مجلة كلية التجارة للبحوث العلمية
كلية التجارة
جامعة الإسكندرية

العدد الثاني– المجلد الرابع والخمسون – يوليو ٢٠١٧

أ/ شيماء عبد القادر محمد أحمد سالم
معيدة بقسم الإحصاء والرياضة والتأمين
كلية التجارة – جامعة الأسكندرية

دراسة مقارنة لمجموعة من اختبارات الاعتدال

A comparative Study To Some Tests of Normality

ملخص البحث

تعتمد العديد من الأساليب التقليدية لتحليل البيانات علي إفتراض إعتدال توزيع مجتمع الدراسة مثل: إختبارات *The t-test" t "*،وتحليل التباين *The analysis of variance (ANOVA)*، وإختبارات معاملات الإنحدار *"regression co-efficients"* في تحليل الإنحدار، وكذلك إختبار F لتجانس التباينات *the F-test for homogeneity of variances*.

لذلك فإن إستخدام أحد الأساليب الإحصائية التي تفترض الإعتدال **يتطلب** أولاً التأكد من تحقق هذا الإفتراض قبل إجرائها، والتجاوز عن ذلك قد يؤدي إلي نتائج مضللة تفقد الثقة في النتائج التي تم التوصل إليها، وذلك لأن هذه الأساليب مبنية في الأساس علي تحقق إفتراض الإعتدال لضمان صحة نتائجها. كذلك فإن التحقق من تبعية البيانات للتوزيع المعتدل أو عدم تبعيتها يؤثر في تحديد حجم العينة الضرورى لبناء فترات الثقة وإختبارات الفروض. ومن هنا تظهر ضرورة إجراء إختبار يُمكننا من تحديد ما إذا كان المجتمع المسحوب منه العينة العشوائية محل الدراسة يتبع توزيعاً معتدلاً أم لا.

Abstract

Many classical methods of data analysis depend on the assumption of normality of the population under study, for example: The T-test ,The analysis of variance,Testing regression coefficients in regression analysis and The F-test for homogeneity of variance.

Hence, the normality assumption should be first verified before utilizing methods that assume normality. Over passing this step may lead the researcher to misleading findings, and make results loose credibility.

In addition, checking whether the data follows a normal distribution affects the determination of the sample size necessary for the construction of confidence intervals and hypothesis testing.

هذا البحث جزء من رسالة ماجستير تحت اشراف ا.د. عادل محمود سليمان حلاوة، ا.د. أسامة عبد العزيز حسين، د. محمد ابراهيم أحمد سليمان

١ – مقدمة

يوجد العديد من الأساليب الإحصائية التي تتطلب تحقق إفتراض الإعتدال ، ولذلك فإن التحقق من هذا الإفتراض قد يكون أمرا ضرورياً لبعض الباحثين للمضي قدمًا في بحثهم، ولكن المشكلة التى قد يواجهها أحدهم هى تواجد عدة إختبارات للإعتدال . ولكل إختبار حجم العينة المناسب لتطبيقه، ولذلك يحتاج الباحث إلى التأكد من أن الإختبار الذي سيقوم بإستخدامه هو الأفضل لحالة بحثه من بين البدائل الأخرى لإختبارات الإعتدال ، فعلى سبيل المثال فإن الإختبار الذي قد يعد الأفضل في ظل حجم عينة كبير قد لا يظل كذلك عند إجراء الإختبار بإستخدام عينة أصغر. ويمكن تلخيص مشكلة الدراسة أنه من بين إختبارات الإعتدال ما هو أفضل إختبار في كل حالة من الحالات التالية؟

- المجتمعات المتماثلة ذات قمة واحدة ، وفي ظل أحجام مختلفة من العينات.
- المجتمعات المتماثلة ذات قمتين ، وفي ظل أحجام مختلفة من العينات.
- المجتمعات الملتوية جهة اليمين ذات قمة واحدة، وفي ظل أحجام مختلفة من العينات.
- المجتمعات الملتوية جهة اليمين ذات قمتين، وفي ظل أحجام مختلفة من العينات.
- المجتمعات الملتوية جهة اليسار ذات قمة واحدة، وفي ظل أحجام مختلفة من العينات.
- المجتمعات الملتوية جهة اليسار ذات قمتين، وفي ظل أحجام مختلفة من العينات.

وتهدف الدراسة بصورة أساسية إلى الإجابة عن التساؤلات السابقة من خلال:

- إستعراض أهم الدراسات التي تناولت الإختبارات الخاصة بالإعتدال ومنها إختبارات تعتمد علي الدالة التراكمية التجريبية Empirical Cumu-lativeFunction(ECF)، و إختبارات الأنحدار، وإختبارات تعتمد علي متوسط الأنحراف المطلق عن الوسيط ، مع ذكر خصائص هذه الإختبارات.
- أن تقدم عن طريق دراسة المحاكاة تصنيفاً لإختبارات الإعتدال وفق الحالات المختلفة ، بحيث تُرشح للباحث الإختبار الأفضل في ظل الحالة التي يواجهها.

وتخلص أهمية البحث إلى إعتماد الكثير من الباحثين على بعض الأساليب الإحصائية التي تتطلب تحقق إفتراض الإعتدال ، ولذا فإن إستخدام إختبارات الإعتدال يعد أمرا ضرورياً للتأكد من تحقق هذا الإفتراض قبل إستخدام تلك الأساليب ، ولذلك تنشأ أهمية تقديم تصنيفٍ لهذه الإختبارات يُمَكّن الباحث من إختيار الإختبار الأفضل للحالة التي يواجهها.

٢– اختبارات الاعتدال محل الدراسة:

أولاُ: إختبارات تعتمد علي الدالة التراكمية التجريبية "*Empirical cumulative function*":

ويعتمد الأساس الإحصائي لإحصائيات هذه الإختبارات علي قياس المسافة الرأسية بين الدالة التراكمية التجريبية Empirical cumulative fu-nction (ECF) لعينة ما ومنحني الدالة التراكمية للتوزيع المعتدل المعياري ويُطلق علي هذه العائلة إحصائيات المسافة الرأسية Vertical Distance Statistics.

