مقـــارنة أربع طـــــرائق للكشـــــف عــن دالـــة الفقـــــرة التفـــاضلية
(دراسة محاكاة)

إعداد

نزار راسم محمد اللبدي

المشرف

الأستاذ الدكتور خليـــــل محمد عليـــــــان

قدمت هذه الأطروحة استكمالا لمتطلبات الحصول على درجة الدكتوراه في
علم النفس التربوي

كلية الدراسات العليا

الجامعة الأردنية

أيار، ٢٠٠٨

All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

المنارة للاستشارات

www.manaraa.com

نوقشت هذه الأطـــــروحة (مقـارنة أربـع طرائـــق للكشـــف عـن دالة الفقـــرة التفـاضلية (دراسة محاكاة)) وأجيـزت بتاريخ ١٢ / ٥ / ٢٠٠٨ م

أعضاء لجنة المنـــــــــاقشة	التوقيـــــــــع
الدكتور خليل محمد عليان، مشرفـاً أستاذ – القياس والتقويم	.....
الدكتور عمر حسن الشيخ، عضواً أستاذ – المناهج وأساليب تدريس العلوم	.....
الدكتور يونس محمد اليونس، عضواً أستاذ مشارك – البحث التربوي والتقويم	.....
الدكتور يوسف محمد سوالمه، عضواً أستاذ – القياس والتقويم (جامعة اليرموك)	.....

All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

المنارة للاستشارات
www.manaraa.com

الإهداء

إلى روح والدي الطاهرة
إلى والدتي تقديرا واحتراما
إلى إخواني وأخواتي
إلى زوجتي وأبنائي
والى جميع الزملاء الذين وقفوا إلى جانبي
إليهم جميعا اهدي هذا الجهد المتواضع

All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

المنارة للاستشارات
www.manaraa.com

شكر وتقدير

الحمد لله الذي أعانني على إنجاز هذه الدراسة، والشكر دوما له على نعمه وعطاياه.

يطيب لي بعد هذا الجهد المتواضع، أن أتقدم بخالص شكري وتقديري إلى الأستاذ الدكتور خليل عليان لتفضله بالإشراف على هذه الأطروحة، ولما بذله من جهد لانجاز الرسالة وإخراجها بهذه الصورة، والذي كان لملاحظاته وإرشاداته وتوجيهاته القيمة الأثر البالغ في إتمام هذا الجهد.

كما أتقدم بشكري وعرفاني للسادة أعضاء لجنة المناقشة، الأستاذ الدكتور عمر الشيخ، والأستاذ الدكتور يوسف سوالمه، والدكتور يونس اليونس لتفضلهم بالموافقة على مناقشة هذه الأطروحة، وللملاحظات القيمة التي أبدوها، والتي ساهمت في إثراء الدراسة.

وأخيرا أتوجه بالشكر والتقدير إلى كل من ساهم في إنجاز هذا العمل، والله ولي التوفيق.

All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

المنارة للاستشارات
www.manaraa.com

فهرس المحتويات

الموضوع	الصفحة
قرار لجنة المناقشة	ب
الإهداء	ج
شكر وتقدير	د
فهرس المحتويات	هـ
قائمة الجداول	ز
قائمة الأشكال	ط
الملخص باللغة العربية	م
خلفية الدراسة وأهميتها	١
مشكلة الدراسة وأهميتها	٣١
أسئلة الدراسة	٣١
مصطلحات الدراسة	٣٢
الدراسات السابقة	٣٣
الطريقة والإجراءات	٥٨
تصميم الدراسة ومتغيراتها	٥٩
معايير تقييم الطرق	٦٢
توليد بيانات الدراسة	٦٢

All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

المنارة للاستشارات
www.manaraa.com

فهرس المحتويات

الموضوع	الصفحة
نتائج الدراسة	٦٦
مناقشة النتائج	١١٥
المراجع	١٢٤
الملخص باللغة الانجليزية	١٣٣

All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

المنارة للاستشارات
www.manaraa.com

قائمة الجداول

الرقم	عنوان الجدول	الصفحة
١	توزيع أفراد المجموعة الضابطة والتجريبية وفقا للإجابة على الفقرة	١٢
٢	ملخص الدراسات التي اعتمدت على بيانات حقيقية	٥٤
٣	ملخص الدراسات التي استخدمت أسلوب المحاكاة.	٥٦
٤	معاملات التمييز للفقرة رقم ٤٠ في المجموعتين المرجعية والمستهدفة	٦١
٥	معالم الفقـــــــرات الاختبار من رقم ١ إلى ٣٩	٦٤
٦	نسبة الخطأ من النوع الأول للطرق الأربع وفقا لتوزيع القدرة وحجم العينة ونسبتها بين المجموعة المرجعية والمجموعة المستهدفة ومعامل التمييز	٦٨
٧	القوة للطرق الأربع وفقا لتوزيع القدرة وحجم العينة ونسبتها بين المجموعة المرجعية والمجموعة المستهدفة ومعامل التمييز في حالة وجود دالة فقرة تفاضلية منتظمة مقدارها ٠,٢٥	٧١
٨	القوة للطرق الأربع وفقا لتوزيع القدرة وحجم العينة ونسبتها بين المجموعة المرجعية والمجموعة المستهدفة ومعامل التمييز عند وجود دالة فقرة تفاضلية منتظمة مقدارها (٠,٥)	٧٥
٩	القوة للطرق الأربع وفقا لتوزيع القدرة وحجم العينة ونسبتها بين المجموعة المرجعية والمجموعة المستهدفة ومعامل التمييز في حالة وجود دالة فقرة تفاضلية غير منتظمة	٧٩
١٠	نتائج تحليل الانحدار اللوجستي لتأثير متغيرات توزيع القدرة والحجم العينة ونسيتها ومعلم التمييز على الخطأ من النوع الأول لطريقة مانتل هانزل	٨١
١١	نتائج تحليل الانحدار اللوجستي لتأثير متغيرات توزيع القدرة والحجم العينة ونسيتها ومعلم التمييز على الخطأ من النوع الأول لطريقة الانحدار اللوجستي	٨٢

All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

المنارة للاستشارات
www.manaraa.com

الرقم	عنوان الجدول	الصفحة
١٢	نتائج تحليل الانحدار اللوجستي لتأثير متغيرات توزيع القدرة والحجم العينة ونسيتها ومعلم التمييز على الخطأ من النوع الأول لطريقة نسبة الأرجحية	٨٤
١٣	نتائج تحليل الانحدار اللوجستي لتأثير متغيرات توزيع القدرة والحجم العينة ونسيتها ومعلم التمييز على الخطأ من النوع الأول لطريقة SIBTEST	٨٦
١٤	نتائج تحليل الانحدار اللوجستي لتأثير متغيرات توزيع القدرة والحجم العينة ونسيتها ومعلم التمييز على القوة لطريقة مانتل هانزل عند وجود دالة فقرة تفاضلية منتظمة مقدارها ٠,٢٥	٨٧
١٥	نتائج تحليل الانحدار اللوجستي لتأثير متغيرات توزيع القدرة والحجم العينة ونسيتها ومعلم التمييز على القوة لطريقة الانحدار اللوجستي عند وجود دالة فقرة تفاضلية منتظمة مقدارها ٠,٢٥	٩٠
١٦	نتائج تحليل الانحدار اللوجستي لتأثير متغيرات توزيع القدرة والحجم العينة ونسيتها ومعلم التمييز على القوة لطريقة SIBTEST عند وجود دالة فقرة تفاضلية منتظمة مقدارها ٠,٢٥	٩٣
١٧	نتائج تحليل الانحدار اللوجستي لتأثير متغيرات توزيع القدرة والحجم العينة ونسيتها ومعلم التمييز على القوة لطريقة مانتل هانزل عند وجود دالة فقرة تفاضلية منتظمة مقدارها ٠,٥	٩٨
١٨	نتائج تحليل الانحدار اللوجستي لتأثير متغيرات توزيع القدرة والحجم العينة ونسيتها ومعلم التمييز على القوة لطريقة الانحدار اللوجستي عند وجود دالة فقرة تفاضلية منتظمة مقدارها ٠,٥	١٠١
١٩	نتائج تحليل الانحدار اللوجستي لتأثير متغيرات توزيع القدرة والحجم العينة ونسيتها ومعلم التمييز على القوة لطريقة SIBTEST عند وجود دالة فقرة تفاضلية منتظمة مقدارها ٠,٥	١٠٧
٢٠	نتائج تحليل الانحدار اللوجستي لتأثير متغيرات توزيع القدرة والحجم العينة ونسيتها ومعلم التمييز على القوة لطريقة الانحدار اللوجستي عند وجود دالة فقرة تفاضلية غير منتظمة.	١٠٩

All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

المنارة للاستشارات
www.manaraa.com

قائمة الأشكال

الرقم	عنوان الشكل	الصفحة
١	دالة الفقرة التفاضلية المنتظمة	٧
٢	دالة الفقرة التفاضلية غير المنتظمة	٨
٣	تأثير حجم العينة على الخطأ من النوع الأول لطريقة الانحدار اللوجستي	٨٢
٤	تأثير معلم التمييز على الخطأ من النوع الأول لطريقة الانحدار اللوجستي	٨٣
٥	تأثير التفاعل بين حجم العينة ونسبتها على الخطأ من النوع الأول لطريقة الانحدار اللوجستي	٨٣
٦	تأثير معلم التمييز على الخطأ من النوع الأول لطريقة نسبة الأرجحية	٨٤
٧	تأثير التفاعل بين معلم التمييز والفروق في توزيع القدرة على الخطأ من النوع الأول لطريقة نسبة الأرجحية	٨٥
٨	تأثير التفاعل بين توزيع القدرة ونسبة حجم العينة بين المجموعة المرجعية والمجموعة المستهدفة على الخطأ من النوع الأول لطريقة نسبة الأرجحية	٨٥
٩	تأثير التفاعل بين معلم التمييز ونسبة العينة على الخطأ من النوع الأول لطريقة SIBTEST	٨٦
١٠	تأثير حجم العينة على القوة لطريقة مانتل هانزل عند وجود دالة فقرة تفاضلية منتظمة مقدارها ٠,٢٥	٨٨
١١	تأثير معلم التمييز على القوة لطريقة مانتل هانزل عند وجود دالة فقرة تفاضلية منتظمة مقدارها ٠,٢٥	٨٨
١٢	تأثير التفاعل بين معلم التمييز والحجم العينة على القوة لطريقة مانتل هانزل عند وجود دالة فقرة تفاضلية منتظمة مقدارها ٠,٢٥	٨٩
١٣	تأثير التفاعل بين معلم التمييز ونسبة العينة على القوة لطريقة مانتل هانزل عند وجود دالة فقرة تفاضلية منتظمة مقدارها ٠,٢٥	٨٩
١٤	تأثير الاختلاف في متوسط توزيع القدرة بين المجموعتين المرجعية والمستهدفة على القوة لطريقة الانحدار اللوجستي عند وجود دالة فقرة تفاضلية منتظمة مقدارها ٠,٢٥	٩١
١٥	تأثير التفاعل بين معلم التمييز وحجم العينة على القوة لطريقة الانحدار اللوجستي عند وجود دالة فقرة تفاضلية منتظمة مقدارها ٠,٢٥	٩١

All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

المنارة للاستشارات
www.manaraa.com

الرقم	عنوان الشكل	الصفحة
١٦	تأثير التفاعل بين معلم التمييز وتوزيع القدرة على القوة لطريقة الانحدار اللوجستي عند وجود دالة فقرة تفاضلية منتظمة مقدارها ٠,٢٥	٩٢
١٧	تأثير توزيع القدرة على القوة لطريقة SIBTEST عند وجود دالة فقرة تفاضلية منتظمة مقدارها ٠,٢٥	٩٤
١٨	تأثير توزيع نسبة العينة بين المجموعة المرجعية والمجموعة المستهدفة على القوة لطريقة SIBTEST عند وجود دالة فقرة تفاضلية منتظمة مقدارها ٠,٢٥	٩٤
١٩	تأثير حجم العينة على القوة لطريقة SIBTEST عند وجود دالة فقرة تفاضلية منتظمة مقدارها ٠,٢٥	٩٥
٢٠	تأثير معلم التمييز على القوة لطريقة SIBTEST عند وجود دالة فقرة تفاضلية منتظمة مقدارها ٠,٢٥	٩٥
٢١	تأثير التفاعل بين معلم التمييز وحجم العينة على القوة لطريقة SIBTEST عند وجود دالة فقرة تفاضلية منتظمة مقدارها ٠,٢٥	٩٦
٢٢	تأثير التفاعل بين حجم العينة وتوزيع القدرة على القوة لطريقة SIBTEST عند وجود دالة فقرة تفاضلية منتظمة مقدارها ٠,٢٥	٩٦
٢٣	تأثير التفاعل بين معلم التمييز ونسبة العينة بين المجموعة التجريبية والمجموعة المستهدفة على القوة لطريقة SIBTEST عند وجود دالة فقرة تفاضلية منتظمة مقدارها ٠,٢٥	٩٧
٢٤	تأثير حجم العينة على القوة لطريقة مانتل هانزل عند وجود دالة فقرة تفاضلية منتظمة مقدارها ٠,٥	٩٩
٢٥	تأثير معلم التمييز على القوة لطريقة مانتل هانزل عند وجود تحيز منتظم مقداره٠,٥	٩٩
٢٦	تأثير التفاعل بين معلم التمييز وحجم العينة على القوة لطريقة مانتل هانزل عند وجود دالة فقرة تفاضلية منتظمة مقدارها ٠,٥	١٠٠

All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

المنارة للاستشارات
www.manaraa.com

الرقم	عنوان الشكل	الصفحة
٢٧	تأثير الاختلاف في توزيع القدرة بين المجموعتين المرجعية والمستهدفة على القوة لطريقة الانحدار اللوجستي عند وجود دالة فقرة تفاضلية منتظمة مقدارها ٠,٥	١٠٢
٢٨	تأثير الاختلاف في نسبة العينة بين المجموعتين المرجعية والمستهدفة على القوة لطريقة الانحدار اللوجستي عند وجود دالة فقرة تفاضلية منتظمة مقدارها ٠,٥	١٠٢
٢٩	تأثير حجم العينة على القوة لطريقة الانحدار اللوجستي عند وجود دالة فقرة تفاضلية منتظمة مقدارها ٠,٥	١٠٣
٣٠	تأثير معلم التمييز على القوة لطريقة الانحدار اللوجستي عند وجود دالة فقرة تفاضلية منتظمة مقدارها ٠,٥	١٠٣
٣١	تأثير التفاعل بين حجم العينة ونسبتها على القوة لطريقة الانحدار اللوجستي عند وجود دالة فقرة تفاضلية منتظمة مقدارها ٠,٥	١٠٤
٣٢	تأثير التفاعل بين معلم التمييز وتوزيع القدرة على القوة لطريقة الانحدار اللوجستي عند وجود دالة فقرة تفاضلية منتظمة مقدارها ٠,٥	١٠٤
٣٣	تأثير التفاعل بين معلم التمييز ونسبة العينة على القوة لطريقة الانحدار اللوجستي عند وجود دالة فقرة تفاضلية منتظمة مقدارها ٠,٥	١٠٥
٣٤	تأثير التفاعل بين توزيع القدرة ونسبة العينة بين المجموعة المرجعية والمجموعة المستهدفة على القوة لطريقة الانحدار اللوجستي عند وجود دالة فقرة تفاضلية منتظمة مقدارها ٠,٥	١٠٦
٣٥	تأثير حجم العينة على القوة لطريقة SIBTEST عند وجود دالة فقرة تفاضلية مقدارها ٠,٥	١٠٧
٣٦	تأثير معلم التمييز على القوة لطريقة SIBTEST عند وجود دالة فقرة تفاضلية منتظمة مقدارها ٠,٥	١٠٨
٣٧	تأثير التفاعل بين معلم التمييز وحجم العينة على القوة لطريقة SIBTEST عند وجود دالة فقرة تفاضلية مقدارها ٠,٥	١٠٨
٣٨	تأثير توزيع القدرة على القوة لطريقة الانحدار اللوجستي عند وجود دالة فقرة تفاضلية غير منتظمة	١٠٩

All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

المنارة للاستشارات
www.manaraa.com

الرقم	عنوان الشكل	الصفحة
٣٩	تأثير حجم العينة على القوة لطريقة الانحدار اللوجستي عند وجود دالة فقرة تفاضلية غير منتظمة	١١٠
٤٠	تأثير معلم التمييز على القوة لطريقة الانحدار اللوجستي عند وجود دالة فقرة تفاضلية غير منتظمة	١١٠
٤١	تأثير التفاعل بين حجم العينة والتمييز على القوة لطريقة الانحدار اللوجستي عند وجود دالة فقرة تفاضلية غير منتظمة	١١١
٤٢	تأثير التفاعل بين حجم العينة ونسبتها بين المجموعة المرجعية والمجموعة المستهدفة على القوة لطريقة الانحدار اللوجستي عند وجود دالة فقرة تفاضلية غير منتظمة	١١١
٤٣	تأثير التفاعل بين حجم العينة وتوزيع القدرة على القوة لطريقة الانحدار اللوجستي عند وجود دالة فقرة تفاضلية غير منتظمة	١١٢
٤٤	تأثير التفاعل بين معلم التمييز وتوزيع القدرة للمجموعة المرجعية والمجموعة المستهدفة على القوة لطريقة الانحدار اللوجستي عند وجود دالة فقرة تفاضلية غير منتظمة	١١٢

All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

المنارة للاستشارات
www.manaraa.com

مقـــارنة أربـــعة طـــــرائق للكشـــــــــف عن دالة الفقـــــــرة التفـــاضلية
(دراسة محاكاة)

إعداد
نزار راسم محمد اللبدي

المشرف
الأستاذ الدكتور خليـــــل عليـــــــان

ملخـــــص

هدفت هذه الدراسة إلى تقييم فاعلية ومقارنة أربعا من طرائق الكشف عن دالة الفقرة التفاضلية، (Differential Item Function, DIF)، اثنتين من هذه الطرائق تستندان إلى النظرية الكلاسيكية وهما طريقة مانتل هانزل وطريقة الانحـــدار اللوجستي، واثنتين تستندان إلى النظرية الحديثة في القياس (نظرية استجابة الفقرة) وهما وطـــريقة اختبار تحيز الفقرة المتزامن (SIBTEST) وطـــريقة نسبة الأرجحية (IRTLRDIF)، وقد تم تقييم فاعلية هذه الطرائق باستخدام أسلوب المحاكاة في توليد البيانات وفقا لاثنين من المعايير هما الخطأ من النوع الأول والقوة، وقد تم توليد بيانات وفقا للنموذج اللوجستي ثلاثي المعالم لاختبار مؤلف من أربعون فقرة، وقد كانت جميع الفقرات متماثلة المعالم في المجموعتين باستثناء الفقرة الأخيرة التي تم من خلالها دراسة دالة الفقرة التفاضلية. ولقد فحصت الدراسة فاعلية هذه الطرائق تحت ظروف مختلفة من توزيع القدرة، وحجم العينة ونسبتها بين المجموعتين المرجعية والمستهدفة وتمييز الفقرة ونوع دالة الفقرة التفاضلية ومقدارها.

وقد أظهرت نتائج الدراسة ما يلي :

- أن معدل الخطأ من النوع الأول كان مقبولا وقريبا من القيمة النظرية ٠,٠٥ لثلاث طرائق هي طريقة الانحدار اللوجستي، وطريقة مانتل هانزل، وطريقة اختبار تحيز الفقرة المتزامن، كما أظهرت النتائج وجود تضخم واضح لمعدل الخطأ من النوع الأول لطريقة نسبة الأرجحية. ولقد كانت أفضل الطرائق فاعلية في الكشف عن دالة الفقرة التفاضلية وفقا لمعيار الخطأ من النوع الأول طريقة الانحدار اللوجستي، تلتها طريقة مانتل هانزل، ثم طريقة اختبار التحيز المتزامن، ثم طريقة نسبة الأرجحية.

All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

المنارة للاستشارات
www.manaraa.com

- وجود تأثير ذو دلالة إحصائية لمتغيـــــر حجـــــم العينة ومعلــم التمييز والتفاعل بين حجم العينة ونسبتها على الخطأ من النوع الأول لطريقة الانحدار اللوجستي، وكذلك وجود تأثير للتفاعل بين معلم التمييز ونسبة حجم العينة على الخطأ من النوع الأول لطريقة اختبار تحيز الفقرة المتزامن.
- أن القوة للطرائق الإحصائية الثلاث التي كان معدل الخطأ من النوع الأول لها مقبولا (الانحدار اللوجستي، مانتل هانزل ، اختبار تحيز الفقرة المتزامن) عند استخدامها للكشف عن دالة الفقرة التفاضلية المنتظمة التي مقدارها ٠,٢٥ كانت ضعيفة بشكل عام، فقد كانت القوة الإحصائية للطرائق الثلاث في معظم الخلايا ≥٠.٦٠، أما عندما كان مقدارها ٠,٥٠ فقد تمتعت الطرق الثلاث بقدرة عالية في الكشف عن الأداء التفاضلي، حيث كانت القوة للطرائق الثلاث في معظم الخلايا ≤ ٠,٨٠، مع كون نتائج طريقتي الانحدار اللوجستي ومانتل هانزل أفضل قليلا من طريقة اختبار تحيز الفقرة المتزامن. كما أظهرت النتائج أن الطريقة الوحيدة التي تصلح للكشف عن دالة الفقرة التفاضلية غير المنتظمة هي طريقة الانحدار وذلك في الخلايا التي كان فيها حجم الخلية كبيرا أو متوسطا ومعلم التمييز مرتفعا أو متوسطا عند تساوي متوسط توزيع القدرة وفي الخلايا التي كان فيها حجم العينة كبيرا ومعلم التمييز مرتفعا أو متوسطا عند عدم تساوي توزيع القدرة .
- أما فيما يتعلق بتأثير متغيرات الدراسة على القوة فقد أظهرت النتائج أن حجم العينة ومعلم التمييز من أهم المتغيرات التي لها تأثير على القوة للطرائق الثلاث التي كان معدل الخطأ من النوع الأول لها مقبولا (الانحدار اللوجستي ، مانتل هانزل ، اختبار تحيز الفقرة المتزامن) حيث كان للمتغيرين تأثير على القوة في جميع الحالات باستثناء الحالة التي كان يوجد فيها دالة فقرة تفاضلية منتظمة مقدارها ٠,٢٥ لطريقة الانحدار اللوجستي، حيث كانت القوة تزداد بزيادة كل من حجم العينة ومعلم التمييز.

All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

المنارة للاستشارات
www.manaraa.com

الفصـــل الأول

خلفيـــــة الدراســـــة وأهميتـــــها

All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

المنارة للاستشارات
www.manaraa.com

يتعلق صدق الاختبار بالدرجة التي تتوافر فيها أدلة علمية ونظرية تدعم تفسير درجات الاختبار وفقا للاستخدام الذي أعدت من أجله، وهو يتطلب معلومات حول ثبات الاختبار وتطبيق الاختبار، وطريقة إعطاء الدرجات وعملية تدريجها ومعادلتها واستخراج المعايير وعدالته بالنسبة لجميع الطلبة (Bert ,2004) . فوجود فقرات متحيزة في الاختبار ، يؤدي إلى اختبار متحيز ويهدد بشكل جدي عدالة الاختبار وصدق نتائجه، ويوثر بالتالي على معاني الدرجات والتفسيرات التي يتم التوصل إليها.

إن الاهتمام بمشكلة التحيز في الاختبارات ظهر بعد فترة قصيرة من استعمال أول اختبار ذكاء عام ١٩١٠م، عندما لاحظ بينيه (Binet) أن أداء أبناء الطبقات الاجتماعية والاقتصادية العليا أفضل من أداء أبناء الطبقات الاجتماعية والاقتصادية الدنيا في فرنسا وفي بلدان متعددة، مما جعل بينيه يعتقد أن هنالك فقرات في الاختبار تتأثر بالمستويين الاجتماعي والاقتصادي للمفحوصين، ولقد أشار بينيه أن اللغة والخلفية الثقافية وخبرات المفحوصين السابقة هي أمور مهمة في عملية قياس الذكاء دون أن يشير إلى مفهوم التحيز، وفي محاولة منه لتقليل أثر المستويين الاقتصادي والاجتماعي على نتائج الاختبار حذف في طبعته المنقحة للاختبار عام ١٩١١م مجموعة من الفقرات التي تتطلب من المفحوصين تسمية أربع قطع نقدية، إلا أن حذف هذه الفقرات لم يؤدِّ إلى تأثير ملموس على الفروق في اختبار بينيه للذكاء بين الطبقات الاجتماعية والاقتصادية (Jensen, 1980) .

وقد كان لحركة حقوق الإنسان في أواخر الستينات وأوائل السبعينات دورا في تسليط الضوء على موضوع تحيز الاختبارات وبالذات في الاختبارات العقلية، حيث عملت على تحقيق العدالة والمساواة بين الأفراد في فرص التعليم والعمل، بجعل الاختبارات خالية من التحيز ما أمكن (Conoley , 2003) .

ولقد استخدم مصطلح التحيز للدلالة على الأداء التفاضلي للفقرات في السابق ولم يشع استخدام هذا المصطلح إلا في بداية الثمانينات، حيث يذكر كاميلي و شيبارد Camilli & Shepard (1994) أن مصطلح دالة الفقرة التفاضلية (Differential Item Function, DIF) استخدم في الأدب التربوي منذ بداية الثمانينات للتعبير عن الطرائق والمعالجات الإحصائية المستخدمة للكشف عن تحيز فقرات الاختبارات، فهي دالة مشتقة إحصائيا للتعبير عن الفرق في الاستجابة للفقرة بين مجموعتين عند مستوى قدرة واحد، إلا أن الكثير من

All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

المنارة للاستشارات
www.manaraa.com

المؤلفين يستخدمون هذا المصطلح كمرادف لمصطلح التحيز، مثل هذا الاستخدام ليس لـــه مـــا يبرره حتى اليوم، فمصطلح الأداء التفاضلي للفقرة يستخدم لتحديد الفقرات التي يكون احتمـــال الاستجابة الصحيحة عليها مختلفا بين الأفراد في مجموعتين عند مستوى القدرة معين .

ونظرا للاهتمام المتزايد بصدق الاختبارات في البيئات الاجتماعية المختلفـــة وجوانـــب الضعف في الطرائق التي كانت تستخدم للكشف عن دالة الفقرة التفاضلية في منتصف الثمانينات وبداية التسعينات فقد تم القيام بالعديد من البحوث بهدف تطوير طرائـــق وإجـــراءات جديـــدة للكشف عن دالة الفقرة التفاضلية DIF أسفرت عن تطوير طرائق وأساليب جديدة للكشف عن دالة الفقرة التفاضلية حيث قام العديد من الباحثين بتقديم طرائق جديدة للكشف عن دالة الفقـــرة التفاضلية مثل

- Mantel-Haenszel (MH), Holland And Thayer ,1988
- IRT-LR , Thissen ,1988
- Logistic Regression, Swaminthan And Rogers, 1990
- Sibtest ,Shealy And Stout, 1993

(Thissen , 2001)

ويعرف هامبلتون (1991) Hambleton دالة الفقرة التفاضلية سيكومتريا بأنها الفرق في احتمالات الاستجابة الصحيحة للفقرة، في المجموعات المختلفة المتساوية في القدرة.

أما إمبرسون وريس Embertson & Reise (المشار إليهما في Sheilai, Teresi & Gershon, 2005) فيريا أن دالة الفقرة التفاضلية (DIF) تظهر في الفقرة عندما لا تكون للفقرة نفس العلاقة مع السمة الكامنة (القدرة) في اثنتين من المجموعات الفرعية أو أكثر.

كما أشار نونالي وبرنستين Nunnaly & Bernstein (المشار إليهما فـــي Gibson , 1998) إلى أن دالة الفقرة التفاضلية توجد عندما تكون الفقرة أكثر صعوبة أو تمييزا أو يمكـــن تخمينها بسهولة أكثر في أحد المجموعات.

All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

المنارة للاستشارات
www.manaraa.com

ووفقا لنظرية استجابة الفقرة (IRT) فإن دالة الفقرة التفاضلية (DIF) تظهر في الفقرة عندما تكون احتمالية الاستجابة بشكل إيجابي لدى الأفراد من نفس مستوى القدرة غير متساوية لدى الأفراد في المجموعات المختلفة (Teresi, Klenman & Ocepek-Weliksonm , 2000).

أما تحيز الفقرة (Item Bias) من وجهة نظر جنسن Jensen (1980) فهو خطأ منتظم في القياس أو التقدير ، كما يعتبر كروكر وألجينا Crocker & Algina (1986) الفقرة غير متحيزة إذا كانت تأثرها متساوٍ بنفس مصادر التباين في المجموعات لدى الأفراد من نفس المستوى بالنسبة للسمة المقاسة، فتحيز الفقرة يعنى أن احتمال الإجابة الصحيحة على الفقرة لدى مفحوصين من مجموعة معينة أقل من مجموعة أخرى بسبب خصائص معينة في هذه الفقرة أو الموقف الاختباري، وهنالك مفهوم آخر يتداخل مع المفهومين السابقين وهو
تأثير الفقرة (Item Impact) ويقصد به اختلاف احتمال الإجابة عن الفقرة إجابة صحيحة من أفراد من مجموعات مختلفة بسبب وجود فروق حقيقية في القدرة المقاسة بهذه الفقرة .
كما يختلف كل من تحيز الفقرة وتأثير الفقرة في كون الفروق في الأداء تستند إلى فروق ذات علاقة بالاختبار أو القدرة(Zumbo ,1999). ، فالفقرة قد تظهر فيها دالة الفقرة التفاضلية (DIF) لكنها لا تعتبر متحيزة إذا كانت هذه الفروق ناتجة عن فروق حقيقية في القدرة بين المجموعتين. ,(Roever , 2005) ، لذا فإن دراسة (DIF) تتطلب التحقق من تساوي توزيع القدرة لدى المفحوصين، قبل أن يحدد احتمال المجموعتين في الإجابة عن الفقرة إجابة صحيحة.

وهذا يعني أن وجود دالة الفقرة التفاضلية (DIF) يعتبر شرطا ضروريا لكنه غير كاف للحكم على تحيز الفقرة، فعدم وجود (DIF) يعني أن الفقرة لن تكون متحيزة، أما في حالة وجوده فإن ذلك غير كاف من أجل الإعلان عن تحيز الفقرة، لكن ذلك يتطلب تحليل إضافي لتحيز الفقرة (على سبيل المثال تحليل المحتوى، تقييم تجريبي) من أجل التحقق من تحيز الفقرة (Zumbo ,1999) . فلا يمكن اعتبار أي من الطرائق الداخلية وسيلة للتنبؤ الدقيق بالتحيز، لكنها في أحسن الأحوال مؤشر على الفقرات المحتمل تحيزها، كما أنها قد لا تكشف عن التحيز الموجود فعلا في الفقرات (Conoley, 2003) .

ويأخذ التحيز في الفقرات عدة أشكال أبرزها التحيز الجنسي و الثقافي و العرقي و الديني وقد يكون مصدره المحتوى أو اللغة التي قد تكون قد لا تكون مألوفة لدى بعض المجموعات الفرعية (Hambelton & Rodgers ,1995)، و بشكل عام هنالك أسباب عديدة لظهور (DIF) سواء على مستوى الفقرة أو الاختبار، مثل اختلاف معاني الفقرات أو العمليات العقلية

All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

المنارة للاستشارات
www.manaraa.com

المستخدمة في الاستجابة للفقرة أو الاختلاف في سلم الاستجابات، أو قد تكون هنالك مشاكل في ترجمة أو تفسير النتائج أو بالإجراءات المتبعة في جمع البيانات، لذا من الضروري اللجوء إلى طرائق البحث النوعي كالمجموعات المركزة أو المقابلات من أجل تقييم أسباب وجود دالة الفقرة التفاضلية (DIF) (Morales, 2004) .

ويشير هامبلتون وروجرز (Hambelton & Rodgers ,1995) إلى ثلاث خصائص للفقرات يجب أخذها بعين الاعتبار عند تقييم الفقرات من حيث التحيز أو دالة الفقرة التفاضلية ، وهي التحيز (Bias) والعدالة (Fairness) والتنميط (StereoTypeing)، وعادة ما يتم تضمين هذه الخصائص بأي نموذج لتقييم التحيز سواء كان من قبل كاتب الفقرات أو المقيمين الخارجيين لها . وبالرغم من أن هذه الخصائص تبدو قريبة من حيث المضمون، إلا أنها مختلفة ، فهنالك الفرق بين التنميط والتحيز والعدالة، فالتنميط يعني العرض المستمر لمجموعة ما بشكل معين والذي يمكن أن يكون هجوميا(مثل عرض السود بأن لديهم وظائف متدنية، أو التحدث عنهم بلغة غير لائقة) أما التحيز فيشير إلى وجود بعض الخصائص في الفقرة التي تقود إلى أداء مختلف للأفراد من ذوي القدرة نفسها ، بسبب انتمائهم لمجموعات مختلفة،. أما العدالة فيرى جنسن Jensen (المشار إليه في ظاظا، ٢٠٠٦) أنها حكم فلسفي يتعلق بالطريقة التي ينبغي أن تستعمل بها علامات الاختبار، وخاصة في مجال الاختيار التربوي أو الوظيفي، فبعد أن يتقرر استخدام الاختبارات في عملية الاختيار على أسس فلسفية، عندها يمكن تطبيق معايير إحصائية موضوعية، للتأكد من تحقيق الاختبار وطرائق الاختيار لشروط العدالة. فعدالة الاختبار مفهوم واسع يتضمن عدة جوانب منها

-الدقة في قياس السمة .

-التغطية الدقيقة للمحتوى .

-المعيارية أي أن تكون الظروف والمواد وعملية إعطاء الدرجات متماثلة لدى كل المفحوصين.

-عدم وجود فقرات متحيزة في الاختبار .

-الاستخدام الملائم لدرجات الاختبار . (Roever , 2005 ; Gierl, 2001)

وهنالك هدفان أساسيان لدراسة التحيز في الاختبارات

١- التحقق من مصادر التباين المختلفة التي تؤثر في درجات الاختبار لمعرفة ما إذا كانت المجموعات الفرعية تتأثر بنفس مصادر التباين.

٢- تحديد إذا كان أي من مصادر التباين يعطي أفضلية غير عادلة لمجموعة من المجموعات الفرعية. (Crocker & Algina , 1986)

All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

المنارة للاستشارات
www.manaraa.com

ولهذا الغرض تستخدم اثنتين من الطرائق

أ-طريقة المحكمين: وتستند إلى رأي محكم أو أكثر من أجل اختيار الفقرة المحتمل تحيزها.

ب- الطرائق الإحصائية . (Zumbo ,1999).

ويتضمن التركيز الأساسي للمحكمين عند تقويمهم للفقرة أربعة مجالات، مجالين لهما علاقة بالصدق الظاهري للأداة وهما التنميط والعدالة، أما المجال الثالث فله علاقة بصدق المحتوى وهو أن يكون محتوى الفقرات مألوفا لدى كافة المجموعات الفرعية، أما المجال الرابع فيتعلق بتساوي الفرص لتعلم محتوى الفقرات، أي توافق الفقرات مع العملية التعليمية (Berk ,1982).

وينظر Oosterhof (المشار إليه في ظاظا ٢٠٠٦) إلى طرائق التحكيم على أنها من الطرائق الشائعة والمستخدمة على نطاق واسع في الكشف عن التحيز المحتمل للفقرة، فهي تلعب دور في تأسيس صدق البناء للاختبار، حيث يتم في هذه الطريقة استخدام قوائم معدة مسبقا للكشف عن تحيز الفقرة، وعن طريق الأحكام التي يقدمها فريق من المقيمين، والذي يتألف من مجموعة من الحكام ممن لديهم الخبرة والمهارة والتخصص في مجال القياس النفسي والتربوي والمجال الذي يقيسه الاختبار، والمعرفة بخصائص مجتمع الأفراد الذين صمم الاختبار ليستخدم معهم وفقا لهذه القوائم، ويقوم هذا الفريق من الحكام أو ما يطلق عليه هيئة أو لجنة الخبراء بفحص الفقرات للكشف عن التحيز المحتمل وجوده في فقرات الاختبار ككل، بناء على المحكات الواردة في القوائم المعدة للتحكيم التي تتناول الجوانب التي تجعل الفقرة متحيزة للجانب موضع الاهتمام (العرق، أو الجنس، أو المستوى الاقتصادي الاجتماعي، أو الموقع الجغرافي...إلخ) .

ويمكن تصنيف الطرائق الإحصائية التي تستخدم من أجل الكشف عن دالة الفقرة التفاضلية (DIF) بأكثر من بعد

أ- الطرائق المستخدمة في الكشف عن دالة الفقرة التفاضلية بارامترية أو غير بارامترية .

ب- استناد الطريقة إلى السمة الكامنة (غير الملاحظة) أو المتغيرات المشاهدة (مثل الدرجة الكلية للاختبار).

ج- التعامل مع السمة كمتغير متصل أو غير متصل .

د- أحادية البعد أو متعددة الأبعاد .

