طرائق تقدير معالم الفقرة والقدرة في نظرية الاستجابة للمفردة وفق نموذج راش

Methods for estimating paragraph parameters and ability in paragraph response theory According to the Rash model

هاجر شكار [1]، بن نابي نصيرة [2]

[1] جامعة الجزئر2 أبو القاسم سعد الله (الجزائر)، Chakar.hadjer@univ-alger2.dz

[2] جامعة الجزئر2 أبو القاسم سعد الله (الجزائر)، Nacera.Benabi@univ-alger2.dz

تاريخ الاستلام: 2023/8/27 تاريخ القبول: 2023/12/31 تاريخ النشر: 2024/3/31

ملخص:

هدفت هذه الدراسة الى التعرف على مختلف طرق تقدير معالم الفقرة والقدرة في نظرية الاستجابة للمفردة وفق نموذج راش، بالإضافة الى تحديد طرائق التقدير الأكثر دقة في التقدير تحت ظروف مختلفة في عدد الفقرات وحجم العينة والنموذج اللوغاريتمي المستخدم، وتوصلت نتائج الدراسة الى أنه توجد طرق مختلفة وعديدة من أهمها طريقة الارجحية العظمى بأساليبها الثلاث والطريقة "البيزية" بأسلوبيها وتبين من خلال اجراء المقارنة بين طرائق التقدير واستنادا الى مختلف الدراسات السابقة العربية منها والأجنبية الى تفوق الطرق "البيزية" على طريقة الارجحية العظمة خاصة في العينات الصغيرة، أو عند استخدام النموذج اللوغاريتمي الثنائي. كما أنها أعطت تقديرات أقل خطأ مقارنة مع طريقة الارجحية العظمى بالإضافة الى أنها تتمتع بثبات أفضل من تقديرات المعالم والأخطاء المعيارية التي يتم الحصول عليها من خلال طريقة الارجحية العظمى، خاصة عند تقدير القدرات على أطراف متصل القدرة كذلك بزيادة عدد الفقرات وحجم العينة.

الكلمات المفتاحية: طرق التقدير؛ معالم الفقرة، معالم القدرة؛ نموذج راش.

Abstract:

This study aimed to identify the different methods of estimating paragraph parameters and ability in the response theory of the item using the Rasch model, in addition to identifying the most accurate estimation methods in estimation under different conditions in terms of the number of paragraphs, the size of the sample, and the logarithmic model used. The most important of which is the method of great likelihood with its three methods and the Bayesian method with its two methods. It was found through a comparison between the estimation methods and based on various previous studies, both Arab and foreign, that the Bayesian methods are superior to the method of great likelihood, especially in small samples, or when using the binary logarithmic model. It also gave less error estimates compared to the maximum likelihood method, in addition to having better stability than the estimates of parameters and standard errors that are obtained through the maximum likelihood method, especially when estimating the capabilities on the outskirts of the capacity continuum as well as by increasing the number of vertebrae and the size of the sample.

Keywords: Estimate Methods; Item Parameter; Ability Parameter Rash Model

المؤلف المرسل: هاجر شكار .

1. مقدمة :

يعتمد القياس النفسي والتربوي على مدخلين رئيسين في تصميم وبناء المقاييس وكذلك على تحليل البيانات المستمدة منها، ونقصد بالمدخلين هما المدخل التقليدي المتمثل في النظرية الكلاسيكية للاختبارات والمدخل المعاصر المتمثل في نظرية الاستجابة للمفردة. وظلت النظرية الكلاسيكية للاختبار تشكل الاساس النظري والعملي في العلوم السلوكية لسنوات عديدة، وقد استندت هذه النظرية على نموذج بسيط للقياس الذي ينص على أن درجة الفرد الملاحظة في الاختبار تساوي مجموع الدرجة الحقيقية للفرد في السمة المقاسة ودرجة الخطأ في الاختبار. وعلية نجد أن القياس في النظرية الكلاسيكية يتأثر بخصائص عينة المفحوصين وبصعوبة المفردات التي يشمل عليها الاختبار. (Stone and yumoti ,2004 :41)

أما نظرية الاستجابة للمفردة الاختبارية (IRT) تسعى إلى تحقيق الموضوعية وتوفير خطية واستقلالية القياس، وتعتبر من النظريات السيكومترية المعاصرة التي اتضحت فائدتها في التغلب على كثير من مشكلات القياس التي عجزت النظرية الكلاسيكية عن مواجهتها، فهذه النظرية تحاول نمذجة العلاقة القائمة بين مستوى سمة معينة لدى الفرد (Trait Level) التي يقيسها اختبار معين، واستجابة لمفردة من مفردات الاختبار. ونظراً لأن مستويات السمة غير منظوره بطبيعتها وتغير في المتغيرات المنظورة، فإنه يتم تقديرها من هذه المتغيرات، أي من الاستجابات الملاحظة. وعادة يفترض ان السمة التي تنطوي عليها الاستجابات بمثابة قدرة (Ability) من نوع معين. وعلى العكس من نظرية القياس الكلاسيكية، فإن نظرية الاستجابة للمفردة تستند الى افتراضات قوية فيما يتعلق بما يحدث عند استجابة الفرد لمفردات اختبارية (علام،2005: 35)، حيث تفترض أنه يمكن التنبؤ بأداء الأفراد، أو يمكن تفسير أدائهم في اختبار نفسي أو تربوي، في ضوء خاصية مميزة لهذا الأداء تسمى السمة، وتحاول هذه النظرية تقدير درجات الأفراد في هذه السمات، وبالطبع تصعب ملاحظة هذه السمات ملاحظة مباشرة، لذلك يتم تقديرها من أداء الفرد الذي يمكن قياسه من خلال فقرات المقياس، ولهذا يطلق عليها السمات الكامنة. وقد انبثقت عن هذه النظرية مجموعة من النماذج، التي تعرف بنماذج السمات الكامنة، وتهدف جميعها إلى تحديد العلاقة بين أداء الفرد

على المقياس، وهو ما يمكن ملاحظته مباشرة، وبين السمات أو القدرات التي تكمن وراء هذا الأداء وتفسره. (الشريفين، 2009: 20)

تنقسم هذه النماذج إلى مجموعتين رئيسيتين كما يلي: النماذج الاستاتيكية (Statics Models): تهتم هذه النماذج بالقياس في وقت معين كما تهتم بتحديد العمليات التي تنطوي عليها الأداء في الاختبارات التربوية والنفسية، ومنها نموذج راش، نموذج لورد، ونموذج بيرنيوم، النماذج الديناميكية (Dynamic Models): وتهتم هذه النماذج بمشكلة قياس التحسن أو التغير الذي يحدث في السمات النامية المختلفة في فترات زمنية متباعدة، وبعض هذه النماذج يعتبر التغير النمائي عملية تدريجية، أي أن التغير متصل بينما يعتبر البعض الآخر أن هذا التغير انتقالي من حالة إلى أخرى، أي تغير منفصل مثل الانتقال من حالة التمكن إلى حالة عدم التمكن في الاختبارات مرجعية المحك.

وترتكز غالبية هذه النماذج على الحالة الأولى مثل نموذج بوك (Bock,1972) الذي يتعلق بمنحنيات قياس التحصيل في ضوء نموذج بيرنيوم الاستاتيكي، أما نموذج فيشر (Fischer,1973) فقد اعتمد على نموذج راش. (صلاح الدين، 2000 :685)

ويعتبر النموذج الذي وضعه عالم الرياضيات الدنماركي (جورج راش) عام 1960، وطوعه للتطبيق العالم الأمريكي (بن رايت) أهم نماذج السمات الكامنة، حيث يتوفر على متطلبات الموضوعية عندما تستوفي فروض النموذج المتمثلة في: أحادية البعد، استقلالية القياس، توازي المنحنيات المميزة للبنود، وهذا النموذج يفترض تساوي جميع فقرات الاختبار في التمييز بين مستويات القدرة المقاسة وإن الفرد لا يلجأ الى التخمين العشوائي في إجابته عن فقرات الاختبار. (Wright and Stone,1979)

