

### - (3) تحديد المشاهدات المؤثرة – تدابير (IDENTIFYING INFLUENTIAL CASES) (DFFITS , COVRATIO, Cook'S Distance, DFBETAS,)

بعد تحديد المشاهدات القاصية في كل من المتغيرات المستقلة  $X$  ، والمتغيرات التابعة  $Y$  ، تكون الخطوة التالية: التعرف على ما إذا كانت المشاهدات القاصية مؤثرة أم لا. وتعتبر المشاهدة مؤثرة إذا كان استثنائها يسبب تغيرات رئيسية في تقدير معالم دالة الانحدار التوفيقية والإحصاءات المرتبطة بها. وليس من الضرورة أن تكون جميع المشاهدات القاصية مؤثرة .

ويتم قياس فيما إذا كانت المشاهدات القاصية مؤثرة أم لا بعدة مقاييس نذكر منها المقاييس الأربعة التالية:

(DFFITS , COVRATIO, Cook's Distance, DFBETAS,).

#### - (3-1) التأثير على قيم التوفيقية – مقياس (DFFITS).

إحدى المقاييس المفيدة لقياس تأثير المشاهدة رقم  $i$  على القيمة التوفيقية  $\hat{Y}_i$  معطى بإحدى العلاقتين:

$$(DFFITS)_i = \frac{\hat{Y}_i - \hat{Y}_{i(i)}}{\sqrt{MSE_{(i)} h_{ii}}}$$

$$(DFFITS)_i = d_i^* \left[ \frac{h_{ii}}{1 - h_{ii}} \right]^{\frac{1}{2}}$$

حيث أن :

$\hat{Y}_i$  ← القيمة التوفيقية للمتغير التابع للمشاهدة  $i$  عند استخدام جميع المشاهدات في توفيق دالة الانحدار.

$d_i^*$  ← بواقي الحذف المعيرة (بواقي ستودنت المحذوفة) للمشاهدة رقم  $i$  عند استخدام جميع المشاهدات في توفيق دالة الانحدار .

$h_{ii}$  ← قيم العزم للمشاهدة رقم  $i$  الذي يتم استخراجها من العنصر القطري في مصفوفة القبة عند استخدام جميع المشاهدات في توفيق دالة الانحدار .

$\hat{Y}_{i(i)}$  ← قيمة التوفيق للتنبؤ بالمتغير التابع عند حذف المشاهدة رقم  $(i)$  .

$MSE_{(i)} \leftarrow$  متوسط مربعات الخطأ عند حذف المشاهدة رقم (i).

ونلاحظ من العلاقة الثانية أن القيم DFFITS تعتمد على قيم العزم و بواقي الحذف المعيرة أي أنه إذا كانت المشاهدة i قاصية في المتغير المستقل X (قيمة العزم لها كبيرة) وأيضا إذا كانت المشاهدة i قاصية في المتغير التابع Y (قيمة باقي ستودنت كبيرة) أيضا فإن القيمة المطلقة لهذا العامل سوف تصبح كبيرة أيضا ومنها تكون هذه القيمة الكبيرة كدليل نستدل منه فيما إذا كانت مؤثرة أم لا.

- طريقة كشف التأثير لقيمة DFFITS على النموذج .

تعتبر المشاهدة (i) مؤثرة إذا كانت القيمة المطلقة لقيمة DFFITS تحقق العلاقات التالية :

$$|DFFITS| > 2 \sqrt{\frac{p}{n}} \quad 1 - \text{في العينات الكبيرة}$$

$$|DFFITS| > 1 \quad 2 - \text{في العينات الصغيرة والمتوسطة}$$

حيث أن :-

n : عدد المشاهدات .

p : عدد معالم النموذج  $B_0, B_1, \dots, B_{p-1}$

ويقترح آخرون مثل شاترجي وهادي (Chatterge and Hadi (1988)) بمقارنة القيمة المطلقة DFFITS بعدد أكبر قليلا من المقارنة في حالة العينة الكبيرة وهي .

$$|DFFITS| > 2 \sqrt{\frac{p}{n-p-2}}$$

- (3-2) التأثير على معاملات الانحدار- مقياس (DFBETAS) .

يعتبر مقياس DFBETAS من أحد الإحصاءات التي تقيس تأثير المشاهدة رقم i على كل معامل انحدار. وهو يقيس الفرق بين معامل الانحدار  $b_k$  المقدر باستخدام جميع المشاهدات وبين معامل الانحدار المقدر الذي نحصل عليه عند حذف المشاهدة رقم i والذي يرمز له بالرمز  $b_{k(i)}$ ، ويتم حسابه من العلاقة التالية :

$$(DFBTAS)_{k(i)} = \frac{b_k - b_{k(i)}}{\sqrt{MSE_{(i)} C_{kk}}}, k = 0, 1, \dots, p - 1$$

حيث أن :-

$b_k$  : معامل الانحدار رقم  $k$  المقدر باستخدام كل المشاهدات  $(n)$ .

$b_{k(i)}$  : معامل الانحدار رقم  $k$  المقدر بعد حذف المشاهدة رقم  $(i)$  في كل مرة.

$C_{kk}$  : هو العنصر القطري رقم  $k$  من المصفوفة  $(X'X)^{-1}$ .

$MSE_{(i)}$  : متوسط مربعات الخطأ للتقدير لنموذج الانحدار المقدر بعد حذف المشاهدة رقم  $i$  في كل مرة.

- طريقة كشف التأثير لقيمة DFBETAS على النموذج .

تعتبر المشاهدة رقم  $(i)$  مؤثرة على معامل نموذج الانحدار رقم  $k$  إذا كانت القيمة المطلقة لقيمة DFBETAS تحقق العلاقات التالية :

$$\begin{aligned} 1 - \text{في العينات الكبيرة} \quad & |DFBETAS_{k(i)}| > \frac{2}{\sqrt{n}} \\ 2 - \text{في العينات الصغيرة} \quad & |DFBETAS_{k(i)}| > 1 \end{aligned}$$

- (3-3) قياس الأثر على كل معاملات الانحدار - مقياس مسافة كوك.

مقياس مسافة كوك  $D_i$  هو مقياس إجمالي للتأثير المشترك للمشاهدة رقم  $i$  على جميع معاملات الانحدار المقدر . ويختلف مقياس كوك عن مقياس DFBETAS أن مقياس مسافة كوك يقيس أثر حذف المشاهدة رقم  $(i)$  على كل قيم معاملات نموذج الانحدار ، ومقياس DFBETAS يقيس أثر حذف المشاهدة رقم  $(i)$  على كل معامل من معاملات النموذج على حده.

ويأخذ مقياس مسافة كوك العلاقة التالية :

$$D_i = \frac{(b - b_{(i)})' X' X (b - b_{(i)})}{p * MSE}$$

$$D_i = \frac{e_i^2}{MSE_p} \left[ \frac{h_{ii}}{(1 - h_{ii})^2} \right]$$

حيث أن :-

$e_i$  : قيمة البواقي العادية لنموذج الانحدار للمشاهدة رقم (i) .

$h_{ii}$  : قيم العزم التي يتم استخراجها من العنصر القطري في مصفوفة القبة.

$b$  : متجه معاملات الانحدار المقدرة التي نحصل عليها لكل المشاهدات  $n$  .

$b_{(i)}$  : متجه معاملات الانحدار المقدرة التي نحصل عليها عند حذف المشاهدات رقم (i) في كل مرة.

$P$  : عدد معالم النموذج  $B_0, B_1, \dots, B_{p-1}$  .

$MSE$  : متوسط مربعات الخطأ .

ونلاحظ أن مقياس مسافة كوك يعتمد على عاملين هما حجم الراسب  $e_i$  وقيمة العزم  $h_{ii}$  . وكلما كان احدهما كبيرا كلما كان  $D_i$  كبيرا ..

- طريقة كشف التأثير لقيمة مسافة كوك  $D_i$  .

تتم مقارنة مسافة كوك  $D_i$  بالقيمة الاحتمالية لتوزيع  $F$  بالشكل التالي  $F_{\alpha, p, n-p}$  فإذا كانت قيمة مسافة كوك أكبر من هذه القيمة نعتبر المشاهدات رقم (i) مؤثرة على قيم معاملات نموذج الانحدار. أي نعتبر المشاهدات مؤثرة إذا كان:

• طريقة نيترو وآخرون (ص404. 1990. 3<sup>rd</sup> ed) .

$$D_i > F_{\alpha, p, n-p} \quad \text{for} \quad \alpha = 0.5$$

حيث أن :-

$n$  : عدد المشاهدات .

$P$  : عدد معالم النموذج  $B_0, B_1, \dots, B_{p-1}$  .

$F_{\alpha, p, n-p}$  : المئين 50 % لتوزيع  $F$  بدرجات حرية  $p, n-p$  .

• طريقة فوكس (ص 281. (1997) .

$$D_i > \frac{4}{n-p}$$

#### - (3-4) قياس الأثر على الأخطاء المعيارية – مقياس COVRATIO .

نستخدم مقياس الأثر على الأخطاء المعيارية لقياس أثر أي حالة على مصفوفة تباين أو تباين معاملات الانحدار المقدرة . وهي عبارة عن نسبة محددة مصفوفة تباين أو تباين معاملات الانحدار المقدرة (لذلك تمت التسمية )

بعد حذف المشاهدة قم (i) لمحددة مصفوفة تباين أو تباين معاملات الانحدار المقدرة باستخدام جميع الحالات (n) . وقد تم تطوير معادلة حسابية تعتمد على قيم العزم وقيم بواقي ستودنت المحذوفة وهي إلى الشكل التالي :

$$COVRATIO_i = \frac{1}{(1 - h_{ii}) \left[ \frac{n - p - 1 + d_i^{*2}}{n - p} \right]^p}$$

حيث أن :

p : عدد معالم النموذج  $B_0, B_1, \dots, B_{p-1}$  .

n : عدد المشاهدات .

