University of Benghazi
Faculty of Education Almarj

جامعة بنغازي
كلية التربية – المرج
ISSN 2518-5845

Global Libyan Journal
المجلة الليبية العالمية

العدد الخامس والاربعون / اكتوبر/2019

متانة الاختبار المعلمي (t-s'Student) والاختبار اللامعلمي (Mann Whitny-Wilcoxon) في حالة العينتين المستقلتين

د. جلال عبدالله أمعيطي – رئيس قسم الاحصاء –كلية الآداب والعلوم المرج – جامعة بنغازي – ليبيا
أ. راضي عبدالرحيم عثمان – عضو هيئة تدريس بقسم الاحصاء – كلية الآداب والعلوم المرج – جامعة بنغازي – ليبيا

University of Benghazi
Faculty of Education Almarj

جامعة بنغازي
كلية التربية – المرج
ISSN 2518-5845

Global Libyan Journal

المجلة الليبية العالمية

العدد الخامس والاربعون / اكتوبر/2019

متانة الاختبار المعلمي (t-s'Student) والاختبار اللامعلمي (Mann Whitny-Wilcoxon) في حالة العينتين المستقلتين

الملخص:

في هذه الدراسة تمت دراسة متانة (Robustness) الاختبار المعلمي (Student's t-Tes) لعينتين مستقلتين والاختبار الاحصائي اللامعلمي(Mann Whitney Whilcoxon Test) الذي يستخدم عند عدم استيفاء البيانات لافتراضات اختبار (t) ، حيث تم استخدام تقنية المحاكاة (Simulation) لتوليد بيانات الدراسة المتمثلة في ثلاث مجتمعات كبيرة الحجم (100000مشاهدة) لكل مجتمع. ولغرض المقارنة بين الاختبارين ولدت هذه المجتمعات مختلفة من ناحية توزيع المشاهدات (مجتمع طبيعي معياري ، مجتمع متوسط الالتواء والتفرطح ، مجتمع شديد الالتواء والتفرطح). ومن ثم تمت المعاينة من المجتمعات الثلاثة (1000عينة) لكل حالة مع التركيز على حجمي العينتين (متساويتان، الفرق بينها صغير، الفرق بينهما متوسط ، الفرق بينهما كبير).وتم تطبيق الاختبارين على هذه العينات ، وتوصلت الدراسة للنتائج التالية :

- في حالة المعاينة من توزيع طبيعي معياري (تحقق شرط اعتدالية التوزيع للعينتين) اتضح أن احتمالية الوقوع في الخطأ من النوع الاول لاتختلف بشكل كبير بين الاختبارين وكلاهما يقترب من القيمة النظرية (0.05) كلما كان حجم العينات كبير ولايوجد اختلاف كبير بين حجميهما ، ومع ذلك نلاحظ ان القيم المحسوبة لتوزيع (t) هي الاقرب والأفضل في معظم الحالات.
- في حالة المعاينة من توزيع متوسط الالتواء والتفرطح (عدم تحقق شرط اعتدالية التوزيع للعينتين) يتضح أن القيم المحسوبة من اختبار (MWW) افضل من تلك المحسوبة من اختبار(t) والذي اتضح أنه يتأثر عندما تكون المعاينة من توزيع غير طبيعي وخاصة إذا كان هناك اختلاف كبير بين حجمي العينتين المستقلتين ، مع ملاحظة أنه لايتأثر بشكل كبير اذا كان حجمي العينتين كبير ومتقارب في الحجم .
- في حالة المعاينة من توزيع شديد الالتواء والتفرطح (عدم تحقق شرط اعتدالية التوزيع للعينتين) يتضح أن القيم المحسوبة من الاختبار اللامعلمي(MWW) افضل من تلك التي تم حسابها من الاختبار المعلمي (Student's t-Test) والذي اتضح أنه يتأثر بشكل كبير عندما تكون المعاينة من توزيع غير طبيعي وخاصة اذا كان هناك اختلاف كبير بين حجمي العينتين المستقلتين ، مع ملاحظة أنه لا يتأثر بشكل كبير اذا كان حجمي العينتين المستقلتين كبير ومتساوي .

الكلمات المفتاحية: *الاختبارات المعلمية واللامعلمية ، متانة وقوة الاختبار ، اختبار (t) ، اختبار (MWW).*

Abstract :

In this study, the robustness of the (Student's-t Test) for two independent samples and the (Mann Whitney-Whilcoxon Test) were compared. Simulation was used to generate study data from three large populations (100,000 observations) For each population,For the purpose of comparing the two tests, these different populations were born in terms of the distribution of observations standard normal population, Mediator Kurtosis and Skewness population, highly Kurtosis and Skewness population). And then, the sampling of the three populations (1000 samples) for each case with emphasis on the size of the two samples (equal, the difference between them is small, the difference between them is Mediator, the difference between them is large). The tests were applied to these samples, and the study reached the following results:

• In the case of sampling from a standard normal distribution (normality distribution assumption for both samples), it was found that the probability of (type I error) did not differ significantly between the two tests. And both are approaching the theoretical value (0.05) whenever the size of the samples is large and there is no significant difference between their size, However, we note that the calculated values of the distribution of (t) are the closest and best in most cases.

• In the case of sampling from the distribution of the mediator Kurtosis and Skewness (failure to normality distribution assumption for both samples), it is clear that the values calculated from the(MWW) test are better than those calculated from the (t-test), which is shown to be affected when the sampling is abnormal, especially if there is a significant difference between the two independent sample sizes, Note that it is not significantly affected if the size of the two samples is large and convergent in size.

• In the case of sampling from a highly Kurtosis and Skewness distribution (failure to normality distribution assumption for both samples), the values calculated from the non-parametric test (MWW) are better than those calculated from the parametric test (Student's-t) Which was found to be significantly affected when the sampling was abnormal, especially if there was a significant difference between the two independent sample sizes, noting that it was not significantly affected if the two independent sample sizes were large and equal.

Key words:robustness, power, independence assumption, t test, Mann Whitney Whilcoxon Test.

1. مقدمة **(Introduction)**:

الإحصاء الاستدلالي هو الفرع المتعلق بطرق اتخاذ القرارات حول مجتمع الدراسة ويتم ذلك عن طريق دراسة عينة مسحوبة عشوائيا من ذلك المجتمع ، ويهدف الإحصاء الاستدلالي إلى إيجاد تقديرات لمعالم مجهولة أو الإجابة عن بعض الأسئلة البحثية أو التحقق من بعض الفروض المسبقة حول هذه المعالم المجهولة. ويمكن تعريف الإحصاء الاستدلالي بأنه ذلك العلم المنبثق عن علم الإحصاء المهتّم بانتقاء العينات العشوائية سعياً للوصول إلى أهم الاستنتاجات حول معالم المجتمع المجهولة وبشكلٍ أدق فإنه بحث عميق واستدلال حول المجتمع على وجه الخصوص من خلال استخلاص عينة عشوائية منه للكشف عن سلوك شائع الانتشار في مدة زمنية محددة. ويأتي ذلك سعياً لتحقيق الدقة والمصداقية عند إطلاق الأحكام على المجتمعات والإحصاء الاستدلالي(Statistical inference)ينشطر إلى التقدير الإحصائي (Statistical Estimation) الذي يعتمد بالدرجة الأولى على كافة الطرق المتفاوتة والمستخدمة للكشف ورصد المعالم المجهولة لمجتمع ما .والشطر الثاني اختبارات الفروض الإحصائية (Hypothesis Testing) التي تستخدم لإيجاد معلومات تتلخص حول القيم الخاصة بالمجتمع لتسهيل عملية الاستدلال على معالمه (Parameters) من خلال الإجابة على الأسئلة المطروحة ، ويكون ذلك غالباً بواسطة تحليل البيانات. وهذه الدراسة على اختبارات الفروض الاحصائية ، حيث تعرف الفرضية الإحصائية (Statistical Hypothesis) بأنها عبارة عن ادعاء أو تصريح (قد يكون صائباً أو خاطئاً) حول معلمة أو أكثر لمجتمع أو لمجموعة من المجتمعات. وأصبح اختبار الفرضيات الاحصائية من أهم الخصائص التي تميز البحوث الميدانية والتجريبية في مجال التربية وعلم النفس والعلوم الانسانية والتطبيقية بصورة عامة والهدف الاساسي من اختبار الفرضيات هو استنتاج خصائص المجتمع أو بعضها من ملاحظة العينة التي اخذت منه وذلك بهدف تعميم ما نتوصل اليه من دراستنا لعينة صغيرة على ذلك المجتمع الذي تمثله تلك العينة . والإحصاء الاستدلالي يقوم على لغة الاحتمال والفرضيات نوعان :

- **الفرضية الصفرية (Null Hypothesis)**:ويرمز لها(H_0) التي يتم اختبارها احصائياً . وتعني عدم وجود علاقة بين المتغيرات أو عدم وجود فروق بين المجموعات ولذلك تسمى بفرضية العدم .
- **الفرضية البديلة (Alternative Hypothesis)**:ويرمز لها(H_1) تكون عكس الفرضية الصفرية.

