

علاقة حجم العينة بالقيم الحرجة لأساليب المقارنات المتعددة
اللابارامترية وذلك في ضوء عدد المجموعات ومستوى الدلالة

مرشد رده عسكر العتيبي

جامعة ام القرى, المملكة العربية السعودية

كلية التربية – قسم علم النفس

تخصص إحصاء وبحوث

E-Mail : murshedra2014@hotmail.com

No: 966563261778

مستخلص الدراسة

هدفت الدراسة الحالية إلى التعرف على علاقة حجم العينة بالقيم الحرجة لأساليب المقارنات المتعددة اللابارامترية في ضوء عدد المجموعات ومستوى الدلالة ، حيث سارت الدراسة على المنهج الوصفي التحليلي القائم على وصف وتحليل مجتمع الدراسة المكون من بيانات افتراضية غير معلومة التوزيع الاحتمالي ، تشمل عينة تتراوح بين 14 و 34 للعينات المستقلة ، وعينة تتراوح بين 10 و 30 للعينات المترابطة ، وقد تم استخدام مجموعة من الأساليب الإحصائية مع بيانات الدراسة الحالية وهي أسلوب كروسكال واليس ، فريدمان ، نيميني للمقارنات المتعددة ، دن للمقارنات المتعددة وكونوفر آيمان للمقارنات المتعددة وتوصلت الدراسة للنتائج عن طريق استخدام معامل ارتباط بيرسون بين نتائج الأساليب الإحصائية المستخدمة وحساب المتوسطات الحسابية والانحرافات المعيارية لنتائج هذه الأساليب وتوصلت الدراسة للنتائج التالية : وجود علاقة ارتباطية دالة احصائياً بين حجم العينة والقيم الحرجة لأساليب المقارنات المتعددة البعدية اللابارامترية لاختبار كروسكال-واليس في حال استخدام مستوى دلالة 0,01 ، 0,05 وعندما يكون عدد المجموعات 3 ، 4. كما كشفت الدراسة عن وجود علاقة ارتباطية دالة احصائياً بين حجم العينة والقيم الحرجة لأساليب المقارنات المتعددة البعدية اللابارامترية لاختبار فريدمان في حال استخدام مستوى دلالة 0,01 ، 0,05 وعندما يكون عدد المجموعات 5 ، 6. وفي ضوء نتائج الدراسة المستخلصة فقد أوصت الدراسة إلى استخدام اختبار نيميني لإجراء المقارنات البعدية لاختبار كروسكال-واليس وذلك في حالة ما إذا كان عدد المجموعات 3 أو 4 ولاختبار فريدمان وذلك في حالة ما إذا كان عدد المجموعات 5 ، 6 وذلك لكونه أقل وقوعاً في الخطأ من النوع الأول مقارنة بباقي أساليب المقارنات البعدية المتعددة اللابارامترية . كما أوصت الدراسة بعدم استخدام اختبار كونوفر لإجراء المقارنات البعدية اللابارامترية لاختبار كروسكال-واليس واختبار فريدمان، وذلك يعود لأن هذا الاختبار يعد من أكثر أساليب المقارنات البعدية اللابارامترية وقوعاً في الخطأ من النوع الأول .

الكلمات المفتاحية : كروسكال واليس – فريدمان – المقارنات اللابارامترية – حجم العينة – القيم الحرجة – عدد المجموعات – مستوى الدلالة .

Abstract

The current study aimed to investigate the relationship between sample size and critical values of non-parametric Post-hoc multiple comparison procedures in light of number of groups and level of significant , the study followed the descriptive analytical methodology which focuses on description and analysis the data of study that is consisted of assumption data which is unknown probability distribution , including sample consists of 14 to 34 for independent samples and sample consists of 10 to 30 for related samples . The current study have used a variety of statistical methods as Kruskal- wallis test , Friedman , Nemenyi , Dunn and Conover Iman . Also , through using Person correlation for output to above methods and using means with standard deviations the study found the following results : There is correlation with statistical significance between the critical values of Kruskal-Wallis Post-hoc nonparametric multiple comparisons procedures and sample size when the number of the groups is three , four and level of significance is 0.05 and 0.01. More ever , There is correlation with statistical significance between the critical values of Friedman Post-hoc nonparametric multiple comparisons procedures and sample size when the number of the groups is five , six and level of significance is 0.05 and 0.01. In light of study results , the current study recommended to use Nemenyi test for multiple comparisons with Kruskal – wallis test when number of group are 3 , 4 as well as with Friedman test when number of group are 5 , 6 . Due to this multiple comparison procedure considers less likely to fall in one type error . In addition , the study suggested to don't use Conover procedure as multiple comparisons with Kruskal – wallis and Friedman test . Due to this multiple comparison procedure considers more likely to fall in one type error .

Key words: Kruskal-Wallis–Friedman- Non Parametric comparisons- Sample size- Critical values- Number of groups – Level of significant .

أولاً- مقدمة الدراسة:

يهتم علم الإحصاء بالمعلومات والبيانات – ويهدف إلى جمعها وتبويبها وتنظيمها وتحليلها واستخلاص النتائج منها بل وتعميم نتائجها – واستخدامها في اتخاذ القرارات، ولقد أدى التقدم المذهل في تكنولوجيا المعلومات واستخدام الحاسبات الآلية إلى مساعدة الدارسين والباحثين ومتخذي القرارات في الوصول إلى درجات عالية ومستويات متقدمة من التحليل ووصف الواقع ومتابعته ومن ثم إلى التنبؤ بالمستقبل ، فالإحصاء باعتباره أحد أهم خطوات البحث العلمي فإنه بما يقدمه من أساليب إحصائية متعددة . يلجأ الباحثون إلى استخدامها والاعتماد عليها في اتخاذ القرارات حيال قبول أو رفض فروض الدراسة التي ينشد البحث العلمي الإجابة عليها والتأكد من صحتها أو عدمه ومن ثم الوصول إلى نتائج الدراسة اعتماداً على أساس إحصائي منطقي مقبول . ولهذه المبررات وما يترتب عليها من منافع علمية , اكتسب الإحصاء بما يتضمنه من أساليب إحصائية متنوعة أهمية كبيرة لارتباطه الوثيق بالبحث العلمي ووظيفة الأساليب الإحصائية في التعامل مع ما ينتج من البحوث العلمية من كميات كبيرة من البيانات والاعتماد عليها في تلخيص هذه البيانات وتفسيرها , ما جعل استخدام الأساليب الإحصائية شائعاً في مختلف المجالات البحثية في العلوم النفسية والتربوية والاجتماعية . " ويعتمد البحث في مجال العلوم النفسية والتربوية والاجتماعية على الإحصاء باعتباره أسلوباً فعالاً في وصف الظواهر في هذا المجال، فالإحصاء يؤدي دوراً بارزاً ليس في تنظيم البيانات ومعالجتها للخروج منها باستدلالات معينة فحسب، ولكن أيضاً في قيادة التفكير منهجياً نحو ما ينبغي عمله. " (الدردير , 2006، 35).

ويلجأ الباحثون في دراسة الظواهر النفسية والتربوية والاجتماعية إلى استخدام الأساليب والاختبارات الإحصائية في التحقق من صحة الفروض التي قدموها في بداية الدراسة ومن ثم اتخاذ القرارات حول قبولها أو رفضها. وتنقسم هذه الأساليب والاختبارات إلى قسمين رئيسيين وهما الأساليب البارامترية (المعلمية) والأساليب اللابارامترية (اللامعلمية). "أما التمييز بين الإحصاء البارامترية أو المعلمية Parametric statistic كأسلوب إحصائي ، والإحصاء اللابارامترية أو اللامعلمية nonparametric statistic كأسلوب إحصائي آخر، فيتعلق بنوع البيانات المراد تحليلها، ومستوى قياسها، فاستخدام الأسلوب الإحصائي المناسب يعتمد على طبيعة البيانات (عددية / تصنيفية، أو كمية / قياسية)، ومستوى قياس المتغيرات موضع البحث (اسمية، أو رتبية، أو فترية، أو نسبية). " (علام , 2010، 35) .

ويذكر الدردير (2006) "إن حجم العينة ونوع البيانات التي نحصل عليها يحددان نوع الاختبارات الإحصائية الاستدلالية المستخدمة" . ويلمس الباحث أن الأساليب الإحصائية اللابارامترية هي أساليب حرة وغير مقيدة بطبيعة التوزيع للمجتمع الأصلي محل الدراسة، كما أنها لا تتطلب استيفاء افتراضات معينة حول هذا المجتمع من قبيل تجانس التباين والاعتدالية.

" والأساليب الإحصائية اللابارامترية أكثر استخداماً في العلوم النفسية والتربوية والاجتماعية والسلوكية عموماً، وذلك لأنها مناسبة أكثر لطبيعة هذه الظواهر التي يصعب الحصول فيها على قياسات دقيقة من المستوى الفترية على الأقل، بالإضافة أن عملياتها الحسابية غير معقدة، وتجمع البيانات بطرق أقل كلفة نظراً لصغر أحجام العينات في حالة الأساليب اللابارامترية" . (الشربيني , 2001، 99) .

ويستنتج الباحث أن استخدام الأساليب الإحصائية اللابارامترية أكثر ملائمة في دراسة الظواهر السلوكية عموماً، وذلك يعود لرواج حدوث انتهاكات في الافتراضات التي تعتمد عليها الأساليب البارامترية على سبيل المثال افتراضية التوزيع الطبيعي وتجانس التباين، بالإضافة إلى أنها الأساليب الوحيدة التي يمكن عن طريقها دراسة البيانات ذات المستوى الاسمي والمستوى الرتبتي، كما أن الأساليب اللابارامترية تصلح للعينات صغيرة الحجم.

وغالبا ما تستخدم الأساليب الإحصائية بنوعيتها في اختبار الفروض الفارقة للكشف عن الفروق بين متوسطات المجتمع محل الدراسة من خلال إحصاءات العينة، وتتسم الظواهر النفسية والتربوية والسلوكية عموماً بتأثرها بمتغير مستقل له أكثر من مستويين – أو أكثر من متغير مستقل واحد – مما يستدعي وجود أكثر من معالجتين لدراسة أثر المتغير المستقل على المتغير التابع. وهذا يضطر الباحثين إلى استخدام أسلوب تحليل التباين الأحادي للمقارنة بين ثلاث متوسطات أو أكثر في حال قياس المتغير التابع مرة واحدة، أو تحليل التباين ذي القياسات المتكررة في حال قياس المتغير التابع أكثر من مرة. وهناك بدائل لتحليل التباين الأحادي وذي القياس المتكرر في الأساليب الإحصائية اللابارامترية مثل اختبار كروسكال واليس، للعينات المستقلة وهو بديل لتحليل التباين الأحادي، واختبار فريدمان للعينات المترابطة وهو بديل لتحليل التباين ذي القياس المتكرر وهذان الاختباران هما موضع الدراسة الحالية. وفي حال وجود فروق ذات دلالة إحصائية بين هذه المجموعات فلا يمكن الكشف عن اتجاه الفروق بين هذه المجموعات لا عن طريق اختبار كروسكال واليس ولا عن طريق اختبار فريدمان، وذلك ما يستدعي تطبيق المقارنات البعدية المتعددة كما هو الحال في تحليل التباين الأحادي.

ويذكر (Pohlert, 2016) انه لتحديد اتجاه الفروق ذات الدلالة الإحصائية بين المجموعات تطبق الاختبارات البعدية عن طريق المقارنات الثنائية المتعددة من قبيل اختبار نيميني، واختبار دن، واختبار كونوفر آيمان. فالمقارنات المتعددة الثنائية بواسطة إجرائها يمكن معرفة اتجاه الفروق بين كل مجموعتين من مجموعات المقارنة عن طريق مقارنة القيمة الناتجة بالقيم الحرجة فإن كانت تساويها أو أكبر منها دلت على وجود فروق بين المجموعتين محل المقارنة.