$h_{ii}$  : قيم العزم التي يتم استخراجها من العنصر القطري في مصفوفة القبة.

$d_i^*$  :بواقي الحذف المعيرة (بواقي ستودنت المحذوفة) .

ونلاحظ من هذه المعادلة عدم الحاجة إلى توفيق n من معادلات الانحدار وذلك لما ذكرنا لاعتمادها الواضح على قيم العزم وبواقي ستودنت المحذوفة . حيث تزيد قيمة COVRATIO بزيادة قيم العزم وانخفاض بواقي ستودنت المحذوفة ، ويكون ذلك مؤشرا جيدا لاكتشاف تأثير قيم COVRATIO المؤثرة على الأخطاء المعيارية.

#### - طريقة كشف التأثير لقيمة COVRATIO على الأخطاء المعيارية.

تتم مقارنة قيمة المقياس  $COVRATIO_i$  بالقيمة  $1 \pm \frac{3p}{n}$  طبقا للفترة التالية :

$$1 - \frac{3p}{n} < COVRATIO_i < 1 + \frac{3p}{n}$$

أي إذا كانت قيمة  $COVRATIO_i$  خارج هذه الفترة فإن المشاهدة رقم (i) تعتبر مؤثرة على قيم الأخطاء المعيارية لمعاملات نموذج الانحدار .

#### - تطبيق (3-1) :

بالعودة إلى بيانات ومشاهدات تطبيق (1-1) و تطبيق (2-1) . سوف نقيس في هذا التطبيق أثر المشاهدات القاصية على المتغيرات الجهد (الفولت) والسرعة (بوصة/ ثانية) و التمديد (بوصة)

والأخطاء المعيارية في نموذج الانحدار.

### - التحليل الإحصائي :-

باستخدام برامج الحزم الإحصائية Minitab و SPSS سوف نقوم باستخراج المشاهدات المؤثرة على كل من متغير السرعة X1 والتمديد X2 وكذلك المشاهدات المؤثرة على الجهد والأخطاء المعيارية أيضاً.

## Regression

Variables Entered/Removed(b)

Model	Variables Entered	Variables Removed	Method
1	X2, X1(a)	.	Enter

a All requested variables entered.

b Dependent Variable: Y

Model Summary(b)

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.980(a)	.961	.956	.12455

a Predictors: (Constant), X2, X1

b Dependent Variable: Y

ANOVA(b)

Model		Sum of Squares	Df	Mean Square	F	Sig.
1	Regression	5.694	2	2.847	183.540	.000(a)
	Residual	.233	15	.016		
	Total	5.927	17			

a Predictors: (Constant), X2, X1

b Dependent Variable: Y

Coefficients(a)

Model		Unstandardized Coefficients		Standardized Coefficients	T	Sig.
		B	Std. Error	Beta		
1	(Constant)	-1.641	.247		-6.644	.000
	X1	.001	.000	.872	16.115	.000
	X2	-67.396	4.481	-.814	-15.041	.000

a Dependent Variable: Y

## Residuals Statistics(a)

	Minimum	Maximum	Mean	Std. Deviation	N
Predicted Value	.8994	2.9950	1.9222	.57875	18
Std. Predicted Value	-1.767	1.854	.000	1.000	18
Standard Error of Predicted Value	.031	.073	.050	.011	18
Adjusted Predicted Value	.8339	2.9478	1.9191	.57660	18
Residual	-.20556	.19631	.00000	.11699	18
Std. Residual	-1.650	1.576	.000	.939	18
Stud. Residual	-1.735	1.774	.012	1.036	18
Deleted Residual	-.22719	.25614	.00315	.14308	18
Stud. Deleted Residual	-1.875	1.928	.017	1.082	18
Mahal. Distance	.079	4.919	1.889	1.232	18
Cook's Distance	.001	.361	.077	.098	18
Centered Leverage Value	.005	.289	.111	.072	18

## Cumulative Distribution Function

F distribution with 3 DF in numerator and 15 DF in denominator

$$x \quad P(X \leq x)$$

$$0.5 \quad 0.312098 = F_{0.5,3,15}$$

- قيمة (قيمة  $F_{\alpha,p,n-p} = F_{0.5,3,15} = 0.312098$ ) ، وقيمة مقياس مسافة كوك كما في الجدول.

```
MTB > Let k1 = 2 * SQRT(3 / 18)
MTB > Let k2 = 2 / SQRT(18)
MTB > Let k3 = 1 + ((3 * 3) / 18)
MTB > Let k4 = 1 - ((3 * 3) / 18)
MTB > prin k1 k2 k3 k4
```

## Data Display

$$K1 = 2\sqrt{\frac{p}{n}} = 0.816497$$

$$K2 = \frac{2}{\sqrt{n}} = 0.471405$$

$$K3 = 1 + \frac{3p}{n} = 1.50000$$

$$K4 = 1 - \frac{3p}{n} = 0.500000$$

- جدول (3-1) قيم الاختلاف في المتغير التابع الموفق ومعاملات الانحدار ومسافة كوك ومقياس نسبة . COV

I	DFFITS	DF B0	DF B1	DF B2	Di	COVR	Big DF	Di>F	Big covr
1	0.27235	0.17157	-0.13033	-0.12855	0.02581	1.36812			

I	DFFITS	DF B0	DF B1	DF B2	Di	COVR	Big DF	Di>F	Big covr
2	0.70082	0.09835	0.02924	-0.51004	0.14202	0.75026			
3	0.50542	-0.22616	0.30445	-0.38741	0.08515	1.25538			
4	-0.55124	-0.3726	0.29267	0.23874	0.09879	1.13216			
5	-0.81547	-0.7243	0.64425	0.14784	0.21784	1.44889			
6	0.76017	-0.43053	0.53709	-0.55922	0.18456	1.18699			
7	-0.08133	-0.06202	0.0536	-0.00139	0.00235	1.36012			
8	-0.15718	-0.02188	-0.00208	0.0419	0.00859	1.20627			
9	-0.60815	0.28906	-0.36879	0.24677	0.10558	0.69415			
10	0.08363	-0.05866	0.06733	-0.03491	0.00249	1.47309			
11	0.25929	0.17933	-0.17571	0.1473	0.02331	1.30326			
12	-0.12444	-0.0248	0.02094	-0.06481	0.00546	1.27932			
13	-0.27656	0.1042	-0.11473	-0.08166	0.02588	1.14842			
14	-0.55601	0.37521	-0.39519	-0.0564	0.09811	1.01475			
15	1.13059	0.59748	-0.65002	0.93045	0.36073	0.81553	1.13059	0.36073	
16	0.09596	0.02268	-0.02729	0.08039	0.00328	1.49873			
17	0.06008	-0.00777	0.00515	0.04485	0.00129	1.48097			
18	0.09552	-0.04592	0.04277	0.05247	0.00325	1.55418			1.55418

- الاستنتاج :-

### 1 - مقياس DFFITS :-

واضح من جدول (3-1) . أن هناك قيمة واحدة مؤثرة على قيم الجهد الناتج من محركات الغاز والقيمة المطلقة لمقياس  $DFFITS$  أكبر من  $0.816497 = 2\sqrt{P/n}$  وهي المشاهدة رقم 15 حيث أن قيمة الاختلاف لها تساوي:

$$DFFITS_{15} = 1.13059$$

### 2 مقياس DFBETAS :-

نلاحظ من الجدول (3-1) ، أن هناك 4 مشاهدات مؤثرة على معاملات نموذج الانحدار والقيمة المطلقة لمقياس  $DFBETAS$  معاملان وأيضا هناك مشاهدات مؤثرة على جميع المعاملات. حيث أن المشاهدة 2 مؤثرة على  $B_2$  في النموذج لهذه القيم أكبر من  $0.471405 = \frac{2}{\sqrt{n}}$  البعض منها مؤثر على معامل واحد وكذلك على



وكذلك المشاهدة 5 مؤثرة على  $B_0$  و  $B_1$  في النموذج وكذلك المشاهدة 6 مؤثرة على  $B_1$  و  $B_2$  في النموذج أما المشاهدة 15 فهي مؤثرة على جميع معاملات النموذج ( $B_2 B_1 B_0$ ) مع ملاحظة أن المشاهدة 5 كانت مشاهدة قاصية في المتغيران (السرعة والتمديد). كما لم يكن هناك قيم قاصية في المتغير التابع (الجهد).

### 3 - مقياس مسافة كوك $D_i$ :-

نلاحظ من الجدول (3-1) وجود مشاهدة واحدة مؤثرة على معاملات الانحدار وهي المشاهدة 15. حيث أن القيمة المؤثرة بمقياس كوك  $D_i$  تساوي :

$$D_{15} = 0.36073$$

وهي أكبر من  $F_{0.5,3,15} = 0.312098$ .