عند اختبار فرض العدم(H_0) ضد الفرض البديل(H_1) نجد أننا أمام احدى الحالات الاربع الاتية:

- أن يكون فرض العدم صحيحا ويكون القرار بقبوله (القرار صحيح).
- أن يكون فرض العدم صحيحا ويكون القرار برفضه (القرار خاطئ).
- أن يكون فرض العدم خطأ ويكون القرار برفضه (القرار صحيح).
- أن يكون فرض العدم خطأ ويكون القرار بقبوله (القرار خاطئ).

- **الخطأ من النوع الأول (Type I Error)**:رفض (H_0) عندما تكون صحيحة ويرمز لحجم هذا الخطأ بالرمز(α). ويسمى مستوى المعنوية (Level of Significant).

• **الخطأ من النوع الثاني (Type II Error)**:قبول (H_0) عندما تكون خاطئة ويرمز لحجم هذا الخطأ بالرمز (β).

والعلاقة بين الفا وبيتا علاقة عكسية فكلما زادت قيمة الاولى انخفضت قيمة الثانية وبالعكس والاختبار الجيد هو الاختبار الذي يجعل قيمتيهما أقل ما يمكن [4].

ويمكن تقسيم الاختبارات الاحصائية حسب طبيعة بيانات العينة التي تحت الدراسة إلى فرعين هما :

• **الاختبارات المعلمية (Parametric Tests)**:تعتبر الاختبارات الاحصائية المعلمية واحدة من أهم التطبيقات التي قدمها علم الاحصاء وتستخدم في أغلب فروع العلوم التطبيقية والإنسانية ، وهي أحد أنواع الاساليب الاحصائية الاستدلالية التي تهتم بالكشف والاستدلال على معالم المجتمع اعتمادا على ما توافر من بيانات أو معلومات لدى الباحث من العينة المأخوذة من مجتمع الدراسة ، مع مراعاة بعض الافتراضات مثل توزيع البيانات توزيعا طبيعيا ، تجانس التباين ، عشوائية العينات ، استقلال العينات وغيرها من الافتراضات. وتتميز الاختبارات الاحصائية المعلمية في حال توفر افتراضات الاختبار بأنها تكون أدق وأكثر كفاءة من الاحصاء اللامعلمي. ولكن يعاب على الاساليب الاحصائية المعلمية أنها تعتمد على فروض كثيرة تحتاج تفهم واستيفاء لذلك تعتبر أكثر صعوبة من الاختبارات اللامعلمية ، كما أنها مقتصرة فالتطبيق على البيانات الكمية فقط.

• **الاختبارات اللامعلمية (Nonparametric Tests)**:الاساليب اللامعلمية لاتتطلب افتراضات أو معلومات حول توزيع المجتمع ، لذلك تسمى طرق التوزيع الحر(Distribution-free Methods) وتكون أكثر استخداما مع الظواهر التي يصعب فيها الحصول على قياسات دقيقة ، وتصلح للعينات الصغيرة ويمكن الاعتماد على نتائجها بدرجة كبيرة ، وهي سهلة الفهم والتطبيق ، والأسلوب الاحصائي اللامعلمي يمكن اعتماده في حالة البيانات الاسمية والترتيبية ، ولكن يعاب على هذا الاسلوب أنه يكون أقل قوة من الطرق المعلمية عند استيفاء الاختبارات المعلمية لمتطلباتها وافتراضاتها وهذا مايعرف بالكفاءة النسبية (Relative Efficiency) للطرق اللامعلمية بالمقارنة مع الطرق المعلمية ، كما يعاب على هذا الاسلوب أنه عندما يكون حجم العينة كبير جدا فإن استخدام الاسلوب اللامعلمي يكون صعبا. وكذلك في حالة استخدام الاسلوب اللامعلمي لدراسة الفروق بين المجتمعات فلا يمكننا تحديد هذه الفروق حتى ولو كان هذا الفرق حقيقا.

والجدير بالذكر أنه في العديد من الدراسات والبحوث الاحصائية غالبا مايتم تطبيق الاختبارات المعلمية دون فحص البيانات والتأكد من الافتراضات(Assumptions) الاساسية الواجب توفرها لاستخدام مثل تلك الاساليب ، حيث يؤدي انتهاك احد الفروض الى نتائج مضللة لايمكن الوثوق بها لاتصافها بعدم الدقة[2].

ومن هنا جاءت أهمية هذه الدراسة في اختيار الاسلوب الاحصائي المناسب في حالة عدم تحقق احد الشروط المعلمية الاساسية حيث تقوم الدراسة على اجراء مقارنة للمفاضلة بين الطرق المعلمية واللامعلمية في حالة انتهاك أهم الفرضيات الواجب توفرها لاختبار عينتين مستقلتين (فرض تبعية البيانات للتوزيع الطبيعي) مع عدة خيارات لحجمي العينتين المستقلتين (متساويتان، الفرق بينها صغير، الفرق بينهما متوسط ، الفرق بينهما كبير).، ودراسة مدى تأثير عدم تحقق هذا الفرض في كل حالة على قوة ومتانة الاختيار الاحصائي المعلمي(Student's t-Test) لعينتين مستقلتين، والاختبار الاحصائي اللامعلمي (MWW Test).

2.الهدف من الدراسة (The Aim of the Study):

تقوم هذه الدراسة على المقارنة بين كفاءة وقوة الاختبارات المعلمية واللامعلمية حول المتوسط لعينتين مستقلتين ، فعند استيفاء شروط الاختبارات المعلمية فإن الاختبار الاحصائي المناسب هو اختبار(t) لعينتين مستقلتين ، أما في حالة انتهاك احد شروط اختبار (t) فإن الاختبار الاحصائي اللامعلمي المناسب هو اختبار (Mann Whitny-Wilcoxon).

والمعيار الذي تعتمد عليه الدراسة في المقارنة بين الاسلوبين يستند على قياس أو معرفة مدي متانة (Robustness) أو حساسية (Sensitivity) الاختبارين في ثلاث حالات :

- في حالة المعاينة من التوزيع الطبيعي المعياري.
- في حالة المعاينة من توزيع متوسط الالتواء والتفرطح .
- في حالة المعاينة من توزيع شديد الالتواء والتفرطح .

مع مراعاة تغير حجم العينة في كل حالة من الحالات السابقة لمعرفة مدي تأثر هذا الاختبار باختلاف حجمي العينتين المستقلتين .

وللوصول للهدف المرجو من الدراسة والمتمثل فالمقارنة بين الاسلوبين المذكورين أعلاه عن طريق اختبار فرض العدم للحالتين المعلمية واللامعلمية ، تم استخدام اسلوب المحاكاة (Simulation) لتوليد البيانات بحسب ما تقتضيه متطلبات وفروض الدراسة (بيانات متماثلة – بيانات متوسطة الالتواء – بيانات شديدة الالتواء) حيث تم استخدام اسلوب مونتوكارلو (Generated Monte Carlo techniques) باستخدام برنامج (SAS) لتوليد بيانات الدراسة للإجابة على مجموعة من التساؤلات التي تكتنف الدراسة الحالية وهي :

- ماذا سيحدث لو تغير توزيع العينتين (متماثل – متوسط الالتواء – شديد الالتواء).
- ماذا سيحد ث لو كان حجم العينتين متساوي وتغير حجم العينتين معا (صغير – متوسط – كبير).
- ماذا سيحد ث لو كان حجم العينتين مختلف وتغير حجم العينتين معا (صغير – متوسط – كبير).

وكانت أداة المقارنة المستخدمة هي احتمال الوقوع في خطأ النوع الاول (Type I Error) ، فكلما قلت احتمالية الوقوع في خطأ النوع الأول (α) كلما زادت متانة (Robustness) وحساسية (Sensitivity) الاختبار الاحصائي في قبول فرضية العدم وهي صحيحة وهو قرار سليم .

3.الدراسات السابقة (Previous Studies):

هناك الكثير من الدراسات السابقة التي اهتمت بدراسة متانة الاختبارات الاحصائية وهذه الجزئية تعتبر مجال بحث خصب الى يومنا هذا ، سنذكر منها :

- من خلال الدراسات والبحوث المقدمة من [8]،[14]،[15]،[21]،[23]،[25]،[26]،[28]،[29] اتضح من نتائجها أن اختبار(t) عند تساوي التباين يتمتع بالقوة والمتانة حتي إذا انحراف قليلا عن الافتراضات الواجب تحققها عند استخدامه في تحليل العينات.
- من الدراسات الاحصائية التي لها علاقة بموضوع الدراسة الابحاث المقدمة من [9]،[19] والتي تبين من خلال نتائجها أن

الاختبار اللامعلمي (Mann Whitny- Wilcoxon) يتمتع بالقوة والمتانة عند استخدامه مع بيانات لاتتبع التوزيع الطبيعي مثل التوزيعات الاسية واللوغاريتمية وتوزيعات البيانات التي تحتوي على قيم شاذة.

