تحرى الخصائص التوزيعية
لبعض البدائل اللابارامترية لاختبار تحليل التغاير

د/ محسوب عبد القادر الضوى
قسم علم النفس التربوى
كلية التربية بقنا – جامعة جنوب الوادى

ملخص

هدفت الدراسة الحالية إلى تحرى الخصائص التوزيعية (تقديرات الخطأ من النوع الأول، تقديرات الخطأ من النوع الثانى، تقديرات الضلاعة، تقديرات القوة الإحصائية) لسبع بدائل لابارامترية لاختبار تحليل التغاير وهى: Shirley's test, Burnett and Barr's test , Quade's test, McSweeney and Porter's test, Puri-Sen-Harwell-Serlin's test, Hettmansperger's test, Puri-Sen's test باستخدام بيانات إمبريقية تم جمعها من قياس التحصيل فى البرنامج التدريبى: إدارة الضغوط المقدم ضمن أنشطة مشروع الطرق المؤدية للتعليم العالى Pathways for Higher Education المنفذ بجامعة جنوب الوادى .

وقد بلغ حجم عينة الدراسة الأساسية (٧٠١) طالباً وطالبة من منتسبى الفصلين الشتوى والصيفى بالبرنامج التدريبى، وهذه العينة أخذت منها عينات ذات أحجام مختلفة طبقاً لمقتضيات تحليل البيانات، لدراسة تأثير استخدام ثلاث من استراتيجيات التدريب (استرتيجية التعلم التعاونى، استراتيجية KWL ، استراتجية العصف الذهنى) على التحصيل فى تصميم غير متوازن Unbalanced Design .

وتوصلت نتائج الدراسة إلى: تشابه أداء اختبار تحليل التغاير البارامترى والبدائل اللابارامترية السبعة فى مواقف القياس الخمسة حيث كان القرار قبول الفرض الصفرى، وأظهر اختبار تحليل التغاير البارامترى وبدائله اللابارامترية تضخم فى تقدير الخطأ من النوع الأول فى مواقف القياس الخمسة عدا الموقف الثانى، ولم يظهر اختبار تحليل التغاير البارامترى وبدائله اللابارامترية ضلاعة فى مواقف القياس الخمسة عدا اختبار Hettmansperger الذى حقق الحد الأعلى لمحك الضلاعة المتحرر فى موقف القياس الثانى ، وتوصلت نتائج الدراسة أيضاً إلى إطراد القوة الإحصائية للبدائل اللابارامترية لاختبار تحليل التغاير بزيادة حجم العينة ، وأظهر اختبار تحليل التغاير البارامترى قوة (شبه مقبولة، ومتميزة) فقط فى موقفى القياس الثانى والخامس على الترتيب، بينما أظهرت البدائل اللابارامترية خصائص القوة بدرجات متباينة فى مواقف القياس الخمسة .

وأظهرت النتائج المتحصل عليها أهمية وفائدة البدائل اللابارامترية لاختبار تحليل التغاير في الميدان البحثي ، على وجه الخصوص في تحليل نتائج التجارب التي تبتعد توزيعاتها عن الاعتدالية .

تحرى الخصائص التوزيعية
لبعض البدائل اللابارامترية لاختبار تحليل التغاير

د/ محسوب عبد القادر الضوى
قسم علم النفس التربوى
كلية التربية بقنا – جامعة جنوب الوادى

مقدمة

يتعرض المستخدم لأى اختبار إحصائى عند اتخاذه قراراً إحصائياً إلى الوقوع فى الخطأ من النوع الأول Type I Error أو الوقوع فى الخطأ من النوع الثانى Type II Error ، وهما وثيقا الصلة بقوة الاختبار الإحصائى Power of the Statistical Test وضلاعته[⊕] Robustness وفترات الثقة Confidence Intervals .

ويذكر محمد أبو يوسف (١٩٨٩ : ٢٢٧) أنه لا يوجد قرار إحصائى منزه عن الخطأ ، فالقرارات الإحصائية هى دائماً قرارات احتمالية بمعنى أنه لا مفر من وجود احتمال للخطأ فى أى قرار نصدره عن مجتمع عن طريق عينة . ولما كانت هذه القرارات مؤسسة على ما نجريه من اختبارات للفروض وتزيد ثقتنا فيها بزيادة حساسية هذه الاختبارات ، وجب علينا أن ندرس كيف نزيد من هذه الحساسية ؛ أى من قدرة الاختبارات على تمكيننا من اتخاذ القرار السليم الذى لا يشوبه إلا قدر ضئيل من الخطأ ، ويأتى ذلك عن طريق التحكم ما أمكن فى احتمالات الأخطاء التى تنجم حتماً عند استخدام هذه الاختبارات .

وقرار التحليل الأساسى الذى يواجه الباحثين فى مجال التربية وعلم النفس يدور حول الاختيار من بين الاختبارات البارامترية والاختبارات اللابارامترية (Harwell, 1988) ، وتتطلب عملية اختيار الاختبار الإحصائى المناسب مزيداً من المهارة ، فإذا كان الاختبار خاطئاً فإنه بالتبعية لن تكون هناك إفادة من البيانات وتكمن الخطورة فى التوصل إلى استنتاجات غير مقنعة (Dyer, 1995: 384) .

ويعد توافر الافتراضات الأساسية التى تستند إليها الاختبارات الإحصائية البارامترية هو المحك الأساسى لاستخدامها فى تحليل بيانات البحوث فى جميع المجالات ومنها مجال التربية وعلم النفس . وأى تجاهل Violation لافتراضات الاختبار الإحصائى البارامترى يفسد توزيع الاختبار

[⊕] توجد عدة ترجمات لمصطلح Robustness منها : الضلاعة ، المنعة ، النجاعة ، الوقاية ، المقاومة ، الحصانة .

ويغير تقديرات (معدلات) الخطأ من النوع الأول والخطأ من النوع الثانى (Bradley, 1968: 25).

وكلما كان الاختبار الإحصائى قوياً كلما تمكن الباحث من رفض الفرض الصفرى عندما يكون غير صحيح ، وفى حالة العكس فإن الاختبار الضعيف يكلف الباحث أمره الكثير للبحث عن فروق أو اختلافات قد تكون فعلاً موجودة ولكن لضعف الاختبار فإن الباحث لا يتمكن من رفض الفرض الصفرى والإعلان عن دلالة هذه الفروق ويكون فى هذا إهدار لإمكانات البحث (صلاح جلال وآخرون ، ١٩٨٨ : ١٣٧) .

والاختبار الإحصائى المثالى هو الذى يوازن إلى حد كبير بين الخطأ من النوع الأول والثانى . كما أن الوضع الصواب ، هو أن يُحدد من بداية الدراسة ، أو البحث احتمالية حدوث الخطأ من النوع الأول والخطأ من النوع الثانى ، غير أنه فى الممارسة الحقيقية يُحدد معظم الباحثين فقط لحتمالية الخطأ من النوع الأول (على ماهر خطاب ، ٢٠١١ : ٢٨٤ – ٢٨٥) .

وتوصى الطبعة الأخيرة من دليل النشر للجمعية الأمريكية لعلم النفس بأنه عند تطبيق الاختبارات الإحصائية الاستدلالية يجب أن تؤخذ اعتبارات القوة الإحصائية المرتبطة باختبارات الفروض مأخذ الجد ، ومثل هذه الاعتبارات ترتبط باحتمالية الرفض الصحيح للفروض المختبرة ، ومستوى الدلالة الإحصائية ، وحجم التأثير ، وحجم العينة ، وتوصى بأهمية أن تقدم أى دراسة أدلة على أن الاختبارات الإحصائية تتمتع بقوة كافية لتحرى التأثيرات الأساسية موضع الدراسة ، كما توصى باختبار الافتراضات الأساسية للنماذج الإحصائية المتبناة بالإضافة إلى فحص ملاءمة النموذج Model Fitting وحساب فترات الثقة (American Psychological Association, 2010: 30-31) .

وبعد أن سادت الاختبارات الإحصائية البارامترية لعدة عقود ، كان لظهور البدائل اللابارامترية فى بداية السبعينات من القرن العشرين التى تعد مناسبة أكثر للبيانات التى يتم جمعها من تطبيق الأدوات فى مجال القياس النفسى والتربوى أكبر الأثر فى جعل العمل يتسارع فى هذا المجال ، وظهر هذا جلياً فى الدور الذى لعبته اختبارات الرتب (اللابارامترية) فى قياس الدلالة الإحصائية خلال الكتابات المتخصصة فى الإحصاء النفسى والتربوى والعلوم الاجتماعية ، لذا فقد أوصى بها عندما لا يتحقق أحد افتراضى الاعتدالية أو تجانس التباين (Zimmerman, 1996) .

وتتلخص الأسباب المحتملة لندرة استخدام البدائل اللابارامترية فى البحوث النفسية والتربوية فى : صلاحية الحزم الإحصائية المعروفة مثل SPSS, MINITAB, SAS مع عدد صغير جداً من البدائل اللابارامترية ، واستمرار الاعتقاد بأن البدائل اللابارامترية أقل قوة عندئذ

تكون أقل قبولاً مقارنة بنظيراتها البارامترية ، وعدم وعى الباحثين بالبدائل المتاحة للاستخدام مع التصميمات التجريبية المعقدة وكيفية إنجاز تحليل البيانات باستخدام الحزم الإحصائية المتوافرة ، واعتقاد العديد من الباحثين بشكل قاطع أن البدائل اللابارامترية محدودة بالبيانات المستقاة من تصميمات بسيطة نسبياً (Harwell, 1990; Harwell & Serlin, 1988; 2001) .

ويعد اختبار تحليل التغاير* Analysis of Covariance (ANCOVA) من اختبارات الفروق الإحصائية ، وهو واحد من أكثر ثلاثة أساليب إحصائية شيوعاً يستخدمها الباحثون لاختبار الفروض فى تحليل التغير Analysis of Change (Kisbu-Sakarya, MacKinnon & Aiken, 2013) .

وهو إجراء إحصائى مهم لاختبار تعادل Equality مجموعتى معالجة أو أكثر عندما توجد فروق عشوائية فى متغير مصاحب أو أكثر . ويتمتع بمميزات رئيسة عن اختبار تحليل التباين من حيث زيادة القوة الإحصائية وتخفيض التحيز Reduction in Bias (Rheinheimer & Penfield, 2001) .

ويعتمد تحليل التغاير على استخدام العلاقة بين المتغير التابع والمتغير المصاحب Covariate Variable وذلك لضبط أو تعديل Adjust قيم المتغير التابع ، لتقليل تباين الخطأ (التباين داخل المجموعات) الذى يعزى إلى الخطأ التجريبى لتوفير اختبار أكثر حساسية للافتراض الخاص بعدم وجود فروق بين المجموعات (Harwell & Serlin, 1988).

ويخدم تحليل التغاير هدفين محددين ، الأول : فى الدراسات التجريبية التى تنطوى على التعيين العشوائى للوحدات فى المعالجات المشروطة ، فإن المتغير المصاحب ، حين يكون ذا علاقة بالمتغير الأصلى ، يقلل من تباين الخطأ Error Variance مما يزيد من القوة الإحصائية ويوفر حدود ثقة أكثر دقة (ضيقة) Narrower Confidence Intervals ، أى يزيد من الدقة فى تقدير تأثيرات المعالجة (المجموعة) . والثانى : فى الدراسات غير التجريبية والتى لا يستخدم فيها التعيين العشوائى ، فإن المتغير المصاحب يمكن أن يساهم فى تقليل الخطأ فى التمييز بين أثر المتغير الأساسى والمتغير المصاحب Confounding Effect . وتوفر نماذج تحليل التغاير تقديرات لتأثيرات المتغيرات المصاحبة على الناتج (الضبط للمعالجة) Adjusted for Tratment، ويمكن أن تختبر تجانس أثر المعالجة داخل المجموعات الجزئية المحددة فى المتغير

---

* يترجم مصطلح Analysis of Covariance Test إلى اختبار تحليل التباين المشترك ، اختبار تحليل التباين المصاحب ، اختبار تحليل التباين المتلازم .

المصاحب The Homogeneity of the Treatment Effect within Covariate-Defined Subgroups (Keselman et al., 1998; Koch et al., 1998; Zink & Koch, 2012) .

وبمرور السنوات ، أدت الحاجة إلى طرق ملائمة لتحليل البيانات إلى الكثير من البحوث فى إجراءات تحليل التغاير ، على الرغم من أن الكثير من المشكلات المتعلقة بذلك مازالت موجودة ، علاوة على الحاجة إلى تطوير الطرق اللابارامترية لتحليل التغاير لمعالجة المواقف التى لا تفى بافتراض الاعتدالية (Conover & Iman, 1981, 1982) .

وقد قدم المتخصصون فى الإحصاء مجموعة من البدائل اللابارامترية لاختبار تحليل التغاير البارامترى منها على سبيل المثال : Shirley's Test, Burnett and Barr's Test , Quade's Test, McSweeney and Porter's Test, Hettmansperger's Test, HSPS Test, PS Test, Rank Transformation of Conover and Iman, Rogosa's Procedure ، وأيضاً البديل الذى قدمه Koch, Tangen, Jung, and Amara (1998) ,

ويختلف تحليل التغاير البارامترى عن تحليل التغاير اللابارامترى ، حيث يصمم الاختبار البارامترى لاختبار تساوى (تكافؤ) المتوسطات المشروطة Equality of Conditional Means، أما البدائل اللابارامترية فتصمم لتكون اختبارات للتوزيعات المشروطة Equality of Conditional Distributions (Olejnik & Algina, 1985; 2003) .

والدراسة الحالية محاولة لتحرى الخصائص التوزيعية (تقديرات الخطأ من النوع الأول ، تقديرات الخطأ من النوع الثانى ، تقديرات الضلاعة ، تقديرات القوة الإحصائية) لبعض هذه البدائل اللابارامترية لاختبار تحليل التغاير والتى تمثل ثمرة جهود مجموعة من علماء الإحصاء النفسى والتربوى وعلماء الرياضيات فى إطار سعيهم للوصول إلى استنتاجات دقيقة وتعميمات صحيحة فى مواقف القياس المختلفة ووضع استراتيجية محددة وواضحة ربما تسهم فى اختيار الاختبار الإحصائى المناسب مع بيانات تفى أو لا تفى بالافتراضات التى يستند إليها الاختبار .

مشكلة الدراسة

لقد كان ومازال مجال معالجة البيانات إحصائياً مفتوحاً لاضافات جديدة ، وتأتى الاختبارات الإحصائية البارامترية واللابارامترية وما يتصل بها من اختيار الاختبار الإحصائى المناسب ودراسة الدلالة الإحصائية وقوة الاختبار وضلاعته لتمثل أحد الموضوعات المهمة فى مجال تفسير بيانات البحوث التربوية والنفسية .

ويذكر Harwell (1990, 1991) أن اختبارات النظرية الاعتدالية Normal-Theory Tests وهى الاختبارات الإحصائية البارامترية مازالت تستخدم على نطاق واسع فى البحوث

النفسية والتربوية على الرغم من توافر الدليل النظرى والتجريبى بأن البدائل اللابارامترية تتمتع بخصائص قوة متميزة عند دراسة المتغيرات فى مجال البحوث التربوية والنفسية . ويضيف Conover and Iman (1982) أن استخدام الرتب يقلل أهمية افتراض الاعتدالية .

وهكذا فإن الباحث التجريبى الذى يقوم ، وبصورة مستمرة ، باستخدام أساليب ومعالجات تقليدية دون التأكد من توفر الافتراضات التى تنطوى عليها هذه الأساليب والمعالجات الإحصائية ، فإن ذلك الباحث يقوم عن قصد أو بدون قصد بإثراء الأدب التربوى بأبحاث ذات نتائج غير قابلة للإثبات Verification عن طريق الإعادة Replication (سالم عويس ، ٢٠٠٣) .

فجميع الأدوات الإحصائية ذات الطابع الاستدلالى Inferential تنطوى على مجموعة من الافتراضات الأساسية ، وفى حال توفر هذه الافتراضات فإن الأداة الإحصائية ستكون صادقة حيث إنها تقوم بأداء ما تدعى عمله أو قياسه . وفى حالة عدم توفر هذه الافتراضات ، فإن النتائج تكون غير صادقة . أى يعتمد كل اختبار للاستدلال الإحصائى على مجموعة أساسية من الافتراضات ، عندما يتم استيفائها فإن الاختبار سوف يوظف كما هو مستهدف منه ومعد له ، وعندما يتم تجاهل الافتراضات فإن الاختبار ربما يكون مضلل (Keselman et al., 1998) .

وقد فحص Micceri (1989) فى دراسته الشهيرة أكثر من ٤٠٠ دراسة واختبر افتراض الاعتدالية لكل توزيع باستخدام اختبار Kolmogorov-Smirnov ووجد أن جميع التوزيعات تقريباً لا تفى بهذا الافتراض عند مستوى دلالة إحصائية (٠.٠١) ، وأوضح أنه عند التخلى عن افتراض الاعتدالية ، فإن الاختبارات اللابارامترية يمكن أن تكون أكثر قوة عما هو عليه فى حالة اختبار t-test واختبار F-test البارامتريين .

ويذكر ليونا أ. تايلر (١٩٨٣ ، ٣١) أن المقاييس الإنسانية وبخاصة مقاييس السمات النفسية لا تعطى دائماً التوزيع الاعتدالى ، فبعض هذه التوزيعات يكون ملتوياً بمعنى أن يكون هناك الكثير من الدرجات العالية أو الكثير من الدرجات المنخفضة فى المجموعة بحيث لا نحصل على التماثل الموجود فى المنحنى الاعتدالى ، وبعض التوزيعات الأخرى تكون مدببة أو مفرطحة بحيث لا تنطبق عليها معادلة المنحنى الاعتدالى .

ويضيف السيد محمد خيرى (١٩٩٧ : ١١٢) بأن التوزيع الاعتدالى النموذجى لا يمكن الحصول عليه تماماً فى أى بحث من البحوث مهما اتسع نطاقه ، وأن التوزيع الاعتدالى ما هو إلا تجريد لما يجب أن يكون عليه التوزيع .

لذا فالمشكلة التى واجهت الإحصائيين التطبيقيين Applied Statisticians هى ملاءمة (مطابقة) الأساليب الإحصائية البارامترية لمشكلات البيانات الحقيقية فى إطار النظرية الإحصائية الاعتدالية ، لما كان العديد من فئات البيانات التى يتعاملون معها ليست اعتدالية بشكل واضح . وفى شأن هذه المشكلات برز نهجان متمايزان أو مدرستان فكريتان هما : تحويل البيانات إلى هيئة (شكل) أكثر تشابهاً مع التوزيع الاعتدالى ، واستخدام إجراء التوزيع الحر Distribution-Free Procedure . وتتضمن الطريقة الأولى : التحويل اللوغاريتمى Log Transformation ، وتحويل الجذر التربيعى Root square Transformation ، وتحويل جيب الزاوية Arcsin Transformation ، أما الطريقة الثانية فتتضمن عدداً كبيراً من الاختبارات المعتمدة على رتب البيانات (Conover & Iman, 1981) .

ونظراً لاختلاف إجراءات التوزيع الحر فى العديد من الأوجه ظهر اهتمام لمقارنة هذه الإجراءات باعتبار تقديرات للخطأ من النوع الأول والقوة الإحصائية النسبية ، علاوة على مقارنة أداء هذه الإجراءات باختبار تحليل التغاير البارامترى عند تجاهل الافتراضات الأساسية وعند استيفاء كافة الافتراضات (Olejnik & Algina, 1985) .

ويذكر Harwell and Serlin (2001) أنه من المهم للباحثين أن يألفوا مفهوم صدق الاستنتاج الإحصائى Statistical Conclusion Validity الذى يؤكد على أن نتائج التحليل الإحصائى تعتمد كثيراً على المدى الذى تفى خلاله البيانات بالافتراضات التى يستند إليها الاختبار مثل الاعتدالية وتجانس التباين ، ولأن البيانات الحقيقية يمكن أن لا تفى تماماً بالافتراضات التى يستند إليها الاختبار الإحصائى فإن فحص البيانات عند تجاهل أحد الافتراضات يعد أمراً مهماً لابد من أخذه فى الاعتبار .

ويشير MacDonald (1999) إلى ميل الباحثين لاستخدام الاختبارات البارامترية على حساب الاختبارات اللابارامترية خلال العقود الثلاثة الأخيرة من القرن العشرين ، وأن المراجعات توصلت إلى أن ١٤ % من البحوث المنشورة فى الدوريات العلمية النفسية والتربوية استخدمت اختبارات بارامترية ، فى حين أن ٨.٢ % من البحوث النفسية المنشورة فى نفس هذه الدوريات قد استخدمت اختبارات لابارامترية .

ويذكر Sawilowsky and Fahoome (2000) أن الاختبارات اللابارامترية تكون غالباً أكثر قوة من الاختبارات الكلاسيكية البارامترية فى حالة البيانات الحقيقية Real-World Data التى نادراً ما تتوزع اعتدالياً ، ومن ناحية أخرى فإنه ثمة صعوبة فى استخدام هذه الاختبارات اللابارامترية فالمعادلات الخاصة بحسابها متناثرة خلال الأدبيات الإحصائية كما يوجد

نقص فى الجداول الخاصة بالقيم الحرجة Critical Values .

لذا ظهرت الحاجة إلى طرق لتحليل البيانات التى تستبعد افتراض أو شرط الاعتدالية عندما أدرك الباحثون أن عديد من البيانات التربوية والنفسية ليست أحادية المنوال Unimodal ولكنها مفرطحة أو ملتوية ، وتحت هذا الشرط تكون الاختبارات البارامترية غالباً لها خصائص توزيعية منخفضة بالمقارنة مع البدائل اللابارامترية ، كما أنه يوجد القليل من الارشادات للاختيار ما بين الاختبارات البارامترية والاختبارات اللابارامترية المتاحة أمام الباحثين وكذا فى مجال تقويم عمل الباحثين الآخرين عن طريق فحص الدراسات المنشورة فى مجال التربية وعلم النفس للتعرف على الاختبارات الإحصائية المستخدمة ومدى ملاءمتها فى موقف القياس الذى استخدمت فيه . (Harwell, 1988)

ويذكر Seaman, Algina and Olejnik (1985) أن معظم التصميمات البحثية المستخدمة لمقارنة فعالية المعالجات التعليمية تنتج بيانات يمكن تحليلها بكل من الاختبارات البارامترية واللابارامترية . وللاختيار بين الإجراءات المتاحة فإنه من المفيد أن يتم تقدير الخطأ من النوع الأول والقوة مع اعتبار نوع التوزيع ، والباحث يمكن أن يختار الاختبار الأكثر قوة من بين الاختبارات التى تكون مستويات ألفا الفعلية Actual Alpha لها قريبة من مستوى ألفا الإسمى Nominal Alpha الذى يطلق عليه عادة مستوى الدلالة الإحصائية .