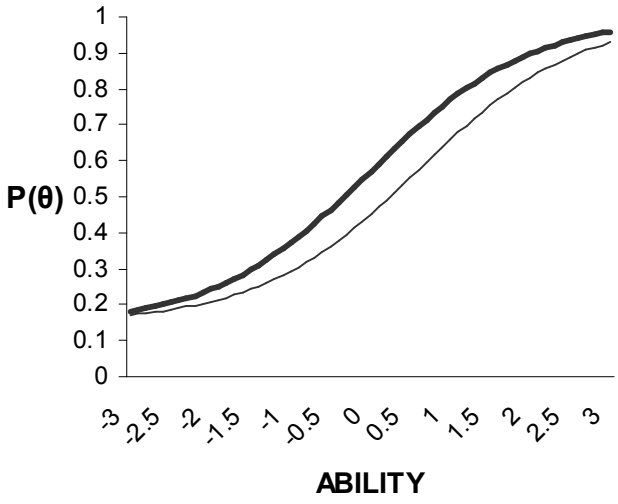
هـ- الدالة التفاضلية المنتظمة أو غير المنتظمة .

و- الفقرات ثنائية أو متعددة الاستجابة .(Teresi, 2004) .

All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

المنارة للاستشارات
www.manaraa.com

ومهما كانت طرائق الكشف عن دالة الفقرة التفاضلية (DIF) فهنالك نوعان منها: دالة الفقرة التفاضلية المنتظمة وغير المنتظمة ، الدالة المنتظمة تظهر عندما تكون احتمالية الاستجابة للفقرة بشكل صحيح أعلى بشكل منتظم عند إحدى المجموعات وعبر كل مستويات القدرة، بمعنى انه لا يوجد تفاعل بين مستوى القدرة والمجموعة التي ينتمي إليها الفرد، أما غير المنتظمة فتظهر عندما يكون هنالك تفاعل بين مستوى القدرة والمجموعة التي ينتمي إليه الفرد، فنمط الفروق في احتمالية الاستجابة للفقرة ليست متشابهة في كل مستويات القدرة، فهذه الفروق قد تكون لصالح إحدى المجموعات في مستوى قدرة معين، وفي مستوى قدرة آخر تكون لصالح مجموعة أخرى، ومن منظور نظرية الاستجابة الفقرة (IRT) ، تظهر دالة الفقرة التفاضلية المنتظمة عندما تكون منحنيات خصائص الفقرة (ICCs) متوازية لكن غير متطابقة كما في الشكل ١ أما منحنيات خصائص الفقرة (ICCs) غير المتوازية أو المتقاطعة كما يتضح من الشكل ٢ فتدل على دالة الفقرة التفاضلية (DIF) غير المنتظمة (BEER, 2004 ; Finch & French, 2007; Tian , 1999) .

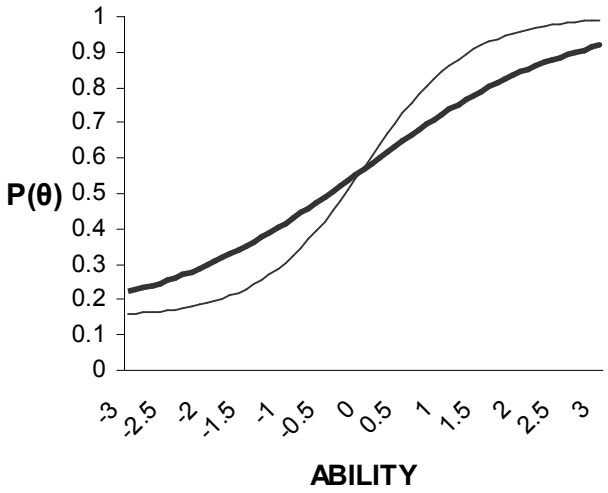


الشكل ١ . دالة الفقرة التفاضلية المنتظمة

الشكل قم (٢)

All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

المنارة للاستشارات
www.manaraa.com



الشكل ٢ . دالة الفقرة التفاضلية غير المنتظمة

وفيما يلي عرض لأهم الطرائق التجريبية الإحصائية المستخدمة في الكشف عن (DIF).

All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

المنارة للاستشارات
www.manaraa.com

أولا طرائق الكشف عن دالة الفقرة التفاضلية المبنية على إجراءات النظرية الكلاسيكية .

طريقة تحليل التباين

بقيت هذه الطريقة من أشهر الطرائق المستخدمة في الكشف عن تحيز الفقرات حتى نهاية الثمانينات ، ويتم الكشف عن التحيز في هذه الطريقة اعتمادا على دلالة تفاعل الفقرة مع المجموعة، حيث تكون المتغيرات المستقلة لتصميم تحليل التباين الثنائي الفقرات والمجموعة ومصادر التباين هي الفقرة والمجموعة والتفاعل بينهما والخطأ.

ويتم التعبير عن المعادلة الخطية لمصادر التباين كالتالي :-

$$\sigma^2 = \sigma^2_i + \sigma^2_j + \sigma^2_{ij} + \sigma^2_{k(ij)}$$

حيث أن

σ^2_i: التباين الناتج عن صعوبة الفقرة .

σ^2_j : التباين الناتج عن اختلاف القدرة بين المجموعات .

σ^2_{ij} :التباين الناتج عن التفاعل بين القدرة والمجموعة .

$\sigma^2_{k\,(ij)}$: تباين الخطأ.

ويسعى الباحث لاختبار الفرضية الصفرية المتعلقة بتفاعل متغيري الفقرات والمجموعة ($H_0 : \sigma^2 ij = 0$)، فإذا وجد تفاعل فإن هذا يعني وجود تحيز لصالح إحدى المجموعات عن المجموعات الأخرى (القرعان ١٩٩٧).

وقد وجه كاميلي وشيبارد المشار إليهما في(القرعان ١٩٩٧ ، Conoley, 2003) بعض الانتقادات لهذه الطريقة وأوصيا بعدم استخدام هذه الطريقة للكشف عن التحيز، فالتفاعل بين الفقرة والمجموعة قد يكون مصدره اختلاف القدرة بين المجموعات، وليس مصدره التحيز في الفقرة، كما أن هذه الطريقة تحتاج لعمل اختبارات مقارنة بعدية.

All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

المنارة للاستشارات
www.manaraa.com

طريقة معامل التمييز

في هـذه الطـريقة يستخدم معامل ارتبـاط الفقـرة مع الـدرجة الكليـة للاختبـار (Point Biserial Correlation) كمؤشر على قدرتها التمييزية، ويستخدم معامل التمييز في اتخاذ قرار بحذف الفقرة أو إبقائها في الاختبار، حيث أن الفقرة ذات الارتباط الـضعيف مـع الاختبار لا تساهم في إضافة تباين حقيقي إلى تباين الاختبار . ويحسب مؤشر التمييز للفقرة لدى مجموعة من الأفراد وفق المعادلة التالية :

$$P_{pbis} = \frac{\bar{x}_i - \bar{x}}{\sigma_x}\sqrt{\frac{P_i}{1-P_i}}$$

وتستخدم هذه المعادلة لكل فقرة ولكل مجموعة بشكل منفصل، حيث تمثل $(\bar{\chi}_i)$ متوسط علامات الأفراد الذين أجابوا على الفقرة إجابة صحيحة ، وتمثل $(\bar{\chi})$ متوسط العلامات لكافة أفراد المجموعة الذين استجابوا إلى الفقرة ، أما (σ_x) فتمثل الانحراف المعياري لعلامات المجموعة ، وتمثل (P_i) صعوبة الفقرة .

يتم ترتب معاملات الارتباط تصاعديا أو تنازليا، والفقرة التي يكون معامل ارتباطها مع الاختبار في النصف العلوي للمجموعة الأولى وفي النصف السفلي للمجموعة الثانية يمكن تصنيفها على أنها فقرة متحيزة، لذا تعتبر الفقرة متحيزة إذا اختلف معامل تمييزها في المجموعة الأولى بدرجة ملحوظة عن المجموعة الثانية(جرادات ٢٠٠٣ ، القرعان ١٩٩٧ ، Berk, 1982).

طريقة مربع كاي χ^2

وفقا لهذه الطريقة يرى Scheuneman (المشار إليه في Crocker & Algina 1986) أن الفقرة تعتبر غير متحيزة إذا كانت نسبة الأفراد الدين استجابوا بشكل صحيح للفقرة متساوية في جميع المجموعات الفرعية وذلك ضمن نفس الفئة من درجات الاختبار . فأسلوب χ^2 هو من نوع المطابقة (Goodness Of Fit) ويقوم على فحص الفرضية الصفرية بعدم وجود فرق ذي دلالة إحصائية ما بين العدد المتوقع والعدد الفعلي للذين يجيبون إجابة صحيحة على الفقرة من مجموعات لها نفس مستوى القدرة ، حيث يتم حساب قيمة الإحصائي χ^2 من أجل اختبار

All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

المنارة للاستشارات
www.manaraa.com

الفرضية الصفرية لفقرة معينة فإذا كانت قيمة هذا الإحصائي ذات دلالة إحصائية فإن الفقرة تعتبر متحيزة (Crocker & Algina 1986) .

وتقوم هذه الطريقة على عدد من الافتراضات

١- أن الاختبار يتمتع بالصدق ، فالعلامة الكلية يمكن أن تستخدم من أجل تقدير القدرة .

٢- أن الاختبار يتمتع بالثبات وفقا للمعايير المعتادة .

٣- تجانس التباين بين المجموعات (Berk, 1982) .

وتتم هذه الطريقة وفقا للخطوات التالية:-

١- تقسيم المدى الكلي لعلامات الاختبار إلى فئات .

٢- تحديد عدد الأفراد ضمن كل فئة ولكل مجموعة من المجموعات .

٣- إيجاد عدد الأفراد الذين أجابوا الفقرة إجابة صحيحة في كل فئة ولكل مجموعة من المجموعات .

٤- حساب القيم المتوقعة لكل مجموعة ومن ثم حساب قيمة كاي تربيع (χ^2) وفقا لمعادلة كاملي (Camilli) كالتالي :

$$x^2_c = \sum_{j=1}^{j} \frac{N_{1j} N_{2j} (P_{1j} - P_{2j})^2}{(N_{1j} + N_{2j}) P_{.j} (1 - P_{.j})} = \sum x^2_j$$

حيث أن

N_{1j} = نسبة المفحوصين من المجموعة الفرعية الأولى في الفئة j .

N_{2j} = نسبة المفحوصين من المجموعة الفرعية الثانية في الفئة j .

p_{1j} = نسبة المفحوصين الذين أجابوا على الفقرة إجابة صحيحة في الفئة j في المجموعة الفرعية الأولى.

P_{2j} = نسبة المفحوصين الذين أجابوا على الفقرة إجابة صحيحة في الفئة j في المجموعة الفرعية الثانية .

$p_{.j}$ = نسبة المفحوصين الذين أجابوا الفقرة إجابة صحيحة في الفئة j من كافة المجموعات الفرعية.

حيث يتم فحص الفرضية الصفرية عند درجة حرية تساوي j (عدد الفئات التي تم تقسيم درجات إليها)، فإذا كانت كاي تربيع دالة إحصائية، فإن ازدياد قيمتها يكون مؤشراً أكبر على تحيز الفقرة (Crocker & Algina 1986).

All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

المنارة للاستشارات
www.manaraa.com

طريقة مانتل هينزل (Mantel-Haenszel Method :HM)

تعتبر طريقة مانتل هينزل من أشهر الطرائق والأكثر استخداما من أجل الكشف عن دالة الفقرة التفاضلية (DIF)، وقدمت هذه الطريقة كطريقة للكشف عن تحيز الفقرة في أواخر الثمانينيات من القرن الماضي من قبل هولند وثاير وتحديدا في عام ١٩٨٨م HOLLAND & THAYER (1988) ، وهي من أشهر طرائق الكشف عن التحيز في الوقت الحاضر وذلك لسهولة حساباتها وإجراءاتها(عبد الله جرادات،٢٠٠٣ Hidalgo & López-Pina , 2004).

وفي هذه الطريقة تتم مقارنة الأداء على الفقرة بين مجموعتين فرعيتين إحداهما تسمى المجموعة المرجعية (الأغلبية) وتسمى الأخرى المجموعة المستهدفة (الأقلية) .
ويستخدم في هذه الطريقة مربع كاي ومعاملات الاقتران كما يوضحها الجدول التالي .

جدول ١. توزيع أفراد المجموعة الضابطة والتجريبية وفقا للإجابة على الفقرة

المجموعة	الإجابة الصحيحة	الإجابة الخطأ	المجموع
المجموعة المرجعية	A_j	B_j	n_{rj}
المجموعة المستهدفة	C_j	D_j	n_{fj}
المجموع	M_{1j}	M_{0j}	T_j

حيث أن

A_j عدد أفراد المجموعة المرجعية الذين أجابوا على الفقرة إجابة صحيحة عند مستوى القدرة J.

B_j عدد أفراد المجموعة المرجعية الذين أجابوا على الفقرة إجابة صحيحة عند مستوى القدرة J.

C_j عدد أفراد المجموعة المستهدفة الذين أجابوا على الفقرة إجابة خاطئة عند مستوى القدرة J.

D_j عدد أفراد المجموعة المستهدفة الذين أجابوا على الفقرة إجابة خاطئة عند مستوى القدرة J.

M_{1j} عدد الأفراد الذين أجابوا على الفقرة إجابة صحيحة عند مستوى القدرة J لكل أفراد العينة.

M_{0j} عدد الأفراد الذين أجابوا على الفقرة إجابة خاطئة عند مستوى القدرة J لكل أفراد العينة.

n_{rj} عدد أفراد المجموعة المرجعية ممن هم في مستوى القدرة J.

N_{fj} عدد أفراد المجموعة المستهدفة ممن هم في مستوى القدرة J

T_j عدد الأفراد ممن هم في مستوى القدرة J.

All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

المنارة للاستشارات
www.manaraa.com

ويتم حساب الإحصائي مربع كاي مانتل هينزل من خلال المعادلة التالية :-

$$MH\chi^2 = \frac{\left(\left|\sum_{j=1}^{s}\{A_j - E(A_j)\}\right| - 0.5\right)^2}{\sum_{j=1}^{s} Var(A_j)}$$

حيث يتم حساب قيمة E(A_J) المتوقعة من خلال التكرارات الهامشية وفقا للمعادلة التالية

$$E(A_j) = (n_{Rj}\, m_{1j}) / T_J$$

أما $Var\ (A_J)$ فيشير إلى التباين ويحسب من المعادلة

$$Var(A_j) = \frac{n_{\mathrm{R}j} n_{\mathrm{F}j} m_{1j} m_{0\mathrm{j}}}{T_j^2 (T_j - 1)}$$

ويفيدنا مربع كاي في فحص الفرضية المتعلقة باستقلالية الإجابة الصحيحة والخاطئة عند كون المجموعة مرجعية أو مستهدفة.

وفي طريقة مانتل هينزل MH يمكن تقدير حجم الأثر (Effect Size) اعتمادا على Constant Odds Ratio من خلال المعادلة التالية

$$\alpha_{MH} = \frac{\sum_{j=1}^{s} A_j D_j / T_j}{\sum_{j=1}^{s} B_j C_j / T_j}$$

هنالك العديد من المزايا لاستخدام طريقة MH ، فحيث أن طريقة MH تعتبر من الطرائق غير البارامترية فهي تحتاج إلى عدد قليل من المفحوصين مقارنة مع الطرائق الأخرى، كما أنها توفر تقدير لحجم الأثر وكذلك اختبار إحصائي لمستوى الدلالة، كما أن طريقة MH من أكثر الطرائق سهولة للفهم وجذبا للأشخاص الذين لديهم القليل من الخبرة في طرائق الكشف عن دالة الفقرة التفاضلية (DIF) (Samuelsen , 2005; Chaimongkol, 2005) .

All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

المنارة للاستشارات
www.manaraa.com

ولقد اقترح هولاند وتاير Holland & Thayer (1988) تحويلا لوجريتميا لـــ α بهدف الحصول على مقياس متماثل ، حيث تدل القيمة صفر على عدم وجود (DIF) أما القيم السالبة فتدل على أفضلية للمجموعة المرجعية ، والقيم الموجبة فتشير إلى أفضلية في الاتجاه المعاكس. ويتم التعبير عن هذا التحويل من خلال الصيغة التالية :

$$\Delta_{\alpha(HM)} = -2.35\ln(\alpha_{(HM)})$$

وبناءا على هذا التحويل اقترح زوك وإركان Zwick & Ercikan (1988) (المشار إليهما في Hidalgo & López-Pina, 2004; Bert & Stoneberg, 2004) دليل التفسير التالي من اجل تقييم حجم الأثر لـــ (DIF) :

النوع A من الفقرات –الفقرات ذات DIF المهمل : $|\Delta_{\alpha(HM)}| > |١|$.

النوع B من الفقرات –الفقرات ذات DIF المتوسط : $|١| \geq |\Delta_{\alpha(HM)}| \geq |١,٥|$.

النوع c من الفقرات –الفقرات ذات DIF الكبير : $|\Delta_{\alpha(HM)}| < |١,٥|$.

ولقد أشارا إلى أن الفقرات من النوع B يمكن استخدامها إذا كانت لا توجد فقرات أخرى يمكن استبدالها بها، وأن الفقرات من النوع C يمكن اختيارها في الحالة التي تكون فيها ضرورية من اجل تحقيق خصائص معينة للاختبار .

طريقة دالة التمييز اللوجستي(Logistic Discriminant Function analysis LDFA)

تم اقتراح هذه الطريقة من قبل كل من ملر وســوري Miller & Soray (1993) (المشار إليهما في su & wang, 2005 ; Pastor & Richarde, 2003) من أجل الكشف عــن دالة الفقرة التفاضلية (DIF) المنتظمة وغير المنتظمة في الفقرات متعددة الاستجابة، وتشبه هذه الطريقة طريقة الانحدار اللوجستي الرتبي، فكما في طريقة الانحدار اللوجستي الرتبي فإنه فــي البداية تتم مطابقـــة ثلاثـــة من النمـــاذج المتداخلة مع البيانـــات وتتـــم مقـــارنة -2 Log Likelihood) من أجل التعرف على (DIF) المنتظم وغير المنتظم، حيث أن مقارنة أول نموذجين تمكننا من التعرف على (DIF) المنتظم ومقارنة آخر نموذجين تمكننا من التعرف على (DIF) غير المنتظم، وفيما يلي توضيح للنماذج الثلاث المستخدمة

model1

$$p(\text{group}=0) = \frac{\exp(\alpha_0 + \alpha_1(\text{total score}))}{1 + \exp(\alpha_0 + \alpha_1(\text{total score}))},$$

model2

$$p(\text{group}=0) = \frac{\exp(\alpha_0 + \alpha_1(\text{total score}) + \alpha_2(\text{item}))}{1 + \exp(\alpha_0 + \alpha_1(\text{total score}) + \alpha_2(\text{item}))},$$

All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

المنارة للاستشارات
www.manaraa.com

model3

$$p(\text{group}=0) = \frac{\exp(\alpha_0 + \alpha_1(\text{total score}) + \alpha_2(\text{item}) + \alpha_3(\text{total score*item}))}{1 + \exp(\alpha_0 + \alpha_1(\text{total score}) + \alpha_2(\text{item}) + \alpha_3(\text{total score*item}))},$$

والنماذج الثلاث المستخدمة في (LDFA) تشبه إلى حد كبير النماذج المستخدمة في تحليل الانحدار اللوجستي لكن المتغيرات التي تستخدم كمتغيرات تابعة تختلف في النموذجين فالمتغير التابع هو المجموعة التي ينتمي إليها الفرد أما المتغيرات التي نتنبأ من خلالها فهي الدرجة الكلية المشاهدة والدرجة على الفقرة والتفاعل بينهما .

طريقة الانحدار اللوجستي (Logistic Regression)

تستخدم طريقة الانحدار اللوجستي(1990) Swaminthan & Roggers في الكشف عن دالة الفقرة التفاضلية (DIF) المنتظمة وغير المنتظمة للفقرة.

والمعادلة التي تحدد بواسطتها احتمالية الإجابة الصحيحة وفق المجموعة والمحك هي

$$P(u=1|x,g) = \frac{\exp\{b_0 + b_1x + b_{2g} + b_3(xg)\}}{1 + \exp\{b_0 + b_1x + b_{2g} + b_3(xg)\}}$$

حيث أن الاستجابة U تأخذ القيم (١= النجاح في الفقرة ، ٠= غير ذلك) .

U وترمز للاستجابة للفقرة .

G ترمز للمجموعة التي يتمي إلها الفرد .

X هو المحك .

Xg يرمز للتفاعل بين المحك والمجموعة التي ينتمي إليها الفرد .

ويمكن إعادة صياغة المعادلة باستخدام اللوجت (Logit)على النحو التالي بحيث تصبح معادلة خطية

$$\ln\left[\frac{\pi i}{(1-\pi i)}\right] = b_0 + b_1x + b_2g + b_3(xg)$$

حيث أن

$$\pi_i = p(u=1|x,g)$$

All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

المنارة للاستشارات
www.manaraa.com

ويدل الرمز X عادة إلى الدرجة الكلية الملاحظة
أما المتغير g فيدل على المجموعة التي ينتمي إليها الفرد .
وتعطى القيمة ١ للمتغير g إذا كان المفحوص أحد أفراد المجموعة الأولى.
وصفر إذا كان المفحوص أحد أفراد المجموعة الثانية .
Xg هي ناتج تفاعل المتغيرين g و X .

ومن أكثر استخدامات الانحدار اللوجستي شيوعا هي التي يتم فيها اعتبار الدرجة الكلية الملاحظة (Observed Total Score) متغيرا شرطيا، فلكل فقرة تتم مطابقة النموذج في المعادلة رقم (٢) أولا من خلال ربط الفقرة مع الدرجة الكلية X .
ثم يتم فحص وجود الدالة التفاضلية للفقرة من خلال استخدام χ^2 بدرجات حرية تساوي ٢ للتحقق من تحسن مطابقة النموذج عند إضافة متغير المجموعة g والتفاعل بين المتغيرين x g بشكل متزامن إلى النموذج(Zumbo, 2001) .

وتمتاز طريقة الانحدار اللوجستي مثل طريقة مانتل هينزل بوجود مقياس للدلالة الإحصائية ومقياس لحجم الأثر (Effict Size)، لكن خلافاً لطريقة مانتل هينزل فإن طريقة الانحدار اللوجستي يمكن استخدامها من اجل الكشف عن دالة الفقرة التفاضلية (DIF) غير المنتظمة دون أي تعديل على النموذج، كذلك يمكن استخدام الدرجة الكلية الملاحظة أو معلم القدرة التي يتم تقديره من خلال نظرية استجابة الفقرة كمتغير للمقارنة (Matching Variable) .
وتمتاز كذلك بأنها تحتاج إلى عينات أصغر وبأنها أقل تعقيدا من نماذج نظرية استجابة الفقرة، وتستخدم دالة اللوجت (Logit Function) في تحويل الاحتمالات من المدى (٠ ، ١) إلى المدى ($-\infty$ ، ∞) في عملية التنبؤ الخطي، وحيث أن النموذج يكون خطيا عند استخدام اللوجت فإن تفسير معالم الانحدار اللوجستي يكون عادة من خلال اللوجت .

و من الممكن أن يتم استخدام نموذج الانحدار اللوجستي في ما يلي :
١- التنبؤ باحتمال حصول حدث معين على أساس المتغيرات المستقلة .
٢- ترتيب المتغيرات المستقلة بناء على الأهمية النسبية لها .
٣- معرفة أثر التفاعل .
٤- فهم تأثير المتغيرات الضابطة المشتركة .

All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

المنارة للاستشارات
www.manaraa.com

وعند استخدام الانحدار اللوجستي في الكشف عن (DIF) فإنه يتم اعتبار الاستجابة على الفقرة كمتغير تابع، وكل من القدرة المشاهدة والمجموعة التي ينتمي إليها الفرد والتفاعل بينهما كمتغيرات مستقلة .

وتتضمن عملية فحص مستوى الدلالة عملية حذف هرمي للمتغيرات .

أولا تتم ملائمة النموذج الكامل الذي يتضمن القدرة والمجموعة والتفاعل بينهما.

ثانيا يتم حذف التفاعل من النموذج ليعطي النموذج الأول المختزل (R1) الذي يضم القدرة والمجموعة فقط.

ثالثا يتم حذف متغير المجموعة من النموذج ليعطي النموذج الثاني المختزل (R2) الذي يضم القدرة فقط.

هذه النماذج الثلاث تتم مقارنتها باستخدام G^2 الذي يمثل التغير في لوجريتم الأرجحية (2loglikelihood) فعند مقارنة اثنين من النماذج المتداخلة (Nested Models) تكون لقيم G^2 توزيع مربع كاي تقريبا. وتكون قيمة درجات الحرية مساوية لعدد المعالم الإضافية في النموذج الأكبر، وبالتالي تكون درجات الحرية تساوي ١ عند مقارنة النموذج (R1) مع النموذج الكامل لفحص(DIF) المنتظم وعند مقارنة النموذج (R1) مع النموذج لفحص (DIF) غير المنتظم، وتساوي اثنان عند مقارنة النموذج الكامل مع النموذج(R2) .

كما أن طريقة الانحدار اللوجستي تعتبر من الطرائق التي تتمتع بالمرونة في بناء المزيد من النماذج المعقدة، على سبيل المثال يمكننا في الانحدار اللوجستي أن نستخدم متغيرات أخرى كمتغيرات محكية إضافة إلى متغير القدرة للحصول على دقة أكثر، كذلك فإن طريقة الانحدار اللوجستي من الممكن أن تستخدم في الفقرات ذات الاستجابة المتعددة (Chaimongkol, 2005) .

ولقد اقترح زومبو وتوماس Zumbo & Thomas (المشار إليهما في Hidalgo & López-Pina, 2004) استخدام (ΔR^2) من أجل الحكم على مقدار (DIF) المنتظم وغير المنتظم، حيث يتم حساب (ΔR^2) من خلال طرح R^2 للنموذج الكامل الذي يتضمن القدرة والمجموعة والتفاعل بينهما من R^2 للنموذج الذي يضم القدرة فقط، حيث اقترحا المعايير التالية :

النوع A من الفقرات –الفقرات ذات DIF المهمل : ΔR^2 < ٠,١٣ .

النوع B من الفقرات –الفقرات ذات DIF المتوسط : ٠,١٣ ≤ ΔR^2 ≤ ٠,٢٦ .

النوع c من الفقرات –الفقرات ذات DIF الكبيـــر : ΔR^2 > ٠,٢٦

All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

المنارة للاستشارات
www.manaraa.com

ثانيا طرائق الكشف عن تحيز الفقرة التي تستند إلى نظرية استجابة الفقرة .

تتبنى الطرائق الكلاسيكية للكشف عن DIF اختبارات إحصائية تهدف إلى التحقق من عدم وجود DIF عند كل قيم x من خلال التعريف التالي :-

$E_R [Y| X] = E_F [Y| X]$

حيث تشير $E_R [Y| X]$ و $E_F [Y| X]$ إلى القيمة المتوقعة (y) من الانحدار على الدرجة المشاهدة x للمجموعة الضابطة و التجريبية .

أما طرائق النظرية الحديثة في القياس فتعتمد على اختبارات مباشرة تهدف إلى التحقق من عدم وجود التحيز وذلك عند كل قيم θ من خلال التعريف التالي

$E_R [Y|\theta] = E_F [Y | \theta]$

حيث تشير $E_R [Y| \theta]$ و $E_F [Y| \theta]$ إلى القيمة المتوقعة (y) من الانحدار على القدرة θ للمجموعة الضابطة و التجريبية .

(Chang & Mazzeo & Rouses 1996)

وتقدم نماذج النظرية الحديثة طرائق جذابة للكشف عن دالة الفقرة التفاضلية، وبناء مقاييس أقل تحيزا، فتحت شروطا معينة فإن هذه النماذج تعطينا العديد من المميزات

- معالم فقرات لا تتغير مع تغير العينة التي تستخدم لتقدير المعالم .
- القدرة لا تتغير بتغير الفقرات التي استخدمت في تقديرها .
- النظرية الحديثة خلافا للنظرية الكلاسيكية تربط معلم الصعوبة مع معلم القدرة على سلم مشترك.
- إمكانية تقدير الثبات من خلال الخطأ المعياري الذي يمكن حسابه عند كل مستوى من مستويات القدرة .
- دالة المعلومات (Information Function) التي تساعد على اختيار أفضل للفقرات في المقاييس التكيفية (Tailor Measures) .
- القدرة على تحديد الفقرات الحساسة للتغير، والتي تقيس السمة عند كل المستويات .

(Teresi, Klenman & Ocepek-Weliksonm , 2000 ; Hambelton 1985)

All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

المنارة للاستشارات
www.manaraa.com

ومن المهم عند الحديث عن النظرية الحديثة أن نأخذ بعين الاعتبار الافتراضات التي تقوم عليها وأن نتحقق من مطابقة النموذج المستخدم مع البيانات حتى نحصل على هذه المزايا النظرية (Gibson , 1998)، فعدم تطابق النموذج يمكن أن يؤدي إلى أخطاء في الكشف عن دالة الفقرة التفاضلية (Teresi, 2004) .

وفيما يلي عرض لافتراضات النظرية الحديثة، وبعض نماذجها وشرح لبعض مفاهيمها :

افتراضات النظرية الحديثة

وتقوم نظرية استجابة الفقرة على عدة افتراضات هي

١- أحادية البعد

أي أن هنالك قدرة واحدة فقط ضرورية من أجل تفسير أداء المفحوصين على الاختبار، وفي الواقع العملي من الصعب تحقيق هذا الافتراض بشكل تام حيث أن هنالك عوامل أخرى تؤثر على الأداء على الاختبار مثل العوامل المعرفية والشخصية ، وهنالك عدة طرائق تستخدم للتحقق من أحادية البعد أهمها التحليل العاملي، حيث يتم الحكم على أحادية البعد من خلال الرسم البياني للقيم المميزة للعوامل التي تفسر التباين في الأداء على الاختبار فكلما كانت القيمة المميزة للعامل الأول أكبر بشكل كبير من بقية العوامل كان ذلك دلالة على أحادية البعد .

٢- الاستقلال الموضعي

وهو افتراض مكافئ لافتراض أحادية البعد، وهو ينص على أن استجابة المفحوص لفقرات الاختبار المختلفة مستقلة إحصائيا، أي أن استجابة المفحوص على أي فقرة من الاختبار لا تتأثر سلبا أو إيجابا بالاستجابة على بقية الفقرات . ويتحقق هذا الافتراض رياضيا عندما تكون احتمالية الحصول على نمط معين من الاستجابة تساوي ناتج حاصل ضرب احتمالية استجابة الفرد على كل فقرة من فقرات الاختبار. ومن الجدير بالذكر أن فرضية الاستقلال الموضعي لا تعني عدم وجود ارتباط بين فقرات الاختبار لدى المفحوصين من كافة مستويات القدرة المختلفة، لكن عدم وجود ارتباط بين الفقرات يكون عند مستوى قدرة ثابت.

٣- منحنى خصائص الفقرة

وهو دالة رياضية تربط بين احتمالية الاستجابة الصحيحة للفقرة مع قدرة المفحوص التي يتم تقديرها من خلال الاختبار، وهو يمثل الانحدار غير الخطي لدرجة الفقرة على القدرة أو السمة التي يقيسها الاختبار.

All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

المنارة للاستشارات
www.manaraa.com

٤- عامل السرعة

وينص هذه الافتراض على أن الاختبار لم يتم تطبيقه تحت عامل السرعة ، أي أن المفحوص فشل في الاستجابة على الفقرة بسبب قدرته وليس بسبب فشله في الوصول إلى هذه الفقرة. ومن أساليب التحقق من ذلك دراسة العلاقة بين درجات الاختبار عند وضح حدود زمنية للإجابة مع الدرجات على الاختبار عند عدم وضح محددات زمنية للإجابة .
(Hambleton & Swaminathan, 1985)

نماذج نظرية استجابة الفقرة

تمثل الطريقة التي يستجيب فيه المفحوص على الفقرة إحدى الأساليب التي يتم من خلالها تصنيف نماذج نظرية استجابة الفقرة، فهنالك ثلاثة مستويات شائعة للاستجابة على الفقرة

- الاستجابة الثنائية مثل الأسئلة الموضوعية(كأسئلة الاختيار المتعدد أو أسئلة الصح والخطأ) التي تكون فيها الاستجابة صحيحة أو خاطئة .
- الاستجابة المتعددة كالمقاييس التي تبنى وفقا لتدريج ليكرت.
- نظام الاستجابة المتصل مثل أن يطلب من المستجيب أن يضع إشارة صح على نقطة من متصل لتقدير ظاهرة معينة، وفيما يلي عرض لبعض نماذج الاستجابة الثنائية .

١- النموذج اللوجستي ثلاثي المعالم

ويتم فيه التعبير عن منحنى خصائص الفقرة من خلال المعادلة التالية :

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

حيث أن

bi = معلم الصعوبة، ويمثل النقطة على سلم القدرة التي يكون احتمال الاستجابة الصحيحة للمفحوص على الفقرة عندها يساوي $(1+ci)/2$.

ai = معلم التمييز ، وهو يتناسب طرديا مع ميل $Pi(\theta)$ عند النقطة $\theta = bi$، من ناحية نظرية يعرف معامل التمييز على سلم يمتد من $(-\infty , +\infty)$، والفقرات ذات التمييز السالب يتم حذفها من الاختبار، ومع ذلك من غير المعتاد أن يتم الحصول على فقرات يزيد معامل تمييزها عن ٢، القيمة العالية لمعلم التمييز تؤدي إلى انحدار قوي لمنحنى خصائص الفقرة والقيمة المنخفضة تؤدي إلى انحدار خفيف للمنحنى.

D = ١,٧ .

All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

المنارة للاستشارات
www.manaraa.com

ci = معلم التخمين، وهو الخط التقاربي السفلي لمنحنى خصائص الفقرة، ويمثل احتمالية الاستجابة الصحيحة على الفقرة للأفراد منخفضي القدرة .

٢-النموذج اللوجستي ثنائي المعالم

تم اقتراح هذا النموذج من قبل بيرنبام (Birnbaum, 1968)

ويأخذ فيه منحنى خصائص الفقرة دالة ثنائية المعالم كما في المعادلة التالية :

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

٣- النموذج الأحادي (نموذج راش)

هذا النموذج يمثل حالة خاصة من النموذج الثلاثي والثنائي ، حيث يفترض بالإضافة إلى الافتراضات التي تم الحديث عنها سابقا أن

-الفقرات لها نفس معامل التمييز

-التخمين في حدوده الدنيا يساوي صفر.

ويتم التعبير عنه من خلال المعادلة

$$Pi(\theta) = \frac{1}{1+e^{-D(\theta-bi)}}$$

(Hambleton & Swaminathan, 1985)

ومن الصعوبات التي تحد من استخدام الطرائق التي تستند إلى نظرية استجابة الفقرة حاجتها إلى برامج كمبيوتر خاصة نظرا لكونها معقدة رياضيا ، وقد أشار كاملي وشبرد Camilli and Shepard (1994) إلى أن العينات الصغير التي تقل عن ٥٠٠ يمكن أن تعطينا نتائج غير موثوقة ، وإلى ضرورة ألا تقل العينة في المجموعة الأصغر عن ١٠٠٠ فرد حتى نحصل على نتائج مستقرة . (Roever , 2005)

وقد قام شبرد Shepard (1981) وإرنسون Ironson (1982, 1983) (المشار إليهم في Hambelton, 1985) بمراجعة الطرائق التي تستند إلى النظرية الحديثة للكشف عن DIF وجد أن هذه الطرائق تتبع الإجراءات الأساسية التالية: للكشف عن DIF .

١- مقارنة منحنيات خصائص الفقرة . (Item Characteristic Curve).

٢- مقارنة معالم الفقرات .

٣- مقارنة مطابقة النموذج للبيانات في المجموعات المختلفة.

All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

المنارة للاستشارات
www.manaraa.com

وفيما يلي عرض لبعض طرائق الكشف عن دالة الفقرة التفاضلية التي تستند إلى النظرية الحديثة في القياس

طريقة راش

يستخدم نموذج راش أحادي المعلمة لفحص دالة الفقرة التفاضلية(DIF) عندما تكون العينة صغيرة نسبيا لاستخدام النموذج الثنائي أو عندما لا تطابق البيانات النموذج الثنائي، إلا انه من الضروري استخدام طرائقا أخرى مع هذه الطريقة إذا كان من المتوقع ظهور دالة الفقرة التفاضلية (DIF) غير المنتظمة(ذلك أن النموذج الأحادي يفترض أن معامل التمييز متساو لدى كل الفقرات وبالتالي فإن منحنيات خصائص الفقرة لا يمكن أن تتقاطع، وبالتالي لا يمكن أن تكشف عن دالة الفقرة التفاضلية(DIF) الناتجة عن الاختلاف في معلم التمييز) ، المبدأ الأساسي في حساب (DIF) باستخدام النموذج الأحادي يعتمد على استخدام اختبار T فلو فرضنا أن الفقرة I لها اثنـــان من التقديرات لمعلم الصعوبة (d_{i1}- d_{i2}) ، وذلك للمجموعة الأولى والمجموعة الثانية بخطأ معياري S1 ، S2 على التوالي فإن صيغة اختبار (T TEST) سوف تكون كالتالي

$$T_{12} = \frac{di_1 - di_2}{\sqrt{\{(Si_1 \times Si_1) + (Si_2 \times Si_2)\}}}$$

حيث أن $(S_{i1}^2 + S_{i2}^2)^{1/2}$ تمثل تقدير للخطأ المعياري للفرق بين d_{i1} و d_{i2} ثم تتم بعد ذلك مقارنة قيمة T_{12} مع القيمة الحرجة لـــ t .