### 4 - مقياس COVRATIO :-

نلاحظ من الجدول (3-1) أنه يوجد قيمة واحدة مؤثرة فقط على الأخطاء المعيارية وهي المشاهدة رقم 18 وتساوي :

$$COVRATIO_{18} = 1.55418$$

### • جدول تلخيص للمشاهدات القاصية والمؤثرة في المتغير التابع والمتغير المستقل .

المشاهدة i	قاصية في المتغير التابع Y	قاصية في المتغير المستقل X	مؤثرة DFFITS	مؤثرة DFBETAS	مؤثرة كوك $D_i$	مؤثرة COVRATIO
2	×	×	×	✓	×	×
5	×	✓	×	✓	×	×
6	×	×	×	✓	×	×
15	×	×	✓	✓	✓	×
18	×	×	×	×	×	✓

### • تطبيق (3-2) :-

بالعودة إلى بيانات ومشاهدات تطبيق (1-2) وتطبيق (2-2). سوف نقيس في هذا التطبيق أثر المشاهدات القاصية على متغير العمالة. وأيضا القيم المؤثرة على التكلفة (التعويضات) المقدر وكذلك المشاهدات المؤثرة على

معاملات النموذج والأخطاء المعيارية .

### • التحليل الإحصائي :-

باستخدام برامج الحزم الإحصائية Minitab و SPSS سوف نقوم باستخراج المشاهدات المؤثرة على متغير العمالة X3 وكذلك المشاهدات المؤثرة على التكلفة والأخطاء المعيارية أيضا.

## Regression

Variables Entered/Removed(b)

Model	Variables Entered	Variables Removed	Method
1	X3(a)	.	Enter

a All requested variables entered.

b Dependent Variable: Y

Model Summary(b)

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.588(a)	.345	.324	102.53415

a Predictors: (Constant), X3

b Dependent Variable: Y

ANOVA(b)

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	172041.453	1	172041.453	16.364	.000(a)
	Residual	325910.789	31	10513.251		
	Total	497952.242	32			

a Predictors: (Constant), X3

b Dependent Variable: Y

Coefficients(a)

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	356.982	22.215		16.069	.000
	X3	.001	.000	.588	4.045	.000

a Dependent Variable: Y

#### Residuals Statistics(a)

	Minimum	Maximum	Mean	Std. Deviation	N
Predicted Value	361.4351	775.6284	410.4848	73.32323	33
Std. Predicted Value	-.669	4.980	.000	1.000	33
Standard Error of Predicted Value	17.860	92.012	21.930	12.693	33
Adjusted Predicted Value	361.7796	1096.6870	420.3437	125.52021	33
Residual	-162.15150	276.79572	.00000	100.91933	33
Std. Residual	-1.581	2.700	.000	.984	33
Stud. Residual	-1.716	2.742	-.030	1.037	33
Deleted Residual	-398.68707	285.66348	-9.85884	124.87139	33
Stud. Deleted Residual	-1.774	3.100	-.020	1.079	33
Mahal. Distance	.001	24.800	.970	4.283	33
Cook's Distance	.000	6.088	.202	1.057	33
Centered Leverage Value	.000	.775	.030	.134	33

a Dependent Variable: Y

#### Cumulative Distribution Function

F distribution with 2 DF in numerator and 31 DF in denominator

$$x \quad P(X \leq x) \\ 0.5 \quad 0.388662 = F_{0.5, 2, 31}$$

```
MTB > Let k1 = 2 * SQRT(2 / 33)
MTB > Let k2 = 2 / SQRT(33)
MTB > Let k3 = 1 + ((3 * 2) / 33)
MTB > Let k3 = 1 - ((3 * 2) / 33)
MTB > prin k1 k2 k3 k4
```

#### Data Display

$$K1 = 2 \sqrt{\frac{p}{n}} = 0.492366 \\ K2 = \frac{2}{\sqrt{n}} = 0.348155 \\ K3 = 1 + \frac{3p}{n} = 1.18182 \\ K4 = 1 - \frac{3p}{n} = 0.818182$$

- جدول (3-2) قيم الاختلاف في المتغير التابع الموفق ومعاملات الانحدار ومسافة كوك ومقياس نسبة . COV

I	DFFITS	DF B0	DF B1	Di	COVR	Big DF	Di>F	Big covr
1	0.05786	0.04342	0.00488	0.00172	1.09362			

I	DFFITS	DF B0	DF B1	Di	COVR	Big DF	Di>F	Big covr
2	-0.06977	-0.0458	-0.01502	0.0025	1.092			
3	-0.01186	-0.01168	0.00534	0.00007	1.10969			
4	-0.20869	-0.1938	0.05317	0.02157	1.01375			
5	0.55486	0.49146	-0.08565	0.12048	0.63214	0.55486		0.63214
6	0.11587	0.10845	-0.03178	0.00685	1.07535			
7	0.20105	0.18279	-0.04156	0.02006	1.01711			
8	0.15375	0.12673	-0.00551	0.01191	1.04776			
9	0.17897	0.16809	-0.05072	0.01605	1.03809			
10	0.03107	0.03091	-0.01585	0.0005	1.11171			
11	-0.13459	-0.12951	0.04765	0.0092	1.0698			
12	-0.19697	-0.19633	0.10413	0.01947	1.05178			
13	0.2095	0.17665	-0.01468	0.02167	1.00547			
14	0.17426	0.15393	-0.02603	0.01521	1.03537			
15	-0.07135	-0.06911	0.02691	0.00262	1.0967			
16	-0.01571	-0.0157	0.00883	0.00013	1.11686			
17	-0.07989	-0.07981	0.04464	0.00328	1.10676			
18	-0.29741	-0.29707	0.16551	0.04294	0.98609			
19	-0.23329	-0.23267	0.12484	0.02701	1.02919			
20	-0.33754	-0.33477	0.16461	0.05392	0.93299			
21	0.14944	0.08023	0.05353	0.0113	1.06189			
22	-0.14016	-0.03609	-0.08733	0.01002	1.09571			
23	-0.16711	-0.11114	-0.03409	0.01403	1.04236			
24	-0.05441	-0.05418	0.02823	0.00153	1.1089			
25	-3.60816	1.545	-3.53963	6.08765	4.49185	-3.60816	6.08765	4.49185
26	-0.2059	-0.14494	-0.03121	0.02098	1.01073			
27	0.35762	0.25809	0.04525	0.05826	0.85647			
28	-0.14821	-0.14398	0.05747	0.01113	1.06469			
29	0.15046	0.02646	0.10325	0.01155	1.10503			
30	-0.15741	-0.15562	0.07365	0.01255	1.06677			
31	0.16017	0.07315	0.07093	0.01297	1.0627			
32	-0.02188	-0.00602	-0.01332	0.00025	1.12108			
33	0.08117	0.07983	-0.03571	0.00339	1.09709			

• الاستنتاج :-

1 - مقياس DFFITS :-

واضح من جدول (3-2) . أن هناك قيمتان مؤثرتان على قيم التكلفة لهذا المشروع والقيمة المطلقة لمقياس DFFITS أكبر من 0.492366 وهي المشاهدة رقم 5,25 حيث أن قيم الاختلاف لها تساوي:

$$DFFITS_5 = 0.55486 , DFFITS_{25} = -3.60816$$

2 مقياس DFBETAS :-

نلاحظ من الجدول (3-2) ، أن هناك مشاهدتين مؤثرتين على معاملات نموذج الانحدار البعض منها مؤثر على معامل واحد وكذلك على معاملان وأيضا هناك مشاهدات مؤثرة على جميع المعاملات. حيث أن المشاهدة 25 مؤثرة على جميع معاملات النموذج  $(B_2 B_1 B_0)$ . والمشاهدة رقم 5 مؤثرة على المعامل  $(B_0)$  .

مع ملاحظة أن المشاهدة رقم 25 كانت مشاهدة قاصية في المتغير المستقل (العمالة) والمشاهدة رقم 5 مشاهدة قاصية في المتغير التابع (التكلفة) .

3 مقياس مسافة كوك Di :-

نلاحظ من الجدول (3-2) . وجود مشاهدة واحدة مؤثرة على معاملات الانحدار وهي المشاهدة 25 . حيث أن القيمة المؤثرة بمقياس كوك Di تساوي :

$$D_{25} = 6.08765$$

وهي أكبر من  $F_{0.5,2,31} = 0.388662$

4 - مقياس COVRATIO :-

نلاحظ من الجدول (3-2) أنه يوجد قيمتان مؤثرتان على قيم الأخطاء المعيارية وهي المشاهدتان رقم 5,25 حيث أن قيمها تساوي :

$$COVRATIO_5 = 0.55486 , COVRATIO_{25} = 4.49185$$

• جدول تلخيص للمشاهدات القاصية والمؤثرة في المتغير التابع والمتغير المستقل .

المشاهدة i	قاصية في المتغير التابع Y	قاصية في المتغير المستقل X	مؤثرة DFFITS	مؤثرة DFBETAS	مؤثرة كوك Di	مؤثرة COVRATIO
5	✓	×	✓	✓	×	✓
25	×	✓	✓	✓	✓	✓

### • تطبيق (3-3) :-

بالعودة إلى بيانات ومشاهدات تطبيق (1-3) وتطبيق (2-3). سوف نقيس في هذا التطبيق أثر المشاهدات القاصية على متغير درجة تخثر الدم  $X_1$ ، و دليل الإنذار القياسي  $X_2$ ، و درجة اختبار وظيفة الأنزيم  $X_3$ . وأيضا القيم المؤثرة على الفترة التي يعيشها المريض بعد الجراحة المقدر  $\gamma$ ، وكذلك المشاهدات المؤثرة على معاملات النموذج والأخطاء المعيارية.