ويمكن الحصول علي الدالة التراكمية التجريبية $F_n(x)$ لعينة حجمها (n) من خلال ترتيب مشاهداتها $(x_1, x_2, \dots, x_n)$ ترتيباً تصاعدياً أي وضعها في صورة احصاءات مرتبة بحيث $[x_1 < x_2 < x_3 \dots\dots < x_n]$، وتأخذ الدالة التراكمية التجريبية في هذه الحالة الشكل التالي:

$$F_n(x) = \frac{Number\ of\ Observations \le x}{n} \quad -\infty \le x \le \infty \qquad (2-1)$$

أي أن :-

$$F_n(x) = 0 \quad x < x_{(1)}$$

$$F_n(x) = \frac{i}{n}\ x_{(i)} \le x < x_{(i+1)}$$

$$, i = 1,2, \dots\dots, n-1$$

$$F_n(x) = 1 \qquad x \ge x_{(n)}$$

ووفقاً لهذا التعريف يمكن القول بان الدالة التراكمية التجريبية هي دالة سلمية Step Function. ويمكن تقسيم هذه العائلة من احصائيات إختبارات الإعتدال إلي مجموعتين اساسيتين : **المجموعة الأولي** تُسمي مجموعة الاحصائيات العليا Supremum Statistics، **أما المجموعة الثانية** فيُطلق عليها مجموعة الإحصائيات التربيعية Quadratic Statistics. **ويتم عرض هاتين المجموعتين فيما يلي:**

- **الإحصائيات العليا Supremum Statistics**

تعتبر احصائية اختبار Kolomogorov [6] من أهم وأشهر احصائيات هذه الفئة حيث تعتمد علي قياس أكبر مسافة رأسية Maximum Vertical Distance بين الدالة التراكمية التجريبية لمشاهدات العينة $F_n(x)$ وبين الدالة التراكمية $F(x)$ للتوزيع المعتدل المعياري **وذلك من خلال قياس دالتي مسافة يُرمز لهما علي التوالي D^+ ، D^- حيث:**

المسافة D^+ : تعبر عن أكبر مسافة رأسية عندما تزيد قيمة الدالة التراكمية التجريبية للعينة عن قيمة الدالة التراكمية للتوزيع المعتدل المعياري، ويمكن التعبير عن هذه المسافة رياضياً كما يلي:

$$D^+ = sup_x[F_n(x) - F(x)] \qquad (2-2)$$

المسافة D^- : تعبر عن أكبر مسافة رأسية عندما تزيد قيمة الدالة التراكمية للتوزيع المعتدل المعياري عن قيمة الدالة التراكمية التجريبية ، ويمكن التعبير عنها رياضياً كما يلي :

$$\mathrm{D}^- = sup_x[F(x) - F_n(x)] \qquad (2-3)$$

في حين يُرمز لإحصائية Kolomogorov بالرمز D وتُعرف بأنها أقصي فرق مطلق بين منحني الدالة التجريبية والدالة التراكمية للتوزيع المعتدل المعياري وتمثل رياضياً كالتالي:

$$D = sup_x|F_n(x) - F(x)| \qquad (2-4)$$

وبشكل أكثر وضوحاً يمكن القول بأن احصائية Kolomogorov هي القيمة الأكبر من بين المسافتين D^- ، D^+ أي أن :

$$\mathrm{D} = \mathrm{Maximum}\ (D^+, \mathrm{D}^-) \qquad (2-5)$$

وبعد حساب احصائية الإختبار (D) يتم مقارنتها بقيمة حرجة معيارية مُستخرجة من جداول خاصة بهذا الإختبار تعتمد علي حجم العينة (n)،

ومستوي المعنوية (α) وذلك لإتخاذ القرار الإحصائي بالرفض أو القبول .

وقدم Lilliefors [7] تعديلاً علي احصائية Kolomogorov وذلك عندما يكون كلاً من المتوسط والتباين مجهولين واللذان يتم تقديرهما من العينة، وتأخذ احصائية اختبار Lilliefors الشكل التالي :

$$D_{Lill} = \max_{1 \le i \le n} \left[\Phi(x_i, \bar{x}, s^2) - \frac{(i-1)}{n} \ ; \ \frac{i}{n} - \Phi(x_i, \bar{x}, s^2) \right] \qquad (2-6)$$

حيث $\Phi(x_i, \bar{x}, s^2)$ هي الدالة التراكمية للتوزيع المعتدل المعياري ذو المعالم المقدرة من العينة، و $\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}$، و $s^2 = \frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{n-1}$ ويتم رفض فرض العدم للقيم الكبيرة للإحصائية D_{Lill}. كما قدم Lilliefors جداول تحتوي علي أحجام مختلفة للعينات من ٤ إلي ٣٠ مشاهدة ومستويات معنوية مختلفة 0.01 و0.05 و 0.1 و0.15 و0.2 للإحصائية D_{Lill} معتمداً علي المحاكاة، وسيرمز لهذا الأختبار بالرمز Lill .

ويتميز اختبار Kolomogorov بسهولة تطبيقه وانتشار استخدامه في حزم البرامج الإحصائية، إضافة لإمكانية استخدامه بكفاءة في حالة العينات الصغيرة حيث أن فرض العدم لهذا الإختبار يعتمد علي معالم معلومة لكلاً من μ و σ^2، ولكن **يُعاب**عليه أنه يُستخدم لإختبار البيانات المستمرة فقط.