وقد أكد (William, 2013, 130) أن الباحثين يقعون في أخطاء عند اختيارهم لأساليب المقارنات المتعددة وذلك من حيث عدم مناسبتها لطبيعة بيانات بحثهم وهذا راجع إلى نقص المعرفة الإحصائية لدى هؤلاء الباحثين. كما أن هناك مشكلة بالنسبة لاستخدام أساليب المقارنات المتعددة والتي تتعلق بعدم انتباه الباحثين إلى الفروض التي تستند إليها أساليب المقارنات المتعددة، حيث إنه في حالة عدم تحقق أحد هذه الفروض فإن نتائج استخدام أساليب المقارنات المتعددة تصبح ليست ذات معنى (Richard, 2004, 128), (Zeighami, 1973, 3). ويؤكد Mary (2014) إن أساليب المقارنات المتعددة لا تصلح جميعها لكل الأغراض فقد يكون أسلوباً ما مناسباً لدراسة معينة وغير مناسب لدراسة أخرى.

ومن هذا المنطلق سعت الدراسة الحالية للكشف عن العلاقة بين أساليب المقارنات المتعددة البعدية لاختباري كروسكال واليس واختبار فريدمان والمتمثلة في أساليب نيميني، دن، كونوفر آيمان وحجم العينة في كل اختبار وذلك في ضوء عدد المجموعات ومستوى الدلالة المستخدم.

ثانياً- مشكلة الدراسة:

من الملاحظ أن الأساليب المستخدمة لإجراء المقارنات الثنائية المتعددة بعد القيام بإجراء اختبار كروسكال واليس واختبار فريدمان كثيرة ومتعددة، الأمر الذي يجعل الباحثين يقعون في حيرة عند اختيار أساليب المقارنات المتعددة لمعرفة دلالة الفروق بين وسطي كل مجموعتين من المجموعات الخاضعة للتجريب، وقد تم تأكيد ذلك من كل من (Pohlert,2016)، (حمدان , 2015)، (Mary,2014)، (الشمراي , 2000)، (Alyones , 1999). وقد تبنت كل دراسة من الدراسات السابقة اعتماد أسلوب واحد أو عدد من أساليب المقارنات المتعددة الذي يفضل استخدامه بعد كل من الاختبارين اللابارامترين في حال وجود فروق ذات دلالة إحصائية. أيضاً فقد لاحظ الباحث عدم وجود دراسات عربية أو أجنبية تناولت العلاقة بين حجم العينة والقيم الحرجة لأساليب المقارنات المتعددة اللابارامترية وذلك في ضوء عدد المجموعات ومستوى الدلالة المستخدم، وتبرز مشكلة الدراسة في التساؤل الرئيس التالي: ما العلاقة بين حجم العينة والقيم الحرجة لأساليب المقارنات المتعددة البعدية اللابارامترية في ضوء عدد المجموعات ومستوى الدلالة؟ ويتفرع عن هذا السؤال الرئيس عدة أسئلة فرعية وهي:

1. ما علاقة حجم العينة بالقيم الحرجة لأساليب المقارنات المتعددة البعدية اللابارامترية لاختبار كروسكال-واليس في حال استخدام مستوى دلالة 0,01 , 0,05 وثلاث مجموعات؟ .
2. ما علاقة حجم العينة بالقيم الحرجة لأساليب المقارنات المتعددة البعدية اللابارامترية لاختبار كروسكال-واليس في حال استخدام مستوى دلالة 0,01 , 0,05 وأربع مجموعات؟ .
3. ما علاقة حجم العينة بالقيم الحرجة لأساليب المقارنات المتعددة البعدية اللابارامترية لاختبار فريدمان في حال استخدام مستوى دلالة 0,01 , 0,05 وخمس مجموعات؟ .
4. ما علاقة حجم العينة بالقيم الحرجة لأساليب المقارنات المتعددة البعدية اللابارامترية لاختبار فريدمان في حال استخدام مستوى دلالة 0,01 , 0,05 وست مجموعات؟ .

ثالثاً-أهداف الدراسة:

هدفت الدراسة الحالية إلى التعرف على:

1. علاقة حجم العينة بالقيم الحرجة لأساليب المقارنات المتعددة البعدية اللابارامترية لاختبار كروسكال-واليس في حال استخدام مستوى دلالة 0,01 , 0,05 وثلاث مجموعات.
2. علاقة حجم العينة بالقيم الحرجة لأساليب المقارنات المتعددة البعدية اللابارامترية لاختبار كروسكال-واليس في حال استخدام مستوى دلالة 0,01 , 0,05 وأربع مجموعات.
3. علاقة حجم العينة بالقيم الحرجة لأساليب المقارنات المتعددة البعدية اللابارامترية لاختبار فريدمان في حال استخدام مستوى دلالة 0,01 , 0,05 وخمس مجموعات.
4. علاقة حجم العينة بالقيم الحرجة لأساليب المقارنات المتعددة البعدية اللابارامترية لاختبار فريدمان في حال استخدام مستوى دلالة 0,01 , 0,05 وست مجموعات.

رابعاً- أهمية الدراسة:

1. تسليط الضوء على أساليب المقارنات البعدية اللابارامترية الخاصة باختبار كروسكال واليس واختبار فريدمان حيث ان المراجع العربية لم تتناول هذه الأساليب.
2. إن معرفة أكثر أساليب المقارنات الثنائية البعدية اللابارامترية تأثيراً بحجم العينة ومستوى الدلالة وعدد المجموعات يفيد الباحثين في اختيار الأسلوب المناسب لطبيعة دراستهم.
3. تقديم إطار عام يساعد الباحثين في اختيار أسلوب المقارنات المتعددة اللابارامترية المناسب لدراساتهم من حيث حجم العينة ومستوى الدلالة وعدد المجموعات.

خامساً- مصطلحات الدراسة:

1- القيم الحرجة لأساليب المقارنات المتعددة البعدية:

هي تلك القيم التي يمكن من خلالها الحكم على دلالة الفروق بين المتوسطات الرتيبة للمجموعات موضع المقارنة في اللابارامترية لتحليل التباين، وهي تعتمد على حجم العينة ومستوى الدلالة وعدد المجموعات. (Pohlert , 2016).

2- أساليب المقارنات الثنائية البعدية المتعددة اللابارامترية:

هي الأساليب التي تستخدم بعد إجراء الاختبارات اللابارامترية لعدد من المجموعات لتحديد دلالة الفروق بين المتوسطين الرتيبين لكل مجموعتين على حده. (Pohler , 2016).

3- مستوى الدلالة:

هو احتمال رفض الفرض الصفري عندما يكون صحيحاً. (Andy & Jeremy, 2010).

سادساً- حدود الدراسة:

اقتصرت الدراسة الحالية على أشهر اختبارات لابارامتريين، وهما بديلان لتحليل التباين الأحادي، والتباين ذي القياس المتكرر وهما:

- 1- اختبار كروسكال واليس للعينات المستقلة. 2- اختبار فريدمان للعينات المترابطة.

كما اقتصرت الدراسة الحالية على دراسة أثر ثلاثة عوامل فقط على القيم الحرجة لأساليب المقارنات البعدية المتعددة اللابارامترية وهي كالتالي:

- 1- حجم العينة. 2- مستوى الدلالة. 3- عدد المجموعات.

سابعاً: فروض الدراسة:

نشدت الدراسة الحالية الإجابة عن الفروض التالية:

- 1) لا توجد علاقة ارتباطية دالة إحصائية بين حجم العينة والقيم الحرجة لأساليب المقارنات المتعددة البعدية اللابارامترية لاختبار كروسكال-واليس عندما يكون عدد المجموعات ثلاثة ومستوى الدلالة 01,0 و 0,05.

- (2) لا توجد علاقة ارتباطيه دالة إحصائيا بين حجم العينة والقيم الحرجة لأساليب المقارنات المتعددة البعدية اللابارامترية لاختبار كروسكال- واليس عندما يكون عدد المجموعات أربعة ومستوى الدلالة 0,01 و 0,05
- (3) لا توجد علاقة ارتباطيه دالة إحصائيا بين حجم العينة والقيم الحرجة لأساليب المقارنات المتعددة البعدية اللابارامترية لاختبار فريدمان عندما يكون عدد المجموعات خمسة ومستوى الدلالة 0,01 و 0,05.
- (4) لا توجد علاقة ارتباطيه دالة إحصائيا بين حجم العينة والقيم الحرجة لأساليب المقارنات المتعددة البعدية اللابارامترية لاختبار فريدمان عندما يكون عدد المجموعات ستة ومستوى الدلالة 0,01 و 0,05

منهج الدراسة وإجراءاتها:

أولاً-منهج الدراسة: اعتمد الباحث في الدراسة الحالية على استخدام المنهج الوصفي (التحليلي) وهو من أكثر المناهج البحثية شيوعا في البحث العلمي , حيث يساعد الباحثين في وصف وتحليل الكثير من الظواهر , وهو يعتمد في تفسير نتائجه بشكل كبير على زمن الدراسة , وحجم العينة , وأدوات القياس , والتحليلات الإحصائية , وغيرها من أجل التوصل إلى تفسير الظاهرة واكتشاف المعاني والعلاقات الخاصة بها (المنيزل , العنوم , 2010) .

ثانيا: مجتمع الدراسة وعينتها:

أمجتمع الدراسة: تكون مجتمع الدراسة الحالية من بيانات افتراضية غير معلومة التوزيع الاحتمالي (أي لا تتحقق فيها افتراضات الأساليب الإحصائية المعلمية) حيث تم قياسها وفق المقياس الرتبي والذي يتناسب فقط مع الأساليب اللابارامترية . ولقد تم توليد البيانات باستخدام برنامج الاكسيل **Microsoft Excel 2010** وذلك بما يتماشى مع متطلبات الدراسة .

ب-تم سحب عينات الدراسة من مجتمع الدراسة بشكل عشوائي بشروط تتناسب مع أسئلة الدراسة ومتطلبات اختبار كروسكال-واليس واختبار فريدمان كما يلي :

1. توليد بيانات لثلاث مجموعات بحيث يكون اجمالي العينة 14-20-26 (مع مراعاة عدم تساوي المجموعات)، وتطبيق اختبار كروسكال واليس، أما بالنسبة لاختبار فريدمان فقد تم توليد بيانات لخمس مجموعات بحيث يكون اجمالي العينة 10-20-30 (مع تحقيق الترابط بين العينات لفريدمان).
2. توليد بيانات لأربع مجموعات بحيث يكون اجمالي العينة 18-26-34 (مع مراعاة عدم تساوي المجموعات)، وتطبيق اختبار كروسكال واليس، أما بالنسبة لاختبار فريدمان فقد تم توليد بيانات لست مجموعات بحيث يكون اجمالي العينة 10-20-30 (مع تحقيق الترابط بين العينات لفريدمان).

ثالثاً-أداة الدراسة:

تم تفرغ بيانات الدراسة من خلال برنامج الحزمة الإحصائية للعلوم الإنسانية (SPSS) وهو نفس البرنامج المستخدم في تحليل بيانات الدراسة .

رابعاً- إجراءات الدراسة:

1. تم حساب قيم اختبار كروسكال واليس عند حجوم العينات المختلفة وعندما يكون عدد المجموعات ثلاثة وذلك عند مستوى دلالة 0,01 و 0,05.
2. تم حساب قيم اختبار فريدمان عند حجوم العينات المختلفة وعندما يكون عدد المجموعات خمسة وذلك عند مستوى دلالة 0,01 و 0,05.

إجراء أساليب المقارنات البعدية اللابارامترية كلا على حده بعد اختبار كروسكال واليس، واختبار فريدمان في الإجراء رقم (1, 2) ، وذلك بأسلوب: نيميني، دن، كونوفير، تبعاً لاختبار كروسكال-واليس. وأسلوب: نيميني، كونوفير تبعاً لاختبار فريدمان، حيث تم استبعاد اختبار مان-ويتني واختبار ويلكوكسون نظراً لأن هذين الاختبارين بالنظر إلى تطبيقهما في برنامج SPSS فإنهما يعتمدان في الكشف على دلالة الفروق بين المجموعات موضع المقارنة على قيم P- value أي على الدلالة الإحصائية ذاتها وبالتالي لا يمكن تتبع مؤشر القيم الحرجة لمعرفة تأثيرها بحجم العينة من عدمه , كما أنهما يعدان من أكثر الاختبارات البعدية اللابارامترية وقوعاً في الخطأ من النوع الأول حيث أشار لذلك كل من (Siegel Castellan (1988)، (Daniel (1990)، (Hollander & Wolfe (1999)، (Demsar (2006) .

