جامعة سوهاج

كلية التربية

المجلة التربوية

مقارنة معاملات ثبات درجات الاختبار في ظل مجموعة من الاشتراطات: دراسة محاكاة مونت كارلو

إعداد

د/ نسرين محمد سعيد زارع

أستاذ علم النفس التربوي المساعد

قسم علم النفس – كلية التربية

جامعة القصيم – المملكة العربية السعودية

تاريخ الاستلام : ٢١ أبريل ٢٠٢١م – تاريخ القبول : ١١ مايو ٢٠٢١م

DOI: 10.12816/EDUSOHAG.2021.

المستخلص

تمت المقارنة في هذه الدراسة بين قيم ونسب تحيز ثلاث عشرة نوع مختلف من معاملات الثبات وهي معاملات ثبات الحدود الدنيا لجتمان(λ_1، λ_2، λ_3، λ_4، λ_5، λ_6) ومعامل ثبات ألفا كرونباخ (α) ، ومعامل ثبات أوميجا الكلية (ω_t)، ومعامل ثبات أوميجا الهرمية التقاربي (ω_h)، ومعامل ثبات أكبر حد أدنى (glb)، ومعامل ثبات ألفا الرتبية للأقسام المتعددة (α_{poly})، ومعامل ثبات أسوء تجزئة نصفية بيتا (β) ، ومعامل ثبات ألفا الطبقية (α_{strata})، ومعامل الثبات الأقصى (Maximal Reliability) وذلك من خلال بيانات مولدة بطريقة مونت كارلو بإطارين للقياس من نماذج أحادية البعد وأخرى متعددة الأبعاد عبر أربعة شروط للبيانات (نوع بيانات القياس — خيارات الاستجابة- طول الاختبار – حجم العينة) ، وأتضح من النتائج أن هناك أربعة أنواع من معاملات الثبات تعطي أعلى معامل تقدير للثبات ويمكن اعتبارها معاملات ثبات غير متحيزة حيث تفوقت على معامل ألفا كرونباخ التقليدي مما يجعلها الأفضل للاستخدام في حالات البيانات المختلفة، وهي معامل ثبات أوميجا الكلية (ω_t) ومعامل ثبات أكبر حد أدنى (glb) ومعامل ثبات الحد الأدنى لجتمان (λ_4) الأقصى، ومعامل ثبات ألفا للأقسام المتعددة (α_{poly}) وكان أفضلهم أداء معامل ثبات أوميجا الكلية (ω_t) حيث أعطى أعلى قيم للثبات بأقل تحيز نسبي لغالبية التداخلات بين حالات وأنواع البيانات والاختبارات وتنوع احجام العينات.

الكلمات المفتاحية: معامل ثبات ألفا – الحدود الدنيا لجتمان – معامل ثبات أوميجا – معامل ثبات بيتا – معامل الثبات الأقصى– معامل ثبات ألفا الطبقية – معامل ثبات ألفا الرتبية للأقسام المتعددة–– دراسة محاكاه

Print:(ISSN 1687-2649) Online:(ISSN 2536-9091)

Comparison of Test Scores Reliability Coefficients under a Set of Variables: A Monte Carlo Simulation Study

Nisreen Mohamed Said Zarea
Educational Psychology Assistant Professor
Psychology Department – Faculty of Education
Qassim University - KSA

Abstract

In this study, a comparison was made between the values and relative bias ratios of thirteen different types of reliability coefficients, which are the Guttman reliability lower bounds coefficients (λ_1, λ_2, λ_3, λ_4, λ_5, λ_6), and the Cronbach alpha reliability coefficient (α), and the Total Omega coefficient (ω_t), Hierarchical Omega-asymptotic coefficient (ω_h), Greatest lower bound coefficient (glb), Polychoric alpha coefficient (α_{poly}), Worst split half coefficient (β), Alpha-stratified coefficient (α_{strata}), and Maximal Reliability Coefficient (H) through data generated by Monte Carlo method with two measurement frames unidimensional and multidimensional models through four data conditions (Data Type - Response Options - Test Length - Sample Size), and it became clear from the results that there are four types of reliability coefficients that give the highest estimation coefficient of measurement reliability and can be considered non-bias reliability coefficients as it outperforms the traditional Cronbach alpha coefficient, making it the best for use in different data cases, namely the total omega reliability coefficient (ω_t), the greatest lower bound reliability coefficient (glb), the maximum Guttman lower bound reliability coefficient (λ_4), and the polychoric alpha reliability coefficient (α_{poly}). The best performance among them was Total Omega Coefficient performance (ω_t), which gave the highest values of reliability with the lowest relative bias for the majority of interactions between cases of data types and tests, and the diversity of sample sizes.

Key words: Alpha reliability coefficient - Guttman lower bounds - Omega reliability coefficient - Beta reliability coefficient - Maximal reliability coefficient - Stratified alpha reliability coefficient - Polychoric alpha reliability coefficient - Simulation study

مقدمة

تستخدم الاختبارات والمقاييس والاستبيانات في العلوم السلوكية والاجتماعية لقياس المعرفة والمهارات والاستجابات للمواقف لدى المشاركين أو المفحوصين، وبمجرد تطبيق مقياس أو استبيان يتم استخدام درجة الاختبار لتلخيص معرفة أو موقف أو أداء المستجيب. ومن الخصائص المرغوبة لدرجات الاختبار أنه إذا تم إعادة الاختبار عدة مرات في ظل ظروف الاختبار المتطابقة بالنسبة لكل مفحوص فإنه سيحقق دوما نفس النتيجة. ويشار إلى هذه الخاصية بمصطلح الثبات لدرجات الاختبار ويعتبر حساب معامل الثبات من أهم ما يؤخذ في الاعتبار عند حساب الخصائص السيكومترية للمقاييس والاختبارات النفسية (McDonald, 1999; Revelle & Zinbarg, 2009).

ومن المشاكل الأساسية للقياس في جميع العلوم حساب الثبات Reliability حيث أن القياسات دائما ما تكون مشوشة بأخطاء القياس. ولأن القياس النفسي أكثر صعوبة من القياسات الخاصة بالعلوم الطبيعية الأخرى ، فقد استفاض علماء النفس في دراسة مشكلة قياس الثبات (Guttman, 1945; Lord, 1955; Cronbach, 1951; McDonald, 1999) ولا يزال مشاكل قياس الثبات مسار بحث وجدل بين الباحثين حتى الأن (Sijtsma, 2009; Revelle & Zinbarg, 2009; Bentler, 2009; McNeish, 2017; Novak, 2020) وعلى الرغم من أن التطورات الحديثة في نظريات القياس ووسائل قياس الثبات قد تجاوزت بكثير مساهمات الباحثين الأوائل، فإن الكثير من هذه الأدبيات تناولت الموضوع من ناحية تقنية أكثر من كونها تطبيقية وكانت معظم الدراسات موجهة للمتخصصين بدلاً من المستخدمين لأدوات القياس المختلفة.

ويعتبر موضوع حساب الثبات أساسي لفهم الارتباطات بين المتغيرات المشاهدة وتأثيرها على العلاقات بين المفاهيم النفسية الأساسية، وكيف أن التقديرات المشاهدة لدرجة الشخص تتجاوز أو تنخفض عن تقديرات درجاتهم الحقيقية الكامنة، وكيفية تقدير فترات الثقة حول أي قياس بعينه. يتيح فهم الطرق العديدة لتقدير الثبات بالإضافة إلى طرق استخدام هذه التقديرات إمكانية تقييم الأفراد بشكل أفضل وتقييم تقنيات الاختيار والتنبؤ بالأداء، ويمكننا أن ننسب الفضل إلى سبيرمان (Spearman, 1904) في إضفاء الطابع الرسمي الأول على الثبات، حيث طور معامل الارتباط الرتبي وأنشأ المبادئ الأساسية لنظرية

Print:(ISSN 1687-2649) Online:(ISSN 2536-9091)

الثبات. وكانت رؤية سبيرمان الأساسية هي أن درجة الاختبار المشاهدة يمكن أن تقسم إلى بنائين غير قابلين للرصد: الدرجة الكامنة المطلوب قياسها ودرجة الخطأ المتبقية. ويتطلب تحديد جودة أي مقياس أو اختبار نفسي أن يظهر المقياس كلاً من الصدق والثبات المناسب لغرض القياس. فبينما الصدق يتعامل مع مدى دعم الدلائل ومدى الدعم النظري لتفسير درجات أداة القياس مقابل الاستخدامات المقترحة للمقياس، فإن الثبات يعالج مدى توقع أن تكون درجات الأداة قابلة للتكرار والتعميم. يمكن حساب الثبات، على سبيل المثال، عن طريق قياس مدى ارتباط درجات المقياس بالتطبيق المتكرر (ثبات إعادة الاختبار)، أو بين محكمي المقياس (الثبات بين المحكمين)، أو عبر نصفين من الأداة (ثبات التجزئة النصفية) أو عبر كل نصفين محتملين من الأداة (الاتساق الداخلي). وبالنسبة للمقاييس التي يقدم فيها المفحوص استجابات على الأسئلة الفردية أو المفردات التي تهدف إلى معرفة التركيبات النفسية الرئيسية، فإن هذا النوع الأخير من الثبات هو الأكثر شيوعًا في الاستخدام (Hancock & An, 2020).

مشكلة الدراسة

إن مفهوم الثبات ببساطة هو ارتباط الاختبار باختبار مثله تماماً، أو يمكن القول بأنه جزء الاختبار الذي ليس بسبب خطأ، أو إلى أي مدى ترتبط الدرجة المقاسة بالدرجة الحقيقية. ولكن ليس هناك معامل ثبات واحد فقط يجب حسابه وتسجيله، بل هناك مجموعة متنوعة من المعاملات، كل منها هو الأكثر ملائمة لأغراض معينة، مثلا هل نحاول التعميم على الفقرات، أم بمرور الوقت، أو على المحللين؟ هل نحن نقدر الأبعاد من حيث الأحادية أو تشبع العامل العام أو ثبات التباين الكلي؟ كل من هذه الأسئلة يؤدي إلى تقدير مختلف، لذا بدلاً من أن نسأل ما هو الثبات، يجب أن نتساءل أي ثبات سنستعمل ومن أجل ماذا سيتم حساب معامل الثبات؟ (Revelle & Condon, 2019).

معامل ألفا α هو معامل الثبات الأكثر استخدامًا لتقدير الثبات في البحوث التطبيقية كما ذكر Sijtsma (2009)، نظرا لشعبيته الكبيرة ومع ذلك ، فإن حدوده معروفة جيدًا (Lord and Novick, 1968; Cortina, 1993; Yang and Green, 2011) ، ومن أهمها افتراضات الأخطاء غير المترابطة وافتراضات تكافؤ تاو الأساسية والتوزيع الطبيعي لدرجات العينات. افتراض الأخطاء غير المترابطة (درجة الخطأ لأي زوج من الفقرات

Print:(ISSN 1687-2649) Online:(ISSN 2536-9091)

غير مرتبطة) هو افتراض نظرية الاختبار الكلاسيكية Classical Test Theory (Lord and Novick, 1968)، وقد يعني وجود بيانات لا تتبع هذه الافتراضات وجود تركيبات متعددة الأبعاد أكثر تعقيدا، ولتقدير معامل الثبات لهذه البيانات تتطلب إجراءات مختلفة تأخذ هذه التعقيدات في الاعتبار كما ذكر (Tarkkonen and Vehkalahti, 2005; Green and Yang, 2015) ومن المهم أيضا تغيير الاعتقاد الخاطئ بأن معامل ألفا هو مؤشر جيد على أحادية البعد لأن قيمته ستكون أعلى إذا كان المقياس أحادي البعد. ولكن العكس هو الصحيح، كما أوضح Sijtsma (2009) ، حيث قد يؤدي تطبيقه في مثل هذه الظروف إلى تضخم قيمة تقدير الثبات (Raykov, 2001) وبالتالي ، قبل حساب ألفا ، من الضروري التحقق من أن البيانات تتناسب مع النماذج أحادية البعد. ويعد افتراض تكافؤ تاو الأساسي Essential Tau-equivalence (أي نفس الدرجة الحقيقية لجميع فقرات الاختبار، أو أن تشبع العامل يكون متساوي لجميع الفقرات في النموذج العاملي) شرطًا لأن تكون ألفا مكافئة لمعامل الثبات (Cronbach, 1951). فإذا تم انتهاك افتراض تكافؤ تاو، فذلك سوف يقلل من قيمة الثبات الحقيقية (Raykov, 1997a;Graham, 2006)، وأيضا البيانات التي تتوافق مع افتراض تكافؤ تاو بشكل عام لا يمكن الحصول عليها عمليًا (Teo and Fan, 2013)، ويكون النموذج المتجانس Congeneric Model (أي عوامل تشبعات المفردات على العامل تكون مختلفة) هو الأكثر واقعية.(Trizano-Hermosilla and Alvarado, 2016)

وبالنظر إلى الأدبيات الوفيرة التي تناولت قيود استخدام معامل ألفا والتحيز الذي يظهر بنتائجه (Revelle and Zinbarg, 2009; Sijtsma, 2009, 2015; Cho and Kim, 2015; Sijtsma and van der Ark, 2015) فإن السؤال الذي يطرح نفسه لماذا يستمر الباحثون في استخدام ألفا بالرغم من وجود معاملات بديلة تتغلب على هذه القيود والاشتراطات اللازمة لاستخدام معامل ألفا، ومن الممكن أن تكون أسباب نجاح معامل α وبقائها في الأدبيات العلمية واسعة النطاق، أنه يعتمد تطبيقه على طريقة بسيطة ومستقرة للقياس مثل مجموع أو متوسط استجابات المفردة ؛ ومن السهل مشاركتها مع المراجعين وقراء تقارير العلوم الاجتماعية والصحية ؛ ويمكن الحصول عليها باستخدام تصميم بسيط يعتمد على تطبيق أحادي للاستبيان ؛ ويتم حسابه بسهولة في حزم برمجية إحصائية

متنوعة مثل SPSS أو SAS أو Stata . ويشير Sijtsma (2009) إلى أن التطور في الطرق والأساليب الرياضية للقياسات النفسية ودراسة علم النفس قد تباعدت عن بعضها البعض حيث أصبحت طرق القياسات النفسية أكثر إحصائية وبقي علماء النفس بعيدا عن النظرة الرياضية المتعمقة للقياس النفسي. وبدون مناقشات واضحة للبدائل وتوفر البرامج المتاحة بسهولة للعثور على تقديرات بديلة للثبات، سيستمر معظم علماء النفس في استخدام معامل ألفا لحساب الثبات. واتفق Revelle & Zinbarg (2009) مع Sijtsma (2009) على أنه ينبغي تشجيع الباحثين على استعمال تقديرات أفضل للثبات بالإضافة إلى معامل ألفا، واختلفوا فيما هي المعاملات الأكثر مناسبة للاستخدام.

ونظرا لندرة الأبحاث العربية التي تناولت معاملات الثبات المختلفة واعتماد معظمها في حساب الثبات على معامل الثبات ألفا مع عدم التحقق من شروط تطبيقه، وحيث أن طرق محاكاة مونت كارلو، هي فئة واسعة من الخوارزميات الحسابية التي تعتمد على أخذ العينات العشوائية المتكررة للحصول على نتائج عددية. فالمفهوم الأساسي للمحاكاة هو استخدام التجارب العشوائية لحل المشكلات. فيكون استخدامها لتوفير التحكم في خصائص البيانات في الدراسة الحالية هو الأنسب. ومن هنا برزت أهمية هذه الدراسة وتبلورت مشكلة الدراسة في التساؤل التالي:

- ما هو معامل الثبات الأفضل للاستخدام في حالة اختلاف نوعية المقاييس التي يتم قياسها (أحادية البعد أو تعدد الأبعاد)، واختلاف حجم العينة، واختلاف عدد العبارات (طول المقياس)، واختلاف نوعية بيانات القياس (بيانات مكافئة لتاو – بيانات متجانسة)، واختلاف خيارات الاستجابة (استجابة ثنائية – استجابة متعددة من نوع ليكارت)؟

مصطلحات الدراسة

الثبات: اتساق القياس، أي الدرجة التي يكون فيها الاختبار أو أداة القياس خاليًا من الأخطاء العشوائية الناتجة من عملية القياس، مما يؤدي إلى نفس النتائج عبر تطبيقات متعددة لنفس العينة.

معامل الثبات: مؤشر خالٍ من الوحدات وهو يصف تناسق الدرجات، وتوضح قيمته ، التي تتراوح من ٠ إلى ١ ، تقديرًا لمقدار تباين النتيجة الذي تم الحصول عليها والذي يرجع إلى

التباين الحقيقي بدلاً من الخطأ، وكلما زاد هذا المعامل، زادت الثقة في أن الدرجات التي تم الحصول عليها في أوقات مختلفة في ظل ظروف مماثلة مع نفس المشاركين ستكون متشابهة. (APA Dictionary of Psychology, 2007)

أهداف البحث:

- المقارنة بين مجموعة من معاملات الثبات لتحديد افضليتها ومناسبتها للاستخدام حسب موقف القياس واشتراطات البيانات .
- تحديد معامل الثبات الأفضل للاستخدام في حالة اختلاف نوعية المقاييس التي يتم قياسها من ناحية الأبعاد.
- تحديد معامل الثبات الأفضل للاستخدام في حالة اختلاف حجم العينة أو اختلاف عدد العبارات.
- تحديد معامل الثبات الأفضل للاستخدام في حالة اختلاف خيارات الاستجابة (استجابة ثنائية – استجابة متعددة من نوع ليكارت) أو اختلاف نوعية بيانات القياس (بيانات مكافئة لتاو – بيانات متجانسة).

أهمية البحث:

تبرز أهمية هذه الدراسة في كونها تلقي الضوء على معاملات الثبات المختلفة وإتاحة البدائل المختلفة لحساب الثبات نظرا لأهميته في قياس الخصائص السيكومترية للمقاييس والاختبارات وتوضيح الأخطاء يقع فيها الباحثون عند استخدام معاملات الثبات، حيث أن الاستخدام غير السليم لهذه المعاملات يؤدي إلى نتائج خاطئة لا يمكن تعميمها أو الاعتماد عليها.

الإطار النظري والدراسات السابقة:

في العديد من مجالات علم النفس والعلوم السلوكية، لا يمكن ملاحظة المتغيرات النفسية بشكل مباشر ويتم قياسها بمقاييس أو أدوات تتكون من مجموعة من المفردات أو الأسئلة والعبارات. تقيس هذه الفقرات بشكل غير مباشر المتغير النفسي موضع الاهتمام من خلال استنتاج أن بعض البنية الأساسية تتجلى من خلال هذه المفردات. حيث لا يمكن لدراسة ظاهرة ما قياس مقدار الظاهرة في نفسية الشخص المفحوص بشكل مباشر. بدلاً من

ذلك، يتم إنشاء مفردات المقياس للفرد وإذا كان لدى الفرد نسبة عالية من السمة المنوط قياسها، فإن هذه السمة تتجلى من خلال ردود معينة على هذه العبارات. ونظرًا لأن معظم القياس في علم النفس يتم من خلال استخدام أدوات القياس غير المباشرة، غالبًا ما يستعين الباحثون بمعامل الثبات لإثبات أن العبارات المكونة للمقياس موثوقة ويمكن الاعتماد عليها، مما يعني أن الدرجات المستندة إلى العبارات متسقة بشكل معقول، والاستجابات على المقياس قابلة للتكرار، وأن الاستجابات لا تتكون ببساطة من ضوضاء عشوائية. بعبارة أخرى، يوفر تحليل الثبات دليلاً على أن المقياس يقيس باستمرار الشيء نفسه.

الثبات

بدأ البحث في ثبات المقاييس في علم النفس مع Spearman (1904, 1910)، الذي قدم مفاهيم الدرجة الحقيقية وخطأ القياس، وكذلك تصحيح التوهين وحساب ثبات التجزئة النصفية. ووفقًا له، فإن الثبات هي وسيلة لإيجاد الارتباط الصحيح بين القيم "الحقيقية" من خلال تصحيح التوهين (تقليل التأثيرات) التي تحدث للارتباط المشاهد، والتي تنتج عن الأخطاء العشوائية الناتجة أثناء عملية القياس.

تم تعريف الثبات على أنه مربع الارتباط بين الدرجات الحقيقية والدرجات المشاهدة (Lord & Novick, 1968, p. 61) أو كنسبة تباين الدرجة الحقيقية للمقياس إلى تباينه الإجمالي (McDonald, 1999). تكون هذه التعريفات متكافئة رياضيًا عندما يكون تباين الدرجات المشاهدة موجب (Raykov & Marcoulides, 2011) ويفترض كلا التعريفين الوصول إلى تباين الدرجات الحقيقية للمقياس. ونظراً لأن التباين الدرجات الحقيقية غير معروف ولا يمكن تقديره إلا من خلال البيانات المشاهدة، فإن معظم تقديرات الثبات تعتمد على الافتراض القائل بأن التباينات المشتركة (التغاير) المشاهدة تمثل بالضرورة تباين الدرجات الحقيقية (Geldhof, Preacher& Zyphur, 2014).

لقد اتخذ تقدير الثبات أشكالًا عديدة منذ تصور سبيرمان للأخطاء في الملاحظة (Spearman, 1904). على الرغم من التاريخ الطويل، فقد كافح الباحثون للعثور على معامل للثبات مناسب في البيئات الواقعية. فعلى الرغم من دراسة Cronbach (1951) حول معامل ألفا هي واحدة من أكثر الدراسات التي يتم الاستشهاد بها بشكل كبير في نظرية القياس النفسي، إلا أنه غالبًا ما يتم إساءة استخدام معامل الثبات ألفا باستعماله مع عدم

Print:(ISSN 1687-2649) Online:(ISSN 2536-9091)

التحقق من اشتراطاته قبل التطبيق (Sijtsma, 2009). كان كرونباخ قادرًا على إثبات أن معامل ألفا هو متوسط جميع عوامل الثبات المتحصل عليها عن طريق التجزئة النصفية من خلال وضع افتراضات صارمة بأن كل فقرة من الفقرات في المقياس يُظهر تكافؤ تاو (الفقرات لها تشبعات متكافئة على العامل) وهي أحادية الأبعاد (Novick and Lewis, 1967). عندما يلبي المقياس هذه الافتراضات وتكون البيانات على مستوى المجتمع، ستكون جميع معاملات ثبات التجزئة النصفية متساوية. لذلك، بينما يمكن للباحث أن يصل إلى المعلمة المناسبة عن طريق اختيار الحد الأقصى أو الأدنى لثبات التجزئة النصفية، فإن المتوسط يكون معامل غير متحيز في العينة. وعلى الرغم من صحة الافتراضات التي وضعها Cronbach (1951)، إلا أنه لا يتم الوفاء بها في الظروف الواقعية.

