنتحصل على مخرجات كما بالشكل الآتي :



لإدخال البيانات التي قمنا بتحديد مداها ونوعها نقوم بتسمية متغيرات الدراسة وتفريغ بيانات كل سلسلة موافقة للإسم المعطى لها (X, Y) بإتباع إحدى الطرق الثلاثة كما جاء في الفصل الأول وسنختار نحن إحدى تلك الطرق كمايلي:

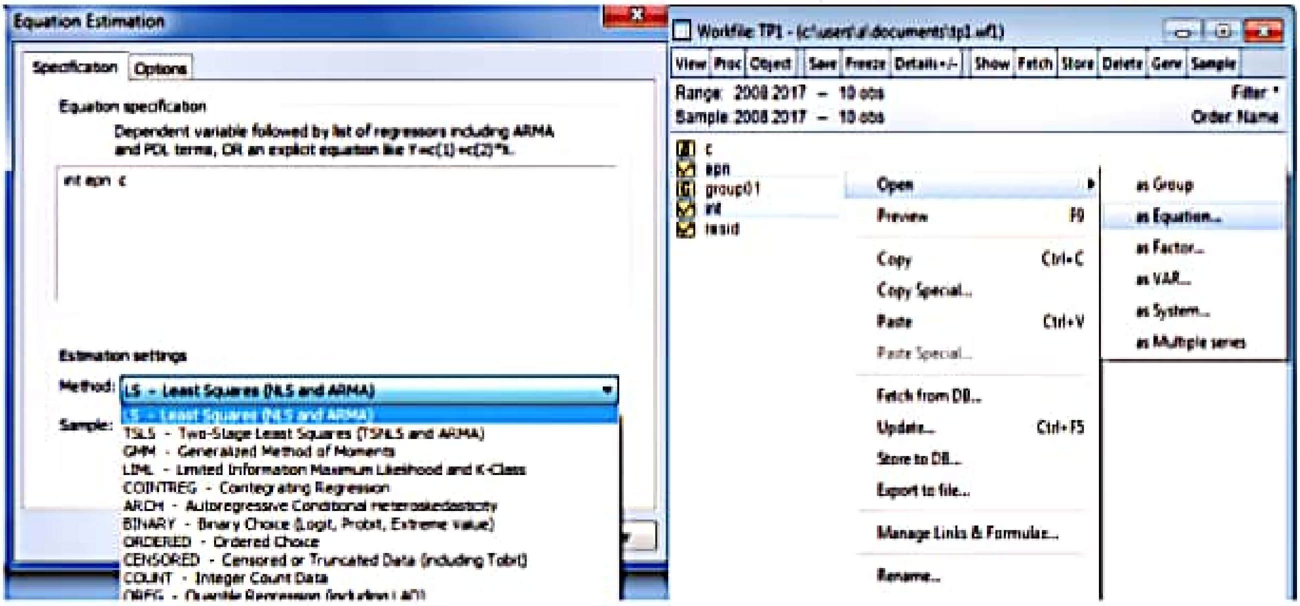
نكتب في الفراغ التي تحت شريط القوائم (نافذة الأوامر) أمر DATA ثم نكتب اسم متغيرات الدراسة أين نترك فراغ بين كل اسم واسم آخر:(DATA INT EPN) بحيث INT تعبر عن سعر الفائدة و EPN تعبر عن الإدخار الخاص بالعائلات، ثم نضغط على ENTER في لوحة المفاتيح ونملأ بياناتنا كما في الشكل الموالي:



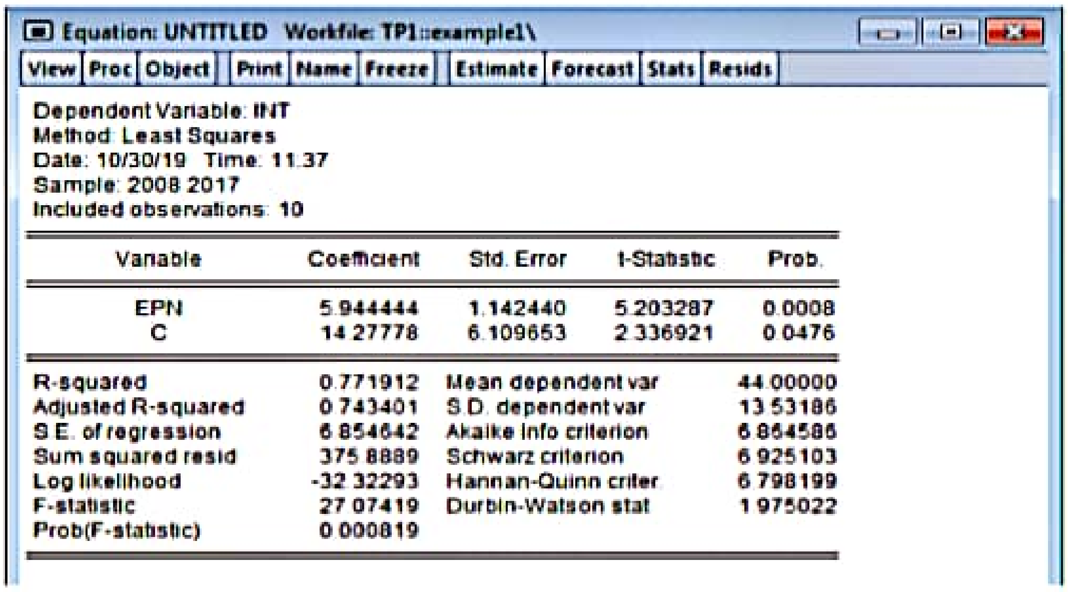


بعد الإنتهاء من ملأ جميع البيانات نقوم بحفظ هذه المجموعة (Groupe) بالضغط على Name وإعطاء إسم لها مثلا Groupe1 فتسجل على واجهة صفحة الـ Workfile ثم تسجيل خروج ، سوف تظهر

السلسلتين EPN,INT في الصفحة Workfile TP1 نقف بمؤشر الماوس أولا على السلسلة التي تمثل التابع INT ونضغط باليسار ليتم التظليل ونقف على السلسلة EPN أيضا وذلك على الترتيب، وبدون تغير مكان الماوس نضغط باليمين ليظهر قائمة نختار منها OPEN ثم AS EQUATION نضغط عليها لتحصل على صندوق بعنوان Specification Equation في أسفل الإطار نختار طريقة التقدير Least Squares (طريقة المربعات الصغرى) من Method كمايلي :



ثم نضغط OK لتظهر نتيجة التقدير في جدول على الشكل التالي:


Equation: UNTITLED Workfile: TP1::example1\

View | Proc | Object | Print | Name | Freeze | Estimate | Forecast | Stats | Resids

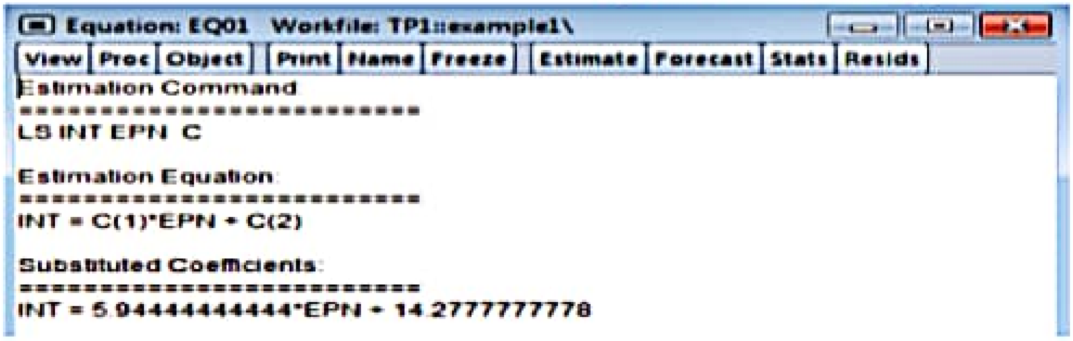
Dependent Variable: INT
Method: Least Squares
Date: 10/30/19 Time: 11:37
Sample: 2008 2017
Included observations: 10

Variable	Coefficient	Std. Error	t-Statistic	Prob.
EPN	5.944444	1.142440	5.203287	0.0008
C	14.27778	6.109653	2.336921	0.0476

R-squared	0.771912	Mean dependent var	44.00000
Adjusted R-squared	0.743401	S.D. dependent var	13.53186
S.E. of regression	6.854642	Akaike info criterion	6.864586
Sum squared resid	375.8889	Schwarz criterion	6.925103
Log likelihood	-32.32293	Hannan-Quinn criter.	6.798199
F-statistic	27.07419	Durbin-Watson stat	1.975022
Prob(F-statistic)	0.000819		

نحتفظ بالمعادلة المقدرة بالضغط على الزر name ونختار اسما مثلا eq1 ثم نضغط على OK فتظهر داخل صفحة Workfile TP1 .