وكذلك أشار كلا من [12]،[18]،[28]،[33] في حالة عدم تساوي كلا من التباين وحجوم العينات، أو يكون التباين غير متساوي مع توزيعات لاتتبع التوزيع الطبيعي وباستخدام التباين المدمج (variances are combined) ، فأن معدل الخطأ من النوع الاول (Type I error rate) لاختباري (t) و(MWW) ينحرف عن القيمة النظرية (0.05) وفي مثل هذه الحالات ينصح [11]،[24]،[32] بإجراء التباين المفكك (separate-variance) واستخدام اختبار(Welch test) الذي يعطي معدل خطأ من النوع الاول أقل انحرافا عن القيمة النظرية لاختباري (t) و(MWW) .

- من خلال مراجعة الدراسات السابقة تبين أن متابعة ودراسة تأثير حجم العينة على أداء اختباري (t) و (MWW) أمر صعب ومعقد ، حيث علق [9] أنه في بعض الحالات كانت النتائج بالنسبة لأحجام العينات الصغيرة (أقل من 10) عكس الخاصة بأحجام العينات المعتدلة. كما أن البحث المقدم من [17] أشار فيه إلى العديد من الكتب والمقالات التي تساند الرأي القائل "أن الاختبارات اللامعلمية مفضلة عندما يكون حجم العينة صغيرًا وأن اختبار(t) يصبح متفوقًا على اختبار (MWW) عند زيادة حجم العينة نتيجة لنظرية النهاية المركزية (Central Limit Theorem)".
- ومع ذلك ، أظهرت دراساتي المحاكاة التي أجراها [20]،[22] أن (MWW) تحقق قوة اختبار متزايدة مع زيادة حجم العينة مما يشير إلى أنه ينبغي استخدامها ليس فقط للأحجام الصغيرة ولكن أيضًا لأحجام العينات الكبيرة.
- وفي العديد من الابحاث والدراسات الاحصائية ومنها البحوث المنشورة من قبل [10]،[13]،[34] تبين أنه يوجد خلاف حول الفرضية التي يتم اختبارها باستخدام (MWW)، حيث غالبًا ما يتم تفسير (MWW) كاختبار لمتوسطات متساوية أو كاختبار لتوزيعات متساوية.

والتفسير الصحيح الذي قدمه كلا من [10]،[13]،[34]هو أن اختبار(MWW) يكون مطابق لإجراء اختبار(t) بعد تحديد رتب العينات، ولذلك نجد أن اختبار (t) يقيم الاختلافات في المتوسطات، في حين يقوم اختبار(MWW) بتقييم الاختلافات في رتب المتوسط.

- وفي دراسة مقدمة من [16]كان الهدف من هذه الدراسة هو معرفة ما إذا كان يجب استخدام اختبار (t) أو (MWW) عند مقارنة عينتين مستقلتين لبيانات مقياس ليكارت الخماسي (Five-Point Likert Items) ، حيث خلصت الدراسة بأن الاختبارين بشكل عام يتمتعان بقدرة مماثلة ، ولاتوجد فروق بينهما .

لاحظنا من خلال الدراسات السابقة أن جل البحوث المتعلقة بمتانة الاختبار وقوته (Robust and Power Test) تفترض عدم تجانس التباين ، وعدم استقلالية العينات ، لذلك تم اجراء الدراسة الحالية تحت فرض استقلال العينات وتجانس التباين .لدراسة المتانة لاختبار (Student's-t) واختبار(Mann Whitny- Wilcoxon) لعينتين مستقلتين.

4.منهجية الدراسة (Methodology of Study):

إن أي بحث علمي ينبغي أن يلتزم فيه الباحث بعدد من الخطوات الاساسية والتي تعرف بخطوات البحث العلمي ، وتبدأ هذه الخطوات بتحديد المشكلة وصياغتها بأسلوب مبسط ، ومن ثم مراجعة الدراسات المتعلقة بها وصياغة الأسئلة والفرضيات وإعداد

وتصميم البحث ، ومن ثم جمع البيانات وتحليلها احصائيا وتفسير النتائج والوصول للقرارات التي تحل تلك المشكلة.

ومن أهم الخطوات السالفة الذكر والتي يعتمد عليها الباحث في إتخاذ القرار النهائي هي خطوة اختيار الأسلوب الاحصائي المناسب لبيانات البحث وستيفاء كامل افتراضاته وشروطه ، وإغفال هذه الخطوة الاساسية قد يلغي دقة نتائج البحث ، وبالتالي تصبح القرارات الناتجة عنه مظللة وليست ذالت قيمة علمية وعملية يعتمد عليها.

وبما أن نواة الدراسة وجوهرها متمثل في المفاضلة بين الاختبارات المعلمية واللامعلمية المستخدمة حول متوسط عينتين مستقلتين ، فإن الاسلوب المعلمي المناسب هو اختبار (Student's-t) لعينتين مستقلتين والأسلوب اللامعلمي المناسب هو اختبار (Mann Whitny- Wilcoxon). لذلك الفقرة التالية سوف تركز على عرض مفصل للإختبارين من ناحية الشروط والافتراضات الواجب توفرها قبل استخدام احدهما .

4.1. اختبار t (t-test):

يرجع الفضل في اشتقاق توزيع (t) الى العالم الإيرلندي (W.S Gosset, 1908) حيث نشر بحث باسم مستعار هو (Student) ، لذلك عرف التوزيع (Student's t- distribution) ودالة كثافته الاحتمالية بدرجة حرية $(v-1)$ هي :

$$f(t)=\frac{1}{\sqrt{v}\beta\left(\frac{1}{2},\frac{1}{v}\right)}\left(1+\frac{t^2}{v}\right)^{\frac{(v+1)}{2}} \qquad ,-\infty<t<\infty$$

حيث (t) متغير عشوائي متصل ينتمي لمجموعة الاعداد الحقيقية . ويمكن تلخيص خصائص توزيع (t) في النقاط التالية :

– يوجد عدد لانهائي من توزيعات (t) تختلف درجة تدببها وتفرطحها بحسب قيمة درجة الحرية $(v=n-1)$.

– توزيع (t) متصل (Continuous) ، وبالتالي منحناه يكون ممهد (Smooth) ولذلك يمكن حساب الاحتمالات بإيجاد المساحات تحت هذا المنحنى .

– شكل منحنى التوزيع (t) يشبه الجرس المقلوب وهو متماثل حول المحور الرأسي (حول الصفر) لأن متوسطه دائما صفر .

– تحدد درجات الحرية (v) شكل التوزيع ، فكلما زادت درجات الحرية كلما نقص التباين وكانت قيم المتغير أكثر تجانس .

– كلما زادت درجات الحرية (v) كلما اقترب التباين من الواحد الصحيح ، وأقترب توزيع (t) من التوزيع الطبيعي المعياري .

يستخدم اختبار (t) لمقارنة متوسطي مجتمعين ، وهل هما متساويان أم لا ؟ وهل الفرق بينهما دال احصائيا أم غير ذلك ؟ إلا أن استخدام اختبار(t) يتفرع إلى ثلاث أنواع هي :

– اختبار حول المتوسط (المقارنة بين متوسط عينة ومتوسط مجتمع) .

– اختبار الفرق بين متوسطي عينتين مستقلتين (المقارنة بين متوسطي عينتين مستقلتين) .

– اختبار الفرق بين متوسطي عينتين مرتبطتين (المقارنة بين متوسطي عينتين مرتبطتين) .

وفي هذه الدراسة سوف نناقش اختبار الفرق بين متوسطي عينتين مستقلتين ، يستخدم هذا الاختبار لقياس الفرق المعنوي (Significant Difference) بين متوسطي عينتين مستقلتين ، وللاختبار لحالتين هما:

(افتراض أن تباين العينتين متساوي ، أو أن تباين العينتين غير متساوي) .ويشترط هذا الاختبار تحقق الفرضيات التالية :

4.1.1. اعتدالية التوزيع (Normality):

حيث يفترض الاختبار بأن تكون بيانات المتغير التابع (بيانات العينتين) تتبع التوزيع الطبيعي وهناك طرق عدة للتأكد من اعتدالية التوزيع منها :

- **حساب مقاييس النزعة المركزية**: الوسط (Mean) والوسيط (Median) والمنوال (Mode) فالتوزيع الطبيعي يكون فيه (المنوال ≈ الوسيط ≈ الوسط).
- **حساب مقاييس الالتواء (Skewness, α_3) والتفرطح (Kurtosis, α_4):** يحسب مقياسي الالتواء والتفرطح لبيانات العينة وتتم مقارنته بمقاييس الالتواء والتفرطح للتوزيع الطبيعي وهي على النحو التالي :

$$(-2 \le \alpha_3 \le 2) \quad \text{and} \quad (-3 \le \alpha_4 \le 3)$$

- **باستخدام اختبارات اعتدالية التوزيع** : وأهمها اختبار كولموجروف – سميرنوف (K.S. test) واختبار شابيرو ولك (Shapiro- Wilk) ، وكلا الاختبارين متوفر بأغلب البرامج الاحصائية .
- **بالتمثيل البياني لمتغيرات المتغير التابع عن طريق الاختبارات البيانية**: مثل اختبار(Probability - Plots) ، واختبار (QQ-norm) أو عن طريق المدرج التكراري (Histogram) .