في البداية اقترح سبيرمان صيغة تستند إلى افتراض أن هذه الأخطاء العشوائية يمكن قياسها باستخدام مقدار التباين بين القياسات المتتالية لنفس البنى النفسية، وبالنظر إلى أن الارتباط الحقيقي المفترض يتضاءل، فإذا تم قسمة الارتباط الذي تم الحصول عليه على جذر ناتج حاصل ضرب ثبات المتغيرين، فإن الارتباط غير المتأثر يمثل نسبة المتغيرات المقاسة بدون أخطاء. يشير تقييم الثبات في المقام إلى نتائج الارتباطات المشاهدة x وy مع أشكالهما المتوازية x' وy' التي تم الحصول عليها من الاختبارات الموازية التي تقيس نفس البنية وبالتالي لها نفس المتوسطات الحسابية والتباين للنتائج والأخطاء غير المرتبطة للتباينات المتساوية (Lord & Novick, 1968)، وكذلك نفس الترابطات المتساوية مع المتغيرات الخارجية ولتحديد القيم الحقيقية من الضروري إيجاد متوسط الارتباط بين مجموعتين أو أكثر من مجموعات النتائج التي تم الحصول عليها بشكل مستقل.

$$\mathbf{r_{x'y'}} = \frac{\mathbf{r_{x,y}}}{\sqrt{\mathbf{r_{xx'}}}\mathbf{r_{yy'}}} \qquad (1)$$

ولتقييم الثبات في حالة توفر اختبار واحد فقط كان الحل هو تجزئة الاختبار إلى نصفين مع إضافة تصحيح (k=2) يضع في الاعتبار هذه التجزئة (Brown, 1910; Spearman, 1910)

$$\boldsymbol{\rho XX' *= \frac{kr_{XX'}}{1+(k-1)r_{XX'}}} \qquad (2)$$

Print:(ISSN 1687-2649) Online:(ISSN 2536-9091)

عندما يكون تباين نتائج نصف الاختبار غير متكافئة، فلا مبرر لاستخدام معادلة سبيرمان– براون. لذلك، اقترح Flanagan (1937) و Rulon (1939) صياغة مختلفة لثبات التجزئة النصفية، والتي تأخذ في الاعتبار تباين النتائج. وشكلت تفسيراتهم للثبات مع مفهوم سبيرمان الأساس للتفسير الكلاسيكي لمفهوم الثبات، حيث قدموا صياغة مختلفة باستخدام الارتباط والانحرافات المعيارية. وتعتبر معادلة Flanagan–Rulon هي شكل خاص من المعامل ألفا الأكثر استخدامًا (Cronbach, 1951).

$$\rho_{\ Flanagan} = 2\left(1 - \frac{var_1 + var_2}{var_X}\right) \qquad (3)$$

$$\rho_{\ Rulon} = 1 - \frac{var_e}{var_x} \qquad (4)$$

$$\rho_{\ Flanagan-Rulon} = \frac{\left(4r_{1,2}SD_1SD_2\right)}{var_x} \qquad (5)$$

الثبات من وجهة نظر النظرية الكلاسيكية للقياس:

استنادًا إلى نظرية القياس الكلاسيكية CTT (Nunnally, 1978) ، تتكون الدرجة المشاهدة X من مكونين: الدرجة الحقيقية T بالإضافة إلى بعض أشكال الخطأ e مع توقع أن الخطأ هو عشوائي وليس منهجي.

$$X = T + e \qquad (6)$$

تفترض نظرية الاختبار الكلاسيكية أنَ مكونات الخطأ غير مرتبطة بالأجزاء الحقيقية وكذلك مع بعضها البعض. نتيجة لذلك، فإن مصفوفة التغاير (Γ) للفقرات المشاهدة هي مجموع مكونين: مصفوفة التغاير ($\Gamma\tau$) للأجزاء الحقيقية ومصفوفة التغاير ($\Gamma\xi$) لمكونات الخطأ:

$$\Gamma = \Gamma\tau + \Gamma\xi \qquad (7)$$

افتراض عدم ترابط الأخطاء يعني أن مصفوفة التغاير للأخطاء $\Gamma\xi$ هي مصفوفة قطرية، وبالتالي فإن أقطار مصفوفة التغاير للفقرات المشاهدة Γ ومصفوفة التغاير للدرجات الحقيقية $\Gamma\tau$ متطابقين. يعد افتراض الأخطاء المستقلة أمرًا ضروريًا للقياسات المشار إليها. توجد العديد من الحالات التي تؤدي إلى عدم تحقق افتراض الأخطاء المستقلة، فعلى سبيل المثال في حالة تطبيق اختبار ما بحدود زمنية معينة وكان هناك فقرات لم يتم الرد عليها، قد

ترتبط أخطاء الفقرات الأخيرة ولا يتحقق استقلال الأخطاء، أو في حالة تطبيق اختبارات طويلة أو صعبة قد تترابط أخطاء الاختبار بسبب تأثير التعب أو انخفاض الدافع أثناء انعقاد الاختبار.

ويمكن التعبير عن المعادلة الأساسية لنظرية الاختبار الكلاسيكية بالصيغة التالية:

$$x_i = \tau_i + \varepsilon_i \qquad (8)$$

حيث x_i هي الدرجة المشاهدة (أو درجة الاختبار)، τ_i هي الدرجة الحقيقية الكامنة التي تكمن وراء x_i ، وε_i هي خطأ القياس العشوائي المحدد على أنه الفرق بين الدرجة الحقيقية والدرجة المشاهدة $x_i - \tau_i$ بافتراض أنه لا يوجد تباين مشترك بين الدرجة الحقيقية وخطأ القياس $cov(\tau_i, \varepsilon_i) = 0$ و $E(\varepsilon_i) = 0$ ويطلق على المعادلة السابقة نموذج الدرجة الحقيقية الكلاسيكية.

ويمكن تحديد الأشكال الثلاثة الرئيسية للنموذج (8) باختبارين:

$$x_i = \beta_i \tau_i + \varepsilon_i \quad and \quad x_j = \beta_j \tau_j + \varepsilon_j \qquad (9)$$

حيث أخطاء القياس ε_i للاختبار الأول وأخطاء القياس ε_j للاختبار الثاني غير مرتبطين بينما الدرجة الحقيقية الكامنة للاختبار الأول τ_i والدرجة الحقيقية الكامنة للاختبار الثاني τ_j متساويين، وعلى الرغم من استخدام مؤشرات مختلفة للدرجات الحقيقية، يتم افتراض درجة حقيقية واحدة للاختبارين $\beta_i = \beta_j$. ويمكن الاشارة إلى التباين في أخطاء القياس ε_i وε_j بالرموز $\sigma^2_{\varepsilon_i}$ و $\sigma^2_{\varepsilon_j}$ على التوالي.

في المعادلة رقم (9):

- في حالة $\beta_i = \beta_j = 1$ وكانت $\sigma^2_{\varepsilon_i} = \sigma^2_{\varepsilon_j}$ تكون الاختبارات متوازية (Parallel Measures).
- في حالة $\beta_i = \beta_j = 1$ ولكن $\sigma^2_{\varepsilon_i} \neq \sigma^2_{\varepsilon_j}$ تكون الاختبارات مكافئة لمعامل تاو (Tau−equivalent Measures).
- في حالة $\beta_i \neq \beta_j$ وكانت $\sigma^2_{\varepsilon_i} \neq \sigma^2_{\varepsilon_j}$ تكون الاختبارات متجانسة (Congeneric Measures).

النموذج الموازي (Spearman, 1904) هو النموذج الأساسي المستخدم في القياس النفسي. نموذج تاو المكافئ (Novick and Lewis, 1967) والنموذج المتجانس (Jöreskog, 1971) هما نموذجان أكثر عمومية له.

ومن خلال النموذج الموازي يمكن تعريف ثبات الاختبار الذي يتكون من فقرات n على النحو التالي:

$$\rho_{tt} = 1 - \frac{\sigma_e^2}{\sigma_t^2} \qquad (10)$$

σ_e^2 هو تباين الخطأ و σ_t^2 هو التباين الكلي لدرجات الاختبار

حيث التباين الكلي:

$$\sigma_t^2 = \sum_{i=1}^{n}\sum_{j=1}^{n}\Gamma_{\mathrm{ij}} \qquad (11)$$

(i) تمثل الفرد و (j) تمثل الفقرة

وتباين الخطأ:

$$\sigma_e^2 = \sum_{i=1}^{n}\Gamma_{\mathrm{eii}} \qquad (12)$$

من المعادلات السابقة يمكن التعبير عن الثبات بالمعادلة التالية

$$\rho_{tt} = 1 - \frac{\sigma_e^2}{\sigma_t^2} = 1 - \frac{\sum_{i=1}^{n}\Gamma_{\mathrm{eii}}}{\sum_{i=1}^{n}\sum_{j=1}^{n}\Gamma_{\mathrm{ij}}} \qquad (13)$$

من خلال النظرية الكلاسيكية للقياس، في حال أن المتوسط الحسابي للخطأ هو صفر، فإن الدرجات المشاهدة للمشاركين في حالة وجود عدد لا حصر له من تطبيقات الاختبار ستكون مساوية للدرجة الحقيقية، وأن تباين الخطأ لا يرتبط بتباين الدرجة الحقيقية كما في حالة تباين الخطأ في القياسات المتوازية (Lord, 1959). ونظراً لأن التباين الحقيقي لا يمكن تحديده بشكل لا لبس فيه، فتستخدم طرقًا مختلفة لتقدير الثبات. وحيث أن هذه الطرق لا تتضمن نفس تباينات الخطأ، لذا فهي تتضمن أنواعًا مختلفة من الثبات: مثل ثبات القياس المتكرر (repeated measurement) أو ثبات إعادة الاختبار (test–retest) التي نحصل من خلالها على استقرار المعاملات، وثبات النماذج المتوازية أو البديلة (parallel/alternative) التي نحصل منها على معاملات التكافؤ وثبات الاتساق الداخلي

(internal consistency) الذي نحصل من خلاله على معاملات الاتساق الداخلي. وهذه المعاملات المختلفة ليست واضحة وقابلة للتبادل فيما بينها ولها قيم مختلفة في حالة تصحيح تأثير الخطأ وتحديد خطأ القياس المعياري والتطبيقات العملية الأخرى (Cronbach, 1947).

- **ثبات القياسات المتكررة (Test–Retest Reliability)**

يشير ثبات القياسات المتكررة إلى تطبيق نفس الاختبار في نقطتين زمنيتين وحساب معامل الارتباط بين النتائج المشاهدة. على الرغم من أنه باستخدام هذه الطريقة، من الممكن تحديد الخصائص الدائمة للعينة (Cronbach, Gleser, Nanda and Rajaratnam, 1972)، إلا أن لها عيوبًا معينة، فقد يؤدي القياس في النقطة الأولى إلى التأثير على النتائج عند نقطة القياس الثانية بسبب الاحتفاظ بالمعلومات أو التمرين على الاستجابات (Cronbach and Furby, 1970). العيب الثاني هو الفاصل الزمني بين نقاط القياس، حيث يمكن أن يزيد الفاصل الزمني القصير من احتمالية انتقال التأثير الحادث في الاختبار الأول بسبب التذكر والتمرين وما إلى ذلك، في حين أن الفاصل الزمني الطويل قد يزيد من احتمال تأثير النمو العقلي والجسمي والاجتماعي لأفراد العينة، وهذه الطريقة لها ما يبررها في حالة تطبيق مقاييس الخصائص المستقرة (Allen and Yen, 1979). علاوة على ذلك، فإن العيب الأكبر هو أنه لا يمكن التمييز ما إذا كان الاختلاف بين التطبيقين عائد إلى تغيير حقيقي في الدرجات أم إلى عدم ثبات الاختبار، ويمثل ارتباط (الاختبار – إعادة الاختبار) الدرجة التي يؤثر بها خطأ القياس على نتائج الاختبار ومقدار التغيير في الدرجات الحقيقية. بالإضافة إلى ذلك، يعتمد هذا الارتباط جزئيًا على الارتباطات التلقائية لعبارات الاختبار، مما يزيد بشكل غير واقعي من تقييم الثبات. أخيرًا، يمكن ملاحظة معامل استقرار مرتفع نسبيًا للاختبار بالرغم من انخفاض الاتساق الداخلي (Nunnally and Bernstein, 1994).

- **ثبات الصور المتوازية أو المتكافئة (Parallel/Alternative Reliability)**

ثبات الصور (الاختبارات) المتوازية أو المتكافئة هو الارتباط بين الدرجات المشاهدة في شكلين متكافئين من الاختبار. نظرًا لأنه غالبًا ما يتعذر تأكيد التوازي بين الاختبارين من الناحية العملية، فقد تم استخدام الاختبارات البديلة بشكل أساسي التي لها متوسطات حسابية

Print:(ISSN 1687-2649) Online:(ISSN 2536-9091)

متساوية تقريبًا، وتباينات وارتباطات متساوية تقريبا مع المقاييس الأخرى، ومن ثم فإنه من المقبول معاملتها على أنها متوازية (Allen and Yen, 1979) والمضي قدمًا في حساب الارتباطات بين نتائجهما.

عيوب هذا النهج هي نفسها بالنسبة لثبات إعادة الاختبار مع إضافة خطأ جديد بسبب أخذ عينات من المفردات، ولكن تقييم الثبات لا يتأثر بتذكر أسئلة الاختبار وعدم توفر نفس ظروف التطبيق أو بالنمو العقلي الذي قد يطرأ على المفحوصين بعد زمن (Cohen and Swerdlik, 2018). في حالة تطبيق الاختبار الأصلي ثم تطبيق الاختبار البديل بعد زمن، من الممكن مقارنة ارتباطات النماذج البديلة التي تم الحصول عليها بعد فترة زمنية أقصر وأطول، حيث يشير الارتباط الأعلى في حالة الفاصل الزمني الأقصر إلى أن التباين في الخصائص بمرور الوقت مصدر مهم لعدم الثبات، أي الزيادة في القياس تشير إلى حالة معينة للاختبار بدلاً من الزيادة في السمة المراد قياسها (Nunnally and Bernstein, 1994).

يسمى تقييم الثبات الذي تم الحصول عليه من خلال هذه الطريقة بمعامل الاستقرار والتكافؤ (Cronbach, 1947) وهو عمومًا أقل من معامل الاستقرار والتكافؤ المحدد للأشكال البديلة للاختبار المطبق على نفس العينة (Crocker and Algina, 1986).

- **ثبات الاتساق الداخلي (Internal Consistency Reliability)**

يشير الاتساق الداخلي إلى تقييم الثبات بناءً على تطبيق اختبار واحد في وقت واحد، ويتم تقسيمه إلى طريقتين، إحداهما تعتمد على طريقة التجزئة النصفية والأخرى على تحليل التباين في هياكل التباين لمفردات الاختبار (Crocker and Algina, 1986). في البداية، كان الأسلوب الأكثر شيوعًا هو ثبات التجزئة النصفية (Split Half)، استنادًا إلى تقسيم الاختبار إلى جزأين متساويين يمكن التعامل معه كاختبار بديل. لا يخضع ثبات الاتساق الداخلي لأوجه القصور الموجودة بأساليب تقدير الثبات الأخرى، ومع ذلك، فإن طريقة التجزئة النصفية لا يفضل استخدامها إذا كان الاختبار لا يمكن تقسيمه إلى أجزاء متوازية بسبب عدد المفردات أو خصائصها (Allen and Yen, 1979). ونظرا لسهولة الحسابات الخاصة بهذه الطريقة، كان التقييم الأكثر شيوعًا للثبات هو ثبات الاتساق الداخلي (Hogan, Benjamin, and Brezinski, 2000).

Print:(ISSN 1687-2649) Online:(ISSN 2536-9091)

وابتكر Kuder and Richardson (1937) طريقة لحساب ثبات الاختبار المكون من مفردات ثنائية التصحيح (٠، ١) دون تجزئته إلى نصفين. كان الهدف هو التخلص من الصعوبات الناتجة عن استخدام معادلة سبيرمان– براون وطريقة التجزئة النصفية. حيث انتقدا كيفية تقسيم الاختبار من خلال طريقة التجزئة النصفية إلى جزأين محددين بشكل عشوائي، وأن طريقة التقسيم يمكن القيام بها بالعديد من الطرق، كل منها يعطي ثبات للاختبار مختلف. وقاما بعرض عدة خيارات لمواقف مختلفة، تتضمن افتراضات وتقديرات مختلفة، كان الافتراض الأساسي الذي ارتكز عليه عملهم هو أن مصفوفة الارتباطات بين المفردات هي مصفوفة أحادية الرتبة.

وهذا يتوافق مع الاختبار الذي تقيس فيه جميع المفردات نفس العامل أو بمعنى أخر افتراض أحادية البعد. وكان منطق حساب معامل الثبات لكودر وريتشاردسون يتلخص فيما يلي: إيجاد متوسط الارتباطات بين المفردات (j) في الاختبار، مع افتراض أن متوسط معامل الارتباط هو الثبات لكل مفردة على حدة ومن ثم تطبيق معادلة سبيرمان–براون لحساب ثبات اختبار يتكون من مفردات ثنائية الاستجابة.ومن بين الصيغ العديدة التي وضعها كودر وريتشاردسون ، كان أحدها مفضلًا، وعلى أساس رقمها في المقالة الأصلية ، فقد أطلق عليها Kuder–Richardson 20 أو معادلة كودر–ريتشاردسون٢٠ (KR20)

$$\rho_{KR-20} = \frac{n}{n-1}\left(1 - \frac{\sum_{n=1}^{n} p_j q_j}{var_x}\right) \qquad (14)$$

باستخدام مفردات اختبار n من j=1 إلى n، حيث p_j هي نسبة الاستجابات الصحيحة لمفردات الاختبار (j) و q_j هي نسبة الاستجابات غير الصحيحة لمفردات الاختبار(j) بحيث تكون $p_j + q_j = 1$

لم يسلم عمل كودر وريتشاردسون أيضًا من الانتقادات، حيث انتقد (Kelley, 1942) عملهم وذكر أن افتراضاتهم بشكل عام مقيدة للغاية، وعلى وجه الخصوص تساءل عن الافتراض القائل بأن مصفوفة الارتباطات بين المفردات هي مصفوفة أحادية الرتبة.

- **الحدود الدنيا للثبات (Reliability Lower Bounds)**

بعد أربعين عامًا من كتابة سبيرمان عن أخطاء القياس، حاول Guttman (1945) توحيد مفهوم الثبات فاشتق ستة حدود دنيا مختلفة للثبات، والتي تشير إلى تقييم الحد الأدنى

من الثبات. تعتمد معاملات الثبات التي طرحها على أنه يتم تحديد خطأ القياس بشكل صريح لكل فرد في محيط القياسات المتكررة ، والتي تنشأ منها ثلاثة مصادر منفصلة للتباين وهي القياسات المتكررة، والأفراد والمفردات، وعدم الثبات ويتم تعريفها على أنها التباين في النتائج في القياسات المتكررة، وبالنسبة للمفردات فلا يمكن أن تتغير في حالة القياسات المتكررة، وفي حالة استقلالية القياسين، فإن معامل الثبات يساوي معامل ارتباط النتائج من القياسات المستقلة، وبسبب قوة تأثير استقلالية القياس فمن الضروري التركيز على المعلومات التي يمكن الحصول عليها من قياس واحد، وهي تقدير للحد الأدنى للثقة. وبالتالي، فقد أزال هذا النهج الافتراضات حول الحاجة إلى القياس في نقطتين زمنيتين على الأقل والصعوبات المنهجية الناشئة عن افتراض استقلالية القياس.

تمثل المعاملات المطورة تحسينات متتالية لتقييم الثبات بناءً على المزايا المذكورة، وقد أثرت منطقية الاشتقاق لكل معامل فردي للثبات على تطور المعاملات التي تم استحداثها لاحقًا. كان المشترك بينهم جميعا هو أنه يمكن استخلاص الثبات من تطبيق واحد للاختبار، ويتجنب الحاجة إلى تطبيق اختبارين مستقلين، وبهذا أصبحت حسابات الثبات أكثر سهولة.

إذا أخذنا في الاعتبار في نقطة قياس واحدة أن هناك تباينًا في نتائج الأفراد لكل مفردة (j)، بالإضافة إلى تباين الدرجة الكلية لجميع الأفراد (var_X)، فيمكن تقدير حد الثبات الأقل بطريقة مبسطة

$$\lambda_1 = 1 - \frac{\sum_{j=1}^{J} var_j}{var_X} \qquad (15)$$

ومع ذلك، على الرغم من سهولة تحقيق استقلالية القياس فيما يتعلق بالأفراد، فإن استقلالية القياس فيما يتعلق بالمفردات ليست حالة واقعية إذا كان الأمر يتعلق بتقدير التباين المشترك الحقيقي. لذلك، فإن البديل الأفضل لـ λ_1 هو معامل الحد الأدنى للثبات λ_2 ، والذي يتضمن ضعف مجموع المربعات $\frac{J(J-1)}{2}$ لتغاير المفردة (j) والمفردات المتبقية (j-1) عند نقطة قياس واحدة.

$$\lambda_2 = \lambda_1 + \frac{\sqrt{\frac{J}{J-1} 2 \sum_{j=1}^{J} cov_{j,j-1}^2}}{var_X} \qquad (16)$$

Print:(ISSN 1687-2649) Online:(ISSN 2536-9091)

وتم اشتقاق معامل الحد الأدنى للثبات التالي λ_3 من معامل الحد الأدنى للثبات λ_2 لتسهيل الحساب، حيث أغفل مجموع مربعات التغاير المشترك للمفردات، ويكون تقدير الثبات له أكثر دقة من λ_1. يُعرف أيضا بإسم معامل ثبات ألفا لكرونباخ (α) لدعوة (Cronbach, 1951) إلى استخدامه بدلاً من معادلة سبيرمان–براون.

$$\lambda_3 = \frac{J}{J-1}\lambda_1 \qquad (17)$$

وبهذا تكون نسبة معاملات الثبات الثلاثة المذكورة أعلاه هي كما يلي:

$$\lambda_1 < \lambda_3 \leq \lambda_2$$

بالنسبة لمعامل الحد الأدنى للثبات λ_4 فإنه مشابه لمعامل ثبات التجزئة النصفية من حيث أنه يتطلب تقسيم النتائج إلى نصفين لهم تباينات منفصلة (var_1) و (var_2). ويمثل الحد الأدنى من الثبات بغض النظر عن طريقة تقسيم الاختبار المستخدمة ولا يتطلب افتراضات حول أحادية البعد وتحقق شروط تكافؤ تاو (نفس تشبعات المفردات على البنية العاملية مع عدم تساوي تباينات الأخطاء)، ومع ذلك، فمن المستحسن تقسيم الاختبار بطريقة يتم فيها زيادة قيمة المعامل إلى أقصى حد لأن هذا يمثل التقدير الأكثر دقة للحد الأدنى للثقة، وبالتالي يُسمى الحد الأقصى لمعامل التجزئة النصفية. إذا كانت تباينات النصفين متساوية، فإن λ_4 يساوي عدديًا القيمة التي يمكن الحصول عليها من خلال تطبيق معادلة سبيرمان–براون على هذين النصفين.

$$\lambda_4 = 2\left(1 - \frac{var_1 + var_2}{var_X}\right) \qquad (18)$$

يعتمد متوسط قيمة معامل الحد الأدنى للثبات λ_4 على جميع التقسيمات الممكنة للاختبار، وبالتالي فإن λ_4 يمثل تقديرًا أكثر دقة للثبات من λ_3.

بينما يستبدل معامل الحد الأدنى الخامس للثبات لجتمان λ_5 القيم القطرية لمصفوفة التغاير المشاهد لعدد Z من المفردات بضعف الجذر التربيعي للحد الأقصى (عبر المفردات) لمجموع مربع التغايرات المشتركة بين المفردات.

$$\lambda_5 = \lambda_1 + \frac{\sqrt[2]{\sum_{z=1}^{Z} \overline{cov}^2_{j,j-1}}}{var_X} \qquad (19)$$

Print:(ISSN 1687-2649) Online:(ISSN 2536-9091)

ويعتمد معامل الحد الأدنى السادس والأخير للثبات لجتمان λ_6 على الارتباطات المتعددة التي يمكن حسابها باستخدام مجموع أخطاء التباين للمفردة (j) بناءً على الانحدار الخطي المتعدد للمفردات المتبقية (j−1).

$$\lambda_6 = 1 - \frac{\sum_{j=1}^{J} var_{e_{j,j-1}}}{var_X} \qquad (20)$$

- **معامل الثبات (α) ألفا (Alpha Reliability Coefficient)**

معامل ثبات ألفا α لكرونباخ (Cronbach, 1951)، هو في الواقع معامل ثبات جتمان $\lambda 3$ ومعامل ثبات كودر– ريتشاردسون العام ρ_{KR20} ، وهو أحد أهم وأشهر معاملات الثبات انتشاراً (Cortina, 1993; Schmitt, 1996; Sijtsma, 2009; Yang and Green, 2011). ويستند إلى الافتراضات التالية: افتراض تكافؤ تاو الأساسي، أي افتراض أن تقيس جميع المفردات نفس البنية على نفس المقياس بنفس درجة الدقة، وافتراض التوزيعات الطبيعية المستمرة للمفردات، وافتراض التوزيع الطبيعي للنتائج الإجمالية، وافتراض عدم ارتباط أخطاء المفردات، وافتراض أحادية البعد، وافتراض البيانات المتصلة (Graham, 2006; Novick and Lewis, 1967; Raykov, 1997a, 1997b; Revelle and Zinbarg, 2009; Yang and Green, 2011). يمكن تفسير معامل الثبات ألفا على أنه متوسط جميع معاملات ثبات التجزئة النصفية الممكنة (متوسط معاملات الارتباط بين المفردات) المقدرة بناءً على مفهوم ثبات Flanagan–Rulon ، وبالتالي يمثل الارتباط المتوقع لاثنين من الاختبارات ذات الطول المتساوي بقياس نفس البنية مع مفردات مختلفة مأخوذة من نفس نطاق المفردات. بعبارة أخرى، فإن منطق تقييم ثبات التجزئة النصفية يذهب خطوة إلى الأمام من خلال التعامل مع كل مفردة باعتبارها اختبارًا فرعيًا. تتمثل ميزة هذا المعامل مقارنة بتقديرات ثبات التجزئة النصفية السابقة في أنه كثبات للتجزئة النصفية يكون الأكثر استقرارًا لأن قيمته لا تختلف اعتمادًا على طريقة تقسيم الاختبار إلى جزأين. إذا تم تطبيق معامل ثبات ألفا على اختبار يتكون من مفردات ثنائية، فسيؤدي إلى قيمة مماثلة للقيمة التي نحصل عليها من معادلة كودر–ريتشاردسون ٢٠ (Feldt, 1969).