إن الخطوة الأساسية والأهم في تطبيق نظرية الاستجابة للفقرة هي تقدير معالم نموذج نظرية الاستجابة للفقرة، التي تحدد خصائص هذا النموذج، كما أن النجاح في عملية التقدير يتوقف على توفير إجراءات مناسبة لتقدير هذه المعالم، سواء في النماذج المعملية أو النماذج اللامعلمية لنظرية الاستجابة للفقرة. وتعتمد الدقة في تقدير المعالم التي يتم الحصول عليها على مبدأين أساسيين، هما: علم التحيز في التقدير (unbiased)، ومعدل مربعات الأخطاء المتوقعةExpected Mean squared Error: EMSE))، حيث يشير التحيز (Bias) إلى توقع الفرق بين تقدير المعلم، وقيمة المعلم الفعلية، ولا يعني

تحقق هذا المبدأ أن التقدير دقيق تماما، وذلك لإمكانية وجود تقديرات عدة غير متحيزة، لهذا فإن مبدأ معدل مربعات الأخطاء التقدير الذي يقاس بواسطة تباين القيمة المقدرة يعطي أسلوب لاختيار تقدير جيد للمعالم. (Casella & Berger , 1990)

يعد تقدير المعالم من الخطوات المهمة في بناء الاختبارات والمقاييس وتحليل مفرداتها عند استخدام نظرية الاستجابة للمفردة، ويطلق عادة على عملية تقدير المعالم مصطلح التعيير (Calibration) غير أن الحصول على هذه التقديرات غالبا لا يتم بطريقة يدوية، لذا تم تطوير العديد من البرامج الحاسوبية للقيام بالعملية باستخدام أساليب إحصائية مختلفة (علام، 2005: 91)، وهناك أسلوبان لتقدير معالم الفقرات والأفراد، إما بشكل مشترك (Joint)، أي يتم تقدير معالم الفقرات والأفراد معا، معناه يتم تقدير معالم الفقرات أولا بعد حذف معالم قدرات الأفراد عن طريق تكامل اقتران الاحتمالية الهامشي (Marginal)، ثم إيجاد معالم تقديرات القدرة للأفراد. فالأسلوب الأول: يعتمد على تقديرات الأرجحية العظمى(Maximum Likelihood) وهي طريقة لتقدير معالم تقديرات القدرة للأفراد من خلال إجراءات تعظيم الاحتمالية للمعلمة المراد تقديرها، ويظم هذا الأسلوب ثلاثة طرق رئيسية: طريقة الارجحية العظمى المشتركة (Join Maximum Likelihooh Estimation)، طريقة الارجحية العظمى الشرطية (Conditional Maximum Likelihooh Estimation) وطريقة الارجحية العظمى الهامشية (Marginal Maximum Likelihooh Estimation)، بينما يعتمد الأسلوب الثاني: يعتمد على نظرية بييز (Bayes Estimation) والتي تمثل طريقة لتقدير معالم قدرات الافراد، عن طريق استخدام معلومات أولية عن العينة متوافرة من خلال الخبرات سابقة، تتم وفق طريقتين هما: طريقة بيز باستخدام أسلوب توقع الاقتران البعدي، وطريقة بييز باستخدام أسلوب تعظيم الاقتران البعدي، تستخدم طريقة بييز في العادة عندما لا نستطيع تطبيق الأسلوب الأول. هذا ما بينته دراسة قام بها وانغ وفيسبول (Wang & Vispoel,1988) هدفت الى تقويم النتائج التي يتم الحصول عليها من استخدام أربع طرق في تقدير القدرة وهي: طريقة الارجية العظمى وثلاث طرق مرتبطة بأسلوب بييز هي: طريقة أوينر (Owen's Method) وتوقع التوزيع البعدي (EAP) وطريقة تعظيم التوزيع البعدي(MAP)، إذ تم استخدام أسلوب المحاكاة لاختبار تكيفي، وقد أظهرت النتائج

اختلافات واضـــحة بين طريقة الارجحية العظمى و طريقة بييز من بينها أن طريقة الارجحية العظمة تعطي أخطاء عاليا في تقدير معلمة القدرة مقارنة بطرق بييز، وفيما يتعلق بطرق بييز فان طريقة أوينر كانت أســـوأ طريقة من حيث دقة التقديرات التي تقدمها. أما في دراســـة قام بها كيم (Kim,2001) هدف الى مقارنة دقة التقديرات التي تم الحصـــول عليها باســـتخدام طريقة الارجحية العظمى الشـــرطية وطريقة الارجحية العظمى الهامشـية وطريقة الارجحية العظمى المشـتركة وطريقة بييز التي تعتمد على توقع التوزيع البعدي وذلك من خلال استخدام أسلوب المحاكاة، إذ توصـلت الدراسـة الى أن نتائج تقدير معلمات الفقرة ومعلمة القدرة كانت متشابهة للطرق الأربعة.

في ضل وجود العديد م الطرق التي تسعى للدقة في تقدير معالم الفقرات وقدرة الافراد فان بحثنا هذا يسعى الى التعرف على مختلف طرق تقدير معالم المفردات وقدرات الافراد ضمن نظرية الاستجابة للمفردة وفق نموذج راش، من خلال التساؤلات التالية:

– ماهي الطرائق المختلفة لتقدير معالم المفردات وقدرات الافراد ضـمن النظرية الحديثة في القياس وفق نموذج راش؟

– ماهي طرائق التقدير الأكثر دقة في التقدير تحت ظروف مختلفة في عدد الفقرات وحجم العينة والنموذج اللوغاريتمي المستخدم؟

2. التحديد المفاهيمي:

1.2 طرق التقدير (Estimation Methods):

هي مجموعة من الأساليب الرياضية الاحتمالية بغية تقدير معالم المفردة والافراد في ضوء افتراضات نظرية الاستجابة للمفردة، حيث تناولت الدراسة (طريقة الارجحية العظمى (Maximum Likelihood Estimation)، طريقة التقدير القائمة على نظرية بييز (Bayeian Estimation)، طريقة بروكسل (Proxy).

2.2 معالم الفقرة (Item Parameter):

هي قيم إحصائية يتم تقديرها باستخدام معادلات رياضية وتشمل (الصعوبة، التمييز، التخمين)، وفي هذا البحث تم استخدام نموذج راش، بالتالي معالم الفقرة هي مستوى القدرة لدى الافراد الذين يستجيبون لمفردات الاختبار يتم تقديرها باستخدام النموذج الأحادي البارامتر (نموذج راش).

3.2 معالم القدرة (Ability Parametre):

هو مستوى القدرة لدى الافراد الذين يستجيبون لمفردات الاختبار يتم تقديرها في هذا البحث باستخدام النموذج الأحادي البارامتر (نموذج راش).

4.2 النموذج اللوغاريتمي احادي المعلم (One-Parameter Logistic Model):

وهو من اهم نماذج النظرية الحديثة الذي يتميز بسهولة مقارنته بالنماذج الاخرى من حيث عدد الفروض اللازم توافرها من البيانات لكي يعطي النموذج تقديرات دقيقة، وبسهولة تقدير معلم الصعوبة والقدرة احصائيا، ويشير فيشر Fisher الى أن الدرجة الخام تعبر عن المعلومات الوحيدة المطلوبة في البيانات التقدير القدرة، أي أن كل الأشخاص الذين يشتركون في درجة ما ستقدر لهم القدرة نفسها. (Fisher , 1973, 100)

3. الإجابة على التساؤل الأول الذي ينص على:

ماهي الطرائق المختلفة لتقدير معالم المفردات وقدرات الافراد ضمن النظرية الحديثة في القياس وفق نموذج راش؟

لقد انبثق عن نظرية الاستجابة للمفردة طرائق رياضية لتقدير معالم نماذج هذه النظرية ومنها:

1.3 طريقة الارجحية العظمى Maximum Likelihoood Method (ML):

تعتبر من الطرق الواسعة الانتشار حيث يتم تقدير المعالم فيها من خلال تعظيم الاحتمالية للمعلمة المراد تقديرها، ويكون شكل دالة الارجحية (Likelihood Function) (L)، إذا افترضنا أن Pi(θs) احتمال استجابة فرد ذو قدرة (θs) على الفقرة (i) إجابة صحيحة ومنه يمكن إعطاء العلاقة التالية كتعبير عن احتمال الاستجابة عن الفقرة كالتالي:

يرمز بـ**Qi(θs)** لاحتمال الإجابة الخاطئة للفرد ذو القدرة(θs) على الفقرة (i) حيث أن:

$$Q_i(\theta s) = 1 - P_i(\theta s)$$

u_i يأخذ القيمة 1 إذا كانت الإجابة صحيحة، والقيمة 0 إذا كانت الإجابة خاطئة عن الفقرة.