- **الإحصائيات التربيعية Quadratic Statistics**

بينما اعتمدت مجموعة الاحصائيات العليا علي قياس الفرق الرأسي بين الدالة التراكمية التجريبية للعينة والدالة التراكمية للتوزيع المعتدل المعياري ، فإن هذه المجموعة من الإحصائيات التربيعية تعتمد علي قياس مربع الفرق بين هاتين الدالتين أي بين$F_n(x)$ ، و$F(x)$ مُرجحاً باستخدام دالة ترجيح يُرمز لها $\varphi_{(x)}$. وتأخذ الإحصائيات التربيعية شكلاً عاماً علي الصورة :

$$Q = n \int_{-\infty}^{\infty} [F_n(x) - F(x)]^2 \ \varphi_{(x)} \ dF(x) \quad (2-7)$$

وهي مجموعة عامة من الإحصائيات تتضمن احصائية اختبار Anderson–Darling والتي يُرمز لها بالرمز (A^2) كحالة خاصة ويعتمد الشكل الرياضي لهذه الإحصائيات علي الصيغة الرياضية لدالة الترجيح $\varphi_{(x)}$، ويتم عرض هذه الإحصائية فيما يلي:

احصائية اختبار Anderson–Darling[1] :

ويُستخدم هذا الإختبار للعينات التي تزيد أحجامها عن ست مشاهدات ،ويمكن الحصول علي هذه الإحصائية بوضع دالة الترجيح علي الشكل التالي:

$$\varphi_{(x)} = \left[F(x)[1 - F(x)]\right]^{-1} \quad (2-8)$$

ومن ثم تكون إحصائية إختبار Anderson–Darlingعلي الشكل التالي:

$$A^2 = n \int_{-\infty}^{\infty} [F_n(x) - F(x)]^2 \left[F(x)[1 - F(x)]\right]^{-1} dF(x) \qquad (2-9)$$

ويُفضل استخدام الإحصائية المُعدلة A^* لإختبار الإعتدال Stephens [10] علي الشكل:

$$A^* = A^2 \left(1 + \frac{0.75}{n} + \frac{2.25}{n^2}\right) \qquad (2-10)$$

ويُعد اختبار Anderson – Darling أشهر اختبارات هذه المجموعة وأكثرها انتشاراً واستخداماً في البرامج الإحصائية الجاهزة ، وسيرمز لهذا الأختبار بالرمز AD . **ويتميز** هذا الإختبار بإمكانية استخدامه لجميع أنواع البيانات سواء كانت متماثلة أو ذات الإلتواء موجب أو سالب ، كما يتميز هذا الإختبار في كونه يعطي وزناً أكبر للمشاهدات علي أطراف التوزيع وبذلك يتمتع هذا الإختبار بحساسية أكبر عند أطراف التوزيع والتي غالباً ما تحتوي علي القيم الشاذة بالمقارنة بإختبار Kolomogorov والذي يتمتع بحساسية أكبر قرب مركز التوزيع ، وهذه الحساسية للقيم الشاذة تُعني قدرة الإختبار علي اتخاذ قرار صحيح يتعلق بقبول أو رفض فرضية الدراسة بالرغم من وجود القيم الشاذة .

فيما **يُعاب** علي هذا الإختبار الحاجة إلي جداول لحساب القيمة الحرجة والتي تُستخدم لإتخاذ القرار الإحصائي بالقبول أو الرفض.

ثانياُ: اختبارات تعتمد علي مفهوم الإنحدار:

- توصلا كلاً من Shapiro and Wilk [8] إلي إحصائية إختبار W تمثل هذه الإحصائية النسبة بين توليفة خطية مناسبة في الإحصاءات المرتبة للعينة order statistic وبين تباين العينة . وتأخذ إحصائية Shapiro–Wilk الصورة التالية:

$$W = \frac{\left[\sum_{i=1}^{[n/2]} a_{n-i+1}\left(x_{(n-i+1)} - x_{(i)}\right)\right]^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2} \qquad (2-11)$$

وتعتمد المعاملات (a_i) والتي تمثل أحسن معاملات خطية تم معايرتها –"The normalized best linear unbiased coefficients" علي مصفوفة التباينات والتغايرات للإحصاءات المرتبة للتوزيع المعتدل المعياري؛ وكلما اقتربت قيمة W من الواحد الصحيح كلما زاد اقتراب توزيع البيانات من التوزيع المعتدل. وهذه الإحصائية تحتاج لجداول لحساب القيم الحرجة وذلك لإحجام العينات *n=3(1)50*.

- قدما Shapiro and Wilk [9] تقريب لهذه المعاملات لا يتطلب استخدام هذه المصفوفة وذلك لإحجام العينات n=7(1)50 يأخذ الشكل التالي :

$$i = 2,3, \dots\dots, n-1 \qquad (2-12)$$

$$\hat{a}_i^* = \begin{cases} 2\mu_{(i)} & i = 2,3,\dots,n-1 \\ \left\{\frac{\hat{a}_1^2}{1-2\hat{a}_1^2}\sum_{i=2}^{n-1}\hat{a}_i^{*2}\right\}^{\frac{1}{2}} & i = 1, i = n \end{cases}$$

حيث :

$$\hat{a}_1^2 = \hat{a}_n^2 \begin{cases} \frac{\Gamma\left(\frac{1}{2}n\right)}{\sqrt{2}\Gamma\left\{\frac{1}{2}(n+1)\right\}} & n \le 20 \\ \frac{\Gamma\left\{\frac{1}{2}(n+1)\right\}}{\sqrt{2}\Gamma\left(\frac{n}{2}+1\right)} & n > 20 \end{cases}$$

ويجب ملاحظة أن $\hat{a}_1^*$ يجب أن تؤخذ بإشارة سالبة كما أن المعاملات المقربة يجب أن يتم معايرتها بحيث يكون $\sum_{i=1}^{n}\hat{a}_i^{*2} = 1$. ويمكن الحصول علي قيم دقيقة للقيم المتوقعة للإحصاءات المرتبة للتوزيع المعتدل المعياري $\mu_{(i)}$ (لجميع قيم $n \le 400$) من خلال الجداول التي قدمها Harter [5] وسيرمز لهذا الأختبار بالرمز SW .

ثالثا: اختبارات تعتمد علي الانحراف المطلق عن الوسيط :

- قدم Gel,Y.R. et al [4] اختبار جديد للإعتدال يُطلق عليه directed test ، حيث يعتمد علي النسبة بين الإنحراف المعياري للعينة S_n ومقياس متين Robust للتشتت يُطلق عليه Robust Average Absolute Deviation from The Sample Median (MAAD) أي متوسط الإنحراف المطلق عن الوسيط ،ويُعرف كالتالي (Gastwirth,J.L [2]) :

$$J_n = \frac{c}{n}\sum_{i=1}^{n}|X_i - M| \qquad (2-13)$$

حيث M هو وسيط العينة أما $c = \sqrt{\pi/2}$ ،وتأخذ احصائية هذا الإختبار الشكل التالي :

$$SJ = \frac{S_n}{J_n} \qquad (2-14)$$

حيث :

$$S_n = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2}$$

ويقترب توزيع هذه الإحصائية من التوزيع الطبيعي $\sqrt{n}(SJ-1) \sim N\left(0, \sigma_{SJ}^2\right)$، حيث :

$$\sigma_{SJ}^2 = \frac{\pi - 3}{2}$$

وبإفتراض صحة فرض العدم فإن هذه النسبة " SJ " تقترب من الواحد الصحيح . وسيرمز لهذا الأختبار بالرمز Gel .