هناك صيغتان مختلفان من معامل ثبات ألفا α:

Print:(ISSN 1687-2649) Online:(ISSN 2536-9091)

$$\alpha = \frac{J}{J-1}\left[1 - \frac{\sum_{j=1}^{J} var_j}{var_X}\right] \qquad (21)$$

حيث var_j هو تباين المفردات j و var_X هو التباين الكلي للدرجات المشاهدة والصيغة الأخرى تعتمد في حسابها على التباين المشترك (التغاير) $(cov_{j,j-1})$ بين المفردات (j) والمفردات المتبقية $(j-1)$

$$\alpha = \frac{J}{J-1}\left[1 - \frac{\sum_{j=1}^{J} cov_{j,j-1}}{var_X}\right] \qquad (22)$$

وكلا الصيغتين تؤديان إلى نفس النتيجة.

- **معامل ثبات أكبر حد أدنى ρ_{glb} (Greatest Lower Bound Reliability Coefficient)**

اشتق Jackson and Agunwamba (1977) معامل ثبات أكبر حد أدنى ρ_{glb} (greatest lower bound) بناءً على الافتراض الرئيسي للنظرية الكلاسيكية للقياس $cov_X = cov_T + cov_E$ ، أي أن مصفوفة التباين المشترك (التغاير) بين المفردات يمكن فصلها إلى مجموع مصفوفة التباين المشترك للدرجات الحقيقية، ومصفوفة التباين المشترك للخطأ (Ten Berge and Sočan, 2004) ولا يتطلب ρ_{glb} تحقق شرط أحادية البعد، ولكنه يتطلب أن تكون أخطاء القياس غير مرتبطة (Green and Yang, 2009a) .

ويمثل أدنى ثبات ممكن فيما يتعلق بمصفوفة الدرجات المشاهدة، مع تقييد أن مجموع تباينات الخطأ قد تم تعظيمها للأخطاء غير المرتبطة بالمفردات الأخرى. يتم طرح مجموع تباينات خطأ المفردة، المقدرة بناءً على خوارزميات معينة، من التباين الكلي للمفردات (Bentler and Woodward, 1980; Ten Berge, Snijders, & Zegers, 1981)

$$\rho_{glb} = 1 - \frac{\sum_{j=1}^{J}(var_{e_j})}{var_X} \qquad (23)$$

- **معامل ثبات بيتا (β) (Beta Reliability Coefficient)**

يعتمد تقدير معامل ثبات بيتا (β) الذي قام باشتقاقه Revelle (1979) أو تقدير الثبات لأسوأ تجزئة نصفية (worst split-half reliability) على التحليل العنقودي

Print:(ISSN 1687-2649) Online:(ISSN 2536-9091)

الهرمي ويجمع بين مبادئ القياس النفسي وإجراءات التجميع العنقودي، حيث يتم إجراؤه عن طريق تحديد مصفوفة التشابه أو الارتباط بين المفردات، ثم العثور على أزواج المفردات أو المتغيرات الأكثر تشابهًا من المصفوفة، ثم دمج أزواج المتغيرات في متغير مركب جديد، وبعد ذلك حساب تشابه المتغير المركب الجديد مع المتغيرات المتبقية ، ويتم تكرار الخطوات من بعد خطوة تحديد مصفوفة التشابه، مع الأخذ في الاعتبار المتغيرات الأولية والمتغيرات المركبة وأخيرا وقف الخطوات عندما لا يكون من الممكن دمج المتغيرات أو عند الوصول إلى المعيار المحدد.

على النقيض من معامل ثبات ألفا، يمكن تحديد معامل ثبات بيتا (β) بتقسيم الاختبار إلى جزأين (J2 , J1) بحيث يتم تقليل التباين المشترك بين النصفين إلى أقل ما يمكن، أو زيادة مجموع التباينات داخل الجزأين إلى أعلى ما يمكن. معامل ثبات بيتا (β) ليس حساسًا لطول الاختبار ويسمح بتجميع الاختبارات الفرعية وبالتالي يعطي تقييمًا أقل تحيزًا لتجانس الاختبار من معامل ثبات ألفا. بالإضافة إلى ذلك، يمكن استخدامه كقاعدة لاتخاذ قرار بدمج اختبارين فرعيين في اختبار واحد أعلى رتبة.

$$\beta = \frac{(J_1 + J_2)^2 \overline{cov_{1,2}}}{cov_X} \qquad (24)$$

وحيث أن معامل ثبات ألفا هو متوسط كل التقسيمات الممكنة للتجزئة النصفية للاختبار، فإن $\beta \leq \alpha$.

- **معامل الثبات (ω) أوميجا (Omega Reliability Coefficient)**

اقترح McDonald (1978, 1999) معامل ثبات أوميجا (ω) ويعرف أيضا باسم معامل ثبات أوميجا الكلية (ω_t)، استنادًا إلى طريقة التحليل العاملي للمفردات ويشبه حسابات هذا المعامل أوميجا، معامل الحد الأدنى للثبات λ_6 لجتمان، ولكنها تستخدم التباين الفريد للمفردات (μ^2) من تحليل العوامل للعثور على تباين الخطأ (e_j^2). يعتمد هذا على تحليل التباين var_X في درجات الاختبار، إلى أربعة أجزاء: التباين بسبب العامل العام (g) ، والتباين بسبب مجموعة من عوامل المجموعة (f)، وهي العوامل المشتركة لبعض المفردات ولكن ليست لكل المفردات، والتباين بسبب عوامل محددة (s) فريدة لكل مفردة، والتباين بسبب الخطأ العشوائي للقياس. ويمكن حسابها باستخدام القيم التي تم الحصول عليها من

خلال تحليل العوامل الاستكشافية والتأكيدية (Revelle & Condon, 2019; Zinbarg, Yovel, Revelle, & McDonald, 2006) ونظرًا لأنه لا يمكن التمييز بين كل نوع من التباين والخطأ العشوائي ما لم يتم إجراء الاختبار مرتين على الأقل ، فإن McDonald (1999) جمعهم معًا في حساب الخطأ.

هذه الطريقة لحساب معامل الثبات أوميجا تقوم بتصحيح التحيز في تقدير ثبات معامل ألفا، والذي يحدث عندما تنتهك الافتراضات حول تكافؤ تاو وأحادية البعد.

باعتبار

$$X = cg + Af + Ds + e \qquad (25)$$

وتكون تشاركيات المفردة (j) بناءا على العامل العام والعوامل المشتركة

$$h_{\mathrm{j}}^{2} = \mathrm{c}_{\mathrm{j}}^{2} + \sum f_{ij}^{2} \qquad (26)$$

ويكون التباين الفريد للمفردة

$$\mu_j^2 = \sigma_j^2\left(1 - h_j^2\right) \qquad (27)$$

في حالة اعتماد تشاركية المفردة (j)

إذا كانت (h_j^2) هي تشاركية المفردة (j)، بناء على العامل العام وكذلك عوامل المجموعة، فعندئذٍ بالنسبة للمفردات المعيارية تكون

$$\sigma_j^2 = \left(1 - h_j^2\right) \qquad (28)$$

وعليه يكون معامل أوميجا (ω_t) كالتالي:

$$\omega_t = 1 - \frac{\sum\left(1 - h_j^2\right)}{var_X} = 1 - \frac{\sum \mu^2}{var_x} \qquad (29)$$

وبالتعويض في المعادلة السابقة بتشبعات المفردات (λ_j) على العامل (k) يمكن صياغتها بالشكل التالي

$$\omega_t = \frac{\left(\sum_{j=1}^{J} \lambda_{j,k}\right)^2}{\left[\left(\sum_{j=1}^{J} \lambda_{j,k}\right)^2 + \left(\sum_{j=1}^{J} 1 - \lambda_{j,k}^2\right)\right]} \qquad (30)$$

عند وجود ارتباط بين الأخطاء، أو وجود أكثر من بُعد كامن في البيانات، يتم تقدير مساهمة كل بُعد في التباين الإجمالي المفسر، والحصول على ما يسمى بأوميجا الهرمية (ω_h) التي تمكننا من تصحيح أسوأ تحيز في التقدير α في حالة وجود بيانات متعددة الأبعاد

Print:(ISSN 1687-2649) Online:(ISSN 2536-9091)

Tarkkonen and Vehkalahti, 2005; Zinbarg et al., 2005; Revelle)
.(and Zinbarg, 2009

$$\omega_h = \frac{\left(\sum_{j=1}^{J} \lambda_{j,g}\right)^2}{\left[\left(\sum_{j=1}^{J} \lambda_{j,k}\right)^2 + \left(\sum_{j=1}^{J} 1 - \lambda_{j,k}^2\right)\right]} \qquad (31)$$

يختلف هذا المعامل عن المعامل السابق في أنه يستخدم مجموع التشبع على العامل العام (g) في البسط، وفي حالة البيانات أحادية البعد، يكون معامل ثبات أوميجا الكلية (ω_t) ومعامل ثبات أوميجا الهرمية (ω_h) مكافئتان.

- **معامل الثبات الأقصى (Maximal Reliability Coefficient–*H*)**

في كثير من الأحيان، كان الباحثون منشغلين بحساب ثبات مجموع الدرجات الإجمالية للاختبارات والمقاييس، والتي يشار إليها في كثير من الدراسات تحت مسمى الثبات المركب (Composite Reliability)، من ناحية أخرى ارتبطت طريقة بديلة لحساب الثبات بمفهوم التركيبة الخطية المثلى (optimal linear combination (OLC، والتي توفر أقصى قدر من الثبات. يتم تحقيق أقصى قدر من الثبات (Maximal Reliability) أو ما يسمى الثبات الأقصى، من خلال إعطاء وزن مثالي لكل فقرة. يُطلق على المجموعة الخطية المكونة من هذه الدرجات الإجمالية الموزونة، والتي تؤدي إلى أقصى قدر من الثبات، اسم التركيبة الخطية المثلى. (Raykov, 2004; Raykov, Gabler, & Dimitrov, 2016)

بافتراض أن مجموعة من الفقرات p المتصلة، و (x_1, x_2, x_3, ...x_p) هي مؤشرات لمقياس نفسي أحادي البعد. وبافتراض أن هذه الفقرات تقيس متغير كامن ξ و δ_1 و δ_2 و... ، δ_p هي أخطاء القياس للمؤشرات. يمكن التعبير عن نموذج التحليل العامل التأكيدي وتقاطعاته بالمعادلة التالية (Joreskog, 1971)

$$x_i = \tau_i + \lambda_i \xi + \delta_i \qquad (32)$$

حيث λ_iهو عامل تشبع الفقرة i على المتغير الكامنξ ، و τ_i هو التقاطع المصاحب، و δ_i هو خطأ القياس للفقرة i. وبافتراض عدم ارتباط أخطاء القياس ببعضها البعض وكذلك عدم ارتباطها مع المتغير الكامنξ، ولتحديد نموذج القياس، يتم اعتبار تباين العامل الكامن يساوي الواحد الصحيح ($var(\xi) = 1$) . وللحصول على حل غير معياري، نفترض أن تشبعات الفقرات (λ_1, λ_2, ...λ_p)، وتباينات الخطأ (θ_1,θ_2,....θ_p) حيث ($\theta =$

($var(\delta_i)$) على التوالي. وحيث أن الدرجة المركبة Z، للفقرات p لنموذج العامل الواحد يتم حسابها على النحو التالي:

$$Z = X_1 + X_2 + \cdots + X_P \qquad (33)$$

فيمكن حساب معامل الثبات المركب للقياس المتجانس (Congeneric measure) كالتالي:

$$\rho = \frac{(\lambda_1 + \lambda_2 + \cdots + \lambda_P)^2}{[(\lambda_1 + \lambda_2 + \cdots + \lambda_P)^2 + \theta_1 + \theta_2 + \cdots + \theta_P]} \qquad (34)$$

(McDonald, 1999; Raykov & Grayson, 2003)

من ناحية أخرى يتم تعريف التركيبة الخطية المثلى للقياس (Optimal Linear Combination) على أنها المجموعة التي تمتلك أعلى ثبات بين التركيبات الخطية الممكنة

$$Z^* = w_1X_1 + w_2X_2 + \cdots + w_PX_P \qquad (35)$$

ويتم تحديد الوزن الأمثل على النحو التالي (Raykov et al., 2016):

$$w_i = \frac{\lambda_i}{\theta_i} \qquad (i = 1, 2, \dots, P) \qquad (36)$$

وعند الجمع بين المعادلتين (35) & (36)، تكون التركيبة الخطية المثلى (OLC) أو الدرجة الإجمالية الموزونة مع أقصى قدر من الثبات (Maximal Reliability) كما يلي:

$$Z^* = \frac{\lambda_1}{\theta_1}X_1 + \frac{\lambda_2}{\theta_2}X_2 + \cdots + \frac{\lambda_P}{\theta_P}X_P \qquad (37)$$

وبهذا يمكن الحصول على الثبات الأقصى (Maximal Reliability) للمجتمع (ρ*) من المعادلة (Raykov et al., 2016)

$$\rho^* = \frac{\frac{\lambda_1^2}{\theta_1} + \frac{\lambda_2^2}{\theta_2} + \cdots + \frac{\lambda_P^2}{\theta_P}}{\left[1 + \left(\frac{\lambda_1^2}{\theta_1} + \frac{\lambda_2^2}{\theta_2} + \cdots + \frac{\lambda_P^2}{\theta_P}\right)\right]} \qquad (38)$$

من المعادلات السابقة نلاحظ أنه دائما ما يتحقق أن $\rho \leq \rho^*$ ، أي أنه لا توجد تركيبة خطية لها ثبات أكبر من ρ^* على وجه الخصوص، ويكون معامل الثبات الذي تم تحقيقه من التركيبة الخطية الموزونة للوحدات في المعادلة (36) دائمًا أقل من، إن لم يكن مساويًا، معامل الثبات الذي تم الحصول عليه من التركيبة الخطية المثلى (OLC). (Penev & Raykov, 2006)

Print:(ISSN 1687-2649) Online:(ISSN 2536-9091)

- **معامل ثبات ألفا الطبقية (Stratified Reliability Coefficient Alpha)**

أشار Feldt and Brennan (1989) انه حتى بالنسبة لاختبار واحد، فإن المفردات الموجودة في الاختبار نادرا ما تكون متجانسة، ففي كثير من الأحيان يتم تجميع المفردات التي في الاختبار لقياس أبعاد مختلفة بدرجة ما عن مجال الأساسي للمحتوى .معامل ثبات ألفا هي واحدة من أكثر المعاملات المستخدمة بشكل متكرر لقياس الثبات لاختبار يتم تطبيقه مرة واحدة أوضح Lord and Novick (1968) أن مفردات الاختبار يجب أن تحقق تكافؤ تاو لكي يكون تقدير معامل ألفا للثبات غير متحيز (unbiased) ونادرًا ما يتم الوفاء بهذا الشرط في البيانات العملية لأنه يتطلب قوة تمييز متساوية لجميع مكونات الاختبار وبعد أحادي للاختبار بالكامل، والتي يتم تمثيلها بالتشبعات المتساوية لجميع المكونات على عامل واحد في إطار النموذج التحليلي أحادي العامل (McDonald, 1999; Kamata *et al*, 2003)لذلك، يمكن افتراض أن الاختبار يتكون من طبقات من المفردات، مع افتراض أن المفردات في كل طبقة أحادية البعد. يمكن حساب ثبات الدرجة المركبة بناءً على ثبات الطبقات، حيث يتم التعامل مع كل منها على أنها اختبار فرعي واحد (Robitzsch, 2020a). يمكن التعبير عن معامل ثبات ألفا الطبقية، على النحو التالي (Feldt and Brennan, 1989)

$$r_{STRAT,\alpha} = 1 - \frac{\sum_i \sigma_i^2 (1 - r_i)}{\sigma_c^2} \qquad (39)$$

حيث $r_{STRAT,\alpha}$ هو ثبات الدرجات المركبة و r_i هو ثبات الطبقة (i) ، و σ_i^2 هو تباين الطبقة (i)، و σ_c^2 هو تباين الدرجات المركبة.

- **معامل ثبات ألفا الرتبية للأقسام المتعددة (Polychoric Ordinal Alpha Coefficient)**

قام زومبو وآخرون Zumbo et al. (2007) باختبار معامل ألفا لبيانات رتبية (متعددة الاستجابات) والذي عرف بإسم ألفا الرتبية للأقسام المتعددة لمعالجة إساءة استخدام معامل ألفا وانخفاض تقديرات الثبات الناتجة من تقدير معامل ألفا كرونباخ عند انتهاك الافتراضات الخاصة بتطبيقه، مثل تحقق تكافؤات تاو الأساسية للبيانات وأحادية البعد، ، يستخدم المعامل مصفوفة الارتباط متعدد الأقسام polychoric correlation matrix)

Print:(ISSN 1687-2649) Online:(ISSN 2536-9091)

(Zumbo et al., 2007 ، والتي تأخذ في الاعتبار بنية البيانات الفئوية المرتبة بدلاً من مصفوفة ارتباط بيرسون، والتي تفترض بنية بيانات على مستوى الفترة (Haldago–Tello et al.، 2008) وهذه البنية تقلل بشدة من العلاقة الحقيقية بين متغيرين متصلين عندما يظهر المتغيرين في توزيع ملتوي (skewed) للاستجابات المشاهدة (Gadermann et al., 2012).

تم اقتراح مصفوفة ارتباط الأقسام المتعدد من قبل بيرسون، ويمكن تعريفة بأنه الارتباط الخطي للتوزيع الطبيعي المقترح المشترك للبيانات (Ekström, 2010). وحساب هذه المصفوفة معقد للغاية وتعد الاختلافات الرئيسية بين ارتباط بيرسون للبيانات المستمرة وارتباط الأقسام المتعدد للبيانات الرتبية هي التوزيعات الكامنة التي يتم تقدير المصفوفة من خلالها، حيث يفترض كل من معامل ارتباط بيرسون ومعامل الارتباط متعدد الأقسام أن المتغيرات لها توزيع طبيعي أساسي ثنائي المتغير ؛ ومع ذلك ، يعتمد التوزيع متعدد الأقسام على السمة الكامنة التي تمثلها فئات الترتيب بينما يفترض معامل ارتباط بيرسون التوزيع الطبيعي القياسي المستمر ويمثل قوة العلاقة الخطية بين متغيرات الصف والعمود. (Zumbo, Gadermann, and Zeisser, 2007)

ومن مزايا استخدام معامل ثبات ألفا الرتبية للأقسام المتعددة. أولاً، من ناحية المفهوم، تعادل ألفا الرتبية معامل ألفا كرونباخ، ولكنها تستند إلى مصفوفة الارتباط متعدد الأقسام polychoric correlation matrix بدلاً من مصفوفة ارتباط بيرسون. لذلك، تشير الأدلة التجريبية إلى أنه تقدير أكثر دقة للقياسات التي تتضمن بيانات متعددة الاستجابات (Gadermann et al., 2012; Zumbo et al., 2007). ثانيًا، معامل ألفا الرتبي للأقسام المتعددة يأخذ في الاعتبار الاستجابات المتعددة كتعبيرات عن السمة الكامنة ويفسر ثبات الدرجات الرتبية المشاهدة باستخدام استجابات الفقرات المشاهدة، بينما تفسر ألفا كرونباخ ثبات الدرجات المشاهدة من خلال معاملتها على أنها بيانات متصلة (Gadermann et al., 2012). ثالثًا، تقدمت حزم برامج الكمبيوتر مثل R إلى النقطة التي يمكن فيها حساب أو إدخال مصفوفة ارتباط متعددة الأقسام لاستخدامها في تقدير ألفا الرتبية متعددة الأقسام ويمكن أيضا تفسير معامل ألفا الرتبي بسهولة نظرًا لأن القياس الناتج

يتراوح بين ٠ و١ حيث يدل الصفر على عدم وجود ثبات والواحد على ثبات كامل للمقياس (Lewis, 2007; Zumbo et al., 2007).

الدراسات السابقة

على الرغم من استخدام معامل ألفا المتكرر بين الباحثين، فإنها لا تكاد تخلو من النقد، فقد حذر Cronbach (1951) نفسه من سوء الاستخدام للمعامل، حيث أقر بأنه لا يمكن استخدامه للاختبارات القصيرة جدا، وللاختبارات التي لا يمكن تقسيمها إلى مجموعات فرعية متميزة. وناقش Green, Lissitz and Mulaik (1977) قيود معامل ألفا كمؤشر لأبعاد الاختبار. وتحدث Sijtsma (2009) عن الفائدة المحدودة للغاية لألفا كرونباخ وأفرد McNeish (2017) ورقة بحثية كاملة لنقد معامل ألفا وحث الباحثين بالاستغناء عنها لصالح المعاملات الأخرى حيث قال أن ألفا مليئة بالمشاكل، وأشار البعض الآخر إلى سوء استخدام معامل ألفا في الأبحاث النفسية (Cho & Kim, 2015; Green, Lissitz & Mulaik, 1977; Sijtsma, 2009; Schmitt, 1996). وتحدث Sijtsma (2009) عن أن معامل ألفا من أكثر المعاملات الإحصائية تعرضا لسوء الفهم والارتباك بين الباحثين. من ناحية أخرى مازال بعض الباحثين يعتبرون معامل ألفا معاملاً مفيدًا، حيث اعترض (Raykov & Marcoulides, 2017) على الادعاءات بافتقار معامل ألفا إلى قابلية الاستخدام، وصرحوا بأن معامل ألفا لا يزال يجب أن يحتل مكانًا بارزًا في مجموعة أدوات الإحصاء لدينا، لأنه في ظل افتراضات معينة يعتبرونها شائعة إلى حد ما في التطبيق، يتطابق الثبات من وجهة نظرهم ومعامل ألفا.

ويبدو أن كلا الجانبين في هذا السجال ما بين صلاحية معامل ألفا من عدمه، متفقان على أن استخدام معامل ألفا دائمًا، كما يبدو للكثيرين، غير مبرر (Cho, 2016). ويبدو أن الاختلاف الرئيسي هو أنه في حين أن البعض يفضل التخلص من استخدام معامل ألفا تمامًا (McNeish, 2017) يعتقد آخرون (Raykov & Marcoulides, 2017) أن استخدام معامل ألفا عند استيفاء بعض الشروط له ما يبرره، وفي بعض الأحيان يفضل استخدام بدائلها.

إن الاستخدام الدائم لمعامل ألفا لا يتعلق بقوة المعامل الرياضية وإنما يتعلق أكثر بالناحية التسويقية للمعامل فهو الاسم الأكثر تداولا بين الباحثين، فبالرغم من حصول معامل

Print:(ISSN 1687-2649) Online:(ISSN 2536-9091)

ألفا على مرتبة منخفضة باستمرار في دراسات المقارنات السابقة التي فحصت دقة معاملات الثبات (Kamata, et al., 2003; Osburn, 2000; Revelle & Zinbarg, 2009; Tang & Cui, 2012; van der Ark, van der Palm, & Sijtsma, 2011). فإن الوعي بالاسم وتداوله يفوق أي معاملات ثبات أخرى (Hoekstra et al., 2019)، وقد تمت كتابة الكثير عن التفسيرات الخاطئة أو المعتقدات التي تعتبر غير صحيحة حول معامل ألفا بين الباحثين، على سبيل المثال، فقد ناقش (Cho & Kim, 2015; Cortina, 1993; Sijtsma, 2009) الاعتقاد الخاطئ بأن معامل ألفا يساوي ثبات درجة الاختبار، وفند (Cortina, 1993) الاعتقاد الخاطئ بأن قيمة معامل ألفا مستقلة عن عدد مفردات الاختبار ولا تتأثر بها، وقدم (Cortina, 1993; Nunnally & Bernstein, 1994; Schmitt, 1996; Sijtsma, 2009) الدليل على خطأ الاعتقاد بأن معامل ألفا هو إشارة لأحادية البعد، وناقش (McNeish, 2017; Cho & Kim, 2015; Sijtsma, 2009) الاعتقاد الخاطئ بأن معامل ألفا هو المعامل الأفضل بين معاملات الثبات المختلفة، وعارض (Cortina, 1993; Schmitt, 1996; Cho & Kim, 2015) الاعتقاد الخاطئ بأن هناك مستوى معين من معامل ألفا مرغوب فيه أو مناسب لثبات الاختبار، ودحض (Cho & Kim, 2015) الاعتقاد بأنه إذا أدى حذف المفردة إلى زيادة قيمة معامل ألفا، فإن الاختبار يكون أفضل بدون هذه المفردة.