بصورة عامة فان الصيغة الرياضية لدالة الارجحية (L) لأنماط الاستجابات لعدد من الافراد (N) بواسطة مجموعة من الفقرات (n) ذوي مستويات قدرة ثابتة.

$$L = \prod_{s=1}^{N} \prod_{i=1}^{n} P_i(\theta_s)^{u_{si}} \; Q_i(\theta_s)^{(1-u_{si})}$$

من خلال هذه الصيغة التي تبين احتمال إجابات الافراد من ذوي مستوى قدرة معين هو حاصل ضرب الاحتمالات لجميع المفردات والجدير بالذكر ان عملية ضرب الاحتمالات لا تتم الا إذا كان استقلالية في الإجابة على هذه المفردات أي إجابات الفرد على فقرتين أو أكثر تكون مستقلة (لذا يجب تحقق شرط الاستقلال الموضعي (Local Indepndence) (علام، 2005:93).

ويوجد ثلاث طرق لتقدير معالم الفقرات ضمن طريقة الارجحية القصوى وهذه الطرق تختلف في كيفية التعامل مع مستويات القدرة غير المعلومة، وهي:

1.3. 1 طريقة الارجحية العظمى المشتركة (JML) Joint Maximum Likelihood:

يمكن تطبيق هذه الطريقة في النماذج اللوجستية الأحادية والثنائية والثلاثية ويتم وفقا لهذه الطريقة تقدير معالم القدرة والفقرة في أن معا خلال رفع الحد الأعلى لدالة الارجية المشتركة. هذه الطريقة مستخدمة في برامج تقدير المعالم لنظرية الاستجابة للمفردة مثل: LOGIST, FACETS, BIGSTEPS, WINSTAPS,QUEST.

يتم تقدير معالم الفقرات وفق هذه الطريقة على مرحلتين: في البداية افتراض قيم للقدرة (θs) (قيم مبدئية لمعالم الافراد) ومنه يتم تقدير معالم الفقرات من خلال البدء بمعالم أولية للفقرات، وفي المرحلة الثانية يتم تقدير القدرة (معالم الافراد) اعتمادا علة معالم الفقرات التي تم تقديرها في المرحلة الأولى ونقوم بعملية تكرار المرحلتين السابقتين حتى نحصل على قيم تقديرية للمعالم لا تختلف بين خطوتين متتاليتين بأقل من مقدار ثابت معتمد مسبقا (محك التقارب) أي هناك نوع من الثبات في تقديرات المعالم. (علام، 2005:96)

تتمثل دالة الارجحية العظمى للأفراد والفقرات من خلال الصيغة الرياضية التالية:

$$\ln L\ (\underline{u}|\theta_s, a, \underline{\delta}) = \sum_{s}^{N}\sum_{i}^{n}[u_{si} \ln P_i(\theta_s) + (1 - u_{si})\ \ln Q_i(\theta_s)]$$

حيث:

N : عدد الافراد.

n : عدد المفردات.

θs: قدرة الفرد. $1 - P_i(\theta s) = Q_i(\theta s)$

In: اللوغاريتم الطبيعي.

u: متجه ذو Nn استجابة N من الافراد على n من الفقرات.

u_{si}: متغير عشوائي يأخذ القيمة 1 إذا كانت الإجابة صحيحة، والقيمة 0 إذا كانت الإجابة خاطئة عن الفقرة.

لتقدير معلمة القدرة ومعالم الفقرة التي يكون الاقتران اللوغاريتم الطبيعي للأرجحية عندها قيمة عظمى من خلال تطبيق خوارزمية نيوتن – رافسون (Newton Raphson) التكرارية التي تعتمد على إيجاد المشتقة الأولى والمشتقة الثانية لاقتران لوغاريتم الارجحية العظمى المبنية في المعادلة السابقة ومساواتها بالصفر، ويمكن تلخيص خطوات هذه الطريقة في ما يلي :

– الخطوة الأولى : يتم تحديد أولي لقدرات الأفراد.

– الخطوة الثانية : يتم تقدير معالم الفقرات تبعا للقدرات الأولية التي يتم تحديدها في الخطوة الأولى، حيث يتم في هذه الخطوة افتراض قيم لدرجة معلمة الفقرة التي يتم تقديرها، ثم إيجاد المشتقة الأولى والثانية عند هذه القيمة وبقسمة المشتقة الأولى على المشتقة الثانية نحصل على قيمة التي بواسطها يتم تحسين قيمة المعلمة للدورة التالية حيث : درجة المعلمة الجيدة = درجة المعلمة الحالية – القيمة.

– الخطوة الثالثة : تقدير قدرات الافراد بناء على معالم الفقرات التي تم تقديرها في الخطوة السابقة.

– الخطوة الرابعة: نعتمد على التقديرات التي تم الحصول عليها في الخطوة الثالثة الخاصة بالأفراد ثم نضعها مكان القيم الأولية التي تم وضعها في الخطوة الأولى الى الخطوة الثالثة الى غاية الحصول على تقديرات ثابتة لكل من معالم الفقرة وقدرة الافراد. (التقي، 2009: 133– 132)

2.1.3طريقة الارجحية العظمى الهامشية Marginal Maximum Likelihood (MML):

هي من طرق تقدير معالم الفقرة تعتمد على مجموعة من الخطوات لإجراءات التقدير، حيث أن برمجية (BILOG -MG3) تعتمد عليها في تقدير معالم الفقرة.

ويمكن تطبيق هذه الطريقة على النماذج اللوجستية الأحادية والثنائية والثلاثية حيث يتم وفقا لهذه الطريقة إيجاد اقتران الاحتمالية الهامشي Marginal Maximum Likelihood) Estimation) لمعالم الفقرة من خلال تكامل اقتران الكثافة الاحتمالية على معلم القدرة ثم إيجاد

تقديرات معالم الفقرة وعندما تكون تقديرات معلمات الفقرة معروفة فانه عند إذ يمكن تقدير معلمة القدرة((Hambleton, R. K. & Swaminathan, H, 1985)

معادلات طريقة الارجحية العظمى لتقدير معالم الفقرة تمت كتابة صـــيغتها الرياضـية من خلال تربيع جاوس لتسـهيل وتقدير معالم الفقرة وضـحت من قبل بيكر وكيم(Baker & Kim,2004) ، وتعتمد عملية التقدير على عملية التكرار (Iteration) بمرحلتين في كل دورة، يتم تقدير أولي للمعلم ويتم تحســـينه في دورة الة ان تصـــل المعادلة التالية :

$$-\sum_{r}^{R} \overline{C}_{ri} - \overline{n}_{ri} * P_i(X_r) = 0$$

الى قيمة الصـــفر وفق محك تقارب يحدد دقتها، وللقيام بهذه الطريقة نســتخدم خوارزمية EM (EM algorithm) وهو الأسلوب الذي قام بتطويره بوك وايتكن (Bock& Aitkin,1981)، وتقوم خوارزمية EM في تقدير معالم المفردة على ما يلي:

المرحلة الأولى(E): وهي مرحلة التوقع (Expectation) حيث يتم حساب العدد المتوقع من الافراد في كل مستوى من مستويات القدرة وفق الصيغة:

$$\overline{n}_{ri} = \sum_{s=1}^{N} \frac{L(X_r)w(X_r)}{\sum_r^R L(X_r)w(X_r)}$$

توقع عدد الافراد الذين يجيبون إجابة صـــحيحة عند كل نقطة تربيع من خلال المعادلة:

$$\overline{C}_{ri} = \sum_{s=1}^{N} \frac{u_{si} * L(X_r)w(X_r)}{\sum_r^R L(X_r)w(X_r)}$$

المرحلة الثانية(M): وهي مرحلة التعظيم (Maximization) في هذه المرحلة يتم إيجاد معالم الفقرات التي تعظم دالة الارجحية، من خلال المعادلة التالية باســـتخدام قيم الخطوة الأولى التوقع (E) .

$$-\sum_{r}^{R} \overline{C}_{ri} - \overline{n}_{ri} * P_i(X_r) = 0$$

بعد القيام بالمرحلة الثانية (M) نقارن التقديرات المعالم المتوصـل اليها بتقديرات المرحلة الســـابقة، فاذا تقاربت تنتهي عملية التقدير (الفرق لا يتجاوز محك التقارب)، خلاف ذلك يتم تكرار الخطوتين الأولى والثانية (إذا لم نصـــل بعد الى الحد الأقصـــى لدورات التكرار).

