المحاضرة السابعة

التنبؤ والتقدير بفترة للقيمة الحقيقية ومتوسط القيمة الحقيقية

ملاحظة: يتبع المحاضرة السابعة مثال شامل لمحتويات الفصل الاول (الانحدار الخطي البسيط)

بما ان: $t=\frac{(\hat{\beta}-\beta)\sqrt{\Sigma x^2}}{\delta_u}$

عند: β=0 فان $t=\frac{\beta\sqrt{\Sigma x^2}}{\delta_u}$ اي ان t الاخير هو اختبار t عند β=0: H_0 و بالتالي فان التابع الاختياري يتحول الى التابع الاختباري لفرضية اللاعلاقة و هو $t=\frac{\beta\sqrt{\Sigma x^2}}{\delta_u}$ و هو اختبار فرضية للاعلاقة (معامل الانحدار) كما يدل ان اختيار معامل الارتباط يتطابق تماما مع معامل الانحدار اي عند اختيار معامل الانحدار النتيجة تنطبق تماما على معامل الارتباط لان $t=\frac{\beta\sqrt{\Sigma x^2}}{\delta_u}=\frac{r\sqrt{n-2}}{\sqrt{1-r^2}}$

النتيجة الثانية: اذا العبارة الاخيرة لـ t فانها تعطي t^2 التي هي متماثلة تماما مع اختيار تحليل التباين $t^2=\beta^2\frac{\Sigma x^2}{\delta_u^2}=\frac{SCE}{SCR/n-2}=F$ اي مربع اختيار t^2 يتحول الى اختيار فيشر لتحليل التباين.

اذن نلاحظ ان الاختيارات الثلاثة متطابقة تماما:

اختيار معامل الانحدار (استيودنت)، اختيار معامل الارتباط(استيودنت)، اختيار تحليل التباين (فيشر).

التنبؤ في نموذج الانحدار البسيط la prévision dans le modèle de régression simple

عندما يتم تقدير معاملات نموذج الانحدار البسيط يمكن حساب التنبؤ في افاق

لنفرض ان النموذج البسيط المقدر خلال الفترة t :1,2,3,-----n

$$\hat{y}_t=\hat{\alpha}+\beta x_t$$

فاذا علمنا قيمة المتغير المستقل خلال الفترة (n+1) اي x_{n+1} فان التنبؤ يعطى بـ

$$\hat{y}_{n+1}=\hat{\alpha}+\beta x_{n+1}$$

نستطيع ان نبين بان هذا التنبؤ غير متحيز

خطا التنبؤ: $e_{n+1}=y_{n+1}-\hat{y}_{n+1}$

$$e_{n+1}=(\alpha+\beta x_{n+1}+u_{n+1})-(\hat{\alpha}+\beta x_{n+1})$$

$$e_{n+1}=(\alpha-\hat{\alpha})+(\beta-\beta)x_{n+1}+u_{n+1}$$

بالرجوع الى فرضيات النموذج $E(e_{n+1})=0$.

*التنبؤ غير المتحيز تحصلنا عليه بالتطبيق المباشر لنموذج الانحدار المقدر. الان في المجال التطبيقي لا توجد الا فائدة ضئيلة لمعرفة التنبؤ اذا لم نعرف الى اي درجة من الثقة يمكن منحها لهذا التنبؤ. و عليه سنسعى الى حساب تباين خطا التنبؤ الذي يسمح لنا بتحديد مجال الثقة الذي يتراوح خلاله التنبؤ.

تباين خطا التنبؤ يعطى بالعلاقة التالية:-

$$Var(e_{n+1})=var(y_{n+1}-\hat{y}_{n+1})=\delta_u^2[\frac{1}{n}+\frac{(x_{n+1}-\bar{x})^2}{\Sigma(x_t-\bar{x})^2}+1]$$

*نلاحظ من هذه الصيغة بان تباين خطا التنبؤ هو دالة للانحراف المربع بين المتغير المستقبل المتوقع (x_{n+1}) و متوسط نفس هذا المتغير. حيث كلما كانت القيمة المتوقعة بعيدة عن المتوسط كلما كان خطر الخطا اكبر.