وتعد المراجعة لتاريخ معامل ألفا مفيدا لتحديد المصدر الأساسي لسوء استخدامها، فلم تنبع شعبيتها من تفوقها التقني. لقد أصبح معامل ألفا المعيار الأكثر استخداما حتى مع وجود معامل أقدم متفوق رياضيًا مثل λ_2 (Guttman, 1945) لعدة أسباب بدت مهمة في وقت Cronbach (1951) ولكنها تم نقدها من وجهة نظر حديثة (Cho & Kim, 2015). أولاً، كان حساب ألفا أبسط. ثانيًا، وضع Cronbach (1951) ألفا كمعامل ثبات للمقياس، في حين كان وصف Guttman (1945) λ_i كتقديرات الحدود الأدنى للثبات، والتي كانت صحيحة رياضيًا لكنها كحدود دنيا لم تحظى بالشعبية. ثالثًا، قدم Cronbach (1951) معامل ألفا على أنه يساوي متوسط قيم الثبات λ_4 (Guttman, 1945) التي تم حسابها لجميع أنصاف التجزئة النصفية الممكنة ووضعها كمعامل ثبات عام. على الرغم من أنه يبدو جيدا بشكل بديهي، إلا أن المتوسط ليس ذا مغزى مثل الحد الأقصى (Osburn,

(2000 أو الحد الأدنى (Revelle, 1979) لقيم λ_4 التي تم الحصول عليها من جميع أنصاف التجزئة النصفية المحتملة. وعلى الرغم من تطوير أساليب أكثر تعقيدًا لحساب الثبات، فلقد تسببت كثرة التعليقات الإيجابية والشعبية الطاغية والاستخدام غير المشروط فيما وصلت إليه الأبحاث اليوم من الاستخدام المفرط لمعامل ألفا، بينما فضل Guttman (1945) استخدام معامل الثبات λ_4 ، وهو معامل ثبات للتجزئة النصفية، وذلك لأنه لا يفترض تكافؤ تاو أو أحادية البعد. وأوصى جتمان بتقسيم مفردات الاختبار بطريقة تزيد من هذا المعامل، لكنها لم تقدم طريقة محددة للقيام بذلك. ومع ذلك، قام الكثير من الباحثين بإجراء تطويرات متعددة في هذا الصدد على مدار الستين عامًا الماضية.

واقترح Bentler (1972) أكبر حد أدنى (glb) أو (ρ+)، وهو أن أكبر ثبات يمكن تحقيقه عن طريق تقليل أثر مصفوفة التغاير (التباينات) مع الحفاظ على المصفوفة موجبة شبه محددة (جميع قيم الجذور الكامنة تكون موجبة). تم تطوير معامل الثبات $\lambda 4$ بواسطة Callender & Osburn (1977) حيث استخدما خوارزمية جبرية لتعيين المفردات إلى نصفين مع عدد مكافئ من المفردات. تم تطوير معامل أوميجا (ω_t) عن طريق ماكدونالدز (Mcdonalds, 1978)، وهو معامل لتقدير الثبات الإجمالي للمقياس، يشبه معامل الثبات λ_6 الذي قام باقتراحه جتمان ويستخدم تفرد الفقرات لتقدير تباين الخطأ في المقياس (Revelle & Condon, 2019).

وتجدر الإشارة أيضًا إلى أن Osburn (2000) قدم دليلًا على أن معامل الثبات (λ_4) الأقصى(Maximized) الذي قام بتطويره كان دقيقًا باستمرار في القياسات المتوازية، والقياسات التي تحقق تكافؤ تاو، والقياسات المتجانسة، ولكن كانت البيانات التي حصل عليها مبنية على المحاكاة التي استخدم فيها بيانات المجتمع وليس بيانات العينة.

وقدمت دراسة Zumbo et al (2007) ودراسة Oliden and Zumbo (2008) ودراسة Gadermann et al (2012) ودراسة Bonanomi et al (2013) دليلاً مفاهيميًا وتجريبيًا وعمليًا لتقدير معاملات الثبات الرتبية لبيانات من نوع ليكرت ، باستخدام معامل ثبات ألفا متعدد الأقسام ($alpha_{polychoric}$) حيث تستخدم مصفوفة الارتباط متعدد الأقسام، وأشارت نتائج دراسة المحاكاة إلى أن معامل ثبات ألفا متعدد الأقسام ($\alpha_{polychoric}$) يقدر الثبات بدقة أكبر من تقدير معامل ألفا كرونباخ عندما تأتي البيانات من

عبارات ذات خيارات استجابة قليلة أو تظهر التواءا (skweness) وعارض ذلك دراسة Chalmers (2018) التي قامت بتقديم أربعة مفاهيم خاطئة مهمة بخصوص معامل ثبات ألفا متعدد الأقسام التي ظهرت في الدراسات السابقة ومناقشتها باستفاضة. وفي دراسات كلا من Kamata et al (2003) و Osburn(2000) تم فحص العلاقة بين ألفا الطبقية ($alpha_{strata}$) وثبات الاختبار باقتراح استخدام معامل ثبات ألفا الطبقية كبديل عن استخدام ألفا كرونباخ وعارضهم في ذلك دراسة Gordon (2007) حيث أثبت رياضيًا أنه عندما يكون هناك تكافؤ متجانس داخل الطبقات أو الاختبارات الفرعية، فإن الاختلاف بين المعاملات هو دالة على تباينات التشبعات داخل الطبقات. عندما تكون العناصر الموجودة داخل كل طبقة مكافئة بشكل أساسي لـ tau، تكون هذه الفروق تساوي صفر، وتكون ألفا الطبقية ودرجة الثبات الحقيقية متساوية، بشرط أن تكون أخطاء القياس غير مرتبطة. إذا كانت أخطاء القياس مرتبطة بشكل إيجابي وكان هناك تكافؤ أساسي لتاو داخل الطبقات، فإن ألفا الطبقية ستبالغ في تقدير الثبات. وأشارت بيانات المجتمع الافتراضية المقدمة في هذه الدراسة إلى أنه في ظل ظروف معينة، يمكن أن تكون ألفا الطبقية أكبر بكثير من ألفا كرونباخ وأقرب إلى الثبات الحقيقي، فمن المستحسن أنه بالنسبة للاختبارات الطبقية، يجب على الباحثين حساب معاملات ألفا الطبقية ومعاملات ألفا كرونباخ بشكل روتيني.

ومن ناحية أخرى ناقش (Bentler, 2007; Geldhof et al., 2014; Raykov, 1997a, 1997b) ارتباط الثبات المركب من الناحية المفاهيمية بألفا كرونباخ من حيث أنها تقيم الثبات عبر نسبة التباين الموضحة بالعناصر مقارنةً بالتباين الكلي للمقياس بأكمله وقدم McDonald (1978, 1999) معامل ثبات أوميجا وتم التوصية باستخدامه بشكل شائع للثبات المركب وتوفرت حساباته في برامج متعددة. وقدم كلا من Viladrich, Angulo-Brunet, & Doval (2017) بناءً على التطورات السيكومترية الحديثة ، دليلاً مفاهيميًا وعمليًا لتقدير ثبات الاتساق الداخلي للقياسات التي تم الحصول عليها كمجموع عنصر أو متوسط.

تم تقديم معامل ثبات الاتساق الداخلي كمنتج ثانوي لنموذج القياس الكامن وراء استجابات الفقرة، وشملت الصيغ المقدمة معاملات ألفا كرونباخ وأوميجا للقياسات أحادية الأبعاد مع مقاييس استجابة الفقرة الكمية و معاملات ثبات أوميجا الرتبية، وثبات ألفا الرتبية

وغير الخطية للقياسات أحادية الأبعاد مع العناصر الرتبية والثنائية ، ومعاملات أوميجا وأوميجا الهرمية للمقاييس وتم تعميم الإجراء على مقاييس المجموع الموزون، والمقاييس متعددة الأبعاد، من ناحية أخرى، افترض (Peters, 2014; Revelle & Zinbarg, 2009) بالنسبة لمعامل ثبات أوميجا الكلية (ω_t) أن المقياس أحادي البعد ويقدر الثبات لمركبات الفقرات على المقياس والذي يشبه معامل ثبات ألفا كرونباخ من ناحية المفهوم حيث تستخدم مواصفات تميل عمومًا إلى أن تكون أكثر تحفظًا وتنتج تقديرات أقرب إلى ألفا كرونباخ.

أراد الباحثون Bentler, 2007; Hancock & Mueller, 2001; Raykov, (2004) استخدام المعلومات الموجودة من تشبعات العوامل لإنشاء مقياس يتم وزنه على النحو الأمثل حيث تساهم كل فقرة بكميات مختلفة من المعلومات في مجموع نقاط المقياس بدلاً من إعطاء كل فقرة نفس الوزن ، فيكون معامل الثبات الأقصى هي مقياس أكثر ملاءمة لثبات المقياس، وفي الاتجاه المعاكس فقد حذر Aguirre-Urreta, Rönkkö, & McIntosh (2019) من تصاعد الدعوات في الأونة الأخيرة لاستبدال معامل ألفا كرونباخ بمعاملات الثبات المستندة إلى نموذج الاستجابة للمفردة الأكثر حداثة في البحث النفسي. في ظل افتراض مقاييس القياس أحادية البعد وأخطاء القياس المستقلة، حيث أوضحوا وجود بديلان رئيسيان هما الثبات المركب ومعامل الثبات الأقصى. من بين هذين، حظيت إحصائية الثبات الأقصى (Maximal Reliability)، بقدر كبير من الاهتمام في السنوات الأخيرة. حيث تم اشتقاق معادلة أوزان الثبات الأقصى باستخدام حجم المجتمع؛ ومع ذلك، فإن سلوكهم في العينة المحدودة لم يتم فحصه على نطاق واسع. وأوضح الباحثون وجود نوعان من التحيز عندما يتم حساب إحصائية الثبات الأقصى من بيانات العينة، هما أن قيمة الثبات الأقصى للعينة هو مقدار متحيز إيجابيًا للثبات الأقصى للمجتمع؛ والثبات الحقيقي للمركبات المكونة بأوزان الثبات الأقصى المحسوبة من بيانات العينة تكون في المتوسط أقل من ثبات المجتمع، ويكون كلا التأثيرين أكثر وضوحًا في العينات الصغيرة.

التعليق على الدراسات السابقة

لقد كان الاستخدام الأمثل لمعاملات الثبات المختلفة مصدرًا للارتباك والخلاف بين الباحثين، حيث تأرجح الباحثون ما بين القبول والرفض لكل معامل ثبات تم إقتراحه، وأظهرت الدراسات السابقة التباين الشديد بين الباحثون وإن اتفقوا جميعا على عدم صلاحية استخدام معامل الثبات ألفا بالطريقة التي يتم استعمالها به الأن، ورأي بعضهم ضرورة تجنب استعماله نهائيا (McNeish, 2017; Cho & Kim, 2015; Sijtsma, 2009) بينما كان من رأي الفريق الأخر بضرورة الابقاء عليه ولكن بعد توفر شروط تطبيقه (Raykov & Marcoulides, 2017)

من ناحية أخرى أوصى Sijtsma (2009) باستخدام أكبر حد أدنى (glb)على الرغم من التحيز الإيجابي الكبير للتقدير في العينات الصغيرة والكبيرة نسبيًا (Ten Berge and Socan, 2004). وإقترح Revelle & Zinbarg (2009) استخدام معامل أوميجا لماكدونالدز (ω_t) وقدموا دليلاً على أن تقديرات الثبات التي يمكن الحصول عليها من معامل أوميجا (ω_t) كانت أعلى من معامل ثبات أكبر حد أدنى (glb) على الرغم من أنه لم تتم دراسة معامل أوميجا بشكل صريح من حيث التحيز، إلا أن نتائج Revelle & Zinbarg (2009) تشير إلى أنه يحتوي أيضًا على تحيز إيجابي كبير.

ولتقديم نتائج علمية قوية، فإن النتائج التجريبية تكون في الغالب هي الأفضل، وبالنظر إلى حقيقة أنه غالبًا ما يُزعم أن معامل ألفا قد أسيء فهمه واستخدامه، ولحسم الجدل الدائر حول استخدام معامل ألفا، تأتي هذه الدراسة لتقدم دليل تجريبي مع تقديم المزيد من المعلومات حول قابلية استخدام معامل ألفا أو بدائلها المختلفة، وتم اختيار دراسة المحاكاة لتوفير التحكم في خصائص البيانات والسماح بمقارنات التحيز النسبي في تقديرات الثبات داخل وعبر أطر القياس. استندت الدلائل الإرشادية لتقدير وقياس التحيز النسبي في معاملات الثبات إلى دراسات كلا من (Geldhof et al., 2014; Muthen et al., 1987; Traxler, 2018).

الطريقة والإجراءات

تصاميم أخذ العينات وشروط البيانات للدراسة باستخدام محاكاة مونت كارلو

تم إجراء محاكاة مونت كارلو متعددة المتغيرات لتوليد ١٠٠٠٠ مجموعة بيانات لكل نوع من أنواع تقدير معاملات الثبات موضع الدراسة في ظل الظروف المختلفة التالية. تشير محاكاة بيانات مونت كارلو إلى توليد عينات من توزيع أساسي محدد للمجتمع استنادا على المعلومات المقدمة حول توزيع النموذج وهيكله لبنتلر(Bentler, 2006).

تم فحص عدد من تقديرات الثبات باستخدام تقنيات محاكاة مونت كارلو، حيث تم إنشاء البيانات لتمثيل النماذج بإطارين للقياس نماذج أحادية البعد وأخرى متعددة الأبعاد عبر ٤ شروط للبيانات (نوع بيانات القياس — خيارات الاستجابة– طول الاختبار – حجم العينة). تمثل هذه الشروط الثماني وأربعون تصميمًا متقاطعًا بالكامل (٢ × ٢ × ٢ × ٦) من خلال إطارين أحدهما أحادي البعد والأخر متعدد الأبعاد وكلاهما يتبع التوزيع الطبيعي وتفصيلها بجدول (١).

تم أخذ القرارات المتعلقة بشروط البيانات مثل (أحادية البعد للسمة الكامنة أو تعددها)، ونوعية البيانات (مكافئة لتاو – متجانسة)، والعينات التي تم سحبها (٢٠–٥٠–١٠٠–٢٠٠–٥٠٠–١٠٠٠)، ونوعين لعناصر الاختبار أحدهما قصير والأخر طويل (٦ أبعاد – ١٢ بعد)، وعدد خيارات الاستجابة (ثنائية الاستجابة ومتعددة الاستجابة)، وذلك في ضوء الدراسات السابقة التي أجراها كل من (Gadermann et al., 2012; Lozano et al., 2008; Muthén & Muthén, 2000; Zumbo et al., 2007).

جدول (١)
تصاميم أخذ العينات وشروط البيانات للدراسة باستخدام محاكاة مونت كارلو

التوزيع	البعد	نوع البيانات	طول الاختبار	خيارات الاستجابة	حجم العينة
التوزيع الطبيعي	أحادي البعد	مكافئة لتاو	قصير	٢	٢٠-٥٠-١٠٠-٢٠٠-٥٠٠-١٠٠٠
				٥	٢٠-٥٠-١٠٠-٢٠٠-٥٠٠-١٠٠٠
			طويل	٢	٢٠-٥٠-١٠٠-٢٠٠-٥٠٠-١٠٠٠
				٥	٢٠-٥٠-١٠٠-٢٠٠-٥٠٠-١٠٠٠
		متجانسة	قصير	٢	٢٠-٥٠-١٠٠-٢٠٠-٥٠٠-١٠٠٠
				٥	٢٠-٥٠-١٠٠-٢٠٠-٥٠٠-١٠٠٠
			طويل	٢	٢٠-٥٠-١٠٠-٢٠٠-٥٠٠-١٠٠٠
				٥	٢٠-٥٠-١٠٠-٢٠٠-٥٠٠-١٠٠٠
	متعدد الأبعاد	مكافئة لتاو	قصير	٢	٢٠-٥٠-١٠٠-٢٠٠-٥٠٠-١٠٠٠
				٥	٢٠-٥٠-١٠٠-٢٠٠-٥٠٠-١٠٠٠
			طويل	٢	٢٠-٥٠-١٠٠-٢٠٠-٥٠٠-١٠٠٠
				٥	٢٠-٥٠-١٠٠-٢٠٠-٥٠٠-١٠٠٠
		متجانسة	قصير	٢	٢٠-٥٠-١٠٠-٢٠٠-٥٠٠-١٠٠٠
				٥	٢٠-٥٠-١٠٠-٢٠٠-٥٠٠-١٠٠٠
			طويل	٢	٢٠-٥٠-١٠٠-٢٠٠-٥٠٠-١٠٠٠
				٥	٢٠-٥٠-١٠٠-٢٠٠-٥٠٠-١٠٠٠

توليد البيانات

تم تقييم التأثيرات والتفاعلات الرئيسية المتعلقة بتقديرات الثبات والتي تم إنشاؤها باستخدام برنامج R حزمة (R-Psych) و حزمة (sem) وحزمة (lambda4) وحزمة (lavaan) وحزمة (GPArotation) وتم توليد البيانات بحيث تتبع التوزيع الطبيعي، وبالنسبة للاشتراطات الخاصة بالنموذج أحادي البعد، تم توليد بيانات مكافئة لتاو (تساوي التشبعات واختلاف تباين أخطاء القياس)، حيث تم تعيين تشبع المتغيرات على عامل واحد عند (٠.٦) وتم تعيين التباينات عند $(٠.٦)^٢$، $(٠.٧)^٢$، $(٠.٨)^٢$، $(٠.٩)^٢$، وتم استخدام البيانات لتوليد مصفوفة الارتباط ومصفوفة التغاير لتمثيل بنية البيانات المكافئة لتاو. ولتوليد بيانات متجانسة (اختلاف التشبعات واختلاف تباين أخطاء القياس) أحادية البعد تم تعيين

تشبع المتغيرات على عامل واحد لتكون (٠.٥، ٠.٦، ٠.٧، ٠.٨) وتم تعيين التباينات عند (٠.٦)٢، (٠.٧)٢، (٠.٨)٢، (٠.٩)٢، تم استخدام البيانات لتوليد مصفوفة التغاير لتمثيل بنية البيانات المتجانسة، وبالنسبة للنموذج متعدد الأبعاد تم توليد نموذج ثلاثي العوامل لبيانات مكافئة لتاو، حيث تم تعيين تشبع جميع المتغيرات على العوامل عند (٠.٦) وتم تعيين التباينات عند (٠.٦)٢، (٠.٧)٢، (٠.٨)٢، (٠.٩)٢، وتم تثبيت الارتباطات بين المتغيرات الكامنة عند (٠.٣) واستخدمت البيانات لتوليد مصفوفة الارتباط ومصفوفة التغاير لتمثيل بنية البيانات المكافئة لتاو. ولتوليد بيانات متجانسة متعددة الأبعاد تم تعيين تشبع المتغيرات على العوامل لتكون (٠.٥، ٠.٦، ٠.٧، ٠.٨) على كل عامل، وتم تعيين التباينات عند (٠.٦)٢، (٠.٧)٢، (٠.٨)٢، (٠.٩)٢، وتم تثبيت الارتباطات بين المتغيرات الكامنة عند (٠.٣) واستخدمت البيانات لتوليد مصفوفة الارتباط ومصفوفة التغاير لتمثيل بنية البيانات المتجانسة.

تحليل البيانات

للإجابة على الأسئلة البحثية، وتحديد أفضل معامل ثبات للاستخدام من قبل الباحثين لكل حالة من حالات المحاكاة عبر جميع شروط البيانات وتصميمات العينات وأطر القياس، تم إنشاء البيانات التي تمثل ٤٨ حالة لمعاملات الثبات موضع الدراسة في النموذج أحادي البعد و ٤٨ حالة لمعاملات الثبات في النموذج متعدد الأبعاد باستخدام برنامج R بحزمه المختلفة حيث تم حساب معاملات الثبات باستخدام حزمة (psych) التي طورها Revelle (2020) لحساب معامل ثبات أوميجا الكلية (ωt) ومعامل ثبات أوميجا الهرمية (ωh) ، وحزمة (psychTools) التي طورها Revelle (2020) لحساب معاملات ثبات ألفا (α) ، والحدود الدنيا لجتمان (λ_i) وأكبر حد أدنى (glb) وحزمة (lambda4) التي طورها Hunt (2013) لحساب معامل ثبات بيتا (β) وحزمة (semTools) التي طورها Jorgensen et al (2020) لحساب معامل الثبات الأقصى(H) وحزمة (conogive) التي طورها Moss (2020) لحساب معامل ثبات ألفا الرتبية ($\alpha_{polychoric}$) وحزمة(sirt) التي طورها Robitzsch (2020b) لحساب معامل ثبات ألفا الطبقية (α_{strata}) وذلك لكل حالة بيانات. تم استخدام مصفوفة التغاير ومصفوفة الارتباط بعد توليد البيانات بالشروط الموضوعة للمحاكاة لحساب معاملات الثبات موضع الدراسة. وبالنسبة لحساب معامل ثبات ألفا الرتبية

للأقسام المتعددة تم توليد مصفوفة الارتباط للأقسام المتعددة (Polychoric Correlation Matrix) من البيانات المتحصل عليها ومن ثم حساب معامل الثبات الرتبي. ولحساب معامل ألفا الطبقية تم توليد البيانات الخام من ثلاث طبقات للنموذج الأحادي البعد والنموذج متعدد الأبعاد، ثم تم حساب معامل الثبات للطبقات المفترضة وللنموذج الكلي. وبمجرد حساب جميع معاملات الثبات وتصديرها إلى Excel، تم عمل المقارنة بينها للوقوف على أفضل المعاملات استخداما في ظروف المحاكاة المختلفة، وتم فحص النتائج للبيانات المولدة حسب ما سيتم عرضه لاحقا، ثم تم حساب التحيز النسبي باستخدام المعادلة الموضحة أدناه (معادلة ٤٠) حيث كانت معاملات الثبات التي يتم المقارنة بها هي معاملات ثبات المجتمع حيث يعتبر التحيز النسبي أقل من أو يساوي ١٠٪ مقبولاً (Traxler, 2018)، و تضمن التحليل الإضافي المتعلق بالتحيز النسبي تقييم ما إذا كانت تقديرات الثبات مبالغ فيها (Over Estimated) أو تم التقليل من شأنها (Under Estimated) وذلك استنادا إلى دراسة (Geldhof et al., 2014)

%التحيز النسبي = (ثبات المجتمع)/(ثبات المجتمع – ثبات العينة) (40)

النتائج

- **معاملات الثبات للبيانات ذات الاشتراطات (اختبار أحادي البعد - بيانات مكافئة لتاو - اختبار قصير -خيارات استجابة ثنائية)**

بالرجوع إلى جدول (٢) لقيم معاملات الثبات للبيانات وجدول (٣) للنسب المئوية للتحيز النسبي لمعاملات الثبات نجد أن معامل ثبات أكبر حد أدنى (glb) في حالة العينة (٢٠) هو أعلى معامل ثبات، وقد أظهر نسبة تحيز ٥% في الاتجاه السالب أي أنها مبالغ فيها، ولكن كانت القيمة أقل من ١٠% مما يعتبر نسبة مقبولة يمكن تجاهلها، وفي حالة العينة (٥٠) نجد أن معامل ثبات أوميجا الهرمية التقاربي (ω_h) هو الأعلى ونسبة التحيز النسبي لها كانت صفر، بينما في حالة العينة (١٠٠) والعينة (٢٠٠) نجد أن معامل ثبات ألفا للأقسام المتعددة (α_{poly}) هو الأعلى بنسب تحيز ٤% و١% على الترتيب مما يعني نسب تحيز مقبولة تقلل من شأن تقديرات الثبات ، وفي حالة العينات الكبيرة (٥٠٠، ١٠٠٠) نجد أن

Print:(ISSN 1687-2649) Online:(ISSN 2536-9091)

معامل ثبات أوميجا الهرمية التقاربي (ω_h) يعطي أعلى تقدير للثبات بنسبة تحيز ثابتة ٥% في اتجاه المبالغة نوعا ما في معامل الثبات ولكن بنسبة مقبولة أقل من ١٠%.