وتجدر الاشارة هنا ، إلى أنه في الكثير من أبحاث الاحصائيين المختصين أمثال [1]، [3]،[14] يرون امكانية التغاضي عن اعتدالية التوزيع إذا كان حجم العينة اكبر من (30 مفردة). أي أنه عند استخدام عينات كبيرة يمكن التنازل عن فرض اعتدالية التوزيع ، وهذه إحدى مظاهر قوة ومتانة (Robustness) اختبار (t) وذلك يتفق مع نظرية النهاية المركزية أو نظرية تقارب التوزيعات. إلا أن نفس الاحصائيين السابقين يرون ضرورة اجراء اختبارات التحقق من اعتدالية التوزيع إذا كانت العينة صغيرة (اقل من 30 مفردة) ، فإذا كان المتغير التابع لايتبع التوزيع الطبيعي (توزيعه موجب أو سالب الالتواء) ، فيمكن اللجوء للاختبارات الاحصائية اللامعلمية (Non-Parametric Testing) البديلة والتي لاتشترط أي افتراضات حول توزيع المتغير وشكله.

4.1.2. تجانس التباين (Homogeneity of Variance):

ويقتضي هذا الشرط أن تباين المشاهدات في العينة الأولى (S_1^2) لايختلف عن تباين المشاهدات في العينة الثانية (S_2^2). ويقال في هذه الحالة أن العينتان متجانستان ويمكن اختبار تجانس التباين بعدة طرق منها :

- اختبار ليفن لتجانس التباين (Levene’s Test for Homogeneity of variance).
- اختبار القيمة العظمى لهارتلي (Hartley F_{max}).
- اختبار بارتليت (Bartlett’s Test for Homogeneity of variance).
- اختبار كوجران (Cochran Test).

وقد اتفق الكثير من الاحصائيين في دراساتهم أمثال [7]،[6]،[31]،[27] وغيرهم ، على أن افتراض تجانس التباين يمكن التغاضي عنه عند تساوي حجم العينتين ($n_1 = n_2$) .

4.1.3. الاستقلالية أو الارتباط (Independence or Correlation):

يجب على الباحث عند استخدام اختبار (t) لعينتين مستقلتين التأكيد على أهمية استقلالية مشاهدات العينتين بمعنى أنه لا يكون هناك أي ارتباط بين مشاهدات العينتين.

وذكر [1:ص373] " الاستقلال يعني ببساطة أن البيانات التي نجمعها سواء بين المجموعات أو داخل المجموعات ليست متزاوجة أو متكررة أو متداخلة أو معتمدة على بعضها البعض على أي نحو، ولا يتوافر ذلك إلا إذا كان اختيار العينات عشوائيا تماما ، أي تحكمه عوامل المصادفة من ناحية ، وكذلك أن يكون الباحث قد استخدم وسائل الضبط التجريبي من ناحية أخرى ، فإذا تزاوجت الدرجات على نحو أو أخر سواء كان ذلك عن طريق تكافؤ المجموعات أو تكرار الملاحظات على نفس الافراد فإن المجموعات حينئذ تكون مرتبطة وفي هذه الحالة لابد من استخدام اختبار (t) للمجموعات المرتبطة أو للقياسات المتكررة ".

ويجب التنويه بأننا في هذه الدراسة ينصب اهتمامنا على اختبار (t) لعينتين مستقلتين في حالة تجانس التباين ويستخدم هذا الاختبار إذا كان المجتمعان اللذان سحبت منهما العينتان يتوزعان توزيعا طبيعيا وتباينهما متساو $(\sigma_1^2 = \sigma_2^2)$ ولكن غير معلوم .

فإذا كان $(\bar{x}_1)$ متوسط العينة الأولى بحجم (n_1) سحبت من مجتمع طبيعي متوسطه (μ_1)، وكان $(\bar{x}_2)$ متوسط العينة الثانية بحجم (n_2) سحبت من مجتمع طبيعي متوسطه (μ_2)، فأن التباين المشترك (Pooled Variance) الذي يرمز له بالرمز (S_p^2) يمكن الحصول عليه بالعلاقة :

$$S_p^2 = \frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}$$

حيث S_1^2 تباين العينة الأولى و S_2^2 تباين العينة الثانية. والاحصاءة المناسبة لإجراء الاختبار هي :

$$t = \frac{(\bar{x}_1-\bar{x}_2)-(\mu_1-\mu_2)}{\sqrt{S_p^2\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}}$$

عند مستوى معنوية (α) يمكننا اختبار الفرق بين متوسطي مجتمعين من خلال مقارنة قيمة احصاءة الاختبار بالقيمة الجدولية التي يمكن الحصول عليها من جدول (t) بعد تحديد قيمة مستوى المعنوي (α) ودرجة الحرية $(n_1 + n_2 - 2)$ ، وتصاغ فرضية الاختبار على إحدى الصور التالية :

- إذا كانت فرضية العدم $H_0: \mu_1 \leq \mu_2$

ضد الفرضية البديلة $H_1: \mu_1 > \mu_2$

فإننا نرفض فرض العدم إذا كان :

$$t = \frac{(\bar{x}_1-\bar{x}_2)-(\mu_1-\mu_2)}{\sqrt{S_p^2\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}} > t_{1-\alpha(n_1+n_2-2)}$$

- إذا كانت فرضية العدم $H_0: \mu_1 \geq \mu_2$

ضد الفرضية البديلة $H_1: \mu_1 < \mu_2$

فإننا نرفض فرض العدم إذا كان :

$$t = \frac{(\bar{x}_1-\bar{x}_2)-(\mu_1-\mu_2)}{\sqrt{S_p^2\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}} < -t_{1-\alpha(n_1+n_2-2)}$$

- إذا كانت فرضية العدم $H_0: \mu_1 = \mu_2$

ضد الفرضية البديلة $H_1: \mu_1 \neq \mu_2$

فإننا نرفض فرض العدم إذا كان:
$$t = \left| \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{S_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \right| < t_{1-\alpha/2\,(n_1+n_2-2)}$$

4.2. اختبار مان وتني ويلكوكسن (Mann Whitney Whilcoxon Test) :

اختبار مان وتني – ويلكوكسون (Mann Whitney Whilcoxon Test) ، ويطلق عليه ايضا اختبار يلكوكسن لمجموع الرتب (Wilcoxon rank sum test) هو اكثر الاختبارات اللامعلمية انتشارا واستخداما ، ويستخدم الاختبار للمقارنة بين عينتين مستقلتين لمعرفة مإذا كان العينتين تنتميان لمجتمع واحد أو مجتمعين متماثلين. ويعتبر أول اختبار يتعامل مع عدم التساوي بالنسبة للعينات. ويتطلب اختبار (MWW) الآتي:

- أن تكون بيانات العينتين المستقلتين مأخوذة عشوائيا.
- أن تكون بيانات العينتين رتبيه على الأقل.
- أن يكون المتغير محل البحث متصلا .
- أن تختلف دوال توزيع مجتمعي العينتين فالمركز فقط.

- **وتعد فرضية الاختبار على النحو التالي :**

فرض العدم (H_0): مجتمعي العينتين متماثلين ، أي أن العينتين مسحوبتان من مجتمعين لهما مركز متساوي.

الفرض البديل (H_1): قد يكون الفرض البديل لفرض العدم ذو اتجاهين ، أي أن العينتين تنتميان لمجتمعين مختلفين فالمركز ، أو قد يكون الفرض البديل لفرض العدم ذو اتجاه واحد موجب ، أي أن وسط مجتمع إحدى العينتين أكبر من وسط مجتمع العينة الاخرى ، أو قد يكون الفرض البديل لفرض العدم ذو اتجاه واحد سالب ، أي أن وسط مجتمع إحدى العينتين أصغر من وسط مجتمع العينة الاخرى.

- **ويجرى الاختبار على النحو التالي :**

يأخذ الاختبار ثلاثة اتجاهات حسب حجم العينتين المستقلتين ، فإذا كان حجم بيانات العينة الاولى (n_1)وحجم بيانات العينة الثانية (n_2)بحيث يساوي مجموعهما (n) فإن الاتجاهات الثلاثة للاختبار هي :

- في حالة العينات صغيرة جدا (عندما تكون n_1 و n_2أقل من أو تساوي 8) .
- في حالة العينات صغيرة (عندما تكون n_1 وn_2 أكبر من 8 وأصغر من أو تساوي 20).
- في حالة العينات الكبير (عندما تكون n_1 وn_2 أكبر من 20).