### • التحليل الإحصائي :-

باستخدام برامج الحزم الإحصائية Minitab و SPSS سوف نقوم باستخراج المشاهدات المؤثرة على المتغيرات المستقلة وكذلك المشاهدات المؤثرة على المتغير التابع والأخطاء المعيارية أيضا.

## Regression

Variables Entered/Removed(b)

Model	Variables Entered	Variables Removed	Method
1	$X_3, X_2, X_1(a)$	.	Enter

a All requested variables entered.

b Dependent Variable: Y

Model Summary(b)

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.912(a)	.833	.822	61.21854

a Predictors: (Constant),  $X_3, X_2, X_1$

b Dependent Variable: Y

ANOVA(b)

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	931546.037	3	310515.346	82.855	.000(a)
	Residual	187385.463	50	3747.709		
	Total	1118931.500	53			

a Predictors: (Constant),  $X_3, X_2, X_1$

b Dependent Variable: Y

#### Coefficients(a)

Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.
	B	Std. Error	Beta		
1 (Constant)	-659.179	55.674		-11.840	.000
X1	38.323	5.326	.423	7.196	.000
X2	4.568	.500	.531	9.144	.000
X3	4.485	.400	.656	11.208	.000

a. Dependent Variable: Y

#### Residuals Statistics(a)

	Minimum	Maximum	Mean	Std. Deviation	N
Predicted Value	-100.5444	520.8362	197.1667	132.57579	54
Std. Predicted Value	-2.246	2.441	.000	1.000	54
Standard Error of Predicted Value	8.566	32.976	15.663	5.736	54
Adjusted Predicted Value	-120.1273	501.9748	194.9479	131.73378	54
Residual	-91.90346	317.11111	.00000	59.46070	54
Std. Residual	-1.501	5.180	.000	.971	54
Stud. Residual	-1.525	5.617	.017	1.040	54
Deleted Residual	-94.85796	372.84814	2.21877	68.29340	54
Stud. Deleted Residual	-1.546	9.153	.085	1.439	54
Mahal. Distance	.056	14.397	2.944	3.084	54
Cook's Distance	.000	1.386	.040	.190	54
Centered Leverage Value	.001	.272	.056	.058	54

a. Dependent Variable: Y

#### Cumulative Distribution Function

F distribution with 4 DF in numerator and 50 DF in denominator

$$x \quad P(X \leq x) \\ 0.5 \quad 0.264194 = F_{0.5, 4, 50}$$

```
MTB > Let k1 = 2 * SQRT(4 / 54)
MTB > Let k2 = 2 / SQRT(54)
MTB > Let k3 = 1 + ((3 * 4) / 54)
MTB > Let k4 = 1 - ((3 * 4) / 54)
MTB > prin k1 k2 k3 k4
```

#### Data Display

$$K1 = 2\sqrt{\frac{p}{n}} = 0.544331$$

$$K2 = \frac{2}{\sqrt{n}} = 0.272166$$

$$K3 = 1 + \frac{3p}{n} = 1.22222$$

$$K4 = 1 - \frac{3p}{n} = 0.777778$$

• جدول (3-3) قيم الاختلاف في المتغير التابع الموفق ومعاملات الانحدار ومسافة كوك ومقياس نسبة . COV

I	DFFITS	DF B0	DF B1	DF B2	DF B3	Di	COVR
1	-0.11948	0.02749	-0.06211	0.0124	-0.02732	0.0036	1.06686
2	-0.00222	-0.00155	0.00087	0.00039	0.00106	0	1.1171
3	-0.19244	0.05432	-0.13652	0.05689	-0.05323	0.0093	1.06577
4	-0.03081	-0.01027	-0.00224	-0.00787	0.0248	0.00024	1.17591
5	0.37019	-0.26238	0.22227	0.00162	0.2886	0.03428	1.1431
6	0.08791	0.05252	0.00512	-0.07275	-0.01244	0.00197	1.1452
7	0.10367	0.07526	-0.0041	-0.06731	-0.04524	0.00273	1.11788
8	-0.11679	-0.04538	0.09141	-0.02778	0.00052	0.00346	1.12302
9	-0.27724	0.11539	-0.04956	-0.04727	-0.16804	0.0187	0.92486
10	-0.22765	-0.00085	0.14762	-0.10276	-0.07179	0.01304	1.10671
11	-0.03504	0.01873	-0.00561	-0.02598	-0.00735	0.00031	1.13909
12	0.20966	0.12454	0.04082	-0.08034	-0.15616	0.0111	1.13319
13	3.83739	-2.73503	0.13654	2.67653	2.42267	1.38628	<b>0.02364</b>
14	-0.0262	0.01301	0.00025	-0.019	-0.00858	0.00018	<b>1.13978</b>
15	-0.19313	0.02609	-0.13937	0.02244	0.03664	0.00937	1.07217
16	-0.18191	0.06002	-0.10668	-0.06434	0.0325	0.00832	1.07194
17	0.01771	0.00559	-0.00202	0.0079	-0.01447	0.00008	1.27517
18	0.94205	0.87341	-0.54831	-0.22844	-0.69753	0.20134	0.77691
19	-0.15432	0.05214	-0.1019	-0.02197	0.00032	0.006	1.06834
20	-0.08946	-0.00844	0.00108	0.02763	-0.03492	0.00203	1.08414
21	-0.02684	-0.01625	0.00692	0.0138	0.00248	0.00018	1.11419
22	0.32007	0.14205	-0.21956	0.15944	-0.16741	0.02576	1.1737
23	0.31868	0.15508	0.08828	-0.28191	-0.04604	0.02553	1.1677
24	-0.08754	0.01961	-0.00426	-0.01802	-0.03392	0.00194	1.08043
25	-0.08862	0.03455	-0.02889	0.01524	-0.06423	0.002	1.12188
26	-0.0581	-0.01925	0.00039	0.00776	0.01115	0.00086	1.09069
27	0.05203	0.01797	-0.01006	-0.02558	0.01465	0.00069	1.11127
28	0.6022	-0.42875	0.5585	0.07497	0.18256	0.09062	1.35238
29	0.12643	0.10347	-0.03749	-0.04201	-0.08492	0.00405	1.10882
30	0.01102	0.00217	-0.00134	0.00546	-0.00622	0.00003	1.13269
31	0.11867	0.08329	-0.09892	0.01046	-0.04691	0.00358	1.16423



I	DFFITS	DF B0	DF B1	DF B2	DF B3	Di	COVR
32	0.72133	0.2777	0.33642	-0.27417	-0.49849	0.12767	1.17663
33	-0.03516	-0.02118	0.0153	0.0064	0.00826	0.00032	1.1079
34	-0.04388	0.01998	-0.0025	-0.01709	-0.02473	0.00049	1.11811
35	-0.02981	-0.01779	0.00704	0.00778	0.01001	0.00023	1.10739
36	-0.13775	-0.00348	0.00761	0.05961	-0.08546	0.0048	1.10436
37	-0.27357	-0.08338	0.23205	-0.09471	-0.01511	0.01885	1.15199
38	0.54553	0.15865	-0.04962	-0.44223	0.25326	0.07481	1.43982
39	0.01808	0.00948	-0.00951	-0.00122	-0.00225	0.00008	1.11216
40	-0.09538	-0.02266	0.00791	0.04716	-0.03496	0.00231	1.09469
41	0.02569	0.01813	-0.00623	-0.01534	-0.00576	0.00017	1.12214
42	0.17073	0.0872	-0.06567	-0.12326	0.05123	0.00741	1.24128
43	0.11354	-0.08907	0.0827	0.05266	0.03607	0.00328	1.2288
44	-0.10339	-0.00972	-0.04279	0.04098	-0.00502	0.00271	1.07876
45	-0.25011	-0.01408	0.17536	-0.11428	-0.06456	0.01574	1.11756
46	-0.11459	-0.02483	-0.03905	0.08877	-0.02386	0.00334	1.13899
47	-0.05064	0.01494	0.01752	-0.01766	-0.03347	0.00065	1.16072
48	0.14072	-0.06173	-0.03817	0.09451	0.07054	0.00503	1.16316
49	-0.07516	-0.01946	0.03063	-0.01766	0.00454	0.00143	1.08888
50	-0.09205	0.0218	0.04566	-0.05075	-0.04396	0.00216	1.18891
51	-0.12952	-0.02444	-0.01507	-0.0504	0.09293	0.00426	1.14694
52	-0.16907	-0.02684	-0.00112	-0.08874	0.12126	0.00726	1.18813
53	-0.10509	0.01683	-0.04163	0.0251	-0.03829	0.00279	1.0772
54	-0.2449	0.13926	-0.19642	-0.07724	-0.00534	0.01512	1.14331

• الاستنتاج :-

### 1 مقياس DFFITS :-

واضح من جدول (3-3) . أن هناك 5 قيم مؤثرة على قيم الفترة التي يعيشها المريض بعد الجراحة. والقيم المطلقة لمقياس  $DFFITS$  الأكبر من  $2\sqrt{p/(n)} = 0.544331$  هي المشاهدات رقم 13، 18، 28، 32، 38 حيث أن قيمة الاختلاف لها تساوي:

$$DFFITS_{13} = 3.83739 , DFFITS_{18} = 0.94205 , DFFITS_{28} = 0.6022 , \\ DFFITS_{32} = 0.72133 , DFFITS_{38} = 0.54553$$

ولا ننسى أيضا إن المشاهدات 13، 18 كانت مشاهدات قاصية بالنسبة للمتغير التابع في الموضوع السابق.