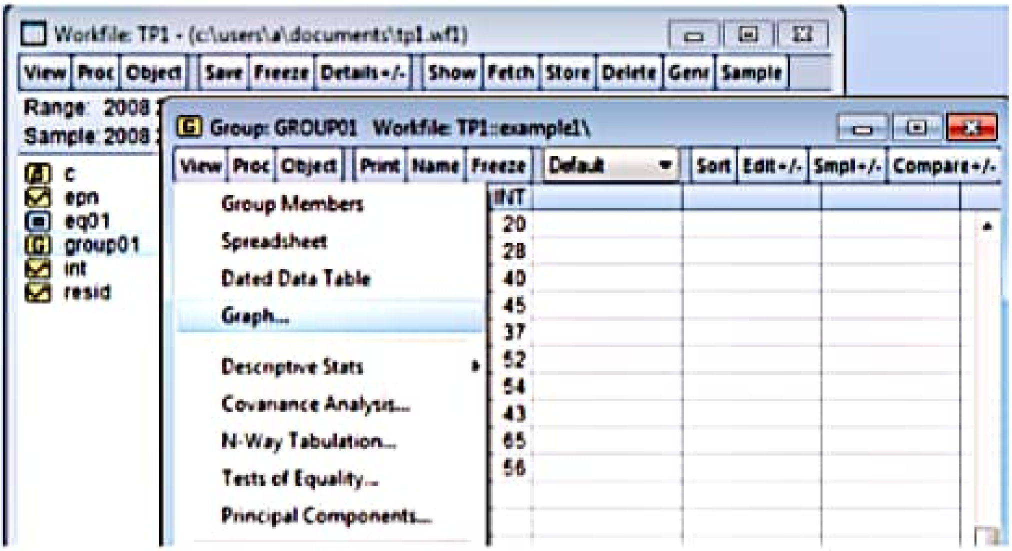
وللحصول على التمثيل الرياضي للعلاقة السابقة نختار من الأمر View الأمر Representations فنتحصل على نموذج الانحدار المقدر كمايلي :



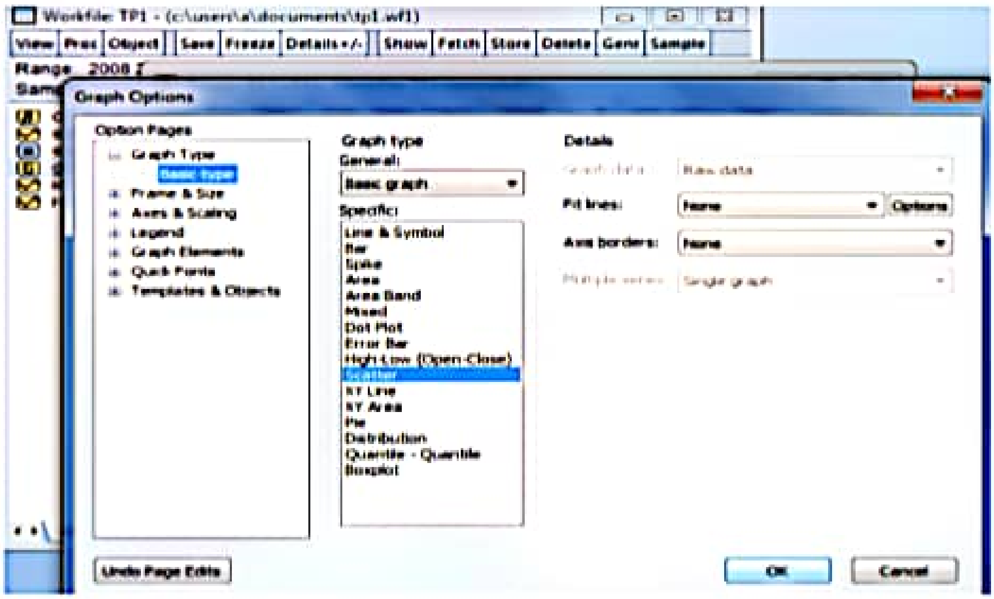
نلاحظ بأننا تحصلنا على نفس النتائج وذلك باستعمال الحل اليدوي بحيث تمثل 5.94 القيمة المقدرة لمعلمة النموذج والتي تعبر عن مقدار التغير في INT لما يتغير EPN بوحدة واحدة وهي قيمة موجبة لتدل بأن العلاقة بين المتغيرين هي علاقة طردية، أما 14.27 فتعبر عن القيمة المقدرة للثابت وهو حجم INT لما تنعدم قيمة المتغير EPN أي في ظل عدم وجود أي إدخار عائلي .