جدول (٢)

بيانات معاملات الثبات لاختبار قصير بخيارات استجابة ثنائية لبيانات مكافئة لتاو احادية البعد

α_{strata}	α_{poly}	H	β	glb	ω_h	ω_t	λ_6	λ_5	λ_4 (max)	λ_3 (α)	λ_2	λ_1	حجم العينة
-	-	0.78	0.77	0.77	0.77	0.79	0.74	0.74	0.77	0.77	0.77	0.64	المجتمع
0.44	0.61	-	0.07	0.81	0.68	0.71	0.59	0.52	0.76	0.45	0.54	0.38	20
0.45	0.57	0.59	0.21	0.56	0.77	0.61	0.45	0.46	0.60	0.43	0.48	0.36	50
0.60	0.74	0.65	0.48	0.71	0.59	0.71	0.58	0.60	0.65	0.60	0.61	0.50	100
0.60	0.76	0.63	0.54	0.69	0.69	0.75	0.59	0.60	0.69	0.62	0.62	0.52	200
0.65	0.78	0.66	0.63	0.68	0.81	0.76	0.60	0.63	0.68	0.64	0.64	0.53	500
0.61	0.74	0.61	0.57	0.63	0.81	0.72	0.56	0.58	0.63	0.60	0.60	0.50	1000

جدول (٣)

النسب المئوية للتحيز لمعاملات ثبات اختبار قصير بخيارات استجابة ثنائية لبيانات مكافئة لتاو احادية البعد

حجم العينة							
1000	500	200	100	50	20		معاملات الثبات
22%	17%	19%	22%	44%	41%	λ_1	
22%	17%	19%	21%	38%	30%	λ_2	
22%	17%	19%	22%	44%	42%	λ_3 (α)	
18%	12%	10%	16%	22%	1%	λ_4 (max)	
22%	15%	19%	19%	38%	30%	λ_5	
24%	19%	20%	22%	39%	20%	λ_6	
9%	4%	5%	10%	23%	10%	ω_t	
-5%	-5%	10%	23%	0%	12%	ω_h	
18%	12%	10%	8%	27%	-5%	glb	
26%	18%	30%	38%	73%	91%	β	
22%	15%	19%	17%	24%	100%	H	
4%	-1%	1%	4%	26%	21%	α_{poly}	
21%	16%	22%	22%	42%	43%	α_{strata}	

- **معاملات الثبات للبيانات ذات الاشتراطات (اختبار أحادي البعد - بيانات مكافئة لتاو - اختبار قصير - استجابة متعددة)**

يتضح من جدول (٤) وجدول (٥) أنه في حالة العينات (٢٠ ، ٥٠ ، ١٠٠، ٢٠٠) يكون معامل ثبات أوميجا الكلية (ω_t) هو الأعلى بنسب تحيز ١% ، ٤% ، ٩%، و ٤% على الترتيب وجميعها تتجه إلى المبالغة في معامل الثبات ولكن بنسب تقل عن ١٠% أي

Print:(ISSN 1687-2649) Online:(ISSN 2536-9091)

مقبولة ويمكن تجاهلها، ويشترك معه معامل ثبات أكبر حد أدنى (glb) في حالة العينات الصغيرة (٢٠ ، ٥٠) ولكن بنسبة تحيز ٦% وهي أيضا مقبولة ولكن بنسبة أعلى قليلا مما يعطي أفضلية لمعامل ثبات أوميجا الكلية، بينما في حالة العينات الكبيرة (٥٠٠، ١٠٠٠) نجد أن معامل ثبات أوميجا الهرمية التقاربي (ω_h) هو الأعلى في القيم ولكن نسب التحيز له كانت ١٧% و١٣% على الترتيب وفي الاتجاه السالب مما يعني أنه تحيز يزيد عن النسبة ١٠% أي أنه تحيز نسبي في اتجاه التضخيم من معامل الثبات عن النسبة المقبولة مما يعني رفضه وبالرجوع إلى جدول (٤) يتضح أن أعلى معامل ثبات غير متحيز هو معامل ثبات أوميجا الكلية (ω_t) للعينات (٥٠٠ ، ١٠٠٠) .

جدول (٤)

بيانات معاملات الثبات لاختبار قصير بخيارات استجابة متعددة لبيانات مكافئة لتاو احادية البعد

αstrata	αpoly	H	β	glb	ω_h	ω_t	λ_6	λ_5	λ_4 (max)	λ_3 (α)	λ_2	λ_1	حجم العينة
0.77	0.77	0.78	0.77	0.77	0.77	0.79	0.74	0.74	0.77	0.77	0.77	0.64	المجتمع
0.66	0.64	0.01	0.40	0.78	0.61	0.78	0.68	0.65	0.76	0.62	0.65	0.52	20
0.70	0.73	0.72	0.61	0.82	0.51	0.82	0.69	0.69	0.78	0.70	0.70	0.58	50
0.79	0.82	0.80	0.76	0.83	0.76	0.86	0.77	0.77	0.83	0.79	0.79	0.66	100
0.73	0.75	0.74	0.68	0.72	0.79	0.82	0.69	0.71	0.75	0.73	0.73	0.60	200
0.72	0.75	0.73	0.68	0.68	0.90	0.75	0.69	0.70	0.73	0.72	0.72	0.60	500
0.77	0.79	0.77	0.75	0.75	0.87	0.81	0.73	0.74	0.78	0.77	0.77	0.64	1000

جدول (٥)

النسب المئوية للتحيز لمعاملات ثبات اختبار قصير بخيارات استجابة ثنائية لبيانات مكافئة لتاو احادية البعد

حجم العينة							معاملات الثبات
1000	500	200	100	50	20		
0%	6%	6%	-3%	9%	19%	λ1	
0%	6%	5%	-3%	9%	16%	λ2	
0%	6%	5%	-3%	9%	19%	λ3 (α)	
-1%	5%	3%	-8%	-1%	1%	λ4 (max)	
0%	5%	4%	-4%	7%	12%	λ5	
1%	7%	7%	-4%	7%	8%	λ6	
-3%	5%	-4%	-9%	-4%	1%	ωt	
-13%	-17%	-3%	1%	34%	21%	ωh	
3%	12%	6%	-8%	-6%	-1%	glb	
3%	12%	12%	1%	21%	48%	β	
1%	6%	5%	-3%	8%	99%	H	
-3%	3%	3%	-6%	5%	17%	αpoly	
0%	6%	5%	-3%	9%	14%	αstrata	

Print:(ISSN 1687-2649) Online:(ISSN 2536-9091)

- **معاملات الثبات للبيانات ذات الاشتراطات (اختبار أحادي البعد - بيانات مكافئة لتاو - اختبار طويل -خيارات استجابة ثنائية)**

يتضح من جدول (٦) وجدول (٧) أن معامل ثبات أكبر حد أدنى (glb) في حالة العينات الصغيرة (٢٠ ، ٥٠) هو أعلى معامل ثبات، بنسب تحيز ٧% في اتجاه المبالغة من قيمة معامل الثبات للعينة ٢٠، ولكن بنسب تقل عن ١٠% أي مقبولة ويمكن تجاهلها، وغير متحيزة على الاطلاق للعينة ٥٠ ،وفي حالة العينات (١٠٠ ، ٢٠٠) نجد أن معامل ثبات ألفا للأقسام المتعددة (α_{poly}) هو الأعلى، بنسب تحيز ٣% في اتجاه المبالغة من قيمة معامل الثبات للعينة ١٠٠، ولكن بنسب تقل عن ١٠% أي مقبولة ويمكن تجاهلها، وغير متحيزة على الاطلاق للعينة ٢٠٠، بينما في حالة العينات الكبيرة (٥٠٠، ١٠٠٠) نجد أن معامل ثبات أوميجا الهرمية التقاربي (ω_h) هو الأعلى بنسب تحيز ٣% و ٥% في اتجاه المبالغة من قيمة معامل الثبات على الترتيب، ولكن بنسب تقل عن ١٠% أي مقبولة ويمكن تجاهلها.

جدول (٦)

معاملات الثبات لاختبار طويل بخيارات ثنائية لبيانات مكافئة لتاو احادية البعد

حجم العينة	λ_1	λ_2	λ_3 (α)	λ_4 (max)	λ_5	λ_6	ω_t	ω_h	glb	β	H	α_{poly}	α_{strata}
المجتمع	0.81	0.88	0.88	0.88	0.85	0.87	0.90	0.88	0.89	0.88	0.89	-	-
20	0.72	0.81	0.78	0.92	0.78	0.93	0.81	0.69	0.95	0.49	0.86	0.84	0.77
50	0.71	0.79	0.77	0.87	0.77	0.82	0.83	0.66	0.89	0.65	0.83	0.86	0.77
100	0.76	0.83	0.83	0.89	0.81	0.84	0.86	0.89	0.88	0.76	0.85	0.91	0.83
200	0.72	0.79	0.79	0.84	0.77	0.79	0.82	0.66	0.83	0.71	0.80	0.88	0.78
500	0.74	0.81	0.81	0.84	0.78	0.80	0.82	0.91	0.84	0.76	0.81	0.89	0.80
1000	0.73	0.79	0.79	0.82	0.77	0.78	0.80	0.92	0.78	0.76	0.80	0.88	0.79

Print:(ISSN 1687-2649) Online:(ISSN 2536-9091)

جدول (٧)

النسب المئوية للتحيز لمعاملات الثبات

معاملات الثبات		حجم العينة					
		20	50	100	200	500	1000
	λ_1	11%	12%	6%	11%	9%	10%
	λ_2	8%	10%	6%	10%	8%	10%
	λ_3 (α)	11%	13%	6%	10%	8%	10%
	λ_4 (max)	-5%	1%	-1%	5%	5%	7%
	λ_5	8%	9%	5%	9%	8%	9%
	λ_6	-7%	6%	3%	9%	8%	10%
	ω_t	10%	8%	4%	9%	9%	11%
	ω_h	22%	25%	-1%	25%	-3%	-5%
	glb	-7%	0%	1%	7%	6%	12%
	β	44%	26%	14%	19%	14%	14%
	H	3%	7%	4%	10%	9%	10%
	α_{poly}	5%	2%	-3%	0%	-1%	0%
	α_{strata}	13%	13%	6%	11%	9%	10%

- **معاملات الثبات للبيانات ذات الاشتراطات (اختبار أحادي البعد - بيانات مكافئة لتاو - اختبار طويل -خيارات استجابة متعددة)**

من جدول (٨) وجدول (٩) أن معامل ثبات الحد الأدنى لجتمان (λ_4) الأقصى في حالة العينات (٢٠، ٥٠) هو أعلى معامل ثبات، بنسب تحيز ١% و ٠% على الترتيب، ونجد أن معامل ثبات (glb) يعطي نفس قيمة الحد الأدنى الرابع لجتمان ولكن بنسبة تحيز ١% في اتجاه التقليل من قيمة معامل الثبات، وفي حالة العينة (١٠٠) كان هو الأعلى، بنسبة تحيز غير مؤثرة ٢% ، بينما في حالة العينة (٢٠٠) نجد أن معامل ثبات الحد الأدنى لجتمان (λ_4) الأقصى هو الأعلى، بنسبة تحيز ٢% غير مؤثرة في اتجاه المبالغة في قيمة المعامل، وفي حالة العينات الكبيرة (٥٠٠، ١٠٠٠) نجد أن معامل ثبات أوميجا الهرمية التقاربي (ω_h) هو الأعلى، بنسب تحيز ٢% و ٥% على الترتيب في اتجاه المبالغة من قيمة معامل الثبات.

Print:(ISSN 1687-2649) Online:(ISSN 2536-9091)

جدول (٨)

بيانات معاملات الثبات لاختبار طويل بخيارات استجابة متعددة لبيانات مكافئة لتاو احادية البعد

αstrata	αpoly	H	β	glb	ω_h	ω_t	λ_6	λ_5	λ_4(m)	$\lambda_3(\alpha)$	λ_2	λ_1	حجم العينة
-	-	0.89	0.88	0.89	0.88	0.90	0.87	0.85	0.88	0.88	0.88	0.81	المجتمع
0.89	0.84	0.92	0.76	0.96	0.74	0.93	0.94	0.87	0.97	0.89	0.90	0.82	20
0.89	0.91	0.91	0.84	0.94	0.72	0.92	0.91	0.87	0.94	0.89	0.90	0.82	50
0.87	0.88	0.88	0.78	0.92	0.66	0.89	0.88	0.84	0.91	0.86	0.87	0.79	100
0.88	0.89	0.88	0.82	0.87	0.87	0.89	0.87	0.85	0.90	0.87	0.88	0.80	200
0.88	0.89	0.88	0.85	0.89	0.94	0.88	0.87	0.85	0.89	0.88	0.88	0.80	500
0.87	0.88	0.88	0.85	0.86	0.93	0.88	0.86	0.84	0.88	0.87	0.87	0.80	1000

جدول (٩)

النسب المئوية للتحيز لمعاملات ثبات اختبار طويل بخيارات استجابة متعددة لبيانات مكافئة لتاو احادية البعد

حجم العينة							معاملات الثبات
1000	500	200	100	50	20		
-8%	-2%	-10%	-1%	-2%	-1%	λ1	
-5%	-2%	-7%	-1%	-2%	-1%	λ2	
-1%	1%	-3%	2%	1%	2%	λ3 (α)	
0%	0%	-2%	1%	0%	1%	λ4 (max)	
0%	0%	-1%	0%	0%	1%	λ5	
1%	1%	0%	1%	1%	1%	λ6	
-8%	-2%	-10%	-1%	-2%	-1%	ωt	
-5%	-2%	-7%	-1%	-2%	-1%	ωh	
-1%	1%	-3%	2%	1%	2%	glb	
0%	0%	-2%	1%	0%	1%	β	
0%	0%	-1%	0%	0%	1%	H	
1%	1%	0%	1%	1%	1%	αpoly	
-8%	-2%	-10%	-1%	-2%	-1%	αstrata	

- **معاملات الثبات للبيانات ذات الاشتراطات (اختبار أحادي البعد - بيانات متجانسة -اختبار قصير -خيارات استجابة ثنائية)**

يتضح من جدول (١٠) وجدول (١١) أن معامل ثبات أكبر حد أدنى (glb) في حالة العينة (٢٠) هو أعلى معامل ثبات، بنسبة تحيز ٤% التي يمكن تجاهلها في اتجاه المبالغة في تقدير الثبات، بينما في حالة العينة (٥٠) نجد أن أعلى معامل ثبات هو معامل ثبات أوميجا الكلية (ω_t) بنسبة تحيز ٦% في اتجاه التقليل من تقدير الثبات، وفي حالة العينات (١٠٠ ، ٢٠٠، ٥٠٠، ١٠٠٠) نجد أن معامل ثبات ألفا للأقسام المتعددة (α_{poly}) هو الأعلى بتحيز نسبي غير مؤثر يقل جميعه عن ١٠ % وهو على الترتيب ٥%، ٠%، ٢%، ١% ، وبالنسبة لحالة العينة (٥٠٠) نجد أن معامل ثبات أوميجا الهرمية التقاربي (ω_h) يشترك مع معامل

ثبات ألفا للأقسام المتعددة في كونه يعطي أعلى تقدير للثبات بنفس نسبة التحيز ٢% في اتجاه المبالغة من تقدير الثبات.

جدول (١٠)

بيانات معاملات الثبات لاختبار قصير بخيارات استجابة ثنائية لبيانات متجانسة احادية البعد

αstrata	αpoly	H	β	glb	ω_h	ω_t	λ_6	λ_5	λ_4(m)	λ_3(α)	λ_2	λ_1	حجم العينة
-	-	0.82	0.82	0.82	0.82	0.83	0.79	0.80	0.82	0.82	0.82	0.68	المجتمع
0.67	0.74	0.92	0.49	0.85	0.68	0.83	0.74	0.65	0.79	0.62	0.67	0.49	20
0.55	0.74	0.65	0.46	0.74	0.36	0.78	0.60	0.60	0.74	0.60	0.61	0.50	50
0.78	0.86	0.77	0.67	0.82	0.69	0.84	0.73	0.73	0.82	0.75	0.75	0.62	100
0.69	0.82	0.70	0.65	0.73	0.79	0.80	0.65	0.67	0.72	0.69	0.69	0.57	200
0.70	0.84	0.71	0.68	0.72	0.84	0.78	0.67	0.69	0.73	0.71	0.71	0.59	500
0.66	0.81	0.68	0.65	0.68	0.74	0.77	0.64	0.65	0.70	0.68	0.68	0.56	1000

جدول (١١)

النسب المئوية للتحيز لمعاملات ثبات اختبار قصير بخيارات استجابة ثنائية لبيانات متجانسة احادية البعد

حجم العينة							
1000	500	200	100	50	20		معاملات الثبات
18%	13%	16%	9%	26%	28%	λ1	
17%	13%	16%	9%	26%	18%	λ2	
17%	13%	16%	9%	27%	24%	λ3 (α)	
15%	11%	12%	0%	10%	4%	λ4 (max)	
19%	14%	16%	9%	25%	19%	λ5	
19%	15%	18%	8%	24%	6%	λ6	
7%	6%	4%	-1%	6%	0%	ωt	
10%	-2%	4%	16%	56%	17%	ωh	
17%	12%	11%	0%	10%	-4%	glb	
21%	17%	21%	18%	44%	40%	β	
17%	13%	15%	6%	21%	-12%	H	
1%	-2%	0%	-5%	10%	10%	αpoly	
20%	15%	16%	5%	33%	18%	αstrata	

- **معاملات الثبات للبيانات ذات الاشتراطات (اختبار أحادي البعد - بيانات متجانسة -اختبار قصير -خيارات استجابة متعددة)**

من جدول (١٢) وجدول (١٣) نجد أن أعلى معامل ثبات هو معامل ثبات أوميجا الكلية (ω_t) في حالة العينات (٢٠، ٥٠، ١٠٠) بنسب تحيز ٦% في اتجاه التقليل من تقدير الثبات، و ٥%، ٧% على الترتيب في اتجاه المبالغة من تقدير الثبات، ونجد أن معامل ثبات أوميجا الهرمية التقاربي (ω_h) هو الأعلى في حالة العينات (٢٠٠، ٥٠٠، ١٠٠٠) وكان التحيز النسبي للعينات هو ١٠%، ١٣%، ٩% على الترتيب وجميعهم في اتجاه المبالغة في تقدير

Print:(ISSN 1687-2649) Online:(ISSN 2536-9091)

الثبات وبالنسبة للعينة ٥٠٠ كان التحيز النسبي أعلى من حد عدم التأثير ١٠% وبالنظر لجدول ١٢ أن أعلى قيمة تقدير للثبات غير متحيزة هي قيمة معامل ثبات ألفا للأقسام المتعددة (α_{poly}) بنسبة تحيز ٢% في اتجاه التقليل من تقدير الثبات.

جدول (١٢)

بيانات معاملات الثبات لاختبار قصير بخيارات استجابة متعددة لبيانات متجانسة احادية البعد

αstrata	αpoly	H	β	glb	ω_h	ω_t	λ_6	λ_5	λ_4(m)	λ_3(α)	λ_2	λ_1	حجم العينة
-	-	0.82	0.82	0.82	0.82	0.83	0.79	0.80	0.82	0.82	0.82	0.68	المجتمع
0.67	0.62	0.65	0.48	0.76	0.49	0.78	0.62	0.60	0.75	0.59	0.62	0.49	20
0.73	0.79	0.80	0.62	0.86	0.59	0.87	0.77	0.75	0.83	0.76	0.77	0.63	50
0.81	0.84	0.82	0.78	0.85	0.85	0.89	0.80	0.79	0.85	0.82	0.82	0.68	100
0.77	0.80	0.78	0.74	0.80	0.90	0.83	0.74	0.75	0.80	0.77	0.77	0.64	200
0.78	0.80	0.78	0.76	0.78	0.93	0.79	0.75	0.75	0.79	0.78	0.78	0.65	500
0.82	0.84	0.81	0.79	0.83	0.89	0.88	0.78	0.78	0.83	0.81	0.81	0.68	1000

جدول (١٣)

النسب المئوية للتحيز لمعاملات ثبات اختبار قصير بخيارات استجابة متعددة لبيانات متجانسة احادية البعد

حجم العينة							معاملات الثبات
1000	500	200	100	50	20		
0%	4%	6%	0%	7%	28%	λ1	
1%	5%	6%	0%	6%	24%	λ2	
1%	5%	6%	0%	7%	28%	λ3 (α)	
-1%	4%	2%	-4%	-1%	9%	λ4 (max)	
3%	6%	6%	1%	6%	25%	λ5	
1%	5%	6%	-1%	3%	22%	λ6	
-6%	5%	0%	-7%	-5%	6%	ωt	
-9%	-13%	-10%	-4%	28%	40%	ωh	
-1%	5%	2%	-4%	-5%	7%	glb	
4%	7%	10%	5%	24%	41%	β	
1%	5%	5%	0%	2%	21%	H	
-2%	2%	2%	-2%	4%	24%	αpoly	
0%	5%	6%	1%	11%	18%	αstrata	

- **معاملات الثبات للبيانات ذات الاشتراطات (اختبار أحادي البعد - بيانات متجانسة -اختبار طويل -خيارات استجابة ثنائية)**

يتضح من جدول (١٤) وجدول (١٥) أن معامل ثبات أكبر حد أدنى (glb) في حالة العينات (٢٠، ٥٠) هو أعلى معامل ثبات، بتحيز نسبي غير مؤثر ٨%، ٢% على التوالي وفي اتجاه المبالغة في قيمة تقدير الثبات، ويشترك معه معامل ثبات الحد الأدنى لجتمان (λ_4) الأقصى في حالة العينة (٥٠) بنفس قيمة التحيز النسبي وفي نفس الاتجاه، وفي حالة

Print:(ISSN 1687-2649) Online:(ISSN 2536-9091)

العينات (١٠٠، ٢٠٠، ٥٠٠) نجد أن معامل ثبات ألفا للأقسام المتعددة (α_{poly}) هو الأعلى بنسب تحيز غير مؤثرة ٢%، ١%، ٠% على الترتيب وفي اتجاه المبالغة من تقدير الثبات للعينة (١٠٠) وفي اتجاه التقليل من قيمة تقدير الثبات للعينة (٢٠٠)، بينما حالة العينة (١٠٠٠) نجد أن معامل ثبات أوميجا الهرمية التقاربي (ω_h) يعطي أعلى تقدير للثبات بتحيز نسبي ٤% في اتجاه التقليل من الثبات.

جدول (١٤)

بيانات معاملات الثبات لاختبار طويل بخيارات استجابة ثنائية لبيانات متجانسة احادية البعد

αstrata	αpoly	H	β	glb	ω_h	ω_t	λ_6	λ_5	λ_4 (max)	λ_3 (α)	λ_2	λ_1	حجم العينة
-	-	0.90	0.90	0.90	0.90	0.91	0.89	0.87	0.90	0.90	0.90	0.82	المجتمع
0.80	0.86	0.85	0.59	0.97	0.41	0.86	0.95	0.80	0.96	0.81	0.83	0.74	20
0.83	0.90	0.84	0.70	0.92	0.75	0.86	0.86	0.81	0.92	0.82	0.83	0.75	50
0.85	0.92	0.86	0.78	0.90	0.80	0.87	0.86	0.83	0.91	0.85	0.85	0.78	100
0.81	0.89	0.81	0.73	0.87	0.70	0.84	0.81	0.79	0.86	0.81	0.81	0.74	200
0.82	0.90	0.82	0.78	0.86	0.83	0.85	0.81	0.80	0.86	0.82	0.82	0.75	500
0.81	0.90	0.81	0.79	0.84	0.94	0.84	0.80	0.79	0.84	0.81	0.81	0.75	1000

جدول (١٥)

النسب المئوية للتحيز لمعاملات ثبات اختبار طويل بخيارات استجابة ثنائية لبيانات متجانسة احادية البعد

حجم العينة							
1000	500	200	100	50	20		معاملات الثبات
9%	9%	10%	5%	9%	10%	λ1	
10%	9%	10%	6%	8%	8%	λ2	
10%	9%	10%	6%	9%	10%	λ3 (α)	
7%	4%	4%	-1%	-2%	-7%	λ4 (max)	
9%	8%	9%	5%	7%	8%	λ5	
10%	9%	9%	3%	3%	-7%	λ6	
8%	7%	8%	4%	5%	5%	ωt	
-4%	8%	22%	11%	17%	54%	ωh	
7%	4%	3%	0%	-2%	-8%	glb	
12%	13%	19%	13%	22%	34%	β	
10%	9%	10%	4%	7%	6%	H	
0%	0%	1%	-2%	0%	4%	αpoly	
10%	9%	10%	6%	8%	11%	αstrata	

- **معاملات الثبات للبيانات ذات الاشتراطات (اختبار أحادي البعد - بيانات متجانسة -اختبار طويل -خيارات استجابة متعددة)**

يتضح من جدول (١٦) وجدول (١٧) أن معامل ثبات الحد الأدنى لجتمان (λ_4) الأقصى هو أعلى معامل ثبات في حالة العينات (٢٠، ٥٠، ١٠٠، ٢٠٠) وبقيم تحيز نسبي غير مؤثر

٨%، ٧%، ٢%، ٢%، وجميعها في اتجاه المبالغة في تقدير الثبات، ويشترك معه معامل ثبات أكبر حد أدنى (glb) في حالة العينة (٢٠) بنفس قيم الثبات والتحيز النسبي، بينما في حالة العينات (٥٠٠، ١٠٠٠) نجد أن معامل ثبات أوميجا الهرمية التقاربي (ω_h) يعطي أعلى تقدير للثبات ويقيم تحيز نسبي غير مؤثرة ٦%، ٨% في اتجاه المبالغة من تقدير الثبات.