3. حساب معامل الارتباط بين القيم الحرجة لكل أسلوب من هذه الأساليب البعدية وحجم العينة باستخدام معامل ارتباط بيرسون (Person correlation) كل اختبار على حده للإجابة على الفرض الصفري الأول القائل (لا توجد علاقة ارتباطية دالة إحصائية بين حجم العينة والقيم الحرجة لأساليب المقارنات المتعددة البعدية اللابارامترية لاختبار كروسكال واليس عندما يكون عدد المجموعات ثلاثة ومستوى الدلالة 01,0 و 0,05). وتكرار نفس الإجراء للإجابة على الفرض الثاني.
4. تكرار الإجراءات السابقة حينما يكون عدد المجموعات خمسة – ستة مع اختبار فريدمان للإجابة على الفرضين الثالث والرابع.

خامساً- الأساليب الإحصائية المستخدمة في الدراسة:

للتحقق من صحة فروض الدراسة، استخدم الباحث عدداً من الأساليب الإحصائية باستخدام البرنامج الإحصائي (SPSS) ولقد كانت هذه الأساليب تتعلق بجانبين وهي على النحو التالي :

أ- أساليب إحصائية تم استخدامها في إجراءات الدراسة وهي كالتالي :

1- أسلوب كروسكال واليس للعينات المستقلة.

2- أسلوب فريدمان للعينات المترابطة .

3- اختبار نيميني للمقارنات البعدية .

4- اختبار دن للمقارنات البعدية .

5- اختبار كونوفير للمقارنات البعدية .

ب- أساليب إحصائية تم استخدامها للإجابة على فروض الدراسة وهي :

1- معامل ارتباط بيرسون .

2- المتوسطات الحسابية والانحرافات المعيارية .

نتائج الدراسة ومناقشتها

أولاً-نتائج الفرض الأول ومناقشتها:

ينص الفرض الأول على أنه " لا توجد علاقة ارتباطية دالة إحصائية بين حجم العينة والقيم الحرجة لأساليب المقارنات المتعددة البعدية اللابارامترية لاختبار كروسكال-واليس عندما يكون عدد المجموعات ثلاثة ومستوى الدلالة 0.05، 0.01".

وللتحقق من صحة هذا الفرض تم حساب قيم الفروق بين الفرق المطلق D (متوسط رتب المجموعة i – متوسط رتب المجموعة j) وبين القيم الحرجة لأساليب المقارنات المتعددة البعدية المستخدمة في حالة اختبار كروسكال –واليس وهي (اختبار نيميني، اختبار دن، اختبار كونوفير) عند مستويي دلالة 0.05، 0.01 وعند عدد مجموعات يساوي ثلاثة، وتم حساب معامل الارتباط بين ناتج الفروق لكل اختبار من هذه الأساليب وحجم العينة باستخدام معامل ارتباط بيرسون , ويوضح جدول (1) معاملات ارتباط بيرسون ودلالاتها الإحصائية : ويتضح من جدول (1) أن جميع معاملات الارتباط بين ناتج الفروق لكل اختبار من هذه الأساليب وحجم العينة كانت دالة إحصائياً عند مستوى 0.01، مما يدل على أن قيم الفروق بين الفرق المطلق D وبين القيم الحرجة لأساليب المقارنات المتعددة البعدية اللابارامترية تتأثر بحجم العينة .

جدول (1)

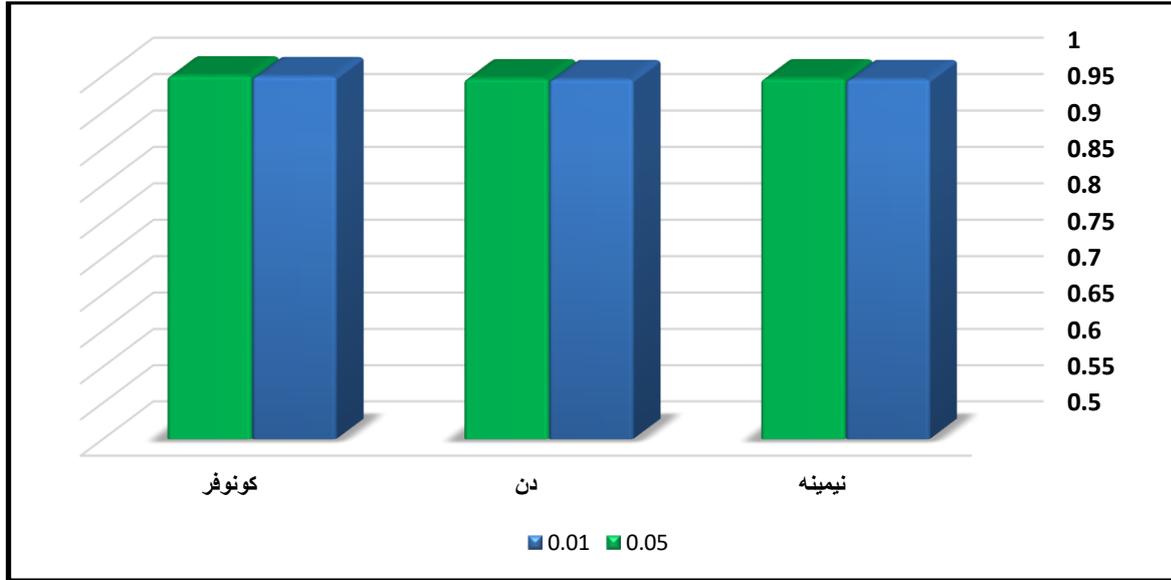
معاملات الارتباط بين ناتج الفروق لأساليب المقارنات المتعددة اللابارامترية لاختبار كروسكال-واليس وحجم العينة عند استخدام مستوى دلالة 0.05، 0.01 وثلاث مجموعات .

أساليب المقارنات البعديّة	معامل ارتباط بيرسون	
	مستوي دلالة 0.01	مستوي دلالة 0.05
نيمينيّه	**0.995	**0.996
دن	**0.995	**0.996
كونوفير	**0.999	**1.000

** دالة عند مستوى 0.01

ومن الملاحظ أن جميع معاملات الارتباط كانت موجبة مما يدل على أن قيم الفروق بين الفرق المطلق D وبين القيم الحرجة لأساليب المقارنات المتعددة البعدية اللابارامترية تزيد كلما زاد حجم العينة، مما يعني أنه كلما زاد حجم العينة كلما زادت قيم الفروق بين المتوسطات الرتبية الفرق المطلق D وبين القيم الحرجة لأساليب المقارنات المتعددة البعدية لاختبار كروسكال-واليس، مما يعطي فرصة أكبر لإعطاء دلالة للمقارنات الثنائية بين متوسطات رتب درجات المجموعات.

كذلك يتضح من جدول (1) أن اختبار كونوفير كان أكثر أساليب المقارنات المتعددة البعدية ارتباطاً بحجم العينة، سواء أكان مستوى الدلالة المستخدم 0.05 أو 0.01 حيث بلغت قيمة معامل الارتباط بين ناتج الفروق لاختبار كونوفير وبين حجم العينة (1.000) أي ارتباط موجب تام وذلك عند استخدام مستوى دلالة 0.05، أيضاً فقد بلغت قيمة معامل الارتباط بين ناتج الفروق لاختبار كونوفير وحجم العينة (0.999) وذلك عند استخدام مستوى دلالة 0.01، مما يدل على أن اختبار كونوفير أكثر أساليب المقارنات البعدية تأثراً بحجم العينة، بينما كان اختبار نيمينيّه ودن أقل أساليب المقارنات المتعددة ارتباطاً بحجم العينة وذلك في حالة استخدام مستوى دلالة 0.05 أو 0.01 حيث بلغت قيمة معامل الارتباط بين ناتج الفروق لاختبار نيمينيّه ودن وحجم العينة (0.996) عند استخدام مستوى دلالة 0.05، بينما بلغت قيمة معامل الارتباط بين ناتج الفروق لاختبار نيمينيّه ودن وحجم العينة (0.995) عند استخدام مستوى دلالة 0.01 مما يدل على أن اختبار نيمينيّه ودن أقل أساليب المقارنات البعدية تأثراً بحجم العينة، ويوضح شكل (1) معاملات ارتباط الفروق الناتجة لأساليب المقارنات البعدية لاختبار كروسكال-واليس وحجم العينة عند استخدام مستوى دلالة 0.05 ، 0.01 وثلاث مجموعات .



شكل (1) معاملات ارتباط الفروق لأساليب المقارنات البعدية لاختبار كروسكال-واليس وحجم العينة عند استخدام مستوى دلالة 0.05 ، 0.01 وثلاث مجموعات .

وقد تم حساب المتوسطات الحسابية والانحرافات المعيارية للفروق بين D والقيم الحرجة لأساليب المقارنات البعدية لاختبار كروسكال-واليس كما هو موضح بالجدول التالي :

جدول (2)

المتوسطات الحسابية والانحرافات المعيارية للفروق بين D والقيم الحرجة لأساليب المقارنات البعدية لاختبار كروسكال-واليس عند استخدام مستوى دلالة 0.05 ، 0.01 وثلاثة مجموعات .

مستوي الدلالة المستخدم 0.01		مستوي الدلالة المستخدم 0.05		أساليب المقارنات البعدية
الانحراف المعياري	المتوسط الحسابي	الانحراف المعياري	المتوسط الحسابي	
2.48	2.79	2.61	4.91	نيمينه
2.487	2.866	2.619	4.98	دن
2.412	7.670	2.545	9.310	كونوفر

ويتضح من جدول (2) ما يلي :

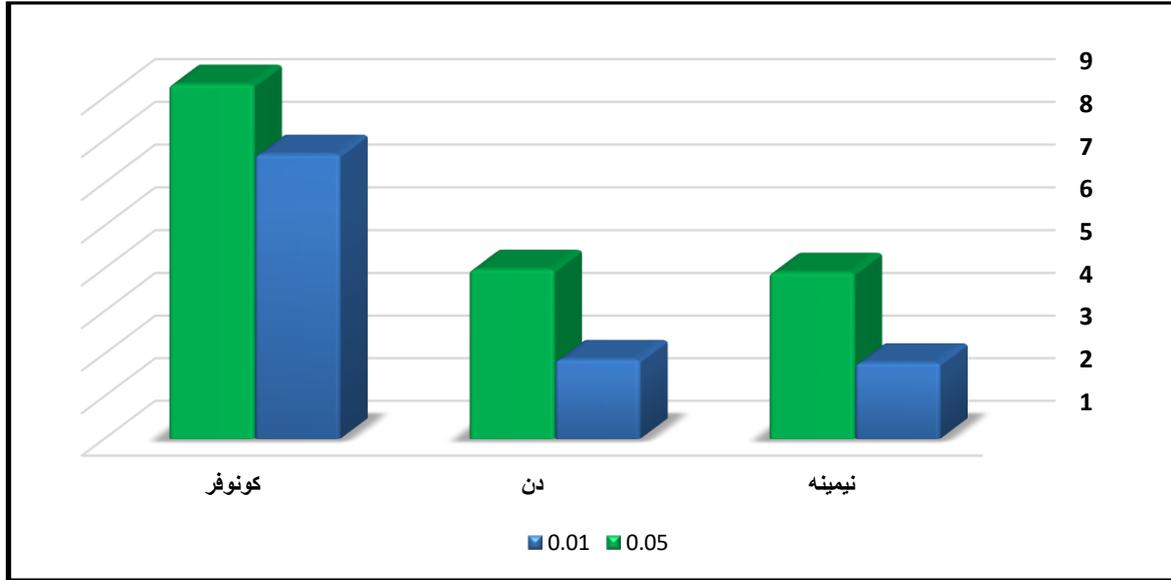
أ- في حالة ما إذا كان مستوى الدلالة المستخدم 0.05 :

يتضح من جدول (2) أن قيمة المتوسط الحسابي للفروق بين D والقيم الحرجة لاختبار نيمينييه (4.91) وهو أقل من قيم المتوسطات الحسابية للفروق بين D والقيم الحرجة لأساليب المقارنات البعدية الباقية، مما يدل على أن اختبار نيمينييه يعطي حدوداً للثقة أكبر من بقية أساليب المقارنات البعدية مما يعطي فرصة أقل لإعطاء دلالة للمقارنات الثنائية بين متوسطات رتب المجموعات، بينما كانت قيمة المتوسط الحسابي للفروق بين D والقيم الحرجة لاختبار كونوفير (9.310) ، ومن الملاحظ أن المتوسط للفروق بين D والقيم الحرجة لاختبار كونوفير أكبر من بقية أساليب المقارنات البعدية ، وهذا يعني أن اختبار كونوفير يعطي حدوداً للثقة أقل من حدود الثقة المحسوبة باستخدام باقي أساليب المقارنات البعدية اللابارامترية، وهذا بالطبع يؤدي إلى أن اختبار كونوفير يعطي فرصة أكبر لإعطاء دلالة للمقارنات الثنائية بين متوسطات رتب درجات المجموعات .