$$Var(e_{n+1})=E(e^2_{n+1})$$

$$e_{n+1}= y_{n+1}-\hat{y}_{n+1}$$

$$\hat{y}_{n+1}=\hat{\alpha}+\beta x_{n+1}$$

و ان

$$y_{n+1}=\alpha+\beta x_{n+1}+u_{n+1}$$

$$e_{n+1}= u_{n+1}-(\hat{\alpha}-\alpha)-(\beta-\beta)x_{n-1}$$

$$= u_{n+1}-[(\hat{\alpha}-\alpha)+(\beta-\beta)x_{n-1}]$$

$$Var(e_{n+1})=E(e_{n+1})$$

$$=E\{u^2_{n+1}+[(\hat{\alpha}-\alpha)+(\beta-\beta)x_{n-1}]^2-2\ u_{n+1}(\hat{\alpha}-\alpha)(\beta-\beta)x_{n-1}\}$$

$$=E(u^2_{n+1})+E((\hat{\alpha}-\alpha)+(\beta-\beta)x_{n-1})^2+E(2\ u_{n+1}(\hat{\alpha}-\alpha)(\beta-\beta)x_{n-1})$$

$$=var(u_{n+1})+var(\hat{\alpha})+x^2_{n+1}var(\beta)+2x_{n+1}cov(\hat{\alpha},\beta)$$

لكن بالفرض u_{n+1} هي مستقلة عن $u_1,u_2,\text{------}u_n$ تباينها المشترك مع: $(\hat{\alpha}-\alpha)$ و$(\beta-\beta)$ اذن فهو مساوي للصفر حيث هذه الاخيرة هي دوال خطية بالنسبة لـ $u_1,u_2,\text{------}u_n$.

و لدينا:-

$$var(\beta)=\delta_u^2\Sigma w_i^2$$

$$=\frac{\delta_u^2}{\Sigma x^2}$$

$$Var(\hat{\alpha})=\delta_u^2\left(\frac{1}{n}+\frac{\bar{x}^2}{\Sigma x^2}\right)$$

$$Cov(\hat{\alpha},\beta)=-\frac{\bar{x}\delta^2{}_u}{\Sigma x^2}$$

و

$$var(u_{n+1})=\delta_u^2$$

و منه:-

$$var(e_{n+1})=\delta_u^2\left[1+\frac{1}{n}+\frac{\bar{x}^2}{\Sigma x^2}+\frac{x_{n+1}^2}{\Sigma x^2}-\frac{2x_{n+1}\bar{x}}{\Sigma x^2}\right]$$

$$=\delta_u^2\left[1+\frac{1}{n}+\frac{(x_{n+1}-\bar{x})^2}{\Sigma x^2}\right]$$

كما نلاحظ بان خطا التنبؤ e_{n+1} هو دالة خطية لعدد من المتغيرات الطبيعية و بالتالي فان هذا الخطا يتبع هو الاخر التوزيع الطبيعي و كذا:

$$\frac{e_{n+1}}{\delta_u\sqrt{1+\frac{1}{n}+\frac{(x_{n+1}-\bar{x})^2}{\Sigma x^2}}}\sim N(0,1)$$

و بما ان

$$\frac{\Sigma e_i^2}{n-2}=\delta^2{}_u$$

و منه:-

$$\frac{\Sigma e_i^2}{\delta_u^2}=\frac{(n-2)\delta_u^2}{\delta_u^2}\sim x_{n-2}^2$$

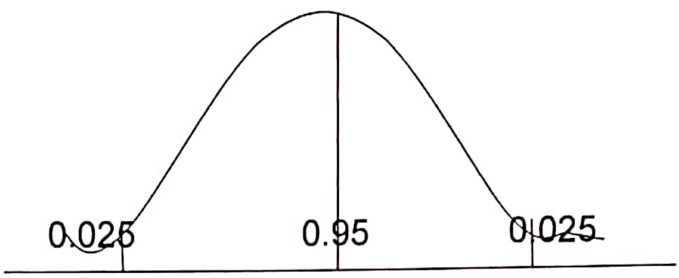
و منه:-

$$\frac{\dfrac{e_{n+1}}{\cancel{\delta_u}\sqrt{1+\frac{1}{n}+\frac{(x_{n+1}-\bar{x})^2}{\Sigma x^2}}}}{\sqrt{\dfrac{\cancel{(n-2)}\delta_u^2}{\cancel{\delta_u^2}/\cancel{(n-2)}}}}\sim t_{n-2}$$

$$t=\frac{e_{n+1}}{\delta_u\sqrt{1+\frac{1}{n}+\frac{(x_{n+1}-\bar{x})^2}{\Sigma x^2}}}$$

$$\frac{y_{n+1}-\hat{y}_{n+1}}{\delta_u\sqrt{1+\frac{1}{n}+\frac{(x_{n+1}-\bar{x})^2}{\Sigma x^2}}}$$