ويعتمد المحك الإحصائى Statistical Criteria كأساس للاختيار ما بين الاختبارات الإحصائية البارامترية واللابارامترية على اثنين من الخصائص التوزيعية هما القوة الإحصائية وقدرة الاختبار على الضبط أو السيطرة على تقديراته للخطأ من النوع الأول عند مستوى دلالة – يحدده الباحث– معروف ومحدد . والاختبار الذى يتحقق له ذلك يكون هو موضع الاختيار الصحيح ، ويعد التحكم فى الخطأ من النوع الأول محكاً ضرورياً وليس كافياً ، ويجب فحص القوة الإحصائياً أيضاً والتى بدورها ترتبط بتقدير الخطأ من النوع الثانى (Harwell, 1988, 1989) .

وبهذا فإن الاختيار الخاطىء لاختبار إحصائى سواء كان بارامترياً أو لابارامترياً ربما يؤدى إلى استخدام اختبار ذو تقدير مرتفع للخطأ من النوع الأول أو ذو قوة منخفضة (ذو تقدير مرتفع للخطأ من النوع الثانى) مما يترتب عليه دلالات زائفة وتعميمات غير مقبولة تبتعد كثيراً عما يعرف بصدق الاستنتاجات الإحصائية ، وهذا فى مجمله يتطلب مزيداً من الدقة والحرص واليقظة وبخاصة فى مجال الدراسات النفسية والتربوية .

ويذكر Hsu (1983) أن التجاهلات المتعددة Multiple Violations للافتراضات التى يستند إليها اختبار تحليل التغاير تعد مشكلة خطيرة . ويضيف Barrett (2011) أن الفشل فى

استيفاء افتراضات اختبار تحليل التغاير يمكن أن يؤدى إلى تفسيرات خاطئة للنتائج Misinterpretation of Results ، فالفشل فى استيفاء افتراض توازى خطوط ميل الانحدار للمجموعة Parallel Group Regression Slopes هو أمر شائع فى العديد من فئات البيانات وهناك العديد من الطرق المتاحة لتحليل فئات البيانات هذه ومنها على سبيل المثال طريقة Johnson-Neyman Technique . وعلى الرغم من ضلاعة الاختبار إزاء تجاهل بعض الافتراضات مثل الاعتدالية وتساوى التباينات فى حالة تساوى أحجام العينات ، فهناك العديد من الاختبارات اللابارامترية المعتمدة على الرتب ، المتاحة كبدائل لابارامترية لاختبار تحليل التغاير .

وعلى الرغم من الندرة الشديدة فى الدراسات والبحوث المتعلقة بموضوع الدراسة الحالية ، إلا أن نتائج ما توافر منها يبين نتائج متباينة فيما يتصل بخصائص القوة الإحصائية وتقديرات أخطاء القرارات الإحصائية فى مواقف قياس مشروطة بنوع معين من التوزيع ، وحجم العينة ، ودرجة التجانس ، ودرجة الاعتدالية ، والقيمة العددية لمعامل الارتباط بين المتغير التابع والمتغير المصاحب ، وعدد الافتراضات التى يتم استيفائها والفروق فى الضلاعة والقوة ، وعلى سبيل تتفق نتائج بعض الدراسات على تميز الاختبار البارامترى ، بينما لا تتفق نتائج دراسات أخرى بشأن تميز الاختبار البارامترى عن البدائل اللابرامترية . كما تبنى آخرون منحى بحثى لاقتراح وتجريب طرق جديدة لتحسين أداء الاختبار البارامترى فى حالة عدم توفر افتراض تجانس ميل الانحدار Heterogeneous Regression Slopes باستخدام إجراءات مثل : Blocking or Stratification, Hollingsworth's F Statistic, Johnson-Neyman Procedure, Wilcox's Trimmed-Mean Method, Potthoff's Modified Johnson-Neyman Procedure, Rogosa's Three Alternatives (Kim, 2007, 2010) .

وقد خلصت البحوث والدراسات السابقة الأجنبية فى مجال مقارنة أداء الاختبار الإحصائى البارامترى وبدائله اللابارامترية إلى أن الخصائص التوزيعية (الخطأ من النوع الأول ، الخطأ من النوع الثانى ، الضلاعة ، القوة الإحصائية) من الموضوعات النامية والمتطورة ، أما الجهود العربية ومنها المصرية فمازالت ترقب وترصد تلك الجهود دون الانطلاق للحاق بما تم إنجازه فى هذا المجال ، وتمثلت الجهود العربية -المتاحة- فى هذا المجال فى دراسة عزو إسماعيل (١٩٩٦) التى هدفت إلى دراسة قوة كفاءة اختبارى (ويلكوكسون ، الإشارة) مقابل اختبار "ت" ، ودراسة الباحث الحالى لنيل درجة الدكتوراه (٢٠٠٤) حول قوة بعض البدائل اللابارامترية لاختبار تحليل التباين العاملى فى مجال الدراسات النفسية .

وفى ضوء ما سبق تتحدد مشكلة الدراسة الحالية فى الاجابة عن السؤالين التاليين :

١. ما أداء اختبار تحليل التغاير البارامترى والبدائل اللابارامترية السبعة تحت شروط مختلفة؟

٢. ما الخصائص التوزيعية (تقديرات الخطأ من النوع الأول ، تقديرات الخطأ من النوع الثانى ، تقديرات الضلاعة ، تقديرات القوة الإحصائية) لاختبار تحليل التغاير البارامترى والبدائل اللابارامترية السبعة تحت شروط مختلفة ؟

أهداف الدراسة

تهدف الدراسة الحالية إلى تحرى :

١. أداء اختبار تحليل التغاير البارامترى والبدائل اللابارامترية السبعة تحت شروط مختلفة .

٢. الخصائص التوزيعية (تقديرات الخطأ من النوع الأول ، تقديرات الخطأ من النوع الثانى، تقديرات الضلاعة، تقديرات القوة الإحصائية) تحت شروط مختلفة .

أهمية الدراسة:

الأهمية النظرية

تكمن الأهمية النظرية للدراسة الحالية فى تقديم إطار نظرى واف عن :

١. اختبار تحليل التغاير البارامترى .

٢. الخصائص التوزيعية للاختبارات الإحصائية (تقديرات الخطأ من النوع الأول ، تقديرات الخطأ من النوع الثانى ، تقديرات الضلاعة ، تقديرات القوة الإحصائية) .

٣. بعض البدائل اللابارامترية لاختبار تحليل التغاير البارامترى التى لم تستخدم بعد فى تحليل نتائج البحوث فى البيئة المصرية والعربية .

الأهمية التطبيقية

تكمن الأهمية التطبيقية للدراسة الحالية فى النقاط التالية :

١. تقديم الصيغ الرياضية لبعض البدائل اللابارامترية لاختبار تحليل التغاير البارامترى .

٢. تحرى الخصائص التوزيعية (تقديرات الخطأ من النوع الأول ، تقديرات الخطأ من النوع الثانى ، تقديرات الضلاعة ، تقديرات القوة الإحصائية) لبعض البدائل اللابارامترية لاختبار تحليل التغاير البارامترى تحت شروط مختلفة .

حدود الدراسة

تحددت نتائج الدراسة الحالية بالعينة والأداة والمعالجة الإحصائية للبيانات . حيث بلغ حجم عينة الدراسة الأساسية (٧٠١) طالباً وطالبة من طلاب جامعة جنوب الوادى المسجلين فى المجال السلوكى Behavioral Batch ببرنامج إدارة الضغوط من مختلف كليات الجامعة وينتمون إلى

فرق دراسية مختلفة ، وطور الباحث اختباراً تحصيلياً فى محتوى البرنامج التدريبى إدارة الضغوط المقدم ضمن أنشطة مشروع الطرق المؤدية للتعليم العالى ، وعولجت البيانات الإمبيريقية باستخدام اختبار تحليل التغاير وبعض البدائل اللابارامترية .

تعريف مصطلحات الدراسة Definition of Terms

الخطأ من النوع الأول A Type I Error

احتمال الرفض الخاطىء لفرض صفرى صحيح(Finch, Cumming & Thompson, 2002). أو هو ذلك الخطأ الذى ينشأ حين يُتخذ قرار برفض الفرض الصفرى بينما يكون هذا الفرض صحيحاً فى الواقع ، ويرمز لاحتمال هذا الخطأ بالرمز α (محمد أبو يوسف ، ١٩٨٩ : ٢٢٨) .

الخطأ من النوع الثانى A Type II Error

احتمال القبول لفرض صفرى خاطىء (Finch, Cumming & Thompson, 2002) . أو هو ذلك الخطأ الذى ينشأ حين يُتخذ قرار بقبول الفرض الصفرى بينما يكون هذا الفرض زائفاً فى الواقع ، ويرمز لاحتمال هذا الخطأ بالرمز β (محمد أبو يوسف، ١٩٨٩: ٢٢٨) .

قوة الاختبار الاحصائى Power of a Statistical Test

احتمال أن يؤدى الاختبار الإحصائى إلى رفض الفرض الصفرى الصحيح عندما يكون الفرض فى حقيقة الأمر خاطئاً (Dyer, 1995: 388) . أو هى احتمال تجنب الخطأ من النوع الثانى ، ويعبر عنها بالمقدار العددى $(1-\beta)$ (محمد أبو يوسف ، ١٩٨٩ : ٢٢٨) .

ضلاعة الاختبار الاحصائى Robustness of a Statistical Test

محافظة الاختبار الإحصائى على أدائه إذا لم تستوف البيانات بعض أو كل الافتراضات التى يستند إليها الاختبار الإحصائى (Bradley, 1978) .

الاطار النظرى للدراسة

اختبار تحليل التغاير Analysis of Covariance Test

قدم اختبار تحليل التغاير بواسطة السير Ronald A. Fisher عام ١٩٣٢م . وهدف تصميم الاختبار إلى استخدام العلاقة بين متغير تابع ومصاحب لتعديل درجات المتغير التابع لغرض تقليل تباين الخطأ غير المفسر Unexplained Error Variance أى يخفض تباين الخطأ فى المتغير التابع بمقدار ما يمكن ضبطه accounted for بواسطة متغير مصاحب لتوفير تقدير أكثر دقة لتأثيرات المعالجة واختبار أكثر قوة للفروض (Nakonezny & Shull, 2007) .

ويذكر صلاح الدين علام (١٩٩٣ : ٣٥٤) أن جميع التصميمات التجريبية تهدف إلى التحقق من أن نتائج التجربة تعزى إلى المتغيرات المستقلة وليس إلى مؤثرات عارضة أو متغيرات

دخيلة . وتحليل التغاير لا يمثل مجموعة من هذه التصميمات التجريبية ، وإنما يعد أسلوباً لزيادة دقة التصميم عن طريق إجراء تعديل فى درجات المتغير التابع وفقاً لدرجات متغير أو أكثر من المتغيرات التى يود الباحث ضبط تأثيرها إحصائياً .

ويضيف مسفر سليم الجعيد (٢٠٠٤) أن الأمر الشائع فى البحوث التربوية والنفسية أنها تجرى على الظواهر الإنسانية التى تتأثر بعوامل عديدة ومتفاعلة يصعب عزلها أو تثبيتها أو السيطرة عليها والتحكم فيها بشكل مطلق لذلك لا يستطيع الباحث أن يطلق حكماً نهائياً على الظاهرة ما لم يتمكن من تحديد أثر كل متغير وحصر العلاقة بين متغيرات الدراسة ، وأن أسلوب تحليل التغاير يقدم دعماً قوياً يساعد فى إزالة أو تقليل أثر المتغير الخارجى بحيث ينسب الأثر الحاصل للمتغير المستقل فقط ، فحاجة الباحث لهذا الأسلوب كطريقة من طرق إزالة آثار المتغيرات الخارجية يكسب الموضوع أهمية .

ويربط اختبار تحليل التغاير بين فلسفة تحليل التباين Analysis of Variance وتحليل الانحدار Regression Analysis ويطلق على المتغير الذى توقع الباحث أهميته اسم المتغير المصاحب أو الملازم Covariate Variable وهو المتغير الذى يسعى الباحث إلى ضبط أثره إحصائياً (أحمد سليمان عودة وخليل يوسف الخليلى ، ٢٠٠٠ : ٥١٢) .

ويمكن تمثيل تحليل التغاير باستخدام إجراء المتوسطات المعدلة (البحث عن الفروق فى المتوسطات المعدلة للمتغير المصاحب) ، ويمكن التعبير عن هذا الإجراء بسهولة بواسطة طريقة الانحدار الخطى المتعدد باستخدام متغيرات منبئة لتمثيل المجموعات (Barrett, 2011) .

وتتفاوت المتغيرات المصاحبة فى أثرها على المتغير التابع بجانب تأثير المعالجة ، لذا فالباحث الماهر هو الأقدر على تحديد المتغيرات المصاحبة ، وعليه أن يقرر أى المتغيرات سيتم ضبطه إحصائياً . والنقطة الجوهرية فى ذلك لو استطاع الباحث أن يضبط هذه المتغيرات المصاحبة تجريبياً من خلال عينة طبقية عشوائية لربما وفر على نفسه جهداً بحثياً لا يستهان به بحيث يجعل منه ميزة يتمتع بها الضبط التجريبى على الضبط الإحصائى ، خاصة وأن الضبط الإحصائى يستند على افتراضات أساسية ، وتعتمد فعاليته على مدى الإيفاء بتلك الافتراضات (Keppel, 1982: 492) .

ويلجأ بعض الباحثين فى مجال علم النفس والتربية إلى تحليل التغاير عندما يضطر للتعامل مع مجموعات لا تخضع لأى معالجة Intact Groups والتى ربما تختلف فى المتغير المصاحب الذى يرتبط بالمتغير التابع . ويعد تحليل التغاير مفيداً للباحث لأنه يمكنه من أن يحدد ويأخذ فى الاعتبار مصدر التباين الذى يعزى إلى المتغيرات المصاحبة لذا يوفر الاختبار ضبط

متميز . فإذا كان التباين فى متغير مصاحب كبير نسبياً وهذا المتغير يرتبط بالمتغير التابع فإنه من الممكن أن تستخدم درجات المفحوصين فى المتغير المصاحب باعتبارها متغير مصاحب . فالتعديل فى المتغير المصاحب سوف يؤدى إلى انخفاض حد الخطأ Reduction in the Error Term وبالتبعية إلى تحليل أكثر حساسية . وتحليل التغاير بالتالى لا يماثل المعالجات فى ضوء المستويات Treatments-by-Levels أو تصميم التجمعات العشوائى Randomized Blocks Design (Kerlinger & Pedhazur, 1973: 266) .

وقد أجرى Shirley and Newnham (1984) دراسة لمقارنة أربع قواعد لتقرير تضمين المتغير المصاحب (الوزن قبل الوفاة) فى دراسات تحليل الأوزان الحقيقية فى مجال علم السموم الحيوانى ، واستنتج أنه إذا كانت المعلومات السابقة Background Information تظهر علاقة خطية بين المتغير التابع والمتغير المصاحب فإنه ينصح بالتعديل (الضبط) تبعاً للمتغير المصاحب to adjust for the covariate .

وعلى الرغم من تقديم لختبار تحليل التغاير منذة مدة طويلة ، لكنه غير مفهوم من قبل العديد من الباحثين ، ولا يستخدم بشكل متكرر . وهو تمديد لاختبار تحليل التباين بتضمين متغير مصاحب أو أكثر ، ومن فوائده مقارنة باختبار تحليل التباين : أنه ذو قوة متزايدة Increased Power ، وانخفاض التحيزات Reduction in Biases الناتجة عن الفروق فى الوحدات التجريبية (المتغيرات المصاحبة) بين المجموعات (Barrett, 2011) .

وتشير المراجعة الدقيقة للأبحاث التربوية محلياً وعالمياً إلى شيوع استخدام هذا الأسلوب فى التحليل الإحصائى لأبحاث فى مجالات متنوعة فى العلوم الإنسانية والتربوية باستخدام مجتمعات دراسية وفئات عمرية مختلفة . لكن وبالرغم من العدد الكبير للأبحاث التى استخدمت تحليل التغاير ، فإن هذا يشكل نسبة صغيرة من الأبحاث التجريبية بشكل عام ، ويرجع السبب فى ذلك إلى عدم فهم وإدراك الباحثين التربويين للفوائد التى يمكن جنيها من استخدام هذا النوع من التحليل الإحصائى ، خاصة فى مجال تقليل تباين الخطأ وما يصاحبه من زيادة فى قوة التحليل الإحصائى إضافة إلى زيادة الدقة فى النتائج (سالم عويس ، ٢٠٠٣ ؛ Kim, 2007, 2010) .

الأساس المنطقى وافتراضات تحليل التغاير

The Logic and Assumptions of Analysis of Covariance

عند حساب الباقى لمتغير Variable is Residualized ، فإن الارتباط بين المتغير المنبىء Predictor Variable والبواقى Residuals يساوى الصفر ، وبتعبير آخر أن المتغير الذي تم حساب الباقى له Residualized Variable قد ارتبط (اشترك) جزء منه بالمتغير المنبىء الذى

تم تنقيته Purged . وهذا هو منطق تحليل التغاير الذى يمكن تلخيصه فى المعادلة التالية :

$$Y_{ij} = \bar{Y} + T_j + b(X_{ij} - \bar{X}) + \varepsilon_{ij}$$

حيث : Y_{ij} درجة المفحوص (i) فى المعالجة (j) ، $\bar{Y}$ المتوسط الكلى للمتغير التابع ، T_j تأثير المعالجة (j) ، b معامل انحدار Y على X ، X_{ij} درجة المفحوص (i) على المتغير المصاحب فى المعالجة (j) ، $\bar{X}$ المتوسط الكلى للمتغير المصاحب ، ε_{ij} الخطأ المتصل بدرجة المفحوص (i) فى المعالجة (j) .

لذا فإن المعادلة السابقة يمكن إعادة كتابتها فى الصورة التالية :

$$Y_{ij} - b(X_{ij} - \bar{X}) = \bar{Y} + T_j + + \varepsilon_{ij}$$

والتى تبين أنه بعد التعديل يمكن تخيل الدرجة $\{Y_{ij} - b(X_{ij} - \bar{X})\}$ كمركب من المتوسط الكلى وتأثير المعالجة وحد الخطأ Error Term ، كما يعد الطرف الأيمن من المعادلة السابقة بمثابة معادلة النموذج الخطى ، وإذا كان معامل الانحدار $(b = 0)$ بمعنى إذا كان المتغير المصاحب لا يرتبط بالمتغير التابع فإن المعادلة السابقة تصبح فى الصورة التالية :

$$Y_{ij} = \bar{Y} + T_j + + \varepsilon_{ij}$$

أى أن فكرة تحليل التغاير تتلخص فى إزالة ذلك الجزء من المشاهدة أو الدرجة فى المتغير التابع والذى يمكن التنبؤ به من المتغير المصاحب أو المتغيرات المصاحبة ، بمعنى أن ضبط تأثير المتغير أو المتغيرات المصاحبة بطريقة إحصائية خاصة تتلخص فى دمج فكرة تحليل التباين مع فكرة تحليل الانحدار ، وعلى هذا فإن الباحث يقرر اختيار المتغير أو المتغيرات المصاحبة فى ضوء العلاقة الارتباطية بينه وبين المتغير التابع ، ولذلك فإن قوة هذه العلاقة بدرجة تصل إلى معامل ارتباط $(r \geq .6)$ يجعل من تحليل التغاير أقوى من تحليل التباين الذى يبنى على تصميم التجمعات العشوائية . وهذا يعنى أن الضبط الإحصائى ليس بديلاً مطلقاً للضبط التجريبى ، وإنما يمكن اللجوء إليه عندما لا تتوافر إمكانية الضبط التجريبى فى حالة إنخفاض القيمة العددية لمعامل الارتباط (Kerlinger & Pedhazur, 1973: 266-267) .

ويشترط تحليل التغاير أيضاً إمكانية قياس المتغير المصاحب دون تأثير المعالجة عليه وذلك إما بقياس المتغير المصاحب قبل المعالجة، أو أن المتغير المصاحب لا يتأثر فعلاً بالمعالجة. ولذلك إذا كان جزء من التباين فى المتغير التابع الذى يعزى إلى المتغير المصاحب نتيجة لتأثير المعالجة عليه أى على المتغير المصاحب فإن الضبط الإحصائى فى هذه الحالة يُضيع جزءاً من هدف التجربة (Keppel, 1982: 502) .