ب- في حالة ما إذا كان مستوى الدلالة المستخدم 0.01 :

يتضح من جدول (2) أن قيمة المتوسط الحسابي للفروق بين D والقيم الحرجة لاختبار نيمينييه (2.79) وهو أقل من قيم المتوسطات الحسابية للفروق بين D والقيم الحرجة لأساليب المقارنات البعدية الباقية، مما يدل على أن اختبار نيمينييه يعطي حدوداً للثقة أكبر من بقية أساليب المقارنات البعدية مما يعطي فرصة أقل لإعطاء دلالة للمقارنات الثنائية بين متوسطات رتب المجموعات ، بينما كانت قيمة المتوسط الحسابي للفروق بين D والقيم الحرجة لاختبار كونوفير (7.670) ، ومن الملاحظ أن المتوسط للفروق بين D والقيم الحرجة لاختبار كونوفير أكبر من بقية أساليب المقارنات البعدية ، وهذا يعني أن اختبار كونوفير يعطي حدوداً للثقة أقل من حدود الثقة المحسوبة باستخدام باقي أساليب المقارنات البعدية اللابارامترية، وهذا بالطبع يؤدي إلى أن اختبار كونوفير يعطي فرصة أكبر لإعطاء دلالة للمقارنات الثنائية بين متوسطات رتب درجات المجموعات .

ويوضح شكل (2) متوسطات الفروق بين D والقيم الحرجة لأساليب المقارنات البعدية لاختبار كروسكال- واليس باختلاف حجم العينة عند استخدام مستوى دلالة 0.05 ، 0.01 وثلاث مجموعات .



شكل (2) متوسطات الفروق بين D والقيم الحرجة لأساليب المقارنات البعدية لاختبار كروسكال-واليس باختلاف حجم العينة عند استخدام مستوى دلالة 0.05 ، 0.01 وثلاث مجموعات .

وقد بينت نتائج الفرض الأول أن القيم الحرجة لأساليب المقارنات البعدية تتأثر جميعها تأثيراً كبيراً بحجم العينة، وقد لوحظ أن اختبار كونوفر من أكثر أساليب المقارنات البعدية تأثيراً بحجم العينة وذلك سواء كان مستوى الدلالة المستخدم 0.05 أو 0.01، كما بينت نتائج الفرض الأول أن اختبار كونوفر من أكثر أساليب المقارنات البعدية وقوعاً في الخطأ من النوع الأول وهذا ما أكدته كل من (Shuster & Boyett, 1979)، (Campbell & Skillings, 1985)، (Neuhäuser, & Bretz, 2001) .

أيضاً فقد بينت نتائج الفرض الأول أن اختبار نيمينه من أكثر أساليب المقارنات البعدية اللابارامترية تحفظاً، حيث أن متوسط القيم الحرجة لاختبار نيمينه كان أقل من قيم متوسطات القيم الحرجة لأساليب المقارنات البعدية اللابارامترية الأخرى وذلك سواء كان مستوى الدلالة 0.05 أو 0.01 .

ثانياً-نتائج الفرض الثاني ومناقشتها:

ينص الفرض الثاني على أنه " لا توجد علاقة ارتباطية دالة إحصائية بين حجم العينة والقيم الحرجة لأساليب المقارنات المتعددة البعدية اللابارامترية لاختبار كروسكال-واليس عندما يكون عدد المجموعات أربعة ومستوى الدلالة 0.05، 0.01" .

وللتحقق من صحة هذا الفرض تم حساب قيم الفروق بين الفرق المطلق D (متوسط رتب المجموعة j – متوسط رتب المجموعة z) وبين القيم الحرجة لأساليب المقارنات المتعددة البعدية المستخدمة في حالة اختبار كروسكال-واليس وهي (اختبار نيمينه، اختبار دن، اختبار كونوفر) عند مستويي دلالة 0.05، 0.01 وعند عدد مجموعات يساوي أربعة ، وتم حساب معامل الارتباط بين ناتج الفروق لكل اختبار من هذه الأساليب وحجم العينة باستخدام معامل ارتباط بيرسون ويوضح جدول (3) معاملات ارتباط بيرسون ودلالاتها الإحصائية :

ويتضح من جدول (3) أن جميع معاملات الارتباط بين ناتج الفروق لكل اختبار من هذه الأساليب وحجم العينة كانت دالة إحصائياً عند مستوى 0.01، مما يدل على أن قيم الفروق بين الفرق المطلق D وبين القيم الحرجة لأساليب المقارنات المتعددة البعدية اللابارامترية تتأثر بحجم العينة .

جدول (3)

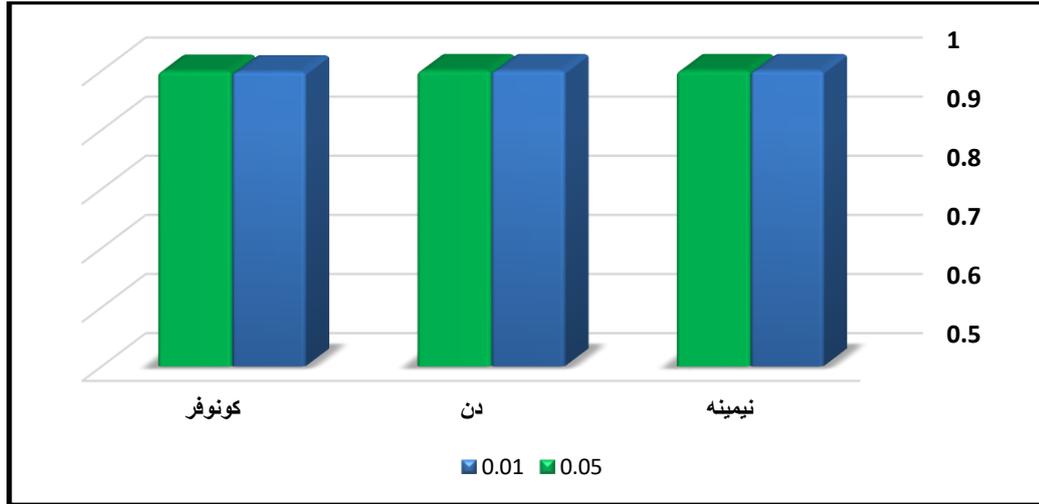
معاملات الارتباط بين ناتج الفروق لأساليب المقارنات المتعددة اللابارامترية لاختبار كروسكال-واليس وحجم العينة عند استخدام مستوى دلالة 0.05، 0.01 واربعة مجموعات .

معامل ارتباط بيرسون		أساليب المقارنات البعدية
مستوي دلالة 0.01	مستوي دلالة 0.05	
**1.000	**1.000	نيمينييه
**1.000	**1.000	دن
**0.998	**0.999	كونوفير

** دالة عند مستوى 0.01

ومن الملاحظ أن جميع معاملات الارتباط كانت موجبة مما يدل على أن قيم الفروق بين الفرق المطلق D وبين القيم الحرجة لأساليب المقارنات المتعددة البعدية اللابارامترية تزيد كلما زاد حجم العينة، مما يعني أنه كلما زاد حجم العينة كلما زادت قيم الفروق بين الفرق المطلق D وبين القيم الحرجة لأساليب المقارنات المتعددة البعدية لاختبار كروسكال-واليس ، مما يعطي فرصة أكبر لإعطاء دلالة للمقارنات الثنائية بين متوسطات رتب درجات المجموعات .

كذلك يتضح من جدول (3) أن اختبار نيمينييه واختبار دن كانا أكثر أساليب المقارنات المتعددة البعدية ارتباطاً بحجم العينة، سواء أكان مستوى الدلالة المستخدم 0.05 أو 0.01 حيث بلغت قيمة معامل الارتباط بين ناتج الفروق لاختبار نيمينييه ودن وبين حجم العينة (1.000) أي ارتباط موجب تام وذلك عند استخدام مستوى دلالة 0.05، وعند استخدام مستوى دلالة 0.01 أيضاً . مما يدل على أن اختبار نيمينييه واختبار دن أكثر أساليب المقارنات البعدية تأثراً بحجم العينة، بينما كان اختبار كونوفير أقل أساليب المقارنات المتعددة ارتباطاً بحجم العينة وذلك في حالة استخدام مستوى دلالة 0.05 أو 0.01 حيث بلغت قيمة معامل الارتباط بين ناتج الفروق لاختبار كونوفير وحجم العينة (0.999) عند استخدام مستوى دلالة 0.05، بينما بلغت قيمة معامل الارتباط بين ناتج الفروق لاختبار كونوفير وحجم العينة (0.998) عند استخدام مستوى دلالة 0.01 مما يدل على أن اختبار كونوفير أقل أساليب المقارنات البعدية تأثراً بحجم العينة، ويوضح شكل (3) معاملات ارتباط الفروق الناتجة لأساليب المقارنات البعدية لاختبار كروسكال-واليس وحجم العينة عند استخدام مستوى دلالة 0.05 ، 0.01 واربعة مجموعات .



شكل (3) معاملات ارتباط الفروق لأساليب المقارنات البعدية لاختبار كروسكال-واليس وحجم العينة عند استخدام مستوى دلالة 0.05 ، 0.01 واربعة مجموعات .

وقد تم حساب المتوسطات الحسابية والانحرافات المعيارية للفروق بين D والقيم الحرجة لأساليب المقارنات البعدية لاختبار كروسكال كما هو موضح بالجدول التالي :

جدول (4)

المتوسطات الحسابية والانحرافات المعيارية للفروق بين D والقيم الحرجة لأساليب المقارنات البعدية لاختبار كروسكال-واليس عند استخدام مستوى دلالة 0.05 ، 0.01 وأربع مجموعات .

مستوي الدلالة المستخدم 0.01		مستوي الدلالة المستخدم 0.05		أساليب المقارنات البعدية
الانحراف المعياري	المتوسط الحسابي	الانحراف المعياري	المتوسط الحسابي	
4.215	6.066	4.525	8.390	نيمينه
4.220	6.140	4.535	8.460	دن
6.016	15.953	5.965	16.903	كونوفر

ويتضح من جدول (4) ما يلي :

أ- في حالة ما إذا كان مستوى الدلالة المستخدم 0.05 :

