مقياس : برمجيات تحليل البيانات

المستوى : السنة الثانية ماستر

تخصص : علم اجتماع الجريمة والانحراف

المحاضرة السابعة

عنوان المحاضرة : تقدير اعتدالية التوزيع في برنامج SPSS

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.

أولا : ماهية التوزيع الطبيعي

يعرف التوزيع الطبيعي بأنه توزيع احتمالي مستمر للمتغيرات العشوائية ، يتميز بمنحنى متماثل يأخذ شكل الجرس ، وتتركز قيمه حوول المتوسط الحسابي ، ويعد هذا التوزيع من أكثر التوزيعات استخداما في الإحصاء والعلوم الاجتماعية والطبيعية نظرا لخصائصه الرياضية المهمة وانتشاره الواسع في الظواهر الواقعية، ومن أهم خصائصه نجد :

- التماثل حول المتوسط
- تساوي مقاييس النزعة المركزية
- انحصار القيم في نطاقات معروفة

ثانيا : تقدير طبيعة التوزيع في برنامج SPSS

1- تقدير طبيعة التوزيع في حالة بيانات كمية (مستمرة)

مثال :

يمثل الجدول التالي نتائج دراسة تمثلت في معرفة مقاومة خليط من المعادن قبل وبعد إضافة المعدن الجديد خلال مدة 08 اشهر ، حيث كانت النتائج على النحو التالي :

الشهر	جانفي	فيفري	مارس	افريل	ماي	جوان	جويلية	اوت
المقاومة قبل الاضافة	12.6	12.9	13.4	12.3	13.6	13.2	12.6	13.1
المقاومة بعد الإضافة	12.9	13.7	13.3	13.9	14.2	13.6	13.5	13.1

تقدير التوزيع عن طريق تمثيل Q-Q Plots

يتمثل Q-Q Plots في تمثيل بياني لقيم المتغير ومطابقتها بمثيلاتها في التوزيع الطبيعي ، ويتم ذلك عن طريق الأوامر التالية :

Analyze >>> Descriptive Statistics >>> Q-Q Plots

بعد ذلك تظهر النافذة الموضحة في الشكل التالي :

تتضمن نافذة Q-Q plots على الأوامر التالية :

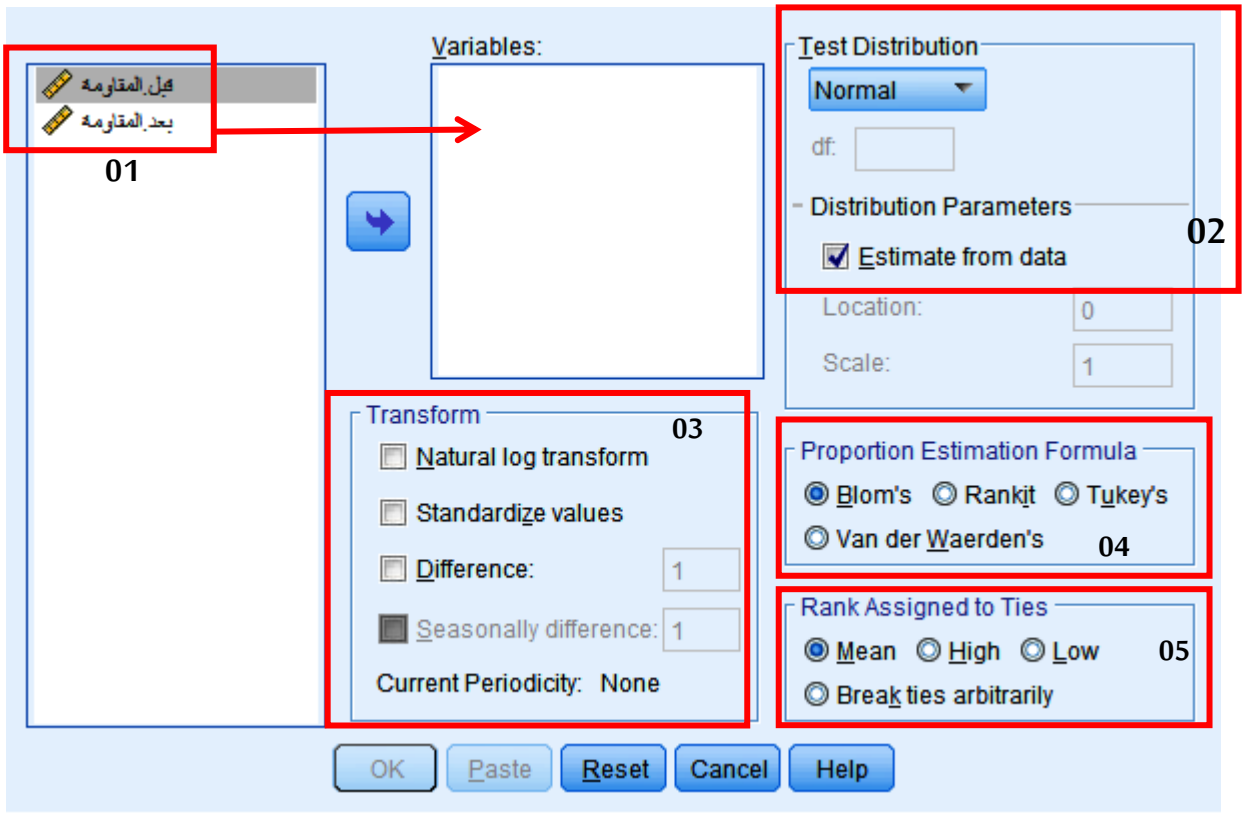
1- المتغيرات : تتمثل في المتغيرات التي نقوم بدراستها في المثال