مع ذلك فإن القيمة الحرجة لمؤشر (DIF) تتأثر بحجم العينة، فعند استخدام عينة كبيرة يمكن الحصول على قيمة دالة إحصائيا لكنها ليست ذات دلالة عملية، لذلك بدلا من استخدام القيمة الحرجة لـــ t التي تساوي ١,٩٦ عند مستوى الدلالة ٠,٠٥ فإننا نستخدم الفرق ٠,٥ لوجت (0.5 Logit) للعينات الكبيرة (N>200) حيث يتم تفسير القيمة (0.5 Logit) على أنها الفرق في المعايرة بين المجموعتين، فإذا كانت الفروق في المعايرة أقل من (0.5 Logit) فإن قيمة تعتبر مهملة، أما إذا كانت قيمة الفرق أكبر من (0.5 Logit) فإن ذلك يدل على وجود (DIF) وبالتالي تتم مراجعة الفقرة أو حذفها من الاختبار (Berk, 1984 ; Sheilai,Teresa, & Gershon 2005) .

All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

المنارة للاستشارات
www.manaraa.com

طريقة لورد / مربع كاي (Lord'S Chi-Sqaure) .

اقترح لورد طريقة لفحص الفرضية تفترض أن معلما الصعوبة والتمييز (a's b's) للفقرة I في المجموعة الأولى تساوي معلما الصعوبة والتمييز (a's b's) في المجموعة الثانية .
وذلك من خلال حساب مربع كاي

$$x^2 = \quad V i^1 \quad \sum_i^{-1} vi$$

حيث أن $V i^1$ يمثل المتجه $\{ b_{i1}-b_{i2} , a_{i1}- a_{i2}\}$

$\sum_i^{-1}$ يمثل معكوس مصفوفة التباين والتباين المشترك لـ $b_{i1} - b_{i2} , a_{i1} - a_{i2}$.

ولأن a_{i1} & b_{i1} مستقلين عن a_{i2} & a_{i2} فإن

$$\sum_i = \sum_{i1} + \sum_{i2}$$

حيث أن $\sum_{i1}$ تمثل مصفوفة التباين والتباين المشترك لـ a_{I1} و b_{I1}

و $\sum_{i2}$ تمثل مصفوفة التباين والتباين المشترك لـ a_{I2} و b_{I2}

(Teresi, Klenman & Ocepek-Weliksonm, 2000)

طريقة مقارنة منحنيات خصائص الفقرة

الفقرة التي تكون فيها منحنيات خصائص الفقرة متماثلة في كل مجموعة هي فقرة لا يوجد فيها (DIF ، وهنالك عدة أساليب تستخدم من أجل مقارنة منحنيات خصائص الفقرة في المجموعات المختلفة فالتقييم البصري لمنحنيات خصائص الفقرة قد يكون وسيلة ناجحة للحكم على وجـود DIF وكذلك تحديد نوعه منتظم أو غير منتظم ، كما يمكن استخدام الأساليب الاستدلالية للحكم على وجود (DIF) مثل

أ-إيجاد القيمة المطلقة للفروق بين المنحنيات .

ب-حساب المساحة التي يكون فيها منحنى خصائص الفقرة (ICC) للمجموعة المستهدفة أعلى من المجموعة المرجعية .

ج-حساب المساحة التي يكون فيها منحنى خصائص الفقرة (ICC) للمجموعة المرجعية أعلى من المجموعة المستهدفة.

د- إيجاد الجذر التربيعي لمجموع مربع الفروق بين منحنيات خصائص الفقرة (ICC).

All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

المنارة للاستشارات
www.manaraa.com

إن حساب المساحة بين منحنيات خصائص الفقرة يتطلب أن يكون معالم الفقرات على نفس السلم، وإلا فإن الفروق في المساحة قد تكون نتيجة لاختلاف تدريج المعالم في كلا المجموعتين. وتعتبر طريقة الربط المكرر (Iterative Linking) أفضل الطرائق المستخدمة في ربط معالم الفقرات في كلا المجموعتين ذلك أنها تعمل على ربط معالم الفقرات بالاعتماد فقط على الفقرات غير المتحيزة، وهذه الطريقة تسير وفقا لعدد من الخطوات كما يلي :

- تقدير المعالم للفقرات والأفراد بشكل منفصل لكل مجموعة،
- ربط معالم الفقرات في كلا المجموعتين على سلم واحد.
- حساب الإحصائيات الخاصة بالتحيز وإيقاف العملية إذا لم يوجد DIF في أي فقرة.
- حذف الفقرات التي وجد فيها DIF ، وإعادة ربط معالم الفقرات على سلم مشترك مرة أخرى اعتمادا على الفقرات التي لم يظهر فيها التحيز فقط .
- حساب مؤشرات التحيز مرة أخرى.
-

وتتكرر عملية إعادة ربط المعالم حتى نصل إلى نفس المجموعة من الفقرات غير المتحيزة في محاولتين متتاليتين (Gibson , 1998; Kim & Cohen, 1993 ,1995).

وقد قدم رندر Runder طريقة لحساب الفرق في المساحة بين منحني خصائص الفقرة وفقا للخطوات التالية :

١-اختيار نموذج استجابة الفقرة الملائم للبيانات .

٢-تقدير معالم الفقرات بشكل منفصل لكل مجموعة من المجموعات .

٣-معادلة التقديرات التي تم الحصول عليها من المعايرات المنفصلة من أجل الحصول على تدريج مشترك للمعالم والقدرة .

٤- تقسيم متصل القدرة (-٣ ،+٣) إلى فئات وباتساع ($\Delta\theta$=0.005) .

٥- تحديد مركز كل فئة (θ_K) وحساب ارتفاع منحني خصائص الفقرة في كلا المجموعتين $P_{i1}(\theta_k)$ & $P_{i2}(\theta_k)$.

٦- إيجاد الفرق بين المساحتين وفق المعادلة التالية:

$$A_{1i} = \sum_{\theta=-3}^{+3} | P_{i1}(\theta_K) - p_{i2}(\theta_K) | \Delta\theta$$

وقد قدم لن وزملاؤه Linn et al. معادلة بديلة لحساب الفرق بين المساحتين كالتالي :

All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

المنارة للاستشارات
www.manaraa.com

$$A_{2i} = \sum_{\theta=-3}^{+3} \{[P_{i1}(\theta_K) - p_{i2}(\theta_K)]^2 \Delta\theta\}^{1/2}$$

(Hambleton & Swaminathan 1985 , 1991)

وقد توصل راجو Raju (1988) إلى اشتقاق صيغ لحساب المساحة الإشارية (Signed Area) والمساحة غير الإشارية (Unsigned Area) (في المساحة الإشارية يتم حساب المساحة اعتمادا على الفروق بين منحنيات استجابة الفقرة أما المساحة غير الإشارية فيتم حساب المساحة اعتمادا على القيمة المطلقة للفروق بين المنحنيات) من خلال اللجوء إلى التكامل كالتالي :

$$SA = (1-c)\,(b_2 - b_1)$$

$$UA = (1-c)\left|\frac{2(a_2-a_1)}{D\,a_1\,a_2}\ln\left(1+\exp\left(\frac{(D\,a_1a_2(b_2-b_1)}{a_2-a_1}\right)\right)-(b_2-b_1)\right|$$

وعند تطبيق النموذج ثنائي المعالم فإنه لا وجود للتخمين (المقدار c)، وعليه تصبح الصيغ كالتالي :

$$SA = (b_2 - b_1)$$

$$UA = \left|\frac{2(a_2-a_1)}{D\,a_1\,a_2}\ln\left(1+\exp\left(\frac{(D\,a_1a_2(b_2-b_1)}{a_2-a_1}\right)\right)-(b_2-b_1)\right| \quad if\ a_1 \neq a_2$$

وفي حالة نموذج راش الأحادي تصبح الصيغ كالتالي:

$$SA = (b_2 - b_1)$$

$$UA = |b2 - b1|$$

مع ملاحظة أن استخدام هذه الصيغ في النموذج الثلاثي لحساب المساحة مقيد بافتراض تساوي معلم التخمين في المجموعة الأولى والثانية (c = c1 = c2) ، ففي حالة عدم تساوي معلمي التخمين في المجموعتين فإن المساحة الإشارية تساوي +∞ أو −∞ والمساحة غير الإشارية =

All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

المنارة للاستشارات
www.manaraa.com

+∞ (Raju ,1988) . وتزودنا المساحة الإشارية بمعلومات عن دالة الفقرة التفاضلية المنتظمة أما المساحة غير الإشارية فتزودنا بمعلومات عن دالة الفقرة التفاضلية غير المنتظمة . (Sheilai, Teresa & Gershon 2005)

وقد اقترح هامبلتون وروجر Hambleton & Rogers (1989) طريقة للحكم على وجود دالة الفقرة التفاضلية أو عدم وجوده بناء على المساحة المحسوبة ، وذلك بإيجاد نقطة قطع (Cut-Off Value) للفروق في المساحة، حيث يتم حساب تلك القيمة بتقسيم المجموعة المرجعية إلى مجموعتين عشوائيتين متكافئتين، ثم يتم تقدير منحنيات خصائص الفقرة لكل مجموعة بشكل منفصل ، ويمكن تكرار عملية اختيار المجموعة وتقسيمها إلى مجموعتين متكافئتين عشوائيا عدد من المرات ، ويتم في كل مرة تقدير معالم الفقرة لكل مجموعة بشكل منفصل، أكبر قيمة للمساحة بين المنحنيين في المجموعات المختارة يتم اعتمادها كنقطة قطع ، فإذا كانت الفرق في المساحة بين المنحنيين أكبر من قيمة نقطة القطع ، دل ذلك على وجود دالة الفقرة التفاضلية (DIF) ، أما إذا كان الفرق في المساحة بين المنحنيين أصغر دل ذلك على عدم وجود دالة الفقرة التفاضلية (DIF) .

لكن هذه المعايير تعتبر بشكل عام اعتباطية (Arbitrary) بسبب عدم وجود اختبارات إحصائية للحكم على دلالة الفروق بين التقديرات المحسوبة للمساحة. (Sheilai, Teresa & Gershon 2005)

All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

المنارة للاستشارات
www.manaraa.com

طريقة نسبة الأرجحية

(Item Response Theory Likelihood-Ratio Tests For Differential Item Functioning IRTLRDIF)

تستخدم هذه الطريقة الإحصـــائي (Likelihood-Ratio Test) من أجل فحص الفرضـية الصفرية التي تنص على عدم وجود فروق في معالم الفقـرات بـين المجموعـات، فالدلالـة الإحصائية لـ (Likelihood-Ratio Test) تشير إلى وجود DIF في الفقرة.
وتعتمد صياغة الفرضية الصفرية لهذا الطريقة على تعريف لورد للتحيز الذي ينص علـى أن وجود فروق في دالة استجابة الفقرة بين مجموعة وأخرى يدل على وجـود التحيـز. ذلـك أن مجموعة معالم الفقرة تناظر دالة استجابة الفقرة أو منحنى خصائص الفقـرة. وهـي تـصلح للكشف عن دالة الفقرة التفاضلية (DIF) في الفقرات ثنائيـة الاسـتجابة أو الفقـرات متعـددة الاستجابة حيث تتماشى هذه الطريقة مع نموذج استجابة الفقرة الثلاثي ونموذج استجابة الفقـرة المتدرج (Samejima,s (1969, 1997) Graded Model) الذي يعتبر النموذج الثنـائي حالة خاصة لكل منهما.

اختبار (Likelihood Ratio) يتكون من الفـرق بـين اثنيـن مـن النمـاذج (Compact & Augmented) كالتالي :

$$Gj2 = -2 \text{ Log L (Compact Model)} + 2 \text{ Log L (Augmented Model)}$$

وله توزيع مماثل لتوزيع مربع كاي χ^2 مع درجات حرية تساوي الفرق بين عدد المعالم في النموذج الأول والنموذج الثاني .

(Compact Model) مبني على افتراض تساوي المعالم في المجموعة المرجعية والمجموعة المستهدفة ، فعند مطابقة البيانات مع معالم منحنيات الاستجابة للفقرات تضبط بحيث تكون متساوية في كلا المجموعتين لكل الفقرات، ويتم اختيار التوزيع الطبيعي الذي متوسطه = صفر وانحرافه المعياري =١ (0,1) N كسلم لمعالم الفقرات في المجموعة المرجعية ، أما في المجموعة المستهدفة فيتم تقدير المتوسط الحسابي والانحراف المعياري للتوزيع الطبيعي للمجتمع بشكل متزامن مع معالم الفقرة .

All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

المنارة للاستشارات
www.manaraa.com

أما النموذج (Augmented Model) فيفترض أن جميع المعالم في المجموعتين المرجعية والمستهدفة متساوية باستثناء الفقرة التي يتم دراستها فيتم تقدير المعالم لها بشكل مختلف في المجموعتين. ومرة أخرى يتم اختيار التوزيع الطبيعي الذي متوسطه = صفر وانحرافه المعياري =١ N (0,1) كسلم لمعالم الفقرات في المجموعة المرجعية ، أما في المجموعة المستهدفة فيتم تقدير المتوسط الحسابي والانحراف المعياري للتوزيع الطبيعي للمجتمع بشكل متزامن مع معالم الفقرة (Thissen , 2001; Kim & Cohen, 1995; Finch & French, 2007) .

اختبار تحيز الفقرة المتزامن (The Simultaneous Item Bias Test SIBTEST)

إن اختبار SIBTEST هو اختبار غير بارامتري يستند إلى نموذج IRT متعدد الأبعاد ، ولا يحتاج إلى تقدير كل من معالم الفقرات و القدرة أو القيام بالعديد من الحسابات المعقدة والمكثفة، ويمكن من خلاله الكشف عن DIF لفقرة واحدة أو للاختبار ككل ، وكذلك تقدير مقداره من خلال المؤشر الإحصائي β.

وقد تم تصميم SIBTEST في الأساس من أجل الكشف عن DIF المنتظم (أي عندما تكون احتمالية الاستجابة الصحيحة لدى إحدى المجموعات أقل أو أكثر في كل مستويات القدرة).

ويتألف نموذج ITR بشكل عام من عنصرين

أ- معلم القدرة حيث تفترض نظرية استجابة الفقرة أحادية البعد .

ب- دالة الاستجابة للفقرة التي تحدد احتمالية الاستجابة الصحيحة للفقرات.

يتضمن النموذج متعدد الأبعاد على الأقل عامل إضافي دخيل (Nuisance Determinant) بالإضافة إلى القدرة المستهدفة (Target Ability) ، ويتم التعبير عن القدرة بالمتجه (θ ,η) حيث تشير θ إلى القدرة المستهدفة و η تشير إلى القدرة الدخيلة ، كما يرمز إلى دالة الاستجابة للفقرة (Item Response Function , IRF) بـ Pi (θ ,η) ، ويفترض هذا النموذج أن كل الفقرات تعتمد على θ وان فقرة أو أكثر تعتمد على η ، ومثل النماذج التي تفترض أحادية البعد يفترض هذا النموذج الاستقلال الموضعي وأن Pi (θ ,η) تزداد بزيادة θ و η عندما تعتمد الفقرة على القدرتين وتزداد بزيادة θ عندما تعتمد الفقرة على θ .

التحيز من خلال النموذج السابق يتشكل من خلال ثلاث مكونات

- احتمالية وجود التحيز تكمن في توزيع القدرة المستهدفة والقدرة الإضافية في كلتا المجموعتين الضابطة والتجريبية .

All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

المنارة للاستشارات
www.manaraa.com

- احتمالية وجود التحيز يتم التعبير عنه من خلال الفقرات التي تعتمد الاستجابة عليها على القدرة الإضافية.
- طريقة إعطاء الدرجات والنظر إليها كتقدير للقدرة المستهدفة ينقل التحيز من الفقرة إلى الاختبار (Shealy & Stout, 1993).

وعلى الرغم من أصول طريقة SIBTEST المرتبطة بالنظرية الحديثة ، إلا أنه يمكن فهمها بشكل أسهل من منظور النظرية الكلاسيكية، حيث سيتم استخدام الدرجة الحقيقية كبديل للسمة الكامنة، فمن الافتراضات الأساسية في النظرية الكلاسيكية حول الدرجة المشاهدة (X) أن $X = T + E$ حيث تشير T إلى الدرجة الحقيقية لأحد المفحوصين الذي تم اختياره بشكل عشوائي ، المتغير $E = X - T$ يشير إلى الخطأ في القياس الذي يفترض أن يكون الوسط الحسابي له يساوي الصفر في كلا المجموعتين .

الفروق بين الاستجابات الصحيحة على الفقرة بين المجموعة المرجعية والمجموعة المستهدفة عند نفس الدرجة الحقيقية والتي يرمز لها بالرمز $B(T)$ يتم التعبير عنها كالتالي

$$B(T) = P_R(T) - P_F(T)$$

حيث أن

$P_R(T)$ = احتمالية الاستجابة الصحيحة على الفقرة لدى المجموعة المرجعية عند الدرجة الحقيقية T .

$P_F(T)$ = احتمالية الاستجابة الصحيحة على الفقرة لدى المجموعة المستهدفة عند الدرجة الحقيقية T .

قيمة DIF الكلية β تعرف بأنها قيمة $B(T)$ المتوقعة عند توزيع ملائم للدرجة الحقيقة T، ويتم تفسير β كمتوسط لأفضلية للمجموعة المرجعية في احتمالية الاستجابة الصحيحة عند درجة حقيقية معينة . ويتم صياغة الفرضية الصفرية لـ SIBTEST كالتالي

$$E[B(T)] = \beta = 0$$

في طريقة SIBTEST يتم توزيع الفقرات إلى مجموعتين ، فقرات المقارنة (Matching Subtest) والفقرة أو الفقرات التي تتم دراستها (Suspect Subtest) ، لكل درجة مشاهدة (Observed Score) لمجموعة فقرات المقارنة يتم تقدير الدرجة الحقيقة $(T_R(s))$ للمجموعة المرجعية والدرجة الحقيقة $(T_F(s))$ للمجموعة المستهدفة بشكل منفصل باستخدام

All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

المنارة للاستشارات
www.manaraa.com

معادلة الانحدار الخطي للدرجة الحقيقية على الدرجة المشاهدة وفقا للنظرية الكلاسيكية .لكن إذا كان هنالك اختلاف في القدرة فإن $T_R(s)$ لن تكون مساوية لـ $T_F(s)$ لذلك تحتاج نسبة الاستجابات الصحيحة (Proportion-Right Scors) المقدرة للفقرة في كلا المجموعتين $P_R[T_R(s)]$ & $P_F[T_F(s)]$ إلى تعديل بحيث يستندا إلى نفس الدرجة الحقيقية $T(s)$ لذلك يتم أخذ الدرجة الحقيقة كمتوسط لـ $T_R(s)$ و $T_F(s)$ ، هذا التحويل يسمى من قبل شيلي وستاوت Shealy & Stout (1993) بـ Regression Correction .

(Rousses & Stout ,1996 ; Chang, Mazzeo & Rouses, 1996).

وللمؤشر β توزيع طبيعي متوسطه يساوي صفر وانحرافه المعياري يساوي ١ ، وهو يعكس كمية DIF الموجودة في الفقرة ، طريقة SIBTEST تسعى لاختبار الفرضية الصفرية التي تنص على أن قيمة المؤشر βuni تساوي صفر($\beta uni = 0$) ، في حين يمكن أن تنص الفرضية البديلة على أن ($\beta uni > 0$) أو ($\beta uni < 0$) أو ($\beta uni \neq 0$). عندما تكون قيمته موجبة يدل ذلك على وجود DIF ضد المجموعة المستهدفة ، أما قيمة β السالبة فتدل على وجود DIF ضد المجموعة المرجعية.

(The William Stout Institute For Measurement ,2002)

ولقد تبنى روز وستاوت Roussos and Stout (1996) الدليل التالي من أجل تصنيف DIF للفقرة الواحدة

النوع A من الفقرات –الفقرات ذات DIF المهمل : $|\beta uni| < ٠,٠٥٩$.

النوع B من الفقرات –الفقرات ذات DIF المتوسط : $٠,٠٥٩ \leq |\beta uni| < ٠,٠٨٨$.

النوع c من الفقرات –الفقرات ذات DIF الكبير: $|\beta uni| \geq ٠,٠٨٨$.

(Gierl, Bisanz , Bisanz , Boughton & Khaliq, 2001)

وفي عام ١٩٩٥ قام كل من Li &Stout (1995) بتعديل طريقة SIBTEST وذلك من خلال عرضهم لطريقة Crossing Sibtest الذي يمكن من خلالها الكشف عن دالة الفقرة التفاضلية غير المنتظمة وذلك بتقدير الفروق بين $E_R[Y|\theta]$ و $E_F[Y|\theta]$ عند كل مستويات القدرة θ ، فبغض النظر عن ثبات قيمة Dif أو عدم ثباتها يظهر DIF عندما تتغير إشارة الفرق بين احتمالية الاستجابة لدى كلا المجموعتين خلال مستويات القدرة (Tian, 1999).

All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

المنارة للاستشارات
www.manaraa.com

مشكلة الدراسة وأهميتها

تلعب الاختبارات دورا هاما في حياتنا اليومية، فالاختبارات تساهم بشكل كبير في القياس الموضوعي للقدرات والمواهب وإمكانات المفحوصين الذين يرغبون في الحصول على الوظائف والتعليم الملائم، ويعتبر صدق المقياس عامل أساسي لإعطاء معنى للدرجات وهو يتأثر بعوامل عدة من أهمها التحيز فأي اختبار يحتوى على فقرات تعطى أفضلية لأي مجموعة من المجموعات على حساب مجموعة لن يكون اختبارا يتسم بالصدق والعدالة، ويعتبر وجود الأداء التفاضلي للفقرة شرطا أساسيا للحكم على تحيز الفقرة، فعدم وجود دالة الفقرة التفاضلية (DIF) يعني أن الفقرة لن تكون متحيزة، أما في حالة وجوده فإن ذلك غير كاف من أجل الإعلان عن تحيز الفقرة، لكن ذلك يتطلب تحليل إضافي لتحيز الفقرة، ومن يقوم بمراجعة الأدب النظري المتعلق بالموضوع والأساليب المستخدمة في الكشف عن دالة الفقرة التفاضلية سوف يجد العديد من الطرائق الإحصائية التي تم تطويرها للكشف عن دالة الفقرة التفاضلية، والتي يمكن أن يلجأ إليها من يقوم ببناء وإعداد الاختبارات، هذا التنوع في طرائق الكشف عن دالة الفقرة التفاضلية يجعل من الصعب على الباحث أن يقرر أي الطرائق أفضل لعملية للتحقق من وجود دالة الفقرة التفاضلية في فقرات الاختبار خاصة أن الأدب التربوي المتوفر لم يعط إجابات شافية حول الطريقة المثلى للكشف عن دالة الفقرة التفاضلية وعن الشروط أو الظروف الملائمة لاستخدامها هذه الطرائق، وتأتي أهمية هذه الدراسة من كونها تهدف إلى مقارنة فاعلية أربعا من طرائق الكشف عن دالة الفقرة التفاضلية (DIF) اثنتين من هذه الطرائق تستندان إلى النظرية الكلاسيكية وهما طريقة مانتل هينزل (Mantel Haenszel) وطريقة الانحدار اللوجستي (Logistic Regression) واثنتين تستندان إلى النظرية الحديثة في القياس (نظرية الاستجابة للفقرة) وهما طريقة اختبار التحيز المتزامن (SIBTEST) وطريقة نسبة الأرجحية (IRTLRDIF) وفقا لعدد من المتغيرات هي توزيع القدرة وحجم العينة ونسبتها ومعامل التمييز ووجود دالة الفقرة ونوعها وبالتحديد فإن هذه الدراسة تهدف إلى الإجابة عن الأسئلة التالية :

١-ما الطرائق الأكثر فاعلية (مانتل هينزل Mantel Haenszel، والانحدار اللوجستي Logistic Regression، اختبار التحيز المتزامن SIBTEST، ونسبة الأرجحية IRTLRDIF) في الكشف عن دالة الفقرة التفاضلية من حيث الخطأ من النوع الأول؟

All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

المنارة للاستشارات
www.manaraa.com

٢-ما الطرائق الأكثر فاعلية(مانتل هينزل Mantel Haenszel، والانحدار اللوجستي Logistic Regression، اختبار التحيز المتزامن SIBTEST، ونسبة الأرجحية IRTLRDIF) في الكشف عن دالة الفقرة التفاضلية من حيث قوة الاختبار الإحصائي؟

٣-هل يوجد تأثير للمتغيرات (حجم العينة ، نسبة العينة بين المجموعة المرجعية والمجموعة المستهدفة، توزيع القدرة، معلم التمييز) على الخطأ من النوع الأول والقوة للطرائق الأربع (مانتل هينزل Mantel Haenszel، والانحدار اللوجستي Logistic Regression، اختبار التحيز المتزامن SIBTEST، ونسبة الأرجحية IRTLRDIF)؟

مصطلحات الدراسة

- **دالة الفقرة التفاضلية (Differential Item Function)**

الفرق في احتمالات الاستجابة الصحيحة للفقرة في المجموعات المختلفة عند مستويات متساوية في القدرة .

- **التحيـــــــز (Bias)**

خطـــأ منتـظم في القيـــاس أو التقـــــدير.

- **الخطـــأ من النوع الأول (Type I error)**

رفض الفرضية الصفرية وهي في واقع الأمر صحيحة، وإجـرائيا اعتبـرت النسبـــة المئوية لعدد مرات الكشـــف عن دالة الفقرة التفاضلية بشكل خطأ عند مستـــوى الدلالة ($\alpha \leq$ ٠,٠٥) الخطأ من النوع الأول .

- **القوة الإحصائية (Statistical Power)**

رفض الفرضية الصفرية وقبول الفرضية البديلة عندما تكون الفرضية البديلة صحيحة، وإجرائيا اعتبرت النسبة المئوية لعدد مرات الكشف عن دالة الفقرة التفاضلية بشكل صحيح عند مستوى الدلالة ($\alpha \leq$ ٠,٠٥) مؤشرا على القوة الإحصائية .

All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

المنارة للاستشارات
www.manaraa.com

الفصل الثاني

الدراســــــــات الســــــــابقة

All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

المنارة للاستشارات
www.manaraa.com

أولا الدراسات التي اعتمدت على بيانات حقيقية

قام ظاظا (٢٠٠٦) بدراسة هدفت إلى التحقق من فاعلية الإجراءات التحكيمية في الكشف عن الفقرات المتحيزة لمتغير الجنس، والمستوى الاقتصادي الاجتماعي، ومقارنتهما بالطرائق التجريبية التي تستند إلى إجراءات النظرية الكلاسيكية في القياس ونظرية الاستجابة للفقرة، ولتحقيق غرض الدراسة تم بناء تجمع من الفقرات مكون من ١٤٠ فقرة تقيس قدرات معرفية للأفراد في الفئة العمرية ما بين ١٥–١٨ سنة ، روعي في بناء ٨٧ فقرة منها أن تكون متحيزة إما للجنس أو للمستوى الاقتصادي الاجتماعي، وقد تم حذف ١٦ فقرة منها بناء على الخصائص السيكومترية للفقرات بعد تجريبها على عينة استطلاعية، وحذف ٣١ فقرة منها لعدم مطابقتها للنموذج أو ارتباطها الضعيف مع الدرجة الكلية، وقد تم تطبيق تجمع الفقرات على عينة مكونة من ٦٠٦ طالبا وطالبة من طلبة الصف التاسع والعاشر والأول ثانوي. اعتمدت إجراءات الكشف عن التحيز المبنية على الإجراءات التحكيمية على عينة مؤلفة من ١٦ حكما، أما الإجراءات التجريبية فقد اشتملت على ثلاث طرائق اثنتان تستندان إلى النظرية الكلاسيكية في القياس وهما طريقة تحويل صعوبة الفقرة (TID) وطريقة مربع كاي (x^2) عند تقسيم مستويات الدرجة الكلية إلى ٣ مستويات وطريقة تستند على نظرية الاستجابة للفقرة.

أظهرت نتائج الدراسة أن طريقة تحويل صعوبة الفقرة كشفت عن أكبر نسبة من الفقرات المتحيزة لكلا المتغيرين (الجنس ، والمستوى الاقتصادي والاجتماعي) والتي تطابق الكشف عنها مع التصنيف المفترض لها عند بنائها بنسبة (٥٧%)، تلتها طريقة الإجراءات التحكيمية بنسبة (٥٢%) ثم جاءت طريقة مربع كاي في المرتبة الثالثة بنسبة (٤٩%)، ثم جاءت طريقة حساب المساحة بين منحني خصائص الفقرة المرسوم وفق النموذج الثلاثي المعلمة في المرتبة الأخيرة بنسبة مقدارها (٤٥%).

كما أظهرت النتائج أن طريقة الإجراءات التحكيمية استطاعت أن تحقق المركز الأول فيما يتعلق بفاعلية الكشف عن الفقرات المتحيزة للجنس تلتها طريقة تحويل صعوبة الفقرة في المركز الثاني، وجاءت طريقة مربع كاي في المركز الثالث، ثم جاءت طريقة المساحة بين منحني خصائص الفقرة في المرتبة الأخيرة. وفيما يتعلق بفاعلية الطرائق في الكشف عن الفقرات المتحيزة لمتغير المستوى الاقتصادي الاجتماعي، فقد احتلت طريقة تحويل صعوبة الفقرة المركز الأول ، تلتها طريقة مربع كاي في المركز الثاني، وجاءت طريقة الإجراءات التحكيمية في المرتبة الثالثة، وجاءت طريقة حساب المساحة بين منحني خصائص الفقرة في المرتبة

All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

المنارة للاستشارات
www.manaraa.com

الأخيرة . وبينت النتائج أن فاعلية الطرائق في الكشف عن الفقرات غير المتحيزة قد اختلفت حيث احتلت طريقة المساحة بين منحنيي خصائص الفقرة المركز الأول تلتها طريقة تحويل صعوبة الفقرة، ثم جاءت طريقة مربع كاي (x^2) والإجراءات التحكيمية في المرتبة الثالثة.

كما قام الباحثان Gao & Wangs (2005) بدراسة هدفت إلى تقييم الأداء التفاضلي في اختبارات (Passage–Based Reading Testlets) بين الذكور والإناث باستخدام خمسة طرائق هي

(Polytomous Logistic Regression PLR ;ZUMBO,1999)

(Graded Response Model GRM ; SAMEJIMA, 1969,1997)

(Mantel-Haenszel MH ; Zwick, Donoghue, & Grima, 1993)

(Poly-SIBTEST; Chang, Mazzeo & Rouses, 1996)

(Testlet Response Theory Model TRTM; Wang, Bradlow & Wainer, 2002)

أجريت الدراسة على عينتين تم اختيارهما بشكل عشوائي من طلاب الصف الثاني ثانوي الذين طبق عليهم اختبار (English 30 Diploma) عام ١٩٩٧م في مقاطعة البرتا بكندا، كل عينة احتوت على ٣٠٠٠ مفحوص مقسمين بالتساوي بين الذكور والإناث (١٥٠٠ ذكر ، ١٥٠٠ إناث) . العينة الأولى تم اختيارها من أجل التحليل الأساسي، أما العينة الثانية فتم اختيارها من أجل توفير الصدق التقاطعي لنتائج الدراسة، وقد أظهرت نتائج الدراسة وجود اتساق في نتائج الكشف عن الأداء التفاضلي في ثلاث من الطرائق هي , PLR , MH POLY-SIBTEST ، وعدم وجود اتساق في نتائج الطريقتين الأخريين.

وفي دراسة قام بها البستنجي (٢٠٠٤) هدفت إلى إجراء مقارنة بين أربع طرائق للكشف عن دالة الفقرة التفاضلية، اثنتين من طرائق النظرية الكلاسيكية (طريقة مانتل هينزل وطريقة الانحدار اللوجستي) واثنتين من طرائق النظرية الحديثة في القياس (طريقة مؤشر الفرق في المساحية لمنحني خصائص الفقرة لمجموعتي الذكور والإناث لنموذج ثنائي المعالم وآخر ثلاثي المعالم، تمّ إعداد اختبار للقدرات العقلية مكون من ٩٠ فقرة موزع في ثلاث اختبارات فرعية تقيس القدرة اللفظية والتفكير الرياضي والقدرة المكانية، وتطبيقه على عينة

All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

المنارة للاستشارات
www.manaraa.com

مكونة من ٩٥٠ طالب، بعد ذالك قام الباحث باستخدام الطرائق التي سبق ذكرها للكشف عن دالة القفرة التفاضلية وقد أظهرت النتائج ما يلي :-

- اتفقت طريقة مانتل هينزل مع الانحدار اللوجستي في اعتبار الفقرة تظهر أداء تفاضليا في اختبار القدرة اللفظية في ٨ فقرات و٣ فقرات في اختبار التفكير الرياضي وفقرتين من اختبار القدرة المكانية ، بينما لم تتفق الطريقتان في أي فقرة من اختبار المفردات بالتضاد .
- اتفقت طريقة مانتل هينزل مع طريقة مؤشر الفرق في المساحة للنموذج ثنائي المعالم في ٣ فقرات لكل من اختبار القدرة اللفظية ، واختبار التفكير الرياضي ، واختبار القدرة المكانية، وفي فقرة واحدة من اختبار المفردات بالتضاد .
- اتفقت طريقة مانتل هينزل مع طريقة مؤشر الفرق في المساحة للنموذج ثلاثي المعالم في ٥ فقرات من اختبار القدرة اللفظية، و ٣ فقرات من اختبار التفكير الرياضي ، و٤ فقرات اختبار القدرة المكانية و٣ فقرات من اختبار المفردات بالتضاد .
- اتفقت طريقة الانحدار اللوجستي مع طريقة مؤشر الفرق في المساحة للنموذج ثنائي المعالم في فقرتين من اختبار القدرة اللفظية، و٣ فقرات من اختبار التفكير الرياضي، وفقرتين من اختبار القدرة المكانية، ولم تتفق الطريقتان في أي فقرة لاختبار المفردات بالتضاد .
- اتفقت طريقة الانحدار اللوجستي مع طريقة مؤشر الفرق في المساحة للنموذج ثلاثي المعلم في ٤ فقرات من اختبار القدرة اللفظية، و٥ فقرات من التفكير الرياضي و٣ فقرات من اختبار القدرة المكانية ، ولم تتفق الطريقتان في أي فقرة لاختبار المفردات بالتضاد.
- اتفقت طريقة مؤشر الفرق في المساحة للنموذج ثنائي المعالم مع طريقة مؤشر الفرق في المساحة للنموذج ثلاثي المعلم في فقرة واحدة في اختبار القدرة اللفظية، وفقرتين من اختبار التفكير الرياضي، و٣ فقرات من اختبار القدرة المكانية، ولم تتفق الطريقتان في أي فقرة لاختبار المفردات بالتضاد .
- كان الأداء التفاضلي بشكل عام في معظمه لصالح الذكور في اختبار التفكير الرياضي وفي اختبار القدرة المكانية، ولصالح الإناث في اختبار القدرة اللفظية .
- اختلف الأداء التفاضلي باختلاف مستوى القدرة في نوعي الأداء التفاضلي المنتظم وغير المنتظم، حيث كانت فروق الأداء التفاضلي في بعض الفقرات أكبر ما يمكن في المستويات العليا من القدرة وأقل ما يمكن في المستويات الدنيا من القدرة، وذلك في كل من الطرائق الأربع المستخدمة في الكشف عن الأداء التفاضلي، وكان الفرق في الأداء التفاضلي في بعض

All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

المنارة للاستشارات
www.manaraa.com

الفقرات التي ظهر فيها الأداء التفاضلي أكبر ما يمكن في المستويات المتدنية، وأقل ما يمكن في المستويات العليا من القدرة .

وفي دراسة قام بها Bertrand & Boiteau (2003) هدفت إلى إيجـــاد معايير لاختيار طريقة جيــدة وغير مكلفة للكشف عن دالة الفقرة التفاضلية كمعدلات الاستقرار الداخــلية (Within-Method Stability Rates) ومعدلات الاتفاق بين الطرائق (Between-Method Agreement Rates). قام الباحثان بدراسة الاستقرار الداخلي لبعض الطرائق التي تستند إلى نظرية استجابة الفقرة وبعض الطرائق التي تستند إلى النظرية الكلاسيكية في اثنين من السياقات الثقافية المختلفة، اعتمادا على بيانات دراسة (Third International Mathematics And Science Study) وبالتحديد على ٤٨ فقرة رياضيات من اختبار (TIMSS) مشتركة بين العامين ١٩٩٥م و١٩٩٩م.