عندما يتم تقدير معالم الفقرات في الخطوات السابقة نمز الى تقدير معالم الافراد أو قدرة الافراد (θs) عن طريق تعظيم اقتران دالة الارجحية العظمى أو استخدام احدى الطرق التي تعتمد على أســـلوب بييز (Bayes) في تقدير قدرات الافراد. (Baker & kim, 2004 ,P.171)

1.3. 3 طريقة الارجحية الشرطيةConditional Maximum Likelihood(CML) :

تنطلق هذه الطريقة في أن هناك شرطا كافيا لتقدير معالم القدرة عند الأفراد هو مجموع الفقرات التي أجاب عنها الفرد، كما أن هناك شرطا كافيا لتقدير معالم الفقرات، وهو مجموع الافراد الذين أجابوا إجابة عن الفقرة المعنية إجابة صحيحة. ومن هنا فان هذه الطريقة تحذف الفقرات التي تكون إجابات الطلبة عنها إجابة كاملة، أو التي تنعدم فيها إجابة جميع الطلبة، وكذلك حذف الأفراد (الحالات) التي تكون فيها إجابات الطلبة إجابة كاملة أو التي تنعدم فيها إجابات الطلبة عن أي من الفقرات.

ونظرا لان نموذج راش هو النموذج الذي تتوافر فيه هذا الشرط، لذلك تنحصر هذه الطريقة في هذا النموذج.

وعند الافتراض بأن الرمز (Sr) يدل على مجموع العلامات التي حصـــل عليها الطالب(S)، فان عملية تقدير درجة الصعوبة للفقرة I التي يتناولها الاختبار من خلال اعتماد نموذج راش يمكن تقديرها من خلال حل المعادلة التالية كما قدمها (امبريتسون وريز، 2000 :217)

$$\sum_{s=1}^{N} u_{is} - \sum_{s=1}^{N} P(u_{is} = 1 \mid r_s, \beta_j) = 0$$

(التقي، 2009: 149-148)

2.3 طرق التقدير القائمة على نظرية بييز (Bayesian Estimation):

تســـتخدم هذه الطريقة عندما لا يكون بالإمكان تطبيق الارجحية العظمة، وذلك بوجود الاستجابات المتطرفة عندما يجيب المفحوص عن جميع فقرات الاختبار إجابة صـــحيحة أو إجابة خاطئة، وتحتاج هذه الطريقة والمتضـــمنة كذلك في برنامج (BILOG -MG3)، بالإضـافة الى اسـتخدامها معلومات عن العينة (Hambelton & Swaminathan,1985)

تعتمد هذه الطريقة في التقدير على نظرية بييز (Bayes Theorem) والتي تربط بين الاحتمالات الشـــرطية (Conditional) والاحتمالات الهامشـــية (**Marginal**) كما

تتطلب كذلك افتراضات قبلية (Prior Probabilities) للمعالم بناء على اعتبارات نظرية أو امبريقية (علام،2005 :101)، والتي يمكن صياغتها على شكل توزيع احتمالي يعرف بدالة الكثافة الاحتمالية السابقة (Prior p. d.f) ويجري التعرف على هذه المعلومات من بيانات و تجارب سابقة أو من النظرية التي تحكم تلك الظاهرة، و كذلك تعتمد نظرية بييز على معلومات العينة الحالية المتمثلة بدالة الإمكان (Likelihood function) الخاصة بالمشاهدات، و عليه بدمج دالة الكثافة الاحتمالية للمعلم P(θ) مع دالة الإمكان P(u/θ) يتم الحصول على دالة الكثافة الاحتمالية اللاحقة للمعالم (Prior p .d.f) وهي: P (θ/u) . (Mislevy, R. T. ,1990)

بعد تحديد نوع التوزيع الاحتمالي وعادة يكون التوزيع القبلي لمعلم القدرة هو توزيع اعتدالي (مع الإشارة الى انه يمكن أن نفترض توزيعا أخر) تطبق الصيغة الرياضية الاحتمالية لنظرية بييز التي تربط الاحتمالات الشرطية بالهامشية بغية الحصول على دالة الكثافة للتوزيع البعدي للقدرة كما يلي:

$$P(\theta|u) = \frac{L(u|\theta)g(\theta)}{P(u)}$$

حيث:

u: متجه استجابة الفرد ذو القدرة (θ) على المفردات.

g(θ): تعبر عن دالة الكثافة للتوزيع القبلي لقدرات الافراد (وهو توزيع افتراضي لا يعتمد على معرفة مسبقة للإجابات الحقيقية للأفراد u).

L(u/θ): دالة الارجحية.

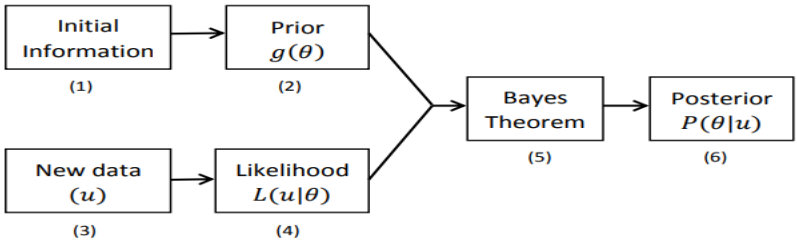
P(u): احتمال إجابة الفرد على مفردات الاختبار.

كما أن دالة كثافة التوزيع البعدي تتناسب مع حاصل ضرب دالة الارجحية مع دالة الكثافة للتوزيع القبلي أي: $\text{Posterior} \propto \text{Likelihood} * \text{Prior}$

ويعبر عن ذلك رياضيا بالصيغة التالية: $P(\theta|u) \propto L(u|\theta).g(\theta)$

والشكل التالي يوضح خطوات عمل طريقة بييز في التقدير (العبد الله، 2012: 36)

الشكل رقم (01): يوضح خطوات عمل طريقة بييز.



يمكننا القول انه بدمج البيانات المشـــاهدة التي يتم الحصـــول عليها من خلال تطبيق أداة القياس بالمعلومات القبلية للتوزيع نحصـل على التوزيع البعدي، وبالاستناد الى هذا التوزيع البعدي نحصل على تقدير قدرة الافراد.

هناك استراتيجيتين رئيسيتين في أسلوب بييز والتقدير مواقع الافراد، الأولى تعتمد على المنوال للتوزيع القبلي كتقدير للقدرة وهي البعدية القصوة (Maximum a posterior MAP)، أما الإستراتيجية الثانية تعتمد على المتوسط للتويع القبلي كتقدير للقدرة وهي البعدية المتوقعة (Expected a Posteriori EAP) . (De Ayala ,2009,P.77)

2.3. 1 أسلوب بييز للتوقع البعدي (EAP) Expected A posteriori:

يعمل هذا الأسلوب على إيجاد تقديرات القدرة للافراد بالاستناد الى تقديرات معالم الفقرات من خلال طريقة الأرجحية العظمة الهامشـــية، ويحســـب التقدير للتوقع البعدي (EAP) من خلال المتوسط الحسابي للتوزيع البعدي للقدرة (θ) دون اللجوء الى تقريب متتابع كما في طريقة الأرجحية العظمى (MLE). وتقاس درجة الدقة من خلال الانحراف المعياري للتوزيع البعدي (Posterior Standard Deviation – PSD).