### 2 مقياس DFBETAS :-

نلاحظ من الجدول (3-3) ، أن هناك 7 مشاهدات مؤثرة على معاملات نموذج الانحدار البعض منها مؤثر على معامل واحد وكذلك على معاملان وأيضا هناك مشاهدات مؤثرة على جميع المعاملات. حيث أن المشاهدة رقم 5 مؤثرة على  $B_3$  في النموذج وكذلك المشاهدة رقم 13 مؤثرة على  $B_0$  و  $B_2$  و  $B_3$  في النموذج وكذلك المشاهدة رقم 18 مؤثرة على  $B_0$  و  $B_1$  و  $B_3$  في النموذج و المشاهدة رقم 23 و 38 مؤثرة على  $B_2$  في النموذج وكذلك المشاهدة رقم 28 مؤثرة على  $B_0$  و  $B_1$  في النموذج و أما المشاهدة رقم 32 فهي مؤثرة على جميع معاملات النموذج  $(B_3 B_2 B_1 B_0)$ .

مع ملاحظة أن المشاهدات 13,28,32,38 كانت مشاهدات قاصية في المتغيرات (درجة تخثر الدم و دليل الإنذار القياسي و درجة اختبار وظيفة الأنزيم) .

### 3 - مقياس مسافة كوك Di :-

نلاحظ من الجدول (3-3) . وجود مشاهدة واحدة مؤثرة على معاملات الانحدار وهي المشاهدة رقم 13. حيث أن القيمة المؤثرة بمقياس كوك Di تساوي :

$$D_{13} = 1.38628$$

$$F_{0.5,2,31} = 0.388662 \text{ وهي أكبر من}$$

### 4 -مقياس COVRATIO :-

نلاحظ من الجدول (3-3) أنه يوجد 7 قيم مؤثرة على قيم الأخطاء المعيارية وهي المشاهدات رقم 13,17,18, 28,38,42,43, حيث أن قيمها تساوي :

$$COVRATIO_{13} = 0.02364 , COVRATIO_{17} = 1.27517 , COVRATIO_{18} = 0.77691$$

$$COVRATIO_{28} = 1.35238 , COVRATIO_{38} = 1.43982 , COVRATIO_{42} = 1.24128$$

$$COVRATIO_{43} = 1.2288$$

• جدول تلخيص للمشاهدات القاصية والمؤثرة في المتغير التابع والمتغير المستقل .

المشاهدة i	قاصية في المتغير التابع Y	قاصية في المتغير المستقل X	مؤثرة DFFITS	مؤثرة DFBETAS	مؤثرة Dj كوك	مؤثرة COVRATIO
5	X	X	X	✓	X	X
13	✓	✓	✓	✓	✓	✓
17	X	✓	X	X	X	✓
18	✓	X	✓	✓	X	✓
23	X	X	X	✓	X	X
28	X	✓	✓	✓	X	✓
32	X	✓	✓	✓	X	X
38	X	✓	✓	✓	X	✓
42	X	X	X	X	X	✓
43	X	X	X	X	X	✓

• تطبيق (3-4) :-

بالعودة إلى بيانات ومشاهدات تطبيق (1-4) وتطبيق (2-4) سوف نقيس في هذا التطبيق أثر المشاهدات القاصية على متغير الحسابات الترويجية  $X_1$  ، وعدد الحسابات المصرفية العاملة  $X_2$  ، وعدد الأصناف المنافسة  $X_3$  . وأيضا القيم المؤثرة على المبيعات المقدرة  $Y$  ، وكذلك المشاهدات المؤثرة على معاملات النموذج والأخطاء المعيارية .

• التحليل الإحصائي :-

باستخدام برامج الحزم الإحصائية Minitab و SPSS سوف نقوم باستخراج المشاهدات المؤثرة على

المتغيرات المستقلة وكذلك المشاهدات المؤثرة على المتغير التابع والأخطاء المعيارية أيضا .

## Regression

### Variables Entered/Removed(b)

Model	Variables Entered	Variables Removed	Method
1	X3, X1, X2(a)	.	Enter

a All requested variables entered.

b Dependent Variable: Y

### Model Summary(b)

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.998(a)	.997	.996	4.97952

a Predictors: (Constant), X3, X1, X2

b Dependent Variable: Y

### ANOVA(b)

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	89274.366	3	29758.122	1200.137	.000(a)
	Residual	272.752	11	24.796		
	Total	89547.117	14			

a Predictors: (Constant), X3, X1, X2

b Dependent Variable: Y

### Coefficients(a)

Model		Unstandardized Coefficients		Standardized Coefficients	T	Sig.
		B	Std. Error	Beta		
1	(Constant)	178.521	8.318		21.463	.000
	X1	2.106	.648	.054	3.250	.008
	X2	3.562	.099	.603	35.819	.000
	X3	-22.188	.529	-.707	-41.979	.000

a Dependent Variable: Y

### Residuals Statistics(a)

	Minimum	Maximum	Mean	Std. Deviation	N
Predicted Value	35.9812	341.3221	168.9467	79.85450	15
Std. Predicted Value	-1.665	2.159	.000	1.000	15
Standard Error of Predicted Value	1.758	3.438	2.497	.636	15
Adjusted Predicted Value	40.1747	342.6833	168.8631	79.24748	15
Residual	-7.54144	6.76184	.00000	4.41387	15
Std. Residual	-1.514	1.358	.000	.886	15
Stud. Residual	-1.694	1.727	.006	1.087	15
Deleted Residual	-9.43791	11.55609	.08355	6.77840	15
Stud. Deleted Residual	-1.879	1.929	.016	1.165	15
Mahal. Distance	.812	5.739	2.800	1.898	15
Cook's Distance	.000	.605	.157	.216	15
Centered Leverage Value	.058	.410	.200	.136	15

a Dependent Variable: Y

## Cumulative Distribution Function

F distribution with 4 DF in numerator and 11 DF in denominator

$$x \quad P(X \leq x) \\ 0.5 \quad 0.263387 = F_{0.5,4,11}$$

```
MTB > Let k1 = 2 * SQRT(4 / 15)
MTB > Let k2 = 2 / SQRT(15)
MTB > Let k3 = 1 + ((3 * 4) / 15)
MTB > Let k4 = 1 - ((3 * 4) / 15)
MTB > prin k1 k2 k3 k4
```

## Data Display

$$K1 = 2\sqrt{\frac{p}{n}} = 1.03280 \\ K2 = \frac{2}{\sqrt{n}} = 0.516398 \\ K3 = 1 + \frac{3p}{n} = 1.80000 \\ K4 = 1 - \frac{3p}{n} = 0.200000$$

- جدول (3-4) قيم الاختلاف في المتغير التابع الموفق ومعاملات الانحدار ومسافة كوك ومقياس نسبة . COV

I	DFFITS	DF B0	DF B1	DF B2	DF B3	Di	COVR
1	0.07521	0.03537	0.00914	-0.06055	0.00701	0.00155	1.87102
2	-0.23616	-0.10543	0.18269	-0.03928	0.06111	0.01502	1.70225
3	-1.12686	0.93139	-0.64157	-0.61448	-0.63951	0.30089	1.44541
4	-0.9423	-0.66758	0.6133	0.0854	0.49533	0.18046	0.54661
5	0.406	0.34442	-0.23734	-0.22341	-0.13858	0.04298	1.5403
6	1.73143	-0.8456	-0.74342	1.18226	0.921	0.60091	0.74639
7	-1.31294	0.08856	-0.76897	0.73536	-0.52681	0.39214	1.25138
8	1.70376	0.13504	1.17398	0.07906	-0.96559	0.60529	0.9246
9	0.07812	0.06274	-0.03205	-0.03933	-0.03062	0.00167	1.67509
10	-0.40956	0.02595	-0.09116	-0.24068	0.2279	0.04506	2.27741
11	-0.01594	0.00988	-0.00204	-0.00862	-0.00944	0.00007	1.73499
12	0.71749	-0.12699	-0.00407	-0.15493	0.50721	0.113	0.71545
13	0.13612	0.0723	0.02969	-0.02045	-0.1026	0.00505	1.73309
14	0.08501	0.03385	-0.00263	-0.05285	0.01541	0.00198	1.63912
15	-0.41713	-0.00496	0.24479	-0.1198	-0.11572	0.04292	1.09079

• الاستنتاج :-

**1 - مقياس DFFITS :-**

واضح من جدول (3-4) . أن هناك 4 قيم مؤثرة على قيم المبيعات. والقيم المطلقة لمقياس  $DFFITS$  الأكبر من  $2\sqrt{p/(n)} = 1.03280$  هي المشاهدات رقم 3,6,7,8 حيث أن قيمة الاختلاف لها تساوي:

$$DFFITS_3 = -1.12686 , \quad DFFITS_6 = 1.73143, \quad DFFITS_7 = -1.31294, \\ DFFITS_8 = 1.70376$$

**2 - مقياس DFBETAS :-**

نلاحظ من الجدول (3-4) ، أن هناك 4 مشاهدات مؤثرة على معاملات نموذج الانحدار ( قيمها أكبر من  $\frac{2}{\sqrt{n}} = 0.516398$  البعض منها مؤثر على معامل واحد وكذلك على معاملان وأيضا هناك مشاهدات مؤثرة على جميع المعاملات. حيث أن المشاهدة رقم 3 والمشاهدة رقم 6 مؤثرة على جميع معاملات النموذج  $(B_3, B_2, B_1, B_0)$ . وكذلك المشاهدة رقم 4 مؤثرة على  $B_0$  و  $B_1$  في النموذج ، وكذلك المشاهدة رقم 7 مؤثرة على  $B_1$  و  $B_2$  و  $B_3$  في النموذج وكذلك المشاهدة رقم 8 مؤثرة على  $B_1$  و  $B_3$  في النموذج .

