الفاعلية النسبية لطريقة بييز البارامترية وطريقة تهذيب النواة اللابارامترية في تقدير معلمة القدرة وفق النموذج اللوجستي الثنائي باختلاف حجم العينة

إعـــــداد

د/ شادي يوسف الشواورة

أستاذ مساعد في القياس والتقويم/ عميد شؤون الطلبة

جامعة العقبة للتكنولوجيا، العقبة، الأردن

الفاعلية النسبية لطريقة بييز البارامترية وطريقة تهذيب النواة اللابارامترية في تقدير معلمة القدرة وفق النموذج اللوجستي الثنائي باختلاف حجم العينة

د/ شادي يوسف الشواورة*

المقدمة والإطار النظري:

يعد القياس والتقويم من أهم العناصر في العملية التدريسة وقياس نتاجات التعلم، ومن أهم وسائل القياس والتقويم لمعرفة القدرة العقلية للطلاب هى الاختبارات التحصيلية والتي تقيس بدورها تحقيق الأهداف (Popham, 2004)، ولابد من توفر دقة القياس في الاختبارات والمقاييس المختلفة، وهذا يحتاج لقوانين رياضية تحكم التأثير والعلاقة بين المتغيرات والتي بدورها تحقق الموضوعية، وهى الغاية المنشودة في القياس النفسي والتربوي، فقياس السمات الإنسانية مثل التحصيل والذكاء والميول والاستعداد والاتجاه يعتبر قياس غير مباشر؛ لأنه يستدل على السمة من خلال السلوك المرتبط بها وهذا يختلف عن القياس الفيزيائي المباشر (علام،٢٠٠٥).

تطورت نظريات القياس والتقويم بشكل متسارع؛ لتعالج مسألة الدقة في القياس فقد سادت النظرية الكلاسيكية ((Classical Test Theory(CTT) والتي كانت تعالج الدقة في القياس من خلال مفهوم الثبات، وقد تعرضت هذه النظرية (CTT) لانتقادات عدة: منها عدم الموضوعية في القياس، ويظهر ذلك من خلال عدم استقلالية خصائص الأفراد عن خصائص الفقرات. ومن الانتقادات الأخرى تأثر تقديرات القدرة بالظروف التي يطبق فيها الاختبار، ومن عيوب هذه النظرية أيضا أنها تقوم على افتراض أن تباين أخطاء القياس هو نفسه لمستويات القدرة جميعها (Hambleton & Swaminathan, 1985).

نتيجة العيوب والانتقادات لنظرية الكلاسيكية في القياس والسعي لتحقيق أعلى درجات الدقة والموضوعية في القياس النفسي والتربوي، ظهر الاتجاه الجديد الذي

* د/ شادي يوسف الشواورة: أستاذ مساعد في القياس والتقويم/ عميد شؤون الطلبة-جامعة العقبة للتكنولوجيا، العقبة، الأردن.

يعتمد على النماذج الرياضية والذي يدعى بنظرية استجابة الفقرة (Item Response Theory (IRT))، التي عالجت جوانب القصور في النظرية الكلاسيكية، حيث لم تعد إحصائيات الفقرة الاختبارية تعتمد على خصائص المفحوصين، وتقديم نماذج تربط بين احتمال الحصول على إجابة صحيحة على الفقرات بمستويات القدرة للأفراد، بالإضافة إلى أن نماذج (IRT) تقدر خطأ قياس عند كل مستوى من مستويات القدرة بعكس النماذج (CTT) التي تفترض خطأ قياس واحد (علام، ٢٠٠٥).

تنقسم نماذج نظرية استجابة الفقرة إلى نوعين رئيسيين هما: النوع الأَول نماذج استجابة الفقرة البارامترية (Parametric Item Response Theory (PIRT)) حيث يكون شكل دالة استجابة الفقرة (Item Response Function (IRF) محدداً، والنوع الثاني يعرف بنماذج استجابة الفقرة اللابارامترية(Non–Parametric Item Response Theory (NIRT) والتي لا تُحدِّد شكل دالة استجابة الفقرة، وبالرغم من الاختلافات بين هذه النماذج إلا أن النماذج البارامترية واللابارامترية تشترك في كثير من الاستخدامات التطبيقية كتطوير المقاييس والاختبارات المختلفة، وتحديد صدق المقياس. وبين فان دير لندن وهامبلتون (Van der Linden & Hambleton,١٩٩٧) أنه يمكن الافتراض أن الدوال بنماذج استجابة الفقرة اللابارامترية (NIRT) أقرب لدوال الاستجابات الواقعية من تلك التي تعطيها النماذج البارامترية (PIRT)؛ لأنها تعتمد على افتراضات أقل حول النموذج الرياضي. وهذا ما أكدته ديهاوس (Dyehouse, 2009) أن الطرق اللابارامترية طُورت لتوازي جميع الطرق البارامترية مع الاستفادة من سهولة افتراضاتها ضمن نظرية استجابة الفقرة، مما يتيح الفرصة لتحليل البيانات الحقيقية (الواقعية) على مقياس رتبي دون اعتبار أن القوانين انتُهِكَت. تعد دالة الاستجابة الفقرة للمفحوصين البارامتري واللابارمتري ركيزة أساسية، إذ تعرف هذه الدالة على أنها العلاقة بين احتمالية حصول المفحوص على الدرجة (١) (الإجابة المفتاحية) لفقرة ثنائية، والسمة الكامنة أو القدرة لذلك المفحوص (Ө). وتفترض النماذج البارامترية (PIRT) قيوداً على شكل الدالة إلا أن نظيرتها اللابارامترية لا تفترض شكلاً معيناً لهذه الدالة، في حين تشترط (NIRT) شرطا وحيداً للدالة هو أن لا

تكون متناقصة مع ازدياد مستوى القدرة، وما عدا ذلك فجميع الأشكال مقبولة (Sijtsma, 1998).

نماذج نظرية استجابة الفقرة البارامترية:

تستند نظريّة استجابة الفقرة (PIRT) على أربع افتراضات رئيسية وهى: افتراض أُحادّية البعد (Unidimentionality)، وافتراض الاستقلال الموضعي (Local Item Independence)، افتراض منحنى الاستجابة للفقرة (Item Response Curve)، افتراض عدم السرعة (Non-Speededness)، (Hambleton, Swaminathan & Rogers, 1991).

أشار بيكر (Baker, 2001) إلى النماذج الرياضية (اللوجستية) لنظرية استجابة الفقرة (PIRT) التي تختلف باختلاف معالم الفقرات المراد تقديرها،وتُعدّ هذه النماذج من أكثر نماذج نظرية استجابة الفقرة شيوعاً، وأكثرها ملاءمة مع الفقرات ثنائية التدريج، وتاليا توضيح للنماذج اللوجستية الثلاثة:

النموذج اللوجستي أحادي المعلمة (One-Parameter Logistic Model): سعى واضع هذا النموذج، العالم الدنماركي جورج راش، إلى تحقيق الموضوعية في القياس النفسي التربوي، من خلال توصله إلى أدوات قياس اتسمت باستقلاليتها عن خصائص الأفراد المراد قياس قدرتهم، فضلاً عن استقلالية قدرة الأفراد المقاسة عن أدوات القياس، ويوصف النموذج اللوجستي أحادي المعلمة بالمعادلة الآتية :

$$p_i(\theta) = \frac{e^{D\bar{a}_i(\theta-\mathbf{bi})}}{1+e^{D\bar{a}_i(\theta-\mathbf{bi})}}$$

(Pi(θ هي احتمال الإجابة الصحيحة للفرد j الذي قدرته Ѳj على الفقرة i التي صعوبتها bi حيث إن ((n, ...,i=1 , 2 ، و n هي عدد الفقرات. وأن e هي الدالة الأسية ذات الأساس الطبيعي (2.71828)، وD ثابت التدريج (Sceling Fector)، ويساوي تقريباً (1.7).