وقد تم استخدام أربع من طرائق الكشف عن دالة الفقرة التفاضلية اثنان منها لا تستند إلى نظرية الاستجابة للفقرة وهي طريقة مانتل هينزل وطريقة الانحدار اللوجستي واثنتان تستندان إلى نظرية الاستجابة للفقرة طريقة المساحة (UPD index) و (Raju's NCDIF index) لمعرفة فيما إذا كانت الفقرات التي ظهر فيها DIF في عام ١٩٩٥م بين الطلاب الأمريكين واليابانيين هي نفس الفقرات التي ظهر فيها DIF في عام ١٩٩٩م .

وقد أظهرت نتائج الدراسة أن الطرائق التي لا تستند إلى نظرية الاستجابة للفقرة وبخاصة طريقة مانتل هينزل (ذات التكلفة المنخفضة) حازت معدلات اتفاق بين الطرائق مماثلة لتلك التي تم الحصول عليها للطرائق التي تستند على نظرية استجابة الفقرة ، كما أن معدلات الاستقرار (Stability Rates) للطرائق التي لا تستند إلى نظرية الاستجابة للفقرة كانت قريبة جدا من معدلات الاستقرار للطرائق التي تستند إلى نظرية الاستجابة للفقرة .

كما قام الجرادات (٢٠٠٣) بدراسة هدفت إلى المقارنة بين طريقة مانتل هينزل (MH) ، وطريقة تحويل صعوبة الفقرة (TID) في كشف الفقرات المتحيزة باختلاف بعد التحيز(الجنس ، والحقل الأكاديمي) ، وإلى تحري درجة اتفاق الطريقتين في تحديد الفقرات المتحيزة لكل مــن الجنس والحقل الأكاديمي، ولتحقيق هذا الهدف استخدمت بيانات تقنين اختبار كاليفورنيا لمهارات التفكيــر الناقــد (California Critical Thinking Skills CCTS) للبيئــة الأردنيــة ، والمكون من (٣٤) فقرة من نوع الاختيار من متعدد لكل منها أربعة بدائل ، والتي تم تطبيقهــا

All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

المنارة للاستشارات
www.manaraa.com

على (١٤٨٥) طالبا وطالبة، من حقلي العلوم والآداب في جامعة مؤتة (٧٨٩أنثى ، ٦٩٦ذكر، ٧٩٨آداب، ٦٨٧ علوم)، وقد أظهرت النتائج أن طريقة الصعوبة المحولة (TID) كانت أكثر فاعلية في القدرة على كشف الفقرات المتحيزة للجنس من طريقة مانتل هينزل استنادا إلى نسبة الفقرات المتحيزة المكتشفة كمعيار للفاعلية، حيث كانت قيمة $\chi 2$ دالة إحصائية عند (α =٠,٠٥)، وكشفت نتائج الدراسة بأن طريقة صعوبة الفقرة أشارت إلى وجود (١٣) فقرة متحيزة للجنس من أصل (٣٤) فقرة وبلغت نسبتها (٣٨,٢%)، وبالمقابل كان عدد الفقرات المتحيزة تبعا للجنس المكتشفة وفق طريقة مانتل هينزل (٦) فقرات من أصل (٣٤) فقرة ، وبلغت نسبتها (١٧,٦%) من فقرات الاختبار، وأن الطريقتين اتفقتا على أن هناك (٦) فقرات متحيزة وفقا الجنس بنسبة (٤٦%)، كما أشارت النتائج إلى أن طريقة صعوبة الفقرة كشفت عن ٦ فقرات متحيزة تبعا للتخصص الأكاديمي من أصل ٣٤ فقرة أي ما نسبته (١٧,٦%)، وأن طريقة الصعوبة المحولة كشفت عن ١٠ فقرات أي ما نسبته (٢٩,٤%)، كما أظهرت نتائج الدراسة أن الطريقتين متماثلتان في القدرة على كشف الفقرات المتحيزة للحقل الأكاديمي، حيث كانت قيمة $\chi 2$ دالة إحصائيا (α = ٠,٠٥) حيث اتفقت الطريقتان على تحيز (٤) فقرات متحيزة وفقا للحقل الأكاديمي، أي بنسبة (١١,٨%) من مجموع فقرات الاختبار
أما فيما يتعلق بتحديد درجة الاتفاق ، فقد أظهرت النتائج أن الطريقتين متفقتان في اعتبار الفقرة متحيزة للجنس حيث بلغ معامل الارتباط فاي (Φ θ= ٠,٥٩) ، ومتفقتان حسب متغير الحقل الأكاديمي ، حيث بلغ معامل ارتباط فاي (Φ θ = ٠,٣٨) .

كما أجرى القرعان (١٩٩٧) دراسة هدفت إلى مقارنة فاعلية ثلاث طرائق للكشف عن دالة الفقرة التفاضلية وهي طريقة تحويل صعوبة الفقرة المحولة و طريقة معامل التمييز و طريقة كاي تربيع، حيث استخدم الباحث اختبار القدرة اللفظية للأطفال الأردنيين في الفئة العمرية (٩-١٢) الذي طوره أبو عناب عام ١٩٩٤، والذي يتكون من ٥٠ فقرة، بعد إضافة ١٠ فقرات اتفق المحكمون واختبار (T-test) على أنها متحيزة (ثمان فقرات متحيزة للمدينة، وفقرتان متحيزتان للقرية)، ليصبح الاختبار في صورته النهائية مكونا من ٦٠ فقرة، وقد تم تطبيق الاختبار على عينة مكونة من ٧٦٢ طالبا وطالبة (٣٨٣ مدينة، ٣٧٩ قرية) ، وقد تم اعتماد ثلاث معايير لمقارنة الطرائق الثلاث

- معيار القدرة على كشف الفقرات المعروفة بتحيزها المسبق .

All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

المنارة للاستشارات
www.manaraa.com

- معيار درجة استقرار مؤشرات التحيز .
- معيار معدل ثبات الفقرة الواحدة التي تصمد أمام طريقة الكشف عن التحيز بعد حذف الفقرات المتحيزة

ولقد أظهرت نتائج الدراسة أن طريقة مربع كاي استطاعت الكشف عن أكبر عدد من الفقرات المعلوم تحيزها المسبق حيث كشفت عن ٨ فقرات، واحتلت المركز الثاني طريقة تحويل صعوبة الفقرة وكشفت عن ٧ فقرات، أما طريقة مؤشر التمييز فاحتلت المركز الثالث وكشفت عن فقرة واحدة فقط ، أما قيما يتعلق بدرجة استقرار الفقرة في الكشف عن الفقرات المتحيزة فقد أظهرت النتائج أن طريقة تحويل صعوبة الفقرة هي أكثر الطرائق استقرارا، ثم طريقة مؤشر التمييز ثم طريقة كاي تربيع حيث بلغ معامل ارتباط فاي بين مرتي التطبيق للطرائق الثلاث على التوالي (٠,٤٠ ، ٠,١٣ ، ٠,٠٣٤) ، كما أظهرت الدراسة عدم وجود فروق دالة إحصائيا بين معدل ثبات الفقرة المتبقية في الطرائق الثلاث بعد حذف الفقرات المتحيزة وفقا لكل طريقة .

كما قام كيم وكوهين (Kim & Cohen (1995 بعمل مقارنة بين ثلاث طرائق تعتمد على إجراءات نظرية الاستجابة للفقرة في الكشف عن التحيز تبعا لمتغير السماح باستخدام الآلة الحاسبة في الفقرات وهي : مربع كاي للورد(Lord's) و اختبار نسبة الأرجحة (Liklelihood Ratio) وقياس المساحة بين منحنيات خصائص الفقرة وفق طريقة راجو (Raju) . اعتمدت هذه الدراسة على بيانات دراسة سابقة للباحثان عام ١٩٩٢م لاختبار رياضيات مكون من (٢٨) فقرة تقيس مهارات رياضية أساسية في التفاضل والتكامل ، تم اشتقاق صورتين من هذه الفقرات الصورة الأولى تتكون من جزأين كل منهما مكون من (١٤) فقرة أما الصورة الثانية فقد تكونت من نفس الجزأين مع عكس ترتيبها ، وقد صممت هذه الفقرات لتحل بدون استخدام الآلة الحاسبة لذلك كان الغرض من عمل تحليل DIF الكشف عن الفروق بين الفقرات التي يمكن أن تعزى لاستخدام الآلة الحاسبة عند تطبيق الاختبار، وقد نصت تعليمات الاختبار على جواز استخدام الآلة الحاسبة في الجزء الثاني من الاختبار فقط ، حيث اعتبرت المجموعة التي تجيب على الفقرة دون استخدام الآلة الحاسبة مجموعة مرجعية، أما المجموعة التي تجيب على الفقرة باستخدام الآلة الحاسبة مجموعة مستهدفة، وبالتالي أصبحت الفقرات في الجزء الأول من الصورة الأولى ضمن المجموعة التجريبية وفقرات الجزء الثاني ضمن المجموعة المستهدفة

All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

المنارة للاستشارات
www.manaraa.com

والعكس بالنسبة للجزأين في الصورة الثانية، وقد تم تطبيق الصورة الأولى على ٧٦٥ طالبا وتطبيق الصورة الثانية على (٧٢٥) طالبا.

وأظهرت نتائج الدراسة أن طريقة المساحة بين منحنيات خصائص الفقرة هي الأكثر فاعلية في الكشف عن التحيز مقارنة بالطرائق الأخرى تلتها طريقة مربع كاي للورد ، وأخيرا نسبة الأرجحية، حيث استطاعت طريقة راجو المعتمدة على المساحة بين المنحنيات في الكشف عن أربع فقرات لصالح المجموعة المستهدفة، كذلك كشفت طريقة مربع كاي للورد عن ثلاث فقرات لكل مجموعة، أما طريقة نسبة الأرجحية فلم تكشف إلا عن فقرتين فقط . وبشكل عام أظهرت النتائج موافقة جيدة بين الطرائق الثلاث في الكشف عن فقرتين فقط ، كما أن استخدام طريقة (Iteration) في الربط بين معالم الفقرات في المجموعتين المرجعية والمستهدفة ساهم في تقليل الفروق بين الطرائق المستخدمة في الكشف عن DIF.

وفي دراسة أجراها راجو وآخرون (Raju, et al. (1993 هدفت إلى مقارنة أربع طرائق للكشف عن التحيز ، تعتمد ثلاث منها على نظرية الاستجابة للفقرة وهي :

طريقة المساحة الإشارية (Singed Area Measure) وطريقة المساحة غير الإشارية(Unsigned Area Measure) (يتم فيها تجاهل الإشارات عند إيجاد مجموع الفروق بين المساحات وذلك باستخدام القيمة المطلقة أو التربيع) وطريقة مربع كاي للورد (Lord's chi-square).

وطريقة واحدة تستند إلى النظرية الكلاسيكية في القياس هي طريقة مانتل هينزل (MH). وذلك باستخدام ٤٥ فقرة تغطي الجانب اللفظي من اختبارات (Gates-MacGinitie Reading Tests, GMRT) المصممة لتستخدم مع الفئة العمرية من (١٠-١٢) سنة، وقد جرى تطبيقها على (٨٣٩) طالبا وطالبة بواقع (٤٤٠) طالبا و (٣٩٩) طالبا يتوزعون على أربع مجموعات فرعية بواقع (٢٤٥) طالبا من السود و (٤٣٦) طالبا من البيض، في حين يتوزع (١٥٨) طالبا على المجموعات العرقية الأخرى. وقد أظهرت نتائج الدراسة أن الطرائق الثلاث التي تستند إلى نظرية الاستجابة للفقرة متماثلة في تحديد الفقرات المتحيزة تبعا للجنس (الذكور، والإناث) من جهة، وتبعا للعرق (الطلاب البيض، والطلاب السود) من جهة أخرى حيث كشفت الطرائق الثلاث عن نفس الفقرات التي يوجد فيها DIF . كما أظهرت الدراسة وجود ارتباط عال بين الطرائق الثلاث التي تستند إلى نظرية الاستجابة للفقرة ، وطريقة مانتل هينزل (MH) فقط في حالة اعتماد الجنس كأساس للكشف عن التحيز ، بينما

All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

المنارة للاستشارات
www.manaraa.com

كان هنالك اختلاف في الفقرات التي تم الكشف عنها باستخدام الطرائق الثلاث وطريقة مانتل هينزل عند اعتماد العرق (البيض والسود) كأساس للكشف عن التحيز .

كما أجرى بدر (١٩٩٣) دراسة هدفت إلى الكشف عن تحيز فقرات اختبار ذكاء جمعي يقيس مكونات أساسية في القدرات العقلية لأفراد أردنيين في الفئة العمرية (٩-١٢) سنة نتيجة لاختلاف هؤلاء الأفراد في مستواهم الاجتماعي والاقتصادي ، وما ينجم عن هذا الاختلاف من تحيز محتمل في الأداء . ولتحقيق أغراض الدراسة تم تطبيق الاختبار الجمعي الذي قام الباحث بإعداده، والمكون من (١٠٠) فقرة ، والموزعة في سبعة اختبارات فرعية تقيس القدرة اللفظية ، والقدرة على التفكير المنطقي على (٧٠٠) طالبا وطالبة في مدارس حكومية وخاصة في مناطق مختلفة من عمان الكبرى ، حيث استخدمت في هذه الدراسة طريقتان من الطرائق المعروفة لكشف التحيز المحتمل في فقرات الاختبار ، تقوم إحداهما على مفاهيم النظرية الكلاسيكية في القياس ، وهي طريقة تحويل صعوبة الفقرة (TID) وأخرى تقوم على مفاهيم النظرية الحديثة في القياس ، وهي طريقة المساحة بين منحنيي خصائص الفقرة المرسوم وفق النموذج ثنائي المعلمة (ICC) بالإضافة لتحليل استجابات أفراد المجموعتين العليا والدنيا على بدائل إجابات الفقرات .

بينت النتائج وجود تحيز متفاوت في الدرجة في عدد من الفقرات لصالح أفراد المجموعة العليا ، تراوحت نسبته من (٣٨%) من مجموع الفقرات وفقا لطريقة طريقة (TID) إلى (٦١%) من مجموع الفقرات وفقا طريقة (ICC) ، كما أظهرت نتائج الدراسة عدم اتفاق الطريقتان في تحديد الفقرات المتحيزة، حيث تبين أن بعض الفقرات المتحيزة بموجب معيار إحدى الطريقتين لم تكن متحيزة بموجب معيار الطريقة الثانية. وقد بلغت نسبة الفقرات المتحيزة بموجب الطريقتين معا (٦٠%) من الفقرات المتحيزة بموجب طريقة (TID) ، (٣٤%) من الفقرات المتحيزة بموجب طريقة (ICC) . كما أن نسبة الفقرات المتحيزة في الفقرات المتحيزة في الاختبارات التي تقيس القدرة اللفظية كانت أعلى من نسبتها في الاختبارات التي تقيس القدرة على التفكير المنطقي بموجب الطريقة الأولى، ولكن هذه النسبة كانت أعلى في اختبارات القدرة على التفكير المنطقي من اختبارات القدرة اللفظية بموجب الطريقة الثانية .

وأجرى سكاجز ولسيتز (Skaggs & Lissitz, (1992 دراسة بعنوان اتساق طرائق الكشف عن تحيز الفقرة خلال التطبيقات المختلفة للاختبار، حيث تم استخدام صورتين لاختبار

All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

المنارة للاستشارات
www.manaraa.com

تحصيلي في الرياضيات يقيس مهارات في الأعداد الصحيحة، والكسور، والهندسة، وأساسيات الجبر ، وحل المشكلات . وقد تم تطبيق الاختبار على طلاب الصف الثامن ثلاث مرات في عام ١٩٨٦ على عينة مكونة من ١٣٠٠ طالب وطالبة وفي عام ١٩٨٧ تم تطبيقه مرتين على عينتين، الأولى مكونة من ٤٠٠ طالب وطالبة والثانية مكونة من ١٣٠٠ طالب وطالب تقريبا، وقد تم اعتماد ٩٦ فقرة مشتركة بين التطبيقات الثلاث جرى حذف ٤ منها لعدم مطابقتها للنموذج المستخدم في تقدير المعالم للنظرية الحديثة. ثم جرى بعدها استخدام سبع طرائق للكشف عن التحيز بالفقرة أربع منها تستند إلى إجراءات النظرية الحديثة في القياس ، وهي طريقة المساحة لرندر (Area) ، وطريقة مربع كاي للورد (Lord's Chi-Square) وطريقة مجموع المربعات غير الإشارية (Unsigned Sum Squares) ، وطريقة مجموع المربعات الإشارية (signed Sum Squares) وثلاث طرائق تستند إلى إجراءات النظرية الكلاسيكية في القياس وهي طريقة مانتل هينزل دلتا (Mantel-Haenszel Delta) ، وطريقة مانتل هينزل باستخدام مربع كاي (M-H Chi-Square) وطريقة مربع كاي (Full Chi-Square) . وأظهرت النتائج التي اعتمدت على الارتباط الرتبي كمؤشر على استقرار الطرائق في الكشف عن الفقرات باختلاف التطبيق، وكمعيار لفحص فعالية الطرائق أن الطرائق المستندة إلى نظرية الاستجابة للفقرة التي تعتمد على مجموع المربعات هي أكثر الطرائق استقرارا في الكشف عن التحيز تلتها طريقة مانتل هينزل وعلى الرغم من ذلك فقد كانت درجة الثبات والاتفاق لهذه الطرائق متوسطة في أفضل الأحوال، في حين كانت طريقة مربع كاي وطريقة المساحة وطريقة مربع كاي للورد أقلها استقرارا مقارنة مع الطرائق الأخرى.

وفي دراسة قامت بها أونيل O'Neal (1991) هدفت إلى المقارنة بين ست طرائق للكشف عن التحيز في الفقرة هي طريقة تحليل الإجابات على البدائل الخاطئة (المموهات) للفقرة، وطريقة مانتل هينزل (MH) وطريقة دلتا المرسومة لأنجوف (Angoff's Delta Plot)، وطريقة دلتا المرسومة المعدلة لأنجوف (Angoff's Modified Delta Plot) ، وطريقة مؤشر الارتباط الجزئي لستريكر (Stricker's Partial Correlation Index)، وطريقة المقارنة المباشرة لصعوبة الفقرة تبعاً لمتغيري الجنس والعرق، وقد استخدمت الباحثة البيانات المتوفرة من خلال عملية تطوير اختبارات (الرياضيات ، والقراءة ، واللغة) من اختبارات ألباما للكفاءة الأساسية (Alabama Basic Competency Tests, ABCT)،

All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

المنارة للاستشارات
www.manaraa.com

حيث جرى تصنيف البيانات المتحققة إلى أربع مجموعات تبعاً لمتغير العرق والجنس وتتمثل هذه المجموعات في الطالبات السود مقابل الطالبات البيض ، الطلاب السود من الذكور مقابل الطلاب البيض ، والطلاب السود مقابل الطالبات السود ، والطلاب البيض من الإناث مقابل الطلاب البيض . حيث أظهرت النتائج أن طريقة المقارنة المباشرة للصعوبة هي أكثر الطرائق استقرارا تلتها مباشرة طريقة تحليل الإجابات على البدائل الخاطئة (المموهات)، كما أظهرت كل من طريقة مانتل هينزل (MH) ، وطريقة مؤشر الارتباط الجزئي لستراليكرز استقرارا تراوح بين الضعيف والمتوسط . أما أكثر الطرائق ضعفاً في الاستقرار ، فكانت طريقة دلتا المرسومة لأنجوف .

All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

المنارة للاستشارات
www.manaraa.com

ثانيا دراسات المحاكاة

قام سو وشانغ (2005) Su & Wang بدراسة هدفت إلى دراسة بعض العوامل التي تؤثر ثلاث طرائق هي طريقة مانتل (Mantel) وطريقة مانتل هينزل العامة (Generalized Mantel-Hanszel (GMH) وطريقة التمييز اللوجستي للفقرات متعددة الاستجابة (Logistic Discriminant Function Analysis, LDFA) ، اشتملت الدراسة على ثمان متغيرات هي

طرائق الكشف عن DIF (٣ مستويات) ، إجراءات تنقية الاختبار (Test Purification Procedure) (٣ مستويات)، نموذج استجابة الفقرة (٣ مستويات) ، الفروق في القدرة بين المجموعات (٤ مستويات)، طول الاختبار (مستويان)، مقدار DIF (٣مستويات) ، نمط DIF (خمس مستويات)، نسبة الفقرات التي يوجد بها Dif (٦ مستويات)، وقد نتج عن هذه المتغيرات ما مجموعه ١١١٧٨ حالة، قام الباحث بتوليد البيانات ١٠٠ مرة لكل حالة من الحالات حيث كان حجم العينة لكل من المجموعة المرجعية والمجموعة المستهدفة ٥٠٠ فرد .

أظهرت النتائج أن مقدار تلوث DIF في درجات المقارنة (Matching Scores) والتي تم حسابها اعتمادا على المساحة الإشارية (Singed Area Measure) بين منحني خصائص الفقرة للمجموعة التجريبية والمجموعة المستهدفة كان أكثر أهمية من نسبة الفقرات التي توجد بها DIF في تحديد الخطأ من النوع الأول للطرائق الثلاث، كما أظهرت الطرائق الثلاث تحكما بالخطأ من النوع الأول عند اقتراب مؤشر المساحة الإشارية من صفر حتى عندما تصل نسبة الفقرات التي يوجد بها DIF إلى ٢٠% ، كما كانت قوة الاختبار (Power) أعلى لطريقتي مانتل و LDFA من طريقة GMH في كل أنماط DIF باستثناء نمط (Balanced DIF).

كما أجرى الباحثون (Kristjansson. , Aylesworth , Mcdowell & Zumbo (2005) دراسة هدفت إلى مقارنة أربع طرائق للكشف عن دالة الفقرة التفاضلية في الفقرات متعددة الاستجابة وهي طريقة مانتل هينزل (MH) وطريقة مانتل هينزل العامة (GMH) وطريقة التمييز اللوجستي (LDFA) وطريقة الانحدار اللوجستي التراكمي (UCLORD) ، تم في هذه الدراسة توليد بيانات لـ لاختبار مكون من ٢٦ فقرة لها أربع مستويات للاستجابة (٠ ، ١، ٢، ٣)وتمت دراسة DIF في الفقرة الأخيرة من الاختبار فقط وهي الفقرة رقم ٢٦.

All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

المنارة للاستشارات
www.manaraa.com

اشتمل تصميم الدراسة على معالجة عدد من المتغيرات لتقييم أثرها على الخطأ من النوع الأول وقوة الاختبار (TYPE I ERROR & POWER) للطرائق الأربع عند استخدامها للكشف عن دالة الفقرة التفاضلية (DIF) المنتظمة وغير المنتظمة ، وهي

- وجود DIF ونوعه (عدم وجود DIF ، و وجود DIF منتظم ،وعدم وجود DIF) .
- معلم التمييز للفقرة حيث تم اختيار ثلاث قيم لمعلم التمييز (a) هي (٠,٨ ، ١,٢ ، ١,٦).
- توزيع القدرة حيث تم اختيار مستويين للفروق في توزيع القدرة بين المجموعة المرجعية والمجموعة المستهدفة الحالة الأولى كان فيها توزيع المجموعتين متساوٍ N(0,1) والحالة الثانية كان فيها توزيع القدرة للمجموعة المستهدفة له متوسط حسابي = -٠,٥ وانحراف معياري =١ N(-0.5,1) ولتوزيع القدرة للمجموعة المرجعية متوسط حسابي = صفر وانحراف معياري = ١ N(0,1) .
- حجم العينة تم اختيار حجم العينة ليساوى ٤٠٠٠ لاثنين من الحالات الحالة الأول كان فيها حجم العينة للمجموعة المرجعية والمستهدفة متساو ٢٠٠٠ لكل مجموعة ،والحالة الثانية كانت فيها النسبة بين المجموعتين أربعة لواحد ٣٢٠٠ فرد للمجموعة المرجعية و٨٠٠ فرد للمجموعة المستهدفة .
- الالتواء في توزيع القدرة تضمنت هذه الدراسة مقارنة اثنين من مستويات الالتواء (وجود التواء متوسط سلبي في كلا المجموعتين المرجعية والمستهدفة ، عدم جود التواء في كلا المجموعتين) .

وقد قام الباحثون بتوليد البيانات ٤٠٠ مرة لمجموع الحالات الناتجة وهي ٩٦ حالة ، كما تم اعتماد مستوى الدلالة (α=0.05). وقد أظهرت نتائج الدراسة ما يلي :-

أولا النتائج المتعلقة بالخطأ من النوع الأول

كانت معدلات الخطأ من النوع الأول للطرائق الأربعة جيدة بشكل عام، فمعدلات الخطأ من النوع الأول كانت قريبة من ٠,٠٥ للطرائق الأربعة، حيث كان معدل الخطأ من النوع الأول لطريقة (UCLOUR) هو الأقل (٠,٠٤٩) بينما كانت أعلى قيمة لطريقة (MH) (٠,٠٥٤)

ثانيا النتائج المتعلقة بقوة الاختبار (POWER)

أظهرت النتائج أن الطرائق الأربعة تتمتع بقوة ممتازة في الكشف عن دالة الفقرة التفاضلية، كما أظهرت نتائج الدراسة أن قوة الطريقتين (GMH & UCLOLR) في الكشف عن دالة الفقرة التفاضلية المنتظمة ترتبط مع معلم التمييز وحجم العينة فقوة الطريقتين في الكشف عن

All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

المنارة للاستشارات
www.manaraa.com

(Uniform DIF) كانت أفضل عندما كان معلم التمييز للفقرات متوسطا ، كما أن قوة الاختبار للطريقتين كانت أقل قليلا عندما كانت نسبة العينة بين المجموعتين (٤:١) .

أما فيما يتعلق بالكشف عن دالة الفقرة التفاضلية غير المنتظمة (Nonuniform DIF) فقد كانت قوة طريقتي (UCLOLR & GMH) تامة تقريبا حيث بلغت على التو الي (١,٠ ، ٠,٩٩)، كما كان أداء طريقة (LDFA) ضعيفا في الكشف عن (Nonuniform DIF) مع معدل قوة = ٠,٥٥، أما طريقة (MH) فلم تكشف عن (Nonuniform DIF) نهائيا .

وفي دراسة بعنوان أداء طريقة SIBTEST عند ارتفاع نسبة الفقرات التي يوجد بها DIF أجراها الباحثون (2004) Gierl , Gotzmann, & Boughton هدفت إلى دراسة بعض المواقف التي يكون فيها نسبة الفقرات التي يوجد بها DIF عاليا ، كالحالات التي يتم فيها ترجمة فقرات الاختبار أو تكييفه ليتلاءم مع بيئة أخرى، وقد تمت دراسة أربعة متغيرات وفقا لأسلوب المحاكاة

١- نسبة DIF في اختبار مكون من ٤٠ فقرة (٢٠% ،٤٠% ،٦٠% من الفقرات بها DIF متوسط أو عالٍ)

٢- اتجاه DIF (Balanced &Unbalanced DIF) .

٣- حجم العينة (٥٠٠، ١٠٠٠، ١٥٠٠، ٢٠٠٠).

٤- الفرق بين توزيع القدرة بين المجموعات (توزيع متساوي وتوزيع غير متساوي) .

ووفقا لهذه المتغيرات تم تكرار كل حالة من الحالات ١٠٠ مرة ، حيث أظهرت نتائج الدراسة أن طريقة SIBTEST تمتاز بمعدلات دقيقة في الكشف عن DIF حتى عندما تصل نسبة الفقرات التي يوجد بها Balanced DIF ٦٠% وذلك عندما لا يقل حجم العينة عن ١٠٠٠ مفحوص لكل مجموعة .

كما أظهرت نتائج الدراسة كذلك أن طريقة SIBTEST تمتاز بمعدلات كشف مرتفعة في الحالة التي يكون فيها حجم العينة يساوي ١٠٠٠ مفحوص لكل مجموعة ونسبة الفقرات التي يوجد فيها Unbalanced DIF ٢٠%، ومع ذلك فقد كان معدل الكشف عن Unbalanced DIF ضعيفا عندما تكون نسبة الفقرات ٤٠% و٦٠% ضعيفة .

All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

المنارة للاستشارات
www.manaraa.com

كما قام الباحث Luppescu, (2002) بدراسة هدفت إلى مقارنة قدرة النماذج الخطية الهرمية في الكشف عن (DIF) مع الطرائق المعيارية المستخدمة في الكشف عن (DIF) مثل طريقة راش للفروق بين معالم الصعوبة . بيانات هذه الدراسة تمت توليدها باستخدام برنامج (Simtest) اعتمادا على النموذج اللوجستي ثلاثي المعالم، حيث تم توليد ١٨٠ مجموعة من البيانات (خمسة مجموعات لكل حالة من الحالات الست والثلاثون المختلفة (٣×٤×٣)) وفقا للمتغيرات التالية :

- حجم العينة ١٠٠، ٢٥٠٠ ، ٥٠٠ .
- مقدار دالة الفقرة التفاضلية (DIF) ٠,٢٥ ، ٠,٥٠ ، ٠,٧٥ ، ١,٥ .
- نسبة الأشخاص في المجموعة الضابطة ١٠% ، ٢٥% ، ٥٠% .

كل مجموعة من البيانات كانت تحتوي على ٥٠ فقرة ، الفقرة الخامسة والعاشرة والخامسة العشرة ...الخ في كل مجموعة كانت تحتوي على (DIF) .

لمقارنة كل من الطريقتين تم حساب (DIF) لكل فقرة بكلا الطريقتين، كما تم حساب الجذر التربيعي لمتوسط مربع الخطأ (Root Mean Squared Error) للمجموعة المرجعية والمستهدفة لمجموعات البيانات الست والثلاثون وفقا لهذه المعادلة :

$$rmse = \sqrt{\frac{\sum (DIFest - DIFact)}{nj}}$$

حيث *DIFest* هو DIF المقدرة بإحدى الطريقتين

DIFact هو DIF الحقيقي

n هو عدد الفقرات

j هو عدد مجموعات البيانات

في معظم الحالات كانت قيمة *rmse* لكلا الطريقتين متشابهة، لكن كانت هنالك نقاط اختلاف في الحالات التي كانت فيها العينة صغيرة ونسبة الأشخاص في المجموعة الضابطة قليل إذ وجد أن قيمة *rmse* كانت أكبر لطريقة النماذج الخطية الهرمية(HML)، أما في الحالات التي نكون فيها العينة كبيرة وحجم DIF صغير ونسبة الأشخاص في المجموعة الضابطة صغيرة فإن قيمة *rmse* لطريقة النماذج الخطية الهرمية (HML) كانت أصغر .

All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

المنارة للاستشارات
www.manaraa.com

وقد قام Ankenmann, Witt & Dunbar, (1999) بدراسة هدفت إلى تحديد مستويات الخطأ من النوع الأول لطريقة (Likelihood Ratio Goodness Of Fit , (LR)) عند استخدامها للكشف عن دالة الفقرة التفاضلية (Differential Item Function (DIF)) لفقرات تم توليدها وفقا لنموذج الاستجابة المتدرج (Samgima's 1969 1972 Graded Response Model)، وقد تمت معالجة أربعة متغيرات في هذه الدراسة هي حجم العينة، وتوزيع القدرة للمجموعتين المرجعية والمستهدفة، ومعالم الصعوبة والتمييز للفقرات، ونوع دالة الفقرة التفاضلية، كما استخدمت أيضا طريقة مانتل (mantel) أيضا للكشف عن دالة الفقرة التفاضلية وذلك لأغراض المقارنة ولقد أظهرت النتائج أن القوة الإحصائية للطـــريقتين تأثرت بحجم العينة كما كان متـــوقعا ، كما أن طريقة (LR) افتقرت إلى الاتساق في الكشف عن (DIF) عندما كان حجم العينة للمجموعتين المرجعية والمستهدفة صغيرا (٥٠٠/٥٠٠)، كما أن طريقة مانتل أظهرت قدرة على السيطرة على معدل الخطأ من النوع الأول وكانت أكثر قوة من طريقة (LR) عندما كان توزيع القدرة متماثلا في المجموعتين عند وجود دالة فقرة تفاضلية ثابتة (Constant DIF)، إلا أنها فقدت السيطرة على الخطأ من النوع الأول عندما اختلف متوسط توزيع القدرة بين المجموعتين، كما أن القوة لطريقة (LR) كانت أفضل من طريقة مانتل عند وجود دالة فقرة تفاضلية متوازنة (Blanced DIF) .

وفي دراسة قام بها تيان Tian, (1999) بعنوان الكشف عن دالة الفقرة التفاضلية في الفقرات متعددة الاستجابة تمت مقارنة فاعلية أربع طرائق للكشف عن دالة الفقرة التفاضلية في الفقرات متعددة الاستجابة وهي (Mantel, GMH, STND, SIBTEST, & LDFA (LDFA-Uniform & LDFA- Nonuniform)) وفقا لعدد من المتغيرات وهي الفروق في توزيع القدرة ، وطول الاختبار، وحجم العينة ونسبتها ، ومعلم التمييز ، ونوع دالة الفقرة التفاضلية ، وقد أظهرت النتائج أن كل الطرائق تتمتع بمستوى جيد للخطأ من النوع الأول عندما كانت المجموعتان لهما نفس مستوى القدرة أو عند وجود فروق بسيطة في القدرة مع قيم غير مرتفعة لمعلم التمييز أو عند وجود فروق كبيرة في القدرة مع قيم منخفضة لمعلم التمييز ، كما أن معدل الخطأ من النوع الأول كان غير مقبول لكل الطرائق بشكل عام عندما كانت الفروق بسيطة بين المجموعتين المرجعية والمستهدفة وقيم معلم التميز مرتفعة، أو عندما كانت الفروق ين المجموعتين كبيرة وقيم معلم التمييز متوسطة ومرتفعة.

All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

المنارة للاستشارات
www.manaraa.com

وقد كانت القوة الإحصائية لكل الطرائق (باستثناء طريقة LDFA-Nonuniform) عالية ، حيث كانت ترتيب أفضل الطرائق LDFA-Uniform ، ثم طريقة Mantel، ثم طريقة SIBTEST، ثم طريقة STND، وأخيرا طريقة GMH.

كما كانت القوة الإحصائية للطرائق (Mantel, STND, SIBTEST & LDFA-Uniform) منخفضة عند استخدامها للكشف عن دالة الفقرة التفاضلية المنتظمة المتوازنة ودالة الفقرة التفاضلية غير المنتظمة (Blanced Uniform DIF & Nonuniform DIF) كما أظهرت النتائج أن الطريقتين (GMH , LDFA-NONUNIFORM) يمكن استخدامهما للكشف عن دالة الفقرة التفاضلية المنتظمة المتوازنة مع كون زيادة قيم القوة لهما عند تساوي توزيع القدرة للمجموعتين ، وكذلك انخفاض القوة لطريقة (LDFA-NONUNIFORM) عند صغر حجم العينة وانخفاض قيمة معلم التمييز.

**** لم يوثق

كما قام Jiang & Stout (1998) بدراسة هدفت إلى تقويم أداء طريقة SIBTEST عند استخدام معادلة انحدار غير خطية بدلا من المعادلة الخطية لانحدار الدرجة الحقيقية على الدرجة المشاهدة في طريقة SIBTEST، حيث تم دراسة أداء الطريقة وفقا لعدد من المتغيرات (حجم العينة ، توزيع القدرة ، معلم التمييز والصعوبة)، وقد أظهرت نتائج الدراسة أن طريقة SIBTEST بعد التعديل حافظت على نفس خصائص طريقة SIBTEST غير المعدلة مع وجود تحسن في معدلات الخطأ من النوع الأول، فعند وجود فروق في القدرة كانت معدلات الخطأ من النوع الأول للطريقة المعدلة أقل من SIBTEST غير المعدلة، كما أظهرت نتائج الدراسة أن قيم القوة لطريقة SIBTEST المعدلة كانت قريبة جدا من قيم القوة للطريقة غير المعدلة حيث كانت هذه القيم ضعيفة عندما كان حجم العينة للمجموعتين المرجعية والمستهدفة صغيرا (٢٥٠، ٢٥٠) وممتازة عندما كان حجم العينة كبيرا (٣٠٠٠،٣٠٠٠)، كما أظهرت الدراسة أن قيم القوة لطريقة مانتل هينزل بشكل عام كانت أعلى قليلا (تم استخدامها لمقارنة أداء طريقة SIBTEST) من طريقة SIBTEST إلا أن معدل تضخم الخطأ من النوع الأول لها كان أعلى من طريقة SIBTEST .

أما روز وستاوت Rousses & Stout (1996) فقد أجريا بحث بعنوان تأثير حجم العينات الصغيرة ومعالم الفقرة على مستوى الخطأ من النوع الأول لطريقتي Mantel-Haenszel MH & SIBTEST) اشتمل على دراستين الدراسة الأولى قامت بدراسة

All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

المنارة للاستشارات
www.manaraa.com

الخطأ من النوع الأول في حالة كون عدد أفراد المجموعة التجريبية والضابطة صغيرا (يمكن أن يصل إلى ١٠٠ مفحوص في كل مجموعة) أما الدراسة الثانية فقد تم فيها اختيار معالم مختلفة للفقرة التي تتم دراستها، ومن أجل جعل الدراسة واقعية قدر الإمكان تم الاعتماد على النموذج اللوجستي ثلاثي المعالم لتوليد البيانات وفقا لمعالم الفقرات التي تم تقديرها لخمس وعشرين فقرة من اختبار (Auto Shop Test) الذي تم اختياره من (Armed Services Vocational Aptitude Battery , ASVAB) والتي وردت في دراسة ((Nislevy & Bock ,(1984) وقد تم اختيار هذا الاختبار كونه قصير نسبيا ويتوفر فيه مدى واسع من معاملات التمييز لفقراته ، وقد تم اختيار فقرة واحدة فقط لدراسة عدم وجود DIF ، حيث تم اختيارها لان معالمها كانت قريبة من المتوسط قدر الإمكان، وقد كانت معالم هذه الفقرة كالتالي (a=1.32 b=0.03 c=0.25) .