جدول (١٦)

بيانات معاملات الثبات لاختبار طويل بخيارات استجابة متعددة لبيانات متجانسة احادية البعد

αstrata	αpoly	H	β	glb	ω_h	ω_t	λ_6	λ_5	λ_4 (max)	λ_3 (α)	λ_2	λ_1	حجم العينة
-	-	0.90	0.90	0.90	0.90	0.91	0.89	0.87	0.90	0.90	0.90	0.82	المجتمع
0.90	0.86	0.93	0.79	0.97	0.63	0.93	0.94	0.88	0.97	0.90	0.90	0.83	20
0.91	0.92	0.92	0.84	0.95	0.77	0.93	0.93	0.88	0.96	0.91	0.91	0.84	50
0.87	0.90	0.89	0.80	0.91	0.75	0.90	0.88	0.85	0.92	0.88	0.88	0.81	100
0.88	0.90	0.89	0.84	0.91	0.84	0.90	0.88	0.85	0.92	0.88	0.88	0.81	200
0.88	0.90	0.88	0.86	0.90	0.95	0.89	0.88	0.85	0.89	0.88	0.88	0.81	500
0.88	0.90	0.88	0.86	0.87	0.97	0.89	0.87	0.85	0.89	0.88	0.88	0.81	1000

جدول (١٧)

النسب المئوية للتحيز لمعاملات ثبات اختبار طويل بخيارات استجابة متعددة لبيانات متجانسة احادية البعد

حجم العينة							معاملات الثبات
1000	500	200	100	50	20		
1%	1%	1%	1%	-2%	-1%	λ1	
2%	2%	2%	2%	-1%	0%	λ2	
2%	2%	2%	2%	-1%	0%	λ3 (α)	
1%	1%	-2%	-2%	-7%	-8%	λ4 (max)	
2%	2%	2%	2%	-1%	-1%	λ5	
2%	1%	1%	1%	-4%	-6%	λ6	
2%	2%	1%	1%	-2%	-2%	ωt	
-8%	-6%	7%	17%	14%	30%	ωh	
3%	0%	-1%	-1%	-6%	-8%	glb	
4%	4%	7%	11%	7%	12%	β	
2%	2%	1%	1%	-2%	-3%	H	
0%	0%	0%	0%	-2%	4%	αpoly	
2%	2%	2%	3%	-1%	0%	αstrata	

- **معاملات الثبات للبيانات ذات الاشتراطات (اختبار متعدد الأبعاد - بيانات مكافئة لتاو - اختبارقصير - استجابة ثنائية)**

من جدول (١٨) وجدول (١٩) نجد أن أعلى معامل ثبات هو أكثر معاملات الثبات ملاءمة هو معامل ثبات أوميجا الكلية (ω_t) في كل أحجام العينة موضع الدراسة بقيم تحيز نسبي ١١%، ٦%، ٦%، ٨%، ٢٢% مقبولة بالمقارنة بباقي نسب معاملات الثبات الأخرى ماعدا حالة العينة (١٠٠) نجد أن معامل ثبات أوميجا الهرمية التقاربي (ω_h) يعطي أعلى

Print:(ISSN 1687-2649) Online:(ISSN 2536-9091)

تقدير للثبات ولكن بنسبة تحيز عالية جدا مما يجعله غير مناسب للاستخدام بإعادة النظر بالجدول نجد أن معامل ثبات أكبر حد أدنى (glb) يعطى أعلى قيمة ثبات بأقل تحيز نسبي غير مؤثر ٣% في اتجاه المبالغة من تقدير الثبات .

جدول (١٨)

بيانات معاملات الثبات لاختبار قصير بخيارات استجابة ثنائية لبيانات مكافئة لتاو متعددة الأبعاد

حجم العينة	λ_1	λ_2	λ_3 (α)	λ_4 (max)	λ_5	λ_6	ω_t	ω_h	glb	β	H	αpoly	αstrata
المجتمع	0.42	0.52	0.51	0.61	0.50	0.49	0.62	0.33	0.62	0.43	0.68	-	-
20	0.26	0.50	0.42	0.66	0.40	0.51	0.69	0.20	0.64	0.22	0.68	0.61	0.53
50	0.12	0.29	0.22	0.51	0.23	0.26	0.58	0.54	0.48	0.01	-	0.47	0.18
100	0.33	0.43	0.40	0.64	0.43	0.41	0.48	0.66	0.64	0.24	-	0.53	0.49
200	0.32	0.42	0.39	0.51	0.41	0.39	0.58	0.44	0.54	0.22	0.50	0.53	0.47
500	0.34	0.42	0.40	0.52	0.42	0.38	0.57	0.39	0.52	0.30	0.46	0.55	0.47
1000	0.22	0.29	0.27	0.38	0.28	0.25	0.46	0.18	0.38	0.19	0.35	0.39	0.36

جدول (١٩)

النسب المئوية للتحيز لمعاملات ثبات اختبار قصير بخيارات استجابة ثنائية لبيانات مكافئة لتاو متعددة الأبعاد

معاملات الثبات		حجم العينة					
		20	50	100	200	500	1000
	λ1	38%	71%	21%	24%	19%	48%
	λ2	4%	44%	17%	19%	19%	44%
	λ3 (α)	18%	57%	22%	24%	22%	47%
	λ4 (max)	-8%	16%	-5%	16%	15%	38%
	λ5	20%	54%	14%	18%	16%	44%
	λ6	-4%	47%	16%	20%	22%	49%
	ωt	-11%	6%	23%	6%	8%	2٢%
	ωh	39%	-64%	-٩٩%	-33%	-18%	45%
	glb	-3%	23%	-3%	13%	16%	39%
	β	49%	98%	44%	49%	30%	56%
	H	0%	99%	99%	26%	32%	49%
	αpoly	-17%	10%	-2%	-2%	-6%	25%
	αstrata	-2%	65%	6%	10%	10%	31%

- **معاملات الثبات للبيانات ذات الاشتراطات (اختبار متعدد الأبعاد – بيانات مكافئة لتاو – اختبارقصير – استجابة متعددة)**

يتضح من جدول (٢٠) وجدول (٢١) أن معامل ثبات أكبر حد أدنى (glb) في حالة العينات (٢٠، ١٠٠) بقيم تحيز نسبي ٣٢%، ٢% في اتجاه المبالغة في تقدير الثبات، ونلاحظ من الجدول أنه بالنسبة للعينة (٢٠) كانت جميع قيم نسب التحيز أعلى من الحد المقبول (١٠%)، وفي حالة العينة (٥٠) نجد أن أعلى معامل ثبات هو معامل ثبات أوميجا

Print:(ISSN 1687-2649) Online:(ISSN 2536-9091)

الهرمية التقاربي (ω_h) ولكن بقيمة تحيز نسبي عالية جدا مما يجعلها غير مقبولة وبإعادة النظر إلى جدول (٢١) نجد أن أعلى قيمة لمعامل ثبات بأقل تحيز نسبي هي قيمة معامل ثبات أكبر حد أدنى (glb)، وفي حالة العينات (٢٠٠، ٥٠٠، ١٠٠٠) نجد أن معامل ثبات أوميجا الكلية (ω_t) يعطي أعلى تقدير للثبات بقيم تحيز نسبي ١٣%، ٨%، ١١% على التوالي في اتجاه المبالغة النسبية في تقدير الثبات.

جدول (٢٠)

بيانات معاملات الثبات لاختبار قصير بخيارات استجابة متعددة لبيانات مكافئة لتاو متعددة الأبعاد

αstrata	αpoly	H	β	glb	ω_h	ω_t	λ_6	λ_5	λ_4 (max)	λ_3 (α)	λ_2	λ_1	حجم العينة
-	-	0.68	0.43	0.62	0.33	0.62	0.49	0.50	0.61	0.51	0.52	0.42	المجتمع
0.74	0.61	0.76	0.66	0.83	0.78	0.82	0.73	0.72	0.81	0.73	0.74	0.61	20
0.57	0.54	0.63	0.36	0.70	0.75	0.63	0.53	0.52	0.70	0.50	0.53	0.41	50
0.54	0.49	0.55	0.23	0.64	0.32	0.63	0.46	0.47	0.57	0.46	0.49	0.38	100
0.46	0.38	0.50	0.12	0.48	0.48	0.54	0.35	0.38	0.48	0.36	0.39	0.30	200
0.62	0.57	0.56	0.42	0.63	0.65	0.67	0.52	0.54	0.63	0.54	0.56	0.45	500
0.58	0.51	0.51	0.38	0.61	0.30	0.69	0.46	0.48	0.61	0.48	0.50	0.40	1000

جدول (٢١)

النسب المئوية للتحيز لمعاملات ثبات اختبار قصير بخيارات استجابة متعددة لبيانات مكافئة لتاو متعددة الأبعاد

حجم العينة							
1000	500	200	100	50	20		معاملات الثبات
5%	-7%	29%	10%	2%	-45%	λ1	
4%	-8%	25%	6%	-2%	-42%	λ2	
6%	-6%	29%	10%	2%	-43%	λ3 (α)	
0%	-3%	21%	7%	-15%	-33%	λ4 (max)	
4%	-8%	24%	6%	-4%	-44%	λ5	
6%	-6%	29%	6%	-8%	-49%	λ6	
-11%	-8%	13%	-2%	-2%	-32%	ωt	
9%	-97%	-45%	3%	-127%	-136%	ωh	
2%	-2%	23%	-3%	-13%	-34%	glb	
12%	2%	72%	47%	16%	-53%	β	
25%	18%	26%	19%	7%	-12%	H	
2%	-10%	27%	6%	-4%	-17%	αpoly	
-12%	-19%	12%	-4%	-10%	-42%	αstrata	

- **معاملات الثبات للبيانات ذات الاشتراطات (اختبار متعدد الأبعاد - بيانات مكافئة لتاو - اختبار طويل -خيارات استجابة ثنائية)**

يتضح من جدول (٢٢) وجدول (٢٣) أن معامل ثبات الحد الأدنى لجتمان (λ_4) الأقصى هو أعلى معامل ثبات في حالة العينات (٢٠، ٥٠، ١٠٠، ٢٠٠) بقيم تحيز نسبي ١٤%،

Print:(ISSN 1687-2649) Online:(ISSN 2536-9091)

٤%، ١%، ١٠% على التوالي وكانت في اتجاه التقليل من قيمة الثبات ماعدا بالنسبة للعينة (٢٠) نجد أن معامل ثبات أكبر حد أدنى (glb) يعطي أفضل قيمة لتقدير الثبات بأقل قيمة لنسبة التحيز غير المؤثرة ١٠% في اتجاه المبالغة في قيمة الثبات، وأيضا يعطي أعلى قيمة في حالة العينة (٥٠، ١٠٠، ٢٠٠) بقيم تحيز نسبي ٤%، ١%، ١٠% على التوالي وكانت في اتجاه التقليل من قيمة الثبات، بينما في حالة العينات (٢٠٠، ٥٠٠، ١٠٠٠) نجد أن معامل ثبات ألفا للأقسام المتعددة (α_{poly}) يعطي أعلى تقدير للثبات بقيم تحيز نسبي غير مؤثرة وهي بالترتيب ٤%، ٣%، ٠% .

جدول (٢٢)

بيانات معاملات الثبات لاختبار طويل بخيارات استجابة ثنائية لبيانات مكافئة لتاو متعددة الأبعاد

αstrata	αpoly	H	β	glb	ω_h	ω_t	λ_6	λ_5	λ_4 (max)	λ_3 (α)	λ_2	λ_1	حجم العينة
-	-	0.62	0.57	0.80	0.45	0.81	0.76	0.72	0.80	0.74	0.75	0.68	المجتمع
0.51	0.72	-	0.17	0.88	0.15	0.75	0.87	0.53	0.91	0.63	0.68	0.42	20
0.63	0.68	0.76	0.15	0.77	0.48	0.59	0.69	0.57	0.77	0.56	0.62	0.50	50
0.69	0.78	0.72	0.53	0.79	0.58	0.72	0.69	0.66	0.79	0.66	0.68	0.61	100
0.64	0.72	0.65	0.38	0.72	0.50	0.66	0.62	0.59	0.72	0.59	0.61	0.54	200
0.70	0.77	0.66	0.48	0.74	0.56	0.71	0.66	0.64	0.74	0.65	0.66	0.60	500
0.68	0.75	0.63	0.44	0.69	0.53	0.68	0.63	0.61	0.71	0.62	0.63	0.57	1000

جدول (٢٣)

النسب المئوية للتحيز لمعاملات ثبات اختبار طويل بخيارات استجابة ثنائية لبيانات مكافئة لتاو متعددة الأبعاد

حجم العينة							معاملات الثبات
1000	500	200	100	50	20		
16%	12%	21%	10%	26%	38%	λ1	
16%	12%	19%	9%	17%	9%	λ2	
16%	12%	20%	11%	24%	15%	λ3 (α)	
11%	8%	10%	1%	4%	-14%	λ4 (max)	
15%	11%	18%	8%	21%	26%	λ5	
17%	13%	18%	9%	9%	-14%	λ6	
16%	12%	19%	11%	27%	7%	ωt	
-18%	-24%	-11%	-29%	-7%	67%	ωh	
14%	8%	10%	1%	4%	-10%	glb	
23%	16%	33%	7%	74%	70%	β	
-2%	-6%	-5%	-16%	-23%	98%	H	
0%	-3%	4%	-4%	9%	4%	αpoly	
9%	7%	15%	8%	16%	32%	αstrata	

Print:(ISSN 1687-2649) Online:(ISSN 2536-9091)

- **معاملات الثبات للبيانات بالاشتراطات (اختبار متعدد الأبعاد - بيانات مكافئة لتاو -اختبار طويل - استجابات متعددة)**

يتضح من جدول (٢٤) وجدول (٢٥) أن معامل ثبات الحد الأدنى لجتمان (λ_4) الأقصى هو أعلى تقدير لمعامل ثبات في حالة العينات (٢٠، ٥٠، ٢٠٠، ٥٠٠، ١٠٠٠) بقيم تحيز نسبي ١٥%، ١١%، ٢%، ٤%، ١% في اتجاه المبالغة في التقدير، ويشترك معه معامل ثبات أكبر حد أدنى (glb) في حالة العينة (٥٠) بقيمة تحيز نسبي ١١% في اتجاه المبالغة في تقدير الثبات، بينما في حالة العينة (١٠٠) يكون هو الأعلى في تقدير الثبات بقيمة تحيز نسبي ١% في اتجاه المبالغة من تقدير الثبات.

جدول (٢٤)

بيانات معاملات الثبات لاختبار طويل بخيارات استجابة متعددة لبيانات مكافئة لتاو متعددة الأبعاد

αstrata	αpoly	H	β	glb	ω_h	ω_t	λ_6	λ_5	λ_4 (max)	λ_3 (α)	λ_2	λ_1	حجم العينة
-	-	0.62	0.57	0.80	0.45	0.81	0.76	0.72	0.80	0.74	0.75	0.68	المجتمع
0.71	0.72	-	0.41	0.68	0.54	0.83	0.91	0.62	0.92	0.74	0.78	0.50	20
0.80	0.79	0.81	0.57	0.89	0.54	0.83	0.83	0.76	0.89	0.76	0.78	0.70	50
0.71	0.65	0.71	0.34	0.81	0.61	0.68	0.69	0.63	0.79	0.63	0.66	0.58	100
0.77	0.74	0.74	0.51	0.80	0.58	0.78	0.75	0.71	0.82	0.72	0.73	0.66	200
0.80	0.76	0.75	0.56	0.82	0.56	0.81	0.77	0.73	0.83	0.74	0.75	0.68	500
0.79	0.75	0.74	0.54	0.79	0.57	0.79	0.75	0.72	0.81	0.73	0.74	0.67	1000

جدول (٢٥)

النسب المئوية للتحيز لمعاملات ثبات اختبار طويل بخيارات استجابة متعددة لبيانات مكافئة لتاو متعددة الأبعاد

حجم العينة							
1000	500	200	100	50	20		معاملات الثبات
1%	0%	3%	15%	-3%	26%	λ1	
1%	0%	3%	12%	-4%	-4%	λ2	
1%	0%	3%	15%	-3%	0%	λ3 (α)	
-1%	-4%	-2%	1%	-11%	-15%	λ4 (max)	
0%	-1%	1%	13%	-6%	14%	λ5	
1%	-1%	1%	9%	-9%	-20%	λ6	
2%	0%	4%	16%	-2%	-2%	ωt	
-27%	-24%	-29%	-36%	-20%	-20%	ωh	
1%	-2%	0%	-1%	-11%	15%	glb	
5%	2%	11%	40%	0%	28%	β	
-19%	-21%	-19%	-15%	-31%	98%	H	
0%	-1%	1%	13%	-5%	4%	αpoly	
-5%	-7%	-3%	5%	-7%	5%	αstrata	

Print:(ISSN 1687-2649) Online:(ISSN 2536-9091)

- **معاملات الثبات للبيانات ذات الاشتراطات (اختبار متعدد الأبعاد - بيانات متجانسة -اختبار قصير -خيارات استجابة ثنائية)**

يتضح من جدول (٢٦) وجدول (٢٧) نجد أن أعلى معامل ثبات هو معامل ثبات أوميجا الكلية (ω_t) في كل أحجام العينة موضع الدراسة بقيم تحيز نسبي ١٤%، ٧%، ١٤%، ١٠%، ١٣% بالترتيب وفي اتجاه التقليل من تقدير الثبات، بينما في حالة العينة (٢٠) نجد أن معامل ثبات أوميجا الهرمية التقاربي (ω_h) يعطي أعلى تقدير للثبات ولكن بقيمة تحيز نسبي كبيرة للغاية مما يستدعي تجاهلها وبإعادة النظر في الجدول (٢٦) والجدول (٢٧) نجد أن أعلى قيمة لمعامل الثبات هي قيمة معامل ثبات ألفا للأقسام المتعددة (α_{poly}) بقيمة تحيز نسبي ٥% غير مؤثرة في اتجاه المبالغة في تقدير الثبات.

جدول (٢٦)

بيانات معاملات الثبات لاختبار قصير بخيارات استجابة ثنائية لبيانات متجانسة متعددة الأبعاد

حجم العينة	λ_1	λ_2	λ_3 (α)	λ_4 (max)	λ_5	λ_6	ω_t	ω_h	glb	β	H	αpoly	αstrata
المجتمع	0.49	0.61	0.59	0.71	0.58	0.59	0.72	0.37	0.72	0.50	0.46	-	-
20	0.42	0.56	0.50	0.73	0.55	0.58	0.62	0.99	0.75	0.24	0.86	0.64	0.62
50	0.25	0.38	0.33	0.59	0.38	0.40	0.62	0.35	0.59	0.08	-	0.48	0.47
100	0.40	0.50	0.48	0.63	0.49	0.48	0.67	0.53	0.61	0.35	0.59	0.62	0.56
200	0.38	0.48	0.45	0.57	0.47	0.45	0.62	0.56	0.58	0.34	0.53	0.60	0.56
500	0.41	0.50	0.49	0.59	0.48	0.46	0.65	0.60	0.59	0.40	0.51	0.64	0.56
1000	0.31	0.40	٠.٣٧	0.54	0.39	0.37	0.63	0.18	0.54	0.23	0.47	0.51	0.49

جدول (٢٧)

النسب المئوية للتحيز لمعاملات ثبات اختبار قصير بخيارات استجابة ثنائية لبيانات متجانسة متعددة الأبعاد

معاملات الثبات	حجم العينة					
	20	50	100	200	500	1000
λ1	14%	49%	18%	22%	16%	37%
λ2	8%	38%	18%	21%	18%	34%
λ3 (α)	15%	44%	19%	24%	17%	37%
λ4 (max)	-3%	17%	11%	20%	17%	24%
λ5	5%	34%	16%	19%	17%	33%
λ6	2%	32%	19%	24%	22%	37%
ωt	14%	14%	7%	14%	10%	13%
ωh	-168%	5%	-43%	-51%	-62%	51%
glb	-4%	18%	15%	19%	18%	25%
β	52%	84%	30%	32%	20%	54%
H	-87%	98%	-28%	-15%	-11%	-2%
αpoly	-5%	21%	-2%	2%	-5%	16%
αstrata	-2%	23%	8%	8%	8%	20%

- **معاملات الثبات للبيانات ذات الاشتراطات (اختبار متعدد الأبعاد - بيانات متجانسة -اختبار قصير -خيارات استجابة متعددة)**

يتضح من جدول (٢٨) وجدول (٢٩) نجد أن أعلى معامل ثبات هو معامل ثبات أوميجا الكلية (ω_t) في كل أحجام العينة موضع الدراسة بقيم تحيز نسبي غير مؤثرة ويمكن تجاهلها ٦%، ٠%، ١%، ٤% ماعدا حالة العينتين (٢٠، ٥٠) نجد أن معامل ثبات أوميجا الهرمية التقاربي (ω_h) يعطي أعلى تقدير للثبات ولكن بقيم تحيز نسبي كبيرة للغاية مما يستدعي استبدالها بمعامل أخر وبإعادة النظر في الجداول (٢٨، ٢٩) يتضح أن أعلى قيمة لمعامل الثبات بأقل قيم للتحيز النسبي لحالة العينات (٢٠، ٥٠) كانت لمعامل ثبات أكبر حد أدنى (glb) بقيم تحيز نسبي ١٠%، ١% في اتجاه التقليل من تقدير الثبات .

جدول (٢٨)

بيانات معاملات الثبات لاختبار قصير بخيارات استجابة متعددة لبيانات متجانسة متعددة الأبعاد

αstrata	αpoly	H	β	glb	ω_h	ω_t	λ_6	λ_5	λ_4 (max)	λ_3 (α)	λ_2	λ_1	حجم العينة
-	-	0.46	0.50	0.72	0.37	0.72	0.59	0.58	0.71	0.59	0.61	0.49	المجتمع
0.47	0.53	-	0.15	0.65	0.82	0.58	0.52	0.42	0.68	0.47	0.51	0.29	20
0.56	0.51	-	0.23	0.71	0.75	0.65	0.54	0.53	0.67	0.50	0.54	0.42	50
0.71	0.66	0.69	0.53	0.73	0.59	0.76	0.63	0.64	0.73	0.64	0.65	0.53	100
0.63	0.55	0.60	0.39	0.66	0.32	0.72	0.52	0.52	0.66	0.52	0.54	0.43	200
0.66	0.53	0.58	0.35	0.68	0.47	0.73	0.53	0.51	0.68	0.50	0.54	0.42	500
0.71	0.63	0.61	0.50	0.74	0.48	0.75	0.60	0.59	0.74	0.60	0.61	0.50	1000

جدول (٢٩)

النسب المئوية للتحيز لمعاملات ثبات اختبار قصير بخيارات استجابة متعددة لبيانات متجانسة متعددة الأبعاد

حجم العينة							معاملات الثبات
1000	500	200	100	50	20		
-2%	14%	12%	-8%	14%	41%	λ1	
0%	11%	11%	-7%	11%	16%	λ2	
-2%	15%	12%	-8%	15%	20%	λ3 (α)	
-4%	4%	7%	-3%	6%	4%	λ4 (max)	
-2%	12%	10%	-10%	9%	28%	λ5	
-2%	10%	12%	-7%	8%	12%	λ6	
-4%	-1%	0%	-6%	10%	19%	ωt	
-30%	-27%	14%	-59%	-103%	-122%	ωh	
-3%	6%	8%	-1%	1%	10%	glb	
0%	30%	22%	-6%	54%	70%	β	
-33%	-26%	-30%	-50%	98%	98%	H	
-3%	13%	10%	-8%	16%	13%	αpoly	
-16%	-8%	-3%	-16%	8%	23%	αstrata	

- **معاملات الثبات للبيانات ذات الاشتراطات (اختبار متعدد الأبعاد - بيانات متجانسة -اختبار طويل -خيارات استجابة ثنائية)**

يتضح من جدول (٣٠) وجدول (٣١) أن معامل ثبات أكبر حد أدنى (glb) هو أعلى معامل ثبات في حالة العينة (٢٠، ٥٠، ١٠٠) بقيم تحيز نسبي ١٣%، ٤%، ١% في اتجاه المبالغة من تقدير الثبات، بينما في حالة العينة (٢٠٠) فإن معامل ثبات الحد الأدنى لجتمان الأقصى (λ_4) هو الأعلى تقديرا للثبات بقيمة تحيز نسبي ١١% في اتجاه التقليل من تقدير الثبات، ونجد أن معامل ثبات ألفا للأقسام المتعددة (α_{poly}) يعطي أعلى تقدير للثبات في حالة العينة (١٠٠) بالاشتراك مع معامل ثبات أكبر حد أدنى ولكن بنسبة تحيز أعلى بقيمة ٦% في اتجاه المبالغة من تقدير الثبات، ويعطي (α_{poly}) أعلى تقدير للثبات في حالة العينات الكبيرة (٥٠٠، ١٠٠٠) بنسب تحيز غير مؤثرة ٣%، و ٠%.