النموذج اللوجستي ثنائي المعلمة (Two-Parameter Logistic Model): يسمح النموذج ثنائي المعلمة بتقدير معلمتين من معالم الفقرات وهما الصعوبة (bi) والتمييز (ai)، وبذلك تم إضافة معلمة جديدة على نموذج راش الأحادي

الذي يقدر معلمة صعوبة الفقرة فقط، وعليه يتم تدريج الفقرات في ضوء صعوبتها وقدرتها على التمييز بين المستويات المختلفة للقدرة ، ويمثل النموذج اللوجستي الثنائي بالدالة الرياضية الآتية:

$$p_i(\theta) = \frac{e^{Dai(\theta - \mathbf{bi})}}{1 + e^{Dai(\theta - \mathbf{bi})}}$$

النموذج اللوجستي ثلاثي المعلمة (Three-Parameter Logistic Model):

من المشكلات التي واجهت النموذج اللوجستي ثنائي المعلمة أن هنالك احتمالاً لإجابة الأفراد ذوي القدرة المنخفضة على الفقرة، وتكون نسبة إجابتهم أكبر من الصفر؛ وذلك بسبب عامل التخمين في الإجابات عن الفقرات الاختبارية. لذلك قام النموذج اللوجستي الثلاثي على ثلاثة معالم وهي: معلمة الصعوبة (bi)، ومعلمة التمييز (ai)، ومعلمة التخمين (Ci)، ويحدد المعلم الأخير على منحنى خصائص الفقرة في الجزء الأسفل منه، ويدعى خط التقارب الأدنى (Lower Asymptote) (Glas, Falcon 2003). ويصاغ النموذج اللوجستي الثلاثي على النحو التالي:

$$p_i(\theta) = Ci + (1 - Ci)\frac{e^{Dai(\theta - bi)}}{1 + e^{Dai(\theta - bi)}}$$

حيث إن $pi(\boldsymbol{\theta})$: احتمال إجابة المفحوص الذي اختير عشوائياً من مستوى القدرة (θ)عن الفقرة (i) إجابةً صحيحة، bi: معلمة الصعوبة، ai: معلمة التمييز، ci: معلمة التخمين، D: عامل التدريج.

نماذج نظرية استجابة الفقرة اللابارامترية (NIRT)

أثبتت نماذج الاستجابة للفقرة اللابارامترية أنها نماذج إحصائية يمكن استخدامها لدراسة جودة وكفاءة المقياس من خلال تحليلات جتمان (Guttman scalogram)، التي تفترض أن كل فقرة وكل مفحوص يمتلكان موقعاً على متصل القدرة، حيث يجيب المفحوص على الفقرة إجابة صحيحة فقط في حال كانت قدرته أعلى من صعوبة تلك الفقرة، مما يتيح التنبؤ بدرجة المفحوص الكلية وتوقع نمط استجابته، وقد لا يتحقق ذلك تجريبياً على الدوام، مما أدى إلى ظهور

الفكرة التي جاءت بها النماذج الحديثة لنظرية الاستجابة للفقرة والقائمة على الاحتمالية لا التحديد، والتي نصت على أن احتمالية الإجابة الصحيحة مرتفعة ولكن لا تكون تامة كلما ازدادت قدرة المفحوص، كما أنها قليلة ولكن لا تكون صفراً عند انخفاض قدرة المفحوص (Linden and Hambleton, 1997).

أشار سيجتسما ومولينار (Sijtsma, K. & Molenaar, I, 2002) إلى أنّ الدرجة التي يستطيع بها النموذج شرح البيانات المُلاحظة تُحدّد عن طريق درجة الواقعيّة للافتراضات حول عمليات الاستجابة التي تؤدي إلى علامة الفقرة، والافتراضات الأربعة للنموذجين اللابارامتريين للفقرات ثنائية التدريج هي:أُحاديّة البعد (Unidimensionality) والاستقلال الموضعي (Local Independence) واطرادية دوال الاستجابة للفقرة (Monotonicity of IRFs) وعدم تقاطع دوال استجابات الفقرة (Nonintersecting IRFs)، الافتراضات الثلاثة السابقة تكفي للعديد من تطبيقات نظريّة الاستجابة للفقرة اللابارامترية (NIR) خاصّةً عندما يكون التركيز على قياس الأشخاص، لكن بعض التطبيقات تتطلب ترتيب الفقرات حسب الصعوبة، وبالتالي تكون الحاجة إلى إضافة افتراض رابع، وهذا الافتراض يقول: إنّ K من دوال الاستجابة للفقرة (IRFs) لا تتقاطع على طول متصل السّمة θ، وبشكل أكثر تحديداً عدم التقاطع يعني أنّ جميع الـ IRFs يمكن ترتيبها كما يلي :

$$p_1(\theta) \le p_2(\theta) \le ,........, \le p_k(\theta)$$

وتظهر أهمية استخدام نماذج استجابة الفقرة اللابارامترية من أنها نماذج تسمح بتقدير معالم هامه مثل مؤشرات الصعوبة، ومعاملات التدريج المختلفة، والتي تشير إلى القوة التمييزية لفقرات الاختبار والاختبار ككل. وتتضمن هذه النماذج ترتيب المفحوصين باستخدام درجاتهم على الاختبار على الرغم من وجود الخطأ العشوائي، إذ إنها تعتبر أن هذه الدرجات تعكس رتب القدرة (***θ).*** ***(Sijtsma and Molenaar, 2002).***

يـ شـ ير الـ درديـ ر، (6002) إلـ ى مـزايـ ا الـ طـرق الـ لابـ ارامـ تريـ ة مـ قارنـ ة بـ الـ طـرق الـ بارامـ تريـ ة، وهـ ي: بـ ساطة الاف تراضات و سـهولـ ة تـ حـ قـ يـ قـها، وإمـ كانـ يـة الـ تطـ بـ يق عـ لـى الـ بـ يانـات الـ و صـ فـ يـة والـ ترتـ يـ بـ يـة، و سـهولـ ة جمـع الـ بـ يانـات وتـ حـ لـ يـ لـهـا .

تستخدم نماذج استجابة الفقرة اللابارامترية في القياس النفسي والعقلي لدراسة جودة وكفاءة المقياس، من خلال تحليلات جوتمان (Guttman Scalogram) التي تفترض أنّ كل فقرة وكل مفحوص يمتلكان موقعا على متصل القدرة، ومن أ شهر نماذج الـ فقرة الـ لابـ ارامـ ترية مايـ لـي:

نموذج موكن (The Mokken Model)

نموذج موكن: هو صياغة تنطوي عليها أغلب نماذج السِّمات الكامنة اللابارامترية الحالية (Mokken & Lewis, 1982)، وهو يفترض أنّ السّمة الكامنة أُحادّية البعد موجودة على متصل، ويفترض أيضا أنّ كلَّ شخص يمتلك قيمة معيّنة غير معروفة على هذا المتصل، والتي تمثّل بمَعْلَمة القدرة (θ)يعمل نموذج موكن على ترتيب الفقرات والأشخاص، حيثُ تُرتّب الفقرات وفقا للزِّيادة بنسبة الاستجابة الصحيحة والأشخاص وفقا للزِّيادة بالعلامة الكليّة.

