جامعة الملك عبد العزيز
كلية العلوم
قسم الاحصاء

المحاضرة الثالثة

خطوات اختبار هل البيانات تتبع التوزيع الطبيعي :

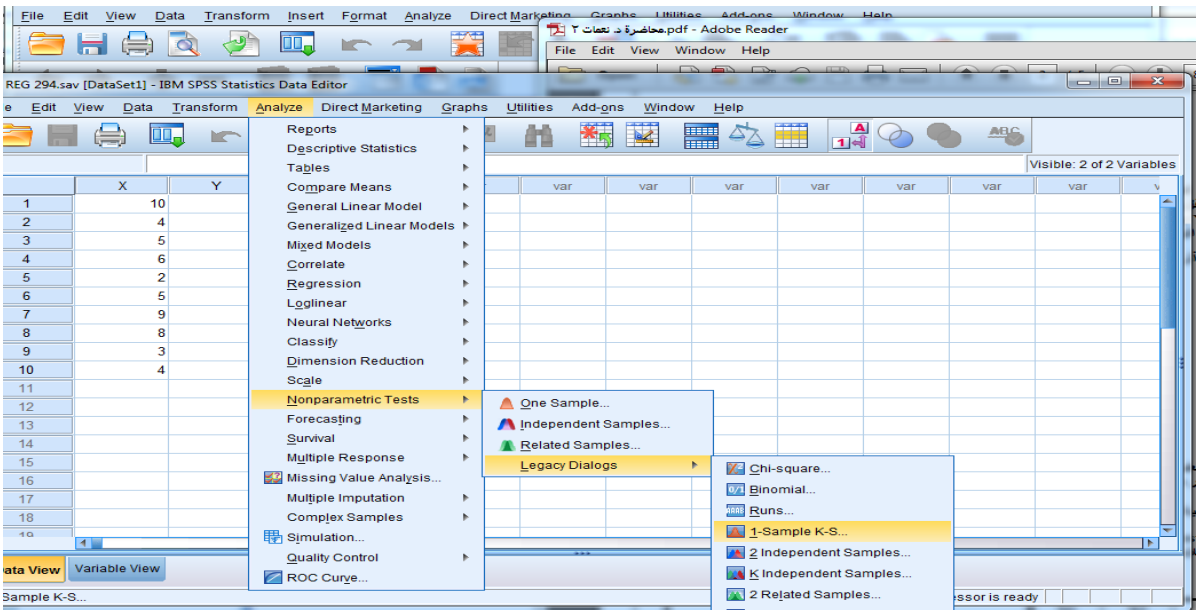
الفروض:

البيانات تتبع التوزيع الطبيعي: H_0

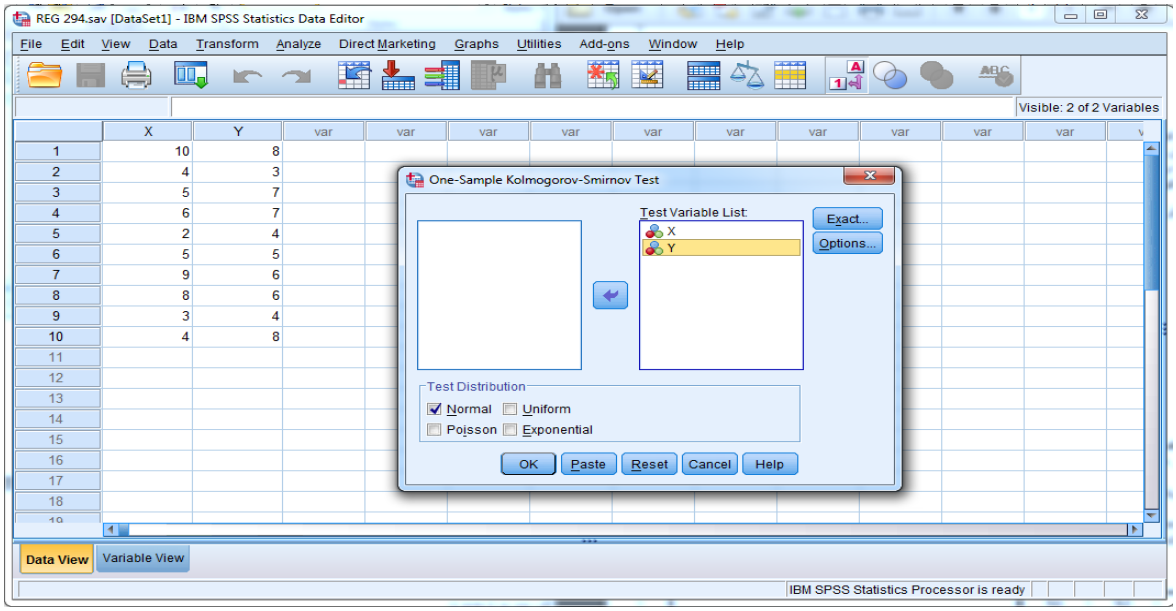
البيانات لا تتبع التوزيع الطبيعي: H_1

نتبع الخطوات التالية كما هو مبين بالشكل:

Analyze → Nonparametric - tests →Legacy dialog → 1- sample k.s test



تظهر لنا الشاشة التالية:



ندخل المتغيرين للخانة test variable list ثم نقوم باختيار Normal فيظهر الجدول التالي

One-Sample Kolmogorov-Smirnov Test

		X	Y
N		10	10
Normal Parameters[a,b]	Mean	5.60	5.80
	Std. Deviation	2.633	1.751
Most Extreme Differences	Absolute	.190	.153
	Positive	.190	.148
	Negative	-.119-	-.153-
Kolmogorov-Smirnov Z		.601	.485
Asymp. Sig. (2-tailed)		.863	.973

a. Test distribution is Normal.

b. Calculated from data.

كيفية تفسير النتائج:

يتم النظر في الجدول الى الرقم ((Asymp. Sig. (2-tailed) ويتم تفسيره كالتالي:

a. اذا كانت قيمة ال .sig اكبر من الرقم (α الفا – مستوى المعنوية) فإننا لانرفض الفرض H_0 أي أن البيانات تتبع توزيع طبيعي

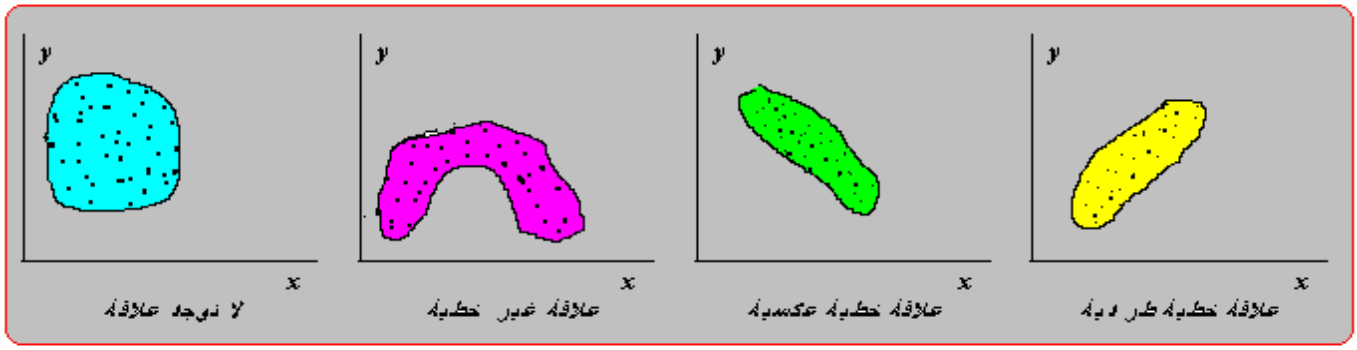
b. اذا كانت قيمة ال .sig اصغر من الرقم (α الفا – مستوى المعنوية) فإن نرفض الفرض H_0 أي ان البيانات تتبع التوزيع الطبيعي