جدول (٣٠)

بيانات معاملات الثبات لاختبار طويل بخيارات استجابة ثنائية لبيانات متجانسة متعددة الأبعاد

αstrata	αpoly	H	β	glb	ω_h	ω_t	λ_6	λ_5	λ_4 (max)	λ_3 (α)	λ_2	λ_1	حجم العينة
-	-	0.64	0.59	0.83	0.47	0.83	0.79	0.74	0.83	0.76	0.77	0.70	المجتمع
0.73	0.65	-	0.11	0.94	0.79	0.72	0.89	0.69	0.85	0.66	0.71	0.61	20
0.71	0.76	0.75	0.39	0.86	0.46	0.75	0.75	0.66	0.85	0.65	0.69	0.60	50
0.75	0.82	0.76	0.54	0.82	0.53	0.77	0.74	0.71	0.82	0.71	0.73	0.65	100
0.67	0.73	0.66	0.37	0.71	0.47	0.68	0.64	0.61	0.74	0.61	0.63	0.56	200
0.72	0.79	0.68	0.49	0.76	0.53	0.73	0.68	0.66	0.77	0.67	0.68	0.61	500
0.70	0.77	0.65	0.49	0.68	0.57	0.70	0.65	0.63	0.72	0.65	0.66	0.59	1000

جدول (٣١)

النسب المئوية للتحيز لمعاملات ثبات اختبار طويل بخيارات استجابة ثنائية لبيانات متجانسة متعددة الأبعاد

حجم العينة							
1000	500	200	100	50	20		معاملات الثبات
16%	13%	20%	7%	14%	13%	λ1	
14%	12%	18%	5%	10%	8%	λ2	
14%	12%	20%	7%	14%	13%	λ3 (α)	
13%	7%	11%	1%	-2%	-2%	λ4 (max)	
15%	11%	18%	4%	11%	7%	λ5	
18%	14%	19%	6%	5%	-13%	λ6	
16%	12%	18%	7%	10%	13%	ωt	
-21%	-13%	0%	-13%	2%	-68%	ωh	
18%	8%	14%	1%	-4%	-13%	glb	
17%	17%	37%	8%	34%	81%	β	
-2%	-6%	-3%	-19%	-17%	98%	H	
0%	-3%	5%	-6%	1%	16%	αpoly	
9%	6%	13%	3%	8%	5%	αstrata	

Print:(ISSN 1687-2649) Online:(ISSN 2536-9091)

- **معاملات الثبات للبيانات ذات الاشتراطات (اختبار متعدد الأبعاد - بيانات متجانسة -اختبار طويل -خيارات استجابة متعددة)**

يتضح من جدول (٣٢) وجدول (٣٣) أن معامل ثبات أكبر حد أدنى (glb) ومعامل ثبات الحد الأدنى لجتمان (λ_4) الأقصى يعطيا سويا أعلى معامل ثبات في حالة كل أحجام العينات بقيم تحيز نسبي ١٤%، ٧%، ٢%، ٤%، ٠%، في اتجاه المبالغة في تقدير الثبات، بينما يتفوق معامل ثبات أكبر حد أدنى في حالة العينة (٢٠) ولكن بقيمة تحيز نسبي أعلى يساوي ١٦%، ويتفوق معامل ثبات الحد الأدنى لجتمان (λ_4) الأقصى بفارق ضئيل في حالة حجم العينة الكبيرة (١٠٠٠) ولكن بنسبة تحيز نسبي أقل يساوي ١% مما يعطيه أفضلية في تقدير الثبات.

جدول (٣٢)

بيانات معاملات الثبات لاختبار طويل بخيارات استجابة متعددة لبيانات متجانسة متعددة الأبعاد

α_{strata}	α_{poly}	H	β	glb	ω_h	ω_t	λ_6	λ_5	λ_4 (max)	λ_3 (α)	λ_2	λ_1	حجم العينة
-	-	0.64	0.59	0.83	0.47	0.83	0.79	0.74	0.83	0.76	0.77	0.70	المجتمع
0.80	0.65	0.86	0.56	0.96	0.89	0.85	0.92	0.78	0.95	0.78	0.80	0.71	20
0.81	0.78	0.79	0.51	0.89	0.64	0.83	0.82	0.75	0.89	0.76	0.78	0.70	50
0.79	0.73	0.80	0.41	0.85	0.59	0.80	0.77	0.71	0.85	0.71	0.73	0.65	100
0.81	0.76	0.80	0.47	0.86	0.49	0.81	0.78	0.73	0.86	0.73	0.75	0.67	200
0.80	0.76	0.75	0.55	0.83	0.55	0.80	0.76	0.72	0.83	0.74	0.75	0.68	500
0.80	0.75	0.74	0.51	0.80	0.51	0.80	0.76	0.72	0.82	0.73	0.75	0.67	1000

جدول (٣٣)

النسب المئوية للتحيز لمعاملات ثبات اختبار طويل بخيارات استجابة متعددة لبيانات متجانسة متعددة الأبعاد

حجم العينة							
1000	500	200	100	50	20		معاملات الثبات
4%	3%	4%	7%	0%	-1%	λ_1	
3%	3%	3%	5%	-1%	-4%	λ_2	
4%	3%	4%	7%	0%	-3%	λ_3 (α)	
1%	0%	-4%	-2%	-7%	-14%	λ_4 (max)	
3%	3%	1%	4%	-1%	-5%	λ_5	
4%	4%	1%	3%	-4%	-16%	λ_6	
4%	4%	2%	4%	0%	-2%	ω_t	
-9%	-17%	-4%	-26%	-36%	-89%	ω_h	
4%	0%	-4%	-2%	-7%	-16%	glb	
14%	7%	20%	31%	14%	5%	β	
-16%	-17%	-25%	-25%	-23%	-34%	H	
3%	1%	1%	5%	-1%	16%	α_{poly}	
-4%	-4%	-5%	-3%	-5%	-4%	α_{strata}	

Print:(ISSN 1687-2649) Online:(ISSN 2536-9091)

ولتعيين أفضل معامل ثبات للاستخدام في حالة اختلاف نوعية المقاييس التي يتم قياسها (أحادية البعد أو تعدد الأبعاد)، واختلاف حجم العينة، واختلاف عدد العبارات (طول المقياس)، واختلاف نوعية بيانات القياس (بيانات مكافئة لتاو – بيانات متجانسة)، واختلاف خيارات الاستجابة (استجابة ثنائية – استجابة متعددة من نوع ليكارت) كما يتضح من جدول (٣٤).

جدول (٣٤)
قيم متوسطات وتحيزات معاملات الثبات في حالة اختلاف نوعية البيانات

نوع البيانات		λ1	λ2	α	λ4	λ5	λ6	ωt	ωh	glb	β	H	αpoly	αstrata
بيانات اختبار أحادي البعد	م	0.7	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.7	0.8	0.8	0.8
	ت	10	9	10	2	9	7	3	10	3	20	10	3	10
بيانات اختبار متعدد الأبعاد	م	0.5	0.6	0.6	0.7	0.6	0.6	0.7	0.5	0.7	0.4	0.6	0.6	0.6
	ت	15	10	13	3	11	8	6	-33	4	30	7	3	4
بيانات اختبار قصير	م	0.5	0.6	0.6	0.7	0.6	0.6	0.7	0.6	0.7	0.4	0.6	0.7	0.6
	ت	17	13	16	6	13	12	3	-19	6	31	17	4	9
بيانات اختبار طويل	م	0.7	0.8	0.8	0.9	0.8	0.8	0.8	0.7	0.9	0.6	0.8	0.8	0.8
	ت	8	6	7	-1	6	2	6	-4	1	19	1	1	5
بيانات مكافىء لتاو	م	0.6	0.7	0.6	0.8	0.6	0.7	0.7	0.6	0.7	0.5	0.6	0.7	0.7
	ت	13	9	12	2	10	7	5	-11	3	27	16	2	8
بيانات متجانسة	م	0.6	0.7	0.7	0.8	0.7	0.7	0.8	0.7	0.8	0.6	0.7	0.8	0.7
	ت	11	10	11	3	9	7	5	-12	3	24	1	3	6
بيانات خيارات ثنائية	م	0.5	0.6	0.6	0.7	0.6	0.6	0.7	0.6	0.7	0.5	0.6	0.7	0.6
	ت	20	16	19	7	17	14	10	-5	8	35	15	3	15
بيانات خيارات متعددة	م	0.6	0.7	0.7	0.8	0.7	0.7	0.8	0.7	0.8	0.6	0.7	0.7	0.8
	ت	4	3	4	-2	3	1	0	-19	-1	15	3	3	-1

ويتضح من الجدول السابق أنه في حالة بيانات الاختبار أحادي البعد، يكون أعلى قيمة لمعامل الثبات والتي تحيزها أقل من ١٠% أي يمكن اعتبارها قيمة غير متحيزة هي لمعامل ثبات أوميجا الكلية (ω_t)، بينما في حالة بيانات الاختبار متعدد الأبعاد كانت أعلى قيمة غير متحيزة لمعامل ثبات الحد الأدنى لجتمان (λ_4) الأقصى، وفي حالة البيانات التي تم الحصول عليها من اختبار قصير كانت أعلى قيمة لمعامل الثبات التي تم اعتبارها كقيمة غير متحيزة أقل من ١٠% هي لمعامل ثبات أوميجا الكلية (ω_t)، بينما في حالة بيانات الاختبار الطويل كانت أعلى قيمة غير متحيزة لمعامل ثبات الحد الأدنى لجتمان (λ_4) الأقصى، وفي حالة البيانات المكافىء لتاو والبيانات المتجانسة كانت أعلى قيمة لمعامل الثبات التي تم

اعتبارها كقيمة غير متحيزة أقل من ١٠% هي لمعامل ثبات الحد الأدنى لجتمان (λ_4) الأقصى، وأعطى معامل ثبات أكبر حد أدنى (glb) نفس القيم بالنسبة للبيانات المتجانسة، أما عند اختلاف عدد خيارات الاستجابة نجد أنه كانت أعلى قيمة لمعامل الثبات التي تم اعتبارها كقيمة غير متحيزة أقل من ١٠% هي لمعامل ثبات الحد الأدنى لجتمان (λ_4) الأقصى للخيارات الثنائية وأيضا للخيارات المتعددة.

جدول (٣٥) قيم متوسطات وتحيز معاملات الثبات في حالة اختلاف التداخلات بين نوعية البيانات التي يتم قياسها

αstrata	αpoly	H	β	glb	ωh	ωt	λ6	λ5	λ4	α	λ2	λ1		
0.6	0.7	0.5	0.4	0.7	0.7	0.7	0.6	0.6	0.7	0.6	0.6	0.5	م	١
27	9	33	46	12	10	6	24	24	13	28	24	27	ت	
0.6	0.7	0.5	0.4	0.7	0.7	0.7	0.6	0.6	0.7	0.6	0.6	0.5	م	٢
5	3	19	16	1	4	-2	4	4	0	6	6	6	ت	
0.6	0.7	0.5	0.4	0.7	0.7	0.7	0.6	0.6	0.7	0.6	0.6	0.5	م	٣
10	0	7	22	3	10	9	5	8	2	10	9	10	ت	
0.6	0.7	0.5	0.4	0.7	0.7	0.7	0.6	0.6	0.7	0.6	0.6	0.5	م	٤
0	0	0	7	-2	8	0	-2	0	-4	0	0	1	ت	
0.6	0.7	0.5	0.4	0.7	0.7	0.7	0.6	0.6	0.7	0.6	0.6	0.5	م	٥
18	2	10	27	8	17	4	15	17	9	18	16	18	ت	
0.6	0.7	0.5	0.4	0.7	0.7	0.7	0.6	0.6	0.7	0.6	0.6	0.5	م	٦
7	5	6	15	1	5	-1	6	8	1	8	7	8	ت	
0.6	0.7	0.5	0.4	0.7	0.7	0.7	0.6	0.6	0.7	0.6	0.6	0.5	م	٧
9	1	8	19	1	18	6	5	8	1	9	8	8	ت	
0.6	0.7	0.5	0.4	0.7	0.7	0.7	0.6	0.6	0.7	0.6	0.6	0.5	م	٨
1	0	0	8	-2	9	0	-1	1	-3	1	1	0	ت	
0.6	0.7	0.5	0.4	0.7	0.7	0.7	0.6	0.6	0.7	0.6	0.6	0.5	م	٩
20	1	51	54	14	-22	10	25	28	12	31	25	37	ت	
0.6	0.7	0.5	0.4	0.7	0.7	0.7	0.6	0.6	0.7	0.6	0.6	0.5	م	١٠
-13	1	14	16	-5	-66	-7	-4	-4	-4	0	-3	-1	ت	
0.6	0.7	0.5	0.4	0.7	0.7	0.7	0.6	0.6	0.7	0.6	0.6	0.5	م	١١
14	2	8	37	4	-4	15	9	17	3	16	14	21	ت	
0.6	0.7	0.5	0.4	0.7	0.7	0.7	0.6	0.6	0.7	0.6	0.6	0.5	م	١٢
-2	2	-1	14	0	-26	3	-3	3	-5	3	1	7	ت	
0.6	0.7	0.5	0.4	0.7	0.7	0.7	0.6	0.6	0.7	0.6	0.6	0.5	م	١٣
11	5	-8	45	15	-45	12	23	21	14	26	23	26	ت	
0.6	0.7	0.5	0.4	0.7	0.7	0.7	0.6	0.6	0.7	0.6	0.6	0.5	م	١٤
-2	7	9	28	3	-55	3	6	8	2	9	7	12	ت	
0.6	0.7	0.5	0.4	0.7	0.7	0.7	0.6	0.6	0.7	0.6	0.6	0.5	م	١٥
7	2	9	32	4	-19	13	8	11	5	13	11	14	ت	
0.6	0.7	0.5	0.4	0.7	0.7	0.7	0.6	0.6	0.7	0.6	0.6	0.5	م	١٦
-4	4	-23	15	-4	-30	2	-1	1	-4	2	1	3	ت	

Print:(ISSN 1687-2649) Online:(ISSN 2536-9091)

١.اختبار قصير بخيارات استجابة ثنائية لبيانات مكافئة لتاو احادية البعد

٢.اختبار قصير بخيارات استجابة متعددة لبيانات مكافئة لتاو احادية البعد

٣.اختبار طويل بخيارات استجابة ثنائية لبيانات مكافئة لتاو احادية البعد

٤.اختبار طويل بخيارات استجابة متعددة لبيانات مكافئة لتاو احادية البعد

٥.اختبار قصير بخيارات استجابة ثنائية لبيانات متجانسة احادية البعد

٦.اختبار قصير بخيارات استجابة متعددة لبيانات متجانسة احادية البعد

٧.اختبار طويل بخيارات استجابة ثنائية لبيانات متجانسة احادية البعد

٨.اختبار طويل بخيارات استجابة متعددة لبيانات متجانسة احادية البعد

٩.اختبار قصير بخيارات استجابة ثنائية لبيانات مكافئة لتاو متعددة الأبعاد

١٠.اختبار قصير بخيارات استجابة متعددة لبيانات مكافئة لتاو متعددة الأبعاد

١١.اختبار طويل بخيارات استجابة ثنائية لبيانات مكافئة لتاو متعددة الأبعاد

١٢.اختبار طويل بخيارات استجابة متعددة لبيانات مكافئة لتاو متعددة الأبعاد

١٣.اختبار قصير بخيارات استجابة ثنائية لبيانات متجانسة متعددة الأبعاد

١٤.اختبار قصير بخيارات استجابة متعددة لبيانات متجانسة متعددة الأبعاد

١٥.اختبار طويل بخيارات استجابة ثنائية لبيانات متجانسة متعددة الأبعاد

١٦.اختبار طويل بخيارات استجابة متعددة لبيانات متجانسة متعددة الأبعاد

ويتضح من جدول (٣٥) أنه في حالة اختلاف التداخلات بين نوعية البيانات التي يتم قياسها، كان معامل ثبات أوميجا الكلية (ω_t) هو المعامل الأفضل في حالة الاختبار القصير بخيارات استجابة ثنائية لبيانات مكافئة لتاو احادية البعد، وحالة الاختبار القصير بخيارات استجابة متعددة لبيانات مكافئة لتاو احادية البعد، وحالة الاختبار القصير بخيارات استجابة متعددة لبيانات متجانسة احادية البعد، وحالة الاختبار الطويل بخيارات استجابة متعددة لبيانات متجانسة احادية البعد، وحالة الاختبار القصير بخيارات استجابة ثنائية لبيانات مكافئة لتاو متعددة الأبعاد، وحالة الاختبار القصير بخيارات استجابة متعددة لبيانات مكافئة لتاو متعددة الأبعاد، وحالة بيانات اختبار قصير بخيارات استجابة ثنائية لبيانات متجانسة متعددة الأبعاد، وحالة بيانات اختبار قصير بخيارات استجابة متعددة لبيانات متجانسة متعددة الأبعاد.

وكان معامل ثبات معامل ثبات ألفا للأقسام المتعددة (α_{poly}) هو المعامل الأفضل في حالة بيانات اختبار طويل بخيارات استجابة ثنائية لبيانات مكافئة لتاو احادية البعد، وحالة بيانات اختبار قصير بخيارات استجابة ثنائية لبيانات متجانسة احادية البعد، وحالة بيانات اختبار طويل بخيارات استجابة ثنائية لبيانات متجانسة احادية البعد.

بينما معامل ثبات الحد الأدنى لجتمان ($\lambda_{4\ max}$) الأقصى هو المعامل الأفضل في حالة بيانات اختبار طويل بخيارات استجابة متعددة لبيانات مكافئة لتاو احادية البعد، وحالة بيانات اختبار طويل بخيارات استجابة ثنائية لبيانات مكافئة لتاو متعددة الأبعاد، وحالة بيانات

Print:(ISSN 1687-2649) Online:(ISSN 2536-9091)

اختبار طويل بخيارات استجابة متعددة لبيانات مكافئة لتاو متعددة الأبعاد، وحالة بيانات اختبار طويل بخيارات استجابة متعددة لبيانات متجانسة متعددة الأبعاد.

ويكون معامل ثبات أكبر حد أدنى (glb) هو المعامل الأفضل في حالة بيانات اختبار طويل بخيارات استجابة ثنائية لبيانات متجانسة متعددة الأبعاد، ويشترك مع معامل ثبات الحد الأدنى لجتمان (λ_4) الأقصى في حالتي اختبار طويل بخيارات استجابة ثنائية لبيانات مكافئة لتاو متعددة الأبعاد وحالة اختبار طويل بخيارات استجابة متعددة لبيانات متجانسة متعددة الأبعاد، ويشترك أيضا مع معامل ثبات أوميجا الكلية (ω_t) في اختبار قصير بخيارات استجابة متعددة لبيانات متجانسة متعددة الأبعاد.

الاستنتاجات

تمت المقارنة في هذه الدراسة بين قيم ونسب تحيز ثلاث عشرة نوع مختلف من معاملات الثبات للبيانات وهي معاملات ثبات الحدود الدنيا لجتمان(λ_1، λ_2، λ_3، λ_4، λ_5، λ_6) ومعامل ثبات ألفا كرونباخ (α) الذي هو نفسه(λ_2) لجتمان، ومعامل ثبات أوميجا الكلية (ω_t)، ومعامل ثبات أوميجا الهرمية التقاربي (ω_h)، ومعامل ثبات أكبر حد أدنى (glb)، ومعامل ثبات ألفا الرتبية للأقسام المتعددة (α_{poly})، ومعامل ثبات بيتا (β) الذي قام باشتقاقه Revelle (1979)، ومعامل ثبات ألفا الطبقية (α_{strata})، ومعامل الثبات الأقصى (Maximal Reliability – Coefficient *H*) ، وبالرجوع إلى النتائج الخاصة بالدراسة فقد أظهرت النتائج تحيز معامل ألفا في اتجاه التقليل من الثبات باختلاف حالات البيانات في حالة مخالفة افتراض أحادية البعد وفي حالة العينات الصغيرة والاختبارات القصيرة وفي أغلب حالات نتائج الاختبارات متعددة الأبعاد، وأكدت النتائج ما تناولته الدراسات السابقة لكل من (Bentler, 2007; Cortina, 1993; Dunn, Baguley, & Brunsden, 2013; Geldhof, Preacher, & Zyphur, 2014; Graham, 2006; Green & Hershberger, 2000; Green & Yang, 2009a, 2009b; McNeish, 2017; Peters, 2014; Raykov, 1997a,1997b, 1998, 2004; Raykov & Shrout, 2002; Revelle & Zinbarg, 2009; Schimtt, 1996; Sijtsma, 2009; Yang & Green, 2011; Zinbarg, Yovel, Revelle, & McDonald, 2006; Zinbarg, Revelle, Yovel, & Li, 2005) عن كيفية تطبيق معامل ألفا

Print:(ISSN 1687-2649) Online:(ISSN 2536-9091)

كرونباخ وأوضحت أن الافتراضات التي وضعها كرونباخ لمعامل ثبات ألفا يتم انتهاكها بشكل شائع في أنواع البيانات والنماذج التي يعمل بها الباحثون النفسيون، وقد أدت ظهور هذه الدراسات وانتشارها بين الباحثين إلى تطوير معاملات قياس بديلة تكون افتراضاتها أكثر انسجامًا مع البيانات النفسية.

يتضح أيضا من النتائج أن هناك أربعة أنواع من معاملات الثبات تعطي أعلى معامل تقدير للثبات وتحيزها أقل من ١٠% أي أنه يمكن اعتبارها معاملات ثبات غير متحيزة مما يجعلها الأفضل للاستخدام في حالات البيانات المختلفة، وهي معامل ثبات أكبر حد أدنى (glb) في حالات البيانات المشار إليها بالنتائج بما يتفق مع دراسة كلا من Jackson & Agunwamba (1977) ودراسة Bentler & Woodward (1980) ، ودراسة Sijtsma (2009)، وعارض ذلك دراسةRevelle & Zinbarg (2009) حيث ناقشا تحيز معامل الثبات في حالة العينات الصغيرة نسبيا، وقام Bendermacher (2017) بعمل خوارزم يقوم بتصحيح هذا التحيز، وحسب نتائج الدراسة الحالية يشترك معامل ثبات أكبر حد أدنى مع معامل ثبات الحد الأدنى لجتمان (λ_4) الأقصى في الكثير من حالات البيانات بفوارق بسيطة لصالح (glb)، وأتفق ذلك مع دراسة Benton (2015)، من ناحية أخرى أشارت دراسة Oosterwijk, et. al (2017) إلى مبالغة هذا المعامل في تقدير الثبات، وجاءت نتائج الدراسة الحالية لتؤكد عدم تحيز معامل ثبات (λ_4) في حالات معينة للبيانات كما ورد ذكره بجداول (٣٤، ٣٥) وأتفقت دراسة Woodruff (2012) ودراسة Hunt & Bentler (2015) مع هذه النتائج ، وأعطى معامل ثبات ألفا للأقسام المتعددة (α_{poly}) نتائج جيدة مع الاختبارات أحادية البعد ذات خيارات الاستجابة الثنائية وكانت النتائج تقريبا بدون أي تحيز نسبي، حيث تفوقت على معامل ألفا كرونباخ التقليدي وذلك لاستخدامها مصفوفة الارتباط المتعددة بديلا عن مصفوفة ارتباط بيرسون وجاءت هذه النتائج متسقة مع نتائج دراسة Gadermann, Guhn & Zumbo (2012) ودراسة Bonanomi, Ruscone & Osmetti (2013) ، ودراسة Traxler (2018) وعارض ذلك دراسة Chalmers (2018) والتي ناقش فيها الفائدة المحدودة وسوء الاستخدام لمعامل ثبات ألفا للأقسام المتعددة.