- وجد كلاً من Gel,Y.R. and Ga-stwirth,J.L [3] أن العزوم المركزية للعينة تتأثر بشكل كبير بالقيم الشاذة – ويكون تباين العينه أكثر تأثراً مقارنة بمتوسطها – ومن ثم يتأثر كلاً من معامل الإلتواء والتفرطح في إحصائية إختبار JB . لذا قاما بإجراء تعديلاً علي إختبار *JB* بإستخدام J_n عُرف بإختبار (*RJB*) أي Robust JB ، وتكون المقدرات المتينة لكلاً من معاملي الإلتواء والتفرطح بإستخدام J_n علي الصورة :

$$\sqrt{b_1} = \hat{\mu}_3 / J_n^3 \qquad (2-15)$$

$$b_2 = \hat{\mu}_4 / J_n^4 \qquad (2-16)$$

حيث $\hat{\mu}_3$ و $\hat{\mu}_4$ هما العزم الثالث والرابع المركزي للعينة علي التوالي .

وتأخذ احصائية اختبار *RJB* الشكل التالي:

$$RJB = \frac{n}{c_1}\left(\frac{\hat{\mu}_3}{J_n^3}\right)^2 + \frac{n}{c_2}\left(\frac{\hat{\mu}_4}{J_n^4} - 3\right)^2 \qquad (2-17)$$

حيث c_1 و c_2 مقادير ثابتة ويتم حسابها بإستخدام المحاكاة ولكن يُوصي بإستخدام $c_1 = 6$ و $c_2 = 64$ وذلك للحفاظ علي المستوي الإسمي "The Nominal Level ٠.٠٥، ويقترب توزيع هذه الإحصائية من توزيع $\chi^2_{(2)}$. وسيرمز لهذا الأختبار بالرمز GGW.

٣- دراسة المحاكاة:

تُحاول دراسة المحاكاة (بإستخدام لغة *R*) المقارنة بين إحصائيات الإختبار محل الدراسة في إختبار الفرض الإحصائي بأن المجتمع المسحوب منه عينة الدراسة يتبع توزيعاً معتدلاً . ولتحقيق هذا الهدف تم توليد ٥٠٠٠٠ عينة عشوائية بأحجام مختلفة " *n* " (٢٠، ٣٠ ، ٥٠، ١٠٠، ٥٠٠) من كل توزيع من التوزيعات الإحتمالية الموضحة بالجدول التالي مع ذكر الرمز المستخدم :

التوزيع الاحتمالي	الرمز المستخدم
توزيع لمدا المعمم	GLD$(\lambda_1, \lambda_2, \lambda_3, \lambda_4)$
توزيع وايبل	Weibull(α)
التوزيع اللوغاريتمي الطبيعي	Lognormal
التوزيع المعتدل المختلط	MN(p,m, b^2)
توزيع ت	T($\boldsymbol{\nu}$)
التوزيع المنتظم المستمر	Uniform
توزيع بيتا	Beta(α,β)
توزيع Arcsine	Arcsine
توزيع Laplace	Laplace
التوزيع اللوجيستي	Logistic
توزيع χ^2	χ^2_v
التوزيع المعتدل المبتور	Tnorm(μ, σ^2)
التوزيع الأسي	Exponential
توزيع جاما ذو المعلمتين	G(α,β)
توزيع Half-Normal	Half N(μ, σ^2)

وتمت المقارنة بين هذه الإختبارات وفقاً لمعيارين هما :

- **مستوي المعنوية التجريبي" *Empirical Level*** " وهو يُمثل نسبة عدد مرات اجتياز قيمة احصائية اختبار معين للقيمة الحرجة المناظرة لمستوى معنوية أسمي*"Nominal Level"* (في دراستنا الحالية ٥%)- أي عندما تكون قيمة احصائية الاختبار أكبر من القيمة الحرجة- وذلك بإفتراض أن البيانات مسحوبة فعلاً من مجتمع توزيعه الاحتمالي معتدلاً .

وبدرجة ثقة ٩٩% فإن عدد العينات المستخدم في دراسة المحاكاة (٥٠٠٠٠ عينة) يجعل الحد الأقصي لخطأ تقدير مستوى المعنوية الإسمي٥% هو ٠.٠٠٢٥

$$\left\{Z_{0.005} * \sqrt{\frac{\alpha(1-\alpha)}{50000}} = 2.576 * \sqrt{\frac{0.05(1-0.05)}{50000}}\right\}$$

،ومن ثم فإن الإختبار الذي يُصاحبه مستوى معنوية تجريبي يزيد عن ٠.٠٥٢٥ سيعتبر اختباراً متحرراً " *Liberal* " ومن ثم سيعتبر اختبار غير صالح وهو ما يوضحه جدول (٣-١).

- **قوة الإختبار التجريبية "*Empirical Power*"** وهي تُمثل نسبة عدد مرات إجتياز قيمة إحصائية اختبار معين للقيمة الحرجة المناظرة لمستوي معنوية إسمي (في دراستنا ٥%) وذلك بإفتراض أن البيانات مسحوبة فعلاً من مجتمع توزيعه الاحتمالي غير معتدل.

جدول (٣-١) مستوي المعنوية التجريبي لإختبارات الاعتدال

N	Lill	AD	GGW	Gel	SW
20	0.04760	0.04864	0.04996	0.05024	0.04986
30	0.04906	0.04984	0.05232	0.05190	0.05056
50	0.04952	0.05130	0.05060	0.05076	0.05040
100	0.05012	0.04904	0.05060	0.05080	0.04954
500	0.05034	0.04866	0.05126	0.05114	0.05008

٤- نتائج دراسة المحاكاة:

أولاً : في حالة التوزيعات المتماثلة:

- **في حالة التوزيعات المتماثلة ثنائية القمة**

➢ يكون لإختبارات الإنحدار والتي تعتمد علي الإحصاءات الترتيبية قوة أعلي مقارنة بالإختبارات التي تعتمد علي متوسط الانحراف المطلق عن الوسيط .