من المثال نجد أن ال (Asymp. Sig. (2-tailed) للمتغيرين X,Y اكبر من ال (α الفا حيث $\alpha = 0.05$) وعليه فإننا لانرفض الفرض H_0 أي أن البيانات تتبع توزيع طبيعي

ملاحظات خاصة بمعاملات الارتباط:

يعتبر معامل الارتباط بيرسون اكثر دقة مقارنة بمعامل ارتباط سبيرمان وذلك لان معامل ارتباط بيرسون يعتمد على البيانات الاصلية بينما يعتمد معامل ارتباط سبيرمان على رتب البيانات

اشكال مختلفة لشكل الانتشار



<u>مثال</u>

إدرسي العلاقة بين المتغيرات التالية وعلقي علي النتائج، حيث المتغير الأول X1 هو درجات تمثل علاقة الطفل بجيرانه – المتغير الثاني X2 درجات لعلاقته بالاصدقاء – المتغير الثالث X3 درجات لعلاقته بالأهل

X1	X2	X3
8	6	16
11	12	18
19	12	20
22	9	11
11	16	9
7	14	20
24	15	4
15	16	8
12	20	24
6	5	23

الحل:قبل تنفيذ أسلوب الارتباط المعلمي بيرسون يجب أولا التأكد من شروط تطبيقه فإذا لم تتحقق ننفذ ارتباط سبيرمان , الشروط هي أ)البيانات كمية ب)البيانات تتبع التوزيع الطبيعي ولايمكن التغاضي عن ذلك الشرط لضمان العلاقة الخطية

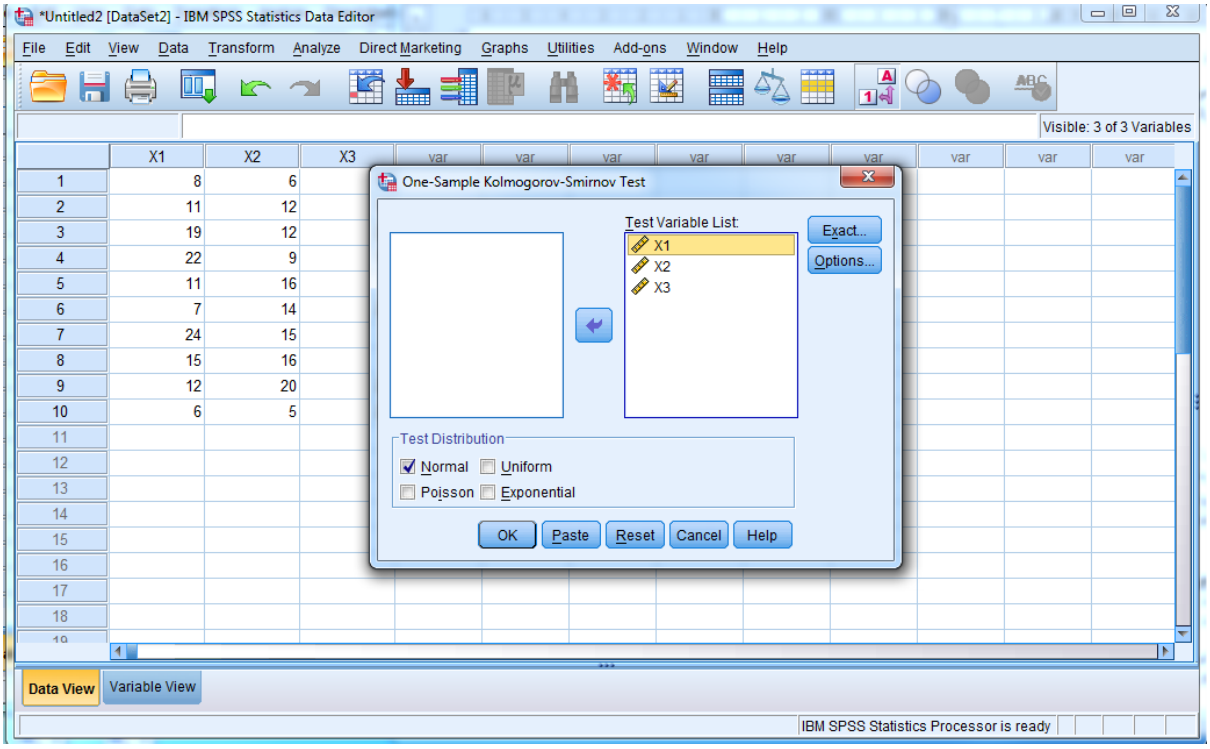
<u>الفروض:</u>

البيانات تتبع التوزيع الطبيعي: H0

البيانات لا تتبع التوزيع الطبيعي: H1

نتبع الخطوات التالية:

Analyze → Nonparametric - tests →Legacy dialog → 1- sample k.s test



النتائج

One-Sample Kolmogorov-Smirnov Test

		X1	X2	X3
N		10	10	10
Normal Parameters[a,b]	Mean	13.50	12.50	15.30
	Std. Deviation	6.311	4.720	6.881
Most Extreme Differences	Absolute	.194	.158	.153
	Positive	.194	.129	.134
	Negative	-.117	-.158	-.153
Kolmogorov-Smirnov Z		.613	.499	.483
Asymp. Sig. (2-tailed)		.846	.965	.974