2.3. 2 أسلوب التقدير البعدي الأعظم Maximum A posterior (MAP) :

يطلق على هذا الأسـلوب اسـم (التعظيم البعدي)، وإجراءاته مشـابها لإجراءات طريقة الارجحية العظمى، حيث يتم القيام بالتقريب المتتالي لإعطاء تقديرات ثابتة لمعلمة القدرة (θ)، ويتم تقدير (θ) عن طريق تعظيم(θ) نفســـها. ويســـتخدم في هذا الأســـلوب معلومات ســـابقة عن توزيع القدرة (θ)، حيث يتم افتراض هذا التوزيع بناء على معلومات ســـابقة، وغالبا ما يتم اســـتخدام التوزيع الطبيعي المعياري، بمعنى أن المفحوصـين قد تم اختيارهم من توزيع متوسـطه ويسـاوي (صـفر) وانحرافه المعياري يساوي (1) (Cooper, 2017). وبعد الحصول على استجابات الافراد على الفقرة الأولى

يجري حساب دالة الارجحية للاستجابات ودمجها مع المعلومات عن التوزيع المعدل للقدرة والذي يسمى التوزيع البعدي للقدرة كتوزيع قبلي ليتم دمجه مع دالة الارجحية لاستجابات الافراد على الفقرة الثانية، وتستمر هذه العملية بالتتابع (فقرة فقرة) عن طريق استخدام التوزيع البعدي للفقرة السابقة كتوزيع قبلي للفقرة التي تليها، ودمج معلومات هذا التوزيع مع دالة الارجحية لاستجابة الافراد على الفقرة حتى تنتهي جميع فقرات الاختبار (Wang & Vispoel , 1998)

وتستخدم طريقة (MAP) القيمة المنوالية للتوزيع البعدي كتقدير للقدرة بطريقة (MAP) لنمط الاستجابة المعطى (u) عن طريق حساب المشتقة الأولى للوغاريتم التوزيع البعدي للقدرة(θ) ومساواتها بالصفر، حسب المعادلة التالية:

$$\frac{\partial log[P(\theta|u)]}{\partial \theta} = \frac{\partial log[L(u|\theta)]}{\partial \theta} + \frac{\partial log[P(\theta)]}{\partial \theta} = 0$$

و تمثل $P(\theta/u)$: التوزيع البعدي للقدرة (θ) لنمط الاستجابة(u) ، وتمثل $L(u/\theta)$ التوزيع القبلي للقدرة (θ).

3.3 طريقة التقدير الموزونة (Estimation Method Biweight):

قدم (Warm , 1989) هذه الطريقة التي تعتمد في تقدير القدرة على توزين دالة الارجحية ببعض الدوال المناسبة، للتقليل من التحيز في تقدير القدرة بطريقة الارجحية العظمى (ML) ويمكن الحصول على تقدير الارجحية العظمى الموزونة للقدرة من خلال حل المعادلة التالية: $\frac{\partial lnL((u|\theta)}{\partial \theta} + \frac{\partial ln\, w(\theta)}{\partial \theta} = 0$

تمثل $L(\theta/u)$ دالة الارجحية لنمط الاستجابة، وتمثل $W(\theta)$ دالة وزن الارجية عند مستوى القدرة (θ) .

تعتبر تقديرات الارجحية العظمى لقدرات الافراد، ذات حساسية مفرطة للتشويش، والاختلال (Disturbance) مثل اللامبالاة والتخمين العشوائي، والذي يحدث غالبا في الاختبارات بأنواعها المختلفة تم تطوير هذه الطريقة لتقدم تقديرات قوية لقدرات الافراد باستخدام الارجحية العظمى، وفي هذه الطريقة فان الارجحية العظمى يتم تعديلها (Modified) لتوزين الفقرات التي تكون قريبة من كفاءة الفرد أو المفحوص، إن عملية توزين الفقرات (weighting) على نحو ملائم يمكن أن يقدم متوسطات

مكافئة للتحيزات المتوقعة، وبالتالي الحصـــول على تقديرات للقدرة أكثر دقة من تلك التي يمكن الحصول عليها، و ذلك من خلال استخدام هذا أسلوب الارجحية العظمى غير الموزونـة (Lewy & Nielsen,2002). وقد أدخـل ميســـلفي وبوك (Bock & Mislevy,1982) تعديلات جوهرية على تقدير الأرجحية العظمى بقصـــد التقليل من حساسيتها لاستجابات المفحوصين التي تتعارض مع نماذج نظرية الاستجابة للفقرة ومن الأمثلة على ذلك اعتبار الاســتجابة المشــوشــة على الفقرة الســهلة إجابة غير صــحيحة، بالنسـبة للمفحوصــين ذوي القدرات العالية، وهذا التعليل اسـتخدم الطريقة الموزونة لغاية تقدير القدرة للأفراد المفحوصين.

4.3 طريقة كوهلين التقريبية (Cohen's Approximation):

وهي طريقة أخرى أكثر اقتصـــادا قدمه كوهلين (1976)، وذلك لتقدير معالم نموذج راش وذلك بافتراض قدرات الفرد يمكن أن تقرب بواسطة الدالة الصريحة للدرجة الكلية وان هذه الدالة معرفة تماما ما عدا واحد فقط من المعالم المضروبة والذي يمكن تقديره بواســطة الترجيح الأكبر ويقوم هذا الافتراض على أن الصــورة الملائمة، التي يمكن أن يوصـــف بها التوزيع التكراري لكل من قدرات الافراد وصـــعوبات البنود هو التوزيع الاعتدالي

ولتطبيق هذه الطريقة التقريبية نتبع ثلاث خطوات رئيسية تجملها فيما يلي:

أ- تعيين تقديرات أولية لمعالم كل من صعوبات البنود وقدرات الافراد وتباينها.

ب- حساب معاملي الامتداد (Expansion Coeffcient) وهما:

- لصعوبة البند ويختص بتصحيح التقدير الاولي لمعلم صعوبة البند من تأثير امتدادا مدى القدرة لأفراد العينة، أي من تأثير تشتت قدرة الافراد.

- معامل الامتداد لفدرة الفرد: ويختص بتصـــحيح التقدير الاولي لمعلم صـــعوبة البند من تأثير امتدادا مدى البنود.

ت-حساب التقديرات النهائية للمعالم وأخطائها المعيارية. (كاظم،1988: 27)

4.3 برامج الحاسوب لتقدير المعالم :

لتقدير معالم نموذج راش لابد من وجود برامج حاسوبية ذات قدرات عالية في التحليل الاحصائي للحصول على تقديرات المعالم، من هذه البرامج:

برنامج Bical من اعداد (wright ,1980) – برنامج Logist من اعداد , Lord) (1980 – برنامج Dicot من اعداد (Masters,1984) – برنامج Microscale من اعداد (Wright & Linacer , 1985)– برنامج Bilog من اعداد (Mislevy & Bock,1990). ومن البرامج الحديثة أيضا برنامج Multilog من اعداد (Thissen,1991) برنامج Big Steps او Win Steps من اعداد (Wright & Linacer,1992) برنامج Parascale.3 من اعداد (Bock & Muraki,1990) برنامج Quest من اعداد (Admas & Khoo,1994) برنامج Rumm 2010 من اعداد (Lue,Andrich & Sheridan, 1996) وكذلك برنامج Bilog-mg 3 من اعداد(Zimowski , Mislevy & Bock , 2003) (عبدالله ،2003 :57)

وتوجد طرق أخرى لا تعتمد على الحاسب الالي وإنها يمكن إجرائها باستخدام الة حسابية، من بينها الطريقة التي تعرف باسم (UFORM) وهي تتطلب جداول خاصة (Wright, 1980) والطريقة التي تعرف باسم (Prox) والتي اقترحها (Wright & Stone ,1979) وهي تعتمد على البواقي المعيارية (علام،2001: 204).

4. الإجابة على التساؤل الأول الذي ينص على:

ماهي طرائق التقدير الأكثر دقة في التقدير تحت ظروف مختلفة في عدد الفقرات وحجم العينة والنموذج اللوغاريتمي المستخدم؟

للإجابة عن هذا التساؤل سنعرض الطرق الأكثر دقة في تقدير معالم القدرة والفقرة من خلال المقارنة بين الطرق المختلفة مع ذكر مختلف المميزات والعيوب لكل طريقة وكذا ما تطرقت اليه الدراسات السابقة.

4. 1 المقارنة بين الطرق المختلفة مع ذكر مختلف المميزات والعيوب لكل طريقة:

من خلال عرضنا لمختلف طرق تقدير معالم الفقرة والقدرة باستخدام نظرية الاستجابة للمفقرة وفق النموذج البارا متري الأحادي المعلمة (نموذج راش) نلاحظ أن طرق الارجحية القصوى (ML) تمتلك عدة خصائص منها: الاتساق والتقارب للقيم الحقيقية مع زيادة حجم العينة، كفاءة الخطأ المعياري النسبي والتوزيع الطبيعي لتقديرات الخطأ. غير ان هذه الطريقة ذات حساسية مفرطة للتشويش والاختلال مما يؤدي الى التحيز في تقدير القدرة ومنه جاءت طريقة التقدير الموزونة (Estimation Method Biweight) لحل هذا المشكل .