يوجد هناك مجهول واحدة في العلاقة السابقة هي y_{n+1}: و منه نستطيع اذن بالطريقة المعروفة ايجاد لـ y_{n+1} فترة ثقة بـ 95% بالعلاقة التالية



$$t_1 \quad p(-t_{0.025}<t\leq+t_{0.025})=0.95$$

$$p\left(-t_{0.025}<\frac{y_{n+1}-\hat{y}_{n+1}}{\delta\hat{}_u\sqrt{1+\frac{1}{n}+\frac{(x_{n+1}-\bar{x})^2}{\Sigma x^2}}}\leq t_{0.025}\right)=0.95$$

$$p\left(\hat{y}_{n+1}-t_{0.025}\delta\hat{}_u\sqrt{1+\frac{1}{n}+\frac{(x_{n+1}-\bar{x})^2}{\sum x^2}}<y_{n+1}\leq\hat{y}_{n+1}+t_{0.025}\delta\hat{}_u\sqrt{1+\frac{1}{n}+\frac{(x_{n+1}-\bar{x})^2}{\sum x^2}}\right)=0.95$$

و منه و بدرجة ثقة $(1-\alpha)\%$ فان مجال الثقة لـ y_{n+1}

$$y_{n+1}=\hat{y}_{n+1}\pm t_{n-2}^{\alpha/2}\delta\hat{}_u\sqrt{1+\frac{1}{n}+\frac{(x_{n+1}-\bar{x})^2}{\sum(x_t-\bar{x})^2}}$$

*في بعض الاحيان تكون مصلحتنا البحث عن التنبؤ للتوقع الرياضي لـ y_{n+1}

$$E(y_{n+1})=\alpha+\beta x_{n+1}$$

يدل البحث عن التنبؤ لـ y_{n+1} لانه لا يمكن باي صورة التنبؤ بدقة عن نتائج سحب عشوائي بدا من $p(u)$ قانون احتمال المتحول العشوائي u.

و منه: $$e_{n+1}=E(y_{n+1})-\hat{y}_{n+1}$$

$$=(\alpha-\hat{\alpha})+(\beta-\beta)x_{n+1}$$

$$=-(\hat{\alpha}-\alpha)+(\beta-\beta)x_{n+1}=-[(\hat{\alpha}-\alpha)+(\beta-\beta)x_{n+1}]$$