**3 - مقياس مسافة كوك Di :-**

نلاحظ من الجدول (3-4) . وجود 4 قيم مؤثرة على معاملات الانحدار وهي المشاهدات رقم 3,6,7,8. حيث أن القيمة المؤثرة بمقياس كوك  $Di$  تساوي :

$$D_3 = 0.30089 , \quad D_6 = 0.60091 , \quad D_7 = 0.39214 , \quad D_8 = 0.60529$$

وهي أكبر من  $F_{0.5,4,11} = 0.263387$ .

**4 - مقياس COVRATIO :-**

نلاحظ من الجدول (3-4) أنه يوجد قيمتان مؤثرتان على قيم الأخطاء المعيارية وهي المشاهدات رقم 1,10 حيث أن قيمها تساوي :

$$COVRATIO_1 = 1.87102, \quad COVRATIO_{10} = 2.27741$$

• جدول تلخيص للمشاهدات القاصية والمؤثرة في المتغير التابع والمتغير المستقل .

المشاهدة i	قاصية في المتغير التابع Y	قاصية في المتغير المستقل X	مؤثرة DFFITS	مؤثرة DFBETAS	مؤثرة Cook Di	مؤثرة COVRATIO
1	X	X	X	X	X	X
3	X	X	X	✓	✓	X
6	X	X	X	✓	✓	X
7	X	X	X	✓	✓	X
8	X	X	X	✓	✓	X
10	X	X	X	X	X	✓

• (4) تشخيصات الارتباطية الخطية المتعددة – تضخم التباين .

• الارتباطية الخطية المتعددة وتأثيراتها :-

في تحليل الانحدار المتعدد ، يهتم المرء أيضا بطبيعة وأهمية العلاقات بين المتغيرات المستقلة فيما بينها . فإذا كانت المتغيرات المستقلة التي يشملها نموذج الانحدار مرتبطة فيما بينها فإنه يقال أنه يوجد ارتباط داخلي أو خطية متعددة فيما بينها . وعندما تكون المتغيرات المستقلة غير مرتبطة ، فإن التأثيرات المنسوبة لها ، بواسطة نموذج انحدار من المرتبة الأولى ، تبقى نفسها بصرف النظر عن أية متغيرات مستقلة يشملها النموذج .

وبصورة عامة ، عندما يكون متغيران مستقلان ( أو أكثر ) غير مرتبطين ( غير مرتبطة ) فإن المساهمة الهامشية لمتغير مستقل واحد في تخفيض مجموع مربعات الخطأ ، مع وجود المتغيرات الأخرى في النموذج ، تبقى نفسها بالضبط كما لو أن النموذج كان يقتصر على ذلك المتغير المستقل بمفرده .

• تشخيصات الارتباطية الخطية المتعددة:

من المشاكل الرئيسية التي تبرز في صورته تقليديه عندما تكون المتغيرات المستقلة المعتبرة في نموذج الانحدار مرتبطة في بينها ارتباطا عاليا :

1. إضافة أو حذف متغير مستقل يغير معاملات الانحدار بشكل كبير ( يتغير تأثير المتغيرات المستقلة الأخرى على المتغير التابع بشكل كبير ) .
2. تصبح الانحرافات المعيارية لمقدرات معاملات الانحدار كبيره عندما تكون المتغيرات المستقلة في نموذج الانحدار مرتبطة فيما بينها ارتباطا عاليا .
3. يتغير مجموع المربعات الإضافي المترافق مع متغير مستقل معتمدا على أي المتغيرات المستقلة الأخرى مشمول في النموذج.

4. فحص مصفوفة معاملات الارتباط الخطي البسيط بين أزواج المتغيرات المستقلة ( $R_{xx}$ ) حيث أن :

$$R_{xx} = \begin{bmatrix} 1 & r_{x1x2} & \cdots & r_{x1xp-1} \\ \vdots & & \ddots & \vdots \\ \cdots & & & 1 \end{bmatrix}, \quad r_{xh,x1} = \frac{\sum (X_{hi} - \bar{X})(X_{k1} - \bar{X})}{\sqrt{\sum (X_{hi} - \bar{X})^2} \sqrt{\sum (X_{k1} - \bar{X})^2}}, h, k = 1, 2, \dots, P-1$$

حيث أن :

$r_{xj,xk}$  : معامل الارتباط الخطي بين  $X_k$  و  $X_j$ .

وبملاحظه قيم معاملات الارتباط إذا وجد أن هناك ارتباطا قويا بين متغيرين مستقلين دل على ذلك على احتمال وجود ارتباط خطي متعدد.

5. توفيق عدد (P-1) من نماذج الانحدار لكل متغير مستقل على بقية المتغيرات المستقلة بالطريقة التالية :

$$X_1 = b_0 + b_2X_2 + \cdots + b_{p-1}X_{p-1} + \varepsilon_i$$

$$X_2 = b_0 + b_1X_1 + \cdots + b_{p-1}X_{p-1} + \varepsilon_i$$

.

.

$$X_{p-1} = b_0 + b_2X_2 + \cdots + b_{p-2}X_{p-2} + \varepsilon_i$$

فإذا كانت قيمه احد معاملات التحديد ( $R^2$ ) لهذه النماذج تقترب من الواحد الصحيح دل ذلك على وجود ارتباط خطي متعدد.

وهذه الطرق بعضها جيد والبعض الآخر اقل جودة ولذلك نلجأ للطريقة الأساسية والواسعة الاستخدام مثل عامل تضخم التباين (VIF).

#### • عامل تضخم التباين (VIF) :-

عامل تضخم التباين هو إحدى الطرق الرسمية المستخدمة على نطاق واسع للكشف عن وجود الارتباطية الخطية المتعددة وتقيس هذه العوامل مدى تضخم تباينات معاملات الانحدار المقدرة بالمقارنة مع حاله عدم وجود صله خطيه بين المتغيرات المستقلة ويتم حساب عامل تضخم التباين بالعلاقة التالية :

$$VIF_k = (1 - R_k^2)^{-1}, k = 1, 2, \dots, p-1, \quad 1 \leq VIF_k \leq \infty, \quad 0 \leq R_k^2 \leq 1$$

حيث أن :-

$VIF_k$  : عامل تضم التباين للمتغير المستقل رقم (k).



$R_k^2$  : معامل التحديد لنموذج انحدار المتغير المستقل رقم  $(k)$  على بقية المتغيرات المستقلة  $(p-2)$  .

وإذا كان عامل تضخم التباين كبيراً كانت هذه دلالة واضحة على وجود ارتباط خطي متعدد.

### • طريقة الكشف عن وجود ارتباط خطي متعدد .

للكشف عن الارتباط الخطي المتعدد نقوم باستخراج عامل تضخم التباين  $(VIF_k)$  ومقارنته مع القيمة (10) . فإذا كان عامل التضخم أكبر من هذه القيمة فإن ذلك يدل على وجود ارتباط خطي متعدد مرتفع . أي إذا كان :

$$VIF_k > 10 , \quad k = 1, 2, \dots, p-1$$

وأيضاً إذا كانت قيمة معامل التحديد  $(R_k^2 = 0.90)$  لنموذج انحدار المتغير المستقل رقم  $k$  على بقية المتغيرات المستقلة  $(p-2)$  نلاحظ ما يلي :

- يأخذ عامل تضخم التباين قيم أكبر أو تساوي الواحد أي أن  $VIF_k > 1$  .
- في حالة وجود ارتباط خطي تام بين المتغير المستقل رقم  $(k)$  وبقية المتغيرات المستقلة  $(p-2)$  فإن عامل تضخم التباين يساوي ما لانهاية  $VIF_k = \infty$  .
- في حالة عدم وجود ارتباط خطي بين المتغير المستقل رقم  $(k)$  وبقية المتغيرات المستقلة  $(p-2)$  فإن قيمة عامل تضخم التباين تساوي واحد  $VIF_k = 1$  .

### - متوسط قيم عوامل تضخم التباين لمعاملات الانحدار .

يقدم متوسط القيم  $VIF$  معلومات عن خطورة الخطية المتعددة بدلالة المسافة بين معاملات الانحدار المعيارية المقدرة  $b'_k$  وبين القيم الحقيقية  $B'_k$  . ويمكن تبين أن القيمة المتوقعة لمجموع مربعات الخطأ هذه  $(b'_k - B'_k)^2$  معطاة بالعلاقة :-

$$1 - E \left\{ \sum_{k=1}^{p-1} (b'_k - B'_k)^2 \right\} = (\sigma')^2 \sum_{k=1}^{p-1} (VIF)_k$$

وعندما لا توجد صلة خطية لأي متغير  $X$  بالمتغيرات المستقلة الأخرى في نموذج الانحدار ، أي  $R_k^2 = 0$  فعندئذ ،  $(VIF)_k = 1$  وتصبح العلاقة كما يلي :

$$2 - E \left\{ \sum_{k=1}^{p-1} (b'_k - B'_k)^2 \right\} = (\sigma')^2 (p-1)$$

وتقدم نسبة النتيجة السابقتين (1) (2) معلومات مفيدة عن تأثير الخطية المتعددة على مجموع مربعات الخطأ :

$$\frac{(\sigma')^2 \sum_{k=1}^{p-1} (VIF)_k}{(\sigma')^2 (p-1)} = \frac{\sum_{k=1}^{p-1} (VIF)_k}{(p-1)}$$

ونلاحظ أن هذه النسبة هي متوسط قيم VIF وسنرمز لها بالرمز  $\overline{(VIF)}$  :

$$\overline{(VIF)} = \frac{\sum_{k=1}^{p-1} (VIF)_k}{(p-1)}$$

فإذا كانت المتغيرات المستقلة متعامدة أي لا يوجد ارتباط خطي ، فإن المتوسط يساوي واحد صحيح .لذلك نجد أنه كلما زادت قيمة متوسط عوامل تضخم التباين عن الواحد دل ذلك على وجود مشكلة الارتباطية الخطية المتعددة بين المتغيرات المستقلة .