وبإلقائنا النظر على طرق الارجحية القصـــوى الثلاث نرى أنه تكون عملية التقدير بطريقة الارجحية العظمى المشـتركة (JML) بشـكل مباشـر بدون وضـع أي افتراضـــات لتوزيع قدرات الافراد او توزيع معالم الفقرة الا انه هناك بعض النقاط تأخذ على هذه الطريقة وقد ذكرها (علام، 2005) نوجزها فيما يلي:

– قد تنتج تقديرات غير متسقة بسبب تقدير معالم الفقرات والافراد في ان واحد ومتحيزة وبخاصـــة في النموذج الثنائي والثلاثي المعلم وطريقة الارجحية القصـــوى الهامشية تعالج بشكل أفضل مشكل اتساق تقديرات من خلال تحديدها لتوزيع تقديرات الافراد وبالتالي تحقق تقديرات معالم المفردات كلما زاد عدد الافراد.

– لا تســـتطيع هذه الطريقة تقدير قدرة الافراد الذين كانت كل اجابتهم على الفقرات إجابة صـــحيحة أو كانت كلها خاطئة، كذلك بالنســبة لمعالم الفقرات حيث لا يمكن تقديرها إذا جاءت جميع إجابات الافراد عليها إجابة صـــحيحة أو إجابة خاطئة لذا في هذه الطريقة يتم حذف هؤلاء الافراد وهذه الفقرات قبل البداية في عملية التقدير، لان في هذه الحالات عدم وجود قيمة عظمى للوغاريتم الارجحية.

– مشكل الاتساق والتحيز لتقديرات معالم المفردات وكذلك اعتمادها على قدرات الافراد التي تكون غير معلومة يؤثر على دقة تقدير الأخطاء المعيارية.

كما أشـــار (علام، 2005: 98) الى أهم مميزات وعيوب طريقة الارجحية العظمى الهامشية (MML) والمتمثلة في الاتي :

– يمكن اســـتخدام هذه الطريقة في تقدير معالم جميع النماذج الأحادية البعد (أحادية، ثنائية، ثلاثية المعالم) إضافة الى النماذج المتعددة الابعاد

– تمتاز هذه الطريقة بالفعالية مهما اختلف طول الاختبار، وإعطاء تقديرات لمعالم الفقرات سواء كانت الإجابة صحيحة أو خاطئة لجميع الفقرات (لا تشترط حذف هذه الفقرات) وبالتالي الاستفادة من جميع المعلومات.

رغم ما تمتاز به هذه الطريقة من مزايا الا انها يشوبها بعض العيوب أهمها:

– تتطلب هذه الطريقة في البداية قبل تقدير معالم المفردات افتراض توزيع لمستويات القدرة أو السمة للأفراد في المجتمع وبالتالي فان عملية التقدير تعتمد على مدى ملائمة هذا التوزيع الافتراضي لمستويات القدرة.

– في بعض الأحيان يكون تقدير غير دقيق لمعلمة التخمين في النموذج الثلاثي المعلم (ارتفاع الخطأ المعياري للتقدير) مما يؤثر سلبا على تقديرات باقي معالم المفردات والافراد.

– تتطلب هذه الطريقة توفر حجم عينة كبير نسبيا من أجل الاقتراب من مستويات القدرة أو السمة والوصول الى الدقة في تقدير معالم الفقرات.

أما طريقة الارجحية العظمى المشروطة (CML) في التقدير فإنها تختص بمجموعة من المميزات أشار اليها كل من أمبريستون وريزر (Embreston & Reise, 2000 ,pp. 217-218)، وعلام (علام ،2005: 100) بما يلي :

– لا تعتمد طريقة الارجحية العظمى المشروطة على أي افتراض حول توزيع محدد لمستويات القدرة أو السمة الكامنة في تقدير معلمة الصعوبة.

– تحقق مبدأ اللاتغاير في تقدير معالم الفقرة (صعوبة الفقرات) عند اختلاف الافراد بسبب عدم تأثر هذه التقديرات بمستوى قدرات عينة الافراد.

– تميز التقدير المحصل عليها بهذه الطريقة بخصائص جيدة مثل الاتساق حيث كلما كبر حجم العينة فان التقديرات تقترب من القيم الحقيقية، التوزيع الاعتدالي، والفاعلية أي أن لهذه التقديرات أقل تباين.

وقدما أيضا بعض الانتقادات لهذه الطريقة في التقدير نوجزها في الاتي:

– اقتصار استخدام هذه الطريقة في النموذج الأحادي المعلم (نموذج راش) وتعميماته فقط، حيث أنه في النماذج الأخرى (الثلاثي المعلم والثنائي المعلم) الدرجات الكلية في الاختبار لا تعد إحصاء كافيا لتقدير مستويات القدرة للأفراد.

– طريقة الارجحية العظمى المشروطة لا تقوم بالتقدير للعلامات الكاملة او الصفرية سواء كانت للأفراد او للمفردات وبالتالي تؤدي الى فقدان بعض المعلومات.

– توجد صعوبات في التحليل العددي عندما يكون عدد البنود أو المفردات في الاختبار كبير مما يؤدي الى فشل في استكمال إجراءات التقدير لبرامج الحاسوب خاصة إذا فاق عدد الفقرات 100 فقرة.

في حين إذا تحدثنا عن الطرق البيزية (Bayesian Estimation) فإنها تقدم حلولا لمشكلات التقدير لمعالم نماذج السمات الكامنة، وذلك عندما يكون حجم عينة الافراد

صـــغيرا وعدد فقرات الاختبار قليلا، ففي مثل هذه الحالات يصـــبح للمعتقدات القبلية أهمية كبيرة، أما إذا كان حجم العينة فان هذه المعتقدات تبدو قليلة الأهمية في مثل هذه التقديرات (علام، 2005)

فأبرز ما يميز الطرق البييزية هو أنها وجدت حلا للمشكلة المتمثلة في حصول المفحوص على علامة كاملة، أي في حال قيام المفحوص بالإجابة عن جميع الفقرات بشكل صحيح، وكذلك في حال حصول المفحوص على علامة الصفر، أي أن يجيب المفحوص على علامة الصـفر عن جميع فقرات الاختبار بشـكل خاطئ. وعلى الرغم من المآخذ على أســـاليب بييز، والمتمثلة في طبيعة المعلومات الأولوية المتوفرة من خبرات ســـابقة لدى الباحثين، إلا أن أســـاليب بييز يتوقع منها تخمين دقة التقدير وخاصـــة إذا كان حجم العينة صـــغيرا، وفي الحالة التي يكون فيها تقدير الارجحية العظمى لا يعمل بشكل جيد.(Gao & Chen,2005)

ولو تطرقنا للمقارنة بين أســلوبيها (أســلوب بييز للتوقع البعدي وأســلوب بييز للتقدير البعدي الاعظم) نجد أنه يكون لتقديرات بييز للتوقع البعدي (EAP) أقل معدل خطـأ معيـاري من التقـديرات الأخرى بمـا فيهـا الارجحيـة العظمى وتعظيم الاقتران. وتمتاز هذه الطريقة بأنها غير متكررة (Non-Iterative)، ســريعة الحســاب، وتعطي تقديرات للقدرة (θ) لجميع أنماط الاســـتجابات، ومن عيوبها عندمها يكون عدد فقرات الاختبار قليلا، فان التقدير يكون متحيزا نحو وســـط التوزيع، وذلك لأنها تعتمد على معلومـات أوليـة عن وســـط المجتمع وانحرافـه المعيـاري (Prasitpong & Arayathanitkul akkapo, 2016)، أما أسلوب التقدير البعدي الأعظم يمتاز في تقدير القدرة لجميع الافراد عكس طريقة الارجحية العظمى التي لا يمكنها تقدير معلمة القدرة للأفراد الذين أجابوا إجابة صحيحة أو خاطئة على كل الفقرات (حمادنة،2011 :11)، كما أن هذه الطريقة تعتمد على المنوال في اختيار قيمة القدرة الأكثر تكرار في التوزيع البعدي (EAP) فتعتمد على حساب معلم القدرة من خلال التوزيع القبلي.