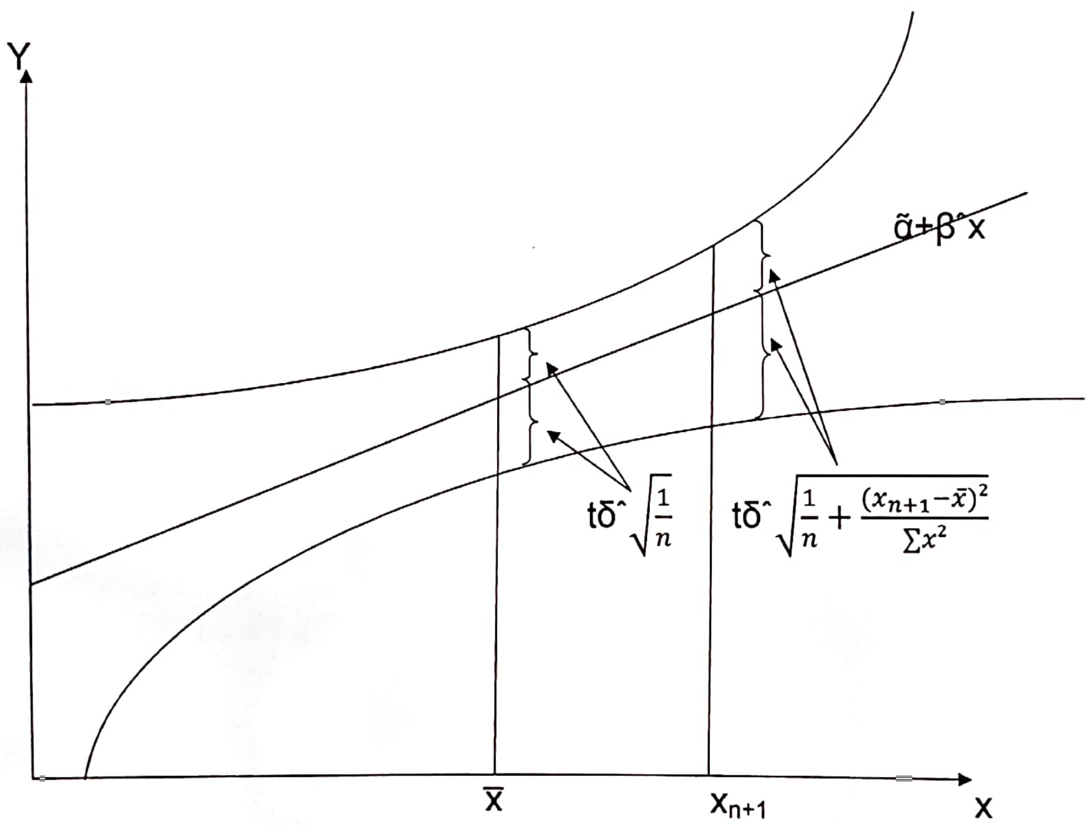
$$Var(e_{n+1})=\delta_u^2\left(\frac{1}{n}+\frac{(x_{n+1}-\bar{x})^2}{\sum x^2}\right)$$

مما يسمح لنا بتكوين مجال ثقة بـ 95% لـ $E(y_{n+1})$ حيث:-

$$\tilde{\alpha}+\hat{\beta}x_{n+1}=\hat{y}_{n+1}$$

$$\hat{y}_{n+1}\pm t_{0.05}\,\hat{\delta}_u\sqrt{\frac{1}{n}+\frac{(x_{n+1}-\bar{x})^2}{\sum x^2}}$$

*ان طول مجال الثقة وفقا لمجال الثقة بالنسبة لـ y_{n+1} و $E(y_{n+1})$ يزيد كلما بعد x_{n+1} عن متوسط $\bar{x}$ للعينة كما هو موضح بالبيان التالي:-



مثال:

تقابل العين المبينة بالجدول ادناه النموذج الخطي $y=\alpha+\beta x+u$

حيث تمثل x الدخل و yالانفاق الاستهلاكي

X	y
20	11
15	10
12	7
10	7
6	4
3	3

المطلوب:

اولا: تقدير مستقيم المربعات الصغربلهذه البيانات

ثانيا: حساب كل من التباين المفسر و التباين غير المفسر لـy و ايجاد تقدير التباين مقدار الخطا u.

ثالثا: اجراء اختبار استيودنت (و حيد الطرف) للفرضية القائلة بان النزعة الحدية للاستهلاك اكبر من 0.5 – $H_0 : \beta > 0.5$

رابعا: تقدير المجال بـ 90%من الثقة من اجل β.

خامسا: اختبار الفرضية القائلة بان الاستهلاك يتناسب تماما مع الدخل: $H_0 : \alpha = 0$ و ذلك بمستوى دلالة 5%.

سادسا: اختبار فرضية اللاعلاقة $H_0: \beta=0$ و ذلك باستخدام كلا من اختبار استبودنت و اختبار تحليل التباين و في هذه الحالة على الطالب اتخاذ مستوى الدلالة المناسب و ابداء رايه حول درجة او شدة العلاقة بين المتغيرين الاقتصاديين xوy.

سابعا: حقق نتائج الاختبارين الاخيرين باجراء اختبار معامل الارتباط.