2- اختبار التوزيع Test Distribution : حيث يمكن من خلاله تعديل تقدير التوزيع الذي ندير معرفته من خلال اختيار تطبيعة التوزيع بتغيير (Normal الى توزيع اخر من القائمة) ، كما نجد أيضا df والتي تتعلق بدرجة حرية الاختبار ، كما نجد اعدادات التوزيع Distribution Parameters ، والتي تكون في حالة تقدير اختبار طبييعي Normal ، Estimate from data بمعنى الاعدادت يتم تقديرها بناءا على بيانات العينة.

3- تحويل Transform : في حالة ما اذا كانت البيانات غير طبيعية فإن أوامر Transform تساعد في تطبيق تحويل رياضي على البيانات لجعل توزيعها اقرب الى الطبيعي ، ثم نقوم بإعادة رسم Q-Q Plots لتقييم ما اذا أصبحت البيانات أقرب الى خط مستقيم.

4- Proportion Estimation Formula : صيغة تقديرية لنسبة حدوث ظاهرة داخل مجتمع كبير بالاعتماد على عينة محددة.

5- Rank Assigned to tied : يتم من خلاله تحديد كيفية التعامل مع القيم المتساوية عند ترتيب البيانات في الاختبارات التي تعتند على الرتب.



الشكل الأول : نافذة Q-Q Plots

بعد ذلك نقوم بنقل المتغيرات الموجودة في المربع (01) والتي تتمثل في (قبم المتغير قبل المقاومة وبعدها) الى مربع المتغيرات Variables (دون تعديل الأوامر الأخرى) ثم نضغط على تأكيد OK.

بعد التأكيد يفتح ملف Output يتضمن :

Model Description: يوضح تفاصيل النموذج الاحصائي المستخدم في انشاء مخطط Q-Q Plots ، حيث يوضع نوع التوزيع المستخدم ، طريقة تقدير المعلمات وتحديد تطبيق التحويلات من عدمها.

Case Processing Summary : يوضح عدد الحالات الصالحة والحالات المستبعدة.

Estimated Distribution Parameters : يوضح هذا الجدول كل من مركز التوزيع الفعلي للبيانات من خلال قيم Location ، إضافة الى مدى انتشار هذه البيانات حول المركز من خلال قيم Scale.