ويستند اختبار تحليل التغاير على مجموعة من الافتراضات هى نفسها الافتراضات التى يستند إليها اختبار تحليل التباين بالإضافة إلى افتراضات أخرى تتعلق بانحدار المتغير التابع على المتغير المصاحب ، ويمكن اجمال تلك الافتراضات فى التالى :

١. تجانس التباين فى المجتمعات التى تنتمى إليها المجموعات الجزئية فى التصميم التجريبى.
٢. تتوزع البيانات الإحصائية للمتغير التابع فى المجتمع الذى تنتمى إليه المجموعة الجزئية الواحدة توزيعاً اعتدالياً .
٣. استقلالية المجموعات موضوع المقارنة أى يعتمد تعيينها على الانتقاء أو التعيين العشوائى.
٤. تعتبر قيم المتغير المصاحب ثابتة وتقاس بدون خطأ ، ويعنى ذلك ضمنياً استقلال المتغير المصاحب عن تأثيرات المعالجة أو المتغير التابع. والضامن الوحيد للإيفاء بهذا الافتراض هو قياس المتغير المصاحب قبل التجربة (أى قبل المعالجة) وتعيين المفحوصين عشوائياً تبعاً لشروط المعالجة مع الاهمال الكامل للدرجات المأخوذة قبل التجربة.
٥. دلالة وخطية العلاقة بين المتغير التابع والمتغير المصاحب، ويطلق على هذا الافتراض مسمى الانحدار الخطى، ويعنى هذا الافتراض أن الانحرافات عن الانحدار سواء المعدلة أو درجات الباقى تتوزع اعتدالياً ومستقلة فى المجتمع بمتوسطات تساوى صفر وتباينات متجانسة وهذا لا يمكن تحقيقه إذا كان الانحدار غير خطى. ومن المحتمل أو استخدم الانحدار الخطى فى التحليل بينما الانحدار الحقيقى غير ذلك أى يكون مثلاً من النوع Curvilinear Regression فإن التعديلات لغرض حذف تأثير المتغير المصاحب لن تكون ذات فائدة، ومن ناحية أخرى فإن الأمر الأكثر أهمية أيضاً أنه يمكن اختبار معنى متوسطات المعالجة التى تم تعديلها أيضاً فى ضوء افتراض الانحدار الخطى .
٦. معاملات الانحدار لخطوط انحدار المتغير التابع على المتغير المصاحب للمجتمعات التى تنتمى إليها المجموعات الجزئية متساوية إحصائياً. أى تجانس الانحدار داخل المجموعات Homogenity of Regression Coefficients أى ميل خطوط (معاملات) الانحدار متساوى أو خطوط الانحدار متوازية. وتأتى أهمية الإيفاء بهذا الافتراض من أن التجانس يمكن الباحث من تقدير خطأ تجريبى واحد (متوسط المربعات داخل المجموعات فى تحليل التباين) وبدون الإيفاء بهذا الافتراض لا يستطيع الباحث حساب معامل انحدار موحد (ممزوج) Pooled لجميع المجموعات، وأن عدم الإيفاء

بهذا الافتراض يقلل من صدق التحليل الإحصائى كطريقة لضبط المتغيرات المصاحبة . بمعنى أن انتهاك الافتراض الأخير أكثر خطورة من انتهاك بقية الافتراضات ولذلك يصبح من الحتمى على الباحث الكشف إحصائياً عن الإيفاء بهذا الافتراض (Kerlinger & Pedhazur, 1973: 267-268; Keppel, 1982: 502-503; Olejnik & Algina, 1984, 1985; Keselman et al., 1998; Kim, 2007, 2010) \.

ويمكن اختبار مدى الإيفاء بالافتراض الأول والثانى بواسطة اختبارات خاصة بذلك مثل اختبار Bartlett للافتراض الأول ، واختبار مربع كاى للافتراض الثانى (Olejnik & Algina, 1984, 1985) .

خطوات حساب اختبار تحليل التغاير

يمكن حساب اختبار تحليل التغاير F فى ثلاث خطوات كالتالى :

١. تقدير معادلة الانحدار الخطى للمتغير التابع Y على المتغير المصاحب X خلال جميع المجموعات :

$$Y = b_{\circ} + b_{Y/X} X$$

وحساب معامل الارتباط المتعدد $R^2_{Y/X}$.

٢. تقدير معادلة الانحدار الخطى للمتغير التابع على المتغير المصاحب والعامل الجمعى Grouping Factor (G) . وفى حالة تصميم مكون من مجموعتين يتم استخدام A Single Dummy-Coded Variable (0,1) للتعبير عن العامل الجمعى :

$$Y = b_{\circ} + b_{Y/X} X + b_{Y/X} G$$

وحساب معامل الارتباط المتعدد $R^2_{Y/X,G}$.

٣. تحديد Increase R^2 المرتبطة بإضافة العامل الجمعى بنموذج الانحدار فى الخطوة (١)، ثم التعويض فى المعادلة التالية :

$$F = \frac{\left(R^2_{Y/X,G} - R^2_{Y/X}\right)/(P_2 - P_1)}{\left(1 - R^2_{Y/X,G}\right)/(N - P_2)}$$

حيث : P_1, P_2 عدد البارامترات المقدرة فى نموذجى الانحدار فى الخطوتين الأولى والثانية ، N حجم العينة الكلية . والنسبة الفائية المحسوبة يمكن مقارنتها مع النسبة الفائية النظرية بدرجات حرية $(N - P_2)$ ، $(P_2 - P_1)$ (Olejnik & Algina, 1985) .

البدائل اللابارامترية لتحليل التغاير

تقوم البدائل على إجراء تحويل Transformation على الدرجات أو المشاهدات الأصلية . ويؤكد Huberty (1994: 148) على فعالية تحويل الرتب فى حساب الاختبارات التى تستخدم بيانات تستقى من توزيعات غير اعتدالية . ويضيف Koti (1995) أن المعاينة القائمة على الرتب Rank-Set Sampling توفر تقديرات محسنة للمتوسط والتباين عندما توجد صعوبة فى القياس الحقيقى للمشاهدات كما أن عمل الرتب لهذه المشاهدات يكون سهلاً نسبياً .

وتمثل نماذج تحليل التغاير اللابارامترية ، وهى الأقل تقيداً بافتراضاتها الإحصائية بديلاً لاختبار تحليل التغاير البارامترى المعتاد . واقترحت العديد من نماذج تحليل التغاير اللابارامترية ، وهى متشابهة بأن كل نموذج يتضمن إجراء لعمل الرتب Procedure of Ranking لتحويل الدرجات الأصلية ، وتختلف تلك النماذج فيما يتصل بالضلاعة والقوة (Nakonezny & Shull, 2007) . وفيما يلى عرض لبعض البدائل اللابارامترية لاختبار تحليل التغاير :

أولاً : اختبار التوزيع الحر لكوادى Quade's Distribution-Free Test

يعد الاختبار أول البدائل اللابارامترية لاختبار تحليل التغاير البارامترى وطوره Quade (1967)، ويتوزع الاختبار تقاربياً Asymptotically Distributed مثل توزيع اختبار F لفيشر، ويتكون من تحليل التباين أحادى الاتجاه للبواقى التى يتم الحصول عليها من استخدام الانحرافات عن متوسط رتب درجات المتغير المصاحب وذلك للتنبؤ بقيم الانحرافات عن متوسط رتب المتغير التابع.

وتتلخص خطوات حساب الاختبار فى حالة متغير مصاحب وحيد فى التالى :

١. ترتب الدرجات الخام الخاصة بالمتغير التابع Y والدرجات الخاصة بالمتغير المصاحب X كل على حده خلال جميع المجموعات لنحصل من هذه الخطوة على رتب درجات المتغير التابع RY ورتب درجات المتغير المصاحب RX .

٢. يحسب الانحدار الخطى البسيط لرتب المتغير التابع RY على رتب المتغير المصاحب RX لنحصل من هذه الخطوة على معادلة الانحدار الخطى والتى تشمل القيم المتنبأ بها $\hat{RY}$ والثابت B وقيمة معامل الانحدار β .

٣. تحسب البواقى (الفروق) بين القيم المتنبأ بها $\hat{RY}$ والقيم RY لجميع قيم RX .

٤. يحسب الانحدار الخطى للبواقى $(RY - \hat{RY})$ التى تم الحصول عليها فى الخطوة السابقة (متغير تابع) على المتغيرات المتمثلة فى المتغير $G1$ والمتغير $G2$.

٥. تحسب قيمة الاختبار من المعادلة التالية :

$$(TS)_{QUADE} = \frac{\left\{R^2_{Y.12}\right\}/(P_1)}{\left\{1-R^2_{Y.12}\right\}/(N-P_2)} = \frac{\left\{R^2_{Y.12}\right\}(N-P_2)}{\left\{1-R^2_{Y.12}\right\}(P_1)}$$

حيث N تمثل حجم العينة الكلية ، P_1 تمثل عدد البارامترات المقدرة فى معادلة الانحدار الخطى الأولى (الخطوة الثانية) ، P_2 تمثل عدد البارامترات المقدرة فى معادلة الانحدار الخطى الثانية (الخطوة الرابعة) ، $R^2_{Y.12}$ تمثل مربع معامل الارتباط المتعدد الذى يعتمد حسابه على معامل الارتباط بين قيم المتغير التابع وقيم بقية المتغيرات وكذلك على معامل الارتباط بين المتغيرات الأخرى غير المتغير التابع .

وإحصاء Quade يمكن تطبيقه فقط لاختبار تأثيرات المعالجة المعدلة تبعاً للمتغير المصاحب to test for covariate-adjusted treatment effects ، ويتجنب افتراض الاعتدالية Normality والخطية Linearity ولكن يتطلب أن المتغير المصاحب له نفس التوزيع خلال جميع مستويات المعالجة (Wang & Akritas, 2006) .

ثانياً : اختبار مكسويلى وبورتر McSweeney and Porter Test

اقترحه McSweeney and Porter (1971) كبديل لابارامترى لاختبار تحليل التغاير البارامترى ، ويعتمد على حساب الانحدار الكلى والانحدار داخل المجموعات ، ويقوم على تحويل الدرجات الخام للمتغير التابع والدرجات الخام للمتغير المصاحب إلى صورة رتب خلال كل المجموعات الجزئية كل على حده ، ومن ثم استخدام الرتب فى حساب الاختبار بنفس الخطوات المتبعة عند حساب اختبار تحليل التغاير البارامترى وذلك من خلال اتباع الخطوات التالية :

١. ترتب الدرجات الخام الخاصة بالمتغير التابع Y والدرجات الخاصة بالمتغير المصاحب X كل على حده خلال جميع المجموعات لنحصل من هذه الخطوة على رتب درجات المتغير التابع RY ورتب درجات المتغير المصاحب RX .

٢. يحسب الانحدار الخطى البسيط لرتب المتغير التابع RY على رتب المتغير المصاحب RX لنحصل من هذه الخطوة على معادلة الانحدار الخطى والتى تشمل القيم المتنبأ بها $\hat{RY}$ والثابت B وقيمة معامل الانحدار β .

٣. يعاد حساب الانحدار الخطى لرتب المتغير التابع على رتب المتغير المصاحب والمتغير $G1$ والمتغير $G2$ لنحصل من هذه الخطوة على معادلة الانحدار الخطى والتى تشمل القيم المتنبأ بها $\hat{RY}$ والثابت B وقيمة معامل الانحدار β الخاصة بالمتغير المصاحب

والمتغير $G1$ والمتغير $G2$.

٤. تحسب قيمة الاختبار من المعادلة التالية :

$$(TS)_{MPF} = \frac{\left\{R^2_{Y.123} - R^2_{Y.1}\right\}(N - P_2)}{\left\{1 - R^2_{Y.123}\right\}(P_2 - P_1)}$$

ثالثاً : اختبار بيرنت وبار Burnett and Barr's Test

طور هذا الاختبار Burnett and Barr (1977) كبديل لابارامترى لاختبار تحليل التغاير البارامترى ، ويسمى اختبار فروق رتب الدرجات Rank Difference Scores ، ويعد البديل الأسهل فى حسابه من بين البدائل الأخرى . وهذا الاختبار هو صورة معدلة من اختبار Quade (1967) الذى يفترض أن معامل ارتباط رتب المتغير التابع ورتب المتغير المصاحب تساوى الواحد الصحيح .

وتتلخص خطوات حساب الاختبار فى التالى :

١. ترتب الدرجات الخام الخاصة بالمتغير التابع Y والدرجات الخاصة بالمتغير المصاحب X كل على حده خلال جميع المجموعات لنحصل من هذه الخطوة على رتب درجات المتغير التابع RY ورتب درجات المتغير المصاحب RX .

٢. يحسب الفرق بين رتب درجات المتغير التابع RY ورتب درجات المتغير المصاحب RX لنحصل من هذه الخطوة على الفروق $W = (RY - RX)$.

٣. يحسب الانحدار الخطى البسيط للفروق بين رتب درجات المتغير التابع RY ورتب درجات المتغير المصاحب RX على المتغير $G1$ والمتغير $G2$ لنحصل من هذه الخطوة على معادلة الانحدار الخطى والتى تشمل القيم المتنبأ بها $\hat{d}$ والثابت B وقيمة معامل الانحدار β الخاصة بالمتغير $G1$ والمتغير $G2$.

٤. تحسب قيمة الاختبار من المعادلة التالية :

$$F = \frac{\left\{R^2_{Y.12}\right\}(N - K)}{\left\{1 - R^2_{Y.12}\right\}(K - 1)}$$

حيث k عدد المجموعات .

وطريقة رتبة درجة الاكتساب Rank Gain Score Approach التى اقترحها Burnett and Barr (1977) يمكن أن تستخدم فى التصميمات العاملية لكنها محددة بمتغير مصاحب وحيد وتفترض أن العلاقة بين القياس البعدى والقبلى تظل كما هى خلال جميع المجموعات (Olejnik & Algina, 1985) .

رابعاً : حل النموذج الخطى العام لشيرلى

Shirley's General Linear Model Solution

طور هذا الاختبار Shirley (1981) كبديل لابارامترى لاختبار تحليل التغاير البارامترى ، ويشبه اختبار McSweeney and Porter (1968-1971) ، ويعتمد الاختبار على بحوث Bennett (1968) الذى طور النموذج الخطى العام لاختبار النماذج الخطية القائمة على الرتب .

ويتم حساب الاختبار بأخذ نسبة مجموع المربعات المعدل (المضبوط) للعامل الجمعى Adjusted Sum of Squares for the Grouping Factor والخطأ الكلى لمربع المتوسط لرتب المتغير التابع غير المعدلة The Total Mean Square Error for the Unadjusted Posttest Ranks .

وتتلخص خطوات حساب الاختبار فى التالى :

١. ترتب الدرجات الخام الخاصة بالمتغير التابع Y والدرجات الخاصة بالمتغير المصاحب X كل على حده خلال جميع المجموعات لنحصل من هذه الخطوة على رتب درجات المتغير التابع RY ورتب درجات المتغير المصاحب RX .

٢. يحسب الانحدار الخطى البسيط لرتب المتغير التابع RY على رتب المتغير المصاحب RX لنحصل من هذه الخطوة على معادلة الانحدار الخطى والتى تشمل القيم المتنبأ بها $\hat{RY}$ والثابت B وقيمة معامل الانحدار β كما يلى :

$$\hat{RY} = B + \beta RX$$

٣. يحسب الانحدار الخطى البسيط مرة أخرى باعتبار انحدار رتب المتغير التابع على رتب المتغير المصاحب والمتغير $G1$ والمتغير $G2$ فنحصل على معادلة الانحدار التالية :

$$\hat{RY} = B + \beta_1 RX + \beta_2 G1 + \beta_3 G2$$

٤. يحسب مجموع المربعات المعدل للعامل الجمعى كناتج لحاصل ضرب مجموع المربعات الكلى المحسوب لرتب القياس البعدى غير المعدلة (رتب المتغير التابع) والفرق بين (مقدار الزيادة) فى قيمة مربع معامل الارتباط المتعدد ومربع معامل ارتباط سبيرمان بين رتب درجات المتغير التابع RY ورتب درجات المتغير المصاحب RX كما بالمعادلة التالية :

$$SS_{total(adjusted)} = (R^2_{Y.123} - R^2_{Y.12})\left\{\frac{N(N^2-1)}{12}\right\}$$

٥. حساب مربع المتوسط الكلى The Total Mean Square لرتب القياس البعدى غير المعدلة كما بالمعادلة التالية :

$$SS_{total} = \left\{\frac{N(N+1)}{12}\right\}$$

٦. تحسب قيمة الاختبار $(TS)_{Shirley}$ بحساب خارج قسمة ناتج الخطوة الرابعة على ناتج الخطوة الخامسة ويكشف عن دلالة الناتج باستخدام القيم الحرجة لاختبار مربع كاى بدرجة حرية تساوى عدد المعالجات مطروحاً منها الواحد .

خامساً : اختبار بيورى وسن Puri and Sen's Test

بعد سنتين من نشر Quade لاختباره كبديل لاختبار تحليل التغاير البارامترى ، طور Puri and Sen (1969a) بديلاً لابارامترياً آخر لاختبار تحليل التغاير يتميز بالعمومية أكثر من بديل Quade . ويمكن أن يستخدم الاختبار مع البيانات المرتبة أو مع أى تحويلات أخرى مثل الدرجات المعيارية بشرط أن ترتبط الدرجات المحولة تزايدياً بالبيانات الأصلية (الدرجات الخام) ، وتقارن قيمة الاختبار المحسوبة مع قيمة اختبار مربع كاى عند درجة حرية تساوى عدد المعالجات (المجموعات) مطروحاً منها الواحد $\chi^2_{(k-1)}$.

وتتلخص خطوات حساب الاختبار فى التالى :

١. تُرتب الدرجات الخام الخاصة بالمتغير التابع Y والدرجات الخام الخاصة بالمتغير المصاحب X كل على حده خلال جميع المجموعات لنحصل من هذه الخطوة على رتب درجات المتغير التابع RY ورتب درجات المتغير المصاحب RX .

٢. تُحسب متوسطات الرتب داخل المجموعات Within-Groups Means وهى $(\overline{r}_{Y.k}, \overline{r}_{X.k})$ ، وكذلك المتوسط العام الموزون وهو $(\overline{r}_Y, \overline{r}_X)$ ، ومن ثم حساب الانحرافات بين المتوسط العام للرتب ومتوسطات الرتب داخل المجموعات وذلك بالنسبة للمتغير المصاحب والمتغير التابع من المعادلة التالية :

$$d_{Y.k} = \overline{r}_{Y.k} - \overline{r}_Y$$

$$d_{X.k} = \overline{r}_{X.k} - \overline{r}_X$$

٣. يُحسب متوسط الباقى The Residual Mean داخل كل مجموعة وذلك بحساب الفرق بين انحراف المتوسط داخل المجموعة وذلك للمتغير التابع $(d_{Y.k})$ وانحراف المتوسط

المتنبأ به للمتغير التابع القائم على انحدار درجات انحراف الرتب للمتغير التابع $(\overline{r}_{Y.k}-\overline{r}_{Y})$ على درجات انحراف الرتب للمتغير المصاحب $(\overline{r}_{X.k}-\overline{r}_{X})$ خلال كل المجموعات ، ويُحسب متوسط الباقى من المعادلة التالية :

$$\overline{E}_k=(d_{Y.k})-r_S(d_{X.k})\ .$$

حيث r_S هو معامل ارتباط سبيرمان بين رتب المتغير التابع والمتغير المصاحب للعينة الكلية

٤. تُحسب مصفوفة التباين-التغاير Variance-Covariance Matrix (V) باستخدام N فضلاً عن $(N-1)$ فى المقام حيث يأتى تباين رتب المتغير التابع أولاً ثم التغاير ليمثل العنصر الثانى فى المصفوفة .

$$V=\begin{pmatrix} Var. & Covar. \\ Covar. & Var. \end{pmatrix}$$

ويُحسب تباين رتب المتغير التابع من المعادلة :

$$Variance_{(RY)}=\frac{\sum(RY)^2}{N}-\left\{\frac{\sum(RY)}{N}\right\}^2$$

ويُحسب التغاير لرتب المتغيرين التابع والمصاحب من المعادلة[⊕] :

$$Covariance_{(RY.X)}=\frac{\sum(r_{RX}-\overline{r}_{RX})(r_{RY}-\overline{r}_{RY})}{N-1}=\frac{\sum(RX)(RY)-\frac{\sum RX\sum RY}{N}}{N-1}$$

٥. يُحسب العنصر V^{11} ويساوى هذا العنصر مقلوب تباين الخطأ الناتج من انحدار الرتب الفردية المحولة ، ويُعزل من العنصر الأول فى القطر الرئيسى للمصفوفة ، وهذا العنصر يتم حسابه من المعادلة التالية :

$$V^{11}=1/\left\{\frac{N(N^2-1)}{12}(1-r_S^2)\right\}$$

لتصبح المصفوفة فى الشكل التالى :

[⊕] إذا كانت أزواج القيم $(RX\ ,RY)$ هى عينات عشوائية من مجتمع ذى متغيرين وأردنا تقدير التغاير فى هذا المجتمع من التغاير فى العينة فإننا نقسم حواصل الضرب على $(N-1)$ بدلاً من N وذلك لكى يكون هذا التقدير تقديراً غير متحيز (محمد أبو يوسف ، ١٩٨٩ : ٤٤٢) .

$$V = \begin{pmatrix} Var. - v^{11} & Co\ var. \\ Co\ var. & Var. \end{pmatrix}$$

وتكون قيمة محدد المصفوفة :

$$\Delta = (v^{11})(v^{22}) + (v^{21})(v^{12})$$

وتكون المصفوفة المصاحبة :

$$V = \begin{pmatrix} Var. - v^{11} & -Co\ var. \\ -Co\ var. & Var. \end{pmatrix}$$

٦. نوجد معكوس أو مقلوب المصفوفة :

$$V^{-1} = \begin{pmatrix} \frac{v^{11}}{\Delta} & -\frac{v^{12}}{\Delta} \\ -\frac{v^{21}}{\Delta} & \frac{v^{22}}{\Delta} \end{pmatrix}$$

٧. تُحسب قيمة الاختبار من المعادلة التالية :

$$L_n = V^{11}(V^{-1})\sum_{k=1}^{k} n_k \bar{E}_k^2$$

واختبار Puri and Sen مثل اختبار Quade يمكن استخدامه فى التصميمات العاملية Factorial Designs ، ويمكن أن يشتمل على متغيرات مصاحبة متعددة Multiple Covariates ، ويفترض أيضاً أن العلاقة بين القياس القبلى والقياس البعدى مقدار ثابت عبر المجموعات (Puri & Sen, 1969a, 1969b; Olejnik & Algina, 1985) .

سادساً : اختبار بيورى-سن-هارول-سيرلن Puri-Sen-Harwell-Serlin (PSHS) Test

اقترح Harwell and Serlin (1988) بديلاً لابارامترياً لاختبار تحليل التغاير ، يعد تطويراً للبديل الذى طوره Puri and Sen (1969a,1969b) ويناسب حالة وجود متغير مصاحب وحيد ومتغير تابع وحيد لاختبار الفرضية التالية :

$$H_o : E(R_1^A) = E(R_2^A) = \ldots\ldots\ldots\ldots = E(R_k^A)$$

والتى يمكن التعبير عنها فى الصورة التالية :

$$H_o : \rho_y D_1, \ldots\ldots\ldots\ldots, D_{k-1} X = 0$$

حيث $E(\bar{R}_k^A)$ هو الرتبة المتوقعة لمتوسط المتغير المصاحب للمجموعة $(k-1)$ ، $\rho_y D_1, \ldots\ldots, D_{k-1}$ هى معاملات الارتباط بين المتغير التابع وعدد من المجموعات على المتغيرات D_k مع استبعاد تأثيرات المتغير المصاحب وتعتمد الفرضية السابقة على أن جميع المجموعات

لها نفس الرتبة المتوقعة لمتوسط المتغير المصاحب .