للحصول على جواب السؤال الأول وهو الشكل الانتشاري لبيانات الجدول نتبع الخطوات الآتية :

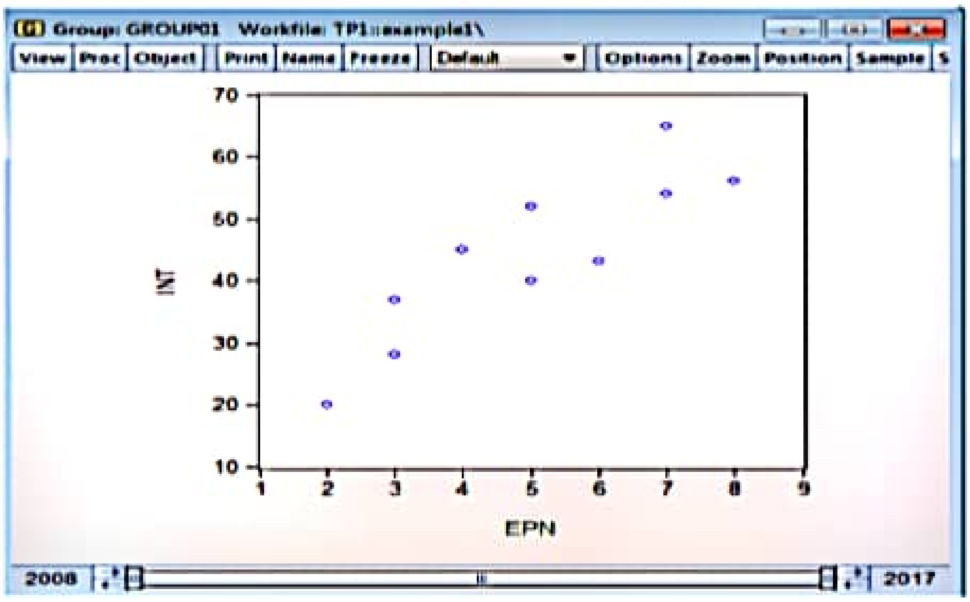
نقوم بفتح Group01 من صفحة Workfile TP1 ثم نختار الأمر Graph من View كما بالشكل الموالي:



ليظهر مربع حواري آخر نقوم من خلاله باختيار الشكل المناسب أين سنختار نحن الأمر Scatter أي الشكل الانتشاري لبيانات الدراسة :



ثم بعدها نضغط على ok لنتحصل على الشكل الموالي:

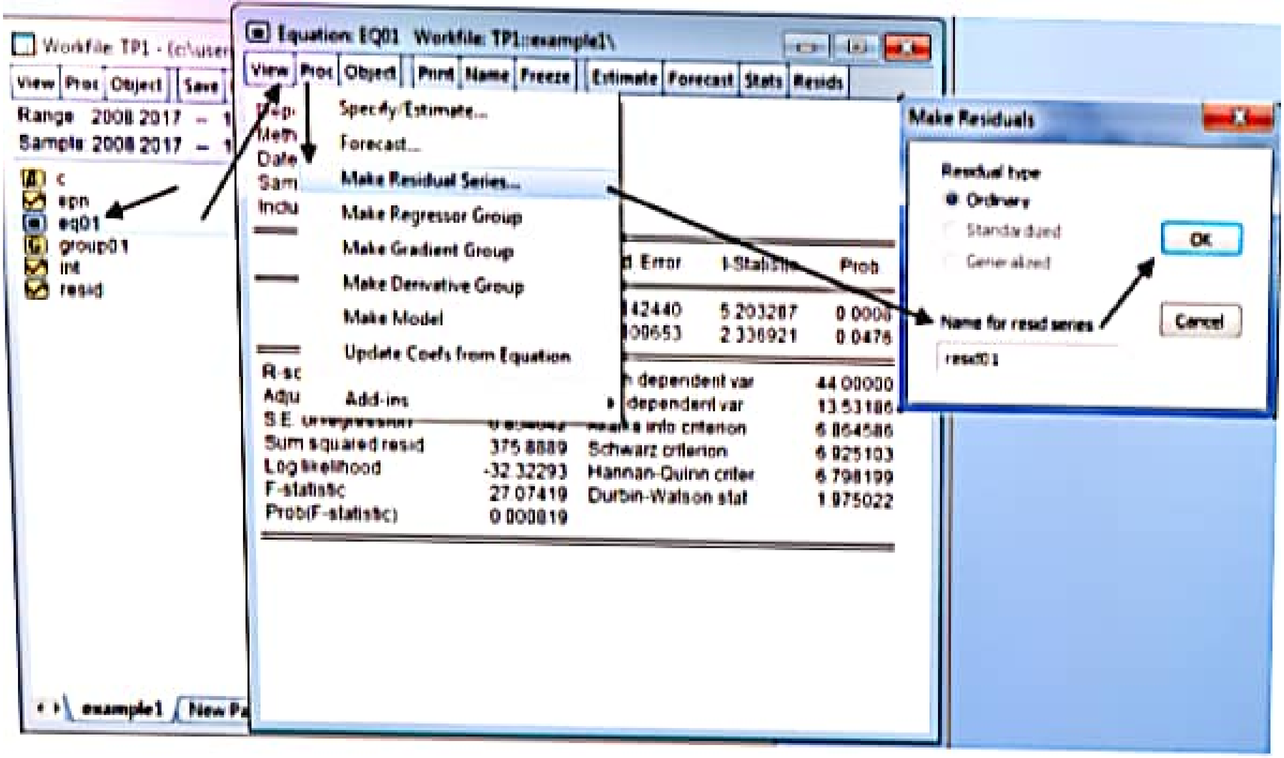


إن الهدف من التمثيل البياني لمختلف قيم متغيرات الدراسة هو توضيح نوع العلاقة واتجاهها، أين يتضح من خلال السحابة النقطية بين المتغيرين بأن العلاقة بينهما هي علاقة خطية موجبة.

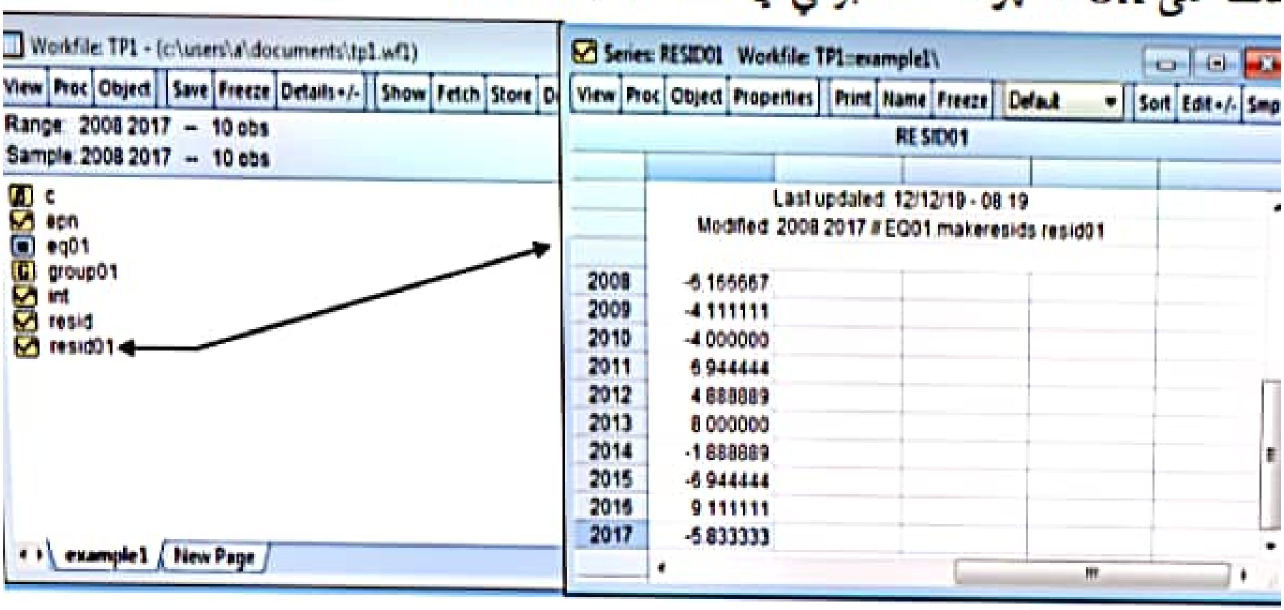
4 ــ حساب مختلف قيم حد الخطأ المقدر $\hat{\varepsilon}_i$:

نتحصل على مختلف قيم حد الخطأ المقدر $\hat{\varepsilon}_i$ عن طريق برنامج Eviews بإتباع الخطوات الآتية:

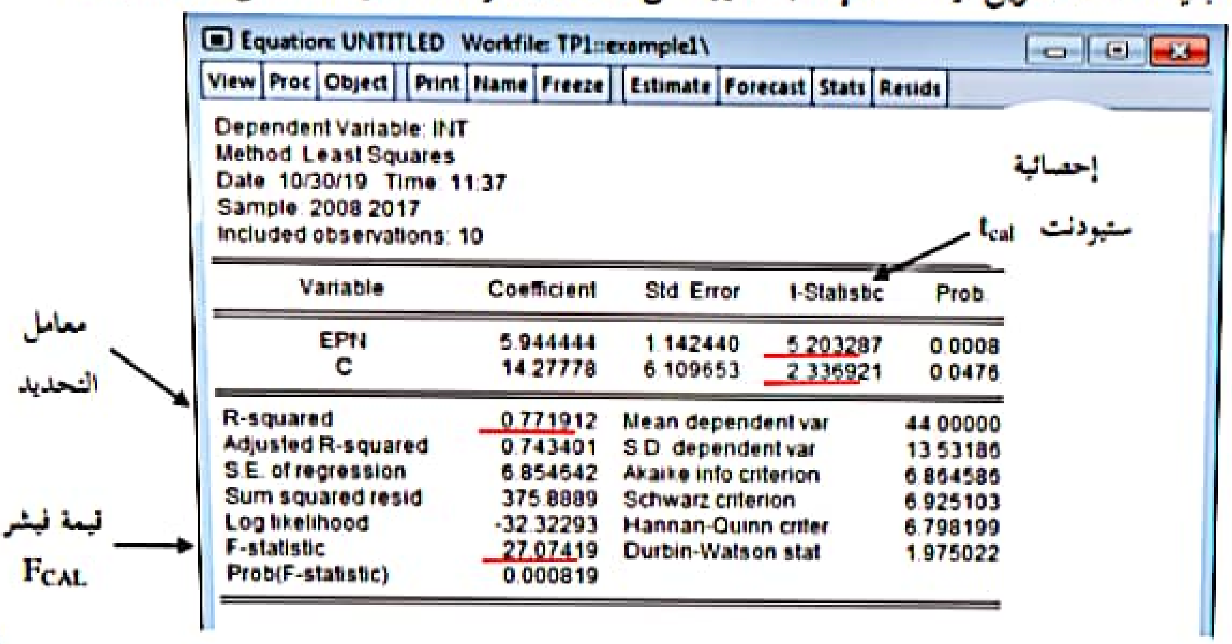
نقوم بفتح eq01 (المعادلة المقدرة المتحصل عليها والتي تم الإحتفاظ بها سابقا) من صفحة Workfile ثم نختار من Proc الأمر Make residual series لنتحصل على مربع حواري نقوم من خلاله بتسمية سلسلة البواقي في خانة Name for resid series كما في الشكل الموالي :



ـغط على OK فتظهر سلسلة البواقي في صفحة الـ Workfile



ما بقية أسئلة التمرين فيمكننا الإجابة عليها من خلال جدول التقدير المتحصل عليه سابقا :



الجدول أعلاه يوضح مختلف القيم الإحصائية التي تساعدنا على الحكم على النموذج المقدر هل هو مقبول إحصائيا أم لا، مثل قيمة معامل التحديد وقيمة إحصائية فيشر بالإضافة إلى إحصائية ستيودنت كما يحتوي الجدول على قيمة إحصائية دربن واتسون التي سنتكلم عنها لاحقا ، كما يمكننا الحكم على معنوية معلمتي النموذج من خلال قيمة إحتمال المقابل لكل معلمة بحيث إذا كانت قيمة الإحتمال أقل من 0.05 (مستوى المعنوية هنا 5%) فإننا نرفض الفرضية الصفرية ونقبل بالفرض البديل الذي ينص على المعنوية الإحصائية للمعلمة وهي النتيجة المتحصل عليها في مثالنا هذا (0.0008 أقل من 0.05 هذا بالنسبة لمعلمة النموذج أما بالنسبة للثابت فلدينا 0.04 أقل من 0.05)، ويمكن الحكم كذلك على المعنوية الكلية للنموذج من خلال إحتمال إحصائية فيشر كما بالجدول Prob(F-statistic)=0.00819 وهي قيمة أقل من 0.05 إذن النموذج ككل معنوي .