#### - تطبيق (4-1) :

بالرجوع إلى تطبيق (1-1) في تجربة لعميل من قسم الميكانيكية لقياس الجهد الناتج عن هذه المحركات في توليفات مختلفة من السرعة X1، وكذلك قياس تمديد الاستشعار X2 لعينة من 18 محرك . سنقيس عوامل تضخم التباين .

#### - التحليل الإحصائي :

لحساب عامل تضخم التباين نقوم ببناء نموذج انحدار كل متغير من المتغيرات المستقلة على الآخر ونحسب معامل تضخم التباين :

#### • الطريقة الأولى :

نحسب عامل تضخم التباين من العلاقة التالية :

$$VIF_k = (1 - R_k^2)^{-1}$$

نوجد معامل التحديد للمتغير X1 والمتغير X2 كما يلي :

1 - انحدار X1 على X2 :

**Regression Analysis: X1 versus X2**

The regression equation is  
 $X1 = 7045 + 42259 X2$

Predictor	Coef	SE Coef	T	P
Constant	7044.7	325.1	21.67	0.000
X2	42259	30718	1.38	0.188

S = 902.918    R-Sq = 10.6%    R-Sq(adj) = 5.0%

#### Analysis of Variance

Source	DF	SS	MS	F	P
Regression	1	1542970	1542970	1.89	0.188
Residual Error	16	13044173	815261		
Total	17	14587143			

إذا كان هناك فقط متغيرين  $X1, X2$  فإن :

$$R_1^2 = R_2^2 = 0.106$$

وبعد إيجاد معامل التحديد لكل متغير نوجد عامل تضخم التباين لكل متغير كما يلي :

$$VIF = (1 - 0.106)^{-1} = 1.118568233$$

• للمتغير  $X1$

$$VIF = (1 - 0.106)^{-1} = 1.118568233$$

• للمتغير  $X2$

#### - الطريقة الثانية : -

1 - باستخدام برنامج SAS نوجد عامل تضخم التباين لكل متغير من المتغيرات المستقلة كما يلي :

وباستخدام الأمر هذا يمكننا استخراج عامل تضخم التباين:

```
proc reg data=Work.e1;
  model Y = X1 X2 / vif;
run;
```

The REG Procedure  
 Model: MODEL1  
 Dependent Variable: y  
 Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
--------	----	----------------	-------------	---------	--------

Model	2	5.69423	2.84711	183.54	<.0001
Error	15	0.23268	0.01551		
Corrected Total	17	5.92691			

<b>Root MSE</b>	0.12455	<b>R-Square</b>	0.9607
<b>Dependent Mean</b>	1.92222	<b>Adj R-Sq</b>	0.9555
<b>Coeff Var</b>	6.47939		

### Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	t Value	Pr >  t	Variance Inflation
Intercept	1	-1.64129	0.24704	-6.64	<.0001	0
x1	1	0.00055571	0.00003448	16.11	<.0001	1.11829
x2	1	-67.39589	4.48082	-15.04	<.0001	1.11829

2 - باستخدام برنامج MINITAB نوجد عامل تضخم التباين لكل متغير من المتغيرات المستقلة كما يلي :

```
MTB > Regress 'Y' 2 'X1' 'X2';
SUBC> Constant;
SUBC> VIF;
SUBC> Brief 2.
```

### Regression Analysis: Y versus X1; X2

The regression equation is  

$$Y = -1.64 + 0.000556 X1 - 67.4 X2$$

Predictor	Coef	SE Coef	T	P	VIF
Constant	-1.6413	0.2470	-6.64	0.000	
X1	0.00055571	0.00003448	16.11	0.000	1.1
X2	-67.396	4.481	-15.04	0.000	1.1

S = 0.124548    R-Sq = 96.1%    R-Sq(adj) = 95.6%

### Analysis of Variance

Source	DF	SS	MS	F	P
Regression	2	5.6942	2.8471	183.54	0.000
Residual Error	15	0.2327	0.0155		
Total	17	5.9269			

Source	DF	Seq SS
X1	1	2.1849
X2	1	3.5094

- جدول (1-4) لحساب معامل تضخم التباين :

$(VIF)_k$	$R_k^2$	$(VIF)_k > 10$	متغير
1.11829	0.106	NO	X1
1.11829	0.106	NO	X2
$\overline{(VIF)} = 1.11829$			$\sum (VIF)_k = 2.23658$

- الاستنتاج :-

نلاحظ أن جميع قيم عوامل تضخم التباين أقل من القيمة 10 لذلك لا يوجد مشكلة الارتباطية الخطية المتعددة بين متغيرات السرعة وتمديد الاستشعار .

- تطبيق (2-4) : -

بالرجوع إلى تطبيق (1-2) في بيانات لمشروع تنفيذي لمعرفة تكلفة أو تعويضات هذا المشروع و بعض المتغيرات المؤثرة عليه مثل المبيعات والعمالة والأرباح. سنقيس عوامل تضخم التباين .

- التحليل الإحصائي : -

1 - باستخدام برنامج SAS نوجد عامل تضخم التباين لكل متغير من المتغيرات المستقلة كما يلي :

وباستخدام الأمر هذا يمكننا استخراج عامل تضخم التباين:

```
proc reg data=Work.e2;
  model Y = X1 X2 X3 / vif;
run;
```

The REG Procedure  
Model: MODEL1  
Dependent Variable: Y Y  
Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
--------	----	----------------	-------------	---------	--------

Model	3	185131	61710	5.72	0.0033
Error	29	312821	10787		
Corrected Total	32	497952			

<b>Root MSE</b>	103.86014	<b>R-Square</b>	0.3718
<b>Dependent Mean</b>	410.48485	<b>Adj R-Sq</b>	0.3068
<b>Coeff Var</b>	25.30182		

### Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	t Value	Pr >  t	Variance Inflation
Intercept	1	344.45486	25.22098	13.66	<.0001	0
X1	1	0.01413	0.01418	1.00	0.3272	10.82538
X2	1	-0.12704	0.13092	-0.97	0.3399	13.71678
X3	1	0.00135	0.00097341	1.39	0.1759	9.79928

2 - باستخدام برنامج MINITAB نوجد عامل تضخم التباين لكل متغير من المتغيرات المستقلة كما يلي :

```
MTB > Regress 'Y' 3 'X1'-'X3';
SUBC> Constant;
SUBC> VIF;
SUBC> Brief 2.
```

### Regression Analysis: Y versus X1; X2; X3

The regression equation is  
 $Y = 344 + 0.0141 X1 - 0.127 X2 + 0.00135 X3$

Predictor	Coef	SE Coef	T	P	VIF
Constant	344.45	25.22	13.66	0.000	
X1	0.01413	0.01418	1.00	0.327	10.8
X2	-0.1270	0.1309	-0.97	0.340	13.7
X3	0.0013506	0.0009734	1.39	0.176	9.8

S = 103.860    R-Sq = 37.2%    R-Sq(adj) = 30.7%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	3	185131	61710	5.72	0.003
Residual Error	29	312821	10787		
Total	32	497952			

Source	DF	Seq SS
X1	1	163582
X2	1	784
X3	1	20766

- جدول (2-4) لحساب معامل تضخم التباين :

$(VIF)_k$	$R_k^2$	$(VIF)_k > 10$	متغير
10.82538	0.908	YES	X1
13.71678	0.927	YES	X2
9.79928	0.898	NO	X3
$\overline{(VIF)} = 11.44714$			$\sum (VIF)_k = 34.34144$

- الاستنتاج :-

نلاحظ من الجدول السابق أن عوامل تضخم التباين للمتغير X1 و X2 أكبر من القيمة 10 وهذا يعني وجود مشكلة الارتباطية الخطية المتعددة في وجود المتغيرين X1 و X2 ، وكذلك نلاحظ أن متوسط عوامل التضخم يساوي 11.44714 وهذا أيضا فيه دلالة على وجود المشكلة لذلك لا بد من حل مشكلة الارتباطية الخطية المتعددة وهناك عدة طرق للتغلب على هذه المشكلة . فعلى سبيل المثال قد يكون من المناسب إلغاء المتغير المستقل (X2) حيث أن له أكبر قيمة معامل تضخم  $VIF_2 = 13.71678$  ثم نقوم بإعادة حساب انحدار Y على المتغيرين الباقيين X1,X3 ونحسب من جديد عن طريق MINITAB معامل تضخم التباين للمتغيرين X1,X3 :

```
MTB > Regress 'Y' 2 'X1' 'X3';
SUBC> Constant;
SUBC> VIF;
SUBC> Brief 2.
```

### Regression Analysis: Y versus X1; X3

The regression equation is  
 $Y = 354 + 0.0059 X1 + 0.000845 X3$

Predictor	Coef	SE Coef	T	P	VIF
Constant	354.47	22.99	15.42	0.000	
X1	0.00594	0.01139	0.52	0.606	7.0
X3	0.0008450	0.0008214	1.03	0.312	7.0

S = 103.759    R-Sq = 35.1%    R-Sq(adj) = 30.8%

#### Analysis of Variance

Source	DF	SS	MS	F	P
Regression	2	174974	87487	8.13	0.002
Residual Error	30	322978	10766		
Total	32	497952			

Source	DF	Seq SS
X1	1	163582
X3	1	11392

بعد حذف المتغير المستقل X2 نجد أن عوامل تضخم التباين للمتغيرين المستقلين X1, X3 أقل من القيمة 10 وهذا يعني أنه لا يوجد مشكلة الارتباطية الخطية المتعددة بين المتغيرات المستقلة (X1, X3) المبيعات وتكلفة العمال).