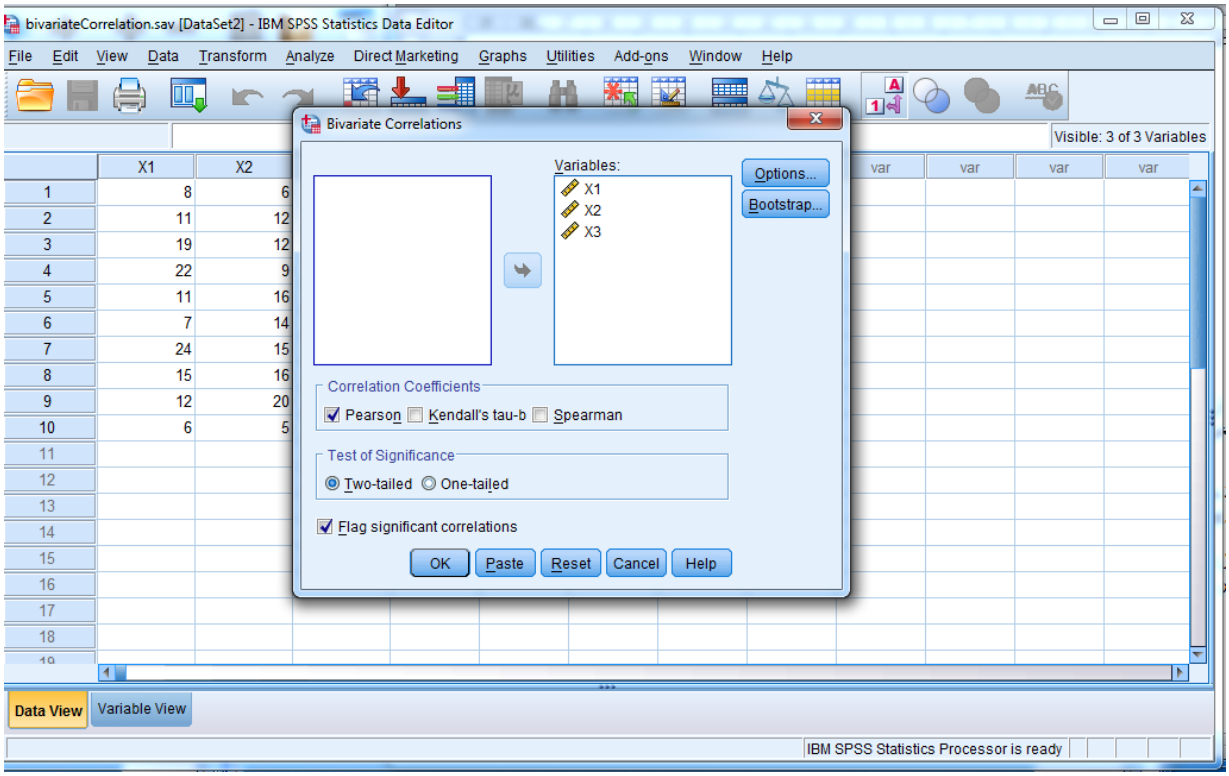
a. Test distribution is Normal.
b. Calculated from data.

نجد أن sig = .846, .965, 974 اكبر من الفا (α) , (حيث α = 0.05), إذن لانرفض الفرض العدمي أي المتغيرات الثلاث تتبع التوزيع الطبيعي وبذلك تحقق شرطي استخدام معامل الارتباط بيرسون

تنفيذ خطوات الارتباط

(1

Analyze → Correlate → bivariate



2)ندخل المتغيرات الثلاث الى الخانة variables

3) نضع علامة √ أمام بيرسون (في حالة عدم تحقق شرط الطبيعية نختار سبيرمان) فتظهر مصفوفة الارتباط التالية، و التي نستطيع معرفة قوة واتجاه العلاقة بين كل متغيرين منها

Correlations				
		X1	X2	X3
X1	Pearson Correlation	1	.233	-.605
	Sig. (2-tailed)		.517	.064
	N	10	10	10
X2	Pearson Correlation	.233	1	-.186
	Sig. (2-tailed)	.517		.606
	N	10	10	10
X3	Pearson Correlation	-.605	-.186	1
	Sig. (2-tailed)	.064	.606	
	N	10	10	10

تفسير النتائج :

1- الأرقام الموجودة علي القطر لاتفسر لأنها تمثل العلاقة بين المتغير ونفسه وهو ارتباط تام

2- بين X1 , X2 علاقة طردية (لأن اشارة معامل الارتباط موجبة)وضعيفة(لأن قيمة معامل الارتباط 0.223)ولكنه ارتباط غيرمعنوي لأن Sig. (2-tailed) = 0.517 وهذه القيمة اكبر من $\alpha = 0.05$

3- بين X1 , X3 علاقة عكسية (لأن اشارة معامل الارتباط سالبة)ومتوسطة (لأن قيمة معامل الارتباط 0.605) ولكنه ارتباط غيرمعنوي لأن Sig. (2-tailed) = 0.064 وهذه القيمة اكبر من $\alpha = 0.05$

4- بين X2 , X3 علاقة عكسية(لأن اشارة معامل الارتباط سالبة)وضعيفة (لأن قيمة معامل الارتباط 0.186) ولكنه ارتباط غيرمعنوي لأن Sig. (2-tailed) = 0.606 وهذه القيمة اكبر من $\alpha = 0.05$

الانحدار Regression

مفهوم الارتباط والتنبؤ - الانحدار

الكتاب صفحة 273

يعتبر تحليل الانحدار أكثر طرق التحليل الإحصائي استخداما. الانحدار بصفه عامه يهدف الى الحصول على أحسن معادله تمثل العلاقة بين المتغير التابع (dependent) والمتغير المستقل (independent). الغرض من استخدام أسلوب تحليل الانحدار الخطي البسيط، هو دراسة وتحليل أثر متغير (مستقل) على متغير آخر(تابع) من خلال معادلة الانحدار. كما تستخدم معادلة الانحدار فى التنبؤ بقيمة المتغير التابع عند قيمة معينه للمتغير (المتغيرات) المستقل. مفهوم التنبؤ في ابسط معنى له هو تقدير قيمة متغير واحد بناء على قيمة متغير اخر

يسمى المتغير Y المتغير التابع (dependent variable)

يسمى المتغير X المتغير المستقل (independent variable) – يمكن ان يكون المتغير المستقل عبارة عن متغير واحد فقط فيسمى الانحدار بالانحدار البسيط ويمكن أن يكون اكثر من متغير مستقل فيكون الانحدار متعدد وسوف نقوم بشرحه لاحقا.

بصفه عامه يمكن ان يكون الانحدار خطيا Linear وقد يكون غير خطيا Non-Linear فاذا كانت لدينا متغيرين وكانت العلاقة بينما خطيه فيمكن تمثيل العلاقة بينهما بمعادلة خط مستقيم.

شروط تطبيق الانحدار

- ومن شروط تطبيق الانحدار أن يكون المتغير المستقل متغير كمي أو تكون متغير ترتيبي, فاذا كانت المتغيرات المستقلة ترتيبية فيجب التعبير عنها بأكواد ولا نتعامل مع بياناتها بدون ترميز.
- المتغير التابع يجب أن يكون متغير كمي (ذو مقياس كمي)
- ان تتبع البيانات التوزيع الطبيعي

امثلة على استخدام تحليل الانحدارالبسيط:

- الإنفاق، والدخل العائلي.
- سعر السلعة، والكمية المطلوبة منها.
- كمية السماد والمحصول الزراعي

Regression function - معادلة انحدار y على x

الكتاب صفحة 275

تسمى العلاقة الرياضية التي تصف علاقة المتغيرات محل الدراسة معادلة انحدار Y على X والتي من خلالها تتحدد قيمة المتغير Y تبعا لقيمة المتغير X لذلك يمكن التعبير عن تلك العلاقة الخطية بالمعادلة التالية:

$$\hat{Y} = a + b\,x$$

حيث

$\hat{Y}$ قيمة المتغير التابع المتنبأ بها

x قيمة المتغير المستقل التي تستخدم للتنبؤ بقيمة $\hat{Y}$

a ثابت الانحدار (الجزء المقطوع من محور Y

b معامل الانحدار (ميل الخط المستقيم)

الكتاب صفحة 279

كيفية حساب معامل الانحدار وثابت الانحدار

يتم حساب معامل الانحدار وثابت الانحدار من خلال القوانين التالية:

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

حيث $\bar{X}$ هومتوسط المتغير المستقل و $\bar{Y}$ هو متوسط المتغير التابع

b يمكن حسابها من القانون التالي

$$b = \frac{n\sum XY - \sum X \sum Y}{n\sum X^2 - (\sum X)^2}$$

حيث n عدد المشاهدات (حجم العينة)