مثال 02: قام أحد المحللين الماليين بحساب الانحرافات المعيارية لمعدلات العائد والتي تمثل درجة المخاطرة (s) ،الخاصة بـ 12 محفظة مالية ذات أحجام مختلفة من الأوراق المالية التي تمثل درجة التنوع (v) خلال سنة، والجدول الموالي يبين النتائج المتحصل عليها :

درجة التنوع (v)	درجة المخاطرة (s)
1	30
2	20
3	15
4	12
5	8.5
6	7.5
7	7
8	6.75
9	6.5
10	6.5
11	6.25
12	6

المطلوب :

1 قدر العلاقة ما بين درجة التنوع ودرجة المخاطرة .

2- بكم تتغير درجة المخاطرة عندما ترتفع درجة التنوع بمقدار ورقتين ماليتين ؟

3- ماذا تعني قيمة $\hat{\alpha}$ و $\hat{\beta}$ إقتصاديا ؟ حيث أن $\hat{\alpha}$:هي الثابت و $\hat{\beta}$:هي الميل.

4- استنتج معادلة حد الخطأ e_i أي $\hat{\varepsilon}_i$ ، ثم أحسب e_5 واشرح ماذا تعني .

5-أحسب الخطأ المعياري $\hat{\sigma}$ لمقدرتي المربعات الصغرى .

6-أوجد قيمة معامل التحديد R^2 ، ماذا تعني لك تلك القيمة ؟

7 - اختبر مدى المعنوية الإحصائية لميل معادلة الانحدار عند مستوى المعنوية $\alpha = 5\%$ مع العلم بأن قيمة $t_{tab} = 2.228$

الحل باستخدام برنامج Eviews :

نقوم أولا وكما تطرقنا في المثال السابق بإنشاء ملف جديد بالبرنامج أين ندخل بيانات متغيرتي الدراسة بحيث نختار بيانات غير مؤرخة unstructured/undated وهي عبارة عن مشاهدات فقط بدون تأريخ أين نحدد عدد المشاهدات (أنظر الفصل الأول) بدلا من Dated-regular frequency ، وهنا يجب أولا تحديد المتغير التابع من المتغير المستقل وفي مثالنا نجد بأن درجة المخاطرة تختلف باختلاف درجة التنوع أي أن S متغير تابع و V متغير مستقل وفي الأخير تحصلنا على الجدول الموالي وهذا باستعمال طريقة المربعات الصغرى في عملية التقدير :

Equation: UNTITLED Workfile: TP2::Untitled\

View | Proc | Object | Print | Name | Freeze | Estimate | Forecast | Stats | Resids

Dependent Variable: S
Method: Least Squares
Date: 11/01/19 Time: 17:49
Sample: 1 12
Included observations: 12

Variable	Coefficient	Std. Error	t-Statistic	Prob.
V	-1.680070	0.370209	-4.538166	0.0011
C	21.92045	2.724664	8.045195	0.0000

R-squared	0.673149	Mean dependent var	11.00000
Adjusted R-squared	0.640464	S.D. dependent var	7.383181
S.E. of regression	4.427055	Akaike info criterion	5.964358
Sum squared resid	195.9882	Schwarz criterion	6.045176
Log likelihood	-33.78615	Hannan-Quinn criter.	5.934437
F-statistic	20.59495	Durbin-Watson stat	0.500824
Prob(F-statistic)	0.001078		

1 - من خلال الجدول نستخرج المعادلة المقدرة التي توضح مدى تأثير درجة التنوع في درجة المخاطرة :

$$\hat{S}_i = 21.92 - 1.68V_i$$

2 - بكم تتغير درجة المخاطرة عندما ترتفع درجة التنوع بمقدار ورقتين ماليتين ؟