كما تم اختيار أربع أحجام للعينات ١٠٠، ٢٠٠ ، ٥٠٠ ، ١٠٠٠ متساوية في كل من المجموعة المرجعية والمجموعة المستهدفة، وقد كان التباين موحدا بين المجموعتين ويساوي واحد ، أما الفروق بين متوسطات توزيعي المجموعتين المرجعية والمستهدفة و فقد كان هنالك ثلاث مستويات (٠,٠ ، ٠,٥ ، ١,٠) .

لكل حالة من الحالات تم توليد البيانات ٤٠٠ مرة، مع اعتماد مستوى الدلالة (٠,٠٥) كمعيار للحكم على وجود أو عدم وجود DIF .

أظهرت نتائج الدراسة أن أداء كل من الطريقتين (MH & SIBTEST) كان جيدا عند مستوى الدلالة (∞ = ٠,٠٥) ، إلا أن معدل الخطأ من النوع الأول كان يزداد بشكل طفيف مع زيادة حجم العينة وزيادة الفرق في المتوسط بين المجموعتين، مع كون معدل الخطأ من النوع الأول أقل قليل في طريقة مانتل هينزل عند يكون الفرق في المتوسط أكبر من صفر .

أما الدراسة الثانية فقد تم فيها اختيار معالم مختلفة للفقرة التي تتم دراستها خلافا للدراسة الأولى كالتالي :

a = ٠,٠٤ ، ١,٠ ، ٢,٥

b = −١,٥ ، −٠,٥ ، ٠,٠ ، ٠,٥ ، ١,٥

c = ٠,٢٠ عندما يكون متوسط المجموعتين المرجعية والضابطة متساويين

c = ٠,٢٠ ، ٠,١٠ ، ٠,٠٥ عندما يكون الفرق بين متوسط المجموعتين = ١

وقد تم توليد البيانات مئة مرة لكل حالة من الحالات السابقة، مع اعتماد مستوى الدلالة ٠,٠٥ كمعيار ، وقد أشارت نتائج الدراسة إلى ما يلي :-

All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

المنارة للاستشارات
www.manaraa.com

-أظهرت طريقتي (MH & SIBTEST) تضخما في الخطأ من النوع الأول في الفقرات ذات التمييز العالي والصعوبة المنخفضة (A = ١,٠ ، B = -١,٥) (A= ٢,٥، B= -١,٥)، مع كون التضخم في طريقة MH أكبر من طريقة SIBTEST .

- أظهرت طريقة MH تضخما في الخطأ من النوع الأول عندما تكون الفقرة ذات تمييز منخفض وصعوبة عالية (A= ٠,٤ B = ٠ إلى ١,٥)

-لم يتغير نمط الخطأ الأول لطريقة SIBTEST مع تناقص التخمين من ٠,٢ إلى ٠,٠٥ ، أما طريقة MH فقد انخفض الخطأ من النوع الأول عند معلم التمييز a = ٠,٤ ، بينما ارتفعت عند التمييز المرتفع والصعوبة المنخفضة (a= ٢,٥ ، b > ٠)

-معدلات الرفض كانت تزداد مع زيادة حجم العينة في كلا الطريقتين .

كما قام كل من Chang, Mazzeo And Rouses (1996) ببحث هدف إلى تقييم طريقة (SIBTEST) المعدلة لتتلاءم مع الفقرات متعددة الاستجابة اشتمل أيضا على دراستين الدراسة الأولى قارنت طريقة مانتل هينزل (Mantel Hanszel) وطريقة (Standard Mean Difference , SMD) مع طريقة (SIBTEST) في الحالات التي كان فيها أداء طريقتي Mantel & SMD جيدا ، أما الدراسة الثانية فقد قامت بمقارنة الطريقتين مع طريقة SIBTEST ضمن قيم واقعية ومتنوعة لمعلم التمييز .

الدراسة الأولى

اعتمدت هذه الدراسة على نفس تصميم دراسة zwick et al. التي تضمن تصميمها ٥٤ حالة وفقا للمتغيرات التالية :-

- اثنين من توزيعات القدرة للمجموعة المستهدفة (N(0,1) & N(-1,1))

- ٢٧ فقرة منها (ثلاث فقرات متعددة الاستجابة في أربع أنواع من DIF (Constant , Balanced, Low-Shift, High Shift) في اثنين من مستويات DIF (0.10 , 0.25) إضافة إلى ثلاث فقرات لا يوجد بها DIF . وقد تم توليد البيانات ٦٠٠ مرة لكل حالة من الحالات .

وقد أظهرت نتائج الدراسة أن معدل الخطأ من النوع الأول لطريقتي (MANTEL & SMD) كان جيدا ، حيث كان على التوالي ٠,٠٤٩ و٠,٠٤٦ أما بالنسبة للطريقة (SIBTEST) فقد كان هنالك تضخم قليل لمعدل الخطأ من النوع الأول حيث كان معدل الخطأ ٠,٠٦٣ .

All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

المنارة للاستشارات
www.manaraa.com

كما أظهرت معدلات الرفض بالنسبة للحالات التي كان يوجد بها DIF أن طريقة (SIBTEST) كانت في مستوى طريقتي (MANTEL & SMD) في الكشف عن وجود DIF المنتظم ، إلا أن أداء الطرائق الثلاث لم يكن جيدا في الكشف عن DIF غير المنتظم (في حالة BALANCED DIF) وقد كان هذا متوقعا حيث أن الطرائق الثلاث مصممة للكشف عن DIF المنتظم .

الدراسة الثـــانية :

الدراسة الثانية كانت امتدادا للدراسة الأولى بحيث تشتمل على قيم واقعية متنوعة للتمييز (٢,٠ ، ١,٥، ١,٣٦ ، ١,١٢ ، ١,٠ ، ٠,٨٦٩ ، ٠,٥٨٨ ، ٠,٣٣ ، ٠,٢٥ ، ٠,٢٣ ، ٠,١٥) العديد من هذه القيم تم اختيارها بناء على معايرة فقرات اختبار (NAEP) للقراءة عام ١٩٩٢حتى تكون القيم واقعية قدر الإمكان .

تضمنت هذه الدراسة على ٢٢ حالة ، ١١ من هذه الحالات كانت الفقرات لا توجد فيها DIF و١١ حالة أخرى كانت الفقرات يوجد بها DIF عند كل قيمة من قيم معلم التمييز التي سبق ذكرها. لكل حالة تم توليد البيانات ١٠٠٠ مرة لعينة مكونة من ٥٠٠ فرد لكل من المجموعة المرجعية والمجموعة المستهدفة .

وقد أظهرت نتائج الدراسة تضخما في معدلات الخطأ من النوع الأول لكل من طريقتي (Mantel & SMD) حيث وصل في بعض الحالات إلى ٤٠% عندما كان معلم التمييز للفقرة يختلف عن متوسط قيم معلم التمييز لفقرات الاختبار(٠,٧٥٨). كما أن معدل الخطأ من النوع الأول كان قريبا من المستوى ٠,٠٥ عندما كان معامل التمييز للفقرات قريبا من متوسط قيم معلم التمييز لفقرات الاختبار، أما طريقة (Sibtest) فقد كانت نتائجها أفضل على الرغـم من أن معـدل الخطأ من النوع الأول كـان يزداد بشكـل معتدل للقيم المتطـرفة لمعلم التمييز، حيث تراوح معدل الخطأ من النوع الأول للفقرات بين ٠,٠٦١ و٠,٠٩٣ .

كما قام Cohen. & Kim (1993) بدراسة هدفت إلى الكشف عن فاعلية اثنين من الاختبارات الإحصائية في الكشف عن دالة الفقرة التفاضلية هما اختبار ز للمساحة الإشارية (Z Test For Exact Signed Area)، واختبار ز للمساحة غير الإشارية (Z Test For Exact Unsigned Area) وفقا لعدد من المتغيرات وهي طول الاختبار ، وحجم العينة ونسبة الفقرات التي يوجد بها دالة الفقرة التفاضلية وشروط تقدير المعالم باستخدام النموذج ثنائي المعالم وقد نتج عن هذا التصميم ١٤ خلية، تم تكرار توليد البيانات لها خمس مرات، وقد تمت

All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

المنارة للاستشارات
www.manaraa.com

مقارنة أخطاء الطريقتين في الكشف عن دالة الفقرة التفاضلية مع أخطاء الكشف عن دالة الفقرة التفاضلية عند استخدام طريقة لورد كاي سكوير (LORD X^2) ، حيث أظهرت نتائج الدراسة أن طريقة لورد كاي سكوير (LORD X^2) كانت أكثر فاعلية في الكشف عن دالة الفقرة التفاضلية من اختباري ز وأن واختبار ز للمساحة الإشارية (Z Test For Exact signed Area) كان أقل الطرائق فاعلية.

All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

المنارة للاستشارات
www.manaraa.com

تلخيص الدراسات السابقة

جدول ٢. ملخص الدراسات التي اعتمدت على بيانات حقيقية

المؤلف	أبعاد التحيز	المعيار	الطرائق المستخدمة
ظاظا (٢٠٠٦)	– الجنس – المستوى الاقتصادية والاجتماعي	– الكشف عن الفقرات المعلوم تحيزها مسبقا	– الإجراءات التحكيمية – تحويل صعوبة الفقرة – مربع كاي – حساب المنحنى
Lingyun Gao & Changjiang Wongs (2005)	– الجنس	–اتساق النتائج	– الانحدار اللوجستي المتعدد الاستجابات. – نموذج الاستجابة المتدرج – مانتل هاتزل – طريقة SIBTEST للفقرات متعددة الاستجابة.
محمود البستنجي (٢٠٠٤)	–الجنس	–الاتفاق بين الطرائق	–مانتل هينزل – الانحدار اللوجستي – مؤشر الفرق في المساحة لمنحنى خصائص الفقرة للنموذج الثلاثي. – مؤشر الفرق في المساحة لمنحنى خصائص الفقرة للنموذج الثنائي.
(Bertrand & boiteau ; 2003)	– الاختلاف في الثقافة (الأمريكية واليابانية)	–الاستقرار الداخلي –الاتفاق بين الطرائق	– مانتل هينزل. – الانحدار اللوجستي. – طريقة المساحة (UPD index) – (Raju's NCDIF index)
الجرادات (٢٠٠٣)	–الجنس –الحقل الأكاديمي (علمي ، أدبي)	– نسبة الكشف عن الفقرات المتحيزة. – الاتفاق بين الطرائق.	– مانتل هينزل (MH) . – تحويل صعوبة الفقرة(TID).
القـرعان (١٩٩٧)	–الجنس – الموقع الجغرافي (مدينة، قرية)	– القدرة على الكشف عن الفقرات المعروف تحيزها المسبق. –استقرار مؤشرات التحيز. – معدل ثبات الفقرة التي تصمد أمام طريقة الكشف عن التحيز بعد حذف الفقرات المتحيزة.	– تحويل صعوبة الفقرة . – مؤشر التمييز . – كاي تربيع .
kim & Cohen (1995)	–استخدام أو عدم استخدام الآلة الحاسبة في	–نسبة الكشف عن -- – الفقرات المتحيزة. – الاتفاق بين الطرائق	–طريقة المساحة بين المنحنيات – مربع كاي للورد(Lord's X^2) – نسبة الأرجحة (Liklelihood

All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

المنارة للاستشارات
www.manaraa.com

المؤلف	أبعاد التحيز	المعيار	الطرائق المستخدمة
	المسائل الرياضية	.	Ratio) –المساحة بين منحنيات خصائص الفقرة وفق طريقة راجو (Raju)
(Raju, et al 1993)	–الجنس –العرق (طلاب بيض ، طلاب سود	– الاتفاق بين الطرائق	– مانتل هينزل – طريقة المساحة الإشارية (Singed Area Measure) –وطريقة المساحة غير الإشارية (Unsigned Area Measure) – مربع كاي للورد (Lord's chi-square)
بدر (١٩٩٣)	–المستوى الاجتماعي والاقتصادي	– الاتفاق بين الطرائق	– تحويل صعوبة الفقرة (TID) –المساحة بين منحنيي خصائص الفقرة المرسوم وفق النموذج ثنائي المعلمة
Skaggs & Lissitz, (1992)	– الجنس	– الاتساق في الكشف عن تحيز الفقرة خلال التطبيقات المختلفة للاختبارات	ـ طريقة المساحة لرندر ـ مربع كاي للورد (Lord's Chi-Square) مجموع المربعات غير الإشارية (Unsigned Sum Squares) – طريقة مجموع المربعات الإشارية (signed Sum Squares) –مانتل هينزل دلتا (Mantel-Haenszel Delta) –طريقة مانتل هينزل باستخدام مربع كاي (M-H Chi-Square) –طريقة مربع كاي (Full Chi-Square)
(O'Neal,1991)	– الجنس – العرق	– استقرار مؤشرات التحيز	–تحليل الإجابات على البدائل الخاطئة (المموهات) للفقرة –طريقة مانتل هينزل (MH) – طريقة دلتا المرسومة لأنجوف (Angoff's Delta Plot) – طريقة دلتا المرسومة المعدلة لأنجوف (Angoff's Modified Delta Plot) – طريقة مؤشر الارتباط الجزئي لستريكر (Stricker's Partial Correlation Index). – طريقة المقارنة المباشرة لصعوبة الفقرة .

All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

المنارة للاستشارات
www.manaraa.com

جدول ٣. ملخص الدراسات التي استخدمت أسلوب المحاكاة.

المؤلف	النموذج المستخدم في توليد البيانات	المتغيرات المستقلة في الدراسة	الطرائق المستخدمة
(su & chung ,2005)	- partial credit model (PCM; Masters, 1982) - graded response model (GRM; Samejima, 1969)	-إجراءات تنقية الاختبار (test purification) - نموذج الاستجابة للفقرة. - الفرق في متوسط القدرة - طول الاختبار -مقدار دالة الفقرة التفاضلية - نوع دالة الفقرة التفاضلية.	- مانتل هينزل (MH) . -مانتل هينزل العامة (GMH) . - التمييز اللوجستي (LDFA) .
(Kristjansson. , Aylesworth , Mcdowell & Zumbo , 2005)	-نموذج الاعتماد الجزئي العام (Generalised Partial Credit Card)	-وجود دالة الفقرة التفاضلية ونوعها . -معلم التمييز . -توزيع القدرة . - حجم العينة. التفلطح في توزيع القدرة .	- مانتل هينزل (MH) . -مانتل هينزل العامة (GMH) . - التمييز اللوجستي (LDFA) . - الانحدار اللوجستي(UCLORD).
(Gierl , Gotzmann, & Boughton ; 2004.)	partial credit model (PCM	-نسبة دالة الفقرة التفاضلية DIF في اختبار مكون من ٤٠ فقرة (٢٠% ،٤٠% ،٦٠% من الفقرات بها DIF متوسط أو عالٍ) -اتجاه DIF (Balanced &Unbalanced DIF) . -حجم العينة -الفرق بين توزيع القدرة بين المجموعات (توزيع متساوي وتوزيع غير متساوي) .	- SIBTEST
(Stuart Luppescu, 2002)	-النموذج اللوجستي ثلاثي المعالم	-عدد الأشخاص ١٠٠، ٢٥٠٠ ، ٥٠٠ . - مقدار دالة الفقرة التفاضلية (DIF) (٠,٢٥ ، ٠,٥٠ ، ٠,٧٥ ، ١,٥) . -نسبة الأشخاص في المجموعة الضابطة ١٠% ، ٢٥% ،	- النماذج الهرمية . - طريقة راش.

All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

المنارة للاستشارات
www.manaraa.com

المؤلف	النموذج المستخدم في توليد البيانات	المتغيرات المستقلة في الدراسة	الطرائق المستخدمة
		٥٠% .	
Fang Tian, 1999	-النموذج اللوجستي ثلاثي المعالم. - generalized partial credit model(GPCM)	-الفروق في توزيع القدرة -طول الاختبار -حجم العينة ونسبتها - معلم التمييز - نوع دالة الفقرة التفاضلية	- Mantel - GMH - STND - SIBTEST - LDFA (LDFAUniform & LDFA-Nonuniform)
Witt & Dumber, 1999	- نموذج الاستجابة المتدرج Samgima's 1969 1972 Graded Response Model	- حجم العينة، - توزيع القدرة للمجموعتين المرجعية والمستهدفة، -معالم الصعوبة والتمييز للفقرات، -نوع دالة الفقرة التفاضلية	- Likelihood Ratio Goodness Of Fit - مانتل
Jiang & Stout (1998)	-النموذج اللوجستي ثلاثي المعالم.	- معلم الصعوبة - معلم التمييز -حجم العينة. - توزيع القدرة	- SIBTEST. - مانتل هينزل (MH)
(Louis A. Rousses, William F. Stout , 1996)	- النموذج اللوجستي ثلاثي المعالم	- حجم العينة - الفرق بين متوسطي القدرة للمجموعتين. - معلم التمييز معلم الصعوبة - معلم التخمين	- مانتل هينزل (MH) - Sibtest.
(Chang H., Mazzeo J., And Rouses, L. , 1996)	-النموذج اللوجستي ثلاثي المعالم. -نموذج الاعتماد الجزئي (Partial credit model)	-توزيع القدرة . -مقدرا دالة الفقرة التفاضلية. -نوع دالة الفقرة التفاضلية .	- مانتل هينزل (MH) - SIBTEST. - الفروق بين المتوسطات المعيارية Standard Mean Difference (SMD)
Cohen & Kim , 1993	-النموذج ثنائي المعالم	-طول الاختبار - -وحجم العينة -نسبة الفقرات التي يوجد بها دالة الفقرة التفاضلية -شروط تقدير المعالم باستخدام النموذج ثنائي المعالم	- اختبار ز للمساحة الإشارية (Z Test For Exact Signed Area). - واختبار ز للمساحة غير الإشارية (Z Test For Exact - Unsigned)

All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

المنارة للاستشارات
www.manaraa.com

الفـصـــــل الثـــــالث

الطريـــــــــــــقة والإجـــــــــــراءات

All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

المنارة للاستشارات
www.manaraa.com

الطريقة والإجراءات

هدفت هذه الدراسة إلى تقييم فاعلية ومقارنة أربعا من طرائق الكشف عن دالة الفقرة التفاضلية، (Differential Item Function, DIF) ، اثنتين من هذه الطرائق تستندان إلى النظرية الكلاسيكية وهما طريقة مانتل هينزل وطريقة الانحدار اللوجستي، واثنتين تستندان إلى النظرية الحديثة في القياس (نظرية استجابة الفقرة) وهما وطريقة اختبار تحيز الفقرة المتزامن (SIBTEST) وطريقة نسبة الأرجحية (IRTLRDIF)، وقد تم تقييم فاعلية هذه الطرائق باستخدام أسلوب المحاكاة في توليد البيانات، حيث تم توليد بيانات للمجموعتين المرجعية والمستهدفة بشكل مستقل وفقا للنموذج اللوجستي ثلاثي المعالم لاختبار مؤلف من أربعين فقرة، وقد كانت جميع الفقرات متماثلة المعالم في المجموعتين باستثناء الفقرة الأخيرة رقم ٤٠ التي تم من خلالها دراسة دالة الفقرة التفاضلية DIF حيث أن معالم هذه الفقرة كانت متماثلة فقط في حالة عدم وجود DIF أما في حالة وجود DIF منتظمة فإن معلم الصعوبة للفقرة يكون مختلفا بين المجموعتين المرجعية والمستهدفة وفي حالة وجود DIF غير منتظمة فإن معلم التمييز للفقرة يكون مختلفا بين المجموعتين المرجعية والمستهدفة. ولقد فحصت الدراسة فاعلية هذه الطرائق تحت ظروف مختلفة من توزيع قدرات أفراد العينة، وحجم العينة ونسبتها بين المجموعتين المرجعية والمستهدفة وتمييز الفقرة ونوع دالة الفقرة التفاضلية ومقدارها .

منهجية الدراسة :

تصميم الدراسة ومتغيراتها :

تضمنت الدراسة أربعة متغيرات مستقلة هي

١-توزيع القدرة (مستويين) .

٢- حجم العينة ونسبتها (٦ مستويات).

٣- معلم التمييز (ثلاث مستويات) .

٤- نوع دالة الفقرة التفاضلية ومقدارها (٤ مستويات) .

نتج عن هذه المتغيرات ١٤٤ خلية (٢×٦×٣×٤) . وفيما يلي توضيح لهذه المتغيرات .

All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

المنارة للاستشارات
www.manaraa.com

١- توزيع القدرة

لدراسة تأثير توزيع القدرة على فاعلية الأساليب الإحصائية المستخدمة في الكشف عن دالة الفقرة التفاضلية DIF تم اختيار مستويين لهذه المتغير المستوى الأول كان فيه توزيع القدرة لكل من المجموعة المرجعية والمستهدفة متماثلا حيث كان التوزيع توزيعا طبيعيا بمتوسط حسابي يساوي صفر وانحراف معياري يساوي واحد N(0,1) .

المستوى الثاني كان فيه توزيع القدرة غير متساو للمجموعتين المرجعية والمستهدفة ، حيث كانت المجموعة المرجعية لها توزيع طبيعي بمتوسط حسابي يساوي صفر وانحراف معياري يساوي ١ N(0,1)، بينما المجموعة المستهدفة لها توزيع طبيعي أيضا لكن بمتوسط حسابي أقل -٠,٥ وانحراف معياري يساوي ١ N(-0.5,1) .

٢- حجم العينة ونسبتها بين المجموعة التجريبية والضابطة

تم اختيار مستويين للنسبة بين المجموعة المرجعية والمجموعة المستهدفة، المستوى الأول كان فيها حجم المجموعتين متساويا، والمستوى الثاني كان فيها حجم المجموعة المرجعية ضعفي حجم المجموعة المستهدفة وذلك كالتالي

أ- النسبة بين حجم المجموعة المرجعية والمستهدفة (١:١)

وتتضمن ثلاث حالات (٢٠٠-٢٠٠، ٦٠٠-٦٠٠ ، ١٢٠٠-١٢٠٠)

ب- النسبة بين حجم المجموعة المرجعية والمستهدفة (١:٢)

وتتضمن ثلاث حالات (٤٠٠-٢٠٠ ، ٨٠٠-٤٠٠ ، ١٦٠٠-٨٠٠)

٣- تمييز الفقرة التي تمت دراستها

عند دراسة الخطأ من النوع الأول (عدم وجود دالة الفقرة التفاضلية في الفقرة) ودالة الفقرة التفاضلية المنتظمة، كان معلم التمييز للفقرة التي تتم دراستها في المجموعتين المرجعية والمستهدفة متماثلا حيث تم اختيار ثلاث قيم لمعلم التمييز ٠,٥ (منخفض) ، و١,٠(متوسط) و١,٥ (مرتفع)، أما عند دراسة دالة الفقرة التفاضلية غير المنتظمة قد تم اختيار القيمة ٠,٥ لمعلم التمييز في المجموعة المستهدفة وثلاث قيم لمعلم التمييز في المجموعة المرجـــعية وهي (١,٠ ،١,٥ ، ٢,٠) ، أي كان هنالك ثلاث مستويات للفرق بين المجموعتين (٠,٥ ، ١,٠ ، ١,٥)

والجدول التالي يبين هذه المستويات

All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

المنارة للاستشارات
www.manaraa.com

الجدول ٤. معاملات التمييز للفقرة رقم ٤٠ في المجموعتين المرجعية والمستهدفة

الرقم	حالة عدم وجود (DIF) أو وجود (DIF) منتظم	حالة وجود (DIF) غير منتظم	
	معلم التمييز للمجموعتين المرجعية والمستهدفة	معلم التمييز للمجموعة المستهدفة (a_F)	معلم التمييز للمجموعة المرجعية (a_R)
-١	$a_R = a_F = 0.5$	$a_F = 0.5$	$a_R = 1.0$
-٢	$a_R = a_F = 1.0$	$a_F = 0.5$	$a_R = 1.5$
-٣	$a_R = a_F = 1.5$	$a_F = 0.5$	$a_R = 2.0$

٤- نوع دالة الفقرة التفاضلية ومقدارها

وكلن لها أربع مستويات

أ- عدم وجود (DIF) .

في هذا المستوى يكون معلمي الصعوبة والتمييز للفقرة متساويين في المجموعة التجريبية والضابطة ($a_F = a_R$, $b_F = b_R$) .

ب- وجود (DIF) منتظم مقداره ٠,٢٥ .

وفي هذا المستوى يكون معلم التميز متساوي في المجموعتين ($a_F = a_R$) ، لكن معلم الصعوبة للمجموعتين غير متساوي، حيث أن معلم الصعوبة للمجموعة المستهدفة (b_F) يزيد عن معلم الصعوبة للمجموعة المرجعية (b_R) بمقدار ٠,٢٥ ($b_F = b_R + 0.25$).

ج- وجود (DIF) منتظم مقداره ٠,٥ .

وفي هذا المستوى يكون معلم التميز متساوي في المجموعتين ($a_F = a_R$) ، لكن معلم الصعوبة للمجموعتين غير متساوي ، حيث أن معلم الصعوبة للمجموعة المستهدفة (b_F) يزيد عن معلم الصعوبة للمجموعة المرجعية (b_R) بمقدار ٠,٥ ($b_F = b_R + 0.5$).

د- وجود (DIF) غير منتظم .

وفي هذا المستوى يكون معلم التميز في المجموعتين المرجعية والمستهدفة غير متساو ($a_R = a_F + 0.5$ or 1.0 or 1.5) ، لكن معلم الصعوبة متساو ($b_R = b_F$).

All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

المنارة للاستشارات
www.manaraa.com

معاييـــــــر تقييـــــم الطرائق

١- الخطأ من النوع الأول (Type1 Error)

ويمثل النسبة التي يتم فيها رفض الفرضية الصفرية التي تنص على عدم وجود دالة الفقرة التفاضلية بين المجموعتيــن المرجعيــة والمستهدفــة خطــأ عند مستــوى الدلالــة ($\alpha \leq$ ٠,٠٥).

وإجــرائيا اعتبــرت النسبـــــة المئوية لعدد مرات الكشـــــف عن دالة الفقرة التفاضلية خطأ ($\alpha \leq$ ٠,٠٥) من المئـــــة مرة وهي عدد مرات توليد البيانات لكل حالة الخطأ من النوع الأول .

٢- القـــــوة (POWER)

وهي تمثل النسبة التي يتم فيها الكشف عن وجود دالة الفقرة التفاضلية بين المجموعتين المرجعية والمستهدفة في الفقرة بشكل صحيح عند مستوى الدلالة($\alpha =$ ٠,٠٥) .

وإجرائيا اعتبرت النسبة المئوية لعدد مرات الكشف عن دالة الفقرة التفاضلية بشكل صحيح من المئة مرة التي تم فيها توليد البيانات مؤشرا علىHGR,M .

توليـــــــد بيـــــــانات الدراســـــة

أ- النموذج الذي تم اعتماده لتوليد البيانات .

تم توليد بيانات الدراسة باستخدام النموذج اللوجستي ثلاثي المعالم الذي يتم من خلاله حساب احتمالية الاستجابة الصحيحة على الفقرة عند مستوى القدرة (θ) وفقا للمعادلة التالية :-

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

حيث أن

a تشير إلى معلم التمييز .

b تشير إلى معلم الصعوبة .

c تشير إلى معلم التخمين .

All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

المنارة للاستشارات
www.manaraa.com

ب- خصائص الفقرات

تم تحديد خصائص الفقرات لجميع فقرات الاختبار باستثناء الفقرة الأخيرة رقم (٤٠) بحيث تكون متماثلة في كلا المجموعتين المجموعة المرجعية والمجموعة ، وقد تم اختيار قيم معالم الفقرات كالتالي :

-معلم الصعوبة (b)

تم اختيار قيم معالم الصعوبة للفقرات بحيث يكون لها توزيع سوي متوسطه صفر وانحرافه المعياري واحد N(0,.1) .

- معلم التمييز (a)

تم اختـــــار قيم معالم التمييـــــز للفقــرات بحيث يكون لها توزيع منتظم تتراوح قيمه بين (٠,٦٤ -١,٥) U(0.64 – 1.5) .

- معلم التخمين (c)

تم اعتماد قيمة ثابتة لمعلم التخمين تساوي (٠,٢) لجميع فقرات الاختبار .

ويوضح الجدول ٥ معالم الفقرات التي تم اختيارها للفقرات التسع وثلاثون المشتركة بين المجموعة المرجعية والمجموعة الضابطة .

All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

المنارة للاستشارات
www.manaraa.com

الجدول ٥. معالم الفقـــــــرات الاختبار من رقم ١ إلى ٣٩

رقم الفقـــرة	معلم الصعـوبة	معلم التمييـــز	معلم التخميـــن
١	٢,١٢-	١,٢٢	٠,٢
٢	٠,٥١-	١,١١	٠,٢
٣	٠,٤٥-	٠,٨٨	٠,٢
٤	٠,٠٨	٠,٩٧	٠,٢
٥	٠,١	١,٠٧	٠,٢
٦	٠,١٧-	٠,٧٤	٠,٢
٧	٠,٧٩	٠,٨٤	٠,٢
٨	٠,٨٣-	١,٠٣	٠,٢
٩	١,٨٤	٠,٩٨	٠,٢
١٠	٠,٩٥	١,٣٩	٠,٢
١١	٢,٣٢	١,٠٧	٠,٢
١٢	٠,٥٧-	٠,٩٢	٠,٢
١٣	٠,٤٦	١,٣٥	٠,٢
١٤	١,٥٥	١,٠٧	٠,٢
١٥	١,٠٩	١,٢٦	٠,٢
١٦	٠,٦٨	٠,٦٥	٠,٢
١٧	١,٠٥-	١,٣٢	٠,٢
١٨	٠,٣٢-	١,١٦	٠,٢
١٩	١,٨٤-	٠,٩٦	٠,٢
٢٠	٠,٠٦-	٠,٨١	٠,٢
٢١	٠,٢٨	١,٣	٠,٢
٢٢	١,٤٣-	١,١٦	٠,٢
٢٣	١,٢٧-	١,٤٣	٠,٢
٢٤	٠,٠٤	٠,٧١	٠,٢
٢٥	١,١٤-	٠,٨٥	٠,٢
٢٦	١,١٧	٠,٨٤	٠,٢
٢٧	٠,٢٢-	١,٠٥	٠,٢
٢٨	٠,٦٤	١,٤٢	٠,٢
٢٩	١,٣٦	٠,٦٤	٠,٢
٣٠	٠,٣٧-	٠,٧٩	٠,٢
٣١	٠,٩٦-	٠,٧٨	٠,٢
٣٢	٠,٥٤	٠,٩٦	٠,٢
٣٣	٠,٨	١,٢	٠,٢
٣٤	٠,٢٢-	١,٢٧	٠,٢
٣٥	١,١١	١,١٨	٠,٢
٣٦	٠,١٤	١	٠,٢
٣٧	١,٢١	١,٤٢	٠,٢
٣٨	٠,٣٦	١,٢٥	٠,٢
٣٩	١,٣٥-	١,٣٣	٠,٢

All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

المنارة للاستشارات
www.manaraa.com

إجراءات توليد البيانات وتحليل دالة الفقرة التفاضلية (DIF)

- تم توليد قيم θ لكل من المجموعة التجريبية والضابطة وفقا لكل من توزيع القدرة وحجم العينة .
- تم توليد الاستجابات على الفقرات لكل قيم θ وفقا لمعالم الفقرات (الصعوبة والتمييز والتخمين) وكذلك تم توليد استجابات للفقرات وفقا لنموذج استجابة الفقرة الثلاثي وبما يتلاءم مع معالم الفقرات.
- تم استخدام طرائق التحليل الأربعة (مانتل هينزل ،الانحـــدار اللوجستي، اختبار التحيز المتزامن (SIBTEST) ، ونسبة الأرجحية (IRTLRDIF)).
- تمت إعادة الخطوة الأولى والثانية والثالثة ١٠٠ مرة لكل حالة من الحالات.

البرامج التي تم استخدامها في هذه الدراسة

١- لتوليد البيانات تم استخدام برنامج PARDISM وهو برنامج يمكن من خلاله توليد بيانات لفقرات ثنائية الاستجابة من خلال النماذج اللوجستية الثلاث (أحادية وثنائية وثلاثية المعالم).

٢- لتحليل النتائج المتعلقة بطريقتي مانتل هينزل والانحدار اللوجستي تم الاعتماد على برنامج spss .

٣- لتحليل النتائج المتعلقة بطريقة اختبار التحيز المتزامن (SIBTEST) تم استخدام برنامج SIBTEST 1.1.

٤- لتحليل النتائج المتعلقة بطريقة نسبة الأرجحية (Likelihood-Ratio Tests) تم اســتخدام برنامج IRTLRDIF v.2.0b .

All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

المنارة للاستشارات
www.manaraa.com

الفصـــــل الرابع

النتـــــــــــــــــائج

All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

المنارة للاستشارات
www.manaraa.com

هدفت هذه الدراسة إلى الإجابة عن الأسئلة التالية :-

١-ما الطرائق الأكثر فاعلية (مانتل هينزل Mantel Haenszel، والانحدار اللوجستي Logistic Regression، اختبار التحيز المتزامن SIBTEST، ونسبة الأرجحية IRTLRDIF) في الكشف عن دالة الفقرة التفاضلية من حيث الخطأ من النوع الأول؟

٢-ما الطرائق الأكثر فاعلية (مانتل هينزل Mantel Haenszel، والانحدار اللوجستي Logistic Regression، اختبار التحيز المتزامن SIBTEST، ونسبة الأرجحية IRTLRDIF) في الكشف عن دالة الفقرة التفاضلية من حيث قوة الاختبار الإحصائي؟

٣-هل يوجد تأثير للمتغيرات (حجم العينة ، نسبة العينة بين المجموعة المرجعية والمجموعة المستهدفة، توزيع القدرة، معلم التمييز) على الخطأ من النوع الأول والقوة للطرائق الأربع (مانتل هينزل Mantel Haenszel، والانحدار اللوجستي Logistic Regression، اختبار التحيز المتزامن SIBTEST، ونسبة الأرجحية IRTLRDIF)؟

للإجابة عن السؤال الأول المتعلق بنسبة الخطأ من النوع الأول تم توليد بيانات لا يوجد فيها دالة الفقرة التفاضلية (DIF) بواقع مئة مرة لكل حالة من الحالات. وللكشف عن دالة الفقرة التفاضلية استخدمت الطرائق الأربعة في كل مرة من مرات توليد البيانات، وتم حساب عدد المرات من المرات المئة التي كانت فيها الفقرة ذات دالة تفاضلية، واعتبرت النسبة المئوية لهذه المرات نسبة الخطأ من النوع الأول، وللإجابة عن السؤال الثاني المتعلق بقوة الاختبار تم توليد بيانات يوجد فيها دالة تفاضلية لفقرة واحدة فقط، وفي كل مرة حسبت الدالة التفاضلية باستخدام الطرائق الأربع، وأعيد توليد البيانات مئة مرة وطبق عليها نفس إجراءات الكشف عن الفقرة ذات الدلالة التفاضلية وعدت المرات التي كانت فيها الدالة تفاضلية بشكل صحيح، واعتبرت النسبة المئوية لهذه المرات دلالة على قوة الاختبار الإحصائي.