الحل:

y	x	$x=x-\bar{x}$	$y=y-\bar{y}$	xy	x^2	y^2
11	20	9	4	36	81	16
10	15	4	3	12	16	9
7	12	1	0	0	1	0
7	10	-1	0	0	1	0
4	6	-5	-3	15	25	9
3	3	-8	-4	32	64	16
42	66			95	188	50

$$\bar{x}=\frac{\Sigma x}{n}=\frac{66}{6}=11$$

$$\bar{y}=\frac{\Sigma y}{n}=\frac{42}{6}=7$$

1-تقدير مستقيم المربعات الصغرى لهذه البيانات:-

*تقدير $\hat{\beta}$:- $$\hat{\beta}=\frac{\Sigma xy}{\Sigma x^2}=\frac{95}{188}\simeq 0.51$$

*تقدير $\tilde{\alpha}$:- $$\tilde{\alpha}=\bar{y}-\hat{\beta}\bar{x}=7-0{,}51.11=1{,}39$$

*تقدير مستقيم المربعات الصغرى:- $$\hat{y}=\tilde{\alpha}+\hat{\beta}x$$

$$\hat{y}=1,39+0,5x$$

2-حساب كل من التباين المفسر و غير المفسر لـ y و ايجاد تقدير تباين حد الاضطراب.

*حساب التباين المفسر $\sum \hat{y}^2=\hat{\beta}^2 \sum x_i^2$

و الاحسن

$$=\hat{\beta}\sum xy$$

$$=0,51.95$$

$$=48,45$$

*حساب التباين غير المفسر $\sum e_i^2=\sum y^2-\sum \hat{y}^2$

$$=50-48,45$$

$$=1,55$$

التحليل:-

$$\sum \hat{y}^2 \longrightarrow \sum \hat{y}^2=48,45$$

$$\sum \hat{y}^2 \longrightarrow \sum e_i^2=1,55$$

و هذا يدل على وجود علاقة شديدة بينxوy و ذلك لان التباين المفسر الذي يقدر بـ 48.4 اكبر بكثير من التباين غير المفسر الذي يقدر بـ1.6

*تقدير تباين حد الاضطراب: $\hat{\delta}_u^2=\frac{\sum e_i^2}{n-2}=\frac{1.6}{4}=0.4$

و هذا يدل على ان النقاط قريبة جدا من مستقيم المربعات اي العلاقة كانها تامة

3-اجراء اختبارات استيودنت القائلة $H_0:\beta>0.5$

و ذلك بمستوى دلالة 0.05

ملاحظة: بما ان المتحول العشوائي t قد ياخذ قيم سالبة

$$t_{0b}=\left|\frac{(\hat{\beta}-\beta)\sqrt{\sum x^2}}{\hat{\delta}_u}\right| < t_{0.05}$$

و منه اذا كان:

$$H0:\beta > 0.5 \Rightarrow |\hat{\beta}-\beta| < t_{0.05}.\frac{\hat{\delta}_u}{\sqrt{\sum x_i^2}}$$ مقبولة

و اذا كان:

$$H0:\beta > 0.5 \Rightarrow |\hat{\beta}-\beta| > t_{0.05}.\frac{\hat{\delta}_u}{\sqrt{\sum x_i^2}}$$ مرفوضة

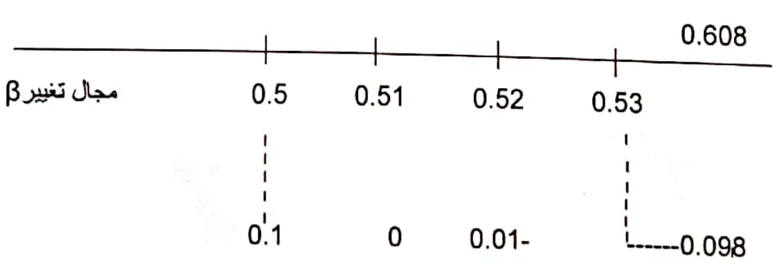
و من اجل ذلك لدينا:

$$t_{0.05}\frac{\hat{\delta}_u}{\sqrt{\sum x_i^2}}=2,13.\frac{0,63}{13,71}=0,098$$

و لذلك اذا كان:

H_0 $|0,51-\beta| < 0,098$ مقبولة.

H_0 $|0,51-\beta| > 0,098$ مرفوضة.



