مقياس : برمجيات تحليل البيانات

المستوى : السنة الثانية ماستر

التخصص : علم اجتماع الجريمة والانحراف

المحاضرة الثامنة (تابع للمحاضرة السابعة)

عنوان المحاضرة : تقدير التوزيع الطبيعي للبيانات عن طريق الاختبارات الإحصائية

أهداف المحاضرة

- التعرف على أنواع التوزيع وتأثيرها في اختيار الاختبارات الإحصائية المناسبة .
- التعرف على مفهوم التوزيع الطبيعي للبيانات .
- معرفة طرق تقدير توزيع البيانات الكمية والكيفية في برنامج SPSS .

تمهيد

يعتبر التوزيع الطبيعي احد اهم المفاهيم الأساسية في العلوم الإحصائية وتحليل البيانات الاجتماعية عموما ، نظرا لكونه النموذج النظري الذي تقترب منه الكثير من الظواهر الإنسانية والاجتماعية ، كما تكمن أهميته في اعتباره الأساس النظري للكثير من الاختبارات الإحصائية اذ تعتمد عليه طرق تحليل البيانات الاستدلالية واختبارات التباين والانحدار الخطي والارتباط وغيرها ، وبالتالي فإن فهم خصائص هذا التوزيع وتحديد طبيعته يعتبر خطوة ضرورية تمكن الباحث من اختيار الأساليب الإحصائية المناسبة لتحقيق نتائج أكثر دقة ومصداقية.

تقدير التوزيع الطبيعي للبيانات عن طريق الاختبارات الإحصائية

1- في حالة العينات الصغيرة (اقل من 30)

في حالة العينات التي يكون فيها العدد صغيرا ، فإن الاختبار الاحصائي المناسب لتقدير التوزيع يتمثل في اختبار Shapiro Wilk ، حيث يتناسب هذا الاختبار مع العينات صغيرة الحجم.

2- العلاقة بين قيمة الاختبار والتوزيع الطبيعي

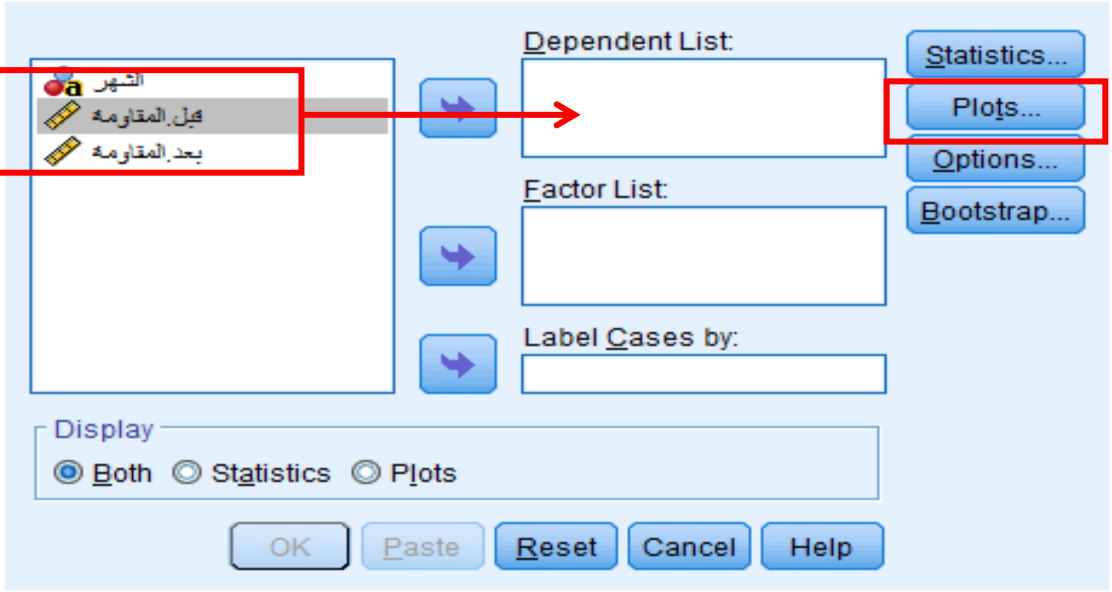
في برنامج SPSS عند تقدير طبيعة التوزيع في حالة عينة صغيرة الحجم نتحصل على جدول من صفحة Output يحمل عنوان **Test Of Normality** ، حيث يتضمن الجدول قيم الاختبار ، درجة الحرية ومعنوية الاختبار ، حيث تمثل المعنوية القيمة التي يتحدد من خلالها اعتدالية التوزيع من عدمها.