$\sum XY$ مجموع حاصل ضرب Y في X

$\sum X$ مجموع X

$\sum Y$ مجموع Y

$\sum X^2$ مجموع مربعات X

- ملاحظة سيتم حساب قيم a , b من خلال استخدام برنامج spss

الانحدار الخطي باستخدام برنامج ال spss

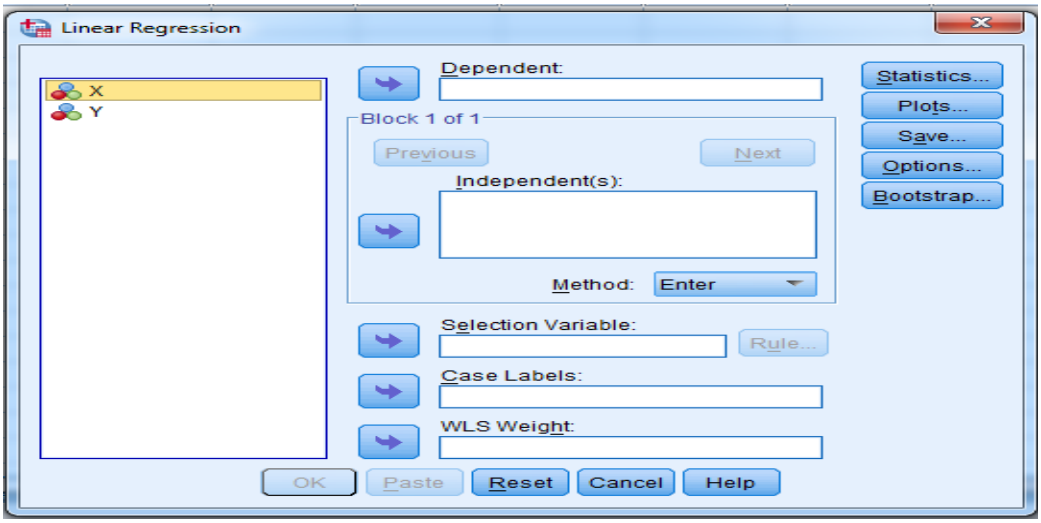
لتطبيق الانحدار الخطي يجب التحقق من الشروط التالية:

1) ان تكون البيانات كمية
2) ان تتبع البيانات التوزيع الطبيعي

مثال على تطبيق الانحدار باستخدام برنامج ال spss الكتاب صفحة 293

نتبع الخطوات التالية:

1- نقوم بادخال البيانات (المتغير التابع {dependent} – المتغير المستقل {independent})
2- نقوم باختبار الطبيعية كما سبق شرحه واذا تم التحقق من الشرط ننتقل الى الخطوة التالية
3- من قائمة analyze نقوم باختيار Regressiom ثم Linear فتظهر الشاشة التالية



4- نقوم بنقل المتغير التابع الى الخانة dependent
5- نقوم بنقل المتغير المستقل الى الخانة independent
6- Ok

تفسير النتائج

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.559[a]	.313	.227	1.540

a. Predictors: (Constant), X

الجدول السابق يوضح ملخص للنموذج (قيمة معامل بيرسون للارتباط (R) ويبين طبيعة الارتباط بين المتغير التابع والمتغير المستقل) وقيمة معامل التحديد وخطأ التقدير(غير مطلوب تفسيره)

ANOVA[a]

Model		Sum of Squares	Df	Mean Square	F	Sig.
1	Regression	8.626	1	8.626	3.637	.093[b]
	Residual	18.974	8	2.372		
	Total	27.600	9			

a. Dependent Variable: Y

b. Predictors: (Constant), X

الجدول السابق يوضح معنوية نموذج الانحدار اي أنه يختبر هل يوجد علاقة خطية بين المتغير التابع والمتغير المستقل من خلال مقارنة ال .Sig مع الفا (α) ويتم تفسيره كالتالي

a. اذا كانت قيمة ال .sig اكبر من الرقم (α الفا = 0.05 – مستوى المعنوية) فإن النموذج غير معنوي أي انه لايوجد علاقة خطية بين المتغيرين
b. اذا كانت قيمة ال .sig اصغر من الرقم (α الفا = 0.05 – مستوى المعنوية) فإن النموذج معنوي أي انه يوجد علاقة خطية بين المتغيرين

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	T	Sig.
		B	Std. Error	Beta		
1	(Constant)	3.718	1.195		3.110	.014
	X	.372	.195	.559	1.907	.093

a. Dependent Variable: Y

ثابت الانحدار

معامل الانحدار

هذا الجدول يحتوي على قيم كل من معامل الانحدار وثابت الانحدار (Unstandardized Coefficients) حيث قيمة B المقابلة ل constant تمثل ثابت الانحدار (3.718) وتمثل قيمة Bالمقابلة ل X معامل الانحدار 0.372

ومنها نستطيع كتابة نموذج الانحدار كالتالي

$$\hat{Y} = 3.718 + 0.372X$$

تنبئي بقيمة Y عنما x=10

$$\hat{Y} = 3.718 + 0.372(10)$$

$$\hat{Y} = 3.718 + 3.72 = 7.438$$

مثال (2): البيانات التالية تبين الكميات المعروضة من سلعة معينة (y) وسعر السلعة (x)

y الكمية المطلوبة	10	15	5	4	3	13
X السعر	7	10	4	3	2	8

المطلوب:

1- تقدير معادلة انحدار الكمية المعروضة على السعر.

2- تقدير الكمية المعروضة عند السعر x=20

نتبع الخطوات التالية:

1- نقوم بادخال البيانات (المتغير التابع Y -{dependent} – المتغير المستقل X -{independent}
2- نقوم باختبار الطبيعية كالتالي:

الفروض:

البيانات تتبع التوزيع الطبيعي: H0
البيانات لا تتبع التوزيع الطبيعي: H1

One-Sample Kolmogorov-Smirnov Test

		Y	X
N		6	6
Normal Parameters[a,b]	Mean	8.33	5.67
	Std. Deviation	5.046	3.141
Most Extreme Differences	Absolute	.246	.202
	Positive	.246	.202
	Negative	-.156-	-.164-
Kolmogorov-Smirnov Z		.601	.495
Asymp. Sig. (2-tailed)		.862	.967

a. Test distribution is Normal.

b. Calculated from data.

تفسير نتائج شرط الطبيعية

بالنظر في الجدول الى الرقم ((Asymp. Sig. (2-tailed) ويتم تفسيره كالتالي:

a. اذا كانت قيمة ال .sig اكبر من الرقم (α الفا – مستوى المعنوية) فإننا لانرفض الفرض H0 أي أن البيانات تتبع توزيع طبيعي
b. اذا كانت قيمة ال .sig اصغر من الرقم (α الفا – مستوى المعنوية) فإن نرفض الفرض H0 أي ان البيانات تتبع التوزيع الطبيعي

من المثال نجد أن ال (Asymp. Sig. (2-tailed) للمتغيرين X,Y اكبر من ال (α الفا حيث α = 0.05) وعليه فإننا لانرفض الفرض H0 أي أن البيانات تتبع توزيع طبيعي

استكمال خطوات الانحدار بعد تحقق الشرط
من قائمة analyze نقوم باختيار Regressiom ثم Linear ثم نقوم بنقل المتغير التابع والمستقل للخانات المناسبة ثم OK