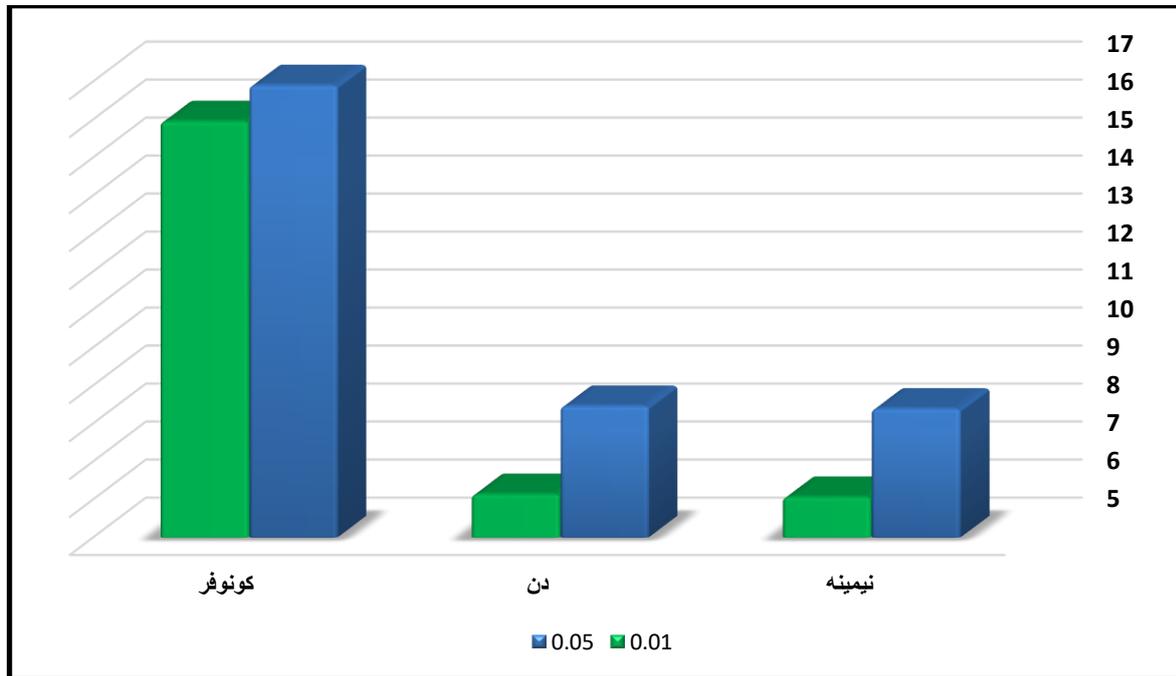
يتضح من جدول (4) أن قيمة المتوسط الحسابي للفروق بين D والقيم الحرجة لاختبار نيمينه (8.39) وهو أقل من قيم المتوسطات الحسابية للفروق بين D والقيم الحرجة لأساليب المقارنات البعدية الباقية، مما يدل على أن اختبار نيمينه يعطي حدوداً للثقة أكبر من بقية أساليب المقارنات البعدية مما يعطي فرصة أقل لإعطاء دلالة للمقارنات الثنائية بين متوسطات رتب المجموعات، بينما كانت قيمة المتوسط الحسابي للفروق بين D والقيم الحرجة لاختبار كونوفر (16.90) ، ومن الملاحظ أن المتوسط للفروق بين D والقيم الحرجة لاختبار

كونوفير أكبر من بقية أساليب المقارنات البعدية ، وهذا يعني أن اختبار كونوفير يعطي حدوداً للثقة أقل من حدود الثقة المحسوبة باستخدام باقي أساليب المقارنات البعدية اللابارامترية، وهذا بالطبع يؤدي إلى أن اختبار كونوفير يعطي فرصة أكبر لإعطاء دلالة للمقارنات الثنائية بين متوسطات رتب درجات المجموعات .

ب- في حالة ما إذا كان مستوى الدلالة المستخدم 0.01:

يتضح من جدول (4) أن قيمة المتوسط الحسابي للفروق بين D والقيم الحرجة لاختبار نيمينييه (6.06) وهو أقل من قيم المتوسطات الحسابية للفروق بين D والقيم الحرجة لأساليب المقارنات البعدية الباقية، مما يدل على أن اختبار نيمينييه يعطي حدوداً للثقة أكبر من بقية أساليب المقارنات البعدية مما يعطي فرصة أقل لإعطاء دلالة للمقارنات الثنائية بين متوسطات رتب المجموعات، بينما كانت قيمة المتوسط الحسابي للفروق بين D والقيم الحرجة لاختبار كونوفير (15.95) ، ومن الملاحظ أن المتوسط للفروق بين D والقيم الحرجة لاختبار كونوفير أكبر من بقية أساليب المقارنات البعدية ، وهذا يعني أن اختبار كونوفير يعطي حدوداً للثقة أقل من حدود الثقة المحسوبة باستخدام باقي أساليب المقارنات البعدية اللابارامترية، وهذا بالطبع يؤدي إلى أن اختبار كونوفير يعطي فرصة أكبر لإعطاء دلالة للمقارنات الثنائية بين متوسطات رتب درجات المجموعات .

ويوضح شكل (4) متوسطات الفروق بين D والقيم الحرجة لأساليب المقارنات البعدية اللابارامترية لاختبار كروسكال-واليس في حالة ما إذا كان مستوى الدلالة 0.01، 0.05 وعدد المجموعات اربعة .



شكل (4) متوسطات الفروق بين D والقيم الحرجة لأساليب المقارنات البعدية لاختبار كروسكال-واليس في حالة ما إذا كان مستوى الدلالة 0.01، 0.05 وعدد المجموعات اربعة .

وقد بينت نتائج الفرض الثاني أن القيم الحرجة لأساليب المقارنات البعدية تتأثر جميعها تأثيراً كبيراً بحجم العينة، وقد لوحظ ان اختبار نيمينييه واختبار دن من أكثر أساليب المقارنات البعدية تتأثر بحجم العينة وذلك إذا

كان مستوى الدلالة المستخدم 0.05 أو 0.01، كما بينت نتائج الفرض الثاني أن اختبار كونوفير من أكثر أساليب المقارنات البعدية وقوعا في الخطأ من النوع الأول وهذا ما أكدته كل من (Shuster & Boyett, 1979)، (Campbell & Skillings, 1985)، (Neuhäuser, & Bretz, 2001).

أيضا فقد بينت نتائج الفرض الثاني أن اختبار نيمينييه من أكثر أساليب المقارنات البعدية اللابارامترية تحفظا، حيث أن متوسط القيم الحرجة لاختبار نيمينييه كان أقل من متوسطات القيم الحرجة لأساليب المقارنات البعدية اللابارامترية الأخرى وذلك سواء كان مستوى الدلالة 0.05 أو 0.01.

نتائج الفرض الثالث ومناقشتها:

ينص الفرض الثالث على أنه " لا توجد علاقة ارتباطية دالة إحصائياً بين حجم العينة والقيم الحرجة لأساليب المقارنات المتعددة البعدية اللابارامترية لاختبار فريدمان عندما يكون عدد المجموعات خمسة ومستوى الدلالة 0.05، 0.01".

وللتحقق من صحة هذا الفرض تم حساب قيم الفروق بين الفرق المطلق D (متوسط رتب المجموعة i – متوسط رتب المجموعة j) وبين القيم الحرجة لأساليب المقارنات المتعددة البعدية المستخدمة في حالة اختبار فريدمان وهي (اختبار نيمينييه، اختبار كونوفير) عند مستويي دلالة 0.05، 0.01 وعند عدد مجموعات يساوي خمسة، وتم حساب معامل الارتباط بين ناتج الفروق لكل اختبار من هذه الأساليب وحجم العينة باستخدام معامل ارتباط بيرسون ويوضح جدول (5) معاملات ارتباط بيرسون ودلالاتها الإحصائية:

ويتضح من جدول (5) أن جميع معاملات الارتباط بين ناتج الفروق لكل اختبار من هذه الأساليب وحجم العينة كانت دالة إحصائياً عند مستوى 0.01، مما يدل على أن قيم الفروق بين الفرق المطلق D والقيم الحرجة لأساليب المقارنات البعدية اللابارامترية تتأثر بحجم العينة.

جدول (5)

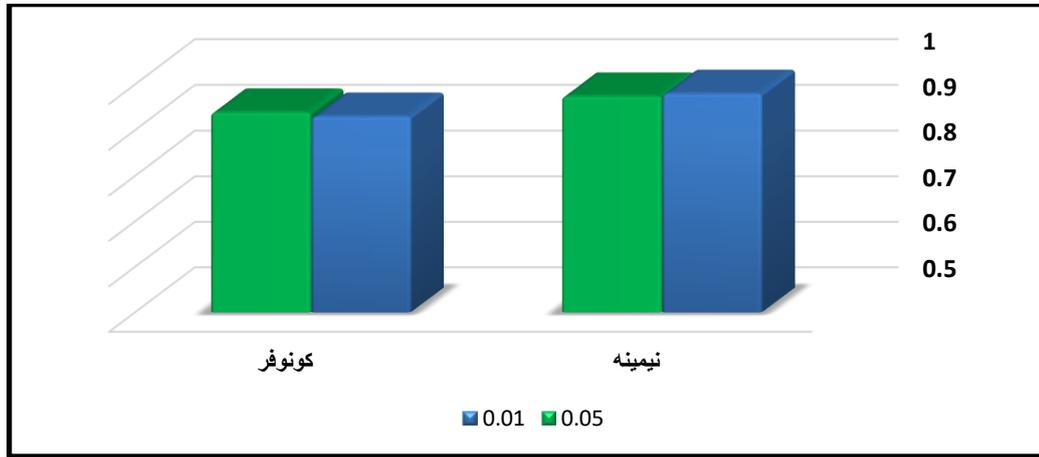
معاملات الارتباط بين ناتج الفروق لأساليب المقارنات المتعددة اللابارامترية لاختبار فريدمان وحجم العينة عند استخدام مستوى دلالة 0.05، 0.01 وخمسة مجموعات.

معامل ارتباط بيرسون		أساليب المقارنات البعدية
مستوي دلالة 0.01	مستوي دلالة 0.05	
**0.980	**0.974	نيمينييه
**0.930	**0.939	كونوفير

** دالة عند مستوى 0.01

ومن الملاحظ أن جميع معاملات الارتباط كانت موجبة مما يدل على أن قيم الفروق بين الفرق المطلق D والقيم الحرجة لأساليب المقارنات البعدية اللابارامترية تزيد كلما زاد حجم العينة، مما يعني أنه كلما زاد حجم العينة كلما زادت قيم الفروق بين الفرق المطلق D وبين القيم الحرجة لأساليب المقارنات البعدية لاختبار فريدمان، مما يعطي فرصة أكبر لإعطاء دلالة للمقارنات الثنائية بين متوسطات رتب درجات المجموعات.

كذلك يتضح من جدول (5) أن اختبار نيمينه كان أكثر أساليب المقارنات المتعددة البعدية ارتباطاً بحجم العينة، سواء أكان مستوى الدلالة المستخدم 0.05 أو 0.01 حيث بلغت قيمة معامل الارتباط بين ناتج الفروق لاختبار نيمينه وحجم العينة (0.974) وذلك عند استخدام مستوى دلالة 0.05، أيضاً فقد بلغت قيمة معامل الارتباط بين ناتج الفروق لاختبار نيمينه وحجم العينة (0.980) وذلك عند استخدام مستوى دلالة 0.01، مما يدل على أن اختبار نيمينه أكثر تأثراً بحجم العينة من اختبار كونوفير، حيث كان معامل الارتباط بين حجم العينة وناتج الفروق لاختبار كونوفير أقل من الارتباط بين حجم العينة وناتج الفروق لاختبار نيمينه وذلك في حالة استخدام مستوى دلالة 0.05 أو 0.01 حيث بلغت قيمة معامل الارتباط بين ناتج الفروق لاختبار كونوفير وحجم العينة (0.939) عند استخدام مستوى دلالة 0.05، بينما بلغت قيمة معامل الارتباط بين ناتج الفروق لاختبار كونوفير وحجم العينة (0.930) عند استخدام مستوى دلالة 0.01 مما يدل على أن اختبار كونوفير أقل تأثراً بحجم العينة من اختبار نيمينه وفيما يلي ويوضح شكل (5) معاملات ارتباط الفروق الناتجة لأساليب المقارنات البعدية لاختبار فريدمان وحجم العينة عند استخدام مستوى دلالة 0.05 ، 0.01 وخمس مجموعات .



شكل (5) معاملات ارتباط الفروق لأساليب المقارنات البعدية لاختبار فريدمان وحجم العينة عند استخدام مستوى دلالة 0.05 ، 0.01 وخمس مجموعات .

وقد تم حساب المتوسطات الحسابية والانحرافات المعيارية للفروق بين D والقيم الحرجة لأساليب المقارنات البعدية لاختبار فريدمان كما هو موضح بالجدول التالي :

جدول (6)

المتوسطات الحسابية والانحرافات المعيارية للفروق بين D والقيم الحرجة لأساليب المقارنات البعدية لاختبار فريدمان عند استخدام مستوى دلالة 0.05 ، 0.01 وخمس مجموعات .