وبالنسبة للسؤال الثالث المتعلق بتأثير للمتغيرات (حجم العينة، نسبة العينة بين المجموعة المرجعية والمجموعة المستهدفة، توزيع القدرة، معلم التمييز) على الخطأ من النوع الأول والقوة فقد استخدم الانحدار اللوجستي للإجابة عنه ، وفيما يلي عرض لهذه النتائج :-

All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

المنارة للاستشارات
www.manaraa.com

الســـــؤال الأول

ما الطرائق الأكثر فاعلية (مانتل هينزل Mantel Haenszel، والانحـــدار اللوجستي Logistic Regression، اختبار التحيز المتزامن SIBTEST، ونسبة الأرجحية IRTLRDIF) في الكشف عن دالة الفقرة التفاضلية من حيث الخطأ من النوع الأول؟

الجدول ٦. نسبة الخطأ من النوع الأول للطرائق الأربع وفقا لتوزيع القدرة وحجم العينة ونسبتها بين المجموعة المرجعية والمجموعة المستهدفة ومعلم التمييز

الطريقة					
توزيع القدرة	حجم العينة ونسبتها	معلم التمييز	مانتل هينزل	الانحدار اللوجستي	نسبة الأرجحية	اختبار تحيز الفقرة المتزامن
N(0,1) N(0,1)	(٣٠٠ ، ٣٠٠)	٠,٥	٠,٠٢	٠,٠٠	٠,١٤	٠,٠٤
		١	٠,٠٥	٠,٠٠	٠,٢١	٠,٠٦
		١,٥	٠,٠٤	٠,٠٢	٠,١٤	٠,٠٥
	(٦٠٠ ، ٦٠٠)	٠,٥	٠,٠١	٠,٠٠	٠,١٦	٠,٠٢
		١	٠,٠٤	٠,٠٢	٠,١٧	٠,٠٢
		١,٥	٠,٠٢	٠,٠٢	٠,١٨	٠,٠٢
	(١٢٠٠ ، ١٢٠٠)	٠,٥	٠,٠٣	٠,٠٢	٠,١٣	٠,٠٦
		١	٠,٠٧	٠,٠٧	٠,٢١	٠,٠٧
		١,٥	٠,٠٦	٠,٠٨	٠,٢١	٠,٠٦
	(٤٠٠ ، ٢٠٠)	٠,٥	٠,٠١	٠,٠٠	٠,١٢	٠,٠٠
		١	٠,٠٢	٠,٠١	٠,١٦	٠,٠٢
		١,٥	٠,٠٣	٠,٠٣	٠,١٧	٠,١١
	(٨٠٠ ، ٤٠٠)	٠,٥	٠,٠٥	٠,٠٤	٠,٢٣	٠,٠٢
		١	٠,٠٢	٠,٠٠	٠,١٦	٠,٠٢
		١,٥	٠,٠٠	٠,٠٤	٠,١٤	٠,٠٥
	(١٦٠٠ ، ٨٠٠)	٠,٥	٠,٠٣	٠,٠٠	٠,٢٢	٠,٠٥
		١	٠,٠٥	٠,٠١	٠,١٤	٠,٠٦
		١,٥	٠,٠٧	٠,٠٣	٠,٢١	٠,٠٧
N(0,1) N(0.5,1)	(٣٠٠ ، ٣٠٠)	٠,٥	٠,٠٦	٠,٠٠	٠,١٧	٠,١١
		١	٠,٠٧	٠,٠٠	٠,٢٣	٠,٠٩
		١,٥	٠,٠٢	٠,٠٠	٠,١٤	٠,٠٠
	(٦٠٠ ، ٦٠٠)	٠,٥	٠,٠٦	٠,٠٠	٠,٢	٠,٠٨
		١	٠,٠٩	٠,٠٧	٠,٢٥	٠,٠٦
		١,٥	٠,٠٤	٠,٠٨	٠,٢٥	٠,٠٨
	(١٢٠٠ ، ١٢٠٠)	٠,٥	٠,٠٤	٠,٠٣	٠,١٢	٠,٠٢
		١	٠,٠٨	٠,٠٨	٠,٢٦	٠,٠٨
		١,٥	٠,٠٣	٠,٠٢	٠,٢٩	٠,٠٧

All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

المنارة للاستشارات
www.manaraa.com

الطريقة			مانتل هينزل	الانحدار اللوجستي	نسبة الأرجحية	اختبار تحيز الفقرة المتزامن
توزيع القدرة	حجم العينة ونسبتها	معلم التمييز				
N(0,1) N(0.5,1)	(٤٠٠ ، ٢٠٠)	٠,٥	٠,٠٦	٠,٠٢	٠,١٢	٠,٠٤
		١	٠,٠٠	٠,٠٠	٠,١٩	٠,٠٤
		١,٥	٠,٠٤	٠,٠١	٠,١٩	٠,١
	(٨٠٠ ، ٤٠٠)	٠,٥	٠,٠٢	٠,٠٠	٠,٢	٠,٠٢
		١	٠,٠٣	٠,٠٧	٠,١٤	٠,٠٢
		١,٥	٠,٠٨	٠,٠٧	٠,٢٢	٠,١٣
	(١٦٠٠ ، ٨٠٠)	٠,٥	٠,٠٥	٠,٠٦	٠,١٢	٠,٠٣
		١	٠,٠٦	٠,٠٠	٠,٢	٠,٠٦
		١,٥	٠,٠٦	٠,٠٦	٠,٢١	٠,١٥

أ-طريقة مانتل هينزل

يتضح من الجدول السابق أنه في حالة تساوي توزيع القدرة للمجموعتين المرجعية والمستهدفة أن معدل الخطأ من النوع الأول كان ≤ ٠,٠٥ في جميع خلايا التصميم باستثناء خليتين كان حجم العينة لهما كبيرا ومتساويا (١:١) ومعلم التمييز لإحداها متوسط وللأخرى مرتفعا ، وخلية واحدة كان فيها حجم العينة كبيرا وغير متساوي بين المجموعتين (٢ : ١) ومعلم التمييز مرتفعا.

كما يتضح أن معدلات الخطأ من النوع الأول عندما كان التوزيعان مختلفان، حيث مستوى القدرة للمجموعة المستهدفة أقل من المجموعة المرجعية كانت < ٠,٠٥ في تسع خلايا، خمس منها كانت فيها نسبة العينة بين المجموعتين المرجعية والمستهدفة متساويا (ثلاث ذات تمييز متوسط واثنتان ذات تمييز منخفض)، وأربع منها كانت فيها نسبة العينة غير متساوية بين المجموعتين المرجعية والمستهدفة (اثنتان ذات تمييز مرتفع وواحدة لكل من للتمييز المتوسط والمنخفض)، ولم يكن هنالك نمط عام يدل على وجود تأثير لحجم العينة .

All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

المنارة للاستشارات
www.manaraa.com

ب- طريقة الانحدار اللوجستي

يتضح من الجدول السابق أنه في حالة تساوي توزيع القدرة للمجموعتين المرجعية والمستهدفة أن معدل الخطأ من النوع الأول كان ≤ ٠,٠٥ في جميع الحالات باستثناء خليتين كان حجم العينة لهما كبيرا ومعلم التمييز لإحداها متوسطا وللأخرى مرتفعا .

وفي حالة عدم تساوي القدرة كان معدل الخطأ من النوع الأول > ٠,٠٥ في سبع خلايا ثلاث منها كان نسبة حجم العينة بين المجموعتين المرجعية والمستهدفة متساويا (اثنتان لهما حجم عينة متوسط إحداهما ذات معلم تمييز متوسط والأخرى مرتفع وواحدة ذات حجم عينة كبير ومعلم تمييز متوسط) ، وأربع خلايا كانت نسبة حجم العينة بين المجموعتين غير متساوية (اثنتان لهما حجم عينة متوسط إحداهما ذات معلم تمييز متوسط والأخرى مرتفع ، واثنتان ذات حجم عينة كبير إحداهما ذات معلم تمييز منخفض والأخرى مرتفع) .

ج-طريقة اختبار تحيز الفقرة المتزامن

يتضح من الجدول السابق أنه في حالة تساوي القدرة للمجموعتين المرجعية والمستهدفة كان عدد الخلايا التي فيها معدل الخطأ من النوع الأول > ٠,٠٥ سبع خلايا، أربع منها كان فيها نسبة العينة بين المجموعتين المرجعية والمستهدفة متساوية وثلاث منها كانت النسبة غير متساوية، وقد كان هنالك خلية واحدة منها معلم التمييز لها منخفضا مقابل ثلاث خلايا لكل من معلم التمييز المتوسط والمرتفع، وقد تركزت هذه الخلايا في العينات ذات الحجم الكبير والمتوسط.

أما عندما كان مستوى قدرة المجموعة المستهدفة أقل من المجموعة المرجعية ، فقد كانت معدلات الخطأ من النوع الأول > من ٠,٠٥ في عشر خلايا سبع منها كانت النسبة فيها بين المجموعة المرجعية والمستهدفة (١:١) وثلاث كانت فيها النسبة (٢ : ١)، ولم يكن هنالك نمط عام يدل على تأثير لحجم العينة أو معلم التمييز.

د- طريقة نسبة الأرجحية

يتضح من الجدول السابق وجود تضخم واضح للخطأ من النوع الأول حيث أن معدلات الخطأ من النوع الأول كانت > ٠,٥ في جميع الخلايا .

All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

المنارة للاستشارات
www.manaraa.com

السؤال الثاني

٢-ما الطرائق الأكثر فاعلية (مانتل هينزل Mantel Haenszel، والانحدار اللوجستي Logistic Regression، واختبار التحيز المتزامن SIBTEST، ونسبة الأرجحية IRTLRDIF) في الكشف عن دالة الفقرة التفاضلية من حيث قوة الاختبار الإحصائي؟

عند استعراض النتائج الخاصة بالقوة سوف تتضمن الجداول النتائج المتعلقة بطريقة نسبة الأرجحية لكن لن يتم التعليق عليها، لأن الخطأ من النوع الأول لهذه الطريقة كان مرتفعا جدا عن القيمة النظرية (٠,٠٥) في كل الحالات، مما يجعل القيم التي نحصل عليها للقوة لهذه الطريقة غير ذات معنى، وفيما يلي عرض لهذه النتائج :

٢-قوة الاختبار الإحصائي عند وجود دالة فقرة تفاضلية منتظمة مقدارها ٠,٢٥.

جدول ٧.القوة للطرائق الأربع وفقا لتوزيع القدرة وحجم العينة ونسبتها بين المجموعة المرجعية والمجموعة المستهدفة ومعلم التمييز عند وجود دالة فقرة تفاضلية منتظمة مقدارها ٠,٢٥

	الطريقة					
توزيع القدرة	حجم العينة ونسبتها	معلم التمييز	مانتل هينزل	الانحدار اللوجستي	نسبة الأرجحية	اختبار تحيز الفقرة المتزامن
N(0,1) (N(0,1	(٣٠٠ ، ٣٠٠)	٠,٥	٠,١١	٠,١٦	٠,٣١	٠,١٣
		١	٠,٢٨	٠,٣٣	٠,٤	٠,٣١
		١,٥	٠,٢٩	٠,٣٥	٠,٥١	٠,٢٤
	(٦٠٠ ، ٦٠٠)	٠,٥	٠,٢٦	٠,١٤	٠,٤١	٠,٢٣
		١	٠,٤٨	٠,٤٦	٠,٨١	٠,٥٥
		١,٥	٠,٥٥	٠,٥٦	٠,٧٥	٠,٥٣
	(١٢٠٠ ، ١٢٠٠)	٠,٥	٠,٣٢	٠,٢٦	٠,٤٢	٠,٣٤
		١	٠,٧٥	٠,٧٣	٠,٩١	٠,٧٢
		١,٥	٠,٨١	٠,٨٧	٠,٩٥	٠,٨١
	(٢٠٠ ، ٤٠٠)	٠,٥	٠,١	٠,١٥	٠,٢٩	٠,١١
		١	٠,٢١	٠,٢١	٠,٣٧	٠,١٩
		١,٥	٠,٢٩	٠,٣١	٠,٥٦	٠,٢٤
	(٤٠٠ ، ٨٠٠)	٠,٥	٠,١٨	٠,١٤	٠,٣٩	٠,٢١
		١	٠,٤٨	٠,٤٨	٠,٦٩	٠,٤٧
		١,٥	٠,٥٤	٠,٦٣	٠,٨١	٠,٥٦
	(٨٠٠ ، ١٦٠٠)	٠,٥	٠,٤	٠,٣١	٠,٥٦	٠,٣٦
		١	٠,٦٤	٠,٧١	٠,٨٦	٠,٦٣
		١,٥	٠,٧٢	٠,٧٧	٠,٩	٠,٧٣

All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

المنارة للاستشارات
www.manaraa.com

	الطريقة		مانتل هينزل	الانحدار اللوجستي	نسبة الأرجحية	اختبار تحيز الفقرة المتزامن
توزيع القدرة	حجم العينة ونسبتها	معلم التمييز				
N(0,1) (N(0.5,1	(٣٠٠ ، ٣٠٠)	٠,٥	٠,١٢	٠,١٥	٠,١٧	٠,١٣
		١	٠,٢٦	٠,٢٥	٠,٤٣	٠,٢٨
		١,٥	٠,٤	٠,٣٣	٠,٥٦	٠,٣٥
	(٦٠٠ ، ٦٠٠)	٠,٥	٠,١١	٠,١٢	٠,٢٧	٠,١٥
		١	٠,٥٣	٠,٤٤	٠,٦٣	٠,٤
		١,٥	٠,٦٥	٠,٧٤	٠,٧٢	٠,٥٩
	(١٢٠٠ ، ١٢٠٠)	٠,٥	٠,٢٤	٠,٢١	٠,٥١	٠,٣
		١	٠,٧١	٠,٧١	٠,٨١	٠,٦٢
		١,٥	٠,٧٧	٠,٨٦	٠,٨٨	٠,٧
	(٢٠٠ ، ٤٠٠)	٠,٥	٠,١٤	٠,١٤	٠,٢٥	٠,١٦
		١	٠,١٩	٠,٢٣	٠,٤٧	٠,٢٦
		١,٥	٠,٣٧	٠,٤٥	٠,٥٤	٠,٣
	(٤٠٠ ، ٨٠٠)	٠,٥	٠,١٤	٠,١٥	٠,٣	٠,٢١
		١	٠,٤٥	٠,٤٤	٠,٦٨	٠,٣٦
		١,٥	٠,٥٦	٠,٨٣	٠,٨	٠,٣١
	(٨٠٠ ، ١٦٠٠)	٠,٥	٠,٣٤	٠,٢٩	٠,٥٥	٠,٣٥
		١	٠,٦٦	٠,٦١	٠,٧٨	٠,٤٣
		١,٥	٠,٦٥	٠,٩١	٠,٨٣	٠,٤٩
المتوسط العام			٠,٤١	٠,٤٣	٠,٥٩	٠,٣٩

يتضح من الجدول السابق ما يلي :

أ- طريقة مانتل هينزل

يتضح من النتائج أنه في حالة تساوي توزيع القدرة بين المجموعتين المرجعية والمستهدفة أن قوة الاختبار الإحصائي كانت ≤ ٠,٨٠ في خلية واحدة فقط وهي الخلية التي كان حجم العينة فيها كبيرا ومتساويا بين المجموعتين (١:١) ومعلم التمييز مرتفعا، كما أن قيمة القوة كانت بين ٠,٧٠-٠,٧٩ في خليتين إحداهما كان فيهما حجم العينة كبيرا ومتساويا بين المجموعتين (١:١) ومعلم التمييز متوسطا، وخلية كان فيها حجم العينة كبيرا وغير متساو بين المجموعتين (٢ : ١) ومعلم التمييز مرتفعا، وكذلك تراوحت قيمة القوة بين ٠,٦٠-٠,٦٩ في خلية واحدة كان فيها حجم العينة كبيرا وغير متساويا ومعلم التمييز متوسطا .

وعند عدم تساوي توزيع القدرة بين المجموعتين المرجعية والمستهدفة لم يكن هنالك أي خلية فيها قوة الاختبار الإحصائي ≤ ٠,٨٠ ، وقد تراوحت قيمة القوة بين ٠,٧٠-٠,٧٩ في خليتين

All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

المنارة للاستشارات
www.manaraa.com

كان فيهما حجم العينة كبيرا ومتساويا ومعلم التمييز لإحداهما مرتفعا وللأخرى متوسطا، كما أن قيمة القوة كانت بين ٠,٦٠-٠,٦٩ في ثلاث خلايا اثنتين فيهما حجم العينة كبيرا وغير متساويا (٢ : ١) ومعلم التمييز لإحداهما مرتفعا وللأخرى متوسطا، أما الخلية الثالثة فقد كان فيها حجم العينة متوسطا ومتساويا بين المجموعة المرجعية والمستهدفة (١:١) ومعلم التمييز مرتفعا .

ب- الانحدار اللوجستي

يتضح من الجدول السابق أنه في حالة تساوي توزيع القدرة بين المجموعتين المرجعية والمستهدفة أن قوة الاختبار الإحصائي كانت ≤ ٠,٨٠ في خلية واحدة فقط وهي الخلية التي كان فيها حجم العينة كبيرا ومتساويا بين المجموعتين (١:١) ومعلم التمييز مرتفعا، كما أن قيمة القوة كانت بين ٠,٧٠-٠,٧٩ في ثلاث خلايا اثنتين كان فيهما حجم العينة كبيرا وغير متساو بين المجموعتين المرجعية والمستهدفة (٢: ١) ومعلم التمييز لإحداها مرتفعا وللأخرى متوسطا والخلية الثالثة كان فيهما حجم العينة كبيرا ومتساويا بين المجموعتين المرجعية والمستهدفة (١:١) ومعلم التمييز متوسطا، كما كانت قيمة القوة بين ٠,٦٠-٠,٦٩ في خلية واحدة فيها حجم العينة متوسطا وغير متساو (٢: ١) ومعلم التمييز مرتفعا .

وعند عدم تساوي توزيع القدرة بين المجموعتين المرجعية والمستهدفة كانت قوة الاختبار الإحصائي ≤ ٠,٨٠ في ثلاث خلايا اثنتين كان فيهما نسبة حجم العينة بين المجموعتين المرجعية والمستهدفة غير متساوية (٢: ١) ومعلم التمييز مرتفعا وحجم العينة للأولى كبيرا والثانية متوسطا، أما الخلية الثالثة فقد كان فيها حجم العينة كبيرا ومتساويا بين المجموعتين (١:١) ومعلم التمييز مرتفعا، وقد تراوحت قيمة القوة بين ٠,٧٠-٠,٧٩ في خليتين نسبة حجم العينة فيهما بين المجموعتين المرجعية والمستهدفة متساوية لكن الأولى كان فيها حجم العينة متوسطا ومعلم التمييز مرتفعا، والثانية كان فيها حجم العينة كبيرا ومعلم التمييز متوسطا، كما أن قيمة القوة تراوحت بين ٠,٦٠-٠,٦٩ في خلية واحدة فقط كان فيها حجم العينة كبيرا وغير متساو بين المجموعتين (٢: ١) ومعلم التمييز متوسطا

ج- طريقة اختبار التحيز المتزامن SIBTEST

يتضح من الجدول السابق أنه في حالة تساوي توزيع القدرة بين المجموعتين المرجعية والمستهدفة أن قوة الاختبار الإحصائي كانت ≤ ٠,٨٠ في خلية واحدة فقط وهي الخلية التي كان فيها حجم العينة كبيرا ومتساويا بين المجموعتين (١:١) ومعلم التمييز مرتفعا، كما أن قيمة

All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

المنارة للاستشارات
www.manaraa.com

القوة وجدت بين ٠,٧٠-٠,٧٩ في خليتين إحداهما كان فيهما حجم العينة كبيرا ومتساويا بين المجموعتين (١:١) ومعلم التمييز متوسطا، وخلية كان فيها حجم العينة كبيرا وغير متساو بين المجموعتين (٢: ١)، ومعلم التمييز مرتفعا. وكذلك كانت قيمة القوة بين ٠,٦٠-٠,٦٩ في خلية واحدة فيها حجم العينة كبيرا وغير متساو ومعلم التمييز متوسطا.

وعند عدم تساوي توزيع القدرة بين المجموعتين المرجعية والمستهدفة لم يكن هنالك أي خلية فيها قوة الاختبار الإحصائي ≤ ٠,٨٠، وقد كانت قيمة القوة بين ٠,٧٠-٠,٧٩ في خلية واحدة فقط كان فيها حجم العينة كبيرا ومتساويا بين المجموعتين المرجعية والمستهدفة(١:١) ومعلم التمييز مرتفعا ، كما كانت قيمة القوة بين ٠,٦٠-٠,٦٩ في خلية واحدة فقط كان فيها حجم العينة كبيرا ومتساويا بين المجموعتين المرجعية والمستهدفة (٢: ١) ومعلم التمييز متوسطا.

All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

المنارة للاستشارات
www.manaraa.com

٢-قوة الاختبار الإحصائي عند وجود دالة فقرة تفاضلية منتظمة مقدارها ٠,٥.

الجدول ٨ . القوة للطرائق الأربع وفقا لتوزيع القدرة وحجم العينة ونسبتها بين المجموعة المرجعية والمجموعة المستهدفة ومعلم التمييز عند وجود دالة فقرة تفاضلية منتظمة مقدارها (٠,٥)

	الطريقة					
توزيع القدرة	حجم العينة ونسبتها	معلم التمييز	مانتل هينزل	الانحدار اللوجستي	نسبة الأرجحية	اختبار تحيز الفقرة المتزامن
N(0,1) (N(0,1	(٣٠٠ ، ٣٠٠)	٠,٥	٠,٣	٠,٢٧	٠,٥٢	٠,٣٥
		١	٠,٦٧	٠,٧	٠,٧٨	٠,٥٧
		١,٥	٠,٧٧	٠,٨٣	٠,٩٥	٠,٧٦
	(٦٠٠ ، ٦٠٠)	٠,٥	٠,٧٦	٠,٧١	٠,٩٦	٠,٧٧
		١	٠,٨٥	٠,٨٢	٠,٩٣	٠,٨٧
		١,٥	٠,٨٩	٠,٩٤	١	٠,٩٣
	(١٢٠٠ ، ١٢٠٠)	٠,٥	٠,٩٢	٠,٨٩	٠,٩٤	٠,٩٤
		١	٠,٩٦	٠,٩٨	١	٠,٩٦
		١,٥	٠,٩٢	٠,٩٧	١	٠,٩١
	(٢٠٠ ، ٤٠٠)	٠,٥	٠,٣٢	٠,٢٩	٠,٥٣	٠,٣٧
		١	٠,٦٢	٠,٥	٠,٧٥	٠,٦١
		١,٥	٠,٧٩	٠,٧٦	٠,٨٤	٠,٧٩
	(٤٠٠ ، ٨٠٠)	٠,٥	٠,٧٥	٠,٦٤	٠,٧٨	٠,٦٨
		١	٠,٨٢	٠,٨	٠,٩٣	٠,٨٧
		١,٥	٠,٨٩	٠,٩	٠,٩٦	٠,٨٨
	(٨٠٠ ، ١٦٠٠)	٠,٥	٠,٨٦	٠,٨٤	٠,٨٩	٠,٨٨
		١	٠,٩٣	٠,٩	١	٠,٩
		١,٥	٠,٨٧	٠,٩٣	١	٠,٩
N(0,1) (N(0.5,1	(٣٠٠ ، ٣٠٠)	٠,٥	٠,٣٢	٠,٢٧	٠,٣٨	٠,٣٦
		١	٠,٥٨	٠,٥٦	٠,٧٣	٠,٤٩
		١,٥	٠,٧	٠,٨٦	٠,٩٥	٠,٧
	(٦٠٠ ، ٦٠٠)	٠,٥	٠,٣٩	٠,٣٣	٠,٥٣	٠,٤٥
		١	٠,٩١	٠,٩٥	١	٠,٨
		١,٥	٠,٨٩	٠,٩٥	١	٠,٨٥
	(١٢٠٠ ، ١٢٠٠)	٠,٥	٠,٨٨	٠,٤٣	٠,٩٦	٠,٨١
		١	٠,٨٦	٠,٩٦	١	٠,٨
		١,٥	٠,٨٢	٠,٩٤	٠,٩٥	٠,٨١
	(٢٠٠ ، ٤٠٠)	٠,٥	٠,٢٥	٠,١٥	٠,٥	٠,٢٤
		١	٠,٥٤	٠,٦١	٠,٧٣	٠,٤٢
		١,٥	٠,٦٢	٠,٧	٠,٨١	٠,٥٥
	(٤٠٠ ، ٨٠٠)	٠,٥	٠,٦٣	٠,٦٥	٠,٨١	٠,٦
		١	٠,٨١	٠,٩٧	٠,٩٨	٠,٧٧
		١,٥	٠,٨٢	٠,٩٧	٠,٩٤	٠,٧٧
	(٨٠٠ ، ١٦٠٠)	٠,٥	٠,٨٨	٠,٩	٠,٩٢	٠,٨٤
		١	٠,٨٧	٠,٩٨	٠,٩٨	٠,٨٦
		١,٥	٠,٨٧	٠,٩٦	٠,٩٨	٠,٨٣
المتوسط العام			٠,٧٤	٠,٧٤	٠,٨٦	٠,٧٢

All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

المنارة للاستشارات
www.manaraa.com

يتضح من الجدول السابق ما يلي :

أ- طريقة مانتل هينزل

يتضح من النتائج أنه في حالة تساوي توزيع القدرات للمجموعتين المرجعية والمستهدفة أن قوة الاختبار الإحصائي كانت ≤ من ٠,٨٠ في كل الخلايا التي كان فيها حجم العينة كبيرا وفي كل الخلايا التي كان فيها حجم الخلية متوسطا باستثناء الخلايا التي كان فيها معلم التمييز منخفضا، ونلاحظ من الجدول وجود أربع حالات تراوحت فيها قيم القوة بين ٠,٧٠-٠,٧٩ اثنتان كان فيها حجم العينة صغيرا ومعلم التمييز مرتفعا واثنتان كان فيها حجم العينة متوسطا ومعلم التمييز منخفضا، كما وجدت خليتان تراوحت فيها قيمة القوة بين ٠,٦٠-٠,٦٩ كان فيهما حجم العينة صغيرا ومعلم التمييز متوسطا .

كما يتضح أن قوة الاختبار عندما كان التوزيعان مختلفان حيث مستوى القدرة للمجموعة المستهدفة أقل من المجموعة المرجعية كان ≤ ٠,٨٠ أيضا في كل الخلايا التي كان فيها حجم العينة كبيرا وفي كل الخلايا التي كان فيها حجم الخلية متوسطا باستثناء الخلايا التي كان فيها معلم التمييز منخفضا، كما وجدت خلية واحدة كانت قيمة القوة فيها بين ٠,٧٠-٠,٧٩ وكان فيها حجم العينة كبيرا ومتساويا في المجموعتين (١:١) ومعلم التمييز مرتفعا، كما وجدت خليتان كانت قيمة القوة لهما بين ٠,٦٠-٠,٦٩ والنسبة فيها بين المجموعتين المرجعية والمستهدفة فيهما غير متساوية (٢: ١) إحداهما ذات حجم عينة كبير وتمييز مرتفع والأخرى ذات حجم عينة متوسط وتمييز منخفض .

ب- طريقة الانحدار اللوجستي

يتضح من الجدول السابق أنه في حالة تساوي توزيع القدرات للمجموعتين المرجعية والمستهدفة أن قوة الاختبار الإحصائي كانت ≤ من ٠,٨٠ في كل الخلايا التي كان فيها حجم العينة كبيرا وفي كل الخلايا التي كان فيها حجم الخلية متوسطا باستثناء الخلايا التي كان فيها معلم التمييز منخفضا، وفي خلية واحدة كان فيها حجم العينة صغيرا ومتساويا بين المجموعتين المرجعية والمستهدفة(١:١) ومعلم التمييز مرتفعا، كما نلاحظ من الجدول وجود ثلاث حالات كانت فيها قيم القوة بين ٠,٧٠-٠,٧٩ اثنتين كان فيهما حجم العينة متساويا بين المجموعتين المرجعية والمستهدفة (١:١) إحداهما ذات حجم صغير ومعلم تمييز متوسط والأخرى ذات حجم متوسط ومعلم تمييز منخفض، والثالثة كان فيها حجم العينة صغيرا وغيـــر متساوي بين المجموعتين

All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

المنارة للاستشارات
www.manaraa.com

(٢: ١) ومعلم التمييز مرتفعا، كما وجدت خليتان كانت فيها قيمة القوة بين ٠,٦٠-٠,٦٩ وحجم العينة صغيرا ومعلم التمييز منخفضا.

كما يتضح أن قوة الاختبار عندما كان التوزيعان مختلفين، حيث مستوى القدرة للمجموعة المستهدفة أقل من المجموعة المرجعية كانت ≤ ٠,٨٠ في كل الخلايا التي كان فيها حجم العينة كبيرا باستثناء الخلية التي كان فيها معلم التمييز منخفضا والنسبة بين المجموعتين متساوية (١:١)، وفي كل الخلايا التي كان فيها حجم الخلية متوسطا باستثناء الخلايا التي كان معلم التمييز لها منخفضا، وفي خلية كان فيها حجم العينة صغيرا ومتساويا بين المجموعتين (١:١) ومعلم التمييز مرتفعا، كما وجدت خلية واحدة كانت قيمة القوة لها بين ٠,٧٠-٠,٧٩ وقد كان لها حجم صغير غير متساو بين المجموعتين المرجعية والمستهدفة (٢: ١) ومعلم تمييز مرتفع، كما وجدت خليتان قيمة القوة لهما بين ٠,٦٠-٠,٦٩ فيهما النسبة بين حجم المجموعتين غير متساوية (٢: ١) إحداهما ذات حجم عينة صغير ومعلم تمييز متوسط والأخرى ذات حجم عينة متوسط ومعلم تمييز منخفض.

ج-طريقة اختبار تحيز الفقرة المتزامن SIBTEST

يتضح من الجدول السابق أنه في حالة تساوي توزيع القدرات للمجموعتين المرجعية والمستهدفة أن قوة الاختبار الإحصائي كانت ≤ من ٠,٨٠ في كل الخلايا التي كان فيها حجم العينة كبيرا وفي كل الخلايا التي كان فيها حجم الخلية متوسطا باستثناء الخلايا التي كان فيها معلم التمييز منخفضا،

كما وجدت ثلاث خلايا كانت قيمة القوة لها بين ٠,٧٠-٠,٧٩ اثنتان لهما حجم صغير ومعلم تمييز مرتفع لكن إحداهما كانت فيها النسبة بين المجموعتين المرجعية والمستهدفة متساوية (١:١) والأخرى (٢: ١) والثالثة ذات حجم عينة متوسط ومتساو (١:١) ومعلم تمييز منخفض، كما كان هنالك خليتان قيمة القوة لهما بين ٠,٦٠-٠,٦٩ غير متساويتين في حجم العينة (٢: ١) لكن إحداهما ذات حجم عينة صغير ومعلم تمييز متوسط والأخرى لها حجم عينة متوسط ومعلم تمييز منخفض.

كما يتضح أن قوة الاختبار عندما كان التوزيعان مختلفين، حيث مستوى القدرة للمجموعة المستهدفة أقل من المجموعة المرجعية كانت ≤ ٠,٨٠ في كل الخلايا التي كان فيها حجم العينة كبيرا وفي خليتين كان حجم العينتين لهما متوسطا ومتساويا بين المجموعتين (١:١) ومعلم التمييز لإحداهما متوسطا والأخرى مرتفعا، كما وجدت ثلاث خلايا تراوحت قيمة القوة لها بين

All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

المنارة للاستشارات
www.manaraa.com

٠,٧٠-٠,٧٩ اثنتيــن لهما حجم متوسط وغير متساو بين المجموعتين المرجعية والمستهدفة (٢: ١) لكن إحداهما ذات تمييز منخفض والأخرى مرتفع والثالثة كانت ذات حجم صغير ومتساو بين المجموعتين المرجعية والمستهدفة (١:١) ومعلم تمييز مرتفع. كما وجدت خلية واحدة لها قوة تتراوح بين ٠,٦٠-٠,٦٩ كانت ذات حجم عينة متوسط وغير متساو بين المجموعتين المرجعية والمستهدفة (٢: ١) ومعلم تمييز منخفض .

All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

المنارة للاستشارات
www.manaraa.com

ج- قوة الاختبار الإحصائي في حالة وجود دالة فقرة تفاضلية غير منتظمة

الجدول ٩. القوة للطرائق الأربع وفقا لتوزيع القدرة وحجم العينة ونسبتها بين المجموعة المرجعية والمجموعة المستهدفة ومعلم التمييز في حالة وجود دالة فقرة تفاضلية غير منتظمة

توزيع القدرة	الطريقة		مانتل هينزل	الانحدار اللوجستي	نسبة الأرجحية	اختبار تحيز الفقرة المتزامن
	حجم العينة ونسبتها	الفرق بين معلمي التمييز				
N(0,1) N(0,1)	(٣٠٠ ، ٣٠٠)	٠,٥	٠,١٠	٠,٢١	٠,٥٩	٠,١٥
		١	٠,١٩	٠,٣٤	٠,٨١	٠,١٨
		١,٥	٠,٢٥	٠,٣٩	٠,٨٤	٠,٢٩
	(٦٠٠ ، ٦٠٠)	٠,٥	٠,٢٢	٠,٢٠	٠,٦٠	٠,١٩
		١	٠,٣٠	٠,٧١	٠,٩٤	٠,٣٥
		١,٥	٠,٤٥	٠,٨١	٠,٩٨	٠,٤٦
	(١٢٠٠ ، ١٢٠٠)	٠,٥	٠,٣١	٠,٥٤	٠,٨٦	٠,٣٢
		١	٠,٥١	٠,٧٤	٠,٩٨	٠,٥٢
		١,٥	٠,٥٤	٠,٩٩	١,٠٠	٠,٥٨
	(٢٠٠ ، ٤٠٠)	٠,٥	٠,١٦	٠,٢١	٠,٣٢	٠,١٥
		١	٠,٢٥	٠,٢٩	٠,٧٠	٠,١٥
		١,٥	٠,٢٦	٠,٤٧	٠,٨٤	٠,٢١
	(٤٠٠ ، ٨٠٠)	٠,٥	٠,٢٠	٠,٣٢	٠,٦٤	٠,٢٣
		١	٠,٢٥	٠,٤٧	٠,٨٤	٠,٢٦
		١,٥	٠,٣٦	٠,٩١	٠,٩٧	٠,٤٢
	(٨٠٠ ، ١٦٠٠)	٠,٥	٠,٤١	٠,٦٤	٠,٩٠	٠,٤٠
		١	٠,٤٧	٠,٨٨	٠,٩٦	٠,٤٣
		١,٥	٠,٤٢	٠,٩٨	١,٠٠	٠,٤٧
N(0,1) N(0.5,1)	(٣٠٠ ، ٣٠٠)	٠,٥	٠,٢٢	٠,١٩	٠,٥٦	٠,٢٢
		١	٠,٢١	٠,٣٣	٠,٦٩	٠,٢٩
		١,٥	٠,٢٩	٠,٣٨	٠,٨٧	٠,٢٢
	(٦٠٠ ، ٦٠٠)	٠,٥	٠,١٦	٠,١٥	٠,٧٣	٠,٢٧
		١	٠,٣٠	٠,٤٠	٠,٩٦	٠,٣٥
		١,٥	٠,٥٩	٠,٥١	٠,٩٨	٠,٤٥
	(١٢٠٠ ، ١٢٠٠)	٠,٥	٠,٢٢	٠,٣٠	٠,٨٢	٠,٣٥
		١	٠,٣١	٠,٥٠	١,٠٠	٠,٣٥
		١,٥	٠,٥١	٠,٨٤	١,٠٠	٠,٥٦
	(٢٠٠ ، ٤٠٠)	٠,٥	٠,٠٨	٠,١٢	٠,٣٦	٠,١٢
		١	٠,٢٠	٠,٢١	٠,٥٧	٠,٢٨
		١,٥	٠,٢١	٠,٣٠	٠,٨٧	٠,٢٢
	(٤٠٠ ، ٨٠٠)	٠,٥	٠,٢٠	٠,٢٠	٠,٦٩	٠,٢٠
		١	٠,٢٣	٠,٣٧	٠,٩٣	٠,٣١
		١,٥	٠,٢٩	٠,٥٥	٠,٩٨	٠,٤٥
	(٨٠٠ ، ١٦٠٠)	٠,٥	٠,٢٤	٠,٣١	٠,٨١	٠,٢٠
		١	٠,٤٦	٠,٧١	١,٠٠	٠,٥٣
		١,٥	٠,٤٦	٠,٨١	١,٠٠	٠,٥٤
المتوسط العام			٠,٣٠	٠,٤٨	٠,٨٢	٠,٣٢

All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

المنارة للاستشارات
www.manaraa.com

يتضح من الجدول السابق ما يلي :-

أ- مانتل هينزل

كما كان متوقعا كانت نسبة الكشف عن دالة الفقرة التفاضلية متدنية، حيث كانت جميع قيم القوة لها أقل من ٠,٦٠ لأن طريقة مانتل هينزل غير مصممة في الأصل للكشف عن دالة الفقرة التفاضلية غير المنتظمة، وقد تراوحت قيم القوة لها بين (٠,٠٨-٠,٥٩) وبمعدل ٠,٣٠ .

ب- الانحدار اللوجستي

يتضح من الجدول السابق أنه في حالة تساوي توزيع القدرات للمجموعتين المرجعية والمستهدفة أن قوة الاختبار الإحصائي كانت ≤ من ٠,٨٠ في الخلايا التي كان فيها معلم التمييز مرتفعا وحجم العينة كبيرا أو متوسطا وفي خلية كان فيها حجم العينة كبيرا ومعلم التمييز متوسطا كما كانت قيمة القوة بين ٠,٧٠-٠,٧٩ في خليتين كانت نسبة العينة فيهما متساوية (١-١) ومعلم التمييز لهما متوسطا إحداهما ذات حجم عينة كبير والأخرى متوسط ، كذلك كانت القوة بين ٠,٦٠-٠,٦٩ في خلية واحدة فقط كان فيها حجم العينة كبيرا ومعلم التمييز متوسطا ونسبة العينة متساوية (١-١). أما في حالة عدم تساوي متوسط توزيع القدرة بين المجموعتين، حيث مستوى القدرة للمجموعة المستهدفة أقل من المجموعة المرجعية فقد كانت قوة الاختبار الإحصائي ≤ من ٠,٨٠ في خليتين فقط كان فيهما حجم العينة كبيرا ومعلم التمييز مرتفعا.