4.2.1. في حالة العينات الصغيرة جدا :

بعد دمج العينتين وترتيبهما من الاصغر فالأكبر يتم حساب قيمة (U) وهي تساوي مجموع قيم العينة الاكبر (إذا كانت العينتان غير متساويتان فالحجم) والتي اصغر من كل قيمة من قيم العينة الاصغر حجما. فإذا كانت بيانات العينة الاصغر هي (X) وبيانات العينة الاكبر حجما هي (Y) فإن:$\sum_{\forall X} n(Y < X)$ يعطي عدد قيم (Y) التي اقل من كل قيمة من قيم (X) وكذلك الحال فإن:$\sum_{\forall Y} n(X < Y)$ يعطي عدد قيم (X) التي اقل من كل قيمة من قيم (Y). وعليه فإن قيمة (U) تساوي:

$$U = min[\sum_{\forall X} n(Y < X), \sum_{\forall Y} n(X < Y)]$$

أي أقل العددين من بين } عدد قيم (Y) التي اقل من كل قيمة من قيم (X) وعدد قيم (X) التي اقل من كل قيمة من قيم (Y) {.وبمعلومية (U، n_1، n_2) فإنه يمكن إيجاد قيمة مستوى المعنوية المصاحب (α^*) من جدول مان وتني للعينات الصغيرة جدا. وتقابل القيمة المعنوية المصاحبة الاختبار ذو الاتجاه الواحد ، وعليه تضاعف القيمة لتقابل الاختبار ذو الاتجاهين. ويمكن بطريقة ثانية استخراج القيمة الحرجة الدنيا (U^-) من جدول مان وتني المعد للعينات الصغيرة جدا وذلك بمعلومية (n_2، n_1) ومستوى المعنوية (α). ومن ثم يمكن الحصول على القيمة الحرجة العليا (U^+) بالعلاقة التالية :

$$U^+ = n_1 n_2 - U^-$$

- **المقارنة والقرار:**

إذا تم حساب مستوى المعنوية المصاحب (α^*) فإنه يقارن مباشرة بمستوى المعنوية (α) فإذا كانت ($\alpha^* < \alpha$) يرفض فرض العدم ويقبل الفرض البديل .

أما بواسطة الطريقة الثانية ، وفي حالة الاختبار ذي اتجاهين ، فإنه يرفض فرض العدم ويقبل الفرض البديل إذا كانت القيمة الحسابية ل (U) أكبر من (U^+) وأقل من (U^-) ، وفي حالة الاختبار ذي الاتجاه الواحد الموجب يرفض فرض العدم ويقبل الفرض البديل إذا كانت القيمة الحسابية ل (U) أكبر من (U^+) ، أما في حالة الاختبار ذي الاتجاه الواحد السالب يرفض فرض العدم ويقبل الفرض البديل إذا كانت القيمة الحسابية ل (U) أصغر من (U^-).

4.2.2. في حالة العينات الصغيرة :

في هذه الحالة تدمج العينتان في مجموعة واحدة وترتب البيانات من الاصغر فالأكبر ، كما في حال العينات الصغيرة جدا ، يعطى بعد ذلك الترتيب المقابل لكل قيمة من قيم العينتين ، بحيث تأخذ أصغر قيمة الرتبة (1) والتي تليها الرتبة (2) وهكذا . وفي حالة تساوي قيمتين أو أكثر فإن الترتيب الذى يعطى هو متوسط الرتب المتتالية .وللتميز بين قيم (ومن ثم رتب) كل عينة فإنه يمكن الرمز لقيم إحدى العينتين بالرمز (X) وقيم العينة الاخرى بالرمز (Y) ومن ثم يكون ($\sum_X r_x$) و ($\sum_Y r_y$) هو مجموع رتب القيم لكل عينة . وباستخدام مجموع الرتب للعينتين وباعتبار (n_1) حجم العينة (X) و (n_2) حجم العينة (Y) فإنه يمكن حساب قيمة (U) باستخدام الصيغة التالية :

$$U = min[(\sum_X r_x - n_1(n_1+1)/2), (\sum_Y r_Y - n_2(n_2+1)/2)]$$

وبمعلومية (n_1 و n_2) ومستوى المعنوية (α) فإنه يمكن استخراج القيمة الحرجة الدنيا (U^-) من جدول مان وتني المعد للعينات الصغيرة ومن ثم يمكن الحصول على القيمة الحرجة العليا (U^+) بالعلاقة السابقة :

$$U^+ = n_1 n_2 - U^-$$

- **المقارنة والقرار:**

في حالة الاختبار ذي اتجاهين ، فإنه يرفض فرض العدم ويقبل الفرض البديل إذا كانت القيمة الحسابية ل (U) أكبر من (U^+)

وأقل من (U^-) ، وفي حالة الاختبار ذي الاتجاه الواحد الموجب يرفض فرض العدم ويقبل الفرض البديل إذا كانت القيمة الحسابية ل (U) أكبر من (U^+) ، أما في حالة الاختبار ذي الاتجاه الواحد السالب يرفض فرض العدم ويقبل الفرض البديل إذا كانت القيمة الحسابية ل (U) أصغر من (U^-). ويمكن فالمقابل حساب مستوى المعنوية المصاحب (α^*) ومقارنته مباشرة بمستوى المعنوية (α) فإذا كانت $(\alpha^* < \alpha)$ يرفض فرض العدم ويقبل الفرض البديل .

4.2.3. في حالة العينات الكبيرة :

في هذه الحالة يتم إتباع ذات الخطوات التي تستخدم في حالة العينات الصغيرة والتي تؤدي على حساب قيمة (U) ، وبمعلومية $(n_1$ و $n_2)$ فإنه يمكن استخدام التوزيع الطبيعي حيث تكون القيمة المتوقعة لقيمة الاختبار (U) تساوي:

$$E(U) = (n_1 n_2)/2$$

أما تباين (U) يساوي:

$$V(U) = (n_1 n_2)(n_1 n_2 + 1)/12$$

وعليه تكون قيمة (Z) تساوي:

$$Z = \frac{U - E(U)}{\sqrt{V(U)}}$$

قد تتساوى اكثر من قيمة في الرتبة الواحدة ، ويمكن أن يحدث ذلك في ذات العينة أو على نطاق العينتين . فالتساوي في القيم أو الرتب داخل العينة الواحدة لا يؤثر في قيمة الاختبار ، ولكن التساوي في القيم أو الرتب على نطاق العينتين هو الذي يؤثر في قيمة الاختبار ، وفي حالة التساوي بالعينات الكبيرة يمكن حساب معامل التصحيح (Correction Function) لقيمة الاختبار حيث يحسب معامل التصحيح (CF) على النحو التالي :

$$CF = \frac{n_1 n_2 [\sum(\tau^3 - \tau)]}{12(n_1 + n_2)(n_1 + n_2 - 1)}$$

حيث (τ) هي عدد حالات التساوي لكل رتبة من الرتب. وتؤثر قيمة (CF) في تباين (U) وعليه تكون قيمة (Z) المعدلة هي:

$$Z = \frac{U - (n_1 n_2)/2}{\sqrt{[(n_1 n_2)(n_1 n_2 + 1)/12] - CF}}$$

المقارنة والقرار:

نقارن قيمة(Z)الحسابية بالقيمة الجدولية للتوزيع الطبيعي المعياري بمستوى معنوية (α). فإذا وقعت قيمة (Z) الحسابية خارج منطقة القبول يرفض فرض العدم ويقبل الفرض البديل . ويمكن إجراء المقارنة عن طريق حساب القيمة المعنوية المصاحبة ل (Z) ومقارنتها بالقيمة المعنوية (α) فإذا كانت القيمة المعنوية المصاحبة (α^*) أقل من القيمة المعنوية (α) يرفض فرض العدم ويقبل الفرض البديل [3].

4.3.المحاكاة (The Simulation) :

هناك الكثير من المشاكل والنظريات الاحصائية التي يصعب تحليلها تحليلا منطقيا بواسطة البرهان الرياضي الأمر الذي يدفع البعض نحو ترجمة هذه النظريات على مجتمعات حقيقية ثم سحب عدد من العينات العشوائية المستقلة منها ليتم الوصول

للحلول المثلى لتلك المشكلات والنظريات. وتعرف المحاكاة بصورة عامة على أنها عبارة عن الحلول للمشكلات والنظريات لذا توجه الباحثين عند مطلع القرن الحالي نحو تطبيق أسلوب المعاينة التجريبية (The Sampling Experimental) وهو ما يعرف اليوم بالمحاكاة (Simulation) من خلال بناء نموذج (Model) مشابه للمشكلة الاصلية ومن ثم تطبيق المعاينة عليه . لقد تم تطبيق هذا الأسلوب بشكل واسع في مجالات الاحصاء المختلفة وذلك من خلال دراسة وتطوير الطرق الاحصائية المختلفة بالمقارنة بينها . إن استخدام أسلوب المحاكاة كأسلوب للتحليل يتم عند استحالة أو صعوبة استخدام اساليب التحليل الاخرى ، ويستخدم أحيانا لتوفير الوقت والجهد والتكاليف[5].

كما وضحنا سابقا أن افتراضات اختبار (t) لها أهمية خاصة وفي حال انتهاك هذه الافتراضات سوف يؤدي الى نتائج غير دقيقة لذلك غالبا ما يلجأ الباحث إلى الطرق اللامعلمية. وفي هذه الدراسة ستتم المقارنة بين الاختبارات المعلمية والاختبارات اللامعلمية في ظل انتهاك شرط (اعتدالية توزيع البيانات) واختيار أفضل اختبار من بين الاختبارين المذكورين سابقا ، ولأن البرهان الرياضي عاجز عن تحقيق هذا الهدف فقد تم اعتماد أسلوب المحاكاة (Simulation) بديلا مناسبا لإجراء المقارنة وذلك من خلال محاكاة الواقع الفعلي بتوليد عدد كبير من العينات العشوائية المستقلة لمجتمعات مختلفة الخصائص ولعدد من الحالات المختلفة التي من شأنها ن تعزز دور المحاكاة في المقارنة وتحقق الهدف المرجو من تطبيقها.