اذن الفرضية $H0 : \beta > 0.5$ مقبولة من اجل $\beta<0.608$ و فيما عدا ذلك فان الفرضية $\beta>0.5$ تكون مرفوضة و 0.95

4-تقدير مجال الثقة بـ 0.90 من اجل β

*بـ 90%

$$\hat{\beta} \pm t_{0.05}\frac{\hat{\delta}_u}{\sqrt{\sum x^2}}$$

0.41   0.61

$\hat{\beta} - t_{0.05}\frac{\hat{\delta}_u}{\sqrt{\sum x^2}}$ $\hat{\beta}=0.51$ $\hat{\beta} + t_{0.05}\frac{\hat{\delta}_u}{\sqrt{\sum x^2}}$

و ذلك كما يلي:-

$$\left.\begin{array}{l} t_{0.05}=2.13 \\ \hat{\delta}_u = 0.63 \\ \sqrt{\sum x^2} = 13.71 \end{array}\right\} t_{0.05}\frac{\hat{\delta}_u}{\sqrt{\sum x^2}} = 0.10$$

الحد الاعلى للمجال:

$$\hat{\beta} + t_{0.05}\frac{\hat{\delta}_u}{\sqrt{\sum x^2}} = 0.51 + 0.10 = 0.61$$

الحد الادنى للمجال:

$$\hat{\beta} - t_{0.05}\frac{\hat{\delta}_u}{\sqrt{\sum x^2}} = 0.51-0.10=0.41$$

اي ان β مجهول ياخذ قيمة مابين 0.41 و 0.61 في 90 عينة مختارة اذا اخذنا عدد من العينات قدره 100عينة.

*من اجل مجال ثقة 0.95

$$\hat{\beta} \pm t_{0.025}\frac{\hat{\delta}_u}{\sqrt{\sum x^2}}$$

$t_{0.025}=2.78$

و منه:

$$t_{0.025}\frac{\hat{\delta}_u}{\sqrt{\sum x^2}} = 2.78\frac{0.63}{13.71}=0.13$$

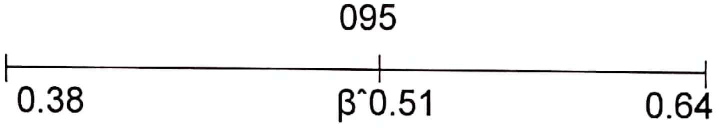
051 ± 0.13

الحد الاعلى:

$051+0.13=0.64$

الحد الادنى:

$051-0.13=0.38$



5-اختبار الفرضية القائلة بان الاستهلاك يتناسب مع الدخل $H_0: \alpha=0$

لدينا التابع الاختباري:

$$t=\frac{(\tilde{\alpha}-\alpha)\sqrt{n\sum x^2}}{\hat{\delta}_u\sqrt{\sum x^2}}$$

و منه: $\alpha=0$

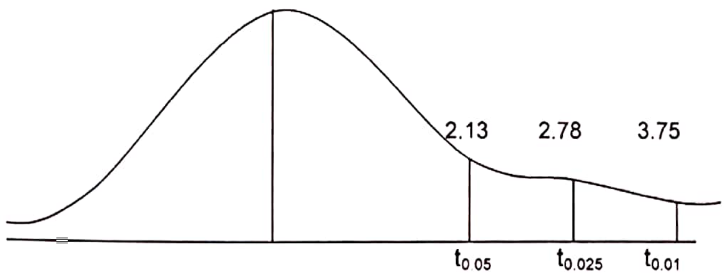
$$t=\frac{\tilde{\alpha}\sqrt{n\sum \bar{x}^2}}{\hat{\delta}_u\sqrt{\sum x^2}}$$

بعد حساب قيمة t نقارنها مع t النظرية $t_t > t_c$ تقبل فرضية ان الدخل يتناسب مع الاستهلاك.

6-اختبار فرضية اللاعلاقة $H0 : \beta = 0$

*باستخدام اختبار استيودنت

$$t=\frac{(\hat{\beta}-\beta)\sqrt{\sum x^2}}{\hat{\delta}_u}$$



$$t=\frac{\hat{\beta}\sqrt{\sum x^2}}{\hat{\delta}_u}=\frac{0,51.13,71}{0,63}=11,10$$

اذن ليس هناك مجال لقبول الفرضية اي نرفض فرضية اللاعلاقة بحزم و نقر بوجود علاقة قوية بين xوy.