2.4 الدراسات السابقة:

بينت جملة من الدراســات الســابقة الاختلاف في دقة تقدير معالم القدرة والفقرة بين مختلف الطرق منها :

دراسة الحمدانية و النصراوين (2020) والتي كانت تحت عنوان "مقارنة بين الطريقة البييزية وطريقة الارجحية العظمى في دقة تقدير معلمة القدرة ومعلمة الصعوبة وفق نموذج راش باستخدام بيانات مولدة محكاة " هدفت الدراسة الى مقارنة دقة تقدير معلمة القدرة ومعلمة الصعوبة باستخدام طريقتين (الطريقة البيزية، وطريقة الارجحية العظمى) باستخدام بيانات مولدة (محاكاة) باستخدام برنامج (WinGen) للحصول على بيانات لاختبارين بطول (20– 50) فقرة لعينتين متكونة من (250 و500) مفحوص، ولتقدير معلمة القدرة والصعوبة تم استخدام برنامج (Bilog-mg) توصلت الدراسة الى عدم وجود فروق ذات دلالة إحصائية بين المتوسطات الحسابية لمؤشر الجذر التربيعي لمتوسط مربع الخطأ في تقدير معلم الصعوبة باختلاف حجم العينة و كذا باختلاف طول الاختبار، بينما كانت هناك فروق دالة احصائيا في تقدير معلمة القدرة باختلاف طريقة التقدير لطالح الطريقة البيزية، وباختلاف حجم العينة لصالح العينة ذات 500 مفحوص.

كذلك **دراسة حمدان غسان حسن (2019)** بعنوان" دراسة مقارنة لطرائق تقدير المعالم في نظرية الاستجابة للمفردة " هدفت الدراسة الى المقارنة بين عدة طرق لتقدير المعالم في نظرية الاستجابة للمفردة بالاعتماد على الخطأ المعياري للتقدير كمؤشر للحكم على دقة التقدير باستخدام برنامج(Bilog-mg)، وباستخدام بيانات مفقودة وبيانات مكتملة من خلال تطبيق اختبار مكون من 60 مفردة، تمت مطابقة البيانات باستخدام النموذج الأحادي والنموذج الثنائي والثلاثي المعلم، بالاعتماد على كل من طرائق التقدير التالية: طريقة بروكس(Prox)، وطريقة الارجحية العظمى(WML)، وطريقة القيمة العظمى للتوزيع البعدي(MAP)، توصلت الدراسة الى تفوق طرائق التقدير القائمة على نظرية بييز(MAP-EAP) على طريقة الارجحية العظمى وطريقة بروكس في تقدير معالم الافراد والمفردات في حالة البيانات المفقودة والمكتملة، كما أن النموذج الأحادي أعطى تقديرات أكثر دقة مقارنة مع بقية النماذج الأخرى (الثنائي، الثلاثي).

و**دراسة شادي يوسف خلف الشواورة (2015)** بعنوان " دقة تقدير معالم الفقرات بطريقة الارجحية العظمى الهامشية وبييز (Bayes) في ظروف مختلفة في عدد الفقرات وحجم العينة والنموذج اللوغاريتمي المستخدم" سعت هذه الدراسة الى

بيان أي من الطريقتين الارجحية العظمى الهامشـــية باســـتخدام برنامج(Bilog-mg)، وطريقة بييز باستخدام برنامج (WinBUGS V1 .4) أكثر دقة في تقدير معالم الفقرات و فقا للنموذج اللوغاريتمي المســـتخدم (الأحادي، الثنائي) بحجم عينة (1500, 1000,500,250) وعدد فقرات (10 , 40) باســـتخدام المنهج التجريبي وللإجابة على تساؤلات الدراسة تم توليد بيانات الدراسة بالاعتماد على برنامج (WinGen) وفقا لكل نموذج تحت افتراض التوزيع الطبيعي لمعلمة القدرة و الصعوبة و التوزيع المنتظم لمعلم التمييز بقيمة ابتدائية وقيمة نهائية وتم الاعتماد على الجذر التربيعي لمتوسطات مربعات الخطأ (RMSE) كمؤشر لدقة التقدير وتوصلت الدراسة الى أفضلية طريقة بييز (Bayes) من طريقة الارجحية العظمة الهامشية(MML) في معظم مواقف الدراسة، وخاصة في العينات الصغيرة ، أو عند استخدام النموذج اللوغاريتمي الثنائي.

دراســـة جاو وجن (Gao & Chen ,2005) بعنوان" Bayesian or Non-Bayesian : A Comparison Study of Item Parameter Estimation in The Three-Parameter Logistic Model " ، هدفت الدراســـة الى المقارنة في دقة التقدير بين طريقة الارجحية العظمى وأســـلوب بييز باســـتخدام طريقة تعظيم الاقتران البعدي (MAP)، و لتحقيق هدف الدراســـة تم توليد بيانات بأحجام (2000,500,100)، واختبار بأطوال (60,30,10) فقرة، توصلت الدراسة الى أن دقة التقدير بين الطريقتين عند حجم عينة (2000,500) كانت متشابهة، كذلك عن تغير عدد الفقرات أما عند حجم عينة 100 فرد كانت الأفضلية لطريقة تعظيم الاقران البعدي في نظرية بييز.

دراســـة وانغ وفيســـبول (Wang & Vispoel,1998) تحت عنوان " Properties of Abiliy Estimation Methods in Computerized Adaptive Testing " هــدفــت الدراســـة الى مقارنة خصـــائص أربع طرق لتقدير القدرة : طريقة الارجحية العظمى (MLE) وثلاث طرق تعتمد على أسلوب بييز (Bayes) طريقة أوين (Owen's، طريقة توقع التوزيع البعدي (EAP) وطريقة القيمة العظمى للتوزيع البعدي (MAP) باســـتخدام أســـلوب المحاكاة في توليد بيانات لاختبار تكييفي محســـوب وتوصـــلت الدراســـة الى افضلية طرق بييز (Bayes) الثلاثة في التقدير حيث أعطت أقل خطأ مقارنة مع طريقة الارجحية العظمى، بينما عادت الأفضلية في أسلوب بييز لطريقة التوزيع البعدي(EAP).

كما أشـــارت نتائج **الأبحاث التي أجراها جيري وفريمنت (Garre & Vermunt, 2006)** المتعلقة بدقة تقدير المعالم، وذلك من خلال اســـتخدام توزيعات قبلية مختلفة لمعلمة القدرة، بحيث تم التحقق الامبريقي من نتائج الاشـتقاقات الرياضـية الخاصـة بطريقتي الارجحية العظمى وبييز بأســـلوب تعظيم الاقتران، من خلال تطبيق بيانات حقيقية بأحجام عينات مقدارها (100 ,1000) وبعدد فقرات (9.5) دلت نتائج الدراسة أن تقديرات المعالم والأخطاء المعيارية التي نحصـــل عليها خلال طريقة بييز، تتمتع بثبات أفضـــل من تقديرات المعالم والأخطاء المعيارية التي يتم الحصـــول عليها من خلال طريقة الارجحية العظمى، خاصة عند تقدير القدرات على أطراف متصل القدرة، كذلك عند حجم العينة 1000 وعدد الفرات9، أي بزيادة عدد الفقرات وحجم العينة.

5. خاتمة:

مما ســـبق نســـتخلص ان هناك طرق مختلفة و متعددة لتقدير معالم الافراد والفقرات في نظرية الاستجابة للمفردة منها : تقديرات الأرجحية العظمى Likelihood) Maximum) ويظم هذا الأســـلوب ثلاثة طرق رئيســـية: طريقة الارجحية العظمى المشـــتركة(Join Maximum Likelihooh Estimation) ، طريقة الارجحية العظمى الشـــرطية (Conditional Maximum Likelihooh Estimation) وطريقة الارجحية العظمى الهامشـــية (Marginal Maximum Likelihooh Estimation)، نظرية بييز (Bayes Estimation) والتي تمثل طريقة لتقدير معالم قدرات الافراد، تتم وفق طريقتين هما: طريقة بيز باســـتخدام أســـلوب توقع الاقتران البعدي، و طريقة بييز باســـتخدام أســـلوب تعظيم الاقتران البعدي ، طريقة التقدير الموزونة Estimation Method) Biweight) وطريقة كوهلين التقريبية(Cohen’s Approximation) .