4.4.خطة الدراسة (Plan of Study) :

لأجراء هذه الدراسة سوف يتم الاعتماد علي أسلوب المحاكاة (Simulation) ويقوم الباحث بتوليد البيانات باستخدام البرنامج الاحصائي (SAS) تحت ظل ظروف مرغوب فيها لغرض الدراسة ، وعليه فإن بيانات الدراسة متمثلة في ثلاث مجتمعات بحجم كبير (100000 مشاهدة) ، المجتمع الاول يتبع التوزيع الطبيعي المعياري ، أما المجتمع الثاني له التواء وتفرطح بشكل متوسط ، أما المجتمع الاخير له التواء وتفرطح بشكل أكثر حده من المجتمع الثاني.

بعد توليد البيانات سنجري عملية معاينة من المجتمعات الثلاثة بأحجام مختلفة بعدد (1000عينة) لكل حالة من حالات الدراسة ،ويتم تطبيق كلا من الاختبار المعلمي(t) لعينتين المستقلتين وكذلك الاختبار اللامعلمي(MWW) علي هذه العينات. وبعد الحصول على نتيجتي الاختبارين تتم المقارنة بين قيمة احتمال رفض الفرضية الصفرية الصحيحة لكل اختبار في كل حالة من الحالات ، والذي يمثل الخطأ من النوع الاول ، والاختبار الذي يقرب قيمة الخطأ من النوع الاول المحسوبة من المحاكاة الي قيمة الخطأ من النوع الاول النظرية (0.05) يعتبر هو الاختبار الافضل من ناحية المتانة (Robustness).

5.النتائج والمناقشة (Results and Discussion) :

في هذا الجزء من البحث سنعرض النتائج ومناقشة التي توصلت لها الدراسة الحالية ، حيث سيتم عرض النتائج المتعلقة بكل مجتمع من مجتمعات الدراسة الثلاث على حده ، وهي على النحو التالي :

– نتائج المعاينة من مجتمع يتبع التوزيع الطبيعي المعياري(المجتمع الأول) .
– نتائج المعاينة من توزيع ملتو ومتفرطح بشكل متوسط (المجتمع الثاني) .
– نتائج المعاينة من توزيع ملتو ومتفرطح بشكل حاد (المجتمع الثالث) .

5.1 نتائج المعاينة من مجتمع يتبع التوزيع الطبيعي المعياري (المجتمع الأول):

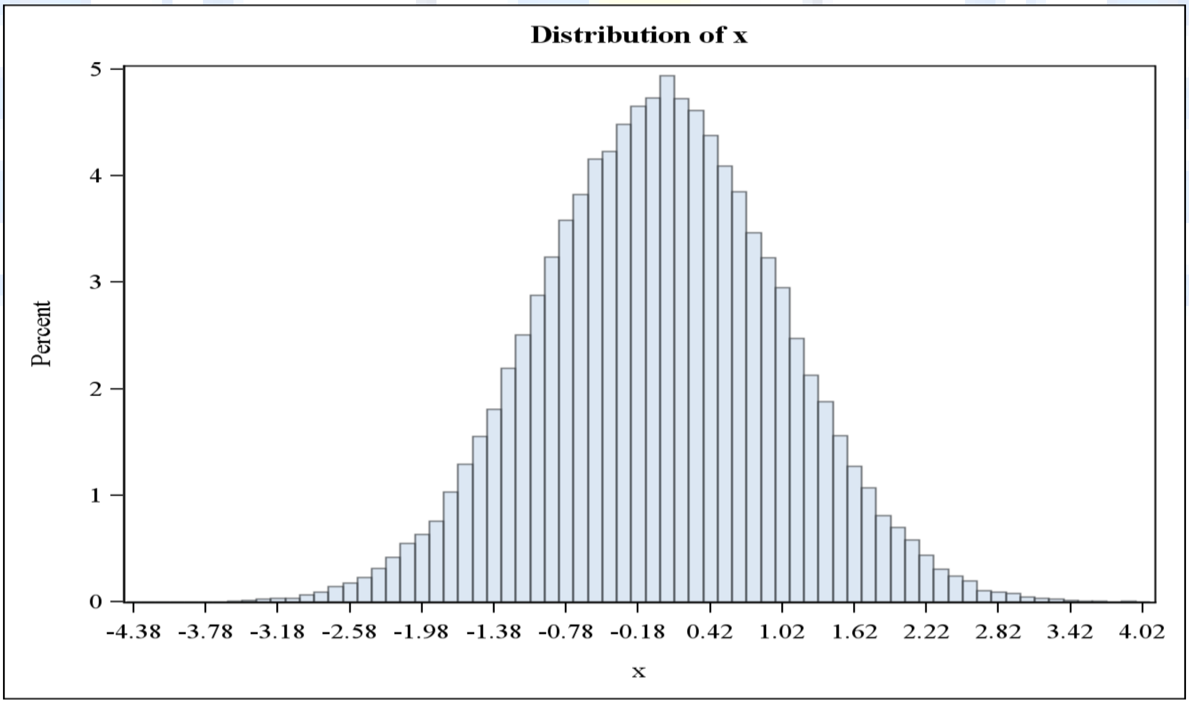
الجدول التالي يعطي الإحصاء الوصفي لمعالم هذا المجتمع وكما نري من بيانات الجدول أن الوسط حسابي قيمته (صفر) والتباين قيمته (واحد تقريبا) ، ومعاملي الالتواء و التفرطح قيمتيهما تقريبا (صفر). كذلك من التمثيل البياني للبيانات الذي يلي هذا الجدول باستخدام (Histogram) يتضح أن هذا التوزيع هو توزيع طبيعي معياري.

جدول (1): الاحصاء الوصفي لمشاهدات مجتمع يتبع التوزيع الطبيعي المعياري (المجتمع الأول).

Moments			
N	100000	**Sum Weights**	100000
Mean	0.00483225	**Sum Observations**	483.225487
Std Deviation	1.00168564	**Variance**	1.00337411
Skewness	0.00469075	**Kurtosis**	0.00128153
Uncorrected SS	100338.743	**Corrected SS**	100336.408
Coeff Variation	20729.1557	**Std Error Mean**	0.00316761

المصدر : الباحث باستخدام برنامج (SAS)

شكل (1): يوضح شكل توزيع مشاهدات مجتمع يتبع التوزيع الطبيعي المعياري (المجتمع الأول)



المصدر : الباحث باستخدام برنامج (SAS)

والجدول التالي يمثل معدل الخطأ من النوع الاول(Type I Error Rate) للاختبارين قيد الدراسة ، الذي تم حسابه من المحاكاة في حالة المجتمع يتبع التوزيع الطبيعي المعياري لعدة حالات مختلفة من أحجام العينتين المستقلتين . يتضح من الجدول أدناه أن احتمالية الوقوع في الخطأ من النوع الاول لا تختلف بشكل كبير بين الاختبارين وكلاهما يقترب من القيمة النظرية (0.05) كلما كان حجم العينات كبير ولا يوجد اختلاف كبير بين حجميهما ومع ذلك نلاحظ ان القيم المحسوبة لتوزيع (t) هي الاقرب و الافضل في معظم الحالات.

جدول (2): معدلات الخطأ من النوع الاول للاختبارين في حالة المعاينة من (المجتمع الأول)

n_1	n_2	Mann Whitney Whilcoxon Test	Student's t- Test
6	6	0.03900	0.03400
10	10	0.04704	0.05000
30	30	0.04800	0.04909
50	50	0.04790	0.04890
6	10	0.04500	0.05000
6	50	0.05300	0.05100
10	50	0.04800	0.05300
30	50	0.05200	0.05100

المصدر : الباحث باستخدام برنامج (SAS)

5.2. نتائج المعاينة من توزيع ملتو ومتفرطح بشكل متوسط (المجتمع الثاني):