ويعطى الاختبار من المعادلة التالية :

$$(TS)_{PSHS} = (N-2)\eta^2$$

حيث η^2 مجموع المربعات بين المجموعات مقسوماً على مجموع المربعات الكلى ويمثل ذلك تباين الرتب المصاحبة المعدلة المفسرة بواسطة المتغير التابع ، N تمثل حجم العينة الكلية ، وفى ضوء الفرض الصفرى ومعادلة الاختبار ، فإن الاختبار يتوزع تقاربياً فى صورة توزيع مربع كاى المركزى بدرجة حرية $(k-1)$ حيث k عدد المجموعات .

وتتبع الخطوتين التاليتين لحساب قيمة الاختبار :

١. ترتب الدرجات الخام الخاصة بالمتغير التابع Y والدرجات الخام الخاصة بالمتغير المصاحب X كل على حده خلال جميع المجموعات لنحصل من هذه الخطوة على رتب درجات المتغير التابع RY ورتب درجات المتغير المصاحب RX .

٢. إدخال رتب المتغيرين والتعامل معهما كدرجات فى برنامج لحساب تحليل التغاير البارامترى وعندئذ يكون مجموع المربعات بين المجموعات ممثلاً لمجموع مربعات التأثير الأساسى ويكون مجموع المربعات الكلى ممثلاً لمجموع المربعات بين المجموعات مضافاً إلى مجموع المربعات داخل المجموعات .

ويمكن تمديد الاختبار ليشمل حالة تحليل التغاير المتعدد (أكثر من متغير مصاحب) (Harwell & Serlin, 1988) .

سابعاً : اختبار هتمانسبرجر Hettmansperger Test

اقترح Hettmansperger (1984: 251-275) اختباراً للرتب يتطلب صف Align المشاهدات الأصلية باعتبار البارامترات الدخيلة ، ويتوزع هذا الإحصاء تقاربياً مثل توزيع مربع كاى بدرجة حرية $(k-1)$ حيث k عدد المجموعات .

ويتم حساب احصاء مربع كاى لـ Hettmansperger (1984) كما يلى :

١. يتم الحصول على البواقى RES من انحدار المتغير التابع على النموذج المخفض Reduced Model الذى يشمل المتغير المصاحب فقط فى حالة تحليل التغاير البسيط أما فى حالة تحليل التباين العاملى فتضاف المتغيرات الجمعية Grouping Variables إلى النموذج المخفض .

٢. ترتيب البواقى RES دون اعتبار لعضوية المجموعة من 1 إلى N وتسمى البواقى المرتبة RRES .

٣. يتم الحصول على البواقى المعيارية المرتبة Standardized Ranked Residuals (SRRES) (الرتب الموزونة) Weighted Ranks من المعادلة التالية :

$$SRRES = \sqrt{12}\left[\left(\frac{RRES}{N+1} - \frac{1}{2}\right)\right]$$

٤. تقدم البواقى المعيارية المرتبة SRRES إلى برنامج إحصائى مناسب لإجراء تحليل التباين .

٥. يُقارن مجموع المربعات بين المجموعات بقيمة اختبار مربع كاى بدرجة حرية $(k-1)$ حيث k عدد المجموعات .

٦. فى حالة تمديد الاختبار لحالة تحليل التغاير العاملى ثنائى الاتجاه Two-Way ANCOVA تقارن قيمة مجموع مربعات التفاعل مع قيمة مربع كاى بدرجة حرية $(I-1)(j-1)$ فى الخطوة رقم (٥) السابقة .

الخصائص التوزيعية للاختبار الإحصائى

Distributional Properties of A Statistical Test

تتمثل الخصائص التوزيعية لأى اختبار إحصائى فى الخطأ من النوع الأول والخطأ من النوع الثانى وقوة الاختبار الإحصائى وضلاعة الاختبار الإحصائى .

(١) الخطأ من النوع (النمط) الأول A Type I Error

يقع الباحث فى الخطأ من النوع الأول عندما يكون بارامتر الأصل مساوياً لإحصاءة العينة ، ومعنى ذلك أن العينة مشتقة من هذا الأصل (أى الفرض الصفرى صحيح) ، ومع ذلك فإن الباحث يرفض هذا الفرض الصفرى (فؤاد أبو حطب ، آمال صادق ، ١٩٩٦ : ٣٤١ – ٣٤٣) . وهذا الاحتمال يطلق عليه خطأ ألفا α أو الخطأ من النوع الأول (عزت عبد الحميد ، ٢٠١١ : ١٩٥) .

وتتحدد درجة احتمالية هذا النوع من الخطأ بمستوى الدلالة الإحصائية α ، والذى يجب أن يُحدده الباحث مسبقاً ، وعادة يختار الباحثون فى العلوم الاجتماعية مستويات دلالة (٠.٠٥) أو (٠.٠١) وذلك ما لم تكن هناك اعتبارات تجعلهم أكثر تحفظاً أو خوفاً من الوقوع فى هذا النوع من الخطأ (على ماهر خطاب ، ٢٠١١ : ٢٨٣) .

والخطأ من النوع الأول هو الخطأ الموجب الذى يحدث عندما يكون القرار رفض الفرض الصفرى وهو فى الحقيقة لا يجب رفضه (Dyer, 1995: 357) . ويشير هذا الخطأ إلى اختبار التحديد السيىء Test of Poor Specificity ، ويمكن أن ينظر إلى الخطأ من النوع الأول على أنه خطأ السذاجة المفرطة Error of Excessive Credulity (عزت عبد الحميد ، ٢٠١١ :

١٩٥) .

والخطأ من النوع الأول هو قرار رفض الفرض الصفرى عندما يكون صحيحاً . وإذا كان القرار هو الفشل فى الرفض عندئذ لا يمكن الوقوع فى الخطأ من النوع الأول ؛ لأن الخطأ من النوع الأول يحدث فقط إذا كان القرار هو الرفض . وإذا كان القرار هو الرفض عندئذ فإنه من غير الممكن أن نعرف على وجه التحديد (اليقين) ما إذا كان أو لم يكن الوقوع فى الخطأ من النوع الأول لأنه من غير الممكن معرفة ما إذا كان الفرض الصفرى صحيح (Algina, 2014) . والاحتمال المكمل لاحتمال الخطأ من النوع الأول يساوى $(1-\alpha)$ ويسمى درجة الثقة Confidence Degree (Park, 2010) .

وعندما يتم إجراء اختبار الفرضية يجب تعيين قيمة ألفا . وإذا كانت جميع افتراضات الاختبار صحيحة يكون معدل (تقدير) الخطأ من النوع الأول يساوى ألفا . وإذا كانت جميع الافتراضات غير صحيحة ربما يكون معدل الخطأ من النوع الأول لا يساوى ألفا . ويمكن حساب معدل الخطأ من النوع الأول من تكامل الدالة الرياضية Integrating the Mathematical Function التى هى توزيع المعاينة Sampling Distribution للاختبار الإحصائى ، وهذا يتطلب : (أ) معرفة صيغة توزيع المعاينة ، (ب) القدرة على حساب تكامل الدالة الرياضية التى هى توزيع المعاينة ، وهذا التكامل يجرى عادة بواسطة الحاسب الآلى (Algina, 2014) .

(٢) الخطأ من النوع (النمط) الثانى A Type II Error

يقع الباحث فى الخطأ من النوع الثانى عندما يكون بارامتر الأصل ليس مساوياً لإحصاءة العينة ، ومعنى ذلك أن العينة مشتقة من أصل مختلف (أى أن الفرض الصفرى خطأ) ، ومع ذلك فإن الباحث يقبل هذا الفرض الصفرى (فؤاد أبو حطب ، آمال صادق ، ١٩٩٦ : ٣٤١-٣٤٣) . وهذا الاحتمال يطلق عليه خطأ بيتا β أو الخطأ من النوع الثانى (عزت عبد الحميد ، ٢٠١١ : ١٩٥) .

وتتحدد درجة احتمالية الخطأ من النوع الثانى بمستوى الدلالة الذى يطلق عليه بيتا β ، والعلاقة بينه وبين الخطأ من النوع الأول علاقة عكسية فعندما تقل احتمالية حدوث أحدهما تزداد احتمالية حدوث النوع الآخر (عبد الرحمن عدس ، ١٩٩٧ : ٣٥ ؛ على ماهر خطاب ، ٢٠١١ : ٢٨٤) .

ويهتم العلماء بتحديد مستوى الدلالة الإحصائية بحيث يمكن التقليل من خطأ احتمالية رفض الفرض الصفرى الحقيقى ، أى تجنب الوقوع فى الخطأ من النوع الأول أكثر من الوقوع فى الخطأ من النوع الثانى . لذلك فالتحفظ الأساسى لدى العلماء يؤدى بهم إلى اختيار مستوى

رفض منخفض حتى يكون حدوث الخطأ من النوع الأول أقل تكراراً وحدوثاً من الخطأ من النوع الثانى . ويشير ذلك إلى أن الباحثين يفضلون خطأ قبول الفرض الصفرى الزائف عن الوقوع فى خطأ رفض الفرض الصفرى الحقيقى (على ماهر خطاب ، ٢٠١١ : ٢٨٤) .

والخطأ من النوع الثانى هو الخطأ السالب الذى يحدث عندما يكون القرار قبول الفرض الصفرى وهو فى الحقيقة يجب رفضه (Dyer, 1995: 384) . ويشير هذا الخطأ إلى اختبار الحساسية السيئة Test of Poor Sensitivity ، ويمكن أن ينظر إلى الخطأ من النوع الثانى على أنه خطأ الشكوكية المفرطة Error of Excessive Skepticism (عزت عبد الحميد ، ٢٠١١ : ١٩٥) .

والخطأ من النوع الثانى هو قرار فشل رفض الفرض الصفرى عندما يكون خاطئاً . وإذا كان القرار هو رفض عندئذ لا يمكن الوقوع فى الخطأ من النوع الثانى لأن الخطأ من النوع الثانى يحدث فقط إذا كان القرار هو فشل الرفض . وإذا كان القرار هو فشل الرفض عندئذ فإنه من غير الممكن أن نعرف على وجه التحديد (اليقين) ما إذا كان أو لم يكن الوقوع فى الخطأ من النوع الثانى لأنه من غير الممكن معرفة ما إذا كان الفرض الصفرى خاطىء (Algina, 2014) .

ويعتمد تقدير الخطأ من النوع الثانى على كيفية اختلاف توزيعات المجتمعات التى اشتقت منها العينات . ويمكن حساب معدل الخطأ من النوع الثانى من تكامل الدالة الرياضية التى هى توزيع المعاينة للاختبار الإحصائى ، وهذا يتطلب : (أ) معرفة الصيغة الرياضية لتوزيع المعاينة ، (ب) القدرة على حساب تكامل الدالة الرياضية التى هى توزيع المعاينة ، وهذا التكامل يجرى عادة بواسطة الحاسب الآلى (Algina, 2014) .

(٣) قوة الاختبار الإحصائى Power of a Statistical Test

تعد القوة خاصية للاختبار الإحصائى عندما يكون بارامتر الأصل ليس مساوياً لإحصاءة العينة (أى أن الفرض الصفرى خطأ) ويرفض الباحث هذا الفرض الصفرى واحتمال رفض الفرض الصفرى الخاطىء هو قرار صحيح بالطبع (فؤاد أبو حطب ، آمال صادق ، ١٩٩٦ : ٣٤١-٣٤٣) . بمعنى أن القوة تشير إلى قدرة الاختبار الإحصائى على اكتشاف العلاقة أو الفروق الحقيقية ، أى أنها مقياساً لقدرة الاختبار الإحصائى على أداء الوظيفة التى صمم من أجلها عندما تتوافر بيانات مناسبة لتطبيقه (Peers, 1996) .

وهى احتمال أن ينتج الاختبار الإحصائى نتائج دالة إحصائياً ، وترتبط بالخطأ من النوع الثانى ويتم حساب قوة الاختبار الإحصائى من المعادلة التالية :

$$Statistical\ Power = 1 - \beta$$

أى أن قوة الاختبار الإحصائى هى الاحتمال المكمل لقيمة بيتا β ، فكلما إزداد حجم بيتا انخفض مقدار قوة الاختبار ، وهو يمتد عادة من (صفر) إلى (١) وتعتبر قوة الاختبار مقبولة فى البحوث الإنسانية حين تمتد قيمها من (٠.٤٠) إلى (٠.٦٠) (زكريا الشربينى ، ١٩٩٠ ، ٦٤-٦٥ ؛ عزت عبد الحميد ، ٢٠١١ : ٢٠٤ - ٢٠٦) .

أى أن قوة اختبار الدلالة هى احتمالية تجنب الخطأ من النوع الثانى ، وهناك ثلاث طرق لزيادتها والحد من احتمال الوقوع فى الخطأ من النوع الثانى وهى : زيادة احتمال الخطأ من النوع الأول α ، زيادة حجم العينات ، وإنقاص التباين المقدر للمجتمع σ^2 (ج. ميلتون سميث ، ١٩٧٨ : ١١١) .

ويستكشف تحليل القوة الإحصائية العلاقة بين أربعة مكونات هى : حجم التأثير Effect Size الذى يعبر عن درجة وجود الظاهرة ، وحجم التأثير الكبير يزيد من قوة الاختبار الإحصائى ، والعكس صحيح . ومستوى الدلالة الإحصائية حيث تنخفض قوة الاختبار الإحصائى عندما تكون مستويات الدلالة أكثر تحفظاً وانخفاضاً ، وترتفع قوة الاختبار الإحصائى بزيادة مستويات الدلالة . وحجم العينة وهو عدد المشاهدات أو الحالات فى العينة ، وتزداد قوة الاختبار الإحصائى بزيادة حجم العينة ، والعكس صحيح . ومقدار التباين Variability حيث تقل قوة الاختبار الإحصائى بارتفاع مقدار التباين ، وتزداد بانخفاض مقدار التباين ، وهذا يعنى أن قوة الاختبار الإحصائى تقل بانخفاض معامل الثبات (Park, 2010) .

(٤) ضلاعة (منعة) الاختبار الإحصائى Robustness of a Statistical Test

يذكر Bradley (1968: 25) أنه من الوجهة التجريبية فإن أى تجاهل لافتراضات الاختبار البارامترى يفسد توزيع الاختبار الإحصائى ويغير تقديرات (معدلات) الخطأ من النوع الأول والخطأ من النوع الثانى حيث يتأثر الخطأ من النوع الأول بالتجاهل كثيراً عن تأثر الخطأ من النوع الثانى ، ويقال أن الاختبار يتمتع بالضلاعة إزاء التجاهل لافتراض معين إذا كان سلوكه (القيمة أو توزيع المعاينة) لا يتأثر كثيراً بالانحراف عن شروط النظرية الاعتدالية .

وتعبر ضلاعة الاختبار الإحصائى عن عدم الحساسية للانحرافات البسيطة عن الشروط أو الافتراضات التى يستند إليها الاختبار (Huber, 1981: 1) . وتمثل قدرة الاختبار الإحصائى على رفض الفرض الصفرى عند مستوى (٠.٠٠٥ ، ٠.٠٠١) فى غياب بعض الافتراضات التى يجب توافرها عند تطبيق الاختبار .

والسلوك الحقيقى لاحتمال الخطأ من النوع الأول فى ضوء تجاهل الافتراض معقد تماماً ويعتمد على العديد من العوامل المتفاعلة ، وعلى هذا فالضلاعة تعد مفهوم محل نقاش وجدال كبير

. وبالرغم من شمولية الأطر النظرية فيما يختص بالضلاعة فإن السيكولوجيين يبدون متأثرين فى البداية بالمعالجة الرياضية والدراسات المتعلقة بالمعاينة وقد عمدوا فى ضوء ذلك إلى عدم استخدام التعريفات الكمية للضلاعة (Bradley, 1978) .

وتعتمد درجة ضلاعة الاختبار الإحصائى إزاء تجاهل افتراض على مجموعة عوامل ليست متضمنة فى افتراضات الاختبار ، وتسبب تلك العوامل تشويه لتقديرات الخطأ من النوع الأول والخطأ من النوع الثانى ، وهذه العوامل هى : العوامل التى تقتصر على القياس ، والعوامل التى تتصل مباشرة بالمعاينة وتكون مقصورة على المعاينة والقياس ، والعوامل التى تتضمن المجتمعات وليس الافتراضات .

وبالاضافة إلى العوامل السابقة فإن هناك عوامل أخرى لا ترتبط بالافتراضات أو الوصف الكلى لتلك الافتراضات وهى : الأشكال النسبية لمجتمعات العينة ، والمقادير النسبية للارتباط بين متوسط العينة وتباين العينة للعينات المختلفة (Bradley, 1968: 26-27) .

وقد اقترح Bradley (1978) المحك الأكثر تشدداً (تعنتاً) More Stringent Criterion للضلاعة إذا حقق الاختبار الإحصائى المتباينة $0.9\alpha \le \alpha^{*} \le 1.1\alpha$ ، حيث α مستوى الدلالة الإحصائية ، α^{*} الاحتمال الفعلى للخطأ من النوع الأول ، كما اقترح المحك التحررى (الأقل شدة) Liberal Criterion للضلاعة إذا حقق الاختبار الإحصائى المتباينة $\left|\alpha^{*} - \alpha\right| \le \frac{\alpha}{2}$ التى يمكن أن تكتب فى الصورة التالية : $0.5\alpha \le \alpha^{*} \le 1.5\alpha$.

دراسات وبحوث سابقة

أجرى Olejink and Algina (1983) دراسة هدفت إلى مقارنة أداء اختبار تحليل التغاير البارامترى واختبار تحويل الرتب من حيث نسبة الخطأ من النوع الأول والقوة الإحصائية فى أربع حالات : كون أخطاء التوزيع المشروط اعتدالية ومتجانسة ، وأخطاء التوزيع المشروط اعتدالية وغير متجانسة ، وأخطاء التوزيع المشروط غير اعتدالية ومتجانسة ، وأخطاء التوزيع المشروط غير اعتدالية وغير متجانسة باستخدام مولدة بالمحاكاة ، ومن بين ما توصلت إليه نتائج الدراسة أن اختبار تحليل التغاير البارامترى يتسم بالضلاعة فيما يتصل بتجاهلات الاعتدالية وتجانس التباين ، ووجد أنه عند تجاهل الافتراضين معاً أن مستويات ألفا المشاهدة تخفض مستوى ألفا الاسمى فى حالة أحجام العينات الصغيرة وقيمة ألفا (٠.٠٥) ، ويصبح اختبار تحويل الرتب اختباراً متحرراً بشكل طفيف عندما يكون توزيع المتغير المصاحب غير اعتدالى والأخطاء غير متجانسة ، وتميز اختبار تحويل الرتب بالقوة النسبية فى حالة أحجام العينات المتوسطة وتوزيعات الخطأ الملتوية .

وهدفت دراسة Seaman (1984) إلى استخدام فنيات مونت كارلو لتقدير احتمالية الخطأ من النوع الأول لاختبار تحليل التغاير البارامترى واختبار تحويل الرتب اللابارامترى لثلاث توزيعات مشروطة فى حالة تصميم من مجموعتين : التوزيعان اعتداليان ، التوزيعان ملتويان فى نفس الاتجاه ولكن بدرجات مختلفة ، التوزيعان ملتويان بنفس الدرجة ولكن فى اتجاهين مختلفين فى حالة تساوى وعدم تساوى المتوسطين ، ومن بين ما توصلت إليه نتائج الدراسة أن اختبار تحليل التغاير البارامترى يفضل استخدامه فقط فى موقف التوزيعين الاعتداليين للمجموعتين .

أما دراسة Olejnik and Algina (1984) فقد هدفت إلى فحص أداء اختبار تحليل التغاير البارامترى واختبار تحويل الرتب RT فى حالة إهمال شرطى الاعتدالية وتجانس التباين باستخدام بيانات محاكاة فى حالة تجانس التباين والاعتدالية للتوزيع المشروط للأخطاء وفى حالة عدم الاعتدالية وعدم تجانس التباين وفى حالة الاعتدالية وعدم تجانس التباين ، ومن بين ما توصلت إليه نتائج الدراسة أن اختبار تحليل التغاير البارامترى يتميز بالضلاعة لتجاهل افتراضى الاعتدالية وتجانس التباين ، وعند تجاهل الافتراضين معاً فإن مستويات المشاهدة تكون منخفضة عند المستوى الاسمى ($\alpha = .05$) عندما تكون أحجام العينات صغيرة وبعد اختبار تحويل الرتب متحرر بشكل ضعيف ، كما وجد أنه عندما يكون توزيع المتغير المضاحب غير اعتدالى وحجم العينة صغير والأخطاء غير متجانسة أن اختبار تحويل الرتب يتميز بالقوة .

كما أجرى Olejnik and Algina (1985) دراسة هدفت إلى فحص الخطأ من النوع الأول والقوة لاختبار تحليل التغاير البارامترى وخمسة بدائل لابارامترية باستخدام بيانات مولدة بالمحاكاة ، ومن بين ما توصلت إليه نتائج الدراسة أن جميع البدائل تعطى نفس الاستنتاجات مثل الاختبار البارامترى فيما يتصل بتأثير المعالجات الثلاث عند مستوى (٠.٠٥) ، كما أن قيم الاحتمال الخاصة بجميع البدائل اللابارامترية (احتمالات الخطأ من النوع الأول) أقل من قيمة الاحتمال الخاص بالاختبار البارامترى ، كما وجد زيادة طفيفة فى القوة الإحصائية للبدائل اللابارامترية عن الاختبار البارامترى .

أما Seaman, Algina and Olejnik (1985) فقد أجرى دراسة هدفت إلى مقارنة احتمالية الخطأ من النوع الأول والقوة لإجراءات تحليل التغاير البارامترى والرتبية تحت شروط تتضمن توزيعات مشروطة تختلف بين المجموعات فى الالتواء ، ومن بين ما توصلت إليه نتائج الدراسة أن تقديرات ألفا الفعلية تقع خارج الفئات المحكية (الحدود) وغير مقبولة فى حالة استخدام الاختبار البارامترى وكل الإجراءات اللابارامترية ، ووجدت ميزة القوة العملية للإجراءات الرتبية فى حالة أحجام العينات المتوسطة للتوزيعات غير الاعتدالية والتباينات الصغيرة والكبيرة .