مستوي الدلالة المستخدم 0.01		مستوي الدلالة المستخدم 0.05		أساليب المقارنات البعدية
الانحراف المعياري	المتوسط الحسابي	الانحراف المعياري	المتوسط الحسابي	
0.602	2.080	0.513	2.356	نيمينييه
0.478	3.096	0.378	3.28	كونوفر

ويتضح من جدول (6) ما يلي :

أ- في حالة ما إذا كان مستوى الدلالة المستخدم 0.05 :

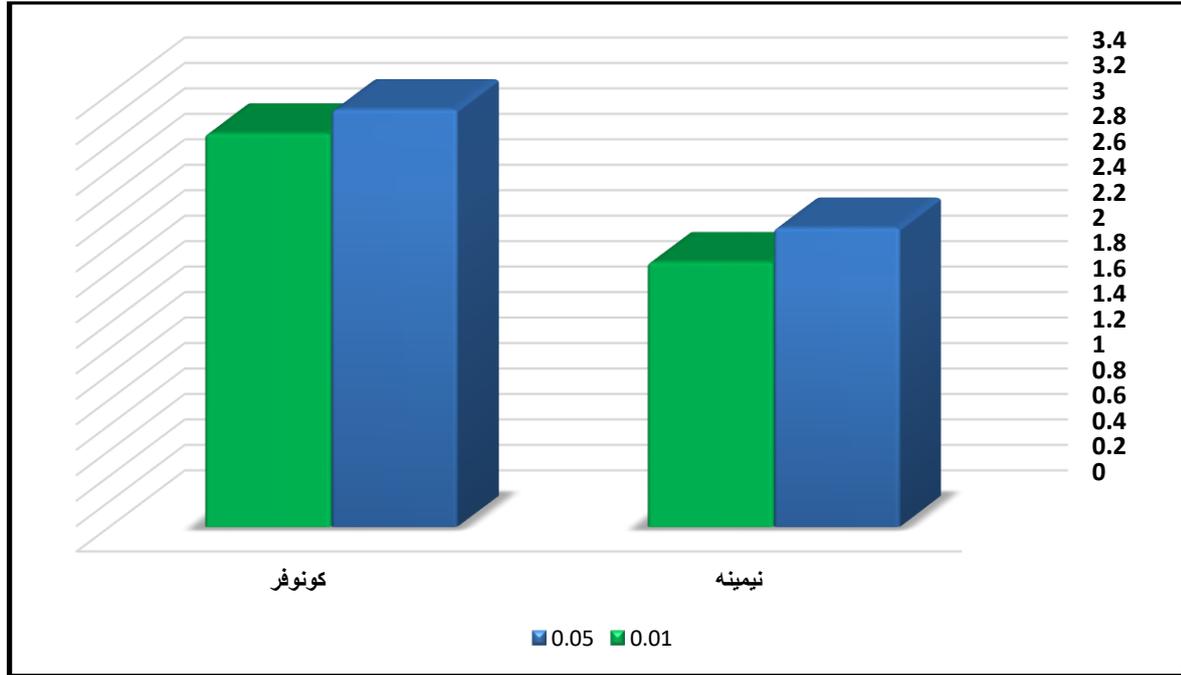
يتضح من جدول (6) أن قيمة المتوسط الحسابي للفروق بين D والقيم الحرجة لاختبار نيمينييه (2.35) هو أقل من قيم المتوسط الحسابي للفروق بين D والقيم الحرجة لأسلوب كونوفر ، مما يدل على أن اختبار نيمينييه يعطي حدوداً للثقة أكبر من أسلوب كونوفر مما يعطي فرصة أقل لإعطاء دلالة للمقارنات الثنائية بين متوسطات رتب المجموعات ، بينما كانت قيمة المتوسط الحسابي للفروق بين D والقيم الحرجة لاختبار كونوفر (3.28) ، ومن الملاحظ أن المتوسط للفروق بين D والقيم الحرجة لاختبار كونوفر أكبر من أسلوب نيمينييه، وهذا يعني أن اختبار كونوفر يعطي حدوداً للثقة أقل من حدود الثقة المحسوبة باستخدام أسلوب نيمينييه، وهذا بالطبع يؤدي إلى أن اختبار كونوفر يعطي فرصة أكبر لإعطاء دلالة للمقارنات الثنائية بين متوسطات رتب درجات المجموعات .

ب- في حالة ما إذا كان مستوى الدلالة المستخدم 0.01:

يتضح من جدول (6) أن قيمة المتوسط الحسابي للفروق بين D والقيم الحرجة لاختبار نيمينييه (2.08) وهو أقل من قيم المتوسط الحسابي للفروق بين D والقيم الحرجة لأسلوب كونوفر ، مما يدل على أن اختبار نيمينييه يعطي حدوداً للثقة أكبر من أسلوب كونوفر مما يعطي فرصة أقل لإعطاء دلالة للمقارنات الثنائية بين متوسطات رتب المجموعات ، بينما كانت قيمة المتوسط الحسابي للفروق بين D والقيم الحرجة لاختبار كونوفر (3.09) ، ومن الملاحظ أن المتوسط للفروق بين D والقيم الحرجة لاختبار كونوفر أكبر من أسلوب نيمينييه، وهذا يعني أن اختبار كونوفر يعطي حدوداً للثقة أقل من حدود الثقة المحسوبة باستخدام أسلوب نيمينييه، وهذا بالطبع يؤدي إلى أن اختبار كونوفر يعطي فرصة أكبر لإعطاء دلالة للمقارنات الثنائية بين متوسطات رتب درجات المجموعات .

ويوضح شكل (6) متوسطات الفروق بين D والقيم الحرجة لأساليب المقارنات البعدية اللابارامترية

لاختبار فريدمان في حالة ما إذا كان مستوى الدلالة 0.05، 0.01 وعدد المجموعات خمسة .



شكل (6) متوسطات الفروق بين D والقيم الحرجة لأساليب المقارنات البعدية لاختبار فريدمان في حالة ما إذا كان مستوى الدلالة 0.05، 0.01 وعدد المجموعات خمسة .

وقد بينت نتائج الفرض الثالث أن القيم الحرجة لأساليب المقارنات البعدية تتأثر جميعها تأثيراً كبيراً بحجم العينة، وقد لوحظ أن اختبار نيمينه يتأثر بحجم العينة أكثر من اختبار كونوفر وذلك سواء كان مستوى الدلالة المستخدم 0.05 أو 0.01، كما بينت نتائج الفرض الثالث أن اختبار كونوفر أكثر وقوعاً في الخطأ من النوع الأول من اختبار نيمينه، حيث أنه كان متوسط الفروق بين D والقيم الحرجة لاختبار كونوفر أكبر من أسلوب نيمينه سواء كان مستوى الدلالة 0.05 أو 0.01 .

أيضاً فقد بينت نتائج الفرض الثالث أن اختبار نيمينه أكثر تحفظاً من اختبار كونوفر، حيث أن متوسط الفروق بين D والقيم الحرجة لاختبار نيمينه كان أقل من أسلوب كونوفر وذلك سواء كان مستوى الدلالة 0.05 أو 0.01 .

رابعاً-نتائج الفرض الرابع ومناقشتها :

ينص الفرض الرابع على أنه " لا توجد علاقة ارتباطية دالة إحصائياً بين حجم العينة والقيم الحرجة لأساليب المقارنات المتعددة البعدية اللابارمترية لاختبار فريدمان عندما يكون عدد المجموعات ستة ومستوى الدلالة 0.05، 0.01" .

وللتحقق من صحة هذا الفرض تم حساب قيم الفروق بين الفرق المطلق D (متوسط رتب المجموعة) – متوسط رتب المجموعة) وبين القيم الحرجة لأساليب المقارنات المتعددة البعدية المستخدمة في حالة اختبار فريدمان وهي (اختبار نيمينه، اختبار كونوفر) عند مستويي دلالة 0.05، 0.01 وعند عدد مجموعات يساوي

سته ، وتم حساب معامل الارتباط بين ناتج الفروق لكل اختبار من هذه الأساليب وحجم العينة باستخدام معامل ارتباط بيرسون ويوضح جدول (7) معاملات ارتباط بيرسون ودلالاتها الإحصائية :

جدول (7)

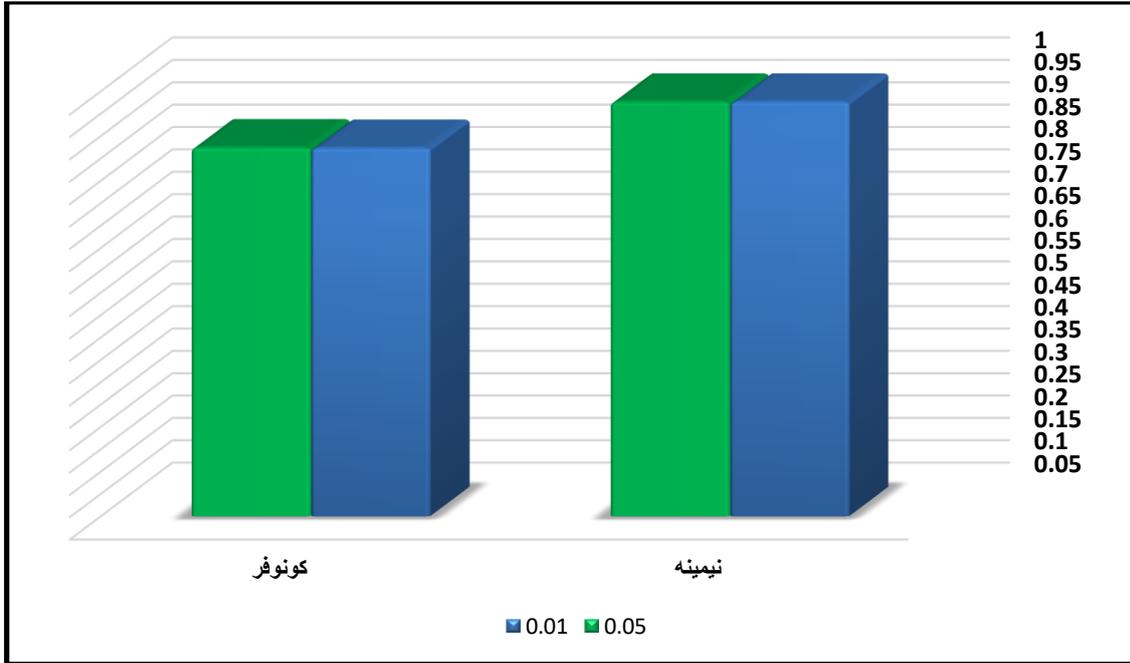
معاملات الارتباط بين ناتج الفروق لأساليب المقارنات المتعددة اللابارامترية لاختبار فريدمان وحجم العينة عند استخدام مستوى دلالة 0.05، 0.01 و ست مجموعات .

معامل ارتباط بيرسون		أساليب المقارنات البعدية
مستوي دلالة 0.01	مستوي دلالة 0.05	
**0.976	**0.976	نيمينه
*0.873	*0.874	كونوفر

* دالة عند مستوى 0.05 ** دالة عند مستوى 0.01

ويتضح من جدول (7) أن جميع معاملات الارتباط بين ناتج الفروق لكل اختبار من هذه الأساليب وحجم العينة كانت دالة إحصائياً عند مستويي دلالة 0.05 و 0.01، مما يدل على أن قيم الفروق بين الفرق المطلق D وبين القيم الحرجة لأساليب المقارنات البعدية اللابارامترية تتأثر بحجم العينة ، ومن الملاحظ أن جميع معاملات الارتباط كانت موجبة مما يدل على أن قيم الفروق بين الفرق المطلق D وبين القيم الحرجة لأساليب المقارنات البعدية اللابارامترية تزيد كلما زاد حجم العينة ، مما يعني أنه كلما زاد حجم العينة كلما زادت قيم الفروق بين الفرق المطلق D وبين القيم الحرجة لأساليب المقارنات البعدية لاختبار فريدمان ، مما يعطي فرصة أكبر لإعطاء دلالة للمقارنات الثنائية بين متوسطات رتب درجات المجموعات .

كذلك يتضح من جدول (7) أن اختبار نيمينه كان أكثر أساليب المقارنات المتعددة البعدية ارتباطاً بحجم العينة، سواء أكان مستوى الدلالة المستخدم 0.05 أو 0.01 حيث بلغت قيمة معامل الارتباط بين ناتج الفروق لاختبار نيمينه وحجم العينة (0.976) وذلك عند استخدام مستوى دلالة 0.05، ومستوى دلالة 0.01 أيضاً، مما يدل على أن اختبار نيمينه أكثر تأثراً بحجم العينة من اختبار كونوفر، حيث كان معامل الارتباط بين حجم العينة وناتج الفروق لاختبار كونوفر أقل من الارتباط بين حجم العينة وناتج الفروق لاختبار نيمينه وذلك في حالة استخدام مستوى دلالة 0.05 أو 0.01 حيث بلغت قيمة معامل الارتباط بين ناتج الفروق لاختبار كونوفر وحجم العينة (0.874) عند استخدام مستوى دلالة 0.05، بينما بلغت قيمة معامل الارتباط بين ناتج الفروق لاختبار كونوفر وحجم العينة (0.873) عند استخدام مستوى دلالة 0.01 مما يدل على أن اختبار كونوفر أقل تأثراً بحجم العينة من اختبار نيمينه وفيما يلي ويوضح شكل (7) معاملات ارتباط الفروق الناتجة لأساليب المقارنات البعدية لاختبار فريدمان وحجم العينة عند استخدام مستوى دلالة 0.05 ، 0.01 و ست مجموعات .