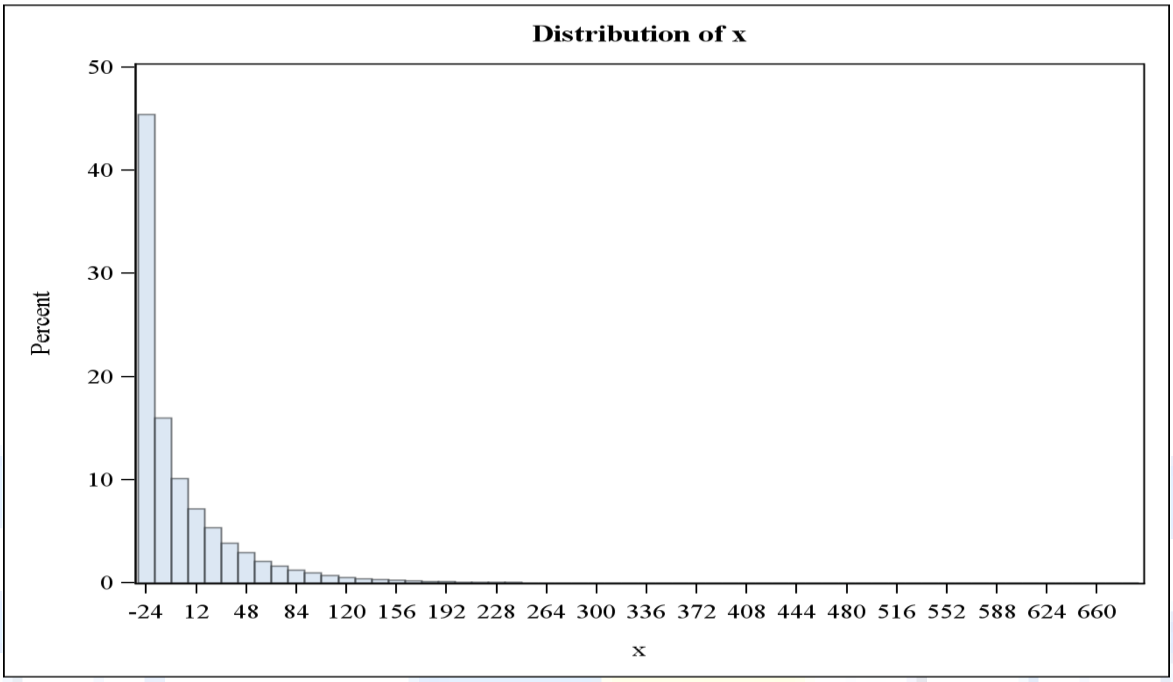
الجدول التالي يعطي الإحصاء الوصفي لمعالم هذا المجتمع و كما نري من بيانات الجدول أن الوسط حسابي قيمته (0.12) والتباين قيمته (1586.68) ، وقيمة معاملي الالتواء و التفرطح علي التوالي هي (2.617) و (10.820). كذلك من التمثيل البياني الذي يلي هذا الجدول يتضح ان هذا التوزيع هو توزيع ملتو ومتفرطح بشكل متوسط .

جدول (3): الاحصاء الوصفي لمشاهدات مجتمع ملتو ومفرطح بشكل متوسط (المجتمع الثاني)

Moments			
N	100000	**Sum Weights**	100000
Mean	0.1202911	**Sum Observations**	12029.1097
Std Deviation	39.8332083	**Variance**	1586.68448
Skewness	2.61773696	**Kurtosis**	10.8209089
Uncorrected SS	158668309	**Corrected SS**	158666862
Coeff Variation	33114.0119	**Std Error Mean**	0.12596366

المصدر : الباحث باستخدام برنامج (SAS).

شكل (2): يوضح شكل توزيع مشاهدات مجتمع ملتو ومفرطح بشكل متوسط (المجتمع الثاني)



المصدر : الباحث باستخدام برنامج (SAS).

من الجدول أدناه والذي يمثل معدلات الخطأ من النوع الاول المحسوب للاختبارين ، يتضح ان القيم المحسوبة من اختبار (MWW) افضل من تلك المحسوبة من اختبار(t) والذي اتضح انه يتأثر عندما تكون المعاينة من توزيع غير طبيعي وخاصة اذا كان هناك اختلاف كبير بين حجمي العينتين المستقلتين ، مع ملاحظة انه لا يتأثر بشكل كبير اذا كان حجمي العينتين كبير ومتقارب في الحجم .

جدول (4): معدلات الخطأ من النوع الاول للاختبارين في حالة المعاينة من (المجتمع الثاني)

n_1	n_2	Mann Whitney Whilcoxon Test	Student's t-Test
6	6	0.04200	0.01800
10	10	0.04800	0.03700
30	30	0.05200	0.04700
50	50	0.04700	0.04670
6	10	0.04870	0.03800
6	50	0.04900	0.14500
10	50	0.05400	0.10100
30	50	0.05200	0.05600

المصدر : الباحث باستخدام برنامج (SAS).

5.3.المعاينة من توزيع ملتو ومتفرطح بشكل حاد (المجتمع الثالث):

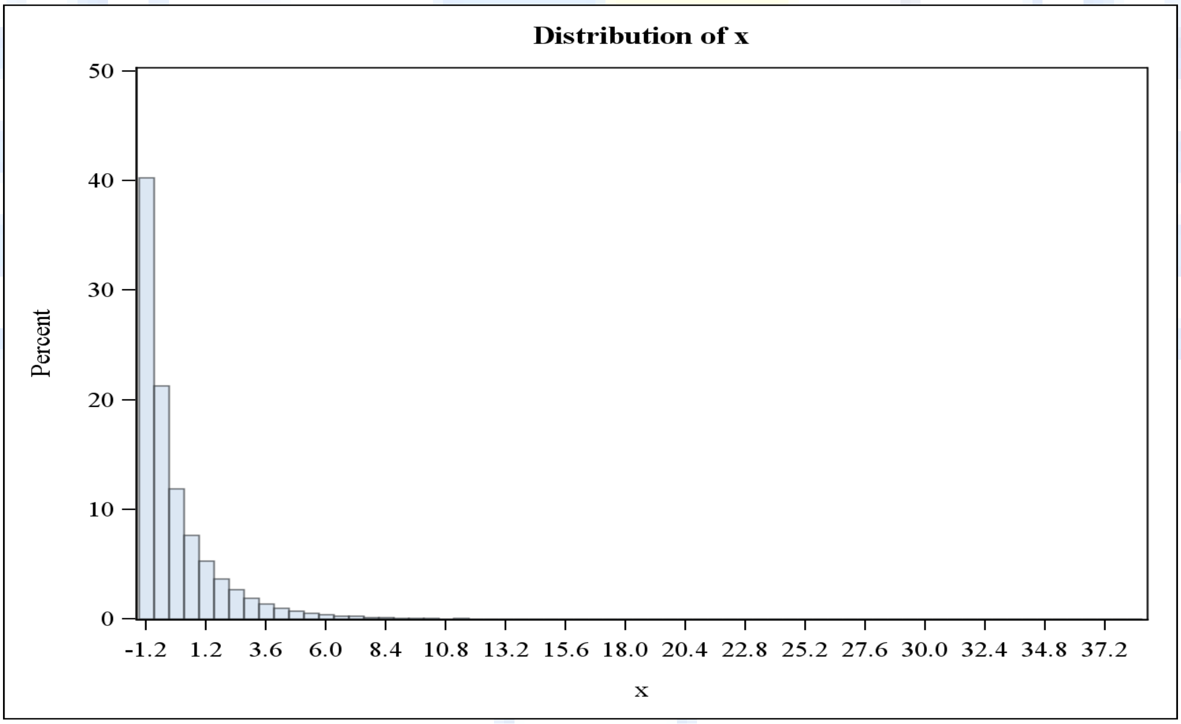
الجدول التالي يعطي الإحصاء الوصفي لمعالم هذا المجتمع الثالث في الدراسة ، و كما نري من بيانات الجدول ان قيمة الوسط حسابي (-0.004) وقيمة التباين (3.198) ، وقيمة معاملي الالتواء و التفرطح هي (3.253) و (19.240)علي التوالي . كذلك من التمثيل البياني الذي يلي هذا الجدول يتضح ان هذا التوزيع هو توزيع ملتو ومتفرطح بشكل كبير.

جدول (5): الاحصاء الوصفي لمشاهدات مجتمع ملتو ومفرطح بشكل حاد (المجتمع الثالث)

Moments			
N	100000	**Sum Weights**	100000
Mean	-0.0040767	**Sum Observations**	-407.66901
Std Deviation	1.78846922	**Variance**	3.19862216
Skewness	3.25358606	**Kurtosis**	19.2404625
Uncorrected SS	319860.679	**Corrected SS**	319859.017
Coeff Variation	-43870.62	**Std Error Mean**	0.00565564

المصدر : الباحث باستخدام برنامج (SAS).

شكل (3): يوضح شكل توزيع مشاهدات مجتمع ملتو ومفرطح بشكل كبير (المجتمع الثالث)



المصدر : الباحث باستخدام برنامج (SAS).

من الجدول الاتي والذي يمثل معدلات الخطأ من النوع الاول المحسوب للاختبارين ، يتضح كما في الحالة السابقة ان القيم المحسوبة من الاختبار اللامعلمي(Mann Whitney Whilcoxon Test) افضل من تلك التي تم حسابها من الاختبار المعلمي

(Student's t-Test) والذي اتضح انه يتأثر بشكل كبير عندما تكون المعاينة من توزيع غير طبيعي وخاصة اذا كان هناك اختلاف كبير بين حجمي العينتين المستقلتين ،مع ملاحظة انه لا يتأثر بشكل كبير اذا كانت حجمي العينتين المستقلتين كبير ومتساوي .

جدول (6): معدلات الخطأ من النوع الاول للاختبارين في حالة المعاينة من (المجتمع الثالث)

n_1	n_2	Mann Whitney Whilcoxon Test	Student's t-Test
6	6	0.04000	0.01100
10	10	0.04100	0.02700
30	30	0.05100	0.05200
50	50	0.04600	0.04870
6	10	0.04770	0.03500
6	50	0.04890	0.12200
10	50	0.05094	0.10589
30	50	0.04900	0.05300

المصدر : الباحث باستخدام برنامج (SAS).