وهدفت دراسة Olejnik, Algina and Abdel-Fattah (1987) إلى تقدير قوة اختبار تحليل التغاير البارامترى واختبار تحويل الرتب تحت شرطى عدم الاعتدالية وعدم تجانس التباين باستخدام بيانات إمبيريقية ، ومن بين ما توصلت إليه نتائج الدراسة أن قوة اختبار تحويل الرتب تأثرت بشرطى عدم الاعتدالية وعدم تجانس التباين ، ولم تتأثر قوة الاختبار البارامترى بذلك ، كما توصلت أيضاً إلى أن الفرق فى القوة بين اختبار تحليل التغاير يمكن أن يكون موجب أو سالب ويمكن أن يكون تافه أو مهم اعتماداً على حجم تأثير المعالجة ودرجة الارتباط بين المتغير المصاحب ودرجات القياس البعدى ومدى الاعتدالية ودرجة عدم التجانس .

وأجرى Harwell and Serlin (1988) دراسة هدفت إلى مقارنة القوة والخطأ من النوع الأول لأربعة بدائل لابارامترية لاختبار تحليل التغاير البارامترى هى : اختبار PSHS ، واختبار Rogosa ، واختبار McSweeney and Porter ، واختبار Hettmansperger ، ومن بين ما توصلت إليه نتائج الدراسة تميز اختبار تحليل التغاير البارامترى فى حالة التوزيع الأعتدالى والتوزيع القريب من الإعتدالى وتساوى خطوط ميل الانحدار عن البدائل اللابارامترية ، وتميز اختبار تحليل التغاير البارامترى عن اختبار Rogosa فى حالة التوزيع غير الاعتدالى ، وتميز اختبار Hettmansperger عن البدائل اللابارامترية الأخرى فى حالة التوزيع غير الاعتدالى حتى فى حالة عدم تساوى أحجام العينات وعدم تساوى خطوط ميل الانحدار ، وتميز اختبار PSHS بالمحافظة على تقدير الخطأ من النوع الأول عند أقل من مستوى الدلالة الإحصائية مقارنة مع اختبار McSweeney and Porter ، كما تتخطى قوة اختبارى PSHS, McSweeney and Porter قوة اختبار تحليل التغاير البارامترى فى حالة التوزيع غير الاعتدالى .

أما دراسة Alyasin (1991) فقد هدفت إلى الكشف عن تأثير حجم العينة وعدم توافر شرطى الاعتدالية وتجانس التباين على الخطأ من النوع الأول وقوة اختبار تحليل التباين أحادى الاتجاه واختبار كروسكال-واليس واختبار L اللابارامترى واختبار تحليل التغاير أحادى الاتجاه اللابارامترى ، ومن بين ما توصلت إليه نتائج الدراسة تميز أداء اختبار PSHS عن اختبار تحليل التغاير البارامترى من حيث القوة وتقديرات الخطأ من النوع الأول .

وأجرى Headrick and Sawilowsky (2000) دراسة هدفت إلى فحص خصائص القوة وتقديرات الخطأ من النوع الأول لاختبار تحويل الرتب اللابارامترى RT كبديل لاختبار تحليل التغاير العاملى ، ومن بين ما توصلت إليه نتائج الدراسة أن كلتا الطريقتين أظهرتا شذوذ Erratic فى تقديرات الخطأ من النوع الأول فى حالة تجانس معاملات الانحدار وإهمال افتراض الاعتدالية ، وفشل اختبار تحويل الرتب فى تفسير التفاعلات فى حالة توافر الشروط التى يستند

إليها اختبار تحليل التغاير البارامترى وذلك نظراً لتضخم تقديرات الخطأ من النوع الأول ، كما وجد تضخم فى تقديرات الخطأ من النوع الأول عندما كان الانحراف عن افتراض الاعتدالية كبير جداً بينما لا يحدث ذلك مع اختبار تحويل الرتب اللابارامترى .

أما دراسة Headrick and Vineyard (2001) فقد هدفت إلى مقارنة تقديرات الخطأ من النوع الأول وخصائص القوة لاختبار تحليل التغاير وثلاثة بدائل لابارامترية هى : اختبار PS، واختبار Blair/Sawilowsky ، واختبار Hettmansperger ، ومن بين ما توصلت إليه نتائج الدراسة أن اختبار PS قد أظهر تقديرات محافظة بشدة Ultra-Conservative للخطأ من النوع الأول وانخفاض فى القوة فى حالة التأثيرات الأساسية ، وأظهر اختبار Blair/Sawilowsky تقديرات متحررة Liberal للخطأ من النوع الأول فى حالة المعاينة القائمة على التوزيعات بسيطة وشديدة الالتواء ، كما أظهر اختبار Hettmansperger تقديرات مقبولة للخطأ من النوع الأول فى حالة كل التوزيعات التى تمت دراستها كما اقترحت الدراسة أن هذا الاختبار يعتبر بديلاً مقبولاً فى حالة العينات المتساوية نسبياً والعينات التى حجمها أكبر من (٢٠) .

وهدفت دراسة Rheinheimer and Penfield (2001) إلى فحص أداء اختبار تحليل التغاير وست بدائل (إجراء Quade ، حل Puri and Sen ، اختبار رتب فروق الدرجات لـــــ Burnett and Barr ، اختبار تحويل الرتب لـــ Conover and Iman ، إجراء Hettmansperger ، اختبار Puri-Sen-Harwell-Serlin) تحت شروط تجريبية متفاوتة (افتراض الاعتدالية وتجانس التباين ، حجم العينة ، عدد مجموعات المعالجة ، قوة العلاقة بين المتغير المصاحب والمتغير التابع ، توليفات متعددة من الشروط) من خلال المحاكاة بطريقة مونت كارلو ، ومن بين ما توصلت إليه نتائج الدراسة أن عدم توافر شرط تجانس التباين وبخاصة فى حالة التوليفة (التصميمات غير المتوازنة وعدم الاعتدالية الشديد) له تأثير كبير على تقديرات الخطأ من النوع الأول ، ويتصف اختبار تحليل التغاير البارامترى بالضلاعة ، وغالباً يكون الأكثر قوة فى حالة التصميمات المتوازنة والتكوينات التوزيعية وأظهر قوة مرتفعة تحت شرط تجانس التباين وفى بعض حالات عدم تجانس التباين ولكنه أقل منافسة كلما انحرفت الشروط عن الاعتدالية .

أما دراسة Shull and Nakonezny (2002) فقد هدفت إلى اشتقاق إمبريقى لمنحنيات القوة وروتينات SAS لبعض البدائل اللابرامترية لاختبار تحليل التغاير البارامترى من خلال وضع Computer Code لكل بديل كتب بلغة (SAS version 8.1) بواسطة المحاكاة لتوليد تصميم متوازن من ثلاث مجموعات يجمع متغير مصاحب ومتغير تابع بأحجام مجموعات من ١٠

إلى ١٠٠ ومعامل ارتباط ($r = .3$) بين المتغير المصاحب والتابع ، ومن بين ما توصلت إليه نتائج الدراسة أن تقديرات الخطأ من النوع الأول كانت داخل حدود كوكران Cochran's Limits، كما وجد أن البدائل (McSweeney and Portet, Rogosa, Conover-Iman, Puri-Sen-Harwell-Serlin) تتميز بالقوة مقارنة بالاختبار البارامترى ، وتميزت البدائل (Quade, Hettmansperger) بالقوة المنخفضة فى حالة أحجام الخلايا أقل من (٥٠) ، وأن البديل Burnett-Barr لا يوصى به

وأجرى Wang and Akritas (2006) دراسة هدفت إلى دراسة دلالة التفاعلات باستخدام اختبار تحليل التغاير البارامترى العاملى وإحصاء NP اللابارامترى الذى اقترحه Akritas et al. (2000) واختبار YB شبه البارامترى الذى قدمه Young and Bowman (1995) واختبار The Drop Test الذى قدمه McKean and Schrader (1980) كبديل لابارامترى باستخدام بيانات مولدة بطريقة مونت كارلو وبيانات إمبيريقية حقيقية مستقاة من مجال العلوم الزراعية ، ومن بين ما توصلت إليه نتائج الدراسة تميز أداء الاختبار البارامترى مقارنة بالبدائل اللابارامترية الثلاث فى حالة توافر الافتراضات الأساسية له وتقاربت قوة اختبار The Drop Test من قوة الاختبار البارامترى .

<u>تعقيب على الدراسات والبحوث السابقة</u>

١. معظم الدراسات والبحوث السابقة مثل : Seaman (1984); Harwell and Serlin (1988); Rheinheimer and Penfield (2001); Headrick and Vineyard (2001); Headrick and Sawilowsky (2000) أجريت باستخدام بيانات محاكاة مولدة عن طريق برامج حاسب متخصصة .

٢. توجد دراسة واحدة استخدمت فئة بيانات إمبيريقية وفئة بيانات مولدة بالمحاكاة هى دراسة Wang and Akritas (2006) ، وتوجد دراسة واحدة فقط استخدمت فئة بيانات إمبيريقية فقط هى دراسة Olejnik, algina and Abdel-Fattah (1987) .

٣. تختلف الدراسة الحالية عن الدراسات والبحوث السابقة فى أنها تستخدم بيانات إمبيريقية وتتم المقارنة بين اختبار تحليل التغاير البارامترى وسبعة بدائل لابارامترية تحت شروط مختلفة .

٤. تباينت نتائج الدراسات والبحوث السابقة -المتاحة- إلى أن اختبار تحليل التغاير البارامترى يصلح فى مواقف معينة وتتوب عنه البدائل اللابارامترية فى مواقف أخرى والعكس بالعكس .

٥. ندرة البحوث والدراسات السابقة التى استهدفت تحرى ضلاعة اختبار تحليل التغاير أو

أحد البدائل اللابارامترية .

٦. بعض الدراسات اعتمدت فقط على المقارنة الثنائية بين اختبار تحليل التغاير البارامترى وأحد البدائل اللابارامترية مثل دراسة : Olejink and Algina (1983, 1984); Seaman (1984); Olejnik, algina and Abdel-Fattah (1987); Headrick and Sawilowsky (2000).

٧. تختلف الدراسة الحالية عن الدراسات والبحوث السابقة فى أنها تستخدم التصميم العشوائى ذى الثلاث مجموعات للقياس القبلى والبعدى The Randomized Pretest-Posttest Design with Three Groups فى حين أن جميع الدراسات السابقة استخدمت التصميم ذى المجموعتين .

فروض الدراسة

بمراجعة الاطار النظرى للدراسة والدراسات والبحوث السابقة تسعى الدراسة الحالية إلى اختبار الفرضين التاليين :

١. لا يختلف أداء اختبار تحليل التغاير البارامترى والبدائل اللابارامترية السبعة تحت شروط مختلفة فى تصميم غير متوازن من ثلاث مجموعات فى حالة استخدام عينات ذات أحجام (صغيرة جداً ، صغيرة ، متوسطة ، كبيرة ، كبيرة جداً) .

٢. لا تختلف الخصائص التوزيعية (تقديرات الخطأ من النوع الأول ، تقديرات الخطأ من النوع الثانى ، تقديرات الضلاعة ، تقديرات القوة الإحصائية) لاختبار تحليل التغاير البارامترى والبدائل اللابارامترية السبعة تحت شروط مختلفة فى تصميم غير متوازن من ثلاث مجموعات فى حالة استخدام عينات ذات أحجام (صغيرة جداً ، صغيرة ، متوسطة ، كبيرة ، كبيرة جداً) .

إجراءات الدراسة

تتحدد إجراءات الدراسة الحالية بالعينة وأداة الدراسة والمعالجة الإحصائية المستخدمة كالتالى :

أولاً : العينة

(أ) العينة الاستطلاعية

بلغ حجم العينة (١١٣) طالباً وطالبة من منتسبى الفصل الصيفى بالبرنامج التدريبى : إدارة الضغوط ضمن أنشطة مشروع الطرق المؤدية للتعليم العالى المنفذ بجامعة جنوب الوادى ، وهذه العينة استخدمت لفحص الخصائص السيكومترية لأداة القياس (ثبات الدرجات، صدق البنود)، وفحص خصائص البنود (معامل الصعوبة ، معامل التمييز) .

(ب) العينة الأساسية

بلغ حجم عينة الدراسة (٧٠١) طالباً وطالبة من منتسبى الفصلين الشتوى والصيفى بالبرنامج التدريبى : إدارة الضغوط ضمن أنشطة مشروع الطرق المؤدية للتعليم العالى المنفذ بجامعة جنوب الوادى ، وهذه العينة أخذت منها عينات ذات أحجام مختلفة طبقاً لمقتضيات تحليل البيانات .

ثانياً : أداة الدراسة

قام الباحث بإعداد اختبار تحصيلى لقياس درجة الاكتساب التى يحققها المتدرب فى البرنامج التدريبى (إدارة الضغوط) المقدم خلال مشروع الطرق المؤدية للتعليم العالى اتبعت فيه الخطوات المعروفة لبناء اختبار تحصيلى يقيس نواتج تعلم الطلاب المنتسبين بالبرنامج فى مستويات معرفية مختلفة .

ثالثاً المعالجة الإحصائية

قام الباحث بقياس التحصيل فى البرنامج التدريبى (إدارة الضغوط) قبل البدء فى التدريب من خلال استخدام ثلاث من الاستراتيجيات لتنفيذ البرنامج التدريبى هى : استراتيجية العصف الذهنى ، واستراتيجية التعلم التعاونى ، واستراتيجية KWL . ثم قاس الباحث التحصيل بنفس الأداة بعد الانتهاء من البرنامج التدريبى فى المدة المحددة وهى ٣ أيام تدريبية لمدة ٦ ساعات فى اليوم . واعتبر القياس القبلى متغير مصاحب X واعتبر القياس البعدى متغير تابع Y . وحدد الباحث خمس مواقف للقياس تمثل تصميم عشوائى غير متوازن من ثلاث مجموعات للمعالجة فى حالة استخدام عينات ذات أحجام (صغيرة جداً ، صغيرة ، متوسطة ، كبيرة ، كبيرة جداً) كما بالجدول التالى :

جدول (١)

أحجام العينات الفرعية فى التصميم التجريبى

موقف القياس	أحجام العينات			N	نوع حجم العينة
	n_1	n_2	n_3		
الأول	٤	٣	٢	٩	صغيرة جداً
الثانى	١٠	١٢	١٢	٣٤	صغيرة
الثالث	٣٦	٤١	٤٧	١٢٤	متوسطة
الرابع	١٢١	١١٠	١٣٣	٣٦٤	كبيرة
الخامس	٢٣١	٢٤٧	٢٢٣	٧٠١	كبيرة جداً

ومن ثم استكشف الباحث الخصائص الإحصائية للبيانات (معامل الالتواء ،

توازى خطوط ميل الانحدار ، تجانس التباينات للمجموعات الفرعية ، العلاقة الخطية بين المتغير التابع والمتغير المصاحب) فى مواقف القياس الخمسة كما بالجدول التالى :

جدول (٢)

الخصائص الإحصائية للبيانات في التصميم

مواقف القياس	معامل الانحراف		معاملات الانحدار			اختبار تجانس التباين								rho
	Y	X	b_1	b_2	b_3	Y				X				
						L	df1	df2	Sig.	L	df1	df2	Sig.	
الأول	٠,١٣٤	٠,٤٢٥	-٠,٠٢٣	-٠,٧٢٠	-٠,٥٩٠	٠,٧٤	٢	٦٠	٠,٥٢	٠,٩٥	٢	٦٠	٠,٤٤٠	٠,٤٣٢
الثانى	٠,٢٠١	٠,٣٣٩	-٠,٠٤٥	-٠,٥٧٠	-٠,٢٨٨	٢,١١	٢	٣٦	٠,١٤	٤,٩٥	٢	٣٦	٠,١٢	٠,٥٧١
الثالث	٠,٢٨٠	٠,٥٣٠	-٠,٢٦٩	-٠,٢٨٣	-٠,٢٦٩	٠,٠٤	٢	١٢١	٠,٩٦	٠,٠٩	٢	١٢١	٠,٩٢	٠,٤٤٩
الرابع	٠,٢٨٣	٠,٥٣٤	-٠,٢٦٤	-٠,٢٧٨	-٠,٢٦٤	٠,١٥	٢٠	٢٦١	٠,٨٦	٠,٠١	٢	٢٦١	٠,٨٩	٠,٤٤٤
الخامس	٠,٢٧٩	٠,٥٢١	-٠,٢٦٦	-٠,٢٦٥	-٠,٢٨٣	٠,٠٠١	٢	٦٩٨	٠,٩٩	٠,٠٢	٢	٦٩٨	٠,٩٨	٠,٤٥٢

ويتضح من الجدول (٢) السابق أن توزيع درجات المتغيرين التابع والمصاحب يبتعد عن التوزيع الاعتدالى بدرجات متفاوته وفى اتجاهين مختلفين (الاتجاه الموجب للمتغير التابع ، الاتجاه السالب للمتغير المصاحب) ، وتختلف القيم العددية لمعاملات الانحدار الثلاثة عبر مجموعات المعالجة الجزئية فى موقف القياس الأول والثانى والخامس ويتساوى زوج من معاملات الانحدار الثلاثة فى موقف القياس الثالث والرابع وهذا بمثابة مؤداه عدم تجانس معاملات الانحدار ، كما أن قيمة الإحصائى Levene Statistic (L) يؤشر لعدم تجانس التباين عبر مجموعات المعالجة الثلاثة فى المتغيرين التابع والمصاحب ، والقيمة العددية لمعامل ارتباط سبيرمان Spearman's rho تتجاوز القيمة (٠,٣٠) تؤشر لارتباط مقبول بين درجات القياس البعدى والقياس القبلى . ويستنتج من ذلك أن البيانات فى موقف القياس الأول حتى موقف القياس الخامس تغيب عنها الافتراضات الأساسية لاستخدام اختبار تحليل التغاير البارامترى ويتطابق ذلك مع نتائج دراسة Micceri (1989) الشهيرة .

وأخيراً استخدم الباحث اختبار تحليل التغاير البارامترى وسبع من البدائل اللابارامترية لاختبار الفرض : لا يختلف متوسط التحصيل الدراسى باختلاف استراتيجية التدريب ، باعتبار أن خصائص البيانات الواردة فى الجدول (٢) هى الشروط التى تتم فى ضوءها مقارنة أداء الاختبار البارامترى وبدائله اللابارامترية .

نتائج اختبار الفرض الأول وتفسيرها

ينص الفرض الأول على: لا يختلف أداء اختبار تحليل التغاير البارامترى والبدائل اللابارامترية السبعة تحت شروط مختلفة فى تصميم غير متوازن من ثلاث مجموعات فى حالة استخدام عينات ذات أحجام (صغيرة جداً، صغيرة، متوسطة، كبيرة، كبيرة جداً) ، ولاختبار هذا الفرض قام الباحث بحساب اختبار تحليل التغاير والبدائل اللابارامترية السبعة كما بالجداول التالية:

جدول (٣ - ١)

نتائج اختبار تحليل التغاير

أحجام العينات			مصدر التباين	مجموع المربعات	درجة الحرية	متوسط المربعات	قيمة ف
n_1	n_2	n_3					
٤	٣	٢	بين المجموعات	٨.٠٤٠	٢	٤.٠٢٠	٠.١٣٤
			داخل المجموعات	١٥١.٧٦١	٥	٣٠.٣٥٢	
			الكلى	١٥٩.٨٠١	٧		
١٠	١٢	١٢	بين المجموعات	٥٩.٢٥٧	٢	٢٩.٦٢٩	٢.٠٢٨
			داخل المجموعات	٤٣٨.٣٦٢	٣٠	١٤.٦١٢	
			الكلى	٤٩٧.٦١٩	٣٢		
٣٦	٤١	٤٧	بين المجموعات	١٣.٩٧٢	٢	٦.٩٨٦	٠.٤٨٣
			داخل المجموعات	١٧٣٥.١٩٧	١٢٠	١٤.٤٦٠	
			الكلى	١٧٤٩.١٦٩	١٢٢		
١٢١	١١٠	١٣٣	بين المجموعات	٨.٢٥٧	٢	٤.١٢٩	٠.٢٩٣
			داخل المجموعات	٥٠٨١.٥٨١	٣٦٠	١٤.١١٦	
			الكلى	٥٠٨٩.٨٣٨	٣٦٢		
٢٣١	٢٤٧	٢٢٣	بين المجموعات	٢.٥٩١	٢	١.٢٩٦	٠.٠٩١
			داخل المجموعات	٩٨٧٢.٣٠٧	٦٩٧	١٤.١٦٤	
			الكلى	٩٨٧٤.٨٩٨	٦٩٩		

$F_{critical}(2,5)=5.79, p \leq .05$

$F_{critical}(2,30)=3.32, p \leq .05$

$F_{critical}(2,120)=3.07, p \leq .05$

$F_{critical}(2,360)=3.07, p \leq .05$

$F_{critical}(2,697)=3.07, p \leq .05$

ويتضح من الجدول (٣-١) السابق أن النتائج تبين دليلاً غير كاف Insufficient Evidence لاستنتاج أن المعالجات الثلاث لها تأثيرات فارقة Differential Effects فى جميع مواقف القياس الخمسة ، أى أنه لا توجد فروق دالة فى

المتغير التابع (التحصيل) بين المعالجات الثلاث بعد ضبطه (تعديله) تبعاً للمتغير المصاحب . وأعلى قيمة لـــ F كانت فى موقف القياس الثانى وأقل قيمة كانت فى موقف القياس الخامس وهذا يناظر (أقل مدى ، أعلى مدى) على الترتيب بين معامل التواء المتغير التابع والمتغير المصاحب .