نتائج الانحدار

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.993[a]	.985	.981	.688

a. Predictors: (Constant), X

الجدول السابق يوضح ملخص للنموذج (قيمة معامل بيرسون للارتباط (R) ويبين طبيعة الارتباط بين المتغير التابع والمتغير المستقل) ونجد أنه ارتباط قوي طردي

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	125.441	1	125.441	265.219	.000[b]
	Residual	1.892	4	.473		
	Total	127.333	5			

a. Dependent Variable: Y

b. Predictors: (Constant), X

الجدول السابق يوضح معنوية نموذج الانحدار اي أنه يختبر هل يوجد علاقة خطية بين المتغير التابع والمتغير المستقل من خلال مقارنة ال .Sig مع الفا (α) ويتم تفسيره كالتالي

a. اذا كانت قيمة ال .sig اكبر من الرقم (α الفا = 0.05 – مستوى المعنوية) فإن النموذج غير معنوي أي انه لايوجد علاقة خطية بين المتغيرين
b. اذا كانت قيمة ال .sig اصغر من الرقم (α الفا = 0.05 – مستوى المعنوية) فإن النموذج معنوي أي انه يوجد علاقة خطية معنوية بين المتغيرين

ومن الجدول نجد أن .Sig اصغر من ال (α الفا حيث α = 0.05) وعليه فإن النموذج معنوي أي انه يوجد علاقة خطية معنوية بين المتغيرين

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	-.703-	.622		-1.130-	.322
	X	1.595	.098	.993	16.286	.000

a. Dependent Variable: Y

هذا الجدول يحتوي على قيم كل من معامل الانحدار وثابت الانحدار (Unstandardized Coefficients) حيث قيمة B المقابلة ل constant تمثل ثابت الانحدار (0.703-) وتمثل قيمة Bالمقابلة ل X معامل الانحدار 1.595

ومنها نستطيع كتابة نموذج الانحدار كالتالي

$$\hat{Y} = -0.703 + 1.595X$$

تنبئي بقيمة Y عنما x=20

$$\hat{Y} = -0.703 + 1.595(20)$$

$$\hat{Y} = -0.703 + 1.595(20) = 31.197$$

الانحدار الخطي المتعدد (غير موجود في الكتاب)

- فى الانحدار المتعدد نفترض أن لدينا متغيرا تابع (dependent) واحد نرمز له بالرمز Y ومجموعة من المتغيرات المستقله (independent) عددها m متغيرا نرمز لها بالرمز $X_1, X_2, \ldots, X_m$
- كحالة خاصة فى الانحدار البسيط يكون m =1 أي انه يوجد متغير مستقل وحيد
- الإنحدار الخطي المتعدد هو عبارة عن إيجاد افضل معادلة رياضية تعبر عن العلاقة بين متغير تابع وعدة متغيرات مستقلة وتستعمل لتقدير قيم سابقة ولتنبؤ قيم مستقبلية ، وهو عبارة عن إنحدار للمتغير التابع (**Y**) على العديد من المتغيرات المستقلة $X_1, X_2, \ldots, X_m$ لذا فهو يستخدم في التنبؤ بتغيرات المتغير التابع الناتجة من تغيرات عدة متغيرات مستقلة.
- معادلة الانحدار

$$\hat{Y} = a + b_1 X_1 + b_2 X_2 + \cdots + b_m X_m \quad , \; j = 1, 2, \ldots, m$$

$\hat{Y}$ قيمة المتغير التابع المتنبأ بها

$X_1, X_2, \ldots, X_m$ قيمة المتغيرات المستقلة التي تستخدم للتنبؤ بقيمة $\hat{Y}$

a ثابت الانحدار

b معاملات الانحدار

شروط تطبيق الانحدار الخطي المتعدد:

1- أن تتبع البيانات التوزيع الطبيعي
2- أن يكون المتغير التابع كمي والمتغيرات المستقلة كمية او ترتيبية او مختلطة (كمية وترتيبية)

طرق تطبيق الانحدار الخطي المتعدد

هناك عدة طرق لتطبيق الانحدار الخطي المتعدد, نذكر منها طريقتي Enter و Stepwise طريقة enter هي طريقة لتقدير معادلة الانحدار التي تدرس تأثير كل متغير من المتغيرات المستقلة على المتغير التابع، بينما طريقة Stepwiseتدرس تأثير المتغيرات المستقلة ذات التأثير المعنوي علي المتغير التابع مع استبعاد أثر المتغيرات المستقلة التي لها تأثيرا غير معنوي علي المتغير التابع .وكلا الطريقتين لهما نفس الخطوات في برنامج ال spss مع اختيار أسم الطريقة كما سيتم شرحه في المثال (المطلوب فقط طريقة enter).

مثال (1)

إدرسي العلاقة بين المتغيرات التالية بأسلوبي الانحدار وعلقي علي النتائج، حيث المتغير الأول X1 هو درجات تمثل علاقة الطفل بجيرانه – المتغير الثاني X2 درجات لعلاقته بالاصدقاء – المتغير الثالث X3 درجات لعلاقته بالأهل .المتغير الرابع Y درجات لعلاقته بالمجتمع بشكل عام ثم تنبئي بدرجة علاقة شخص بمجتمعه إذا كانت درجات علاقته بجيرانه وأصدقائه وأهله علي التوالي هي 25 -19 10

X1	X2	X3	Y
8	6	20	16
11	12	28	18
19	12	22	20
22	9	14	11
11	16	6	9
7	14	18	20
24	15	10	4
15	16	14	8
12	20	23	24
6	5	15	23

لتطبيق الانحدار الخطي يجب التحقق من الشروط التالية:

1) ان تكون البيانات كمية او ترتيبية
2) ان تتبع البيانات التوزيع الطبيعي

خطوات الحل:

1- نقوم بادخال البيانات (المتغير التابع {dependent} – المتغيرات المستقلة {independent})
2- نقوم باختبار الطبيعية كالتالي

الفروض:

البيانات تتبع التوزيع الطبيعي: H0
البيانات لا تتبع التوزيع الطبيعي: H1
نتبع الخطوات التالية:

Analyze → Nonparametric - tests →Legacy dialog → 1- sample k.s test

One-Sample Kolmogorov-Smirnov Test

		X1	X2	X3	Y
N		10	10	10	10
Normal Parameters[a,b]	Mean	13.50	12.50	17.00	15.30
	Std. Deviation	6.311	4.720	6.532	6.881
Most Extreme Differences	Absolute	.194	.158	.123	.153
	Positive	.194	.129	.120	.134
	Negative	-.117-	-.158-	-.123-	-.153-
Kolmogorov-Smirnov Z		.613	.499	.389	.483
Asymp. Sig. (2-tailed)		.846	.965	.998	.974