Model Description

Model Name		MOD_1
Series or Sequence	1	قبل المقاومه
	2	بعد المقاومه
Transformation		None
Non-Seasonal Differencing		0
Seasonal Differencing		0
Length of Seasonal Period		No periodicity
Standardization		Not applied
Distribution	Type	Normal
	Location	estimated
	Scale	estimated
Fractional Rank Estimation Method		Blom's
Rank Assigned to Ties		Mean rank of tied values

Applying the model specifications from MOD_1

Case Processing Summary

		قبل المقاومه	بعد المقاومه
Series or Sequence Length		8	8
Number of Missing Values in the Plot	User-Missing	0	0
	System-Missing	0	0

The cases are unweighted.

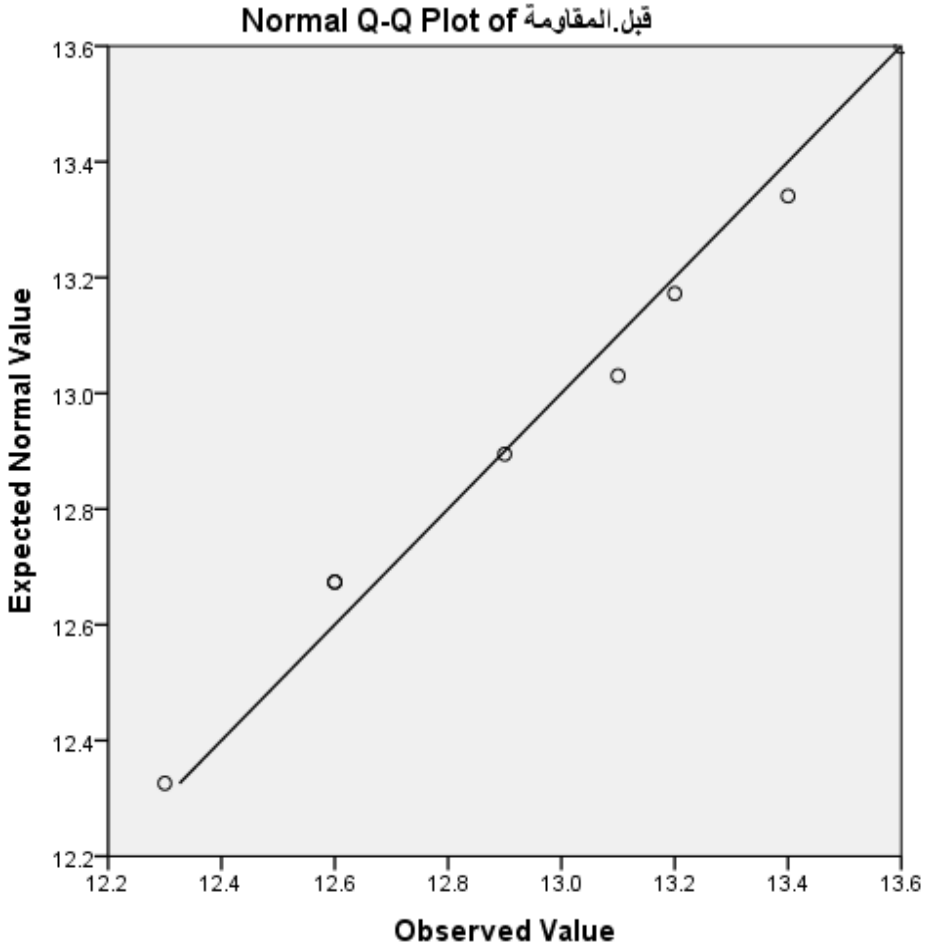
Estimated Distribution Parameters

		قبل المقاومه	بعد المقاومه
Normal Distribution	Location	12.9625	13.5250
	Scale	.44381	.42342