ج- طريقة اختبار تحيز الفقرة المتزامن SIBTEST

كما كان متوقعا كانت نسبة الكشف عن دالة الفقرة التفاضلية لطريقة متدنية أيضا، حيث كانت جميع قيم القوة لها أقل من ٠,٦٠، لأن طريقة SIBTEST غير مصممة في الأصل للكشف عن دالة الفقرة التفاضلية غير المنتظمة، وقد تراوحت قيم القوة لها بين (٠,١٢ - ٠,٥٨) وبمعدل ٠,٣٢ .

All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

المنارة للاستشارات
www.manaraa.com

الســـــــؤال الثـــــــالث

هل يوجد تأثير للمتغيرات (حجم العينة ، نسبة العينة بين المجموعة المرجعية والمجموعة المستهدفة، توزيع القدرة، معلم التمييز) على الخطأ من النوع الأول والقوة للطرائق الأربع (مانتل هينزل Mantel Haenszel، والانحـــدار اللوجستي Logistic Regression، واختبار التحيز المتزامن SIBTEST، ونسبة الأرجحية IRTLRDIF)؟

ا- الخطـــــأ من النـــــوع الأول

أ- طريقة مانتل هينزل

الجدول ١٠ .نتائج تحليل الانحدار اللوجستي لتأثير متغيرات توزيع القدرة وحجم العينة ونسبتها ومعلم التمييز والتفاعل بينها على الخطأ من النوع الأول لطريقة مانتل هينزل

المتغير	B	wald	درجات الحرية	القيمة الاحتمالية
القدرة	٠,٣٥٨	٠,١٩٨	١	٠,٦٥٦
حجم العينة	٠,١٧٨	٠,١٧٧	١	٠,٧٣٢
معلم التمييز	٠,٣٥-	٠,٠٠٤	١	٠,٩٤٨
نسبة العينة	٠,٩٨٨-	١,٢٩٧	١	٠,٢٥٥
حجم العينة * معلم التمييز	٠,٠٣٤	٠,٧٠	١	٠,٧٩١
حجم العينة * نسبة العينة	٠,٠٦٣	٠,٩٠	١	٠,٧٦٤
حجم العينة*توزيع القدرة	٠,١٣٤-	٠,٣٩٠	١	٠,٥٣٢
معلم التمييز*توزيع القدرة	٠,٣١-	٠,٠٢١	١	٠,٨٨٤
معلم التمييز*نسبة العينة	٠,٠٠٤-	٠,٠٠٠	١	٠,٩٨٥
توزيع القدرة * نسبة العينة	٠,٣٦٥	١,٠٥٧	١	٠,٣٠٤
الثابت	٣,١٩١-	٣,٦٢٣	١	٠,٠٥٧

أظهرت النتائج عدم وجود تأثير ذي دلالة إحصائية لأي من متغيرات الدراسة أو التفاعلات الثنائية بين هذه المتغيرات على الخطأ من النوع الأول حيث كانت جميع القيم الاحتمالية للإحصائي Wald والمتعلقة بمعاملات المتغيرات والتفاعلات الثنائية بينها أعلى من (٠,٠٥) .

All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

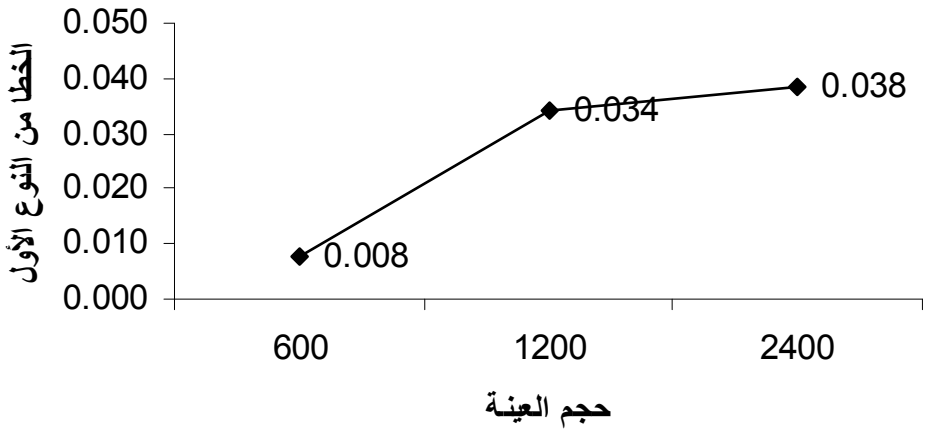
المنارة للاستشارات
www.manaraa.com

ب- طريقة الانحـــدار اللوجســـــتي

الجدول ١١. نتائج تحليل الانحدار اللوجستي لتأثير متغيرات توزيع القدرة وحجم العينة ونسبتها ومعلم التمييز والتفاعل بينها على الخطأ من النوع الأول لطريقة الانحدار اللوجستي

المتغير	B	wald	درجات الحرية	القيمة الاحتمالية
القدرة	٠,٣٩٨	٠,٠٩٠	١	٠,٧٦٤
حجم العينة	٢,٢٥١	٧,٤٣٧	١	٠,٠٠٦
معلم التمييز	١,٩٤٨	٥,٦١٤	١	٠,٠١٨
نسبة العينة	١,٤٤٤	١,٢٨١	١	٠,٢٥٨
حجم العينة * معلم التمييز	-٠,٢٨٠	٢,١٥٣	١	٠,١٤٢
حجم العينة * نسبة العينة	-٠,٧٣٤	٦,٤٨٠	١	٠,٠١١
حجم العينة*توزيع القدرة	٠,٠٨١	٠,٠٧٨	١	٠,٧٨٠
معلم التمييز*توزيع القدرة	-٠,٣٢٦	١,٣٣٩	١	٠,٢٤٧
معلم التمييز*نسبة العينة	-٠,١٧٢	٠,٣٧٤	١	٠,٥٤١
توزيع القدرة * نسبة العينة	٠,٣٦٤	٠,٦٨٣	١	٠,٤٠٩
الثابت	-١١,٣١٨	١٣,٦٣٧	١	٠,٠٠٠

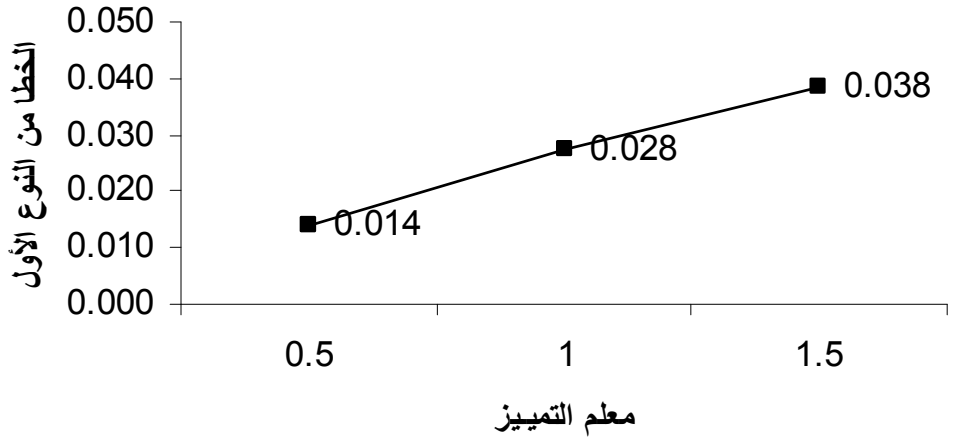
أظهرت النتـــائج وجـــود تأثيـــر ذي دلالـــة إحـــصائية لمتغيـــــر حجـــــم العينـــة (Wald=7.437 P. =0.006) ومعلـــم التمييز (Wald = 5.614, P. =0.018) والتفاعل بين حجم العينة ونسبتها (Wald= 6.48 , P. =0.011) على الخطأ من النـــوع الأول ، حيث أن الخطأ من النوع الأول كان يزداد بزيادة حجم العينة ومعلم التمييز كما يتضح في الرسم البياني في الشكلين رقم (٣) ورقم (٤)



الشكل ٣ . تأثير حجم العينة على الخطأ من النوع الأول لطريقة الانحدار اللوجستي

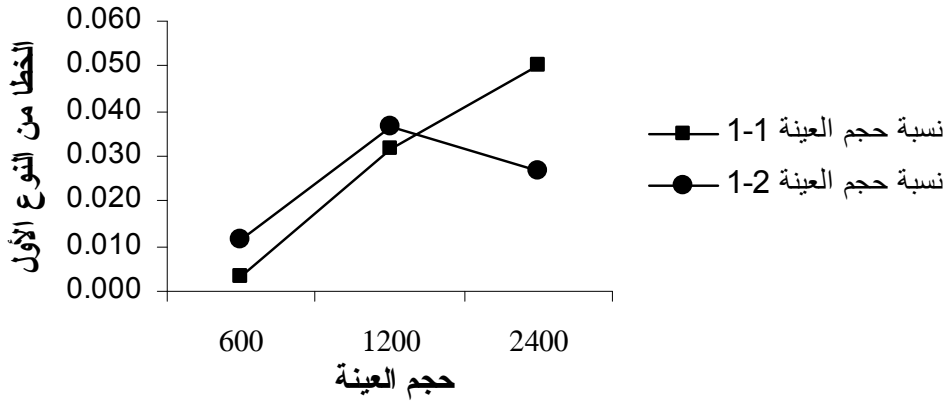
All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

المنارة للاستشارات
www.manaraa.com



الشكل ٤ . تأثير معلم التمييز على الخطأ من النوع الأول لطريقة الانحدار اللوجستي

ويوضح الرسم البياني التالي التفاعل بين حجم العينة ونسبتها



الشكل ٥ .تأثير التفاعل بين حجم العينة ونسبتها على الخطأ من النوع الأول لطريقة الانحدار اللوجستي

يلاحظ من الرسم البياني أن نسبة الخطأ من النوع الأول كانت أكبر عندما كانت النسبة بين حجمي المجموعة المرجعية والمستهدفة ٢- ١ وحجم العينة صغيرا ومتوسطا، وأعلى للنسبة ١-١ عندما كان حجم العينة كبيرا .

All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

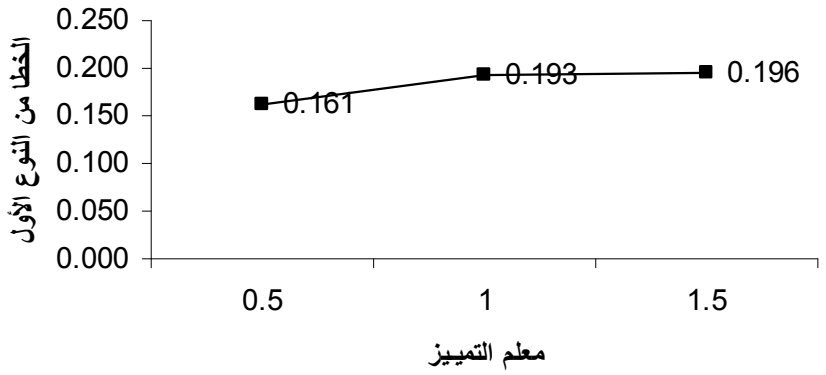
المنارة للاستشارات
www.manaraa.com

ج- طريقة نسبة الأرجحية (Likelihood-Ratio Tests)

الجدول ١٢. نتائج تحليل الانحدار اللوجستي لتأثير متغيرات توزيع القدرة وحجم العينة ونسبتها ومعلم التمييز والتفاعل بينها على الخطأ من النوع الأول لطريقة نسبة الأرجحية

المتغير	B	wald	درجات الحرية	القيمة الاحتمالية
القدرة	٠,٥٧-	٠,٠١٨	١	٠,٨٩٣
حجم العينة	٠,١٧٩	٠,٤٣١	١	٠,٥١١
معلم التمييز	٠,٥٩٠-	٤,٢٠٢	١	٠,٠٤٠
نسبة العينة	٠,٦٨٨	٢,٧٩٠	١	٠,٠٩٥
حجم العينة * معلم التمييز	٠,١١٤	٢,٧٩٦	١	٠,٠٩٥
حجم العينة * نسبة العينة	٠,١١٢-	١,٠٣٨	١	٠,٣٠٨
حجم العينة*توزيع القدرة	٠,١٤٧-	١,٧٥٢	١	٠,١٨٦
معلم التمييز*توزيع القدرة	٠,٤٣٠	١٤,٨٣٨	١	٠,٠٠٠
معلم التمييز*نسبة العينة	٠,٠٢٩-	٠,٠٦٨	١	٠,٧٩٥
توزيع القدرة * نسبة العينة	٠,٤١٦-	٥,٢٥٢	١	٠,٠٢٢
الثابت	١,٦٩٥-	٣,٩٠٦	١	٠,٠٤٨

أظهرت النتائج وجود تأثير ذو دلالة إحصائية لمعلم التمييز (Wald=4.202, P. =0.04) والتفاعل بين معلم التمييز وتوزيع القدرة (Wald = 14.838, P. =0.000) والتفاعل بين توزيع القدرة ونسبة حجم العينة بين المجموعة المرجعية والمجموعة المستهدفة (Wald= 5.252 , P. =0.022) على الخطأ من النوع الأول ، حيث أن الخطأ من النوع الأول كان يزداد بزيادة معلم التمييز كما يتضح من الشكل التالي :

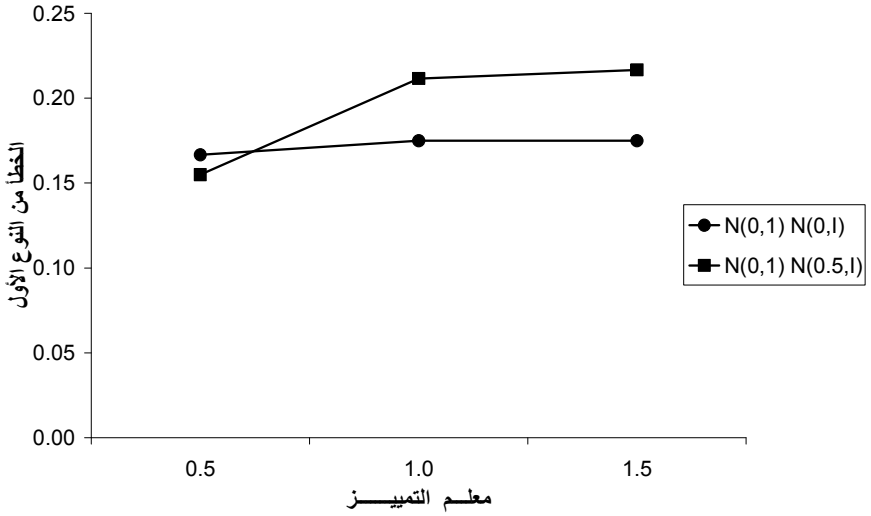


الشكل ٦ . تأثير معلم التمييز على الخطأ من النوع الأول لطريقة نسبة الأرجحية

All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

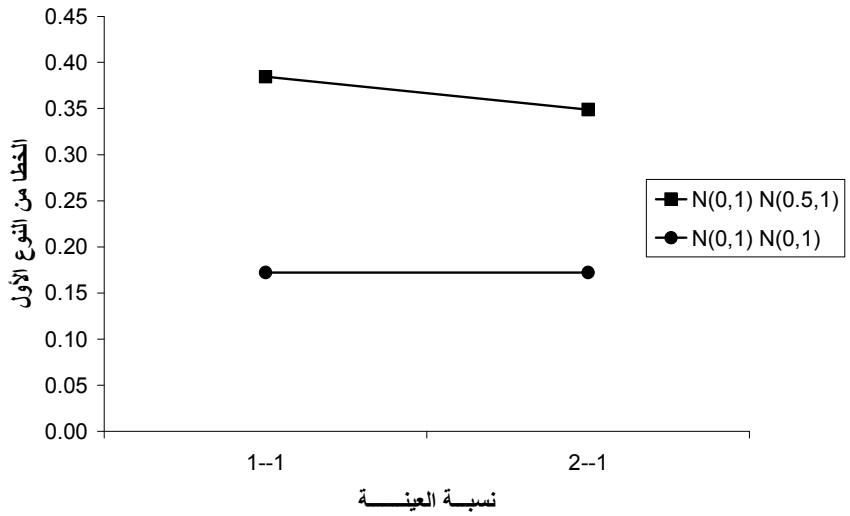
المنارة للاستشارات
www.manaraa.com

وتوضح الأشكال التالية التفاعلات التي كان لها تأثير على الخطأ من النوع الأول لطريقة نسبة الأرجحية



الشكل ٧ . تأثير التفاعل بين معلم التمييز والفروق في توزيع القدرة على الخطأ من النوع الأول لطريقة نسبة الأرجحية

يلاحظ من الرسم البياني أن نسبة الخطأ من النوع الأول كانت أعلى لمعلم التمييز المتوسط (١,٠) والمرتفع (١,٥) في حالة وجود فروق بين متوسط توزيع القدرة للمجموعتين المرجعية والمستهدفة، بينما كانت نسبة الخطأ من النوع الأول أعلى لمعلم التمييز المنخفض (٠,٥) في حالة عدم وجود فروق في توزيع القدرة .



الشكل ٨. تأثير التفاعل بين توزيع القدرة ونسبة حجم العينة بين المجموعة المرجعية والمجموعة المستهدفة على الخطأ من النوع الأول لطريقة نسبة الأرجحية

All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

المنارة للاستشارات
www.manaraa.com

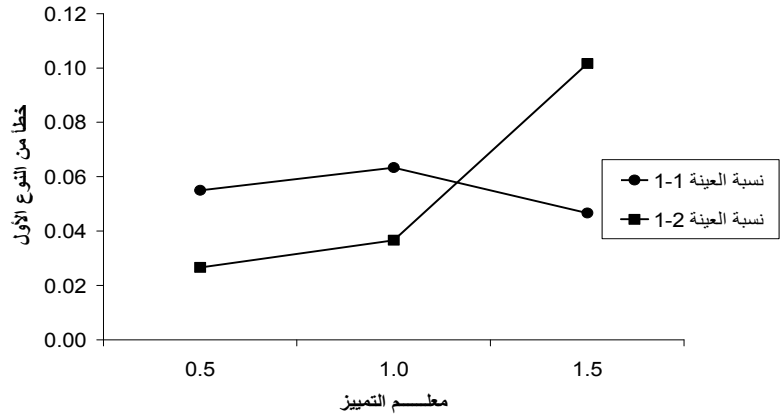
يلاحظ من الشكل السابق أن الفروق في نسبة الخطأ من النوع الأول الناتجة عن الاختلاف في توزيع القدرة بين المجموعتين المرجعية والمستهدفة كانت أعلى في الحالة التي تكون فيها النسبة بين المجموعة المرجعية والمستهدفة (١-١) من الحالة التي تكون فيها النسبية بين المجموعتين (١-٢).

د- طريـــــــقة اختبار تحيز الفقرة المتزامن SIBTEST

الجدول ١٣. نتائج تحليل الانحدار اللوجستي لتأثير متغيرات توزيع القدرة وحجم العينة ونسبتها ومعلم التمييز والتفاعل بينها على الخطأ من النوع الأول لطريقة SIBTEST

المتغير	B	wald	درجات الحرية	القيمة الاحتمالية
توزيع القدرة	١,٢٢٣	٢,٨٦٤	١	٠,٠٩١
حجم العينة	٠,٣٢٧	٠,٤٤٣	١	٠,٥٠٥
معلم التمييز	٠,١٥١	٠,٠٩٢	١	٠,٧٦١
نسبة العينة	١,٣٩٥-	٢,٧٥٤	١	٠,٠٩٧
حجم العينة * معلم التمييز	٠,٠٩٦-	٠,٦٢٤	١	٠,٤٣٠
حجم العينة * نسبة العينة	٠,١٢٧	٠,٤١٨	١	٠,٥١٨
حجم العينة*توزيع القدرة	٠,١٢٧-	٠,٤٢٦	١	٠,٥١٤
معلم التمييز*توزيع القدرة	٠,٣٧٦-	٣,٦١٢	١	٠,٠٥٧
معلم التمييز*نسبة العينة	٠,٥١٨	٦,٨٧٩	١	٠,٠٠٩
توزيع القدرة * نسبة العينة	٠,٠١٦-	٠,٠٠٣	١	٠,٩٥٩
الثابت	٣,٧٤٧-	٥,٥٢٧	١	٠,٠١٩

أظهرت النتائج وجود تأثير ذي دلالة إحصائية للتفاعل بين معلم التميز ونسبة حجم العينة بين المجموعة المرجعية والمجموعة المستهدفة (Wald=6.879 P. =0.009) على الخطأ من النوع الأول. والرسم البياني التالي يوضح هذا التفاعل



الشكل ٩. تأثير التفاعل بين معلم التمييز ونسبة العينة على الخطأ من النوع الأول لطريقة SIBTEST

All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

المنارة للاستشارات
www.manaraa.com

يتضح من الرسم البياني أن نسبة الخطأ من النوع الأول كانت أعلى لمعلم التمييز المنخفض والمتوسط عندما كانت نسبة العينة بين المجموعتين المرجعية والمستهدفة ١-١ بينما كانت نسبة الخطأ لمعلم التمييز المرتفع أعلى عندما كانت النسبة ٢-١.

ثانيا : قوة الاختبار الإحصائي (POWER)

١- قوة الاختبار عند وجود دالة فقرة تفاضلية مقدارها ٠,٢٥ .

أ-طريقة مانتل هينزل

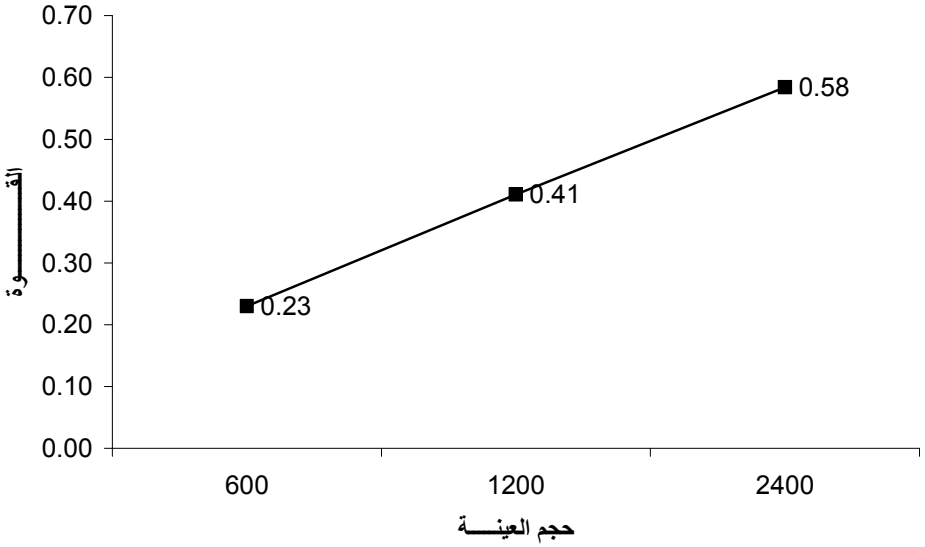
الجدول ١٤. نتائج تحليل الانحدار اللوجستي لتأثير متغيرات توزيع القدرة وحجم العينة ونسبتها ومعلم التمييز والتفاعل بينها على القوة لطريقة مانتل هينزل عند وجود دالة فقرة تفاضلية مقدارها ٠,٢٥

المتغير	B	wald	درجات الحرية	القيمة الاحتمالية
توزيع القدرة	-٠,٠٦٩	٠,٠٣٢	١	٠,٨٥٧
حجم العينة	٠,٨٧٤	١٣,١٧٠	١	٠,٠٠٠
معلم التمييز	٠,٥٧٣	٥,٦٢٤	١	٠,٠١٨
نسبة العينة	٠,٣٠٦	٠,٦٣٦	١	٠,٤٢٥
حجم العينة * معلم التمييز	٠,١٣٦	٥,١٢٧	١	٠,٠٢٤
حجم العينة * نسبة العينة	-٠,٠٣	٠,٠٩٩	١	٠,٧٥٢
حجم العينة*توزيع القدرة	-٠,١٧٢	٣,٢٢٧	١	٠,٠٧٢
معلم التمييز*توزيع القدرة	٠,١٨٥	٣,٧٥٩	١	٠,٠٥٣
معلم التمييز*نسبة العينة	-٠,٢٠٠	٤,٣٨٥	١	٠,٠٣٦
توزيع القدرة * نسبة العينة	٠,٠٠٦	٠,٠٠١	١	٠,٩٧٠
الثابت	-٣,٥٨٦	٢٠,٢٦١	١	٠,٠٠٠

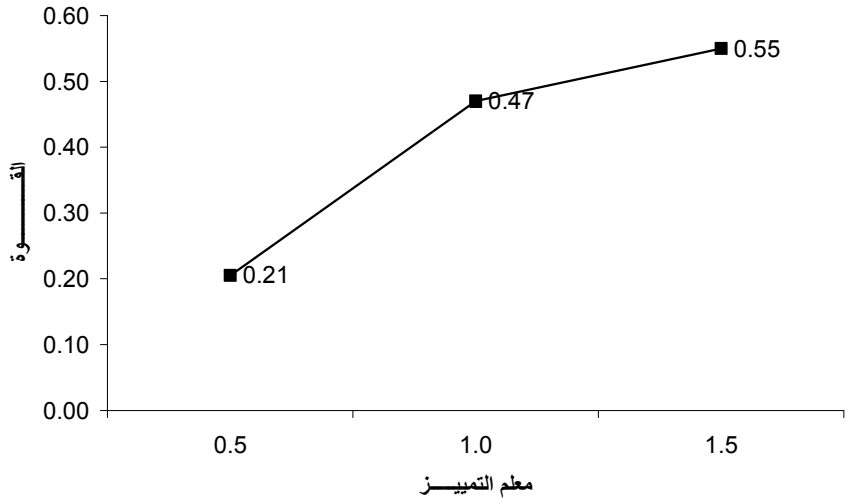
أظهرت النتائج وجود تأثير ذي دلالة إحصائية لمتغير حجم العينة (Wald=13.17 P. =0.000) ومعلم التمييز (Wald = 5.624, P. =0.018) على القوة، حيث كانت القوة تزداد بزيادة حجم العينة ومعلم التمييز كما يظهر في الرسوم البيانية التالية :

All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

المنارة للاستشارات
www.manaraa.com



الشكل ١٠. تأثير حجم العينة على القوة لطريقة مانتل هينزل عند وجود دالة فقرة تفاضلية مقدارها ٠,٢٥

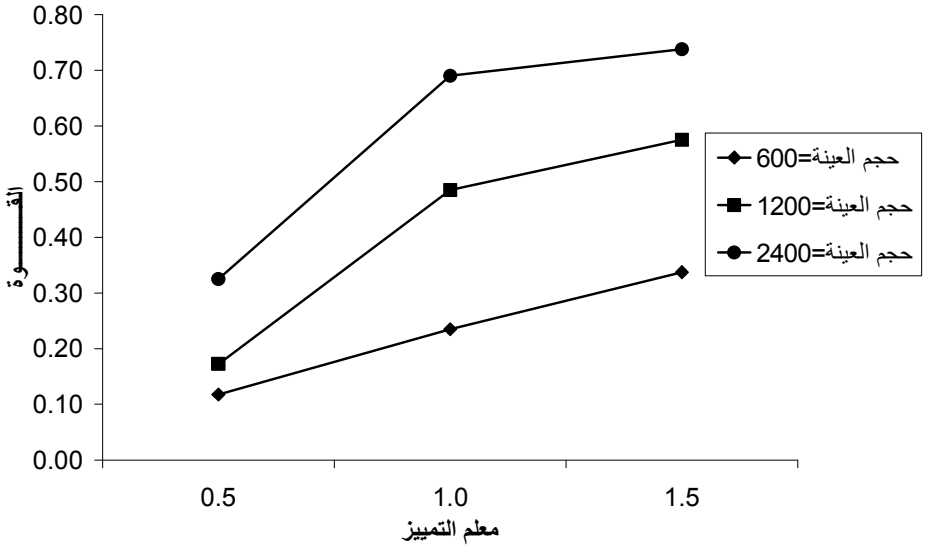


الشكل ١١. تأثير معلم التمييز على القوة لطريقة مانتل هينزل عند وجود دالة فقرة تفاضلية مقدارها ٠,٢٥

كما أظهـــــرت النتائج وجود تأثير للتفاعل بين حجم العينة ومعلم التمييز (Wald = 5.127, P. =0.0.24) والتفاعل بين معلم التمييز ونسبة العينة (Wald=4.385 P. =0.036) على القوة، والرسوم البيانية التالية توضح تأثير التفاعلين

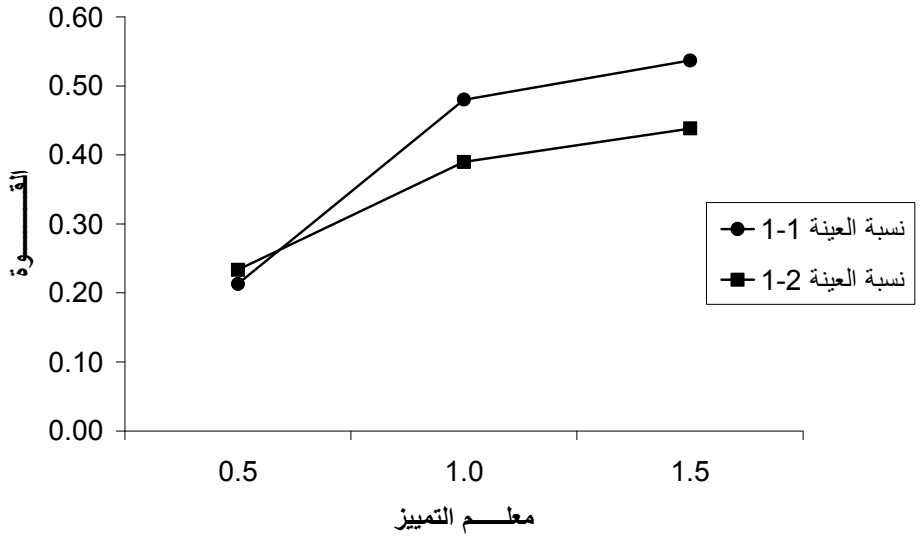
All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

المنارة للاستشارات
www.manaraa.com



الشكل ١٢. تأثير التفاعل بين معلم التمييز والحجم العينة على القوة لطريقة مانتل هينزل عند وجود دالة فقرة تفاضلية مقدارها ٠,٢٥

يتضح من الرسم البياني أن الفروق في القوة الناتجة عن اختلاف حجم العينة تكون أكبر عندما يكون معلم التمييز مرتفعا ومتوسطا من الحالة التي يكون فيها معلم التمييز منخفضا.



الشكل ١٣. تأثير التفاعل بين معلم التمييز ونسبة العينة على القوة لطريقة مانتل هينزل عند وجود دالة فقرة تفاضلية مقدارها ٠,٢٥

All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

المنارة للاستشارات
www.manaraa.com

يتضح من الرسم البياني أن الفروق في القوة الناتجة عن اختلاف نسبة العينة كانت لصالح الحالة التي كانت فيها نسبة العينة ١-١ عندما كان معلم التمييز متوسطا ومرتفعا، بينما كانت لصالح الحالة التي تكون فيها نسبة العينة ٢-١ عندما كان معلم التمييز ٠,٥.

ب- طريقة الانحـــدار اللوجســـتي

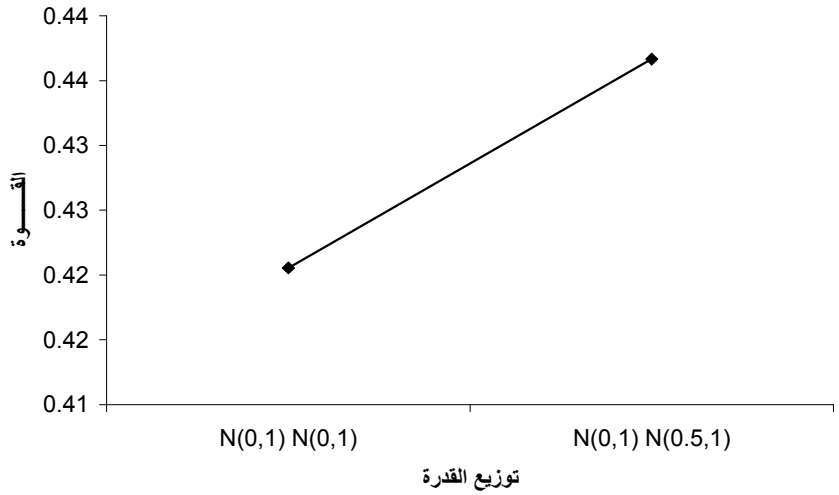
الجدول ١٥. نتائج تحليل الانحدار اللوجستي لتأثير متغيرات توزيع القدرة وحجم العينة ونسبتها ومعلم التمييز والتفاعل بينها على القوة لطريقة الانحـــدار اللوجستي عند وجود دالة فقرة تفاضلية مقدارها ٠,٢٥

المتغير	B	wald	درجات الحرية	القيمة الاحتمالية
توزيع القدرة	١,٠٥٦-	٦,٩١٤	١	٠,٠٠٩
حجم العينة	٠,٠٨٧-	٠,١٢٢	١	٠,٧٢٧
معلم التمييز	٠,٣٤٣-	١,٨٣١	١	٠,١٧٦
نسبة العينة	٠,٤٣٣-	١,١٧٧	١	٠,٢٧٨
حجم العينة * معلم التمييز	٠,٤٣٥	٤٦,٦٣٧	١	٠,٠٠٠
حجم العينة * نسبة العينة	٠,٠٢٨	٠,٠٧٩	١	٠,٧٧٨
حجم العينة*توزيع القدرة	٠,٠٢٧	٠,٠٧٣	١	٠,٧٨٦
معلم التمييز*توزيع القدرة	٠,٣٦٠	١٢,٤٤٨	١	٠,٠٠٠
معم التمييز*نسبة العينة	٠,٠٢٧	٠,٠٦٨	١	٠,٧٩٤
توزيع القدرة * نسبة العينة	٠,٢٣٦	٢,٣٢١	١	٠,١٢٨
الثابت	٠,٨١١-	٠,٩٨٣	١	٠,٣٢١

أظهـــــرت النتـــائج وجـــود تأثيـــر ذي دلالة إحصـــائية لمتغيـــر توزيع القدرة (Wald = 6,914 P. =0.009) على القوة حيث ازدادت القوة مع وجود فروق في توزيع القدرة كما يتضح من الرسم البياني التالي :

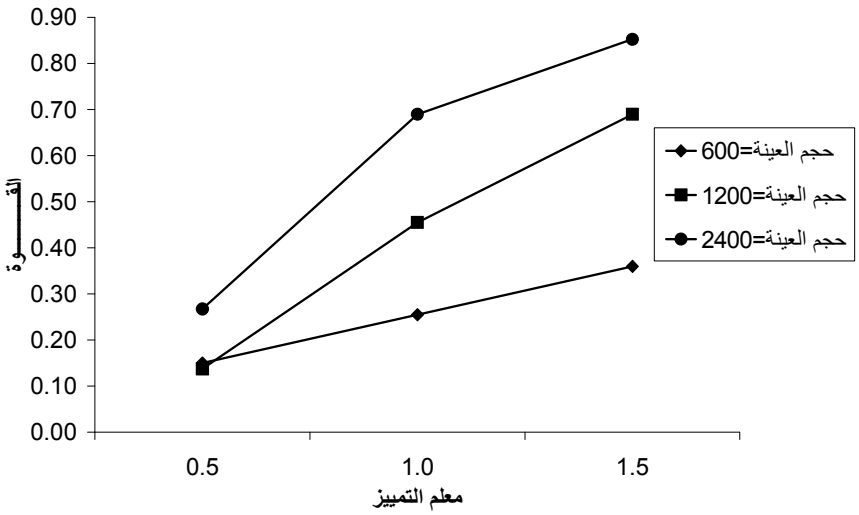
All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

المنارة للاستشارات
www.manaraa.com



الشكل ١٤. تأثير الاختلاف في متوسط توزيع القدرة بين المجموعتين المرجعية والمستهدفة على القوة لطريقة الانحدار اللوجستي عند وجود دالة فقرة تفاضلية مقدارها ٠,٢٥

كما أظهرت النتائج وجود تأثير للتفاعل بين حجم العينة ومعلم التمييز (Wald = 46,637 P. =0.000) والتفاعل بين معلم التمييز وتوزيع القدرة على القوة (Wald = 12.448 P. =0.000) كما يتضح من الرسوم البيانية التالية

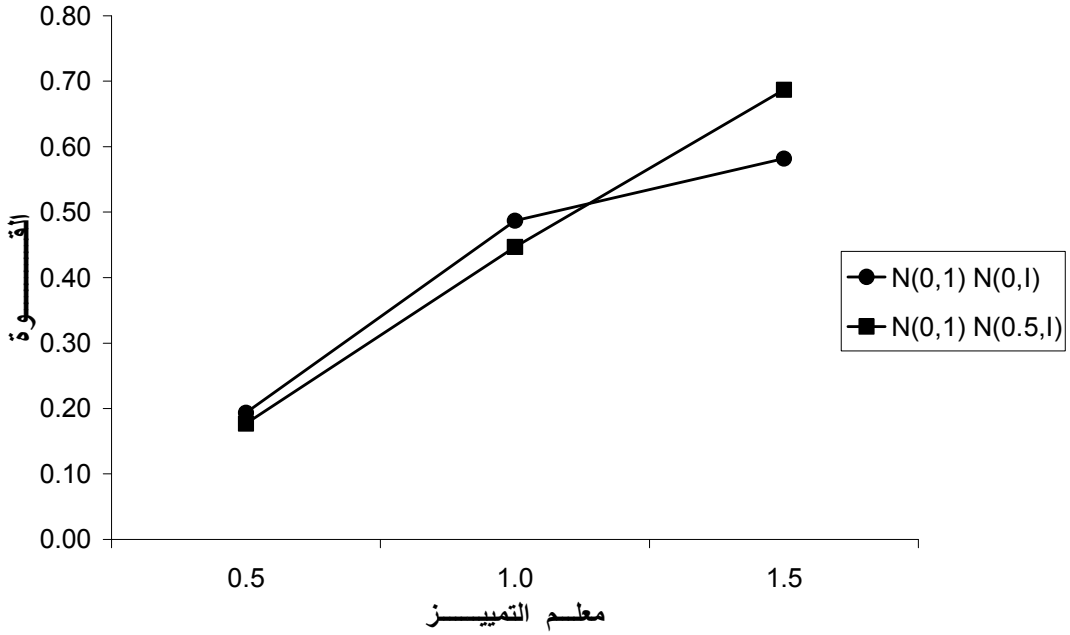


الشكل ١٥. تأثير التفاعل بين معلم التمييز وحجم العينة على القوة لطريقة الانحدار اللوجستي عند وجود دالة فقرة تفاضلية مقدارها ٠,٢٥

يتضح من الرسم البياني أن الفروق في القوة الناتجة عن اختلاف حجم العينة ازدادت مع ازدياد معلم التمييز

All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

المنارة للاستشارات
www.manaraa.com



الشكل ١٦. تأثير التفاعل بين معلم التمييز وتوزيع القدرة على القوة لطريقة الانحدار اللوجستي عند وجود دالة فقرة تفاضلية مقدارها ٠,٢٥

يتضح من الرسم البياني أن الفروق الناتجة عن اختلاف توزيع القدرة بين المجموعة المرجعية والمستهدفة كانت لصالح الحالة التي كان فيها للمجموعتين نفس توزيع القدرة عندما كان معلم التمييز منخفضا ومتوسطا ولصالح الحالة التي كانت فيها اختلاف في متوسط توزيع القدرة بين المجموعتين عندما كان معلم التمييز مرتفعا .