وفيما يلى نتائج التحليل باستخدام اختبار Quade :

جدول (٣ - ٢)

نتائج اختبار Quade

$(TS)_{QUADE}$	$R^2_{Y.12}$	P_2	P_1	N	أحجام العينات		
					n_3	n_2	n_1
٠.٠٣٣	٠.٠٠١١	٣	٢	٩	٢	٣	٤
٠.٣٨١	٠.٠٢٤	٣	٢	٣٤	١٢	١٢	١٠
٠.١٨٢	٠.٠٠٣	٣	٢	١٢٤	٤٧	٤١	٣٦
٠.٠٤١	٠.٠٠٠٢٢٥	٣	٢	٣٦٤	١٣٣	١١٠	١٢١
٠.٠٢٨	٠.٠٠٠٠٨١	٣	٢	٧٠١	٢٢٣	٢٤٧	٢٣١

$F_{critical}(2,5)=5.79, p \leq .05$

$F_{critical}(2,30)=3.32, p \leq .05$

$F_{critical}(2,120)=3.07, p \leq .05$

$F_{critical}(2,360)=3.07, p \leq .05$

$F_{critical}(2,697)=3.07, p \leq .05$

ويتضح من الجدول (٣-٢) السابق أن التحليل يبين دليلاً غير كاف لاستنتاج أن المعالجات الثلاث فاعلة بشكل فارق Differentially Effective ، أى أنه لا توجد فروق دالة فى المتغير التابع (التحصيل) بين المعالجات الثلاث بعد ضبطه (تعديله) تبعاً للمتغير المصاحب. ويشابه أداء هذا البديل اللابارامترى أداء اختبار تحليل التغاير البارامترى فى نفس موقف القياس حيث أن أعلى قيمة لـــ $(TS)_{QUADE}$ كانت فى موقف القياس الثانى وأقل قيمة كانت فى موقف القياس الخامس وهذا يناظر (أقل مدى ، أعلى مدى) على الترتيب بين معامل التواء المتغير التابع والمتغير المصاحب .

وفيما يلى نتائج التحليل باستخدام اختبار McSweeney and Porter :

جدول (٣ – ٣)

نتائج اختبار McSweeney and Porter

$(TS)_{MPF}$	$R^2_{Y.123}$	$R^2_{Y.1}$	P_2	P_1	N	أحجام العينات		
						n_3	n_2	n_1
٠.١٦٠	٠.٢٣٥	٠.١٨٦	٤	٢	٩	٢	٣	٤
٢.٠٢٠	٠.٤٠٦	٠.٣٢٦	٤	٢	٣٤	١٢	١٢	١٠
٠.٤٥٤	٠.٢٠٧	٠.٢٠١	٤	٢	١٢٤	٤٧	٤١	٣٦
٠.٤٤٩	٠.١٩٩	٠.١٩٧	٤	٢	٣٦٤	١٣٣	١١٠	١٢١
٠.٤٣٨	٠.٢٠٥	٠.٢٠٤	٤	٢	٧٠١	٢٢٣	٢٤٧	٢٣١

$F_{critical}(2,5)=5.79, p \leq .05$

$F_{critical}(2,30)=3.32, p \leq .05$

$F_{critical}(2,120)=3.07, p \leq .05$

$F_{critical}(2,360)=3.07, p \leq .05$

$F_{critical}(2,697)=3.07, p \leq .05$

ويتضح من الجدول (٣-٣) السابق وجود دليل غير كاف لرفض الفرض الصفرى بأن توزيعات درجات القياس البعدى تختلف عبر المجموعات الثلاث ، أى أنه لا توجد فروق دالة فى المتغير التابع بين المعالجات الثلاث بعد ضبطه (تعديله) تبعاً للمتغير المصاحب . وأعلى قيمة لـ $(TS)_{MPF}$ كانت فى موقف القياس الثانى وأقل قيمة كانت فى موقف القياس الأول .

وفيما يلى نتائج التحليل باستخدام اختبار Burnett and Barr :

جدول (٣ – ٤)

نتائج اختبار Burnett and Barr

F	$R^2_{Y.12}$	K	N	أحجام العينات		
				n_3	n_2	n_1
٠.١٤٨	٠.٠٤٧	٣	٩	٢	٣	٤
٠.٩٣٧	٠.٠٥٧	٣	٣٤	١٢	١٢	١٠
٠.٠٦١	٠.٠٠١	٣	١٢٤	٤٧	٤١	٣٦
٠.١٨١	٠.٠٠١	٣	٣٦٤	١٣٣	١١٠	١٢١
٠.٠٠٠٣	٠.٠٠٠٠٠١	٣	٧٠١	٢٢٣	٢٤٧	٢٣١

$F_{critical}(2,5)=5.79, p \leq .05$

$F_{critical}(2,30)=3.32, p \leq .05$

$F_{critical}(2,120)=3.07, p \leq .05$

$F_{critical}(2,360)=3.07, p \leq .05$

$F_{critical}(2,697)=3.07, p \leq .05$

ويتضح من الجدول (٣-٤) السابق أن النتائج بينت دليلاً غير كاف لاستنتاج أن التوزيعات غير المشروطة لدرجات القياس البعدى تختلف عبر المجموعات ، أى أنه لا توجد فروق دالة فى المتغير التابع (التحصيل) بين المعالجات الثلاث بعد ضبطه (تعديله) تبعاً للمتغير المصاحب . ويشابه أداء هذا البديل اللابارامترى أداء اختبار تحليل التغاير البارامترى وأداء البديل $(TS)_{QUADE}$ فى نفس موقف القياس حيث أن أعلى قيمة لـــ F كانت فى موقف القياس الثانى وأقل قيمة كانت فى موقف القياس الخامس وهذا يناظر (أقل مدى، أعلى مدى) على الترتيب بين معامل التواء المتغير التابع والمتغير المصاحب .

وفيما يلى نتائج التحليل باستخدام اختبار Shirley :

جدول (٣ - ٥)

نتائج اختبار Shirley

$(TS)_{Shirley}$	MS_{total}	$SS_{total\,(adjusted)}$	$R^2_{Y.123}$	$R^2_{Y.1}$	N	أحجام العينات		
						n_3	n_2	n_1
٠.٣٩٢	٧.٥٠٠	٢.٩٤٠	٠.٢٣٥	٠.١٨٦	٩	٢	٣	٤
٢.٦٣٩	٩٩.١٦٧	٢٦١.٦٦٨	٠.٤٠٦	٠.٣٢٦	٣٤	١٢	١٢	١٠
٠.٧٣٨	١٢٩١.٦٦٧	٩٥٣.٢٥٠	٠.٢٠٧	٠.٢٠١	١٢٤	٤٧	٤١	٣٦
٠.٧٢٦	١١٠٧١.٦٦٧	٨٠٣٨.٠٣٠	٠.١٩٩	٠.١٩٧	٣٦٤	١٣٣	١١٠	١٢١
٠.٧٠٠	٤١٠٠٨.٥٠٠	٢٨٧٠٥.٩٥	٠.٢٠٥	٠.٢٠٤	٧٠١	٢٢٣	٢٤٧	٢٣١

$\chi^2_{(2,.05)}=5.99$

ويتضح من الجدول (٣-٥) السابق أن نتائج التحليل توفر معلومات غير كافية لاستنتاج أن توزيع درجات القياس البعدى تختلف عبر المجموعات الثلاث ، أى أنه لا توجد فروق دالة فى المتغير التابع (التحصيل) بين المعالجات الثلاث بعد ضبطه (تعديله) تبعاً للمتغير المصاحب . وأعلى قيمة لـــ $(TS)_{Shirley}$ كانت فى موقف القياس الثانى وأقل قيمة كانت فى موقف القياس الأول ، ويتشابه أداء هذا البديل مع أداء البديل $(TS)_{MPF}$.

وفيما يلى نتائج التحليل باستخدام اختبار Puri-Sen-Harwell-Serlin :

جدول (٣ – ٦)

نتائج اختبار Puri-Sen-Harwell-Serlin

$(IS)_{FSHS}$	η^2	SS_{total}	$SS_{between(RY)}$	$(N-2)$	N	أحجام العينات		
						n_3	n_2	n_1
٠.٤٢	٠.٠٠٦	٤٦.٧٨٥	٢.٨٠٦	٧	٩	٢	٣	٤
٣.٨٠٨	٠.١١٩	٢١٨٩.٣٥٩	٢٦٠.٧١٦	٣٢	٣٤	١٢	١٢	١٠
٠.٨٥٤	٠.٠٠٠٧	١٢٥٧١٨.٠٥٩	٨٢٨.١٩٩	١٢٢	١٢٤	٤٧	٤١	٣٦
٠.٧٢٤	٠.٠٠٠٢	٣١٩٧١٩٢.٦٢٤	٥٠٥٠.١٩٢	٣٦٢	٣٦٤	١٣٣	١١٠	١٢١
١.٣٩٨	٠.٠٠٠٢	٢٢٦٣٥١٦١.٢٥٣	٥٤٦٩.٦٨٣	٦٩٩	٧٠١	٢٢٣	٢٤٧	٢٣١

$\chi^2_{(2,.05)=5.99}$

ويتضح من الجدول (٣–٦) السابق أن النتائج توفر دليلاً لرفض الفرض البديل (قبول الفرض الصفرى) ، أى أنه لا توجد فروق دالة فى المتغير التابع (التحصيل) بين المعالجات الثلاث بعد ضبطه (تعديله) تبعاً للمتغير المصاحب . وأعلى قيمة لـــ $(TS)_{Shirley}$ كانت فى موقف القياس الثانى وأقل قيمة كانت فى موقف القياس الأول ، ويتشابه أداء هذا البديل مع أداء البديل $(TS)_{MPF}$ ومع أداء البديل $(TS)_{PSHS}$.

وفيما يلى نتائج التحليل باستخدام اختبار Hettmansperger :

جدول (٣ – ٧)

نتائج اختبار Hettmansperger

$(TS)_{Hettmansperger}$	SS_{total}	SS_{error}	$SS_{between}$	N	أحجام العينات		
					n_3	n_2	n_1
٠.٠٧٧	٧.١٣٩	٧.٠٦٢	٠.٠٧٧	٩	٢	٣	٤
٠.٧٦٦	٣١.٩٠٣	٣١.١٣٧	٠.٧٦٦	٣٤	١٢	١٢	١٠
٠.٤١١	١٢١.٣٨٠	١٢٠.٩٦٩	٠.٤١١	١٢٤	٤٧	٤١	٣٦
٠.٠٧٦	٣٦٠.٠٦٦	٣٥٩.٩٩٠	٠.٠٧٦	٣٦٤	١٣٣	١١٠	١٢١
٠.٠٥٦	٦٩٥.٢٨٩	٦٩٥.٢٣٣	٠.٠٥٦	٧٠١	٢٢٣	٢٤٧	٢٣١

$\chi^2_{(2,.05)=5.99}$

ويتضح من الجدول (٣–٧) السابق أن التحليل يبين دليلاً غير كاف لرفض الفرض الصفرى وقبول الفرض البديل ، أى أنه لا توجد فروق دالة فى المتغير التابع (التحصيل) بين المعالجات الثلاث بعد ضبطه (تعديله) تبعاً للمتغير المصاحب . ويشابه أداء هذا البديل اللابارامترى أداء اختبار تحليل التغاير البارامترى وأداء البديل $(TS)_{QUADE}$وأداء اختبار Burnett and Barr فى نفس موقف القياس حيث أن أعلى قيمة لـــ $(TS)_{Hettmansperger}$

كانت فى موقف القياس الثانى وأقل قيمة كانت فى موقف القياس الخامس وهذا يناظر (أقل مدى ، أعلى مدى) على الترتيب بين معامل التواء المتغير التابع والمتغير المصاحب .

وفيما يلى نتائج التحليل باستخدام اختبار Puri-Sen :

جدول (٣-٨)

نتائج اختبار Puri-Sen

L_n	$V^{11}(V^{-1})$	$\sum_{k=1}^{3} n_k \bar{E}_k^2$	$\bar{E}_3^2$	$\bar{E}_2^2$	$\bar{E}_1^2$	N	أحجام العينات		
							n_3	n_2	n_1
٠.٣٣	٠.١٢	٢.٧٥٤	١.٠٦٩	٠.٠٦٨	٠.١٠٣	٩	٢	٣	٤
٢.٠٨	٠.٠٠٨	٢٦٠.٠٦	٩.٤٢	١٢.٠٦	٠.٢٣	٣٤	١٢	١٢	١٠
٠.٤٩٧	٠.٠٠٠٦	٨٢٨.٠٨	١٠.٨٦	٢.٩٨	٥.٤٣	١٢٤	٤٧	٤١	٣٦
٠.٥٠٦	٠.٠٠٠١	٥٠٥٨.٥٩	٨.٣٨	٣١.٧٣	٣.٧٥	٣٦٤	١٣٣	١١٠	١٢١
٠.٠١٦	٠.٠٠٠٠٠٣	٥٤٠٥.١٧	٧.١٠	١.٣٧	١٥.٠٥	٧٠١	٢٢٣	٢٤٧	٢٣١

$\chi^2_{(2,05)}=5.99$

يتضح من الجدول (٣-٨) السابق أن النتائج تبين دليل غير كاف لرفض الفرض الصفرى بأن توزيع درجات القياس البعدى تختلف عبر المجموعات الثلاث ، أى أنه لا توجد فروق دالة فى المتغير التابع (التحصيل) بين المعالجات الثلاث بعد ضبطه (تعديله) تبعاً للمتغير المصاحب . ويشابه أداء هذا البديل اللابارامترى أداء اختبار تحليل التغاير البارامترى وأداء البديل $(TS)_{QUADE}$ وأداء اختبار Burnett and Barr وأداء البديل $(TS)_{Heitmansperger}$ فى نفس موقف القياس حيث أن أعلى قيمة لـــ L_n كانت فى موقف القياس الثانى وأقل قيمة كانت فى موقف القياس الخامس وهذا يناظر (أقل مدى ، أعلى مدى) على الترتيب بين معامل التواء المتغير التابع والمتغير المصاحب .

وبمراجعة النتائج من الجدول (٣-١) إلى الجدول (٣-٨) ، يظهر اختبار تحليل التغاير البارامترى والبدائل اللابارامترية السبعة تشابه فى الأداء من حيث القرار الإحصائى ففى كل مرة حدث قبول للفرض الصفرى ورفض للفرض البديل ، وتتفق هذه النتيجة مع ما توصلت إليه نتائج دراسة Olejnik and Algina (1985) أن البدائل اللابارامترية تعطى نفس الاستنتاجات مثل الاختبار البارامترى فيما يتصل بتأثير المعالجات الثلاث عند مستوى (٠.٠٥) ، كما أن قيم احتمالات الخطأ من النوع الأول الخاصة بجميع البدائل اللابارامترية أقل من قيم احتمالات الخطأ من النوع الأول الخاصة بالاختبار البارامترى .

نتائج اختبار الفرض الثانى وتفسيرها

ينص الفرض الثانى على : لا تختلف الخصائص التوزيعية (تقديرات الخطأ من النوع الأول ، تقديرات الخطأ من النوع الثانى ، تقديرات الضلاعة ، تقديرات القوة الإحصائية) لاختبار تحليل التغاير البارامترى والبدائل اللابارامترية السبعة تحت شروط مختلفة فى تصميم غير متوازن من ثلاث مجموعات فى حالة استخدام عينات ذات أحجام (صغيرة جداً ، صغيرة ، متوسطة ، كبيرة ، كبيرة جداً) ، ولاختبار هذا الفرض قام الباحث بحساب تقديرات الخطأ من النوع الأول والثانى والضلاعة والقوة الإحصائية لاختبار تحليل التغاير والبدائل اللابارامترية السبعة فى مواقف القياس الخمسة كما بالجداول التالية* :

جدول (٤-١)

الخصائص التوزيعية

لاختبار تحليل التغاير والبدائل اللابارامترية السبعة فى موقف القياس الأول

الاختبار	α	α^*	مؤشر (تقدير) الضلاعة				β	تقدير القوة
			المتحرر $0.5\alpha \le \alpha^* \le 1.5\alpha$		الصارم $0.9\alpha \le \alpha^* \le 1.1\alpha$			
			الحد الأدنى	الحد الأعلى	الحد الأدنى	الحد الأعلى		
تحليل التغاير	٠.٠٥	٠.٨٧٩	٠.٠٢٥	٠.٠٧٥	٠.٠٤٥	٠.٠٥٥	٠.٩٣٨	٠.٠٦٢
Quade	٠.٠٥	٠.٨٩١	٠.٠٢٥	٠.٠٧٥	٠.٠٤٥	٠.٠٥٥	٠.٧٤٢	٠.٢٥٨
McSweeney and Porter	٠.٠٥	٠.٧٥٧	٠.٠٢٥	٠.٠٧٥	٠.٠٤٥	٠.٠٥٥	٠.٦٥٦	٠.٣٤٤
Burnett and Barr	٠.٠٥	٠.٨١٦	٠.٠٢٥	٠.٠٧٥	٠.٠٤٥	٠.٠٥٥	٠.٧٢١	٠.٢٧٩
Shirley	٠.٠٥	٠.٨٠٩	٠.٠٢٥	٠.٠٧٥	٠.٠٤٥	٠.٠٥٥	٠.٧١٥	٠.٢٨٥
Puri-Sen-Harwell-Serlin	٠.٠٥	٠.٧٤٩	٠.٠٢٥	٠.٠٧٥	٠.٠٤٥	٠.٠٥٥	٠.٦٤٧	٠.٣٥٣
Hettmansperger	٠.٠٥	٠.٧٢١	٠.٠٢٥	٠.٠٧٥	٠.٠٤٥	٠.٠٥٥	٠.٥٨٩	٠.٤١١
Puri-Sen	٠.٠٥	٠.٧٦٨	٠.٠٢٥	٠.٠٧٥	٠.٠٤٥	٠.٠٥٥	٠.٦٣٣	٠.٣٦٧

ملحوظة : α هو مستوى الدلالة الاسمى ، α^* هو مستوى الدلالة الفعلى (تقدير الخطأ من النوع الأول) ، β هو تقدير الخطأ من النوع الثانى .

ويتضح من الجدول السابق (٤-١) أن اختبار تحليل التغاير البارامترى والبدائل اللابارامترية السبعة أظهروا تضخماً فى تقدير الخطأ من النوع الأول Inflation of the Type I

* يتقدم الباحث بالشكر للأستاذ الدكتور James Algina لمساعدته الباحث فى تقدير معدلات الخطأ من النوع الأول والخطأ من النوع الثانى .

Error Rate ، وأن هذا التقدير لا يساوى قيمة ألفا الاسمية وهذا يتوافق جزئياً مع وجهة نظر Hus (1983); Algina (2014) بأن معدل الخطأ من النوع الأول لا يساوى ألفا فى حالة عدم استيفاء الافتراضات الأساسية لاختبار تحليل التغاير البارامترى . وتتفق هذه النتيجة جزئياً أيضاً مع ما توصلت إليه نتائج دراسة Rheinheimer and Penfield (2001) بأن عدم توافر شرط تجانس التباين وبخاصة فى حالة التوليفة (التصميمات غير المتوازنة وعدم الاعتدالية الشديد) له تأثير كبير على تقديرات الخطأ من النوع الأول .

ويتضح أيضاً أن تقدير مؤشر الضلاعة لاختبار تحليل التغاير البارامترى والبدائل اللابارامترية السبعة لا يحقق المحك الصارم أو حتى المحك المتحرر للضلاعة ، وبهذا فإن الاختبار البارامترى وبدائله اللابارامترية لا يتميزوا بالضلاعة فى موقف القياس الأول ، ويتفق ذلك جزئياً مع ما ورد ببعض الأدبيات مثل Hamilton (1976) بأن اختبار تحليل التغاير البارامترى لا يظهر مستوى من الضلاعة فى حالة استخدام تصميم غير متوازن .

كما يتضح من الجدول السابق (٤-١) أن البدائل اللابارامترية السبعة تتميز بخصائص قوة شبه مقبولة مقارنة باختبار تحليل التغاير البارامترى ، ويتفق ذلك مع ما ورد ببعض الأدبيات مثل : Micceri (1989), Sawilowsky and Fahoome (2000) أنه عند التخلى عن افتراض الاعتدالية فإن الاختبارات اللابارامترية يمكن أن تكون أكثر قوة عما هو عليه فى حالة اختبار t-test واختبار F-test البارامتريين ، وهذا مؤداه أن قدرة البدائل اللابارامترية على رفض الفرض الصفرى الخاطىء تزيد عن قدرة اختبار تحليل التغاير البارامترى فى حالة استخدام عينات ذات أحجام صغيرة جداً ، ويتضح أيضاً أن البدائل اللابارامترية لاختبار تحليل التغاير تتمايز فيما يتصل بالضلاعة والقوة وهذا يتفق جزئياً مع ما قدمه Nakonezny & Shull (2007) . وتتفق هذه النتيجة جزئياً أيضاً مع ما توصلت إليه نتائج دراسة Olejnik, Algina and Abdel-Fattah (1987) بأن الفرق فى تقديرات القوة بين اختبار تحليل التغاير والبدائل اللابارامترية من جهة والفروق فى تقديرات القوة بين البدائل اللابارامترية من جهة أخرى يمكن أن يكون موجب أو سالب ويمكن أن يكون تافه أو مهم اعتماداً على درجة الارتباط بين المتغير المصاحب ودرجات القياس البعدى ومدى الاعتدالية ودرجة عدم تجانس التباين . كما تتفق هذه النتيجة أيضاً مع ما توصلت إليه نتائج دراسة Harwell & Serlin (1988); Shull & Nakonezny (2002) بأن تقدير قوة اختبارى PSHS, McSweeney and Porter أكبر من تقدير قوة اختبار تحليل التغاير البارامترى فى حالة التوزيع غير الاعتدالى . وأيضاً تتفق هذه النتيجة مع ما توصلت إليه نتائج دراسة Alyasin (1991) بتميز أداء اختبار PSHS عن اختبار تحليل التغاير البارامترى من

حيث القوة وتقديرات الخطأ من النوع الأول .