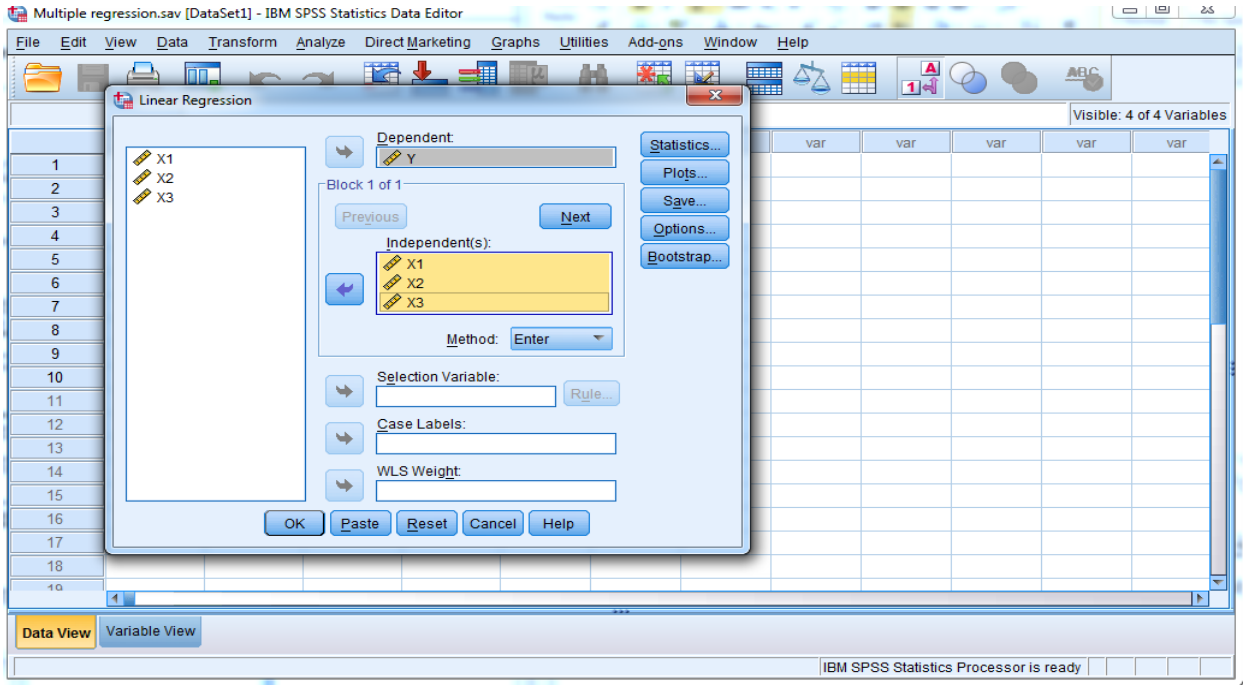
a. Test distribution is Normal.

b. Calculated from data.

تفسير النتائج نجد أن α < 974 ,998 ,.965 ,.846 = sig (لجميع المتغيرات) إذن لانرفض الفرض العدمي أي المتغيرات الاربعة تتبع التوزيع الطبيعي وبذلك تحقق شرط الانحدار

خطوات الانحداربطريقة (Enter)

7- من قائمة analyze نقوم باختيار Regressiom ثم Linear فتظهر الشاشة التالية



8- نقوم بنقل المتغير التابع للخانة (dependent) ونقل المتغيرات المستقلة للخانة (Independent) ونقوم باختيار الطريقة enter من الخانة method ثم ok تظهر النتائج التالية

Variables Entered/Removed[a]

Model	Variables Entered	Variables Removed	Method
1	X3, X2, X1[b]	.	Enter

a. Dependent Variable: Y

b. All requested variables entered.

- من الجدول السابق نلاحظ بأن طريقة الإنحدار المستخدمة وهي طريقة **Enter** حيث يتبين ان البرنامج قام بادخال جميع المتغيرات المستقلة في معادلة الإنحدار الخطي المتعدد .

-

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.824[a]	.679	.519	4.774

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

- من الجدول السابق نلاحظ بأن قيمة معامل الارتباط البسيط R قد بلغ (0.824) ونلاحظ أنه ارتباط قوي طردي

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	289.363	3	96.454	4.232	.063[b]
	Residual	136.737	6	22.789		
	Total	426.100	9			

a. Dependent Variable: Y

b. Predictors: (Constant), X3, X2, X1

- الجدول السابق يتضمن قيم تحليل التباين (ANOVA) والذي يمكن المعرفة من خلاله على معنوية نموذج الانحدارالخطي المتعدد. أي أنه يوضح هل يوجد علاقة خطية ذات دلالة احصائية بين المتغير التابع و المتغيرات المستقلة ام لا ونلاحظ قيمة ال $0.05 = \alpha < Sig=0.063$ أي ان نموذج الانحدار الخطي المتعدد غير معنوي (على الاقل لايوجد علاقة خطية بين المتغير التابع وبين احد المتغيرات الثلاث المستقلة)

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	12.289	7.327		1.677	.145
	X1	-.493-	.267	-.453-	-1.847-	.114
	X2	-.051-	.347	-.035-	-.148-	.887
	X3	.607	.252	.576	2.410	.053

a. Dependent Variable: Y

ثابت الانحدار

معامل انحدار X2على Y

معامل انحدار X1 على Y

معامل انحدار X3 على Y

الجدول السابق يوضح قيمة الثابت ومعاملات الإنحدار ومعنوية تأثير المتغيرات المستقلة على المتغير التابع فنجد أن:

- تأثير المتغيرالمستقل X1 على المتغير التابع Y هو تأثير غير معنوي وذلك لأن ال $\alpha < sig = 0.114$.
- تأثير المتغيرالمستقل X2 على المتغير التابع Y هو تأثير غير معنوي وذلك لأن ال $\alpha < sig = 0.887$.
- تأثير المتغيرالمستقل X3 على المتغير التابع Y هو تأثير معنوي وذلك لأن قيمة ال $sig = 0.053$ تقريبا مساوية ل α .

ويمكن تلخيص هذه الجدول في معادلة الانحدار على الشكل التالي:

$$\hat{Y} = 12.389 - 0.493\,X_1 - 0.051X_2 + 0.607\,X_3$$

تقدير علاقة الطفل بمجتمعه اذا كانت علاقته بجيرانه وأصدقائه وأهله على التوالى هى 25 -19 -10

علاقة الطفل بجيرانه X1

علاقة الطفل بأصدقائه X2

علاقة الطفل بأهله X3

$$\hat{Y} = 12.289 - 0.493(10) - 0.051(19) + 0.607\,(25)$$

$$\hat{Y} = 12.289 - 4.93 - 0.969 + 15.175 \quad = 21.565$$

مقدمة عن اختبارات الفروض

الكتاب صفحة 388

الاستدلال الاحصائي باختبارات الفروض – فرض العدم والفرض البديل

الكتاب صفحة 389

يتعرض الإنسان في كثير من الحالات وفي مجالات العمل الى مواقف معينة تتطلب منه اتخاذ قراريخص معالم المجتمع بناء على معلومات محسوبة من العينه ، وعليه يجب اتخاذ هذا القرار بأقل قدر ممكن من الخطأ. فمثلا نريد أن نقرر مدى فعالية دواء جديد في شفاء مرض ما بناء على معلومات من عينة احصائية.

مثال آخر.. نفرض أن باحثا اجتماعيا ادعى ان متوسط اعمار طلاب الجامعة لايختلف عن متوسط أعمار الطالبات. للتأكد من ذلك فإن الشئ الطبيعي أن نقوم بحصر اعمار الطلاب والطالبات ومنها نحسب المتوسط لكل منهما ثم نقرر من منهما اكبر..ولكن عملية الحصر صعبة ومجهدة لذلك نضطر الى اختيار عينة عشوائية من بين الطلاب وعينة عشوائية من بين الطالبات ونحسب متوسط العمر في كل عينة منهما,,,,فإذا كان متوسط عمر الطالب هو 24 وكان متوسط عمر الطالبة هو 22 فهل يعني ذلك ان متوسط عمر الطالب اكبر من متوسط عمر الطالبة ؟؟ هل الفرق راجع لمجرد الصدفة ؟؟ متى يكون الفرق نتيجة للصدفة؟؟ متى يكون الفرق دالا على وجود اختلاف حقيقي أو جوهري بين متوسطي المجتمعين الأصليين.