➢ يكون اختبار *AD* الأعلي قوة مقارنة بباقي الأختبارات ، ولكن يتفوق عليه اختبار *Gel* في القوة وذلك بإنخفاض *m* وزيادة درجة التفرطح .

➢ تقل قوة الاختبار *GGW* بزيادة حجم العينة وانخفاض مستوي التلوث وذلك عندما $n \leq 50$.

- **في حالة التوزيعات المتماثلة ذات التفرطح المنخفض**

وعندما يكون معامل التفرطح أكبر من (١,٨)، يتفوق دائما اختبار *Gel* علي اختبار *SW* واختبار *AD* في القوة يليهم اختبار *Lilli* .

- **في حالة التوزيعات المتماثلة ذات التفرطح الأعلي قليلاً من تفرطح التوزيع المعتدل**

➢ تكون الإختبارات المُعتمدة علي الإنحراف المطلق عن الوسيط قوة أعلي من غيرها من الإختبارات .

➢ يُحقق اختبار *Lilli* مستويات منخفضة جداً من القوة مقارنة بباقي الإختبارات .

➢ تزداد قوة اختبار *Gel* مقارنة بإختبار *GGW* وذلك بزيادة حجم العينة ،بينما تتساوى القوة عندما $n \leq 100$ ، وذلك في حالة توزيع $t(15)$.

➢ يتفوق اختبار*GGW* علي اختبار *Gel* في القوة وذلك بزيادة درجة التفرطح أي عندما $(\beta_2 \geq 4)$.

- **في حالة التوزيعات المتماثلة ذات التفرطح العالي**

➢ يكون اختبار *Gel* الأقوى مقارنة بباقي الأختبارات .

➢ اختبار *AD* يكون الأكثر قوة من اختبار*SW* عندما يكون التوزيع الأصلي للبيانات هو توزيع Tukey(10) .

ثانياً: في حالة التوزيعات الملتوية

- **في حالة التوزيعات ثنائية القمة (المعتدلة ذات معلمة الموضع الملوثة)**

➢ يكون اختبار *GGW* أعلي قوة عندما $\rho = 0.05$ ، ويتفوق عليه اختبار *SW* وذلك بزيادة ρ .

يليه اختبار *AD*.

➢ بزيادة قيمة ρ تقل درجة التفرطح للتوزيع ،وذلك عند ثبات قيمة *(m)* .

- **في حالة التوزيعات الملتوية وحيدة القمة ذات التفرطح المنخفض**

➢ تؤول الإختبارات التي تعتمد علي متوسط الأنحراف المطلق عن الوسيط إلي أن يكون لها قوة منخفضة مقارنة بإختبار *AD* .

➢ يكون اختبار *SW* الأعلي قوة ويليه إختبار *AD* وذلك بزيادة درجة الالتواء .

- **في حالة التوزيعات الملتوية وحيدة القمة و ذات الالتواء العالي و/ أو التفرطح العالي**

➢ يكون اختبار *SW* الأقوى مقارنة بباقي الاختبارات يليه اختبار *AD* .

➢ بزيادة كلاً من درجتي الالتواء والتفرطح وبزيادة حجم العينة يُحقق كلاً من الاختبارين *Gel* و *GGW* قوة تنافسيه مع باقي الأختبارات .

و يتم عرض النتائج التفصيلية لدراسة المحاكاه في الجداول (١)، (٢) في الملحق

٥-توصيات الدراسة : وتوصي الدراسة بما يلي:

- استكمالاً للدراسة الحالية يمكن ادراج عدداً آخر من اختبارات الاعتدال التي لم تُدرج بعد في دراسة المحاكاة .
- يمكن دراسة تأثير وجود القيم الشاذة في العينة علي نتائج دراسة المحاكاة.
- يمكن استخدام ما تم التوصل إليه من نتائج في اجراء دراسة مقارنة علي عينات مسحوبة من مجتمع متعدد المتغيرات *Multivariate distr-ibution*

قائمة المراجع

[1] Anderson,T.W.and Darling, D.A-.(1954) .A Test of Goodness -of- Fit .Journal of the American Statistical Association,49,756-790.

[2] Gastwirth,J,L.(1982). Statistical Properties of A Measure of Tax Assessment Uniformity.J. Statist. Planning and Inference ,6,135-150 .

[3] Gel,Y.R. and Gastwirth,J.L.(2008).Arobust Modification of The Jarque-Bera Test of Normality .Economics Letters,99,30-32.

[4] Gel,Y.R.,Miao,W. and Gastwirth,J-.L.(2007). Robust Directed Tests of Normality Against Heavy –Tailed Alternatives.Computational Statistics and Data analysis ,51,2734-2746.

[5] Harter,H.L. (1961).Expected Values of Normal Order Statistics. Biometrika,48,1 and 2,151-165.

[6] Kolomogrov,N. (1933).Sulla Determinazione Empirica di una Legge di Distribuzione.Giormale del Istitute it a lia no di Attuari ,4,83-90.

[7] Lilliefors, H.W. (1967) .On the Kolmogorov– Smirnov Test for Normality with Mean and Variance Unknown. Journal of the American Statistical Association,62, 399–402.

[8] Shapiro, S.S. and Wilk, M.B.(1965).An Analysis of Variance Test for Normality (Complete Samples). Biometrika,52,591-611.

[9] Shapiro, S.S. and Wilk, M.B. (1968).Approximations for The Null Distribution of The W statistic. Technometrics ,10,861–866.

[10] Stephens, M.A.(1974). EDF Statistics for Goodness of Fit and Some Comparisons. Journal of the Ame-rican Statistical Association,-69,730– 7.