3- طريقة حساب اختبار Shapiro-Wilk

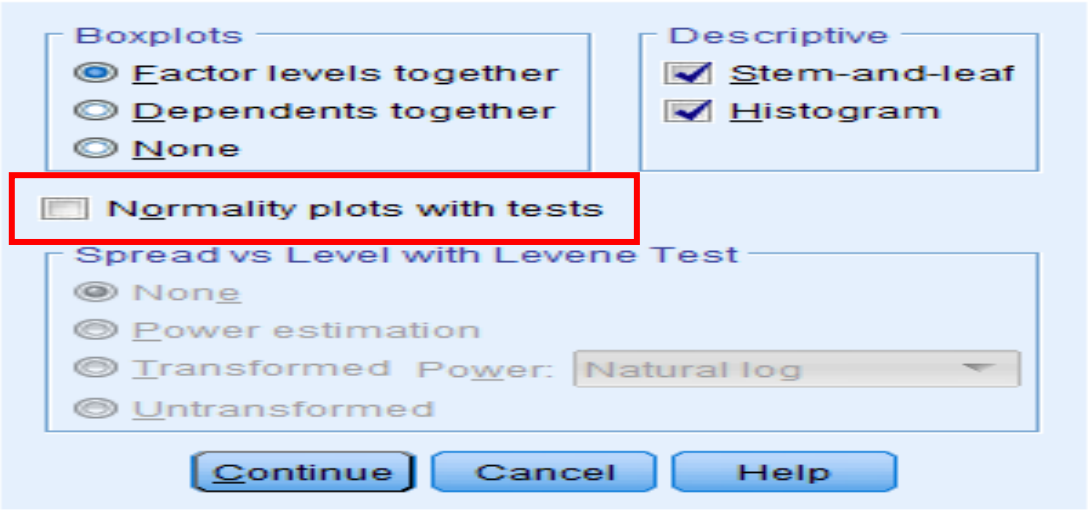
يكون ذلك على طريق الخطوات التالية :

Analyze >>> Descriptive Statistics >>> Explore

بعد تأكيد الاختيار نتحصل على النافذة الموضحة في الشكل التالي :



بعد ذلك نقوم بنقل المتغيرات الى مربع **Dependent List** ، بعد ذلك من مربع **Plots** لنتحصل على النافذة الموضحة في الشكل التالي :



بعد ذلك نقوم بتفعيل امر **Normality plots with tests** والضغط على **Continue** لنعود الى النافذة الأساسية ثم نختار تأكيد **.OK**

من صفحة **Output** نتحصل على استكشاف كامل للبيانات التي نقوم بدراستها مثل الجداول التكرارية والتمثيلات البيانية وغير ذلك ، غير ان المهم في هذه الصفحة هو الجدول الذي يكون بعنوان **Test of Normality** والموضح في الشكل التالي : (ملاحظة تم حساب طبيعة التوزيع بناءا على مثال المحاضرة السابعة)

Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
قبل.المقاومة	.168	8	.200*	.969	8	.892
بعد.المقاومة	.101	8	.200*	.991	8	.997

نلاحظ ان البرنامج يقوم بحساب كل من اختبار Shapiro Wilk واختبار اخر هو اختبار Kolmogrov Smirnov والذي يعتمد في حالة البيانات الكبيرة (اكثر من 30) ، كما نلاحظ ان البرنامج يقوم بحساب التوزيع لكلا المتغيرين (قبل المقاومة ، بعد المقاومة) ، في هذه الحالة كل ما نركز عليه هو قيمة المعنوية Sig والتي تحدد هل الاختبار دال احصائيا ام لا وذلك عن طريق المقارنة التالية :

التوزيع طبيعي = (قيمة محتملة) Sig > 0.05

التوزيع غير طبيعي = (قيمة محتملة) Sig < 0.05

من المثال السابق :