مفاهيم مهمة:

هناك بعض المفاهيم المتعلقة باختبارات الفروض لابد من معرفتها:

الفرض الاحصائى Statistical hypothesis

هو عبارة عن إدعاء او تخمين معين حول معلمة من معالم المجتمع ويكون المطلوب اختبار صحة هذا الادعاء أو التخمين.

انواع الفروض الاحصائية

- **فرض العدم (null hypothesis)**
  يرمز له بالرمز H_o ويصاغ في صورة عدم وجود فرق أو عدم وجود علاقة أو عدم وجود تغير

  مثال : في مثال أعمار الطلاب وطالبات الجامعة فإن فرض العدم هو

  H_o: لايوجد اختلاف بين متوسطي اعمار الطلاب والطالبات

- **الفرض البديل (alternative hypothesis)**
  يرمز له بالرمز H_1 وهو الفرض الذي يجب أن يكون صحيحا اذا كان فرض العدم غير صحيح

  مثال : في مثال أعمار الطلاب وطالبات الجامعة فإن الفرض البديل هو

  H_1: يوجد اختلاف حقيقي وليس ظاهري بين متوسط اعمار الطلاب والطالبات.

- يمكن صياغة الفروض الاحصائية حول احد احد معالم المجتمع كأن تتم صياغة الفروض حول المتوسط (μ)

مثال1 : اذا كانت μ_1 ترمز لأطوال الطالبات و μ_2 ترمز لأطوال الطلاب فأن الفروض الاحصائية حول المتوسط هي:

$H_0: \mu_1 = \mu_2$

ترمز لعدم وجود فرق بين متوسطي اطوال الطالبات والطلاب

$H_1: \mu_1 \neq \mu_2$

ترمز لوجود فرق بين متوسطي اطوال الطالبات والطلاب

مثال 2 : اذا كانت μ تمثل درجات الطلاب في جامعة الملك عبدالعزيز فاختبري الفرض القائل بأن متوسط درجات الطلاب تساوي 85, فإن الفروض هي

$H_0: \mu = 85$

ترمز لعدم وجود فرق بين متوسط درجات الطلاب والقيمة 85

$H_1: \mu \neq 85$

ترمز لوجود فرق بين متوسط درجات الطلاب والقيمة 85

- **مستوى المعنوية α ودرجة الثقة (1 – α):**

- إن القرار الذي سوف نتخذه بناء على الاختبار الإحصائي لا يمكن اعتباره صحيح 100 % فهناك مقدار من الخطأ لأن المعلومات التي نتخذ قرارنا بناء عليها بيانات مأخوذة من عينة وليس من المجتمع الأصلي
- في اختبار فرض معين، فإن مقدار ثقتنا في القرار المتخذ بالرفض أو القبول يسمى بدرجة الثقة ويرمز له بالرمز (1 – α) كما وأن مقدار عدم الثقة أو مقدار الخطأ يسمى بمستوى المعنوية ويرمز له بالرمز α
- عادة يحدد الباحث مستوى المعنوية أو درجة الثقة قبل البدء في عملية الاختبار.
- عند اختبار فرض العدم H_o ضد الفرض البديل H_1 نجد أننا امام احدى الحالات الاربع الاتية :

الحالات الاربع

	H_o صحيح	H_o خطأ
قبول H_o	قرار سليم	خطأ من النوع الثاني β
رفض H_o	خطأ من النوع الاولα	قرار سليم

1) أن يكون فرض العدم صحيحا ويكون القرار بقبوله (قرار سليم)

2) أن يكون فرض العدم صحيحا ويكون القرار برفضه (قرار خاطئ)

3) أن يكون فرض العدم خطأ ويكون القرار برفضه (قرار سليم)

4) أن يكون فرض العدم خطأ ويكون القرار بقبوله (قرار خاطئ)

الخطأ من النوع الأول : هو احتمال رفض H_o عندما يكون H_o صحيحا ويرمز لحجم هذا الخطأ بالرمز α ويسمى مستوى المعنوية

الخطأ من النوع الثاني :هو احتمال قبول H_o عندما يكون H_oخاطئ ويرمز لحجم هذا الخطأ بالرمزβ

إن الاختبار الاحصائي الجيد هو الذي يجعل قيمة كل من α و β اصغر مايمكن

اجراء الاختبار الاحصائي:

لاجراء أي اختبار الاحصائي فإننا نتبع الخطوات التالية:

- صياغة الفرض الاحصائي
- تحديد قيمة احصاءة الاختبار
- تحديد منطقة القبول والرفض
- القيمة الجدولية
- اتخاذ القرار

بينما باستخدام برنامج الSPSS فإننا نتبع الخطوات التالية

- صياغة الفرض الاحصائي المناسب للاختبار
- ادخال البيانات وتطبيق الشروط و خطوات الاختبار
- استخراج القيمة (Sig) ومقارنتها بقيمة مستوى المعنوية
- اتخاذ القرار

يوجد العديد من الاختبارات المعلمية التي تستخدم لاتخاذ قرار بشأن فروض حول متوسط مجتمع أو متوسطي مجتمعين أو أكثروتلك الاختبارات هي:

1. اختبار T لعينة واحدة
2. اختبار T للفرق بين متوسطي عينتين مرتبطتين
3. اختبار T للفرق بين متوسطي عينتين مستقلتين
4. اختبار تحليل التباين ANOVA لمقارنة متوسطات عدة مجتمعات مستقلة

وتلك الاختبارات تتطلب شروط قد يصعب تحقيقها في الواقع وهي:

1- أن تكون البيانات لكل متغير كمية أو ترتيبية متعددة المستويات 4 (مستويات فأكثر.)
2- أن تكون البيانات لكل متغير تتبع التوزيع الطبيعي ويمكن التغاضي عن ذلك الشرط إذا بلغ حجم العينة لكل متغير 30 فأكثر.

أولا : اختبار فرض حول المتوسط μ (اختبار ت للعينة الواحدة)

يدرس ما إذا كان متوسط المجتمع المسحوبة منه العينة مساويا لقيمة معينة أم لا

1) صياغة الفروض الاحصائية كالتالي :

$H_0: \mu = \mu_0$

والذي يفترض عدم وجود اي فرق معنوي بين متوسط العينة والقيمة المفترضة μ_0

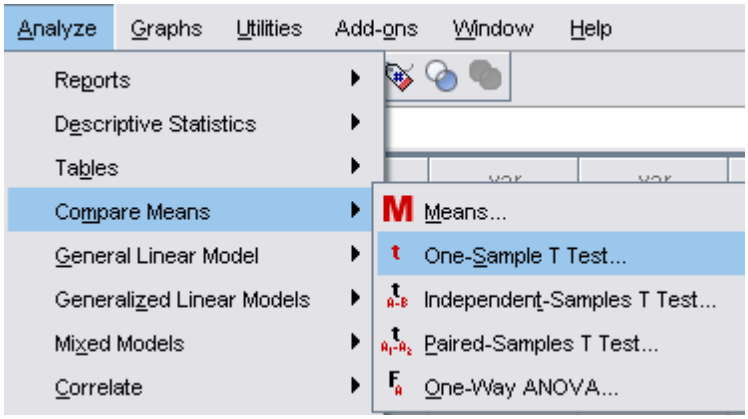
$H_1: \mu \neq \mu_0$ أو $H_1: \mu \geq \mu_0$ أو $H_1: \mu \leq \mu_0$