وفيما يلى الخصائص التوزيعية لاختبار تحليل التغاير والبدائل اللابارامترية السبعة فى موقف القياس الثانى :

جدول (٤-٢)

الخصائص التوزيعية

لاختبار تحليل التغاير والبدائل اللابارامترية السبعة فى موقف القياس الثانى

الاختبار	α	α^*	مؤشر (تقدير) الضلاعة				β	تقدير القوة
			المتحرر $0.5\alpha \le \alpha^* \le 1.5\alpha$		الصارم $0.9\alpha \le \alpha^* \le 1.1\alpha$			
			الحد الأدنى	الحد الأعلى	الحد الأدنى	الحد الأعلى		
تحليل التغاير	٠.٠٥	٠.١٤٩	٠.٠٢٥	٠.٠٧٥	٠.٠٤٥	٠.٠٥٥	٠.٦١٥	٠.٣٨٥
Quade	٠.٠٥	٠.١٤٤	٠.٠٢٥	٠.٠٧٥	٠.٠٤٥	٠.٠٥٥	٠.٦٠٢	٠.٣٩٨
McSweeney and Porter	٠.٠٥	٠.١٣١	٠.٠٢٥	٠.٠٧٥	٠.٠٤٥	٠.٠٥٥	٠.٥١٩	٠.٤٨١
Burnett and Barr	٠.٠٥	٠.١٣٩	٠.٠٢٥	٠.٠٧٥	٠.٠٤٥	٠.٠٥٥	٠.٥٢٣	٠.٤٧٧
Shirley	٠.٠٥	٠.١٣٧	٠.٠٢٥	٠.٠٧٥	٠.٠٤٥	٠.٠٥٥	٠.٦١٤	٠.٣٨٦
Puri-Sen-Harwell-Serlin	٠.٠٥	٠.٠٩٩	٠.٠٢٥	٠.٠٧٥	٠.٠٤٥	٠.٠٥٥	٠.٥٧٣	٠.٤٢٧
Hettmansperger	٠.٠٥	٠.٠٨١	٠.٠٢٥	٠.٠٧٥	٠.٠٤٥	٠.٠٥٥	٠.٤٤٩	٠.٥٥١
Puri-Sen	٠.٠٥	٠.١١٥	٠.٠٢٥	٠.٠٧٥	٠.٠٤٥	٠.٠٥٥	٠.٥٨٨	٠.٤١٢

ملحوظة : α هو مستوى الدلالة الاسمى ، α^* هو مستوى الدلالة الفعلى (تقدير الخطأ من النوع الأول) ، β هو تقدير الخطأ من النوع الثانى .

ويتضح من الجدول السابق (٤-٢) أن اختبار تحليل التغاير البارامترى والبدائل اللابارامترية السبعة أظهروا انخفاضاً شديداً فى تقدير الخطأ من النوع الأول وأن هذا التقدير لا يبتعد كثيراً عن قيمة ألفا الاسمية ، وتميز اختبار Hettmansperger مقارنة بالبدائل الأخرى ، وتتفق هذه النتيجة مع ما توصلت إليه نتائج دراسة Headrick & Vineyard (2001) بأن اختبار Hettmansperger أظهر تقديرات مقبولة للخطأ من النوع الأول وأن الاختبار يعتبر بديلاً مقبولاً فى حالة العينات التى حجمها أكبر من ٢٠ .

ويتضح أيضاً أن تقدير مؤشر الضلاعة لاختبار تحليل التغاير البارامترى والبدائل اللابارامترية السبعة لا يحقق المحك الصارم أو حتى المحك المتحرر للضلاعة عدا اختبار

Hettmansperger الذى يقارب كثيراً نحو تحقيق الحد الأدنى لمحك الضلاعة المتحرر ، وبهذا فإن الاختبار البارامترى وستة من البدائل اللابارامترية لا يتميزوا بالضلاعة فى موقف القياس الثانى ، ويتفق ذلك جزئياً مع ما ورد ببعض الأدبيات مثل Hamilton (1976) بأن اختبار تحليل التغاير البارامترى لا يظهر مستوى من الضلاعة فى حالة استخدام تصميم غير متوازن .

كما يتضح من الجدول السابق (٤-٢) أن اختبار تحليل التغاير البارامترى والبدائل اللابارامترية السبعة تميزوا بخصائص قوة جيدة تناظر تقديرات متوسطة للخطأ من النوع الثانى فى حالة استخدام عينات ذات أحجام صغيرة ، وأن اختبار Hettmansperger أكثر تميزاً مقارنة بالبدائل اللابارامترية الأخرى وهذا يتفق مع نتائج دراسة Harwell and Serlin (1988) التى بينت تميز اختبار Hettmansperger عن البدائل اللابارامترية الأخرى فى حالة التوزيع غير الاعتدالي حتى فى حالة عدم تساوى أحجام العينات وعدم تساوى خطوط ميل الانحدار ، وتتفق هذه النتيجة بشكل جزئى مع نتائج دراسة Olejnik & Algina (1984) بأن البديل اللابارامترى يتميز بالقوة عندما يكون توزيع المتغير المصاحب غير اعتدالي وحجم العينة صغير والأخطاء غير متجانسة . وأيضاً تتفق هذه النتيجة جزئياً مع ما توصلت إليه نتائج دراسة Olejnik, Algina and Abdel-Fattah (1987) بأن للفرق فى تقديرات القوة بين اختبار تحليل التغاير والبدائل اللابارامترية من جهة والفروق فى تقديرات القوة بين البدائل اللابارامترية من جهة أخرى يمكن أن يكون موجب أو سالب ويمكن أن يكون تافه أو مهم اعتماداً على درجة الارتباط بين المتغير المصاحب ودرجات القياس البعدى ومدى الاعتدالية ودرجة عدم تجانس التباين . كما تتفق هذه النتيجة مع ما توصلت إليه نتائج دراسة Harwell & Serlin (1988); Shull & Nakonezny (2002) بأن تقدير قوة اختبارى PSHS, McSweeney and Porter أكبر من تقدير قوة اختبار تحليل التغاير البارامترى فى حالة التوزيع غير الاعتدالى . وتتفق هذه النتيجة أيضاً مع ما توصلت إليه نتائج دراسة Alyasin (1991) بتميز أداء اختبار PSHS عن اختبار تحليل التغاير البارامترى من حيث القوة وتقديرات الخطأ من النوع الأول .

وفيما يلى الخصائص التوزيعية لاختبار تحليل التغاير والبدائل اللابارامترية السبعة في موقف القياس الثالث :

جدول (٤-٣)

الخصائص التوزيعية

لاختبار تحليل التغاير والبدائل اللابارامترية السبعة فى موقف القياس الثالث

الاختبار	α	α^*	مؤشر (تقدير) الضلاعة				β	تقدير القوة
			المتحرر $0.5\alpha \le \alpha^* \le 1.5\alpha$		الصارم $0.9\alpha \le \alpha^* \le 1.1\alpha$			
			الحد الأدنى	الحد الأعلى	الحد الأدنى	الحد الأعلى		
تحليل التغاير	٠.٠٥	٠.٦١٨	٠.٠٢٥	٠.٠٧٥	٠.٠٤٥	٠.٠٥٥	٠.٨٧٢	٠.١٢٨
Quade	٠.٠٥	٠.٦٠٩	٠.٠٢٥	٠.٠٧٥	٠.٠٤٥	٠.٠٥٥	٠.٤٨١	٠.٥١٩
McSweeney and Porter	٠.٠٥	٠.٥٦٥	٠.٠٢٥	٠.٠٧٥	٠.٠٤٥	٠.٠٥٥	٠.٤٣١	٠.٥٦٩
Burnett and Barr	٠.٠٥	٠.٦١٤	٠.٠٢٥	٠.٠٧٥	٠.٠٤٥	٠.٠٥٥	٠.٤٩٦	٠.٥٠٤
Shirley	٠.٠٥	٠.٦١٣	٠.٠٢٥	٠.٠٧٥	٠.٠٤٥	٠.٠٥٥	٠.٥١٣	٠.٤٨٧
Puri-Sen-Harwell-Serlin	٠.٠٥	٠.٥٧٣	٠.٠٢٥	٠.٠٧٥	٠.٠٤٥	٠.٠٥٥	٠.٤١٤	٠.٥٨٦
Hettmansperger	٠.٠٥	٠.٥٦١	٠.٠٢٥	٠.٠٧٥	٠.٠٤٥	٠.٠٥٥	٠.٤٠٨	٠.٥٩٢
Puri-Sen	٠.٠٥	٠.٥٨٦	٠.٠٢٥	٠.٠٧٥	٠.٠٤٥	٠.٠٥٥	٠.٤١٣	٠.٥٨٧

ملحوظة : α هو مستوى الدلالة الاسمى ، α^* هو مستوى الدلالة الفعلى (تقدير الخطأ من النوع الأول) ، β هو تقدير الخطأ من النوع الثانى .

ويتضح من الجدول السابق (٤-٣) أن اختبار تحليل التغاير البارامترى والبدائل اللابارامترية السبعة أظهروا تضخماً فى تقدير الخطأ من النوع الأول ، وأن هذا التقدير لا يساوى قيمة ألفا الاسمية وهذا يتوافق جزئياً مع وجهة نظر Hsu (1983); Algina (2014) بأن معدل الخطأ من النوع الأول لا يساوى ألفا فى حالة عدم استيفاء الافتراضات الأساسية لاختبار تحليل التغاير البارامترى . كما تتفق هذه النتيجة جزئياً مع ما توصلت إليه نتائج دراسة Rheinheimer and Penfield (2001) بأن عدم توافر شرط تجانس التباين وبخاصة فى حالة التوليفة (التصميمات غير المتوازنة وعدم الاعتدالية الشديد) له تأثير كبير على تقديرات الخطأ من النوع الأول .

ويتضح أيضاً أن تقدير مؤشر الضلاعة لاختبار تحليل التغاير البارامترى والبدائل اللابارامترية السبعة لا يحقق المحك الصارم أو حتى المحك المتحرر للضلاعة ، وبهذا فإن

الاختبار البارامترى وبدائله اللابارامترية لا يتميزوا بالضلاعة فى موقف القياس الثالث ، ويتفق ذلك جزئياً مع ما ورد ببعض الأدبيات مثل Hamilton (1976) بأن اختبار تحليل التغاير البارامترى لا يظهر مستوى من الضلاعة فى حالة استخدام تصميم غير متوازن .

كما يتضح من الجدول السابق (٤-٣) أن البدائل اللابارامترية السبعة تميزت بخصائص قوة جيدة جداً مقارنة باختبار تحليل التغاير البارامترى نظراً لانخفاض تقديرات الخطأ من النوع الثانى، وهذا مؤداه أن قدرة البدائل اللابارامترية على رفض الفرض الصفرى الخاطىء أعلى بكثير من قدرة اختبار تحليل التغاير البارامترى في حالة استخدام عينات ذات أحجام متوسطة، وهذا يتفق جزئياً مع نتائج دراسة Olejink & Algina (1983) بأن البديل اللابارامترى يتميز بالقوة النسبية فى حالة استخدام أحجام عينات متوسطة وتوزيعات ملتوية ، وهذا يتفق أيضاً جزئياً مع نتائج دراسة Seaman, Olejnik & Algina (1985) بأن تقديرات ألفا الفعلية تقع خارج الفئات المحكية (للحدود) وغير مقبولة فى حالة استخدام الاختبار البارامترى وكل الإجراءات اللابارامترية ، ووجدت ميزة للقوة العملية للإجراءات الرتبية فى حالة أحجام العينات المتوسطة للتوزيعات غير الاعتدالية والتباينات الصغيرة والكبيرة . وتتفق هذه النتيجة جزئياً مع ما توصلت إليه نتائج دراسة Olejnik, Algina and Abdel-Fattah (1987) بأن للفرق فى تقديرات القوة بين اختبار تحليل التغاير والبدائل اللابارامترية من جهة والفروق فى تقديرات القوة بين البدائل اللابارامترية من جهة أخرى يمكن أن يكون موجب أو سالب ويمكن أن يكون تافه أو مهم اعتماداً على درجة الارتباط بين المتغير المصاحب ودرجات القياس البعدى ومدى الاعتدالية ودرجة عدم تجانس التباين. وأيضاً تتفق هذه النتيجة مع ما توصلت إليه نتائج دراسة Harwell & Serlin (1988); Shull & Nakonezny (2002) بأن تقدير قوة اختبارى PSHS, McSweeney and Porter أكبر من تقدير قوة اختبار تحليل التغاير البارامترى فى حالة التوزيع غير الإعتدالى. وتتفق هذه النتيجة أيضاً مع ما توصلت إليه نتائج دراسة Alyasin (1991) بتميز أداء اختبار PSHS عن اختبار تحليل التغاير البارامترى من حيث القوة وتقديرات الخطأ من النوع الأول .

وفيما يلى الخصائص التوزيعية لاختبار تحليل التغاير والبدائل اللابارامترية السبعة فى موقف القياس الرابع :

جدول (٤-٤)

الخصائص التوزيعية

لاختبار تحليل التغاير والبدائل اللابارامترية السبعة في موقف القياس الرابع

الاختبار	α	α^*	مؤشر (تقدير) الضلاعة				β	تقدير القوة
			المتحرر $0.5\alpha \le \alpha^* \le 1.5\alpha$		الصارم $0.9\alpha \le \alpha^* \le 1.1\alpha$			
			الحد الأدنى	الحد الأعلى	الحد الأدنى	الحد الأعلى		
تحليل التغاير	٠.٠٥	٠.٧٤٧	٠.٠٢٥	٠.٠٧٥	٠.٠٤٥	٠.٠٥٥	٠.٩٠٤	٠.٠٩٦
Quade	٠.٠٥	٠.٧٣٧	٠.٠٢٥	٠.٠٧٥	٠.٠٤٥	٠.٠٥٥	٠.٣٦٦	٠.٦٣٤
McSweeney and Porter	٠.٠٥	٠.٦٨٧	٠.٠٢٥	٠.٠٧٥	٠.٠٤٥	٠.٠٥٥	٠.٢٦٩	٠.٧٣١
Burnett and Barr	٠.٠٥	٠.٨٠٤	٠.٠٢٥	٠.٠٧٥	٠.٠٤٥	٠.٠٥٥	٠.٢٨٤	٠.٧١٦
Shirley	٠.٠٥	٠.٧٨١	٠.٠٢٥	٠.٠٧٥	٠.٠٤٥	٠.٠٥٥	٠.٣٢٣	٠.٦٧٧
Puri-Sen-Harwell-Serlin	٠.٠٥	٠.٦٢٣	٠.٠٢٥	٠.٠٧٥	٠.٠٤٥	٠.٠٥٥	٠.٣٢٩	٠.٦٧١
Hettmansperger	٠.٠٥	٠.٦١٩	٠.٠٢٥	٠.٠٧٥	٠.٠٤٥	٠.٠٥٥	٠.٢٨٧	٠.٧١٣
Puri-Sen	٠.٠٥	٠.٦٢٥	٠.٠٢٥	٠.٠٧٥	٠.٠٤٥	٠.٠٥٥	٠.٣٣٥	٠.٦٦٥

ملحوظة : α هو مستوى الدلالة الاسمى ، α^* هو مستوى الدلالة الفعلى (تقدير الخطأ من النوع الأول) ، β هو تقدير الخطأ من النوع الثانى .

ويتضح من الجدول السابق (٤-٤) أن اختبار تحليل التغاير البارامترى والبدائل اللابارامترية السبعة أظهروا تضخم فى تقدير الخطأ من النوع الأول ، وأن هذا التقدير لا يساوى قيمة ألفا الاسمية وهذا يتوافق جزئياً مع وجهة نظر Algina (2014) ;(1983) Hsu بأن معدل الخطأ من النوع الأول لا يساوى ألفا فى حالة عدم استيفاء الافتراضات الأساسية لاختبار تحليل التغاير البارامترى ، وتتفق هذه النتيجة مع ما توصلت إليه نتائج دراسة Headrick & Sawilowsky (2000) بوجود تضخم فى تقديرات الخطأ من النوع الأول لاختبار تحليل التغاير البارامترى عندما كان الإنحراف عن افتراض الاعتدالية كبير جداً . وتتفق هذه النتيجة جزئياً مع ما توصلت إليه نتائج دراسة Rheinheimer and Penfield (2001) بأن عدم توافر شرط تجانس التباين وبخاصة فى حالة التوليفة (التصميمات غير المتوازنة وعدم الاعتدالية الشديد) له تأثير كبير على تقديرات الخطأ من النوع الأول .

ويتضح أيضاً أن تقدير مؤشر الضلاعة لاختبار تحليل التغاير البارامترى والبدائل اللابارامترية السبعة لا يحقق المحك الصارم أو حتى المحك المتحرر للضلاعة ، وبهذا فإن

الاختبار البارامترى وبدائله اللابارامترية لا يتميزوا بالضلاعة فى موقف القياس الرابع ، ويتفق ذلك جزئياً مع ما ورد ببعض الأدبيات مثل Hamilton (1976) بأن اختبار تحليل التغاير البارامترى لا يظهر مستوى من الضلاعة فى حالة استخدام تصميم غير متوازن .

كما يتضح من الجدول السابق (٤-٤) أن البدائل اللابارامترية السبعة تميزت بخصائص قوة جيدة إلى حد ما مقارنة باختبار تحليل التغاير البارامترى الذى أظهر تدهور حاد فى تقدير القدرة ، وهذا مؤداه أن قدرة البدائل اللابارامترية على رفض الفرض الصفرى الخاطىء أعلى بشكل ملحوظ من قدرة اختبار تحليل التغاير البارامترى فى حالة استخدام عينات ذات أحجام كبيرة . وتتفق هذه النتيجة جزئياً مع ما توصلت إليه نتائج دراسة Olejnik, Algina and Abdel-Fattah (1987) بأن الفرق فى تقديرات القوة بين اختبار تحليل التغاير والبدائل اللابارامترية من جهة والفروق فى تقديرات القوة بين البدائل اللابارامترية من جهة أخرى يمكن أن يكون موجب أو سالب ويمكن أن يكون تافه أو مهم اعتماداً على درجة الارتباط بين المتغير المصاحب ودرجات القياس البعدى ومدى الاعتدالية ودرجة عدم تجانس التباين . كما تتفق هذه النتيجة مع ما توصلت إليه نتائج دراسة Harwell & Serlin (1988); Shull & Nakonezny (2002) بأن تقدير قوة اختبارى PSHS, McSweeney and Porter أكبر من تقدير قوة اختبار تحليل التغاير البارامترى فى حالة التوزيع غير الاعتدالى . وأيضاً تتفق هذه النتيجة مع ما توصلت إليه نتائج دراسة Alyasin (1991) بتميز أداء اختبار PSHS عن اختبار تحليل التغاير البارامترى من حيث القوة وتقديرات الخطأ من النوع الأول .

وفيما يلى الخصائص التوزيعية لاختبار تحليل التغاير والبدائل اللابارامترية السبعة فى موقف القياس الخامس :

جدول (٤-٥) الخصائص التوزيعية

لاختبار تحليل التغاير والبدائل اللابارامترية السبعة فى موقف القياس الخامس

الاختبار	α	α^*	مؤشر (تقدير) الضلاعة				β	تقدير القوة
			المتحرر $0.5\alpha \le \alpha^* \le 1.5\alpha$		المهارم $0.9\alpha \le \alpha^* \le 1.1\alpha$			
			الحد الأدنى	الحد الأعلى	الحد الأدنى	الحد الأعلى		
تحليل التغاير	٠.٠٥	٠.٠٩١٢	٠.٠٢٥	٠.٠٧٥	٠.٠٤٥	٠.٠٥٥	٠.٣٩٠	٠.٦٤٠
Quade	٠.٠٥	٠.٠٧١٩	٠.٠٢٥	٠.٠٧٥	٠.٠٤٥	٠.٠٥٥	٠.٩٣٨	٠.٩٦٢
McSweeney and Porter	٠.٠٥	٠.٠٦١٥	٠.٠٢٥	٠.٠٧٥	٠.٠٤٥	٠.٠٥٥	٠.٢٤٣	٠.٧٥٧
Burnett and Barr	٠.٠٥	٠.٠٧٢٢	٠.٠٢٥	٠.٠٧٥	٠.٠٤٥	٠.٠٥٥	٠.٢٥٣	٠.٧٤٧
Shirley	٠.٠٥	٠.٠٧٠٦	٠.٠٢٥	٠.٠٧٥	٠.٠٤٥	٠.٠٥٥	٠.٣١١	٠.٦٨٩
Puri-Sen-Harwell-Serlin	٠.٠٥	٠.٠٦٦٦	٠.٠٢٥	٠.٠٧٥	٠.٠٤٥	٠.٠٥٥	٠.٤١٨	٠.٦٨٢
Hettmansperger	٠.٠٥	٠.٠٦١٨	٠.٠٢٥	٠.٠٧٥	٠.٠٤٥	٠.٠٥٥	٠.٢٤٩	٠.٧٥١
Puri-Sen	٠.٠٥	٠.٠٦٤١	٠.٠٢٥	٠.٠٧٥	٠.٠٤٥	٠.٠٥٥	٠.٤١١	٠.٦٨٩

ملحوظة : α هو مستوى الدلالة الاسمى ، α^* هو مستوى الدلالة الفعلى (تقدير الخطأ من النوع الأول) ، β هو تقدير الخطأ من النوع الثانى .

ويتضح من الجدول السابق (٤-٥) أن اختبار تحليل التغاير البارامترى والبدائل اللابارامترية السبعة أظهروا تضخم فى تقدير الخطأ من النوع الأول ، وأن هذا التقدير لا يساوى قيمة ألفا الاسمية وهذا يتوافق جزئياً مع وجهة نظر Hsu (1983); Algina (2014) بأن معدل الخطأ من النوع الأول لا يساوى ألفا فى حالة عدم استيفاء الافتراضات الأساسية لاختبار تحليل التغاير البارامترى . وتتفق هذه النتيجة جزئياً مع ما توصلت إليه نتائج دراسة Rheinheimer and Penfield (2001) بأن عدم توافر شرط تجانس التباين وبخاصة فى حالة التوليفة (التصميمات غير المتوازنة وعدم الاعتدالية الشديد) له تأثير كبير على تقديرات الخطأ من النوع الأول .

ويتضح أيضاً أن تقدير مؤشر الضلاعة لاختبار تحليل التغاير البارامترى والبدائل اللابارامترية السبعة لا يحقق المحك الصارم أو حتى المحك المتحرر للضلاعة ، وبهذا فإن الاختبار البارامترى وبدائله اللابارامترية لا يتميزوا بالضلاعة فى موقف القياس الخامس ، ويتفق ذلك جزئياً مع ما ورد ببعض الأدبيات مثل Hamilton (1976) بأن اختبار تحليل التغاير البارامترى لا يظهر مستوى من الضلاعة فى حالة استخدام تصميم غير متوازن .