ملحق

جدول (١) القوي التجريبية لتوزيعات متماثلة غير معتدلة

bimodal distributions توزيعات ثنائية القمة						
	n	Lill	AD	GGW	SW	Gel
MN (0.5, 10, 1)	20	0.99986	0.99998	0.04200	0.99998	0.56294
	30	1.00000	1.00000	0.02346	1.00000	0.69384
	50	1.00000	1.00000	0.01304	1.00000	0.87898
	100	1.00000	1.00000	1.00000	1.00000	0.98444
	500	1.00000	1.00000	1.00000	1.00000	1.00000
Arcsine	20	0.31190	0.61706	0.00508	0.72538	0.42094
	30	0.50168	0.85968	0.00180	0.94260	0.69648
	50	0.79726	0.99192	0.00054	0.99944	0.94000
	100	0.99444	1.00000	0.83940	1.00000	0.99934
	500	1.00000	1.00000	1.00000	1.00000	1.00000
MN (0.5, 5,1)	20	0.64556	0.81360	0.01740	0.75586	0.51240
	30	0.87122	0.96724	0.00624	0.94204	0.71444
	50	0.99274	0.99966	0.00238	0.99842	0.90700
	100	1.00000	1.00000	0.99874	1.00000	0.99508
	500	1.00000	1.00000	1.00000	1.00000	1.00000
MN (0.5, 4, 1)	20	0.34292	0.46104	0.01172	0.40140	0.38152
	30	0.53670	0.70508	0.00402	0.63152	0.62106
	50	0.82386	0.94034	0.00130	0.90034	0.87930
	100	0.99472	0.99970	0.41178	0.99930	0.99400
	500	1.00000	1.00000	1.00000	1.00000	1.00000
Tukey (1.5)	20	0.11516	0.21166	0.00346	0.25300	0.21276
	30	0.17078	0.36798	0.00122	0.47374	0.40160
	50	0.31318	0.67806	0.00036	0.84066	0.71620
	100	0.69496	0.97984	0.08156	0.99934	0.97376
	500	1.00000	1.00000	1.00000	1.00000	1.00000
MN (0.5, 3,1)	20	0.12218	0.15066	0.00942	0.12906	0.17416
	30	0.17816	0.24112	0.00406	0.20298	0.31726
	50	0.31348	0.44442	0.00120	0.37634	0.59656
	100	0.63888	0.81516	0.00674	0.74852	0.92650
	500	1.00000	1.00000	1.00000	1.00000	1.00000
Short tailed distributions توزيعات وحيدة القمة ذات تفرطح منخفض						
	n	Lill	AD	GGW	SW	Gel
uniform	20	0.09820	0.17290	0.00370	0.20184	0.0032
	30	0.14316	0.29908	0.00112	0.38390	0.0004
	50	0.25684	0.57640	0.00026	0.75140	0.0000
	100	0.59324	0.94956	0.03952	0.99676	0.0000
	500	1.00000	1.00000	1.00000	1.00000	0.0000
	20	0.07218	0.10860	0.00408	0.11704	0.13558

Tukey(0.7)	30	0.09568	0.17546	0.00148	0.21144	0.25030
	50	0.16192	0.35292	0.00032	0.47912	0.49566
	100	0.37680	0.77206	0.00578	0.93692	0.86058
	500	0.99924	1.00000	1.00000	1.00000	1.00000
Tukey(3)	**n**	**Lill**	**AD**	**GGW**	**SW**	**Gel**
	20	0.04710	0.06246	0.00448	0.06536	0.08246
	30	0.05494	0.08894	0.00166	0.10910	0.14020
	50	0.07512	0.16428	0.00040	0.26232	0.27436
	100	0.14906	0.43952	0.00018	0.76528	0.57884
	500	0.90180	0.99998	0.99972	1.00000	0.99950
Beta(2,2)	20	0.05018	0.05700	0.00648	0.05266	0.0088
	30	0.05910	0.07788	0.00260	0.07674	0.0040
	50	0.07946	0.13192	0.00052	0.15200	0.0009
	100	0.15332	0.31602	0.00014	0.45066	0.0004
	500	0.83018	0.99804	0.99742	1.00000	0.0000
Tnorm (0,1,-2,2)	20	0.04106	0.03974	0.01198	0.03412	0.05832
	30	0.04194	0.04214	0.00528	0.03680	0.07196
	50	0.04676	0.05552	0.00176	0.05272	0.10772
	100	0.06172	0.08988	0.00036	0.12304	0.21912
	500	0.21742	0.62784	0.60908	0.99560	0.82548
MN (0.5, 2,1)	20	0.0485	0.0501	0.0210	0.0449	0.0697
	30	0.0546	0.0582	0.0154	0.0520	0.0928
	50	0.0656	0.0778	0.0090	0.0649	0.1465
	100	0.1006	0.1328	0.0031	0.1113	0.2824
	500	0.4496	0.6719	0.4573	0.6365	0.8996
MN (0.5, 1,1)	**n**	**Lill**	**AD**	**GGW**	**SW**	**Gel**
	20	0.0484	0.0476	0.0533	0.0478	0.0828
	30	0.0477	0.0480	0.0513	0.0467	0.0863
	50	0.0473	0.0457	0.0463	0.0436	0.0863
	100	0.0499	0.0491	0.0390	0.0450	0.0916
	500	0.0505	0.0520	0.0267	0.0480	0.1070

Distributions with kurtosis slightly higher than the normal توزيعات وحيدة القمة ذات تفرطح أعلي قليلاً من تفرطح التوزيع المعتدل						
	n	Lill	AD	GGW	SW	Gel
t(15)	20	0.0602	0.0711	0.1121	0.0774	0.1278
	30	0.0627	0.0748	0.1310	0.0871	0.1473
	50	0.0659	0.0857	0.1625	0.1094	0.1712
	100	0.0758	0.0993	0.2153	0.1429	0.2162
	500	0.1335	0.2287	0.4854	0.3740	0.4808
GLD(0,0.0545,0.0317,0.0317)	20	0.0714	0.0861	0.1440	0.0943	0.1587
	30	0.0764	0.0968	0.1778	0.1121	0.1903
	50	0.0872	0.1165	0.2295	0.1450	0.2370
	100	0.1096	0.1578	0.3233	0.2084	0.3323
	500	0.3035	0.5098	0.7527	0.6284	0.7849
t(10)	20	0.0723	0.0884	0.1428	0.0987	0.14008
	30	0.0767	0.1014	0.1806	0.1198	0.16394
	50	0.0864	0.1210	0.2314	0.1558	0.20884
	100	0.1097	0.1601	0.3335	0.2279	0.29750
	500	0.2817	0.4836	0.7611	0.6553	0.74454
Logistic	20	0.0837	0.1061	0.1781	0.1166	0.1725
	30	0.0935	0.1245	0.2268	0.1459	0.2141
	50	0.1129	0.1610	0.3019	0.1954	0.2839
	100	0.1560	0.2376	0.4366	0.3027	0.4265
	500	0.5132	0.7521	0.9111	0.8384	0.9219