شكل (7) معاملات ارتباط الفروق لأساليب المقارنات البعدية لاختبار فريدمان وحجم العينة عند استخدام مستوى دلالة 0.05 ، 0.01 وست مجموعات .

وقد تم حساب المتوسطات الحسابية والانحرافات المعيارية للفروق بين D والقيم الحرجة لأساليب المقارنات البعدية لاختبار فريدمان كما هو موضح بالجدول التالي :

جدول (8)

المتوسطات الحسابية والانحرافات المعيارية للفروق بين D والقيم الحرجة لأساليب المقارنات البعدية لاختبار فريدمان عند استخدام مستوى دلالة 0.05 ، 0.01 وست مجموعات .

مستوي الدلالة المستخدم 0.01		مستوي الدلالة المستخدم 0.05		أساليب المقارنات البعدية
الانحراف المعياري	المتوسط الحسابي	الانحراف المعياري	المتوسط الحسابي	
0.701	2.693	0.6145	3.016	نيميئه
0.360	4.066	0.297	4.243	كوففر

ويتضح من جدول (8) ما يلي :

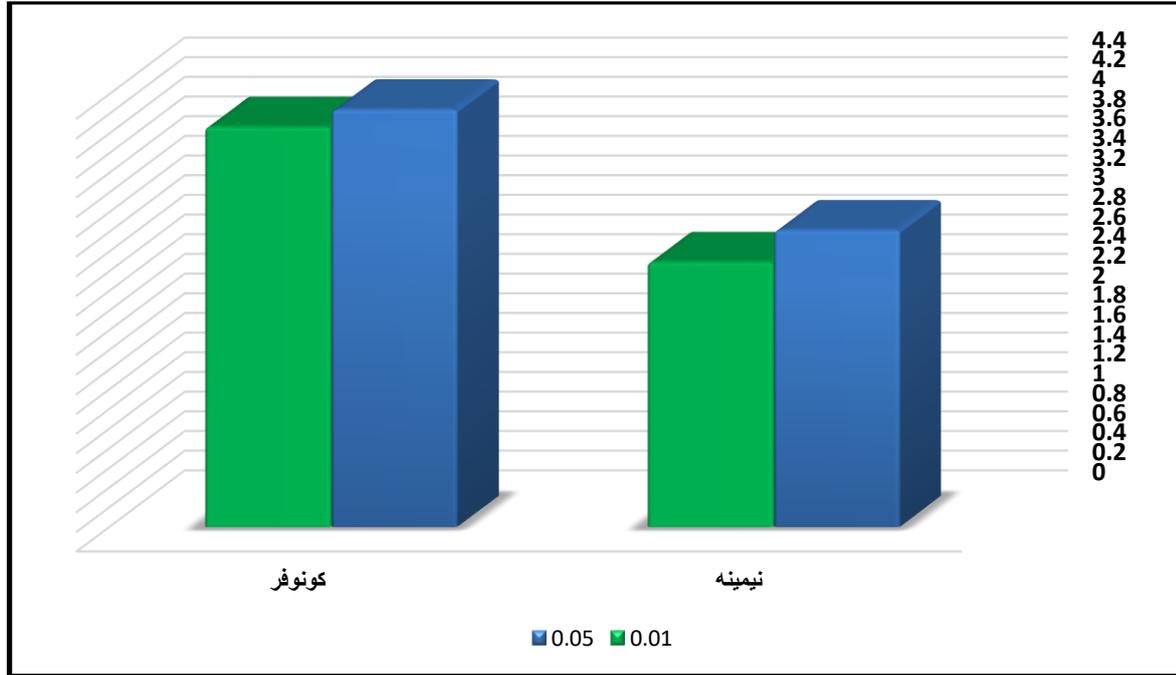
أ-في حالة ما إذا كان مستوى الدلالة المستخدم 0.05:

يتضح من جدول (8) أن قيمة المتوسط الحسابي للفروق بين D والقيم الحرجة لاختبار نيمينييه (3.01) وهو أقل من قيم المتوسط الحسابي للفروق بين D والقيم الحرجة لأسلوب كونوفير ، مما يدل على أن اختبار نيمينييه يعطي حدوداً للثقة أكبر من أسلوب كونوفير مما يعطي فرصة أقل لإعطاء دلالة للمقارنات الثنائية بين متوسطات رتب المجموعات، بينما كانت قيمة المتوسط الحسابي للفروق بين D والقيم الحرجة لاختبار كونوفير (4.24) ، ومن الملاحظ أن المتوسط للفروق بين D والقيم الحرجة لاختبار كونوفير أكبر من أسلوب نيمينييه، وهذا يعني أن اختبار كونوفير يعطي حدوداً للثقة أقل من حدود الثقة المحسوبة باستخدام أسلوب نيمينييه، وهذا بالطبع يؤدي إلى أن اختبار كونوفير يعطي فرصة أكبر لإعطاء دلالة للمقارنات الثنائية بين متوسطات رتب درجات المجموعات .

ب-في حالة ما إذا كان مستوى الدلالة المستخدم 0.01 :

يتضح من جدول (8) أن قيمة المتوسط الحسابي للفروق بين D والقيم الحرجة لاختبار نيمينييه (2.69) وهو أقل من قيم المتوسط الحسابي للفروق بين D والقيم الحرجة لأسلوب كونوفير ، مما يدل على أن اختبار نيمينييه يعطي حدوداً للثقة أكبر من أسلوب كونوفير مما يعطي فرصة أقل لإعطاء دلالة للمقارنات الثنائية بين متوسطات رتب المجموعات، بينما كانت قيمة المتوسط الحسابي للفروق بين D والقيم الحرجة لاختبار كونوفير (4.06) ، ومن الملاحظ أن المتوسط للفروق بين D والقيم الحرجة لاختبار كونوفير أكبر من أسلوب نيمينييه، وهذا يعني أن اختبار كونوفير يعطي حدوداً للثقة أقل من حدود الثقة المحسوبة باستخدام أسلوب نيمينييه، وهذا بالطبع يؤدي إلى أن اختبار كونوفير يعطي فرصة أكبر لإعطاء دلالة للمقارنات الثنائية بين متوسطات رتب درجات المجموعات .

ويوضح شكل (8) متوسطات الفروق بين D والقيم الحرجة لأساليب المقارنات البعدية اللابارامترية لاختبار فريدمان في حالة ما إذا كان مستوى الدلالة 0.05، 0.01 وعدد المجموعات ستة .



شكل (8) متوسطات الفروق بين D والقيم الحرجة لأساليب المقارنات البعدية لاختبار فريدمان في حالة ما إذا كان مستوى الدلالة 0.05، 0.01 وعدد المجموعات ستة .

وقد بينت نتائج الفرض الرابع أن القيم الحرجة لأساليب المقارنات البعدية تتأثر جميعها تأثيراً كبيراً بحجم العينة، وقد لوحظ أن اختبار نيمينيه يتأثر بحجم العينة أكثر من اختبار كونوفر وذلك سواء كان مستوى الدلالة المستخدم 0.05 أو 0.01، كما بينت نتائج الفرض الرابع أن اختبار نيمينيه أقل وقوعاً في الخطأ من النوع الأول من اختبار كونوفر وهذا ما أكدته كل من (Shuster & Boyett, 1979)، (Campbell & Skillings, 1985)، (Neuhäuser, & Bretz, 2001).

أيضاً فقد بينت نتائج الفرض الرابع أن اختبار نيمينيه أكثر تحفظاً من اختبار كونوفر، حيث أن متوسط الفروق بين D والقيم الحرجة لاختبار نيمينيه كان أقل من متوسط القيم الحرجة لاختبار كونوفر وذلك سواء كان مستوى الدلالة 0.05 أو 0.01 .

ملخص نتائج الدراسة :

أظهرت نتائج فروض الدراسة الحالية أن أساليب المقارنات البعدية اللابارامترية تتأثر جميعها بحجم العينة، فكلما زاد حجم العينة زادت الفروق بين متوسطات رتب المجموعات والقيم الحرجة المقارنة بها لصالح الفروق بين المتوسطات مما يعطي دلالة للمقارنات الثنائية بين متوسطات رتب المجموعات وهذا يتعلق بأساليب المقارنات البعدية سواء لاختبار كروسكال واليس في حالة استخدام مستويات دلالة (0.05 أو 0.10) واستخدام عدد مجموعات (3 أو 4) ، أو اختبار فريدمان في حالة استخدام مستويات دلالة (0.05 أو 0.01) واستخدام عدد مجموعات (5 أو 6) على حد سواء .

بالنسبة إلى أساليب المقارنات البعدية لاختبار كروسكال واليس، فقد أظهرت نتائج الدراسة الحالية أن متوسط الفروق بين D والقيم الحرجة لاختبار كونوفير كان أكبر من متوسط الفروق بين D والقيم الحرجة لاختبار نيميني وذن وذلك في حالة ما إذا كان مستوى الدلالة المستخدم 0.05 او 0.01 وعند استخدام عدد مجموعات 3, 4.

بالنسبة إلى أساليب المقارنات البعدية لاختبار فريدمان، فقد أظهرت نتائج الدراسة الحالية أن متوسط الفروق بين D والقيم الحرجة لاختبار كونوفير كان أكبر من متوسط الفروق بين D والقيم الحرجة لاختبار نيميني وذلك في حالة ما إذا كان مستوى الدلالة المستخدم 0.05 او 0.01 وعند استخدام عدد مجموعات 5, 6 .

توصيات الدراسة:

- 1- استخدام اختبار نيميني لإجراء المقارنات البعدية لاختبار كروسكال-واليس وذلك في حالة ما إذا كان عدد المجموعات 3 او 4. وذلك لكونه أقل وقوعاً في الخطأ من النوع الأول مقارنة بباقي أساليب المقارنات البعدية المتعددة اللابارامترية
- 2- استخدام اختبار نيميني لإجراء المقارنات البعدية لاختبار فريدمان وذلك في حالة ما إذا كان عدد المجموعات 5, 6 حيث أن اختبار نيميني أقل وقوعاً في الخطأ من النوع الأول من اختبار كونوفير.
- 3- لوحظ في استخدام اختبار نيميني، أنه أكثر تحفظاً في قبول دلالة المقارنات الثنائية بين متوسطات رتب المجموعات سواء مع اختبار كروسكال واليس أو اختبار فريدمان .
- 4- عدم استخدام اختبار كونوفير لإجراء المقارنات البعدية اللابارامترية لاختبار كروسكال-واليس واختبار فريدمان، وذلك يعود لأن هذا الاختبار يعد من أكثر أساليب المقارنات البعدية اللابارامترية وقوعاً في الخطأ من النوع الأول.

- البحوث المقترحة:

- 1- إجراء دراسة للتعرف على أثر القيم المتطرفة على القيم الحرجة لأساليب المقارنات المتعددة اللابارامترية وذلك في ضوء عدد المجموعات ومستوى الدلالة.
- 2- إجراء دراسة للتعرف على واقع استخدام أساليب المقارنات البعدية البارامترية واللابارامترية في البحوث التربوية المنشورة بجامعة ام القرى.
- 3- إجراء دراسة للتعرف على أثر حجم العينة على القوة الإحصائية لأساليب المقارنات البعدية اللابارامترية.

قائمة المراجع:

- أبو حطب، فؤاد وصادق، امال. (1991)، **مناهج البحث وطرق التحليل الاحصائي في العلوم النفسية والتربوية والاجتماعية**. القاهرة: مكتبة الانجلو المصرية.
- ادم، امين إبراهيم (2005م). **المبادئ الأساسية الإحصائية في الطرق التطبيقية للمعلمية**. مكة المكرمة: مكتبة الملك فهد الوطنية.
- بابطين، عادل احمد. (1422). **مشكلات الدلالة الإحصائية في البحث التربوي وحلول بديلة**. رسالة ماجستير، كلية التربية، جامعة ام القرى : مكة .