معامل التدريج (Coefficients Of Scalability):

مُعامِلات التدريج عديدة هي H_{ij} لكل زوج من الفقرات في التدريج، H_i لكل فقرة في التدريج بالنسبة لباقي الفقرات. و Hمُعامِل تدريج لجميع الفقرات.

مُعامِل التدريج H_{ij} لكل زوج من الفقرات يعرف كما يلي

$$H_{ij} = \frac{P_{ij} - P_i P_j}{(1 - P_j) P_i}$$

حيثُ $P_i < P_j$

P_{ij} : نسبة الأشخاص الذين استجابوا على الفقرتين معا i,j استجابة صحيحة.

P_i : نسبة الأشخاص الذين استجابوا على الفقرة i استجابة صحيحة.

P_j : نسبة الأشخاص الذين استجابوا على الفقرة j استجابة صحيحة.

معايير التدريج (Criteria For Scalability)

ذكر كنجما وتينفيرجت (Kingma & Tenvergert, 1985) أنّ مجموعة من الفقرات تشكل معا تدريج موكن إنْ حققت الشروط الآتية:

١. منحنيات خصائص الفقرة (ICCs) يجب أن تحقق افتراضات الاطراديّة المضاعفة (Double Monotony)

٢. جميع معاملات H_{ij} يجب أن تكون أكبر من الصفر.

٣. جميع معاملات تدريج الفقرة H_i يجب أن تكون أكبر من أو تساوي ثابت موجب C يدعى ثابت تعريف

التدريج (scale- clefining constant)، حيثُ ($0<C<1$). قدم موكن تصنيفات للتدريج من حيثُ قيمة معامل التدريج H كمايلي: تدريج قوي $H > 0.5$ وتدريج متوسط $0.4 \leq H \leq 0.5$ وتدريج ضعيف $0.3 \leq H \leq 0.4$ أما الغير قابل للتدريج عندما $H < 0.3$، (Kingma & Tenvergert, 1985).

طرق تقدير معلمة القدرة البارامترية:

يوجد عدة طرق لتقدير معلمة القدرة وفق نظرية استجابة الفقرة البارامترية منها طريقة الارجحية العظمى (Maximum Likelihood Estimation)، وطريقة بييز (Bayesian Modal Estimation)، فقد أشار هامبلتون وسوامينثان (Hambeleton & Swaminithan, 1985) إلى طريقة الأرجحية العظمى للتقدير، وتقوم هذه الطريقة على إيجاد تقدير معلمة القدرة من خلال إجراءات تعظيم الاحتمالية لمعلمة القدرة عندما يتوافر لدينا معلومات عن عينة الأفراد.وهناك طريقة بييز (Bayes) لتقدير معلمة القدرة في الاستدلال الإحصائي التي تعود للعالم البريطاني توماس بييز(Thomas Bayes)، (Bellhouse, 2004) وتعتبر هذه الطريقة من الأساليب الإحصائية المستخدمة لتقدير معالم الفقرات والقدرة في نماذج نظرية استجابة الفقرة البارامترية، إضافة إلى طريقة الأرجحية العظمى سابقة الذكر، وفي طريقة بييز يتم اعتبار المعلمة التي سيتم تقديرها بأنها متغير عشوائي تتبع توزيع احتمالي معين، أو ما يطلق عليه التوزيع القبلي (Prior Distribution)، وذلك من خلال الأخذ بعين الاعتبار المعلومات السابقة للمعلمة غير المعروفة في عملية التقدير، ويتم تحديد المعلومات السابقة عادة من قبل الباحث معتمداً على معتقد شخصي أي بناءً على خبرته السابقة، أو من خلال الخصائص الإحصائية لذلك المعلم الذي سيتم تقديره (lord,1986) .Swaminithan & Gifford,1982))

ولمعرفة أي الطرق البارامترية أدق في تقدير معالم الفقرات والقدرة، أجريت عدة دراسات منها دراسة أجراها الشواورة (٢٠١٣) لمقارنة طريقتين من الطرق البارامترية في دقة تقدير معالم الفقرات وقد أظهرت طريقة بييز (Bayesian Modal Estimation) أفضلية في دقة التقدير مقارنة مع طريقة الارجحية العظمى (Maximum Likelihood Estimation) في ظروف مختلفة في عدد

الفقرات وحجم العينة والنموذج اللوغاريتمي المستخدم، ويعزي سبب الأفضلية (Bayes) لمراعاة التوزيعات السابقة.

أشار هامبلتون وسوامينثان(Hambeleton & Swaminithan, 1985) لمعادلة بييز على النحو الآتي:

$$f(\theta,a,b,c|u)\infty L(u|\theta,a,b,c)\left\{\prod_{i=1}^{n} f(a_i)f(b_i)f(c_i)\right\}\prod_{a=1}^{N} f(\theta_a)$$

وتشير($f(\theta a)f(ai)f(bi)f(ci)$) إلى التوزيعات البعدية للمعالم، حيث أن: قدرة الأفراد(θa)، وتمييز الفقرة (ai)، وصعوبة الفقرة (bi)، وتخميين الفقرة (ci)، ويمكن إضافة أو حذف المعلم بناءً على النموذج اللوجستي المستخدم، ويشترط تحديد توزيعات القدرة (θ) لإيجاد معالم الفقرة باستخدام طريقة بييز، ويتم إيجاد تقديرات قيم المعالم ((θ , a , b ,c التي تعظم التوزيع اللاحق المشترك، وبذلك نحصل على تقدير لمعالم الفقرات والقدرة. وستعتمد هذه الدراسة على طريقة بييز (Bayes) في تقدير معالم الفقرة البارامترية باستخدام برنامج (WinBUGS).

طرق تقدير معلمة القدرة اللابارامترية:

يشير سيجتسما ومولينار (٢٠٠٢) أن هناك نماذج لنظريّة استجابة الفقرة اللابارامترية (Two Nonparametric Item Response Models) منها نموذج التجانس الاطرادي (The Monotone Homogeneity Model (MHM)، والذي يعتمد على افتراض أُحاديّة البعد والاستقلال الموضعي والاطراديّة. يصف نموذج التجانس الاطرادي (MHM) بيانات الاستجابة للفقرة الناتجة من خلال مجموعة من الفقرات المتجانسة (أُحاديّة البعد) التي تمتلك دوال الاستجابة للفقرة (IRFs) تكون مطّردة (monotonically) ذات الصّلة بالسّمة الكامنة. وهناك نموذج أخر هو نموذج الاطراديّة المضاعفة (The Double Monotonicity Model (DMM)) ويعتمد هذا النموذج على افتراضات نماذج استجابة الفقرة اللابارامترية الأربعة وهي: أُحاديّة البعد، الاستقلال الموضعي، الاطراديّة وعدم التقاطع لدوال الاستجابة للفقرة، يُعتَبر نموذج الاطراديّة المضاعفة (DMM) حالة خاصة من نموذج التجانس الاطرادي (MHM)، وتقدر دالة استجابة الفقرة في الانحدار اللابارامتري دون افتراض أن يكون شكلها لوجستي

كما هو الحال في نماذج الاستجابة للفقرة البارامترية، وقد أشار دوجلس (Douglas, 1997) إلى الانحدار اللابارامتري (Non-Parametric Regression) في تقدير معالم الفقرات والقدرة وله طريقتين لتقدير شكل دالة استجابة الفقرة وهما: طريقة تهذيب النواة ((Kernel_Smoothing(KS)) وطريقة ((Isotonic Regression Estimation (IRE))، وستعتمد هذه الدراسة على طريقة (KS) في تقدير معلمة القدرة باستخدام برنامج (TESTGRAF).