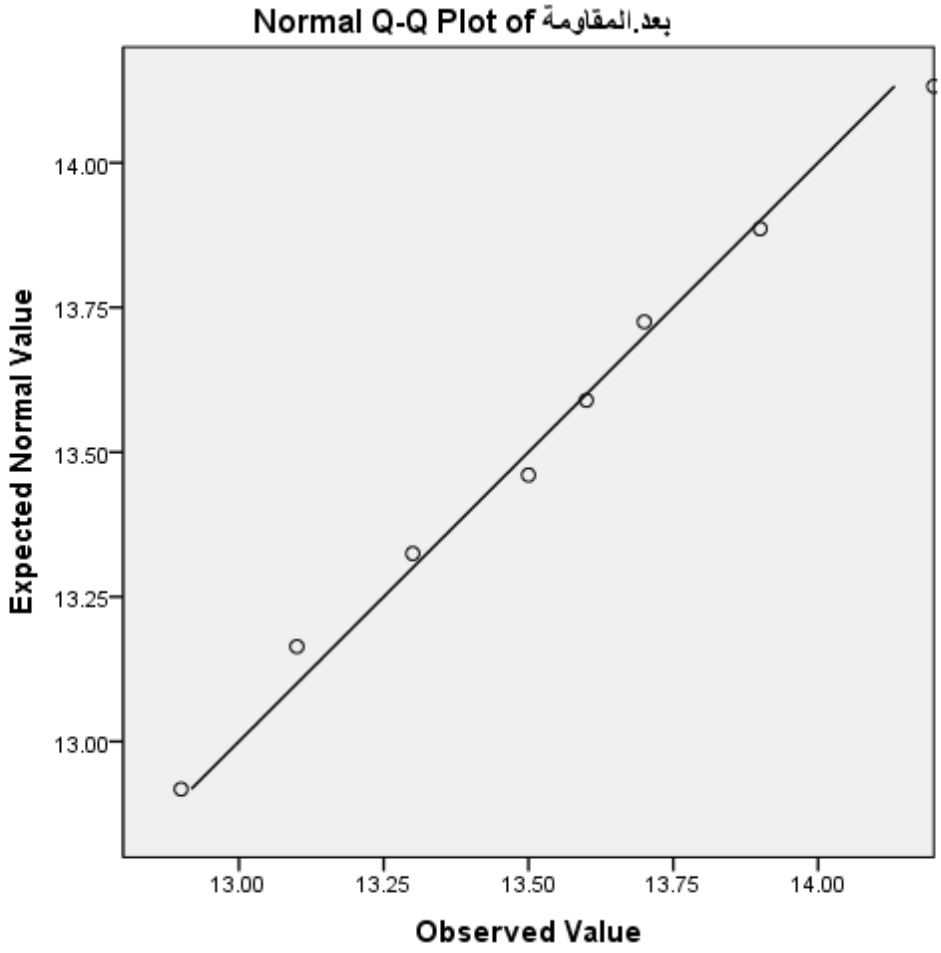
The cases are unweighted.

بعد هذه الجداول من صفحة Output نجد تمثيلات Q-Q Plots والتي سوف نعتمد عليها في تقدير اعتدالية التوزيع ، حيث نجد ان كل متغير من المتغيرات له تمثيله الخاص به :

1- تمثيل Q-Q Plots لمتغير قبل المقاومة



2- تمثيل Q-Q Plots لمتغير بعد المقاومة



لمعرفة اعتدالية التوزيع نقوم بمقارنة كل من الرسم البياني للمتغير الأول مع الرسم البياني للمتغير الثاني ، حيث ان المقارنة تكون من خلال تحديد عدد النقاط التي تكون اقرب للخط المستقيم في التمثيل ، حيث نلاحظ من خلال هذا المثال ان :

كلا المتغيرين يتبعان توزيعا طبيعيا ، غير ان المتغير الثاني (بعد المقاومة) اكثر اعتدالية مت التوزيع الأول (قبل المقاومة) وذلك لأن قيم المتغير الثاني اكثر اقترابا من الخط المستقيم في التمثيل مقارنة بقيم المتغير الأول.

قاعدة المقارنة :

كلما اقتربت نقاط المتغير من الخط المستقيم في التمثيل كلما دل ذلك على اقتراب توزيع المتغير من التوزيع الطبيعي