كل طريقة من هذه الطرق تستخدم أسلوب رياضـي معين وتمتاز بمجموعة من المميزات والعيوب، كذلك تســـتخدم تحت شـــروط معينة هذا ما يجعلها في موطن تفاضـــل، حيث نجد بعض الطرق أفضـــل من طرق أخرى من حيث دقة التقدير تحت ظروف مختلفة في عدد الفقرات وحجم العينة والنموذج اللوجستي المستخدم فقد تفوقت طريقـــة الارجحيـــة العظمى الهـــامشـــيـــة Marginal Maximum Likelihooh) Estimation) بين طرق الارجحية العظمى وهذا لأنه يمكن اســـتخدام هذه الطريقة في

تقدير معالم جميع النماذج الأحادية البعد(أحادية، ثنائية، ثلاثية المعالم) إضـــافة الى النماذج المتعددة الابعاد بالإضـــافة الى أنها تمتاز بالفعالية مهما اختلف طول الاختبار، وإعطاء تقديرات لمعالم الفقرات ســـواء كانت الإجابة صـــحيحة أو خاطئة لجميع الفقرات (لا تشترط حذف هذه الفقرات) وبالتالي الاستفادة من جميع المعلومات، ورغم أن طريقة الأرجحية العظمى تتسـمم الاتسـاق والتقارب للقيم الحقيقية مع زيادة حجم العينة، كفاءة الخطأ المعياري النسـبي والتوزيع الطبيعي لتقديرات الخطأ. لا انها ذات حساسية مفرطة للتشويش والاختلال مما يؤدي الى التحيز في تقدير القدرة ومنه جاءت طريقة التقدير الموزونة (Estimation Method Biweight) لحل هذا المشكل.

كما أنها تواجه بعض المشـاكل الأخرى التي اسـتطاعت الطريقة البيزية حلها والتغلب عليها مما خولها لأن تكون أفضـــل الطرق و أكثر دقة في التقدير من طرق التقدير الأخرى، الامر الذي أكدته مختلف الدراســات الســابقة العربية منها والأجنبية التي اجرت أبحاث قارنت فيها بين طرائق تقدير المعالم في نظرية الاســتجابة للمفردة بالأخص طريقة الارجحية العظمى وأساليبها الثلاثة وطرقة بييز بأسلوبيها هادفة الى معرفة أي الطرق في تقدير معالم الافراد والمفردات أكثر دقة في التقدير، حيث توصلت الدراسات الى عدم وجود فروق ذات دلالة إحصائية بين المتوسطات الحسابية لمؤشر الجذر التربيعي لمتوسط مربع الخطأ في تقدير معلم الصعوبة باختلاف حجم العينة و كذا باختلاف طول الاختبار، بينما كانت هناك فروق دالة احصائيا في تقدير معلمة القدرة باختلاف طريقة التقدير لطالح الطريقة البيزية، وباختلاف حجم العينة كما توصـــلت الدراســـات الى تفوق طرائق التقدير القائمة على نظرية بييز-(MAP EAP) على طريقة الارجحية العظمى في تقدير معالم الافراد والمفردات في حالة البيانات المفقودة والمكتملة، كما أن النموذج الأحادي نموذج راش أعطى تقديرات أكثر دقة مقارنة مع بقية النماذج الأخرى (الثنائي، الثلاثي).

أيضا بينت الدراسات أفضلية طريقة بييز (Bayes) من طريقة الارجحية العظمة الهامشية (MML) في معظم مواقف الدراسة، وخاصة في العينات الصغيرة، أو عند استخدام النموذج اللوغاريتمي الثنائي. و افضلية طرق بييز (Bayes) في التقدير حيث أعطت أقل خطأ مقارنة مع طريقة الارجحية العظمى، بينما عادت الأفضـــلية في

أسلوب بييز لطريقة التوزيع البعدي (EAP) ، كما أن طريقة بييز تتمتع بثبات أفضل من تقديرات المعالم والأخطاء المعيارية التي يتم الحصول عليها من خلال طريقة الارجحية العظمى، خاصة عند تقدير القدرات على أطراف متصل القدرة كذلك بزيادة عدد الفقرات وحجم العينة.

6. قائمة المراجع :

1.6 المراجع باللغة العربية :

- أحمد، محمد التقي. (2009). النظرية الحديثة في القياس، ط1. عمان: دار المسيرة للنشر والتوزيع.
- أمينة، محمد كاظم. (1988). استخدام نموذج راش في بناء اختبار تحصيلي في علم النفس وتحقيق التفسير الموضوعي. الكويت: مطبوعة جامعة الكويت.
- حمادنة، مروان عبد الله. (2011). فاعلية أسلوب تحسين مطابقة الفرد القائم على تصحيح تقدير القدرة وتوزيعها المرجعي عند الاختلاف في حجم العينة والنموذج اللوجستي (رسالة دكتوراه). جامعة اليرموك، اربد.
- علام، صلاح الدين محمود. (2001). القياس والتقويم التربوي أساسياته وتطبيقاته وتوجهاته المعاصرة. ط5. القاهرة: دار الفكر العربي.
- علام، صلاح الدين محمود. (2005). نماذج الاستجابة للمفردة الاختبارية أحادية البعد ومتعددة الابعاد وتطبيقاتها في القياس النفسي والتربوي، القاهرة: دار الفكر العربي للطباعة والنشر.
- علام، صلاح الدين. (1986). أساسياته وتطبيقاته وتطورات معاصرة في القياس النفسي والتقويم التربوي. الكويت: مطابع التجارية.
- علام، صلاح الدين. (2000). القياس والتقويم التربوي والنفسي: أساسياته وتطبيقاته وتوجهاته المعاصرة، ط1. القاهرة: دار الفكر العربي.
- علام، صلاح الدين. (2006). القياس والتقويم التربوي والنفسي: أساسياته وتطبيقاته وتوجهاته المعاصرة، ط2. القاهرة: دار الفكر العربي.
- نضال، كمال الشريفين. (2009). أثر نمط صياغة الفقرة في مقياس التجاه الخصائص السيكومترية للفقرات وللمقياس وتقديرات القدرة وفق نظرية الاستجابة للفقرة. العرق: مجلة العلوم التربوية والنفسية، جامعة اليرموك، المجلد (10)، العدد (4).
- عبد الله، اعتدال غازي. (2003). استخدام نموذج راش في تدريج مقياس للقدرة العقلية لدراسة بعض العوامل المؤثرة على دقة القياس، رسالة دكتوراه، مصر: جامعة عين الشمس.

2.6 المراجع باللغة الأجنبية :

- De Gruijter,D.N and Van der kamp, L.J.(2005).Statistical test Theory for education and psychology.Retrived October,2005 from: www.lei-denuniv,n1/griuijterdnme.
- Baker, F. B., & Kim, S.-H. (2004). Item response theory: Parameter estimation techniques: CRC Press.
- Casella, G., & Berger, R.L., (1990). Statistical Inference, Belmont, CA, Duxbury. Coombs, C. H. (1964). A theory of data, New York, Wiley.
- Cooper, H. (2017). Research synthesis and meta-analysis: A step-by-step approach. (5th ed), Thousand Oaks, CA: Sage.
- Fischer, G. H. (1973): The Linear Logistics model as an instrument in educational Research and psychological, voI .37.
- Hambleton, R. & Swaminatham, H. (1985). Item Respose theory: principles and application. Kluwek NIJ puplishing.
- Hambleton, R. K. & Swaminathan, H. (1985). Item Response Theory, Principles and Applications. Bston: Kluwer. Nijhoff Publishing amembers of the Kluwer Academic Publishers Group.
- Mislevy, R. T. (1990): Modeling Item Response When Different Subjects Employ Different Solations Strategies, Netherlands: Psychometrica, Vol. 55, No., PP. 195- 215.
- Rakkapo, S. & Arayathanitkul, K. (2016). Analysis test of understanding of vectors with the three-parameter logistic model of item response theory and item response curves technique. Physical Review Physics Education Research, 12(2), 1-10.
- Stone, M. &Yumoto, F. (2004). The Effect of Sample Size for Estimating Rasch/ IRT Parameters with Dichotomous Items Journal of Applied Measurement. Vol. 5, No. 1, pp.48 – 61.
- Embretson, S. E., & Reise, S. P. (2000). Item response theory for psychologists. Maheah. In: New Jersey: Lawrence Erlbaum Associates, Publishers.
- Gao, F. & Chen, L. (2005). Bayesian or Non-Bayesian: A Comparison Study of Item parameter Estimation in the Three-Parameter Logistic Model. Applied Measurement in Education, 18(4), 351–380.
- Wang, T, & Vispoel, W.P. (1998) Properties. Of ability estimation methods in computerized adaptive testing. Journal of Educational Measurment, 35(2): 109-135.