أشار دوجلس (Douglas, 1997) ورامسي (Ramsay, 2000) إلى طريقة (KS) بحساب دالة استجابة الفقرة $p_i(\theta_q)$ والذي يدل على احتمالية الإجابة الصحيحة على الفقرة i عند مستوى القدرة q من خلال المعادلة الآتية:

$$p_i(\theta_q) = \sum_{a=1}^{N} w_{aq} y_{ima}$$

$$w_{aq} = \frac{k\left[\dfrac{\theta_a - \theta_q}{h}\right] y_i}{\sum_{b=1}^{N} k\left[\dfrac{\theta_a - \theta_q}{h}\right]}$$

حيث إن:

w_{aq} : متجة القدرة للمفحوص a عند مستوى القدرة q، والذي يتم تقديره تبعا لرتبة المفحوص a مع رتب باقي المفحوصين.

y_{ima} : متجة خيار الفقرة الثنائي بطول يساوي N (عدد المفحوصين)، والذي يأخذ القيمة ١ في حال اختار المفحوص a الخيار m .

K: دالة كيرنل والتي يمكن تقديرها بعدة طرق باستخدام برنامج (TESTGRAF).

h : معلم التهذيب (Smoothing Parameter) وهو يعتمد بشكل أساسي على عدد المفحوصين ويساوي ($1.1N^{1/5}$) في برنامج (TESTGRAF).

أجرى عدد من الباحثين دراسات لتقديرات النماذج (NIRT,PIRT)، حيث قام لورد (Lord, 1986) بدراسة؛ لتوضيح طريقة (Bayes) في تقدير معالم

الفقرات والقدرة مقارنة مع الطرق البارامترية الأخرى، باختلاف طول الاختبار (40,15,10) فقرة، وأشار إلى أن تقديرات بييز المعتمدة على التقديرات السابقة تقلل من متوسط مربع الخطأ (MSE) للتقدير مقارنة مع قيمة الأرجحية العظمى، ويعود سبب ذلك إلى معرفتنا المسبقة بالمعالم، من هنا تظهر أهمية طريقة بييز للاختبارات الواقعية حيث يمكن الحصول على التوزيعات السابقة لاختبارات تم تطبيقها بشكل متكرر على المختبرين سنة بعد سنة، حيث يصبح من الممكن استنتاج توزيعات سابقة لمعالم الفقرة والقدرة من نتائج سابقة وفي هذه الحالة يمكن أن تعطي إجراءات بييز تقديرات أفضل من الطريق البارمترية الأخرى مثل طريقة الأرجحية العظمى.

في دراسة أجرها ميجر وسيجتسما وسميد (Meijer, Sijtsma & Smid, 1990) هدفت إلى المقارنة بين النماذج اللابارامترية (نموذجي موكن للتجانس الاطرادي، ونموذج الاطراديّة المضاعفة)، ونموذج راش البارامتري لمعرفة خصائص التدريج ودقّة المقياس حيثُ تمّ تطبيق اختبار الذكاء اللفظي الهولندي(DVIT) المكون من (40) فقرة على عينة من طلبة المرحلة الثانوية، وأشارت نتائج إجراءات حسن المُطابقة أفضلية نموذج موكن للتجانس الاطرادي، بينما تطابق عدد أقل من الفقرات مع النموذجين الاطرادي المضاعف اللابارامتري، وأحادي المَعْلَمة البارامتري، كما خلصت الدراسة أيضا إلى إمكانيّة اعتبار نموذج موكن للتجانس الاطرادي اللابارامتري بديلا مناسبا لتطبيقات الاختبار الأساسية، إلا أنّ بعض التطبيقات الأكثر تعقيدا كمعادلة الاختبار، والاختبارات التكيفية، تتطلب استخدام نماذج بارامترية.

كما قام كل من كونينغ وسيجتسما وهامرز (Koning, Sijtsma & Hamers, 2002) بدراسة قارنت بين نموذجين بارامتريين وآخرين لابارامتريين من نماذج استجابة الفقرة، بهدف مقارنة نواتج التحليل باستخدام النماذج اللابارامترية مع نظريتها البارامترية. وقد قام الباحثون بتطبيق اختبار للاستنتاج الاستقرائي على عينة مكونة من (487) طالبًا وطالبة من طلبة الصف الثالث الأساسي. وقد تم استخدام نموذج راش البارامتري، ونموذج فيرهيلست البارامتري (Verhelst model)، ونموذج التجانس الاطرادي، ونموذج موكن المضاغف الاطرادي. وأظهرت النتائج أفضلية الجمع بين النوعين من النماذج البارامترية واللابارامترية،

إذ قدمت النماذج اللابارامترية تدريجات رتبيّة للفقرات والأفراد، كما قدمت النماذج البارامترية معلومات مفيدة حول خصائص الفقرات بالإضافة لفائدتها في بعض الجوانب التطبقية كمعايرة درجات الاختبار والاختبارات التكيفيّة.

أجرى ميجر وبانيك (Meijer & Baneke, 2004) دراسة هدفت إلى توضيح فوائد استجابة الفقرة اللابارامترية في بناء وتحليل مقاييس الشخصيّة والمعالجة النفسيّة والاختبارات، حيثُ ناقش الباحثان قابليّة تطبيق نماذج الفقرة اللابارامترية في بناء وتحليل مقاييس الشخصيّة والمعالجة النفسيّة واختلاف هذه النماذج مع نماذج الاستجابة للفقرة البارامترية، ولغايات جمع البيانات تم استخدام وتحليل بيانات من الترجمة الهولندية الرسمية لمقياس مينيسوتا متعدد الأوجة لقياس الشخصيّة للمراهقين (MMPI-2) والذي يتألف من (33) فقرة تقيس مستويات مختلفة من الاكتئاب، وتضمنت العيّنة (439) فرداً ، وأظهرت نتائج الدراسة أنّه عبر استخدام النماذج اللابارامترية لنظريّة الاستجابة للفقرة يمكن الحصول على معلومات حول الفقرات الأكثر صعوبة مقارنة بالنماذج البارامترية. كما أظهرت النتائج أنّ النماذج اللابارامترية مفيدة في استكشاف بنية البيانات، وقد اوصى الباحثون عند تحليل بيانات الشخصيّة والعلاج النفسي استخدام نماذج نظريّة الاستجابة للفقرة اللابارامترية.