كما يتضح من الجدول السابق (٤-٥) أن اختبار تحليل التغاير البارامترى والبدائل اللابارامترية السبعة تميزوا بخصائص قوة متميزة نظراً لانخفاض تقديرات الخطأ من النوع الثانى ، وهذا مؤداه أن قدرة اختبار تحليل التغاير والبدائل اللابارامترية السبعة على رفض الفرض الصفرى الخاطىء عالية فى حالة استخدام عينات ذات أحجام كبيرة جداً . وتتفق هذه النتيجة جزئياً مع ما توصلت إليه نتائج دراسة Olejnik, Algina and Abdel-Fattah (1987) بأن الفرق فى تقديرات القوة بين اختبار تحليل التغاير والبدائل اللابارامترية من جهة والفروق فى تقديرات القوة بين البدائل اللابارامترية من جهة أخرى يمكن أن يكون موجب أو سالب ويمكن أن يكون تافه أو مهم اعتماداً على درجة الارتباط بين المتغير المصاحب ودرجات القياس البعدى ومدى الاعتدالية ودرجة عدم تجانس التباين . وتتفق هذه النتيجة مع ما توصلت إليه نتائج دراسة Harwell & Serlin (1988); Shull & Nakonezny (2002) أن تقدير قوة اختبارى PSHS, McSweeney and Porter أكبر من تقدير قوة اختبار تحليل التغاير البارامترى فى حالة التوزيع غير الاعتدالى . وتتفق هذه النتيجة مع ما توصلت إليه نتائج دراسة Alyasin (1991) بتميز أداء اختبار PSHS عن اختبار تحليل التغاير البارامترى من حيث القوة وتقديرات الخطأ من النوع الأول .

مستخلص نتائج الدراسة

- بزيادة حجم العينة تطرد قوة اختبار تحليل التغاير البارامترى Parameric ANCOVA عبر مواقف القياس الخمسة .
- تبقى تقديرات الخطأ من النوع الأول للبدائل اللابارامترية السبعة أقل من نظيراتها الخاصة باختبار تحليل التغاير البارامترى فى كافة مواقف القياس الخمسة ، وتقديرات الخطأ من النوع الأول الفعلى تعد أكبر من تقديرات للخطأ من النوع الأول الاسمى (تقديرات مستوى للدلالة الإحصائية الاسمى) .
- ظهر اختبار Hettmansperger بأنه اختبار متحرر Liberal Test بلغة الضلاعة فى موقف القياس الثانى .
- تحسن تقدير القوة الإحصائية لاختبار تحليل التغاير البارامترى فى موقف القياس الثانى بشكل ملحوظ واقترن ذلك مع أقل انحراف عن الاعتدالية لتوزيع درجات المتغيرين التابع والمصاحب مقارنة ببقية مواقف القياس (القيمة المطلقة للفرق = ٠.٥٤٠) .
- تدهور تقدير القوة الإحصائية لاختبار تحليل التغاير البارامترى فى موقف القياس الرابع بشكل حاد واقترن ذلك مع أكبر انحراف عن الاعتدالية لتوزيع درجات المتغيرين التابع والمصاحب مقارنة ببقية مواقف القياس (القيمة المطلقة للفرق = ٠.٨١٧) .

توصيات الدراسة:

فى نهاية هذه الدراسة يعرض الباحث الحالى لبعض الممارسات الواجب القيام بها والأخرى الواجب تجنبها فى عملية إجراء البحوث التربوية والنفسية كما يلى :

١. الحذر واليقظة : على الباحثين تجنب إجراء التحليل الإحصائى بصورة أوتوماتيكية . والسبب وراء ذلك هو التغيرات التى تحصل وبتسارع مستمر فى مجال الأبحاث الإحصائية ، وعلى الباحث أن يبقى على اطلاع دائم ومستمر بالتحسينات والإضافات فى مجال الإحصاء بشكل عام .
٢. على الباحث الاهتمام بالقوة الإحصائية للاختبارات المستخدمة إلى حد القيام بالتخطيط لدراسته مسبقاً من حيث حجم العينة والاختيار والتعيين العشوائيين .
٣. استخدام تصميم متوازن يكفل وجود خلايا لها نفس حجم العينة والابتعاد عن استخدام اختبار تحليل التغاير مع تصميمات غير متوازنة .
٤. استخدام البدائل اللابارامترية لاختبار تحليل التغاير البارامترى فى حالة عدم إستيفاء البيانات بالافتراضات التى يستند إليها الاختبار .

٥. الإفادة من تراكم نتائج البحوث والدراسات السابقة لتحديد المتغيرات المصاحبة لأجل تعديل درجات المتغير التابع وليس إعادة دراسة ارتباطها به للوصول إلى نتيجة تتفق مع الدراسات والبحوث السابقة .

٦. تقديم الضبط التجريبى على الضبط الإحصائى ، وفى حالة الاقتضاء يمكن اللجوء إلى اختبار تحليل التغاير أو أحد بدائله اللابارامترية دون الحاجة إلى عمليات تكافؤ العينات قبل إجراء البحث .

٧. عند اتخاذ القرار الإحصائى ، يحسن الأخذ فى الاعتبار الأهمية النسبية للمخاطر (الأخطاء) المتضمنة ، إذا كانت هناك طرق متوافرة لتقويم المخاطر .

٨. على الباحث تجنب أخطاء منطقية (أخطاء من الدرجة الأولى والثانية) تنطوى عليها تناقضات فى تحليل البيانات وتفسيرها .

بحوث مقترحة:

١. تمديد الدراسة الحالية إلى حالة أكثر من متغير مصاحب .

٢. تمديد الدراسة الحالية من تحليل التغاير الأحادى إلى تحليل التغاير العاملى .

٣. مقارنة أداء البدائل الأخرى التى لم تتضمنها الدراسة الحالية مثل اختبار Koch, Tangen, Jung, and Amara (1998) ، وإجراء Rogosa Procedure ، وإحصاء NP الذى اقترحه Akritas et al. (2000) ، واختبار YB الذى قدمه Young and Bowman (1995) ، واختبار The Drop Test الذى قدمه McKean and Schrader (1980)

المراجع:

١- السيد محمد خيرى (١٩٩٧) . الإحصاء النفسى . القاهرة : دار الفكر العربى .

٢- أحمد سليمان عودة ، خليل يوسف الخليلى (٢٠٠٠) . الإحصاء للباحث فى التربية والعلوم الإنسانية . ط٢. عمان : دار الأمل .

٣- ج. ميلتون سميث (١٩٧٨) . الدليل إلى الإحصاء فى التربية وعلم النفس . (ترجمة : إبراهيم بسيونى عميرة) . القاهرة : دار المعارف .

٤- زكريا الشربينى (١٩٩٠) . الإحصاء اللابارامترى فى العلوم التربوية والنفسية والاجتماعية . القاهرة : مكتبة الأنجلو المصرية .

٥- سالم عويس (٢٠٠٣) . إشكاليات البحث التربوى وآلياته الإحصائية : مراجعة نقدية لطرق البحث وآلياته الإحصائية . مجلة التربية . العدد (١٤٦) ، ١٤٢ - ١٥٤ .

٦- صلاح الدين محمود علام (١٩٩٣) . الأساليب الإحصائية الإستدلالية البارامترية واللابارامترية فى تحليل بيانات البحوث النفسية والتربوية . القاهرة : دار الفكر العربى .

٧- صلاح جلال، عصام الطويل، عبد الحليم عشماوى (١٩٨٨). الإحصاء الحيوى ومقدمة فى تصميم التجارب (الجزء الأول). القاهرة: مركز التنمية البشرية والمعلومات .

٨- عبد الرحمن عدس (١٩٩٧). مبادىء الإحصاء فى التربية وعلم النفس (الجزء الثانى : الإحصاء التحليلى). ط٢. عمان: دار الفكر للطباعة والنشر والتوزيع .

٩- عزت عبد الحميد محمد حسن (٢٠١١). الإحصاء النفسى والتربوى تطبيقات باستخدام برنامج SPSS 18 . القاهرة : دار الفكر العربى .

١٠- عزو إسماعيل عفانة (١٩٩٦). قوة كفاءة بعض الاختبارات الإحصائية اللابارامترية مقابل الاختبار الإحصائى البارامترى ستيودنت . مجلة التقويم والقياس النفسى والتربوى . العدد (٧) ، ٣٣-٦٧ .

١١- على ماهر خطاب (٢٠١١). الإحصاء الاستدلالى فى العلوم النفسية والتربوية والاجتماعية. القاهرة. مكتبة الأنجلو المصرية .

١٢- فؤاد أبو حطب ، آمال صادق (١٩٩٦) . مناهج البحث وطرق التحليل الإحصائى فى العلوم النفسية والتربوية والاجتماعية (ط٢). القاهرة: مكتبة الأنجلو المصرية

١٣- ليونا أ. تايلر (١٩٨٣) . الاختبارات والمقاييس (ترجمة : سعد عبد الرحمن ، مراجعة : محمد عثمان نجاتى) . القاهرة : دار الشروق

١٤- محسوب عبد القادر الضوى (٢٠٠٤) . قوة وحجم تأثير بعض البدائل اللابارامترية لاختبار تحليل التباين العاملى، رسالة دكتوراه غير منشورة. كلية التربية بقنا، جامعة جنوب الوادى .

١٥- مسفر سليم الجعيد (٢٠٠٤). تحليل التباين المصاحب واستخدامه فى ضبط المتغيرات الإحصائية فى البحوث التربوية. رسالة ماجستير غير منشورة . كلية التربية. جامعة أم القرى.

١٦- محمد أبو يوسف (١٩٨٩) . الإحصاء فى البحوث العلمية . القاهرة : المكتبة الأكاديمية .

17- Akritas, M. G., Arnold, S. F., & Du, Y. (2000). Nonparametric Models and Methods for Nonlinear Analysis of Covariance. Biometrika. 87, 507-526.

18- Algina, J. (2014, January 19). Re: Help [Electronic Mailing List Message]. Retrieved from https://us-mg5.mail.yahoo.com/neo/launch?.rand=d8vb1d9cagmhe.

19- Algina, J., & Olejnik, S. (2003). Conducting Power Analysis for ANOVA and ANCOVA in Between-Subjects Designs. Evaluation & The Health Professions, 26(3), 288-314.

20- Alyasin, M. M. (1991). An Empirical Investigation of the Behavior of Some Parametric and Non-parametric Tests for Frequently Encountered Data in Educational Research. Dissertation Abstracts International, 5A(52), 721.

21- American Psychological Association (2010). Publication Manual of the American Psychological Association (6th ed.). Washington, DC: Author.

22- Barrett, T. J. (2011). Computations Using Analysis of Covariance. WIREs Computational Statistics. 3, 260-268.

23- Bradley, J. V. (1968). Distribution-Free Statistical Tests. Englewood Cliffs, New Jersey: Prentice-Hall, Inc.

24- Bradley, J. V. (1978). Robustness?. British Journal of Mathematical and Statistical Psychology. 31, 144-152.

25- Burnett, T. D., & Barr, D. R. (1977). A Nonparametric Anology of Analysis of Covariance. Educational and Psychological Measurement. 37, 341-348.

26- Conover, W. J., & Iman, R, L. (1981). Rank Transformations as A Bridge between Parametric and Nonparametric Statistics. The American Statistician. 35(3), 124-129.

27- Conover, W. J., & Iman, R, L. (1982). Analysis of Covariance Using the Rank Transformation. Biometrics. 38(3), 715-724.

28- Dyer, C. (1995). Beginning Research in Psychology: A Practical Guide to Research Methods and Statistics. Oxford: Blackwell Publishers Ltd.

29- Erceg-Hurn, D. M., & Mirosevich, V. M. (2008). Modern Robust Statistical Methods: An Easy Way to Maximize the Accuracy and Power of Your Research. American Psychologist. 63(7), 591-601.

30- Finch, S., Cumming, G., & Thompson, N. (2002). Past and Future APA Guidelines for Statistical Practice. Theory & Psychology. 23, 1-33.

34- Hamilton, B. L. (1976). A Monte Carlo Test of the Robustness of Parametric and Nonparametric Analysis of Covariance against Unequal Regression Slopes. Journal of the American Statistical Association. 71(356), 864-869.

35- Harwell, M. R. (1988). Choosing Between Parametric and Non-Parametric Tests. Journal of Counseling and Development, 67, 35-38.

36- Harwell, M. R. (1990). A General Approach to Hypothesis Testing for Nonparametric Tests. Journal of Experimental Education. 58(2), 143-156.

37- Harwell, M. R. (1991). Completely Randomized Factorial Analysis of Variance Using Ranks. British Journal of Mathematical and Statistical Psychology. 44, 383-401.

38- Harwell, M. R., & Serlin, R. C. (1988). An Empirical Study of a Proposed Test of Non-parametric Analysis of Covariance. Psychological Bulletin. 104 (2), 268-281.

39- Harwell, M. R., & Serlin, R. C. (1989). A Non-parametric Test Statistic for the General Linear Model. Journal of Educational Statistics. 14(4), 351-371.

40- Harwell, M. R., & Serlin, R. C. (2001). Review of Non-parametric Tests for Complex Experimental Designs in Educational Research. Paper Presented at the Annual Meeting of the American Educational Research Association (Seattle, WA, April 24-28).

41- Headrick, T. C. (1997). Type I Error and Power of the Rank Transform Analysis of Covariance (ANCOVA) in a 3 x 4 Factorial Layout. Unpublished Doctoral Dissertation, Wayne State University, Detroit, MI.

42- Headrick, T. C., & Sawilowsky, S. S. (2000). Properties of the Rank Transformation in Factorial Analysis of Covariance. Communications in Statistics: Simulation and Computation. 29(4), 1059-1087.

43- Headrick, T. C., & Vineyard, G. (2001). An Investigation of Four Tests for Interaction in the Context of Factorial Analysis of Covariance. Multiple Linear Regression

Viewpoints, 27(2), 3-15.

44- Hettmansperger, T. P. (1984). Statistical Inference Based on Ranks. New York: Wiley.

45- Hsu, T. (1983). The Robustness of ANCOA to the Violation of Various Assumptions: A Review of Recent Studies. Paper Prsented at the Annual Meeting of the American Educational Research Association (Monteral, Canada).

46- Huber, P. J. (1981). Robust Statistics. New York: John Wiley & Sons, Inc.

47- Huberty, C. J. (1994). Applied Discriminant Analysis. New York: John Wiley & Sons.

48- Keppel, G. (1982). Design & Analysis: A Researcher's Handbook (2nd ed.). Englewood Cliffs, New Jersey: Prentice-Hall, Inc.

49- Kerlinger, F. N., & Pedhazur, E. G. (1973). Multiple Regression in Behavioral Research. New York: Holt, Rinehart and Winston, Inc.

50- Keselman, H. J., Huberty, C. J., Lix, L. M., Olejnik, S., Cribbie, R. A., Donahue, B., Kowalchuk, R., Lowman, L. L., Petoskey, M. D., Keselman, J. C., & Levin, J. R. (1998). Statistical Practices of Educational Researchers: An Analysis of Their ANOVA, MANOVA, and ANCOVA Analysis. Review of Educational Research, 68(3), 350-386.

50- Kim, S. (2007). Type I Error Rate and Statistical Power of ANCOVA, Potthoff's Modified Johnson-Neyman Method, and Wilcox Trimmed-Mean Method in Randomized Two-Group Experiment Design. Unpublished Doctoral Dissertation, University of Georgia, Georgia.

51- Kim, S. (2010). Alternatives to Analysis of Covariance for Heterogeneous Regression Slopes in Educational Research. Korean Journal of Teacher Education. 26(1), 73-91.

52- Kisbu-Sakarya, Y., MacKinnon, D. P., & Aiken, L. S. (2013). A Mote Carlo Comparison Study of the Power of the Analysis of Covariance, Simple Difference, and Residual Change Scores in Testing Two-Wave Data. Educational and Psychological Measurement. 73(1), 47-62.

52- Koch, G. G., Tangen, C. M., Jung, J.-W., & Amara, I. A. (1998). Issues for Covariance Analysis of Dichotomous and Ordered Categorical Data from Randomized Clinical Trials and Non-Parametric Strategies for Addressing Them. Statistics in Medicine. 17, 1863-1892.

53- Koti, K. M. (1995). A Note on Rank-Set Sampling. Technical Report (95-01). Department of Statistics. The Pennsylvania State University.

54- MacDonald, P. (1999). Power, Type I Error, and Type II Error Rates of Parametric and Non-parametric Statistical Tests. The Journal of Experimental Education. 67(4), 367-379.

55- McKean Scharder (1980).

56- McSweeney, M., & Porter, A. C. (1971). Small Sample Properties of Nonparametric Index of Response and Rank Analysis of Covariance. Paper Presented at the Annual Meeting of the American Educational Research Association (New York, January).

57- Micceri, T. (1989). The Unicorn, the Normal Curve, and other Imrobable Creatures. Psychological Bulletin. 105(1), 156-166.

58- Nakonezny, P. A., & Shull, R. D. (2007). JAMSM26: Hettmansperger and McKean Linear Model Aligned Rank Test for the Single Covariate and One-Way ANCOVA Case (SAS). Journal of Modern Applied Statistical Methods. 6(1), 336-340.

59- Olejnik, S. F., & Algina, J. (1983). Parametric ANCOVA vs. Rank Transform ANCOVA When Assumptions of Conditional Normality and Homoscedasticity Are Violated. ERIC Number: ED231882.

60- Olejnik, S. F., & Algina, J. (1984). Parametric ANCOVA and the Rank Transform ANCOVA When the Data are Conditionally Non-normal and Heteroscedastic. Journal of Educational Statistics. 9(2), 129-149.

61- Olejnik, S. F., & Algina, J. (1985). A Review of Non-parametric Alternatives to Analysis of Covariance. Evaluation Review, 9(1), 51-83.

62- Olejnik, S. F., Algina, J., & Abdel-Fattah, A. (1987). An Analysis of Statistical Power for Parametric ANCOVA and Rank Transform ANCOVA. Communications in Statistics: Theory and Methods, 16(7), 1923- 1949.

63- Park, H. M. (2010). Hypothesis Testing and Statistical Power of a Test. Working Paper. The University Information Technology Services (UITS) Center for Statistical and Mathematical Computing. Indiana University. http://www.indiana.edu/.

64- Puri, M. L., & Sen, P. K. (1969a). Analysis of Covariance Based on General Rank Scores. The Annals of Mathematical Statistics, 40 (2), 610-618.

65- Puri, M. L., & Sen, P. K. (1969b). A Class of Rank Order Test for General Linear Hypothesis. The Annals of Mathematical Statistics. 40, 1325-1343.

66- Quade, D. (1967). Rank Analysis of Covariance. Journal of the American Statistical Association. 62(320), 1187-1200.

67- Rheinheimer, D. C., & Penfield, D. A. (2001). The Effects of Type I Error Rate and Power of the ANCOVA "F" Test and Selected Alternatives under Non-normality and Variance Heterogeneity. Journal of Experimental Education, 69(4), 373-391.

68- Rogosa, D. (1980). Comparing Nonparallel Regression Lines. Psychological Bulletin. 88 (2), 307-321.

69- Sawilowsky, S. S., & Fahoome, G. (2000). Review of Twenty Non-parametric Statistics and Their Large Sample Approximations. Paper Presented at the Annual Meeting of the American Educational Research Association (New Orleans, LA, April 24-28).

70- Seaman, S. (1984). Type I Error Probabilities and Power of the Rank and Parametric ANCOVA Procedures. Paper Presented at the Annual Meeting of the American Educational Research Association (68th, New Orleans, LA, April 23-27).

71- Seaman, S., Algina, J., & Olejnik, S. F. (1985). Type I Error probabilities and Power of the Rank and Parametric ANCOVA Procedures. Journal of Educational Statistics, 10(4), 345-367.

72- Shirley, E. A. C. (1981). A Distribution-Free Method for Analysis of Covariance Based on Ranked Data. Journal of the Royal Statistical Society, 30 (2), 158-62.

73- Shirley, E. A. C., & Newnham, P. (1984). The Choice between Analysis of Variance and Analysis of Covariance with Special Reference to the Analysis of Organ

Weights in Toxicology Studies. Statistics in Medicine. 3(1), 85-91.

74- Shull, R. D., & Nakonezny, P. A. (2002). Power Curves and SAS Code to Aid in Selecting and Calculating Non-parametric ANCOVA Procedures. Statistics in an Area of Technological Change. 12(1), 13-25.

75- Wang, L., & Akritas, M. G. (2006). Testing for Covariate Effects in the Fully Nonparametric Analysis of Covariance Model. Journal of the American Statistical Association. 111(474), 722-736.

76- Young, S. G., & Bowman, A. W. (1995). Non-parametric Analysis of Covariance. Biometrics. 51, 920-931.

77- Zimmerman, D. W. (1996). A Note on Homogeneity of Variance of Scores and Ranks. The Journal of Experimental Education. 64, 351-362.

78- Zink, R. C., & Koch, G. G. (2012). NParCov3: A SAS/IML Macro for Nonparametric Randomization-Based Analysis of Covariance. Journal of Statistical Software. 50(3): Published Online.

Investigating Distributional Properties of Some Nonparametric Alternatives to the Parametric Analysis of Covariance Test

By

Mahsoub Abdel Kader Al Dowy Hassan

Associated Professor of Educational Psychology

Educational Psychology Department

Qena Faculty of Education

South Valley University

The current study aimed at investigating the distributional properties (type I, II error rates, robustness, and statistical power) of seven nonparametric alternatives to the parametric ANCOVA test those suggested by (1) Quade (1967); (2) Shirley (1981); (3) Burnett and Barr (1977); (4) McSweeney and Porter (1971); (5) Puri-Sen-Harwell-Serlin (1985); (6) Hettmansperger (1984); and (7) Puri-Sen (1969a).

The researcher used an empirical data collected by measuring achievement in a training program (stress management) within the Pathways for Higher Education project. The sample size was delimited to (701) students who participated in winter and summer sessions. The researcher derived many sample sizes according to the requirements of the data analysis to study the effects of using three training strategies (cooperative learning strategy, KWL strategy, brain storming strategy) on achievement by using unbalanced design including varity of sample sizes (very small, small, medium, large, and very large).

The study results indicated that: the parametric ANCOVA and the seven nonparametric alternatives showed similar performance leading to reject the null hypothesis. The parametric ANCOVA and the seven nonparametric alternatives showed inflation in the Type I error rates under all the conditions as the actual Type I error rate was considerably greater than the nominal Type I error rate except the second condtion. The parametric ANCOVA and the seven nonparametric alternatives appeared not to be robust for unequal group sizes except Hettmansperger test which appeared to be near the maxmum bound of liberal criteria for robustness under the second condtion. Also, the results indicated that the statistical power of the seven nonparametric alternatives increased as the sample size increased, is usually more powerful than the parametric ANCOVA under all the conditions, and the parametric ANCOVA test showed only (acceptable, valuable) statistical power respectively under the second and fifth conditions.

The obtained results showed the importance and usefulness of the nonparametric ANCOVA alternatives in the research field, particularly for the analysis of experiments, where the errors are not normally distributed.