والذي يفترض وجود فرق معنوي بين متوسط العينة والقيمة المفترضة μ_0

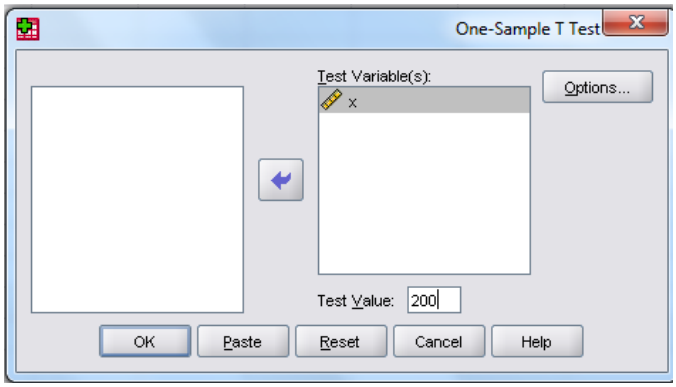
و سنكتفي بالحالة $H_1: \mu \neq \mu_0$

2) اختبار الشروط وخطوات الاختبار كالتالي:

Analyze → compare means → one sample T test



تظهر النافذة التالية:



نقوم بنقل المتغير لنافذة (test variables) ثم نكتب قيمة الادعاء في الخانة test value ثم ok

القرار هو :

c. اذا كانت قيمة ال .sig اكبر من الرقم (α الفا – مستوى المعنوية) فإننا لانرفض الفرض H0 أي أنه لايوجد فرق معنوي بين متوسط المجتمع والقيمة المفترضة

d. اذا كانت قيمة ال .sig اصغر من الرقم (α الفا – مستوى المعنوية) فإن نرفض الفرض H0 أي ان يوجد فرق معنوي بين متوسط المجتمع والقيمة المفترضة

مثال : الكتاب ص 420

ارادت احدى الجهات الصحية باحدى المدن التأكد من كون متوسط كمية البكتريا لكل وحدة حجم من مياه البحيرة داخل حدود الامان (200). فحص باحث 10 عينات من مياه البحيرة فوجد بها كمية البكتريا التالية:

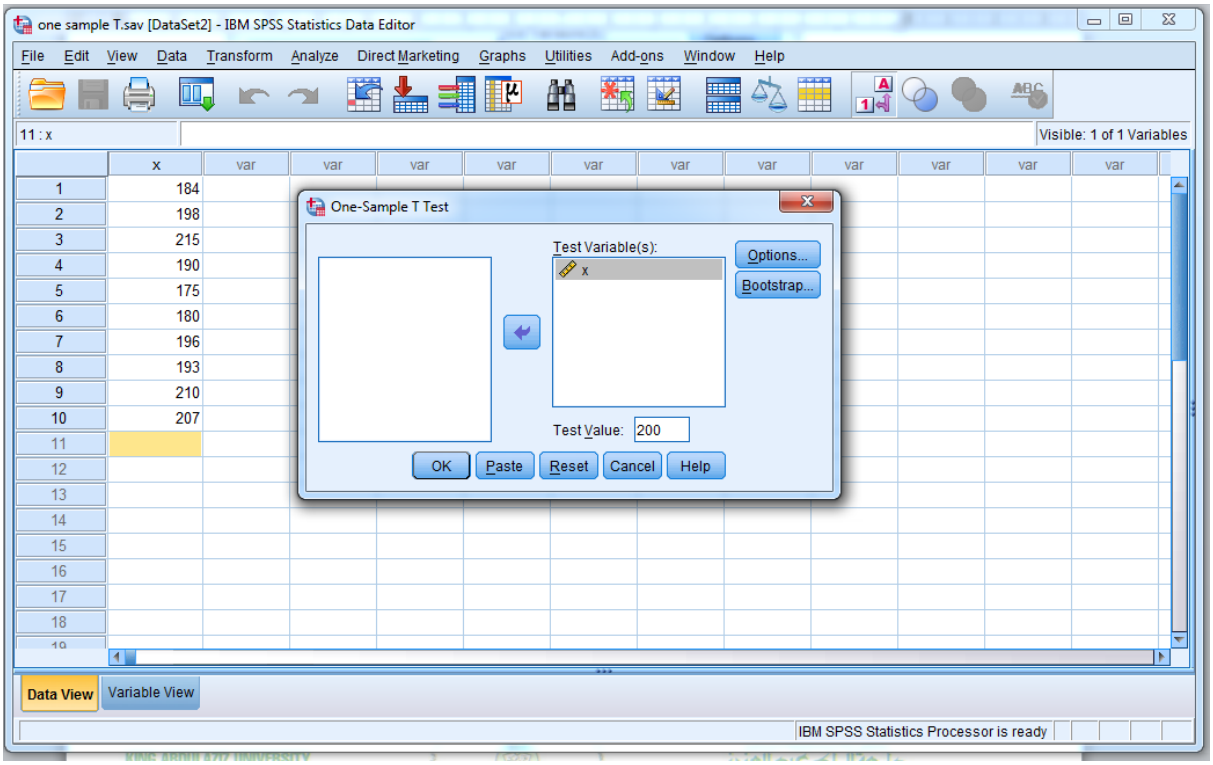
184- 198-215-190-175-180- 196-193-210-207

فهل تدعم هذه البيانات وجود امان لكمية البكتريا داخل البحيرة عند مستوى معنوية 0.05؟

الحل

1) ادخال البيانات : بحيث يتم تعريف متغير كمي ذو مقياس (scale)
2) نتحقق من الشروط
   a. متغير كمي
   b. البيانات تتبع التوزيع الطبيعي (كما تم شرحه سابقا)
3) نتبع خطوات اجراء الاختبار كالتالي

Analyze → compare means → one sample T test



النتائج:

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
X	10	194.80	13.139	4.155

One-Sample Test

	Test Value = 200					
	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference	
					Lower	Upper
X	-1.252-	9	.242	-5.200-	-14.60-	4.20

ستظهر نتائج اختبار T للعينة الواحدة حيث تحتوي جدولين. الجدول الاول يحتوي احصائيات العينة (عدد الحالات, المتوسط , الانحراف المعياري, الخطأ المعياري) بينما الجدول الثاني يحتوي نتائج اختبار ت للعينة الواحدة. في جدول اختبار ت للعينة الواحدة يتم اظهار قيمة احصاءة الاختبار(t) ودرجات الحرية للاختبار (df) وقيمة دلالة الاختبار (.Sig) و الفرق في المتوسطات (Mean difference و هو الفرق بين متوسط العينة والقيمة المفترضة) وفترة الثقة.

القرار هو :

بما أن قيمة ال sig.= 0>242 اكبر من الرقم (α الفا – مستوى المعنوية) فإننا لانرفض الفرض H0 أي أنه لايوجد فرق معنوي بين متوسط المجتمع والقيمة المفترضة... أي ان العينات تدعم وجود امان في مياه البحيرة