أجرى كل من سيجتسما وايمونز وبوميستر ونكليشك ورودرا (Sijtsma, Emons, Boumeester, Nyklicek & Rodra, 2007) دراسة هدفت للكشف عن قدرة نموذج موكن الاطرادي اللابارمتري في تقويم وإنشاء تدريجات أحادية البعد مأخوذه من مقياس متعدد الأبعاد هو مقياس جودة ورفاهية الحياة لمنظمة الصحة العالمية، لمقارنتها مع نماذج نظرية استجابة الفقرة البارامترية، وأظهرت النتائج أن النموذج اللابارامتري للتجانس الاطرادي لموكن كان أكثر مطابقة لبيانات مقياس جودة ورفاهية الحياة.

قام فو (Fu, 2010) بدراسة تهدف إلى تعرف دقة تقدير معلمة القدرة ومعلمة صعوبة الفقرة باستخدام (5) نماذج من نظرية استجابة الفقرة البارامترية واللابارامترية، وكانت النماذج المستخدمة في إطار الدراسة متباينة من حيث مستوى التخمين، وأحجام العينة المستخدمة، وطول الاختبار. وتم في هذه الدراسة توليد مجموعة من الاستجابات بلغت (50) مجموعة من البيانات المولدة باستخدام ظروف اختبار مختلفة. وأشارت نتائج الدراسة إلى أن هناك تبايناً في دقة تقدير

معالم الفقرة والمفحوصين حسب مستوى التخمين الموجود في الاختبار، وحجم العينة المستخدمة، وطول الاختبار. كما أشارت النتائج أيضا إلى أن دقة تقدير معلمة القدرة، ومعالم الفقرة تعتمد على معيار الدقة المستخدم في كل واحد من نماذج استجابة الفقرة.

قام القيسي (٢٠١٣) بدراسة لمقارنة طريقة الارجحية العظمى البارامترية وطريقة تنعيم النواة اللابارامترية، باختلاف حجم العينة (١٠٠، ٢٥٠، ٥٠٠، ١٠٠٠) فرد، وعدد فقرات الاختبار(٢٠، ٤٠، ٦٠)، وللإجابة عن اسئلة الدراسة، استخدم برمجية BILOG–MG؛ لتقدير معالم الفقرة والقدرة، باستخدام طريقة الأرجحية العظمى الهامشية البارامترية، واستخدم برمجية TESTGRAF؛ لتقدير معالم الفقرة والقدرة باستخدام طريقة تنعيم النواة اللابارامترية، وأظهرت النتائج أن المعالم المقدرة بالطريقة البارامترية كانت أفضل من المعالم المقدرة وفق الطريقة اللابارامترية.

مشكلة الدراسة وأسئلتها:

تكمن مشكلة الدراسة في الكشف عن أثر تباين حجم العينة في دقة تقدير معلمة القدرة باختلاف طريقة التقدير(Bayes البارامترية وKS اللابارامترية) لنماذج نظرية استجابة الفقرة، وتحددت مشكلة الدراسة بالإجابة عن الأسئلة الآتية:

السؤال الأول: "هل تختلف متوسطات دقة تقدير معلمة القدرة وفقًا لطريقة التقدير (Bayes البارامترية وKS اللابارامترية) باختلاف حجم العينة؟"

السؤال الثاني: "ما أثر طريقة التقدير (Bayes البارامترية وKS اللابارامترية) في الكفاءة النسبية لتقدير معلمة القدرة في ضوء اختلاف حجم العينة؟"

السؤال الثالث: "هل توجد فروق دالة إحصائيا عند مستوى الدلالة (α=٠.٠٥) في قيم معاملات الارتباط بين معلمة القدرة (θ) الحقيقية وبين القدرة(θ) المقدرة وفق طريقتي التقدير (Bayes البارامترية وKS اللابارامترية) باختلاف حجم العينة؟"

أهمية الدراسة:

تأتي أهميّة هذه الدراسة في محاولتها الكشف عن أثر حجم العينة في دقة تقدير معلمة القدرة وفق نماذج نظرية الفقرة البارامتري واللابارامتري، إذ أن غاية عملية القياس النفسي والتربوي تكمن في دقة تقديرات معالم الفقرات والقدرة، حيث

تتأثر دقة التقدير بالنماذج الرياضية المستخدمة والتي تقوم على افتراضات تميز بعضها عن الآخر. ويلعب حجم العينة دوراً في دقة القياس، لذلك جاءت هذه الدراسة استكمالا لدعم الجهود البحثية في هذا المجال.

مصطلحات الدراسة:

دقة التقدير(Accuracy of Estimation): هو تعبير يشير إلى جودة التقدير، من خلال اقتراب قيمة معلمة القدرة من القيمة الحقيقية، حيث يمكن الوصول إلى ذلك باختيار التقدير الذي يتصف تباينه بأنه أقل تبايئًا من أي تقدير آخر ويقاس باستخدام مربعات الأخطاء للتقدير، أو الخطأ المعياري في التقدير، أو الفعالية النسبية، أو عدم التحيز.

معلمة القدرة ((Ability(θ)): قيم يتم تقديرها بتعظيم ارجحية Maximum Likelihood Function استجابة الافراد على فقرات المقاييس أو الاختبارات التي استجابوا اليها.

نماذج نظرية استجابة الفقرة البارامترية (PIRT): هي النماذج التي تحدد شكل دالة استجابة الفقرة، وتتطلب افتراضات أكثر تشدداً، بالإضافة إلى أنها لا تصلح إلا للبيانات الفئوية، ومن البرنامج المستخدمة لهذه النماذج برنامج (WinBUGS) الذي يستخدم طريقة بييز في التقدير (Bayesian estimation).

نماذج نظرية استجابة الفقرة اللابارامترية (NIRT): هي النماذج التي لا تحدد شكل دالة استجابة الفقرة، والتي تتطلب افتراضات أقل تشدداً، وتصلح للبيانات الفئوية والرتيبة، ومن البرامج المستخدمة لهذه النماذج برنامج (TESTGRAF) والذي يطبق طريقة تهذيب النواة في التقدير (Kernel Smoothing (KS)).

حجم العينة (sample size): تعني عدد الأفراد الذين سيطبق عليهم المقياس أو الاختبار النفسي أو التربوي، وفي هذه الدراسة سيتم استخدام ثلاث مستويات للعينة وهى (1500 ,750, 250) فردا.

الكفاءة النسبية ((Relative Efficiency(RE)): هو حاصل قسمة تباين معلمة القدرة (θ (بطريقة (Bayes) البارامترية على تباين القيم المقدرة لمعلمة القدرة (θ) بطريقة (KS) اللابارامترية، فإذا زاد ناتج القسمة عن 1 تكون الكفاءة

النسبية لصالح (KS) وإذا كان ناتج القسمة أقل من 1 صحيح تكون الكفاءة النسبية لصالح (Bayes)، والصيغة الرياضية للكفاءة النسبية على النحو الآتي:

$$RE_{\theta} = \frac{V\ \hat{\theta}_{Bayes}}{V\ \hat{\theta}_{KS}}$$

الإحصائي (RMSE): هو الجذر التربيعي لمعدل مربعات الخطأ (Root Mean Square Error(RMSE))، بين القيمة المقدرة لمعلمة القدرة (θ) والقيمة الحقيقية لمعلمة القدرة (θ) ، وهو مؤشر لدقة تقدير المعالم، حيث يتم إيجاد القيمة للفرق بين قيم المولدة (الحقيقية) والقيم المقدرة بأي طريقة من
طرق التقدير، وفقا للمعادلة الآتية:

$$RMSE\,\theta = \sqrt{\frac{\sum_{i=1}^{n} (\hat{\theta}_{estmated} - \theta_{true})^2}{n}}$$

محددات الدراسة:

- التوزيع القبلي لمعلمة القدرة (θ) يتوزع توزيعا طبيعيًا.
- تتحدد نتائج هذه الدراسة بالأسلوب الذي يستخدمه الباحث والقائم على توليد البيانات باستخدام برنامج (WinGen v1.4).
- سيتم تقدير معلمة القدرة (θ (بطريقة (Bayes) البارامترية باستخدام برنامج (WinBUGS).
- سيتم تقدير معلمة القدرة (θ) بطريقة (KS) اللابارامترية باستخدام برنامج (TESTGRAF).