- البارقي , طلال . (2012) . واقع الدلالة الإحصائية والدلالة العملية للبحوث المنشورة بمجلة جامعة أم القرى للعلوم التربوية والاجتماعية والإنسانية في المدة من 1425-1430 . رسالة ماجستير غير منشورة . قسم علم النفس , كلية التربية , جامعة أم القرى : مكة .
- باهي، مصطفى حسين. (2010). العلاقة بين الدلالة الإحصائية وحجم التأثير في البحوث التربوية والنفسية. بحث مقدم في مؤتمر مستقبل اعداد المعلم في كليات التربية وجهود الجمعيات العلمية في عمليات التطوير بالعالم العربي، مصر.
- بدر، سالم، عبابنه، عماد (2007). مبادئ الإحصاء الوصفي والاستدلالي. عمان: دار المسيرة للنشر والتوزيع والطباعة.
- حبيب، مجدي عبد الكريم. (2000). الإحصاء اللابارامتري الحديث في العلوم السلوكية. الطبعة الأولى، القاهرة: مكتبة النهضة المصرية.
- حمدان، عادل سمير. (2015) . علاقة حجم العينة بالقيم الحرجة لأساليب المقارنات المتعددة وذلك في ضوء عدد المجموعات ومستوى الدلالة. رسالة ماجستير غير منشورة، قسم علم النفس التربوي، كلية التربية، جامعة أسيوط: مصر.
- الخالدي، محمد عمير (1433هـ). المقارنة بين نتائج بعض الأساليب الإحصائية العملية واللامعلمية في ضوء انتهاك افتراض تجانس التباين. رسالة ماجستير غير منشورة، قسم علم النفس، كلية التربية، جامعة أم القرى: مكة
- الدردير، عبد المنعم. (2006) . الإحصاء البارامتري واللابارامتري في اختبار فروض البحوث النفسية والتربوية والاجتماعية. القاهرة: عالم الكتب.
- الشريني، زكريا. (2001) . الإحصاء اللابارامتري في العلوم النفسية والتربوية والاجتماعية. القاهرة: مكتبة الأنجلو المصرية.
- الشريني، زكريا. (2007) . الإحصاء وتصميم التجارب في البحوث النفسية والتربوية والاجتماعية. القاهرة: مكتبة الأنجلو المصرية.
- الشمراي، موسى (2000). مشكلات استخدام تحليل التباين الأحادي والمقارنات البعدية وطرق علاجها. رسالة ماجستير , غير منشورة , قسم علم النفس، كلية التربية، جامعة أم القرى: مكة .
- العنبي، أشرف (1433). دراسة تقويمية لصحة استخدام أسلوب تحليل التباين في رسائل الماجستير والدكتوراه في كلية التربية في جامعة أم القرى عبر الفترة الزمنية 1421-1430 هـ. رسالة ماجستير غير منشورة , قسم علم النفس، كلية التربية، جامعة أم القرى: مكة.
- علام، صلاح الدين. (2007). القياس والتقويم التربوي في العملية التدريسية. عمان: دار المسيرة للنشر والتوزيع.
- علام، صلاح الدين. (2010) . الأساليب الإحصائية الاستدلالية في تحليل بيانات البحوث النفسية والتربوية والاجتماعية (البارامترية واللابارامترية). القاهرة: دار الفكر العربي.
- علام، صلاح الدين. (2011). القياس والتقويم التربوي والنفسية: اساسياته وتطبيقاته وتوجهاته المعاصرة. الطبعة الخامسة، القاهرة: دار الفكر العربي.

علي، عماد وأحمد ، محمد و عبد المحسن ، علي (2013) . الممارسات الإحصائية الخاطئة في حساب صدق وثبات الأدوات في الرسائل العلمية ، مجلة كلية التربية ، جامعة أسيوط ، 29 (3) ، 325-348 .

غيث، البحر (2012). الاختبارات اللامعلمية وتطبيقاتها في برنامج SPSS. حلب. شركة تبارك للاستشارات ونظم المعلومات.

الكناني، ممدوح (2002م). الإحصاء الوصفي والاستدلالي في العلوم السلوكية والاجتماعية. القاهرة: دار النشر للجامعات.

محمد، عزت عبد الحميد. (2011). الإحصاء النفسي والتربوي: تطبيقات باستخدام برنامج SPSS 18، القاهرة، دار الفكر العربي.

المنيزل ، عبد الله ، العنوم ، عدنان . (2010) . مناهج البحث في العلوم التربوية والنفسية . الأردن : إثراء للنشر والتوزيع .

النبهان، موسى سعيد. (2004). أساسيات القياس في العلوم السلوكية. الطبعة الأولى، عمان: دار الشروق للنشر والتوزيع.

Andy, F. & Jeremy, M. (2010) . **Discovering Statistics Using SAS, London :** SAGE Publication.

Al-Subaihi, A. A. (2000). **A Monte Carlo study of the Friedman and Conover tests in the single-factor repeated measures design.** *Journal of Statistical Computation and Simulation*, 65 (1-4), 203-223.

Barnett, J., Mclean, J. (1999). **Choosing a Multiple Comparison Procedure Based on Alpha**, A paper presented at the Annual Meeting of the American Educational Research Association, Montreal, Ontario, CANADA.

Campbell, G. & Skillings, J. H. (1985). **Nonparametric stepwise multiple comparison procedures.** *Journal of the American Statistical Association*, 80 (392), 998-1003.

Cicchetti, D. V. (1994). **Multiple comparison methods: establishing guidelines for their valid application in neuropsychological research.** *Journal of Clinical and Experimental Neuropsychology*, 16(1), 155-161.

Daniel, W. W. (1990). *Applied nonparametric statistics.*

- David , J . (2014) . **Parametric and non-parametric statistical procedures** ,
Western Connecticut state : Caapman & Hal , CRC .
- Demšar, J. (2006). **Statistical comparisons of classifiers over multiple data sets**. Journal of Machine learning research, 7(Jan), 1-30.
- Dinno, A. (2015). **Nonparametric pairwise multiple comparisons in independent groups using Dunn's test**. Stata Journal.
- Favato, G., & Mills, R. (2007). **Thinking the Unthinkable: Modern Non-Parametric Re-sampling Methods**. Henley Management College, green lands, electronic copy available at: <http://ssrn.com/abstract=1012661>.
- Field, A. (2013). **Discovering statistics using IBM SPSS statistics**. Sage.
- Gibbons, J. D. (1993). **Nonparametric statistics: An introduction** (No. 90). Sage.
- Glantz, S. A. (2011). **Primer of biostatistics**. New York: McGraw-Hill.
- Glass, G. V., & Hopkins, K. D. (1996). **Statistical methods in education and psychology**. Englewood Cliffs, NJ: Prentice-Hall.
- Gravetter, F. J., & Wallnau, L. B. (2014). **Essentials of statistics for the behavioral sciences. 8th ed, Pacific Grove**, CA: Brooks. Cole Publishing Company. Copyright restrictions may apply.
- Hollander, M., & Wolfe, D. A. (1999). **Nonparametric Statistical Methods: By Myles Hollander, Douglas A. Wolfe**. J. Wiley.
- Hollander, M., Wolfe, D. A., & Chicken, E. (2013). **Nonparametric statistical methods**. John Wiley & Sons.
- Hsu, J. (1996). **Multiple comparisons: theory and methods**. CRC Press.

Huston, H. L. (1993). **Meaningfulness, Statistical Significance, Effect Size, and Power Analysis: A General Discussion with Implications for MANOVA**. Paper Presented at Annual Meeting of the Mid-South Educational Research Association (2nd, New Orleans, LA, November 9-12, 1993).

Israel, D. (2009). **Data analysis in business research: A step-by-step nonparametric approach**. Sage Publications.

Jaccard, J., Becker, M. A., & Wood, G. (1984). **Pairwise multiple comparison procedures: a review**. *Psychological Bulletin*, 96(3), 589.

James, P., John, A. & Martha, L. (2002). **Multiple Comparison Methods for Means**. *SIAM Review*, 44(2), 259-278.

Kaplan, A , Marleen . (1983) . **A COMPARISON OF THE KRUSKAL-WALLIS TEST WITH ANOVA UNDER VIOLATIONS OF HOMOGENEITY OF VARIANCE**, University Microfilms International 300 N Zeeb Road, Ann Arbor, MI4810.

Kellow, J. T. (1998). **Beyond statistical significant tests: The importance of using other estimates of treatment effects to interpret evaluation results**. *The American Journal of Evaluation*, 19 (1), 123-134.

Kubiszyn, T., & Borich, G. (1990). *Educational testing and measurement*. Harper Collins Publishers.

Kurtz, T., Link, R., Tukey, J. & Wallace , D. (1965). **Short-cut multiple comparisons for Balanced single and double classifications, Part I, Results** , *Technometrics* .

Lin, F. A., & Haseman, J. K. (1978). **An evaluation of some nonparametric multiple comparison procedures by Monte Carlo methods**. *Communications in Statistics-Simulation and Computation*, 7 (2), 117-128.

- Mary, L. (2011). **Multiple comparison analysis testing in ANOVA**, Biochemia Medica , 21 (3), 203-209.
- McLEAN, J. E., & Ernest, J. M. (1998). **The role of statistical significance testing in educational research**. Research in the Schools, 5(2), 15-22.
- Neuhäuser, M., & Bretz, F. (2001). **Nonparametric All-Pairs Multiple Comparisons**. Biometrical Journal, 43(5), 571-580.
- Nix, T. W., & Barnette, J. J. (1998). **A review of hypothesis testing revisited: Rejoinder to Thompson, Knapp, and Levin**. Research in the Schools, 5(2), 55-57.
- Pohlert, Thorsten. (2016). **the pair wise multiple comparison of mean Ranks package**. Creative commons.org/licenses.
- Richard , M . (2004) . **Statistical Analysis and Data Display**, New York: Springer Science Business Media Inc .
- Richardson, A. M. (2015). **Nonparametric Statistics: A Step-by-Step Approach**. International Statistical Review, 83(1), 163-164.
- Sax, G. (1986). **Multiple comparisons** (No. 61). Sage.
- Schneider, A. L., & Darcy, R. E. (1984). **Policy implications of using significance tests in evaluation research**. Evaluation Review, 8(4), 573-582.
- Shuster, J. J., & Boyett, J. M. (1979). **Nonparametric multiple comparison procedures**. Journal of the American Statistical Association, 74 (366a), 379-382.
- Siegel S, Castellan, J. (1988). **Nonparametric Statistics for The behavioral Sciences**. Second edition. McGraw-Hill, New York.
- Siegel, S., & Castellan, N. J. (1988). **Nonparametric systems for the behavioural sciences**. McGraw Hill International Editions.

Snyder, P., & Lawson, S. (1992). **Evaluating Statistical Significance Using Corrected and Uncorrected Magnitude of Effect Size Estimates**. Paper Presented at Annual Meeting of the American Educational Research Association (San Francisco, CA ,April 20-24, 1992).

Spatz, C. (2007). *Basic statistics: Tales of distributions*. Cengage Learning.

Thompson, B., & Snyder, P. A. (1998). **Statistical significance and reliability analyses in recent Journal of Counseling & Development research articles**. *Journal of Counseling and Development: JCD*, 76(4), 436.

Wilkerson, M., & Olson, M. R. (1997). **Misconceptions about sample size, statistical significance, and treatment effect**. *The Journal of Psychology*, 131(6), 627-631.

William, M. (2013). **Statistics and measurement concepts with open stat**, New York : Springer science Business Media.

Yones, Alyones. (1999). **comparison parametric and nonparametric test and their post hoc comparison procedure**. College of education, Ohio University.

Zeighami, B. (1973). **Models for multiple comparisons in statistics and some nonparametric techniques**, Published Doctor of Philosophy Dissertation, The University of Oklahoma.

Zimmerman, D. W., & Zumbo, B. D. (1990). **Effect of outliers on the relative power of parametric and nonparametric statistical tests**. *Perceptual and Motor Skills*, 71 (1), 339-349.