Distributions with high kurtosis　　**توزيعات وحيدة القمة ذات تفرطح عالي**						
	n	**Lill**	**AD**	**GGW**	**SW**	**Gel**
GLD (0, 0.0262, 0.0148, 0.0148)	20	0.0779	0.0956	0.1620	0.1060	0.1745
	30	0.0848	0.1102	0.2040	0.1301	0.2141
	50	0.1003	0.1381	0.2674	0.1708	0.2746
	100	0.1320	0.1979	0.3829	0.2571	0.3952

	500	0.4097	0.6449	0.8511	0.7527	0.8746
Tukey (10)	20	0.8954	0.9086	0.8490	0.8051	0.9515
	30	0.9788	0.9847	0.9488	0.9415	0.9936
	50	0.9994	0.9998	0.9948	0.9967	0.9999
	100	1.0000	1.0000	1.0000	1.0000	1.0000
	500	1.0000	1.0000	1.0000	1.0000	1.0000
Laplace	20	0.2184	0.2727	0.3851	0.2633	0.4224
	30	0.2925	0.3747	0.5094	0.3604	0.5598
	50	0.4312	0.5441	0.6833	0.5244	0.7499
	100	0.7059	0.8234	0.8979	0.7948	0.9470
	500	1.0000	0.9997	1.0000	1.0000	1.0000
t(4)	20	0.1725	0.2260	0.3257	0.2417	0.3234
	30	0.2166	0.2946	0.4324	0.3247	0.4188
	50	0.3001	0.4167	0.5887	0.4645	0.5760
	100	0.4885	0.6471	0.8136	0.7116	0.8068
	500	0.9856	0.9384	0.9997	0.9992	0.9998
t(2)	20	0.4501	0.5274	0.6299	0.5279	0.6473
	30	0.5869	0.6785	0.7766	0.6836	0.7946
	50	0.7784	0.8579	0.9193	0.8633	0.9312
	100	0.9593	0.9772	0.9938	0.9851	0.9962
	500	1.0000	1.0000	1.0000	1.0000	1.0000
t(1)	20	0.8444	0.8824	0.9130	0.8684	0.9300
	30	0.9450	0.9658	0.9766	0.9591	0.9836
	50	0.9942	0.9972	0.9987	0.9965	0.9994
	100	1.0000	0.9991	1.0000	1.0000	1.0000
	500	1.0000	1.0000	1.0000	1.0000	1.0000

جدول (٢) القوي التجريبية لتوزيعات ملتوية

توزيعات ثنائية القمة (توزيعات معتدلة ذات معلمة الموضع الملوثة)

Bimodal ,Location contaminated normal distributions

	n	Lill	AD	GGW	SW	Gel
LCN (0.2,3)	20	0.2099	0.2641	0.1939	0.2586	0.1867
	30	0.3187	0.4074	0.2539	0.3881	0.2138
	50	0.5127	0.6448	0.3867	0.6161	0.2595
	100	0.8421	0.9364	0.7185	0.9194	0.3550
	500	1.0000	1.0000	1.0000	1.0000	0.8172
LCN (0.05,3)	20	0.1152	0.1592	0.2435	0.1865	0.2249
	30	0.1419	0.2056	0.3305	0.2586	0.2808
	50	0.1979	0.2951	0.4732	0.3870	0.3711
	100	0.3399	0.5053	0.7150	0.6408	0.5423
	500	0.9313	0.9896	0.9989	0.9983	0.9739
LCN (0.1,3)	20	0.1683	0.2284	0.2857	0.2511	0.2566
	30	0.2351	0.3270	0.3958	0.3630	0.3283
	50	0.3649	0.5036	0.5821	0.5568	0.4439
	100	0.6429	0.8065	0.8656	0.8533	0.6688
	500	0.9997	1.0000	1.0000	1.0000	0.9980
LCN (0.2,5)	20	0.7627	0.8810	0.5474	0.8705	0.5379
	30	0.9293	0.9809	0.6935	0.9780	0.6453
	50	0.9959	0.9995	0.8846	0.9997	0.7766
	100	1.0000	1.0000	0.9993	1.0000	0.9274
	500	1.0000	1.0000	1.0000	1.0000	1.0000
LCN (0.2,7)	20	0.9671	0.9855	0.7167	0.9869	0.7476
	30	0.9961	0.9983	0.8371	0.9987	0.8357
	50	0.9999	1.0000	0.9607	1.0000	0.9244
	100	1.0000	1.0000	1.0000	1.0000	0.9890
	500	1.0000	1.0000	1.0000	1.0000	1.0000
LCN (0.1,5)	20	0.5834	0.7260	0.7453	0.7633	0.6958
	30	0.7547	0.8728	0.9069	0.9101	0.8555
	50	0.9202	0.9712	0.9888	0.9863	0.9704
	100	0.9958	0.9995	0.9999	0.9999	0.9994
	500	1.0000	1.0000	1.0000	1.0000	1.0000
	n	Lill	AD	GGW	SW	Gel