منهجية الدراسة وإجراءاتها:

تم استخدام البيانات المولدة (أسلوب المحاكاة) وفق طريقة مونت كارلو (Monte Carlo Methods (MCM، حيث تم توليد (50) فقرة ثنائية التدريج وفق مستويات العينة (250 ,750, 1500) فردا تبعا للنموذج اللوجستي الثنائي باستخدام برنامج (WinGen) وهو من تصميم كل من هان وهامبيلتون (Han& Hambleton,2007)، إذ تم توليد قيم معلمة القدرة (θ) وفقا للتوزيع الطبيعي (0,1) ~normal Distribution، ويظهر الجدول (١) الوسط الحسابي والانحراف المعياري لمعلمة القدرة (θ) المولدة.

جدول (١) المتوسطات الحسابية والانحرافات المعيارية
لمعلم القدرة المولدة وفقاً لمتغير حجم العينة

الإحصائي	حجم العينة		
	250	750	1500
الوسط الحسابي	0.00	0.02	0.06
الانحراف المعياري	0.90	0.96	0.99

تقدير معلمة القدرة:

اعتمدت هذه الدراسة في تقدير معلمة القدرة وفق نماذج (PIRT) بطريقة بييز على برنامج (WinBUGS v 1.4). أما تقدير معلمة القدرة وفق نماذج (NIRT) بطريقة تهذيب النواة تم باستخدام برنامج (TESTGRAF)، وهو برنامج متاح بشكل مجاني كما اشار دوغلس (Douglas, 1997).

برنامج WinBUGS v 1. 4:

تم تقدير معلمة القدرة للأفراد وفق طريقة (Bayes) البارامترية باستخدام برنامج (WinBUGS v 1. 4)، (Cowls, 2004)، إن البرمجيات التي تستخدم هذه الطريقة في تقدير معالم الفقرات والقدرة قليلة جداً، مقارنة مع البرمجيات التي تستخدم طريق الأرجحية العظمى، ويعود سبب ذلك إلى الإجراءات المعقدة لتقديرات بييز الاحتمالية، حيث تم تمرير قيم معلمة القدرة المولدة (الحقيقية) وفق النموذج الثنائي ومستويات حجم العينة الثلاث (1500 ,750, 250) لبرنامج (WinBUGS v 1. 4)، ويظهر الجدول رقم (2) ملخصا لقيم معلمة (θ) المقدرة بطريقة (Bayes).

برنامج TESTGRAF:

تم تقدير معلمة القدرة للأفراد وفق طريقة (KS) اللابارامترية باستخدام برنامج (Testgraf)، حيث تم تمرير قيم معلمة القدرة المولدة (الحقيقية) وفق النموذج الثنائي بمستويات حجم العينة الثلاث (1500 ,750, 250) لبرنامج (Testgraf) للحصول غلى التقديرات المناسبة، ويظهر الجدول رقم (2) ملخصا لقيم معلمة القدرة (θ) المقدرة بطريقة (KS).

عرض النتائج ومناقشتها:

النتائج المتعلقة بالسؤال الأول: هل تختلف متوسطات دقة تقدير معلمة القدرة وفقًا لطريقة التقدير (Bayes البارامترية وKS اللابارامترية)، باختلاف حجم العينة؟ للإجابة عن هذا السؤال تم تقدير قيم القدرة باستخدام البرنامجين المستخدمين Testgraf وWinBUGS، ويظهر الجدول رقم (٢) ملخص قيم معلمة القدرة المقدرة وفق طريقتي التقدير (Bayes وks) حيث تراوح الوسط الحسابي وفق طريقة Bayes باختلاف حجم العينة بين −0.007 و−0.033، وانحراف معياري تراوح بين 1.18 و1.22، أما طريقة التقدير ks فقد تراوحت قيم الوسط الحسابي لمعلمة القدرة باختلاف حجم العينة بين 0.011 و 0.078 وانحراف معياري بين 0.82 و1.37.

جدول (٢) الأوساط الحسابية والانحراف المعياري لتقديرات معلمة القدرة وفقاً لطريقة التقدير (Bayes البارامترية وKS اللابارامترية)، وباختلاف حجم العينة

حجم العينة	الإحصائي	طريقة التقدير			
		Bayes / WinBUGS		Ks / Testgraf	
		الوسط الحسابي	الانحراف المعياري	الوسط الحسابي	الانحراف المعياري
250	θ	−0.007	1.18	0.078	0.82
750	θ	−0.004	1.18	0.035	1.37
1500	θ	−0.033	1.22	0.011	1.33

يظهر الجدول (٣)، أن القيم المطلقة لقيم المتوسط الحسابي لمؤشر التحيز BIAS لتقديرات القدرة حسب النموذج البارامتري Bayes قد كانت أصغر ظاهرياً من قيم المتوسط الحسابي لمؤشر التحيز لتقديرات القدرة حسب النموذج اللابارامتري Ks عندما يكون حجم العينة (250 , 750)، باستثناء ظرفية واحده عندما كان حجم العينة (1500)، مما يعني أن النموذج البارامتري قد كان أكثر دقة ظاهرياً في تقدير معلمة القدرة مقارنة مع النموذج اللابارامتري، وخاصة عندما تكون حجم العينة صغيرة نسبيا حيث جاء المتوسط الحسابي لمؤشر التحيز لتقديرات القدرة وفقاً للنموذج البارامتري أقرب ما يمكن من الصفر. وجاءت قيم الانحراف المعيار لمعلمة القدرة وفق طريقة Bayes أصغر ظاهريا مما يعني أنها أكثر دقة مقارنة بطريقة KS، ويُعزي الباحث ذلك إلى أثر مراعاة التوزيعات السابقة وفق طريقة Bayes في دقة تقدير معالم الفقرات والقدرة، وخاصة عند تقدير المعلمة لحجوم العينات الصغيرة، وقد يعزى السبب أيضا إلى أن توزيع

معلمة القدرة المولدة أقرب للقياس الفئوي منه إلى الرتبي، حيث إنها تتوزع توزيعا طبيعيا (0, 1)، وهو افتراض من افتراضات طريقة التقدير البارامترية Bayes وهذه النتيجة متفقة مع ما أشار إليه لورد (Lord, 1986) وكونينغ وسيجتسما وهامرز (Koning, Sijtsma & Hamers, 2002) وفو (Fu, 2010).

جدول (٣) لمؤشر التحيز لتقديرات معلمة القدرة وفقاً لطريقة التقدير (Bayes البارامترية وKS اللابارامترية)، وباختلاف حجم العينة.