#### - تطبيق (3-4) :-

بالرجوع إلى تطبيق (1-3) في بيانات لوحدة جراحة في مستشفى للفترة التي يعيشها المريض بعد الجراحة و بعض المتغيرات المؤثرة عليه مثل درجة تخثر الدم X1 و دليل الإنذار القياسي X2 و درجة اختبار وظيفة الإنزيم X3 و درجة اختبار وظيفة الكبد X4 . سنقيس عوامل تضخم التباين .

#### - التحليل الإحصائي :-

1 - باستخدام برنامج SAS نوجد عامل تضخم التباين لكل متغير من المتغيرات المستقلة كما يلي :

وباستخدام الأمر هذا يمكننا استخراج عامل تضخم التباين:

```
proc reg data=Work.e3;  
  model Y = X1 X2 X3 X4 / vif;  
run;
```

The REG Procedure  
Model: MODEL1  
Dependent Variable: Y Y  
Analysis of Variance



Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	4	936265	234066	62.79	<.0001
Error	49	182667	3727.89719		
Corrected Total	53	1118932			

<b>Root MSE</b>	61.05651	<b>R-Square</b>	0.8367
<b>Dependent Mean</b>	197.16667	<b>Adj R-Sq</b>	0.8234
<b>Coeff Var</b>	30.96695		

## Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	t Value	Pr >  t	Variance Inflation
Intercept	1	-621.59755	64.80043	-9.59	<.0001	0
X1	1	33.16383	7.01727	4.73	<.0001	1.79901
X2	1	4.27186	0.56338	7.58	<.0001	1.28922
X3	1	4.12574	0.51116	8.07	<.0001	1.67803
X4	1	14.09156	12.52533	1.13	0.2661	2.55533

2 - باستخدام برنامج MINITAB نوجد عامل تضخم التباين لكل متغير من المتغيرات المستقلة كما يلي :

```
MTB > Regress 'Y' 4 'X1'-'X4';
SUBC> Constant;
SUBC> VIF;
SUBC> Brief 2.
```

Regression Analysis: Y versus X1; X2; X3; X4

The regression equation is

$$Y = -622 + 33.2 X1 + 4.27 X2 + 4.13 X3 + 14.1 X4$$

Predictor	Coef	SE Coef	T	P	VIF
Constant	-621.60	64.80	-9.59	0.000	
X1	33.164	7.017	4.73	0.000	1.8
X2	4.2719	0.5634	7.58	0.000	1.3
X3	4.1257	0.5112	8.07	0.000	1.7
X4	14.09	12.53	1.13	0.266	2.6

S = 61.0565    R-Sq = 83.7%    R-Sq(adj) = 82.3%

## Analysis of Variance

Source	DF	SS	MS	F	P
Regression	4	936265	234066	62.79	0.000
Residual Error	49	182667	3728		
Total	53	1118931			

Source	DF	Seq SS
X1	1	155274
X2	1	305511
X3	1	470760
X4	1	4719

### - جدول (3-4) لحساب معامل تضخم التباين :

متغير	$(VIF)_k > 10$	$R_k^2$	$(VIF)_k$
X1	NO	0.444	1.79901
X2	NO	0.224	1.28922
X3	NO	0.404	1.67803
X4	NO	0.609	2.55533
$\sum (VIF)_k = 7.32159$			$\overline{(VIF)} = 1.8303975$

### - الاستنتاج :-

نلاحظ أن جميع قيم عوامل تضخم التباين أقل من القيمة 10 لذلك لا يوجد مشكلة الارتباطية الخطية المتعددة بين المتغيرات المستقلة .

### - تطبيق (4-4) :-

بالرجوع إلى تطبيق (1-4) في بيانات لمنتج ألواح أسفلتية لكسوة السقوف . و المتغيرات المؤثرة عليها مثل الحسابات الترويجية ( X1 بالآلاف الدولارات ) ، عدد الحسابات المصرفية العاملة X2. عدد الأصناف المنافسة X3. وإمكانات المنطقة X4 لكل منطقة من المناطق. سنقيس عوامل تضخم التباين .

### - التحليل الإحصائي :-

1 - باستخدام برنامج SAS نوجد عامل تضخم التباين لكل متغير من المتغيرات المستقلة كما يلي :

وباستخدام الأمر هذا يمكننا استخراج عامل تضخم التباين :

```
proc reg data=Work.e4;
  model Y = X1 X2 X3 X4 / vif;
run;
```

### The REG Procedure

Model: MODEL1

Dependent Variable: Y Y

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	4	89285	22321	851.72	<.0001
Error	10	262.07283	26.20728		
Corrected Total	14	89547			

  

Root MSE	5.11930	R-Square	0.9971
Dependent Mean	168.94667	Adj R-Sq	0.9959
Coeff Var	3.03013		

### Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	t Value	Pr >  t	Variance Inflation
Intercept	1	177.22856	8.78738	20.17	<.0001	0
X1	1	2.17021	0.67370	3.22	0.0092	1.02657
X2	1	3.53801	0.10915	32.41	<.0001	1.16514
X3	1	-22.15834	0.54537	-40.63	<.0001	1.03202
X4	1	0.20354	0.31886	0.64	0.5376	1.17832

2 - باستخدام برنامج MINITAB نوجد عامل تضخم التباين لكل متغير من المتغيرات المستقلة كما يلي :

```
MTB > Regress 'Y' 4 'X1'-'X4';
SUBC> Constant;
SUBC> VIF;
SUBC> Brief 2.
```

### Regression Analysis: Y versus X1; X2; X3; X4

The regression equation is

$$Y = 177 + 2.17 X_1 + 3.54 X_2 - 22.2 X_3 + 0.204 X_4$$

Predictor	Coef	SE Coef	T	P	VIF
Constant	177.229	8.787	20.17	0.000	
X1	2.1702	0.6737	3.22	0.009	1.0
X2	3.5380	0.1092	32.41	0.000	1.2
X3	-22.1583	0.5454	-40.63	0.000	1.0
X4	0.2035	0.3189	0.64	0.538	1.2

S = 5.11930    R-Sq = 99.7%    R-Sq(adj) = 99.6%

#### Analysis of Variance

Source	DF	SS	MS	F	P
Regression	4	89285	22321	851.72	0.000
Residual Error	10	262	26		
Total	14	89547			

Source	DF	Seq SS
X1	1	1074
X2	1	44505
X3	1	43695
X4	1	11

#### - جدول (4-4) لحساب معامل تضخم التباين :

$(VIF)_k$	$R_k^2$	$(VIF)_k > 10$	متغير
1.02657	0.026	NO	X1
1.16514	0.142	NO	X2
1.03202	0.031	NO	X3
1.17832	0.151	NO	X4
$\overline{(VIF)} = 1.1005125$			$(VIF)_k = 4.40205$

#### - الاستنتاج :-

نلاحظ أن جميع قيم عوامل تضخم التباين أقل من القيمة 10 لذلك لا يوجد مشكلة الارتباطية الخطية المتعددة بين المتغيرات المستقلة.

- جون نتير وآخرون "نماذج إحصائية خطية تطبيقية" الجزء الأول (الانحدار). مترجم للعربية بواسطة / أ.د. أنيس كنجو أ.د. عبد الحميد الزيد د. إبراهيم الواصل د. الحسيني راضي. جامعة الملك سعود قسم الإحصاء وبحوث العمليات – كلية العلوم- (1421-2000م) .
- " نماذج الانحدار الخطي " . الطبعة الأولى. الرياض. جامعة الملك سعود. قسم الإحصاء وبحوث العمليات للدكتور/ د. عدنان بري – كلية العلوم - كتاب إلكتروني لمقرر 335 إحص. (1428 هـ).
- عبد العزيز المطيري . مشروع بحث بإشراف أ.د. عبدالله الشبيحة بعنوان " أساليب اكتشاف المشاهدات الشاذة والمؤثرة على الانحدار الخطي " . الرياض. جامعة الملك سعود. قسم الإحصاء وبحوث العمليات. (1431هـ).

## Reference.

- Ronald E. Walpole and Raymond H. Myers “ **Probability and statistics for Engineers and Scientists** ” (fifth edition) .
- John Neter and William Wasserman and Michael H. Kutner “ **Applied Linear Statistical Models** ” (third edition). “

## Web Sits :

- Statisticians Arabs : <http://www.arabicstat.com>
- Wikipedia : <http://www.wikipedia.org>