التوصيات

Print:(ISSN 1687-2649) Online:(ISSN 2536-9091)

كان أفضل معاملات الثبات على مستوى تنوع واختلاف البيانات هو معامل ثبات أوميجا الكلية (ω_t) حيث أعطى أعلى قيم للثبات بأقل تحيز نسبي لغالبية التداخلات بين حالات وأنواع البيانات والاختبارات ويتعدد طول العبارات وتنوع العينات وكان الأقرب للقيم الأعلى في حالات البيانات التي حققت فيها معاملات الثبات الأخرى نتائج جيدة، وجاء تميز هذا المعامل لاعتماده على تحليل العوامل الاستكشافية والتأكيدية، وبالرجوع إلى النتائج بالجدول (٣٥) يمكن اعتبار معامل ثبات أوميجا الكلية غير مناسب للبيانات ذات الخيارات الثنائية حيث كانت قيم معاملات الثبات ذات تحيز نسبي في اتجاه الاقلال من قيم الثبات واتفق ذلك مع دراسة (Flora, 2020)، واعتمدت طريقة حسابات أوميجا الكلية على النموذج البنائي للعوامل ومن ثم عمل تحويل شميدت–ليمان (Schimd–Leiman rotation) للابعاد، واتفقت هذه النتائج مع دراسات (Graham, 2006; McDonald, 1999; Padilla & Divers, 2016; Trizano–Hermosilla & Alvarado, 2016; Zinbarg, Revelle, Yovel, & Li, 2005).

مقترحات البحث

بناءا على نتائج الدراسة الحالية تظهر الحاجة إلى التعرف الجيد على خواص ومواصفات البيانات والاختبارات محل الدراسة والتطوير للباحثين والتأكيد على عمل تحليل عاملي للبيانات قبل البدء في التطبيق وذلك للتعرف الجيد على خصائص البيانات وتحديد معامل الثبات الأنسب للاستخدام، ولفهم التحيز النسبي لمعاملات الثبات بشكل أفضل في النماذج المختلفة للبيانات، تظهر الحاجة إلى عمل المزيد من الدراسات التي تدرس البيانات في حالتي الالتواء والتفرطح والبيانات التي لا تتبع التوزيع الطبيعي ولكن تتبع توزيعات أخرى، وأيضا دراسة معاملات الثبات في إطار نظرية الاستجابة للمفردة IRT من حيث ثبات الأفراد وثبات المفردات وذلك للنماذج أحادية البعد ومتعددة الأبعاد، وتظهر الحاجة إلى إجراء المزيد من الدراسات التي تستخدم بيانات حقيقية للتأكيد على النتائج التي تم الحصول عليها.

Print:(ISSN 1687-2649) Online:(ISSN 2536-9091)

المراجع

أولاً: المراجع العربية:

Aguirre-Urreta, M. I., Rönkkö, M., & McIntosh, C. N. (2019). A cautionary note on the finite sample behavior of maximal reliability. Psychological Methods, 24 (2), 236-252. doi:10.1037/met0000176

Allen, M. J., & Yen, W. M. (1979). *Introduction to measurement theory*. Monterey, CA: Brooks/Cole.

Bendermacher, N. (2017). An unbiased estimator of the greatest lower bound. *Journal of Modern Applied Statistical Methods*, *16(*1), 674-688. doi: 10.22237/jmasm/1493598960

Bentler, P. (2009). Alpha, dimension-free and model-based internal consistency reliability. *Psychometrika*, *74*, 137-143. doi:10.1007/s11336-008-9001-1

Bentler, P. M. (2006). *EQS 6 Structural equations programs manual*. Los Angeles, CA: University of California at Los Angeles.

Bentler, P. M. (2007). Covariance structure models for maximal reliability of unit-weighted composites. In S. Lee (Ed.), Handbook of computing and statistics with applications: Vol. 1. *Handbook of latent variable and related models* (pp. 1–19). New York: Elsevier.

Bentler, P. M. (2017). Specificity-enhanced reliability coefficients. *Psychological Methods*, *22*(3), 527–540. https://doi.org/10.1037/met0000092

Bentler, P. M., & Woodward, J. A. (1980). Inequalities among lower bounds to reliability: With applications to test construction and factor analysis. *Psychometrika*, *45*(2), 249–267. https://doi.org/10.1007/BF02294079

Bentler, P.M. (1972). A lower-bound method for the dimension-free measurement of reliability. *Social Science Research, 1*, 343–357.

Benton, T. (2015). An Empirical Assessment of Guttman's Lambda 4 Reliability Coefficient. *Springer Proceedings in Mathematics and Statistics. 89*. 301-310. 10.1007/978-3-319-07503-7_19.

Brown, W. (1910). Some experimental results in the correlation of mental abilities. *British Journal of Psychology*, 3, 296-322.

Callender, J. C., & Osburn, H. G. (1977). A Method for Maximizing Split-Half Reliability Coefficients. *Educational and Psychological Measurement*, *37*(4), 819–825. https://doi.org/10.1177/001316447703700402

Print:(ISSN 1687-2649) Online:(ISSN 2536-9091)

Chalmers, R. P. (2018). On Misconceptions and the Limited Usefulness of Ordinal Alpha. *Educational and Psychological Measurement*, *78*(6), 1056–1071. https://doi.org/10.1177/0013164417727036

Cho, E. (2016). Making Reliability Reliable: A Systematic Approach to Reliability Coefficients. In *Organizational Research Methods* (Vol. 19, Issue 4). https://doi.org/10.1177/1094428116656239

Cho, E., and Kim, S. (2015). Cronbach's coefficient alpha: well known but poorly understood. *Organ. Res. Methods* 18,207–230.doi:10.1177/109442811 4555994

Cohen, R. J., Swerdlik, M. (2018). *Psychological testing and assessment: An introduction to tests and measurement* (9th ed.). New York, NY: McGraw Hill

Cortina, J. M. (1993). What is coefficient alpha? An examination of theory and applications. *Journal of Applied Psychology*, *78*, 98-104.

Crocker, L. M., & Algina, J. (1986). *Introduction to classical and modern test theory*. New York, NY: Holt, Rinehart, and Winston.

Cronbach, L. J. (1951). Coefficient α and the internal structure of tests. *Psychometrika, 16*, 297-334.

Cronbach, L. J., & Furby, L. (1970). How we should measure "change": Or should we? *Psychological Bulletin, 74*(1), 68–80. https://doi.org/10.1037/h0029382

Cronbach, L. J., Gleser, G. C., Nanda, H., & Rajaratnam, N. (1972). *The dependability of behavioral measurements: Theory of generalizability for scores and profiles*. New York: Wiley.

Cronbach, L.J. (1947). Test "reliability": Its meaning and determination. *Psychometrika* 12**,** 1–16 https://doi.org/10.1007/BF02289289

Dunn, Thomas & Baguley, Thom & Brunsden, Vivienne. (2013). From alpha to omega: A practical solution to the pervasive problem of internal consistency estimation. *British Journal of Psychology*. *105*. 10.1111/bjop.12046.

Ekström, J. (2010). On the relation between the polychoric correlation coefficient and Spearman's rank correlation coefficient. *Department of Statistics Papers. Retrieved from the UCLA* website: https://escholarship.org/uc/item/7j01t5sf

Feldt, L. S. (1969), A test of the hypothesis that cronbach's alpha or kuder-richardson coefficent twenty is the same for two tests, *Psychometrika*, *34*, (3), 363-373

Feldt, L. S., & Brennan, R. L. (1989). Reliability. *Educational Measurement*, 3rd ed., 105-146.

Flanagan, J. C. (1937). A Proposed Procedure for Increasing the Efficiency of Objective Tests. *Journal of Educational Psychology, XXVII*, 17-21.

Flora, D. B. (2020). Your Coefficient Alpha Is Probably Wrong, but Which Coefficient Omega Is Right? A Tutorial on Using R to Obtain Better Reliability Estimates. *Advances in Methods and Practices in Psychological Science*, 251524592095174. https://doi.org/10.1177/2515245920951747

Gadermann, A. M., Guhn, M., & Zumbo, B. D. (2012). Estimating ordinal reliability for likert-type and ordinal item response data: A conceptual, empirical, and practical guide. *Practical Assessment, Research and Evaluation*, *17*(3), 1–13.

Geldhof, G. J., Preacher, K. J., & Zyphur, M. J. (2014). Reliability estimation in a multilevel confirmatory factor analysis framework. *Psychological methods*, *19*(1), 72–91. https://doi.org/10.1037/a0032138

Gordon, R. (2007). A note on using stratified alpha to estimate the composite reliability of a test composed of interrelated nonhomogeneous items. *Psychological method*s. 12. 177-84. 10.1037/1082-989X.12.2.177

Graham, J. (2006). Congeneric and (essentially) tau-equivalent estimates of score reliability. What they are and how to use them. *Educational and Psychological Measurement*, *66*(6), 930–944. doi: 10.1177/0013164406288165.

Green, S. B., & Hershberger, S. L. (2000). Correlated errors in true score models and their effect on coefficient alpha. *Structural Equation Modeling*, *7*(2), 251–270. https://doi.org/10.1207/S15328007SEM0702_6

Green, S. B., Lissitz, R. W., & Mulaik, S. A. (1977). Limitations of coefficient alpha as an index of test unidimensionality. *Educational and Psychological Measurement*, *37*(4), 827–838. https://doi.org/10.1177/001316447703700403

Green, S.B., and Yang, Y. (2009a). Commentary on coefficient alpha: a cautionary tale. *Psychometrika* 74,121–135.doi:10.1007/s11336-008-9098-4

Green, S.B., and Yang, Y. (2015). Evaluation of dimensionality in the assessment of internal consistency reliability: coefficient alpha and omega coefficients. *Educ. Meas. Issues Pract* 34,14– 20. doi:10.1111/emip.12100

Green, S.B., and Yang, Y.(2009b). Reliability of summed item scores using structural equation modeling: an alternative to coefficient Alpha. *Psychometrika* 74,155–167.doi:10.1007/s11336-008-9099-3

Gregory R. Hancock & Ji An (2020) A Closed-Form Alternative for Estimating ω Reliability under Unidimensionality, Measurement: Interdisciplinary Research and Perspectives, 18:1, 1-14, DOI: 10.1080/15366367.2019.1656049

Guttman, L. (1945). A basis for analyzing test-retest reliability. *Psychometrika*, *10*(4), 255–282. https://doi.org/10.1007/BF02288892

Haldago-Tello, F. P., Chacón-Moscoso, S., Barbero-García, I., & Vila-Abad, E. (2008). Polychoric versus Pearson correlations in exploratory and confirmatory factor analysis of ordinal variables. *Quality and Quantity*, *44*(153), 26-34.

Hancock, G. & Mueller, R.O. (2001). Rethinking construct reliability within latent variable systems. *Structural Equation Modeling: Present and Future*. 195-216.

Hancock, G. R., & An, J. (2020). A Closed-Form Alternative for Estimating ω Reliability under Unidimensionality. *Measurement*, *18*(1), 1–14. https://doi.org/10.1080/15366367.2019.1656049

Hoekstra, R., Vugteveen, J., Warrens, M. J., & Kruyen, P. M. (2019). An empirical analysis of alleged misunderstandings of coefficient alpha. *International Journal of Social Research Methodology*, *22*(4), 351–364. https://doi.org/10.1080/13645579.2018.1547523

Hogan, T. P., Benjamin, A., & Brezinski, K. (2000). Reliability Methods: A Plans on the Frequency of Use of Various Types. *Educational and Psychological Measurement*, 60, 523-531. http://dx.doi.org/10.1177/00131640021970691

Hunt, T. (2013). Lambda4: Collection of Internal Consistency Reliability Coefficients. *R Package Version 3.0*. http://cran.r-project.org/package=Lambda4

Hunt, T. D., & Bentler, P. M. (2015). Quantile lower bounds to reliability based on locally optimal splits. *Psychometrika*, *80*(1), 182–195. https://doi.org/10.1007/s11336-013-9393-6

Jackson, P. H., & Agunwamba, C. C. (1977). Lower bounds for the reliability of the total score on a test composed of nonhomogeneous items. I: Algebraic lower bounds. *Psychometrika*, *42*(4), 567–578. https://doi.org/10.1007/BF02295979

Jöreskog, Karl. (1971). Statistical Analysis of Sets of Congeneric Tests. *Psychometrika*. *36*. 109-133. 10.1007/BF02291393.

Jorgensen, T. D., Pornprasertmanit, S., Schoemann, A. M., & Rosseel, Y. (2020). semTools: Useful tools for structural equation modeling. R package version 0.5-3. Retrieved from https://CRAN.R-project.org/package=semTools

Kamata, A., Turhan, A., & Darandari, E. (2003). *Estimating reliability for multidimensional composite score scale scores*. Paper presented at the Annual Meeting of the American Educational Research Association, Chicago.

Kelley, T.L. (1942). The reliability coefficient. *Psychometrika* **7,** 75–83 .https://doi.org/10.1007/BF02288068

Kuder, G.F., & Richardson, M.W. (1937). The theory of the estimation of test reliability. *Psychometrika* **2,** 151–160. https://doi.org/10.1007/BF02288391

Lewis, C. (2007). Classical Test Theory: In C.R. Rao and S Sinharay; Handbook of statistics. *Psychometrics, 26,* 29-43.

Lord, F. M. (1955). Estimating Test Reliability. Educational and Psychological Measurement, 15(4), 325–336. https://doi.org/10.1177/001316445501500401

Lord, F. M. (1959). Tests of the Same Length do Have the Same Standard Error of Measurement. *Educational and Psychological Measurement*, 19(2), 233–239. https://doi.org/10.1177/001316445901900208

Lord, F. M., & Novick, M. R. (1968). *Statistical theories of mental test scores*. Reading, MA: Addison-Wesley Pub. Co.

Lozano, L. M., García-Cueto, E., & Muñiz, J. (2008). Effect of the number of response categories on the reliability and validity of rating scales. *Methodology: European Journal of Research Methods for the Behavioral and Social Sciences, 4*(2), 73-79.

McDonald, R. P. (1978). Generalizability in factorable domains: Domain validity and generalizability. *Educational and Psychological Measurement*, *38*, 75–79.

McDonald, R. P. (1999). *Test theory: A unified treatment*. Mahwah, NJ: Lawrence Erlbaum.

McNeish, D. (2017). Thanks coefficient alpha, We'll take it from here. *Psychological Methods*, *23*(3), 412–433. https://doi.org/10.1037/met0000144

Moss, J. (2020, June 28). Please avoid the standardized alpha and the ordinal alpha. *APsyArXiv preprint,v1.0,* 1-29 https://doi.org/10.31234/osf.io/nvg5d

Muthén, B. O., Kaplan, D., & Hollis, M. (1987). On structural equation modeling with data that are not missing completely at random. *Psychometrika, 52*, 431– 462. doi:10.1007/BF02294365

Muthén, B., & Muthén, L. K. (2000). Integrating person-centered and variable-centered analyses: Growth mixture modeling with latent trajectory classes. *Alcoholism: Clinical and Experimental Research, 24*(6), 882-891. doi:10.1097/00000374- 200006000-00020

Novak, J. (2020). Measurement reliability in psychology: Method development, Cronbach's alpha coefficient infatuation, and recommendations for correct reliability assessment. *Psychological Topics, 29* (2), 427-457. https://doi.org/10.31820/pt.29.2.11

Novick, M. R., & Lewis, C. (1967). Coefficient alpha and the reliability of composite measurements. *Psychometrika*, 32, 1–13.

Novick, M. R., & Lewis, C. (1967). Coefficient alpha and the reliability of composite measurements. *Psychometricka, 32(*1), 1-13.

Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). New York, NY: McGraw-Hill.

Nunnally, J.C. (1978) *Psychometric theory*. 2nd Edition, McGraw-Hill, New York.

Oliden, P. E., & Zumbo, B. D. (2008). Coeficientes de fiabilidad para escalas de respuesta categórica ordenada [Reliability coefficients for ordinal response scales]. *Psicothema*, *20*(4), 896–901.

Oosterwijk, P. R., van der Ark, L. A., & Sijtsma, K. (2017). Overestimation of reliability by Guttman's λ4, λ5, and λ6 and the Greatest Lower Bound. In L. A. van der Ark, S. Culpepper, J. A. Douglas, W-C. Wang, & M. Wiberg (Eds.), *Quantitative psychology research: The 81th Annual Meeting of the Psychometric Society 2016* (pp. 159-172). Springer.

Osburn, H. G. (2000). Coefficient alpha and related internal consistency reliability coefficients. *Psychological Methods, 5*(3), 343–355. https://doi.org/10.1037/1082-989X.5.3.343

Padilla, M. A., & Divers, J. (2016). A Comparison of Composite Reliability Estimators: Coefficient Omega Confidence Intervals in the Current Literature. *Educational and Psychological Measurement, 76*(3), 436–453. https://doi.org/10.1177/0013164415593776

Penev, S., & Raykov, T. (2006). Maximal Reliability and Power in Covariance Structure Models. *British Journal of Mathematical and Statistical Psychology*, *59*, 75-87.

Peters, G. (2014). The alpha and the omega of scale reliability and validity: why and how to abandon Cronbach's alpha and the route towards more comprehensive assessment of scale quality. *The European Health Psychologist*, *16*(2), 56–69. https://doi.org/10.31234/osf.io/h47fv

Raykov, T. (1997a). Estimation of composite reliability for congeneric measures. *Applied Psychological Measurement*, *21*, 173-184

Raykov, T. (1997b). Scale reliability, Cronbach's coefficient alpha, and violations of essential tau-equivalence with fixed congeneric components. *Multivariate Behavioral Research*, *32*, 329-353.

Raykov, T. (2001). Bias of coefficient alpha for fixed congeneric measures with correlated errors. *Appl. Psychol. Meas. 25*, 69–76. doi: 10.1177/01466216010251005

Raykov, T. (2004). Point and Interval Estimation of Reliability for Multiple-Component Measuring Instruments via Linear Constraint Covariance Structure Modeling. *Structural Equation Modeling*, *11*, 342-356.

Raykov, T., & Grayson, D. A. (2003). A Test for Change of Composite Reliability in Scale Development. *Multivariate Behavioral Research*, *38*, 143-159.

Raykov, T., & Marcoulides, G. A. (2011). *Introduction to psychometric theory*. New York: Routledge.

Raykov, T., & Marcoulides, G. A. (2017). Evaluation of True Criterion Validity for Unidimensional Multi-Component Measuring Instruments in Longitudinal Studies. *Structural Equation Modeling, 24*, 599-606.

Raykov, T., & Shrout, P. E. (2002). Reliability of scales with general structure: Point and interval estimation using a structural equation modeling approach. *Structural Equation Modeling*, *9*(2), 195–212. https://doi.org/10.1207/S15328007SEM0902_3

Raykov, T., Gabler, S., & Dimitrov, D. M. (2016). Maximal Reliability and Composite Reliability: A Latent Variable Modeling Approach to Their Difference Evaluation. *Structural Equation Modeling, 23*, 384-391

Revelle, W. (1979). Hierarchical cluster analysis and the internal structure of tests. *Multivariate Behavioral Research*, *14*(1), 57–74. https://doi.org/10.1207/s15327906mbr1401_4

Revelle, W. (2020). psych: Procedures for Personality and Psychological Research, Northwestern University, Evanston, Illinois, USA, https://CRAN.R-project.org/package=psych Version = 2.0.12,.

Revelle, W. (2020). psychTools:Tools to Accompany the 'psych' Package for Psychological Research Northwestern University, Evanston, Illinois, USA, https://CRAN.R-project.org/package=psychTools Version = 2.0.8.

Revelle, W., & Condon, D. M. (2019). Reliability from α to ω: A tutorial. *Psychological Assessment*, *31*(12), 1395–1411. https://doi.org/10.1037/pas0000754

Revelle, W., & Zinbarg, R. E. (2009). Coefficients alpha, beta, omega, and the glb: Comments on sijtsma. *Psychometrika*, *74*(1), 145–154. https://doi.org/10.1007/s11336-008-9102-z

Robitzsch, A. (2020a). Why Ordinal Variables Can (Almost) Always Be Treated as Continuous Variables: Clarifying Assumptions of Robust Continuous and Ordinal Factor Analysis Estimation Methods. *Frontiers in Education*, *5*, 1–10. https://doi.org/10.3389/feduc.2020.589965

Robitzsch, A. (2020b). *sirt: Supplementary Item Response Theory Models*. R package version 3.9-4. https://CRAN.R-project.org/package=sirt

Rulon, P. J. (1939). A Simplified Procedure for Determining the Reliability of a Test by Split-Halves. *Harvard Educational Review, IX*, 99-103.

Schmitt, N. (1996). Uses and abuses of coefficient alpha. *Psychological Assessment, 8*, 350-353. doi:10.1037/1040-3590.8.4.350

Sijtsma, K. (2009). On the use, the misuse, and the very limited usefulness of cronbach's alpha. *Psychometrika*, *74*(1), 107–120. https://doi.org/10.1007/s11336-008-9101-0

Sijtsma, K. (2015). Delimiting coefficient alpha from internal consistency and unidimensionality. *Educational Measurement: Issues and Practice*, *34*(4), 10–13.

Sijtsma, K., and van der Ark, L.A. (2015). Conceptions of reliability revisited and practical recommendations. *Nurs. Res.* 64,128–136.doi: 10.1097/NNR.0000000000000077

Spearman, C. (1904). The Proof and Measurement of Association between Two Things. *The American Journal of Psychology, 15*(1), 72-101. doi:10.2307/1412159

Spearman, C. (1910). Correlation calculated from faulty data. *British Journal of Psychology*, 3, 271-195.

Tarkkonen,L., and Vehkalahti, K. (2005). Measurement errors in multivariate measurement scales. *J. Multivar. Anal.* 96,172–189. doi: 10.1016/j.jmva.2004.09.007

Ten Berge, J. M., Snijders, T. A., & Zegers, F. E. (1981). Computational aspects of the greatest lower bound to the reliability and constrained minimum trace factor analysis. *Psychometrika, 46*(2), 201–213. https://doi.org/10.1007/BF02293900

Ten Berge, J.M., Sočan, G. (2004). The greatest lower bound to the reliability of a test and the hypothesis of unidimensionality. *Psychometrika* 69**,** 613–625. https://doi.org/10.1007/BF02289858

Traxler, K. (2018). Estimating bias in multilevel reliability coefficients: A Monte Carlo simulation. *Dissertation Abstracts International:*

Section B: The Sciences and Engineering, *79* (8-B(E)). http://0-earch.ebscohost.com.wam.city.ac.uk/login.aspx?direct=true&db=psyh&AN=2018-26097-220&site=ehost-live

Trizano-Hermosilla, I., & Alvarado, J. M. (2016). Best alternatives to Cronbach's alpha reliability in realistic conditions: Congeneric and asymmetrical measurements. *Frontiers in Psychology*, *7*, 1–8. https://doi.org/10.3389/fpsyg.2016.00769

van der Ark, L. A., van der Palm, D. W., & Sijtsma, K. (2011). A Latent Class Approach to Estimating Test-Score Reliability. *Applied Psychological Measurement*, 35(5), 380–392. https://doi.org/10.1177/0146621610392911

VandenBos, G. R. (Ed.). (2007). *APA Dictionary of Psychology*. American Psychological Association.

Viladrich, C., Angulo-Brunet, A., & Doval, E. (2017). A journey around alpha and omega to estimate internal consistency reliability. *Anales de Psicologia*, *33*(3), 755–782. https://doi.org/10.6018/analesps.33.3.268401

Woodruff, D. & Wu, Y. (2012). Statistical Considerations in Choosing a Test Reliability Coefficient. *ACT Research Report Series*. Iowa City, Iowa

Yang, Y., & Green, S. B. (2011). Coefficient alpha: A reliability coefficient for the 21st century?. *Journal of Psychoeducational Assessment*, *29*, 377-392.

Yang, Y., & Green, S. B. (2011). Coefficient alpha: A reliability coefficient for the 21st century? *Journal of Psychoeducational Assessment*, *29*(4), 377-392.

Zinbarg, R.E., Revelle, W., Yovel, I., Li, W. (2005). Cronbach's, α Revelle's β and McDonald's ω H: Their relations with each other and two alternative conceptualizations of reliability. *Psychometrika*. *70*: 123-133. DOI: 10.1007/s11336-003-0974-7

Zinbarg, R.E., Yovel, I., Revelle, W.,& McDonald, R.P.(2006). Estimating generalizability to a latent variable common to all of a scale's indicators: A comparison of estimators for ω h. *Applied Psychological Measurement*. 30: 121-144. DOI: 10.1177/0146621605278814

Zumbo, B. D., Gadermann, A. M., & Zeisser, C. (2007). Ordinal versions of coefficients alpha and theta for likert rating scales. *Journal of Modern Applied Statistical Methods*, *6*(1), 21–29. https://doi.org/10.22237/jmasm/1177992180