حجم العينة	الإحصائي	التحيز في القدرة باختلاف طريقة التقدير			
		Bayes / WinBUGS		Ks / Testgraf	
		الوسط الحسابي	الانحراف المعياري	الوسط الحسابي	الانحراف المعياري
250	θ	0.014	0.57	0.095	0.74
750	θ	0.009	0.55	0.043	0.71
1500	θ	0.055	0.57	0.014	0.64

النتائج المتعلقة السؤال الثاني: "ما أثر طريقة التقدير (Bayes البارامترية وKS اللابارامترية) في الكفاءة النسبية لتقدير معلمة القدرة في ضوء اختلاف حجم العينة ؟" للإجابة عن سؤال الدراسة الثاني، فقد تم حساب التباين للقيم المقدرة لمعلمة القدرة (θ) وفقاً لطريقتي التقدير (Bayes وKS) باختلاف حجم العينة، ثم تمت قسمة تباين القيم المقدرة لمعلمة القدرة (θ) وفقاً لطريقة التقدير Bayes على تباين القيم المقدرة لمعلمة القدرة (θ) وفقاً لطريقة التقدير KS؛ كمؤشر كفاءة نسبية في تقدير معلمة القدرة (θ) في المواقف المختلفة وفقاً لطريقتي التقدير والتي تأخذ الصيغة الرياضية الآتية:

$$RE_{\theta} = \frac{V\,\hat{\theta}_{Bayes}}{V\,\hat{\theta}_{KS}}$$

وللوقوف على الكفاءة النسبية لكلا الطريقتين تم إيجاد تباين معلمة القدرة وفق طريقة Bayes وتباين معلمة القدرة وفق طريقة KS، ومن ثم الحصول على مؤشر الكفاءة النسبية وفق الجدول (٤).

جدول (٤) الكفاءة النسبية لطريقة BAYES البارامترية على طريقة KS اللابارامترية

الإحصائي		حجم العينة		
		250	750	1500
معلمة القدرة (θ)	تباين معلمة القدرة وفق طريقة BAYES	0.490	0.950	0.532
	تباين معلمة القدرة وفق طريقة KS	1.170	1.447	0.483
	الكفاءة النسبية لـ BAYES على KS	0.418	0.656	1.101

يلاحظ من الجدول (٤)، أن قيم الكفاءة النسبية لطريقة Bayes على ks قد كانت ظاهرياً لصالح طريقة التقدير Bayes عندما يكون حجم العينة (750,250) في حين كانت لصالح طريقة ks عندما يكون حجم العينة (١٥٠٠)، مما يعني أن النموذج البارامتري قد كان أكثر كفاءة ظاهرياً في تقدير معلمة القدرة مقارنة مع النموذج اللابارامتري في معظم ظرفيات الدراسة، وتعزى أفضلية طريقة Bayes إلى أن توزيع معلمة القدرة المولدة أقرب للقياس الفئوي منه إلى الرتبي، حيث إنها تتوزع توزيعا طبيعيا (1 , 0)، وهو افتراض من افتراضات طريقة التقدير البارامترية Bayes، ومراعاة هذه الطريقة للتوزيعات السابقة للمعالم عند عملية التقدير، في حين أن قيم القدرة التي تم الحصول عليها وفق طريقة KS اللابارامترية كانت على المقياس الرتبي. وهذه النتيجة متفقة مع ما أشار إليه لورد (Lord, 1986) وكونينغ وسيجتسما وهامرز (Koning, Sijtsma & Hamers, 2002 و **فو** (Fu, 2010) والقيسي(٢٠١٣).

وللتأكد من دقة القياس تم حساب مؤشر RMSE لمعلمة القدرة (θ) وفقا لطريقة التقدير (Bayes البارامترية وKS اللابارامترية)، وباختلاف حجم العينة، وذلك بإيجاد الأوساط الحسابية والانحرافات المعيارية الخاصة بمؤشر RMSE كمؤشر دقة التقدير لمعلمة القدرة (θ) ويظهر الجدول (٥) النتائج التي تم الحصول عليها باستخدام المعادلة الآتية:

$$RMSE\,\theta = \sqrt{\frac{\sum_{i=1}^{n} (\hat{\theta}_{estmated} - \theta_{true})^2}{n}}$$

الجدول (٥) الأوساط الحسابية لمؤشر RMSE في دقة تقدير معلمة القدرة (θ) وفقا لطريقة التقدير(Bayes البارامترية و KS اللابارامترية)، وباختلاف حجم العينة

Testgraf Ks	WinBUGS Bayes	مؤشر دقة التقدير RMSE لمعلمة القدرة وفقاً لطريقة التقدير	حجم العينة
0.732	0.536	θ	250
0.810	0.697	θ	750
0.403	0.497	θ	1500

يلاحظ من الجدول (٥)، أن قيم مؤشرRMSE لتقديرات القدرة (θ) والمقدرة وفق طريقة Bayes البارامترية قد كانت أصغر ظاهريا من قيم مؤشرRMSE لتقديرات القدرة (θ) والمقدرة وفق طريقة KS اللابارامترية في معظم ظرفيات الدراسة، مما يعني أن طريقة Bayes البارامترية قد كان أكثر دقة ظاهرياً في

تقدير معلمة القدرة (θ) مقارنة النموذج اللابارامتري ممثل في طريقة KS، حيث جاء المتوسط الحسابي لمؤشر RMSE لتقديرات القدرة وفقاً للنموذج البارامتري وفق طريقة Bayes أقل ما يمكن، باستثناء قيم مؤشر RMSE لتقديرات القدرة حسب النموذج البارامتري قد كانت أكبر ظاهرياً من قيم مؤشر RMSE لتقديرات القدرة حسب النموذج اللابارامتري عندما كان حجم العينة ١٥٠٠ فردا، ويعزى سبب أفضلية طريقة Bayes البارامترية إلى أن قيم القدرة التي تم الحصول عليها وفق طريقة KS اللابارامترية كانت على المقياس الرتبي مقارنة مع طريقة Bayes البارمترية والتي تكون قيمها على المقياس الفئوي. وجاءت هذه النتيجة متفقة مع دراسة القيسي(٢٠١٣).

النتائج المتعلقة بالسؤال الثالث: "هل توجد فروق دالة إحصائيا عند مستوى الدلالة (α=0.05) في قيم معاملات الارتباط بين معلمة القدرة (θ) الحقيقية وبين القدرة (θ) المقدرة وفق طريقتي التقدير (Bayes البارامترية و KS اللابارامترية) باختلاف حجم العينة؟"

للإجابة عن سؤال الدراسة الثالث، فقد تم حساب قيم معاملات الارتباط بين معلمة القدرة (θ) الحقيقية وبين معلمة القدرة (θ) وفق طريقتي التقدير التقدير (Bayes البارامترية و KS اللابارامترية) باختلاف حجم العينة، ثم تم استخدام اختبار t لمعاملات الارتباط للعينات المترابطة للكشف عن جوهرية الفرق في قيمتي (معامل الارتباط بين معلمة القدرة الحقيقية وبين معلمة القدرة المقدرة وفق طريقة التقدير Bayes البارامترية) من جهة و(معامل الارتباط بين معلمة القدرة الحقيقية وبين معلمة القدرة المقدرة وفق طريقة التقدير KS) من جهة أخرى، وذلك كما هو مبين في الجدول (٦).