LCN (0.05,5)	20	0.3639	0.4914	0.5880	0.5443	0.5511
	30	0.4842	0.6358	0.7532	0.7134	0.7000
	50	0.6630	0.8032	0.9023	0.8807	0.8575
	100	0.8921	0.9574	0.9892	0.9860	0.9741
	500	1.0000	1.0000	1.0000	1.0000	1.0000
LCN (0.1,7)	20	0.8197	0.8711	0.8568	0.8786	0.8572
	30	0.9243	0.9548	0.9558	0.9592	0.9542
	50	0.9850	0.9932	0.9949	0.9947	0.9945
	100	0.9998	0.9999	1.0000	1.0000	1.0000
	500	1.0000	1.0000	1.0000	1.0000	1.0000
LCN (0.05,7)	20	0.5753	0.6416	0.6591	0.6520	0.6624
	30	0.7163	0.7844	0.7998	0.7956	0.8010
	50	0.8608	0.9157	0.9273	0.9262	0.9273
	100	0.9756	0.9897	0.9941	0.9940	0.9935
	500	1.0000	1.0000	1.0000	1.0000	1.0000

Unimodal low Kurtosis				**توزيعات وحيدة القمة ذات تفرطح منخفض**			
	n	Lill	AD	JB	GGW	SW	Gel
Weibull (4)	20	0.0463	0.0459	0.0124	0.0391	0.0422	0.0714
	30	0.0485	0.0476	0.0125	0.0322	0.0408	0.0731
	50	0.0499	0.0513	0.0124	0.0246	0.0438	0.0768
	100	0.0521	0.0531	0.0122	0.0263	0.0455	0.0786
	500	0.0593	0.0630	0.0101	0.0166	0.0520	0.0918
Weibull (3.6)	20	0.0448	0.0422	0.0109	0.0346	0.0385	0.0684
	30	0.0457	0.0443	0.0114	0.0287	0.0390	0.0720
	50	0.0460	0.0460	0.0105	0.0199	0.0379	0.0770
	100	0.0501	0.0512	0.0073	0.0113	0.0407	0.0969
	500	0.0731	0.1049	0.0658	0.0324	0.1293	0.2562

Beta (2,1)	20	0.1698	0.2585	0.0288	0.0684	0.3018	0.1084
	30	0.2599	0.4255	0.0427	0.0770	0.5150	0.1410
	50	0.4522	0.7240	0.1049	0.1133	0.8456	0.2105
	100	0.8225	0.9831	0.7465	0.4395	0.9987	0.3695
	500	1.00000	1.0000	1.00000	1.00000	1.0000	0.92542
Beta (3,2)	20	0.0638	0.0727	0.0073	0.0243	0.0722	0.0748
	30	0.0811	0.1034	0.0078	0.0198	0.1081	0.1037
	50	0.1200	0.1749	0.0086	0.0162	0.2032	0.1648
	100	0.2365	0.3972	0.0494	0.0256	0.5289	0.3340
	500	0.9320	0.9993	1.0000	0.9993	1.0000	1.0000
Weibull (2.2)	20	0.0773	0.0938	0.0473	0.0900	0.1079	0.0952
	30	0.0988	0.1275	0.0755	0.1142	0.1565	0.1031
	50	0.1423	0.1997	0.1328	0.1610	0.2691	0.1151
	100	0.2632	0.4033	0.3094	0.3044	0.5716	0.1269
	500	0.9089	0.9972	0.9996	0.9996	1.0000	0.1773
Weibull (2)	20	0.0984	0.1293	0.0688	0.1219	0.1522	0.1108
	30	0.1324	0.1871	0.1135	0.1615	0.2357	0.1210
	50	0.2045	0.3097	0.2104	0.2436	0.4140	0.1352
	100	0.3947	0.6122	0.5035	0.4786	0.7900	0.1451
	500	0.9877	1.0000	1.0000	1.0000	1.0000	0.1563
Half-Normal	20	0.2352	0.3644	0.1620	0.2445	0.4420	0.1773
	30	0.3486	0.5632	0.2825	0.3451	0.6809	0.1996
	50	0.5655	0.8328	0.5395	0.5418	0.9321	0.2229
	100	0.9115	0.9950	0.9617	0.9177	0.9999	0.2635
	500	1.0000	1.0000	1.0000	1.0000	1.0000	0.4243
توزيعات وحيدة القمة وذات التواء عالي و /أو ذات تفرطح عالي Unimodal ,high skweness and							

/or Kurtosis						
	n	Lill	AD	GGW	SW	Gel
χ_4^2	20	0.32098	0.46700	0.39154	0.53368	0.29402
	30	0.46230	0.66294	0.54568	0.74986	0.36536
	50	0.69292	0.88886	0.77450	0.94848	0.46496
	100	0.95286	0.99748	0.98446	0.99988	0.64490
	500	1.00000	1.00000	1.00000	1.00000	0.98680
	n	Lill	AD	GGW	SW	Gel
Gamma(2,3)	20	0.32098	0.46700	0.39154	0.53368	0.29402
	30	0.46230	0.66294	0.54568	0.74986	0.36536
	50	0.69292	0.88886	0.77450	0.94848	0.46496
	100	0.95286	0.99748	0.98446	0.99988	0.64490
	500	1.00000	1.00000	1.00000	1.00000	0.98680
Exponential	20	0.57300	0.77284	0.59290	0.83386	0.46378
	30	0.78050	0.93528	0.77554	0.96772	0.58294
	50	0.96056	0.99674	0.94784	0.99952	0.73820
	100	1.00000	1.00000	1.00000	0.91474	1.00000
	500	1.00000	1.00000	1.00000	1.00000	0.99996
χ_1^2	20	0.87806	0.96838	0.81544	0.98266	0.71648
	30	0.98278	0.99820	0.94322	0.99950	0.84468
	50	0.99988	1.00000	0.99646	1.00000	0.95212
	100	1.00000	1.00000	1.00000	1.00000	0.99742
	500	1.00000	1.00000	1.00000	1.00000	1.00000
Lognormal	20	0.78588	0.90486	0.80784	0.93168	0.72044
	30	0.93024	0.98478	0.93500	0.99188	0.85076
	50	0.99482	0.99968	0.99470	0.99998	0.95510
	100	1.00000	1.00000	1.00000	1.00000	0.99820
	500	1.00000	1.00000	1.00000	1.00000	1.00000
	20	0.98298	0.99778	0.95378	0.99894	0.91980

weibull(0.5)	30	0.99958	0.99996	0.99464	1.00000	0.97912
	50	1.00000	1.00000	0.99998	1.00000	0.99850
	100	1.00000	1.00000	1.00000	1.00000	1.00000
	500	1.00000	1.00000	1.00000	1.00000	1.00000