*اختبار فيشر:

$$F=t^2=\frac{\hat{\beta}^2\sqrt{\sum x^2}}{\hat{\delta}_u}=\frac{(0.51)^2 188}{0.4}$$

$$=\frac{0,2601.188}{0,4}$$

$$=\frac{48,8988}{0,4}=122,247$$

و منه: $F_c=122,247$

و من الجدول نجد:

$$F_{0.05}=7.71$$
$$(1.4)$$

$$F_{0.01}=21.2$$
$$(1.4)$$

نرفض نتيجة اختبار تحليل التباين فرضية اللاعلاقة بحزم اي نقر وجود علاقة قوية بين

xوy.

7-حقق نتائج الاختبارين باجراء اختبار معامل الارتباط

$$t=\frac{л\sqrt{n-2}}{\sqrt{1-л^2}}$$

$$л^2=\frac{\sum y^{\wedge 2}}{\sum y}=\frac{48.45}{50}=0.969$$

$$л^2=0.969$$

و منه:

$$1-л^2=0.031$$

$$\sqrt{1-л^2}=\sqrt{0.031}\simeq 0.18$$

و ايضا:

$$Л=\sqrt{0.969}=0.98$$

و منه:

$$t=\frac{л^2\sqrt{n-2}}{\sqrt{1-л^2}}$$

$$=\frac{0.98\sqrt{6-2}}{0.18}$$

$$t=\frac{0,98.2}{0,18}=10,89$$

8-ايجاد فترة الثقة للتنبؤ بالانفاق عندما الدخل $x_{n+1}=2$ بـ 95% من درجة الثقة.

و منه: $\hat{y}_{n+1}=1.39+0.512$

$\hat{y}_{n+1}=2.41$

و فترة الثقة تبعا لذلك:

$$\hat{y}_{n+1} \pm t_{n-2}^{\alpha/2} \quad \hat{\delta}_u \sqrt{1+\frac{1}{n}+\frac{(x_{n+1}-\bar{x})^2}{\sum x^2}}$$

$$2{,}41 \pm 2{,}78.0{,}63\sqrt{1+\frac{1}{6}+\frac{(2-11)^2}{188}}$$

$$2{,}141 \pm 2{,}78.0{,}63\sqrt{1+0{,}17+0{,}43}$$

$$2{,}41 \pm 2{,}78.0{,}63.1{,}26$$

$$2{,}41 \pm 2{,}21$$

$$2{,}41+2{,}21=4{,}62$$

$$2{,}41-2{,}21=0{,}2$$

$$0{,}2 \leq y_{n+1} \leq 4{,}62$$

اي بدرجة ثقة قدرها 95% ان الانفاق الحقيقي تكون قيمته محصورة بين 4.62 و 0.2 عندما يكون الدخل يساوي 2.

9-ايجاد فترة الثقة للتنبؤ بمتوسط الانفاق عندما $x_{n+1}=2$ بـ 95% من الثقة:

$$\hat{y}_{n+1}+t_{n-1}^{\alpha/2}\hat{\delta}_u\sqrt{\frac{1}{n}+\frac{(x_{n+1}-\bar{x})^2}{\sum x_1^2}} \leq E(y_{n+1}) \leq \hat{y}_{n+1}-t_{n-2}^{\alpha/2}\hat{\delta}_u\sqrt{(\frac{1}{n}\frac{(x_{n+1}-\bar{x})^2}{\sum x_1^2}}$$

$$\hat{y}_{n+1} \pm t_{\frac{\alpha}{2}}(n-2).\hat{\delta}_u\sqrt{\frac{1}{n}+\frac{(x_{n+1}-\bar{x})^2}{\sum x_1^2}}$$

$$2{,}41 \pm 2{,}78.0{,}63\sqrt{\frac{1}{6}+\frac{(2-11)^2}{188}}$$

$$2{,}41 \pm 2{,}78.0{,}63.0{,}77$$

$$2{,}41 \pm 1{,}35$$

$$2{,}41+1{,}35=3{,}76$$

$$2{,}41-1{,}35=1{,}06$$

$$1{,}06 \leq E(y_{n+1}) \leq 3{,}76$$

اي بدرجة ثقة قدرها 95% ان متوسط الانفاق الفعلي تكون قيمته محصورة بين 3.76 و 1.06 عندما يبلغ الدخل يساوي 2 .