الجدول (٦) نتائج اختبار t لمعاملات الارتباط للعينات المترابطة لقيم معاملات الارتباط بين معلمة القدرة الحقيقية وبين معلمة القدرة المقدرة وفق طريقتي التقدير باختلاف حجم العينة

حجم العينة	الارتباط بين معلمة القدرة الحقيقية وبين معلمة القدرة المقدرة وفق طريقة التقدير: BAYES	KE	الارتباط لمعلمة القدرة وفق الطريقتين	قيمة ت المحسوبة	الدلالة الإحصائية
250	0. 9522	0. 9578	1. 00	-8. 701	0. 000
500	0. 9648	0. 9626	1. 00	3. 027	0. 003
1500	0. 9422	0. 9402	1. 00	1. 803	0. 073

أظهرت النتائج في الجدول (٦) وجود فرق دال إحصائياً عند مستوى الدلالة α=٠.٠٥ بين قيمتي (معامل الارتباط بين معلمة القدرة الحقيقية وبين معلمة القدرة المقدرة وفق طريقة التقدير Bayes البارامترية) من جهة و(معامل الارتباط بين معلمة القدرة الحقيقية وبين معلمة القدرة المقدرة وفق طريقة التقدير KS اللابارامترية) من جهة أخرى عندما كان حجم العينة 750,250 فرداً لصالح (معامل الارتباط بين معلمة القدرة الحقيقية وبين معلمة القدرة المقدرة وفق طريقة التقدير Bayes البارامترية). في حين لم يثبت وجود فرق دال إحصائياً عند مستوى الدلالة α=٠.٠٥ بين قيمتي (معامل الارتباط بين معلمة القدرة الحقيقية وبين معلمة القدرة المقدرة وفق طريقة التقدير Bayes البارامترية) من جهة و(معامل الارتباط بين معلمة القدرة الحقيقية وبين معلمة القدرة المقدرة وفق طريقة التقدير KS اللابارامترية) من جهة أخرى عندما كان حجم العينة 1500 فرداً.

التوصيات:

- إجراء دراسة مشابهة على البيانات الحقيقية لنتائج الطلبة في الاختبارات المدرسية والجامعية.
- أجرى دراسة مشابهة للمقارنة بين طرق البارمترتية واللابارمترية وفق ظرفيات مختلفة في عدد الفقرات.
- أجرى مقارنة بين النماذج اللوجستية البارامترية واللابارامترية في دقة تقدير الطرق.

المراجـــع

أولاً– المراجع العربية:

الدردير، عبد المنعم أحمد. (٢٠٠٦): **الإحصاء البارامتري واللابارامتري في اختبار فروض البحوث النفسية والتربوية والاجتماعية.** القاهرة: عالم الكتب.

الشواورة، شادي يوسف.(٢٠١٣): **دقة تقدير معالم الفقرات بطريقتي الأرجحية العظمى الهامشية وبييز في ظروف مختلفة في عدد الفقرات وحجم العينة والنموذج اللوغاريتمي المستخدم.**رسالة دكتوراه غير منشورة، جامعة اليرموك، اربد، الأردن.

علام، صلاح الدين محمود. (٢٠٠٥): **نماذج الاستجابة للمفردة الاختبارية أحادية البعد ومتعددة الأبعاد وتطبيقاتها في القياس النفسي والتربوي.** القاهرة: دار الفكر العربي.

القيسي، حسين عبدالنبي.(٢٠١٣): **دقة تقدير معالم الفقرة والقدرة باستخدام نماذج نظرية استجابة الفقرة البارامترية واللابارامترية باختلاف حجم العينة وطول الاختبار.**رسالة دكتوراه غير منشورة، جامعة اليرموك،اربد، الأردن.

ثانيًا– المراجع الأجنبية:

Baker, F. (2001). The basics of item response theory. ***ERIC Clearinghouse on Assessment and Evaluation***, University of Maryland, College Park, MD.

Bellhouse, D.R. (2004). The Reverend Tomes Bayes, FRS: A Biography to celebrate the Tercentenary of His Birth. **Statistical Science** 19 (1), 3- 43.

Cowles, K. (2004). "Review of WinBuGs 1.4" **The American statistician**, 58, 330- 336.

Douglas, J. (1997). Joint consistency of nonparametric item characteristic curves and ability estimation. **Psychometrika**, 62, 7-28.

Dyehouse , M. (2009). A Comparison of Model-Data Fit for Parametric and Nonparametric Item Response Theory

Models Using Ordinal – Level Ratings. **Dissertation Abstract International**. (UMI No.3379330).

Glas, C. A, W., & Falcon, J. C. S. (2003). A comparison of item – fit statistics for the three–parameter logistic model. **Applied Psychological Measurement**, 27 (2), 89- 106.

Hambleton, R. K, swaminathan, H. & Rogers, Jane. H. (1991). **Fundamental Of Item Response Theory**. Newbury park, CA: sage publications, Inc.

Han, K. T., & Hambleton, R.K.(2007). User's Manual for WinGen: Windows Software that Generated IRT Model Parameter and Item Response. Center for Educational Assessment Research Report No 642, Amherst , MA: University of Massachusetts Center for Educational Assessment.

Koning, E., Sijtsma, K. & Hamers, J. (2002). Comparison of Four IRT Models When Analysing Two Tests for Inductive Reasoning. **Applied psychological measurement**, 26 (3), 302-320.

Kingma, J. & Tenvergert, E. (1985). A Nonparametric Scale Analysis of Development of Conservation. **Applied psychological measurement**, 9, 375-387.

Lord, M. F. (1986). Maximum Likelihood, and Bayesian parameter Estimation in Item Response Theory. **Journal of Educational Measurement**, 23 (2),157-162.

Linden,W. and Hambleton, R. (1997).**Handbook of Modern Item Response Theory**. Springer-Verlag. New York Inc: New YorkBerlin Heideelberg.

Meijer, R. & Baneke, J. (2004). Analyzing Psychopathology Items: A Case for Nonparametric Item Response Theory Modeling. **American Psychological Association**, 9 (3), 354–368.

Meijer, R. R., Sijtsma, K. & Smid, N. G. (1990). Theoretical and Empirical Comparison of the Mokken and Rasch Approach to IRT. **Applied Psychological Measurement**, 14, 283-298.

Mokken, R. J. & Lewis, C. (1982). A Nonparmetric Approach to the Analysis of Dichotomous Item Response. **Applied Psychological Measurement**, 6, 417-430.

Popham, W.J. (2004). **Clasroom Assessment: What Teacher need to Know**. 4th ed. Boston : Allyn and Bacon.

Ramsay,J. (1991). Kernel Smoothing Approaches to Non-para Metric Item characteristic Curve Estimation**. Psychometrika**, 56 (4), 611 – 630.

Sijtsma , K. (1998). Methodology review: Nonparametric IRT approaches to the analysis of dichotomous item scores. **Applied Psychological Measurement** .22, 3-31.

Sijtsma, K. and Molenaar, I. (2002).**Introduction to Nonparametric Item Response Theory. Sage Publication, International Educational and Professional Publisher .**Thousand Oaks: London. NewDelhi.

Van der Linden,W.J,& Hambleton , R.K.(1997).Part IV :Non-Parametric models introduction in W. J. van der Linden & R. K. Hambleton (Eds),Handbook of modern item response theory (pp.347-349).New York :Springer-Verlag.