المتغير يتبع توزيعا طبيعيا = Sig = 0.892 > 0.05

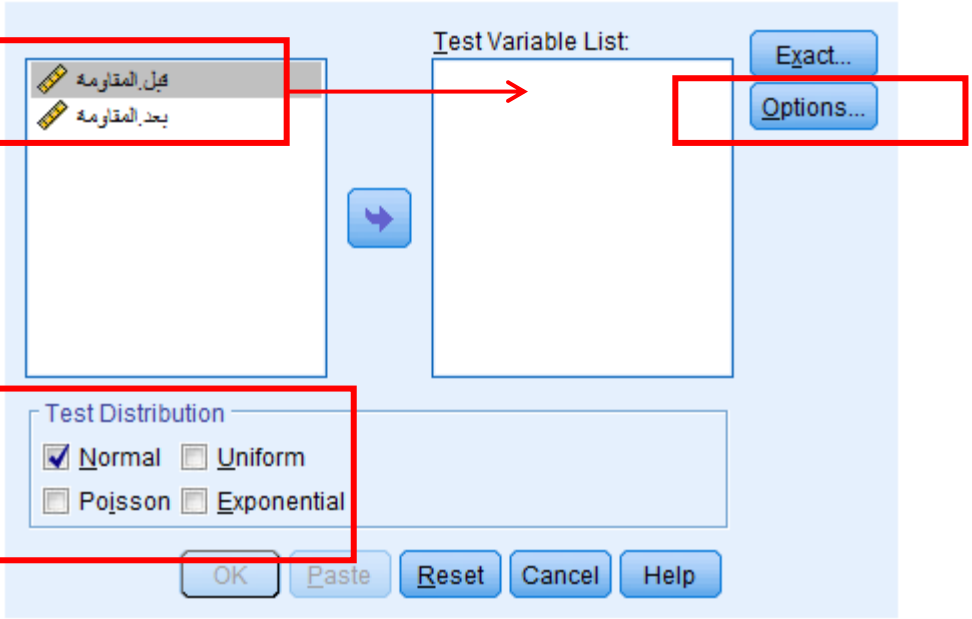
المتغير الثاني يتبع توزيعا طبيعيا = Sig = 0.997 > 0.05

ثانيا : في حالة البيانات الكبيرة (حجم العينة اكبر من 30)

في هذه الحالة يعتبر اختبار K-S هو الاختبار المناسب لتقدير التوزيع ، حيث يتم حساب الاختبار من البرنامج عن طريق الخطوات التالية :

Analyze >>> Nonparametric Test >>> Legacy Dialogs >>> 1-Sample K-S

بعد تأكيد الاختيار نتحصل على نافذة الأوامر الموضحة في الشكل التالي :



بعد ذلك نقوم بالخطوات التالية :

نقل المتغيرات الى مربع **Test Variable list**

تفعيل اختيار **Normal** من قائمة **Test Distribution** ثم التأكيد **OK**

يتوفر **Options** على أوامر واعدادات أخرى مثل ادراج القيم المفقودة في الاختبار إضافة الى ادراج الإحصاءات في صفحة **Output**.

بعد ذلك نتحصل على الجدول التالي من صفحة **Output** الذي يكون بعنوان **One Sample K-S test**.

One-Sample Kolmogorov-Smirnov Test

		بعد.المقاومة	قبل.المقاومة
N		8	8
Normal Parameters[a,b]	Mean	13.5250	12.9625
	Std. Deviation	.42342	.44381
Most Extreme Differences	Absolute	.101	.168
	Positive	.092	.168
	Negative	-.101	-.122
Test Statistic		.101	.168
Asymp. Sig. (2-tailed)		.200[c,d]	.200[c,d]

من خلال الجدول ولتحديد طبيعة التوزيع ، نركز على قيمة Asymp ,Sig للاختبار ونقارنها مع القيمة المحتملة 0.05 حيث نجد :

يمكن القول ان التوزيع طبيعي = 0.05 < 0.2 Sig

ملاحظة : قيمة 0.2 Sig في الجدول غير صحيحة لانه تم حسابها بناءا على عينة حجمها 08 وبالتالي فإن اختبار K-S غير مناسب في هذه الحالة ، حيث تم اعتماد المثال لاهداف الشرح فقط .