6.الخلاصة (Summary) :

من خلال هذه الدراسة اتضح انه اذا كان حجمي العينتين المستقلتين صغير جدا فإن كلا الاختبارين المعلمي واللامعلمي يعطي خطأ من النوع الاول بعيد عن الخطأ النظري ، أي أن كلاهما يتأثر في حالة العينات الصغيرة ، وبغض النظر عن خصائص مجتمع المعاينة. وكذلك تبين من الدراسة أن النتائج التي أعطاها الاختبار المعلمي (Student's t-Test) أفضل وأقرب لقيمة الخطأ النظري (0.05) من النتائج التي أعطاها الاختبار اللامعلمي(MWW) في حالة المعاينة من التوزيع الطبيعي المعياري ، وأنهما لايتأثرا باختلاف حجمي العينتين في هذه الحالة (المعاينة من توزيع طبيعي معياري). أما في حالة المعاينة من مجتمعات بعيدة عن التوزيع الطبيعي فأن اختبار(t) يتأثر بشكل كبير باختلاف حجمي العينتين المستقلتين ، وكذلك لاحظنا بأنه يعطي نتائج جيده في حالة قرب وكبر حجمي العينتين المستقلتين من بعضهما ولا يوجد اختلاف جوهري بينه و بين الاختبار اللامعلمي (MWW) .

7.التوصيات (Recommendations):

من خلال الاستنتاجات التي أشرنا إليها هناك عدد من التوصيات التي يمكن أن تؤخذ بعين الاعتبار وتشمل ما يلي :

- ضرورة نشر الوعي لدى الباحثين من طلاب دراسات عليا وأعضاء هيئة تدريس والقائمين على تحكيم البحوث المنشورة بالمجلات العلمية بأهمية التحقق من الوفاء بافتراضات الاختبارات المعلمية قبل الشروع في استخدامها ومنها اختبار (t).
- تطبيق أسلوب المعاينة التجريبية (The Sampling Experimental) والذي يعرف اليوم بالمحاكاة (Simulation) لدراسة النظريات الاحصائية التي يصعب تحليلها تحليلا منطقيا بواسطة البرهان الرياضي .
- إجراء دراسة مماثلة للمقارنة الأساليب الاحصائية المعلمية والأساليب الاحصائية اللامعلمية من خلال دراسة الاختبار الافضل من ناحية المتانة أو القوة (Robustness or Power).

المراجع (References) :

[1]: أبوحطب،فؤاد،وصادق،امال(1996).مناهج البحث وطرق التحليل الاحصائي في العلوم النفسية والتربوية والاجتماعية .القاهرة:مكتبة الأنجلو المصرية.

[2]: أدم،أمين ابراهيم(2005). المبادئ الأساسية الاحصائية في الطرق التطبيقية اللامعلمية. الرياض: مكتبة الملك فهد الوطنية.

[3]: الشربيني،زكريا(1995). الاحصاء وتصميم التجارب في البحوث النفسية والتربوية والاجتماعية. القاهرة: مكتبة الأنجلو المصرية.

[4]: الصياد،جلال مصطفى وحبيب،محمد الدسوقي(1997). مقدمة في الطرق الاحصائية.الطبعة الرابعة،جدة، دار عكاظ للطباعة والنشر.

[5]: رشيد،ظافر حسين،وحسين،سجى محمد(2007). مقارنة بعض الطرائق المعلمية واللامعلمية لتصميم القطاعات العشوائية للقياسات المكررة. "مجلة العلوم الاقتصادية والإدارية" المجلد(13) العدد (48).

[6]: علام،صلاح الدين محمود(1993). الاساليب الاحصائية الاستدلالية البارامترية واللابارامترية في تحليل بيانات البحوث النفسية والتربوية. القاهرة:دار الفكر العربي.

[7]: عودة،أحمد سليمان،الخليلي،خليل يوسف(1988). الاحصاء للباحث في التربية والعلوم الانسانية. عمان: دار الفكر.

[8]: Baker, B. O., Hardyck, C. D., & Petrinovich, L. F. (1966). Weak measurements vs. strong statistics: An empirical critique of S. S. Stevens' proscriptions on statistics. Educational and Psychological Measurement, XXVI, 291–309.

[9]: Blair, R. C., Higgins, J. J., & Smitley, W. D. S. On the relative power of the U and t-tests. British Journal of Mathematical and Statistical Psychology, 1980, 33(1), 114-120.

[10]:Conover WJ,ImanRL (1981):Rank transformation as a bridge between parametric and nonparametric statistics.AmericanStatistician35, 124–129

[11]: Cribbie, R. A., & Keselman, H. J. (2003). The effects of nonnormality on parametric, nonparametric, and model comparison approaches to pairwise comparisons. Educational and Psychological Measurement, 63, 615–635.

[12]: Fagerland MW,Sandvik L(2009): The Wilcoxon-Mann-Whitney test under scrutiny. Statistics in Medicine28,1487–1497.

[13]: Fagerland MW, Sandvik L(2009):Performance of five two-sample location tests for skewed distributions with unequal variances. ContemporaryClinicalTrials30,490–496.

[14]: Glass, G. V., Peckham, P. D., & Sanders, J. R. (1972). Consequences of failure to meet assumptions underlying the fixed effects analysis of variance and covariance. Review of Educational Research, 42, 237-288

[15]: Heeren, T., & D'Agostino, R. (1987). Robustness of the two independent samples t-test when applied to ordinal scaled data. Statistics in Medicine, 6, 79–90.

[16]: Joost C. F. de Winter and Dimitra Dodou. 2010 .Five-Point Likert Items: t test versus Mann-Whitney-Wilcoxon. Practical Assessment ،Research & Evaluation 15, 11 (October 2010), 1–12.

[17]: Lumley, T., Diehr, P., Emerson, S., & Chen, L. (2002). The importance of the normality assumption in large public health data sets. Annual Review of Public Health, 23, 151–169.

[18]: MacDonald, R. (1999), "Asset Market and Balance of Payments Characteristics: An Eclectic Exchange Rate Model for the Dollar, Mark, and Yen", Open Economies Review 10, 1, 5-30.
[19]: MacDonald, P. (1999). Power, Type I, and Type III error rates of parametric and nonparametric statistical tests. The Journal of Experimental Education, 67, 367–379.
[20]: Nanna, M. J., & Sawilowsky, S. S. (1998). Analysis of Likert scale data in disability and medical rehabilitation research. Psychological Methods, 3, 55–67.
[21]: Posten, H. O., Yeh, H. C., & Owen, D. B. (1982). Robustness of the two-sample t-test under violations of the homogeneity of variance assumptions. *Communications in Statistics: Theory and Methods, 11*, 109-126.
[22]: Prana Ugiana Gio, Elly Rosmaini The (2018): Robustness of Two Independent Samples t Test Using Monte Carlo Simulation with R. Materials Science and Engineering 300 (2018) 012030.
[23]:Rasch, D., & Guiard, V. (2004). The robustness of parametric statistical methods. *Psychology Science, 46*, 175-208.
[24]: Ruxton, G. D. (2006). The unequal variance t-test is an underused alternative to Student's t-test and the Mann-Whitney U test. Behavioral Ecology, 17, 688–690.
[25]: Sawilowsky, S. S., and Blair, R. C. (1992). A more realistic look at the robustness and type II error properties of the t test to departures from population normality. Psychological Bulletin,111, 353-360
[26]: Sawilowsky, S. S., & Hillman, S. B. (1992). Power of the independent samples t test under a prevalent psychometric measure distribution. Journal of Consulting and Clinical Psychology, 60, 240–243.
[27]: Scheffé, H. (1959), The Analysis of Variance, New York: John Wiley & Sons.
[28]: Stonehouse, J. M., & Forrester, G. J. (1998). Robustness of the t and U tests under combined assumption violations. Journal of Applied Statistics, 25, 63-74.
[29]: Sullivan LM,D'Agostino RB(1992).Robustness of the t test applied to data distorted from normality by floor effects.JDentRes71:1938-1943.
[30]: W. S. Gosset, R. A. Fisher and Karl Pearson with notesand comments,Biometrika55, 445 – 457.
[31]: Welch, B. L. (1938). The significance of the difference between two means when the population variances are unequal. *Biometrika, 29*, 350-362.
[32]: Zimmerman, B. J., & Cleary, T. J. (2006). Adolescents' development of personal agency: The role of self-efficacy beliefs and self-regulatory skill. In F. Pajres, & T. Urdan (Eds.),Self-efficacy beliefs of adolescence(pp. 45–69). Mahwah, NJ7Information Age Publishing.
[33]: Zimmerman, D. W. (2006). Two separate effects of variance heterogeneity on the validity and power of significance tests of location. Statistical Methodology, 3, 351–374.
[34]: Zimmerman, D. W., & Zumbo, B. D. (1993). The relative power of parametric and nonparametric statistical methods. In G. Keren & C. Lewis (Eds.), A handbook for data analysis in the behavioral sciences: Methodological issues (pp. 481–517).