All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

المنارة للاستشارات
www.manaraa.com

ج- طـــريقـــة اختبار تحيز الفقرة المتزامن SIBTEST

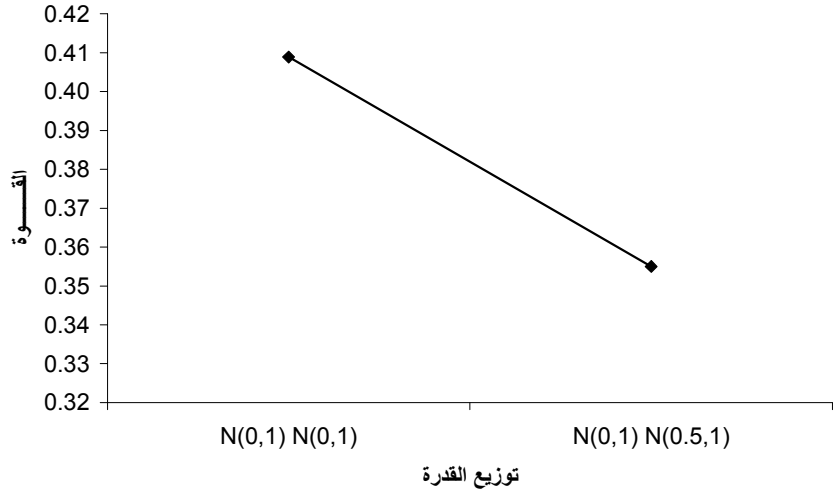
الجدول ١٦. نتائج تحليل الانحدار اللوجستي لتأثير متغيرات توزيع القدرة وحجم العينة ونسبتها ومعلم التمييز والتفاعل بينها على القوة لطريقة SIBTEST عند وجود دالة فقرة تفاضلية مقدارها ٠,٢٥

المتغير	B	wald	درجات الحرية	القيمة الاحتمالية
توزيع القدرة	١,٠١١	٧,٥٨٠	١	٠,٠٠٦
حجم العينة	١,٢٠٨	٢٥,٧١٢	١	٠,٠٠٠
معلم التمييز	٠,٩٢٤	١٥,١٧٧	١	٠,٠٠٠
نسبة العينة	٠,٧٦٨	٤,٣٤٠	١	٠,٠٣٧
حجم العينة * معلم التمييز	٠,١٥٥	٧,١٣٠	١	٠,٠٠٨
حجم العينة * نسبة العينة	١,٣٣-	٢,٠٢٢	١	٠,١٥٥
حجم العينة*توزيع القدرة	٠,٣٨٩-	١٧,٢٤٥	١	٠,٠٠٠
معلم التمييز*توزيع القدرة	٠,١٣١-	١,٩٩٣	١	٠,١٥٨
معلم التمييز*نسبة العينة	٠,٢٧٧-	٨,٦٠٦	١	٠,٠٠٣
توزيع القدرة * نسبة العينة	٠,١٢٥-	٠,٧١٥	١	٠,٣٩٨
الثابت	٥,٠١٥-	٤١,٣٨٢	١	٠,٠٠٠

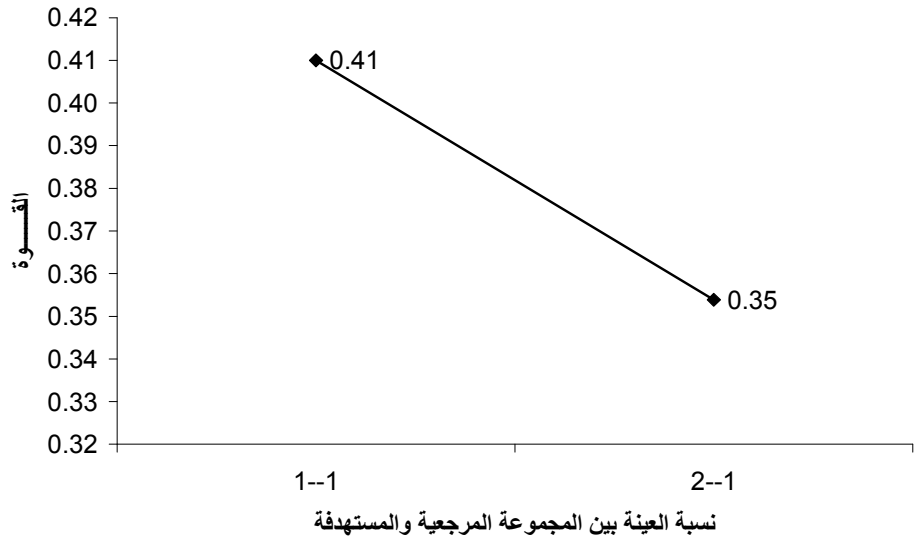
أظهـــــرت النتـــائج وجـود تأثيـــر ذي دلالة إحصـــائية لمتغيـــر توزيـــع القـــدرة (Wald = 7.580 P. =0.006)، ومتغير حجـــم العينة (Wald = 25.712 P. =0.000) ، ومتغير معلم التمييز (Wald = 15.777 P. =0.000)، ومتغير نـــسبة العينـــة (Wald = 4.340 P. =0.037) على القوة حيث انخفضت القوة عند اختلاف توزيع القوة وعند ما كانت نسبة العينة ٢-١، كما أن القوة ازدادت مع ازدياد كل من حجم العينة ومعلم التمييز كما يتضح من الرسوم البيانية التالية :

All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

المنارة للاستشارات
www.manaraa.com



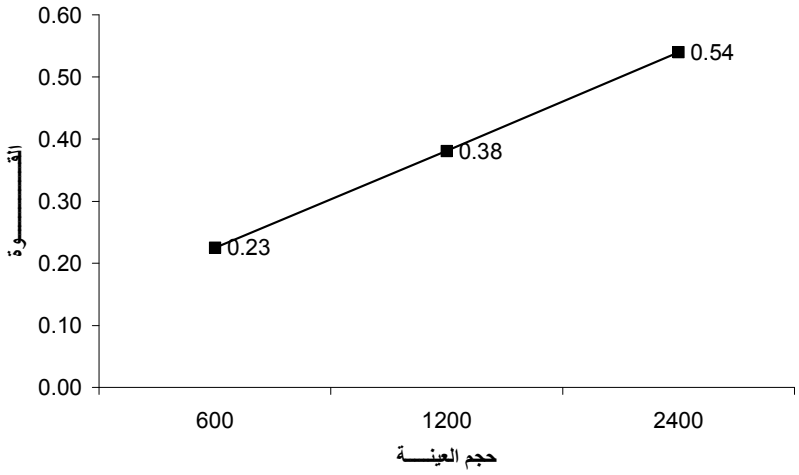
الشكل ١٧. تأثير توزيع القدرة على القوة لطريقة SIBTEST عند وجود دالة فقرة تفاضلية مقدارها ٠,٢٥



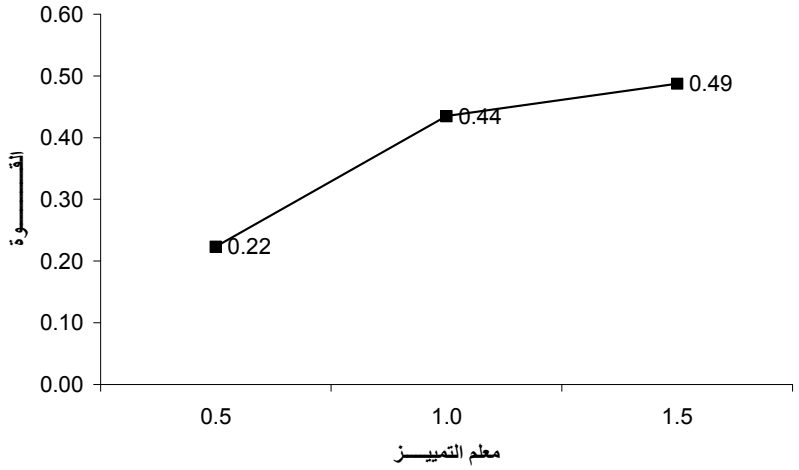
الشكل ١٨. تأثير توزيع نسبة العينة بين المجموعة المرجعية والمجموعة المستهدفة على القوة لطريقة SIBTEST عند وجود دالة فقرة تفاضلية مقدارها ٠,٢٥

All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

www.manaraa.com



الشكل ١٩. تأثير حجم العينة على القوة لطريقة SIBTEST عند وجود دالة فقرة تفاضلية مقدارها ٠,٢٥

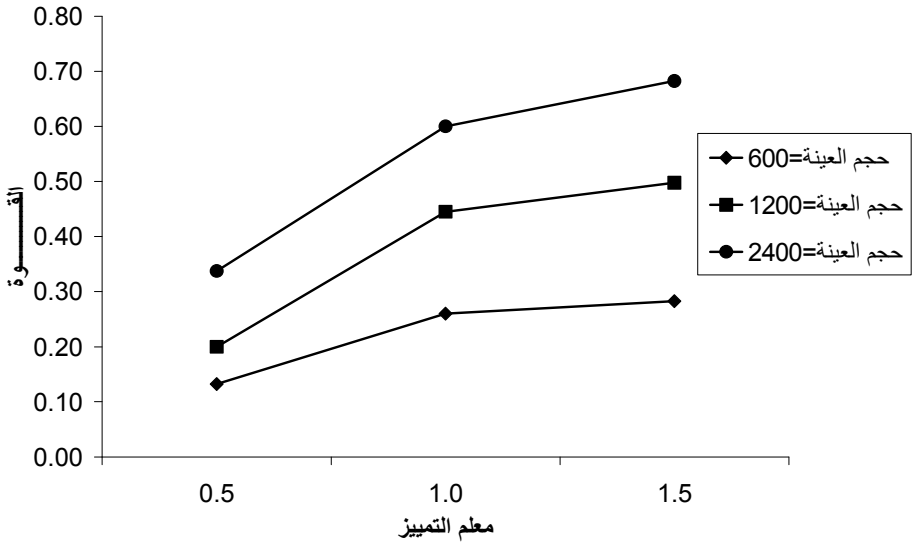


الشكل ٢٠. تأثير معلم التمييز على القوة لطريقة SIBTEST عند وجود دالة فقرة تفاضلية مقدارها ٠,٢٥

كما أظهرت النتائج وجود تأثير للتفاعل بين حجم العينة ومعلم التمييز والتفاعل بين معلم التمييز وتوزيع القدرة على القـــــوة كما يتضح من الرسوم البيانية التالية :

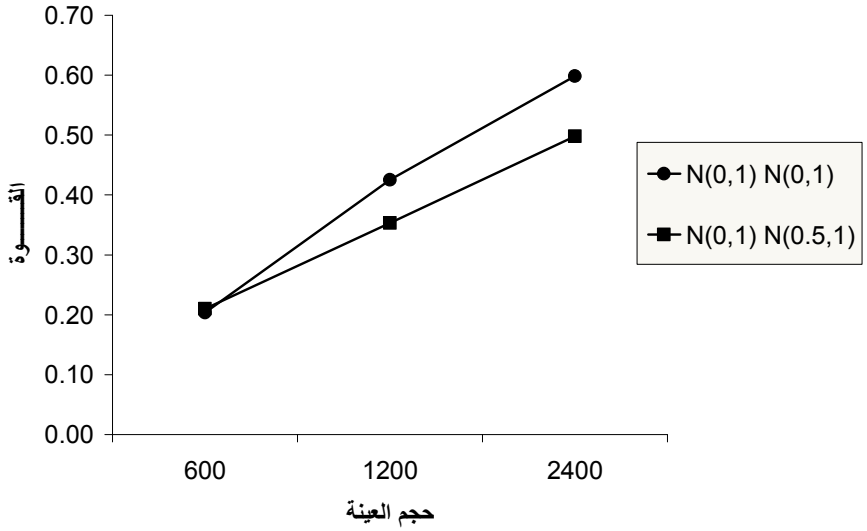
All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

المنارة للاستشارات
www.manaraa.com



الشكل ٢١. تأثير التفاعل بين معلم التمييز وحجم العينة على القوة لطريقة SIBTEST عند وجود دالة فقرة تفاضلية مقدارها ٠,٢٥

يتضح من الرسم البياني أن الفروق في القوة الناتجة عن اختلاف حجم العينة ازدادت مع ازدياد معلم التمييز

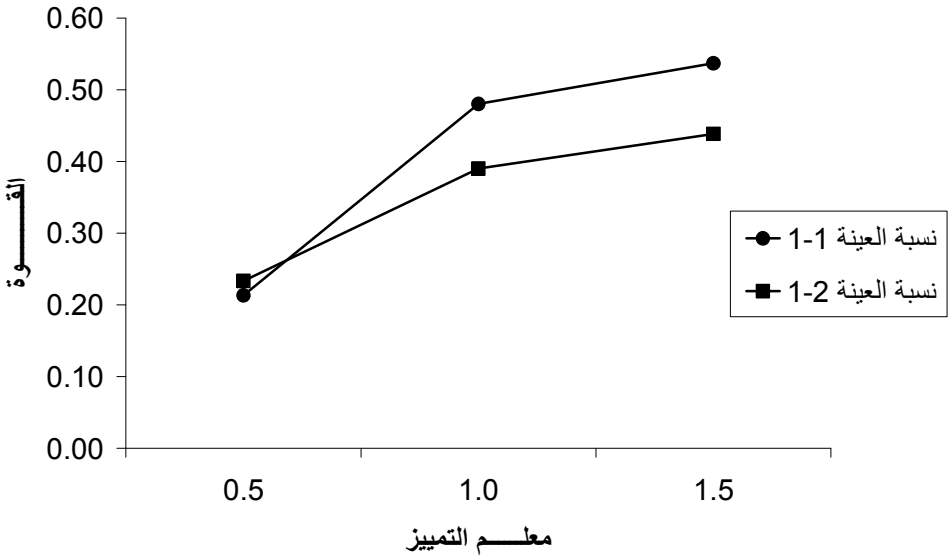


الشكل ٢٢. تأثير التفاعل بين حجم العينة وتوزيع القدرة على القوة لطريقة SIBTEST عند وجود تحيز دالة فقرة تفاضلية مقدارها ٠,٢٥

All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

المنارة للاستشارات
www.manaraa.com

يتضح من الرسم البياني أن الفروق في القوة الناتجة عن اختلاف توزيع القدرة بين المجموعتين المرجعية والمستهدفة ازدادت مع ازدياد حجم العينة، كما أن هذه الفروق كانت لصالح الحالة التي كانت فيها اختلاف في متوسط توزيع القدرة عندما كان حجم العينة صغيرا (٦٠٠) ولصالح الحالة التي كانت فيها متوسط توزيع القدرة متساويا عندما كان حجم العينة متوسطا (١٢٠٠) وعاليا (٢٤٠٠)..



الشكل ٢٣. تأثير التفاعل بين معلم التمييز ونسبة العينة بين المجموعة التجريبية والمجموعة المستهدفة على القوة لطريقة SIBTEST عند وجود دالة فقرة تفاضلية مقدارها ٠,٢٥

يتضح من الرسم البياني أن الفروق في القوة الناتجة عن اختلاف توزيع القدرة بين المجموعتين المرجعية والمستهدفة ازدادت مع ازدياد معلم التمييز، كما أن هذه الفروق كانت لصالح الحالة التي كانت فيها نسبة العينة ٢—١ عندما كان معلم التمييز منخفضا، ولصالح الحالة التي كانت فيها النسبة ١–١ عندما كان معلم التمييز متوسطا ومرتفعا .

All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

المنارة للاستشارات
www.manaraa.com

٢- قوة الاختبار عند وجود دالة فقرة تفاضلية مقدارها ٠,٥

أ- طريقة مانتل هينزل

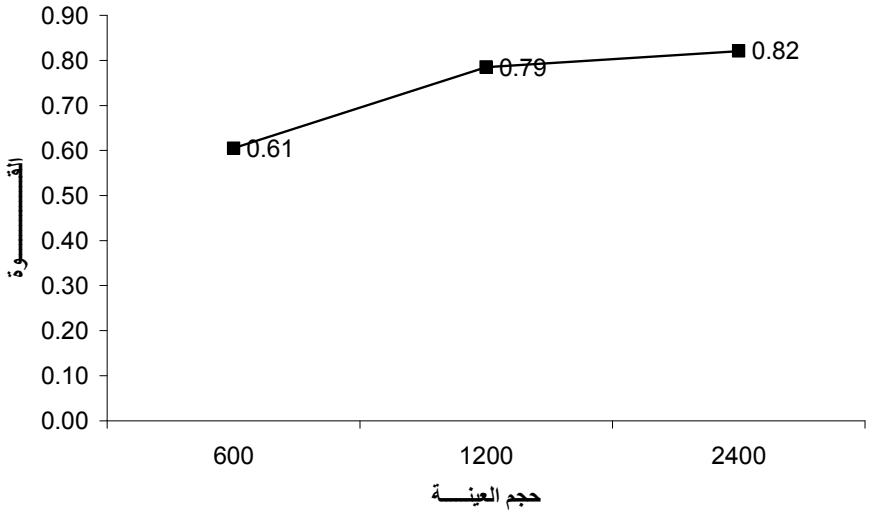
الجدول ١٧. نتائج تحليل الانحدار اللوجستي لتأثير متغيرات توزيع القدرة وحجم العينة ونسبتها ومعلم التمييز والتفاعل بينها على القوة لطريقة مانتل هينزل عند وجود دالة فقرة تفاضلية منتظمة مقدارها ٠,٥

المتغير	B	wald	درجات الحرية	القيمة الاحتمالية
توزيع القدرة	٠,٤٧٤-	١,٥٢٥	١	٠,٢١٧
حجم العينة	١,٩٩٤	٥١,٠٤١	١	٠,٠٠٠
معلم التمييز	١,٦٧٤	٤٠,٩٢٤	١	٠,٠٠٠
نسبة العينة	٠,١٨-	٠,٠٠٢	١	٠,٩٦٢
حجم العينة * معلم التمييز	٠,٤٦٢-	٤٤,١٩١	١	٠,٠٠٠
حجم العينة * نسبة العينة	٠,٠٢٨	٠,٠٦٤	١	٠,٨٠١
حجم العينة*توزيع القدرة	٠,٠٨٧-	٠,٦٠٢	١	٠,٤٣٨
معلم التمييز*توزيع القدرة	٠,٠٠٨	٠,٠٠٦	١	٠,٩٣٧
معلم التمييز*نسبة العينة	٠,١٥٥-	٢,١٦٣	١	٠,١٤١
توزيع القدرة * نسبة العينة	٠,١١٤	٠,٤٥٥	١	٠,٥٠٠
الثابت	٣,١٦٨-	١٧,٥٥١	١	٠,٠٠٠

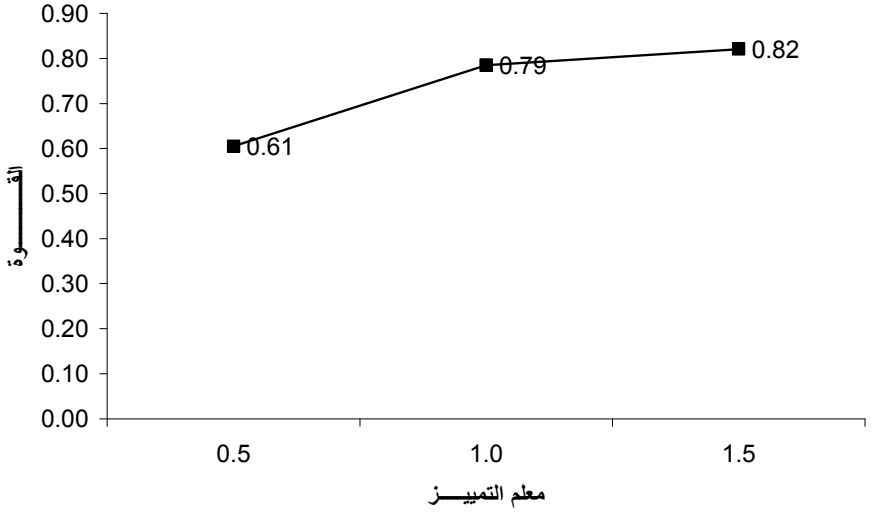
أظهرت النتائج وجود تأثير ذي دلالة إحصائية لمتغير حجم العينة (Wald=51.014 P. =0.000) ومعلم التمييز (Wald = 40,924 P. =0.000) على القوة، حيث كانت القوة تزداد بزيادة حجم العينة ومعلم التمييز كما يظهر في الرسوم البيانية التالية :

All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

المنارة للاستشارات
www.manaraa.com



الشكل ٢٤. تأثير حجم العينة على القوة لطريقة مانتل هينزل عند وجود دالة فقرة تفاضلية مقدارها ٠,٥

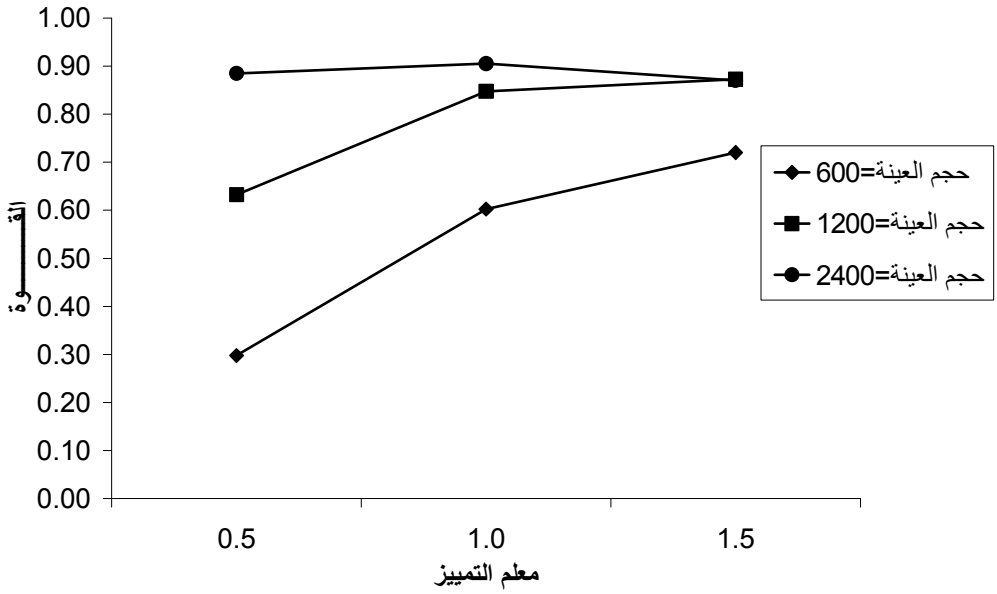


الشكل ٢٥. تأثير معلم التمييز على القوة لطريقة مانتل هينزل عند وجود دالة فقرة تفاضلية مقدارها ٠,٥

كما أظهـــــرت النتائج وجود تأثير للتفاعل بين حجم العينة ومعلم التمييز (Wald =44191, P. =0.000) على القوة، والرسم البياني التالي توضح تأثير التفاعل على القوة

All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

المنارة للاستشارات
www.manaraa.com



الشكل ٢٦. تأثير التفاعل بين معلم التمييز وحجم العينة على القوة لطريقة مانتل هينزل عند وجود دالة فقرة تفاضلية مقدارها ٠,٥

يتضح من الرسم البياني أن الفروق في القوة الناتجة عن اختلاف حجم العينة قلت مع ازدياد قيمة معلم التمييز.

All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

المنارة للاستشارات
www.manaraa.com

ب- طريقة الانحـــدار اللوجســـــتي

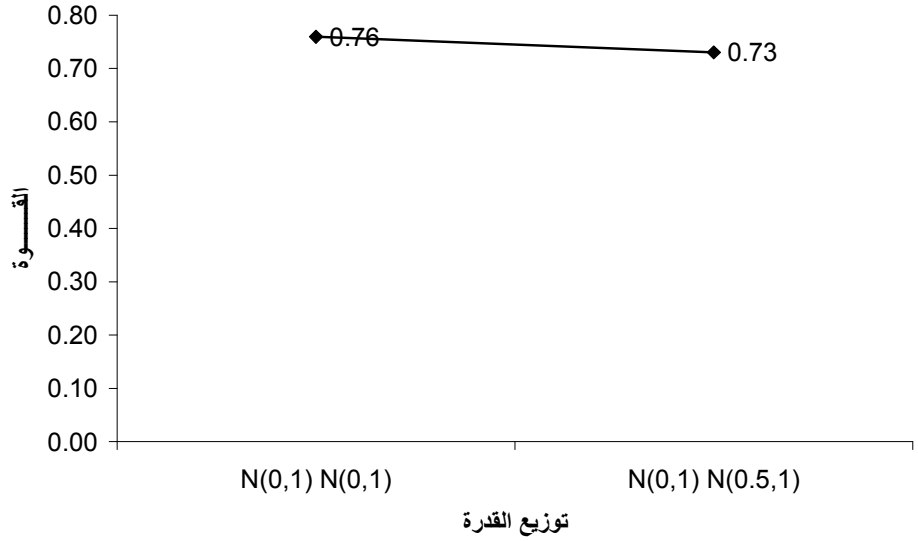
الجدول ١٨. نتائج تحليل الانحدار اللوجستي لتأثير متغيرات توزيع القدرة وحجم العينة ونسبتها ومعلم التمييز والتفاعل بينها على القوة لطريقة الانحدار اللوجستي عند وجود دالة فقرة تفاضلية منتظمة مقدارها ٠,٥

المتغير	B	wald	درجات الحرية	القيمة الاحتمالية
توزيع القدرة	٢,٣٦٩-	٢٨,٢٤٨	١	٠,٠٠٠
حجم العينة	٠,٦١٤	٣,٩٩٧	١	٠,٠٤٦
معلم التمييز	١,١٨٩	١٣,٩٥٩	١	٠,٠٠٠
نسبة العينة	١,٢٨١-	٨,٥٥٢	١	٠,٠٠٣
حجم العينة * معلم التمييز	٠,٠٣٣-	٠,١٥٥	١	٠,٦٩٤
حجم العينة * نسبة العينة	٠,٤٤٧	١٢,٩١١	١	٠,٠٠٠
حجم العينة*توزيع القدرة	٠,٠١١-	٠,٠٠٨	١	٠,٩٣٠
معلم التمييز*توزيع القدرة	٠,٥١٤	١٦,٨٦٨	١	٠,٠٠٠
معلم التمييز*نسبة العينة	٠,٤٣٦-	١٢,٠٧٣	١	٠,٠٠١
توزيع القدرة * نسبة العينة	٠,٨٦٧	٢٢,٨٣٨	١	٠,٠٠٠
الثابت	٠,٠١٣	٠,٠٠٠	١	٠,٩٨٨

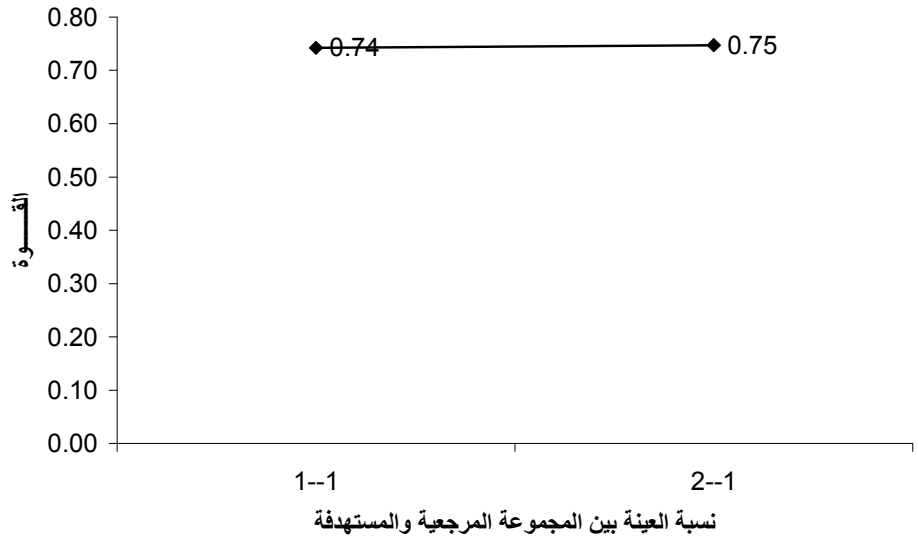
أظهـــــرت النتـــائج وجـود تأثيـــر ذي دلالة إحصـــائية لمتغيـــر توزيـع القـدرة (Wald = 28.248 P. =0.000) ومتغير نسبة العينة (Wald = 8.552 P. =0.003)، وحجم العينة (Wald = 3.997 P. =0.046) ومعلم التمييز (Wald = 13.959 P. =0.000) علـى القوة حيث قلت القوة مع وجود فرق في توزيع القدرة وعندما كانت نسبة العينة ٢—١ ، كمـا ازدادت القوة مع ازدياد كل من حجم العينة ومعلم التمييز، وتوضح الرسوم البيانية التالية تأثير هذه المتغيرات على القوة :

All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

المنارة للاستشارات
www.manaraa.com



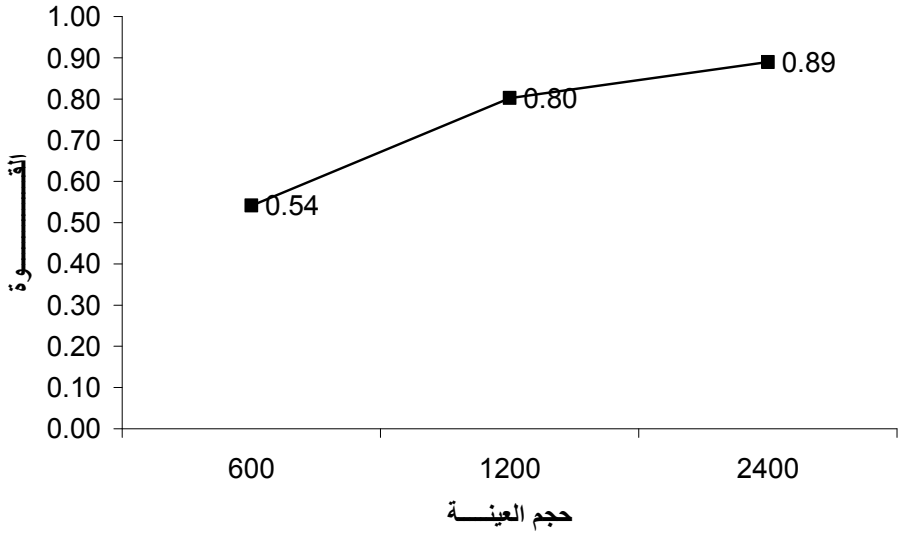
الشكل ٢٧. تأثير الاختلاف في توزيع القدرة بين المجموعتين المرجعية والمستهدفة على القوة لطريقة الانحدار اللوجستي عند وجود دالة فقرة تفاضلية مقدارها ٠,٥



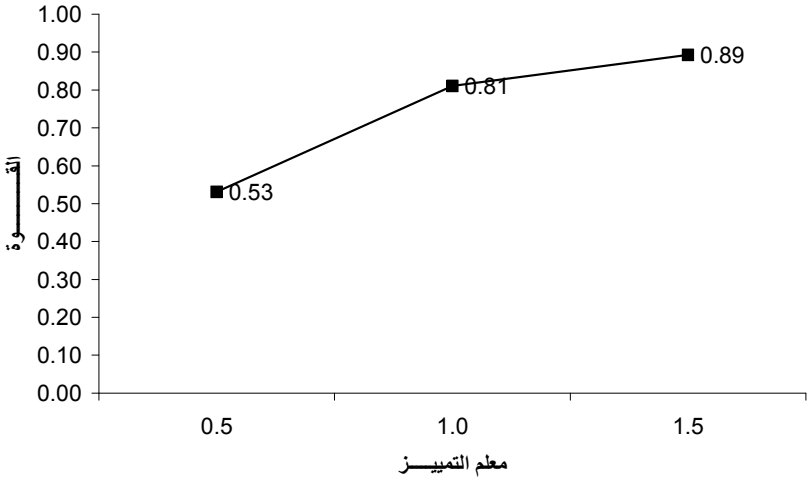
الشكل ٢٨. تأثير الاختلاف في نسبة العينة بين المجموعتين المرجعية والمستهدفة على القوة لطريقة الانحدار اللوجستي عند وجود دالة فقرة تفاضلية مقدارها ٠,٥

All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

المنارة للاستشارات
www.manaraa.com



الشكل ٢٩. تأثير حجم العينة على القوة لطريقة الانحدار اللوجستي عند وجود دالة فقرة تفاضلية مقدارها ٠,٥

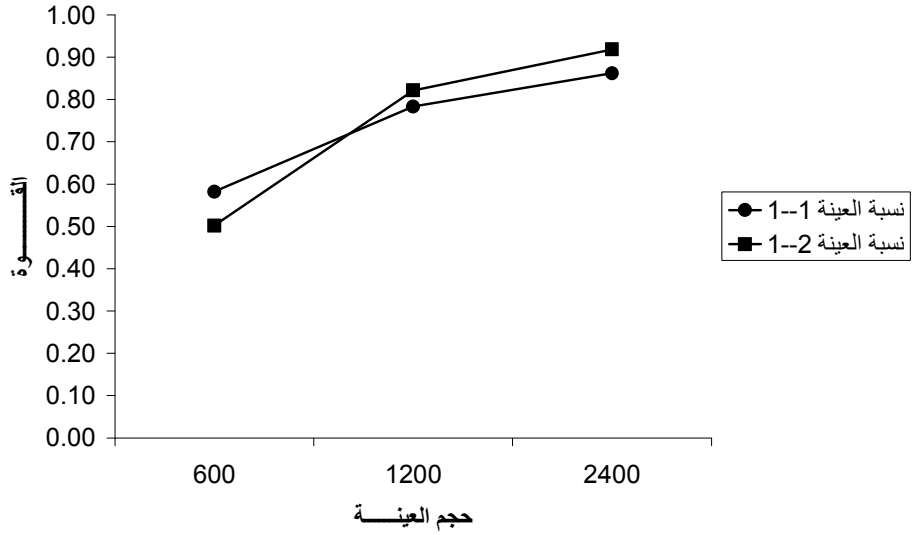


الشكل ٣٠. تأثير معلم التمييز على القوة لطريقة الانحدار اللوجستي عند وجود دالة فقرة تفاضلية مقدارها ٠,٥

كما أظهرت النتائج وجود تأثير للتفاعل بين حجم العينة ونسبتها بـــين المجموعـــة المرجعيـــة والمجموعة المستهدفة (Wald = 12.911 P. =0.000)، والتفاعل بين معلم التمييز وتوزيـــع القدرة (Wald = 16.868 P. =0.000)، والتفاعل بين معلم التمييز ونسبة العينـــة (Wald = 12.073 P. =0.001) ، والتفاعل بين توزيع القدرة ونـــسبة العينـــة (Wald = 22.838 P. =0.000)، على القـــــوة كما يتضح من الرسوم البيانية التالية :-

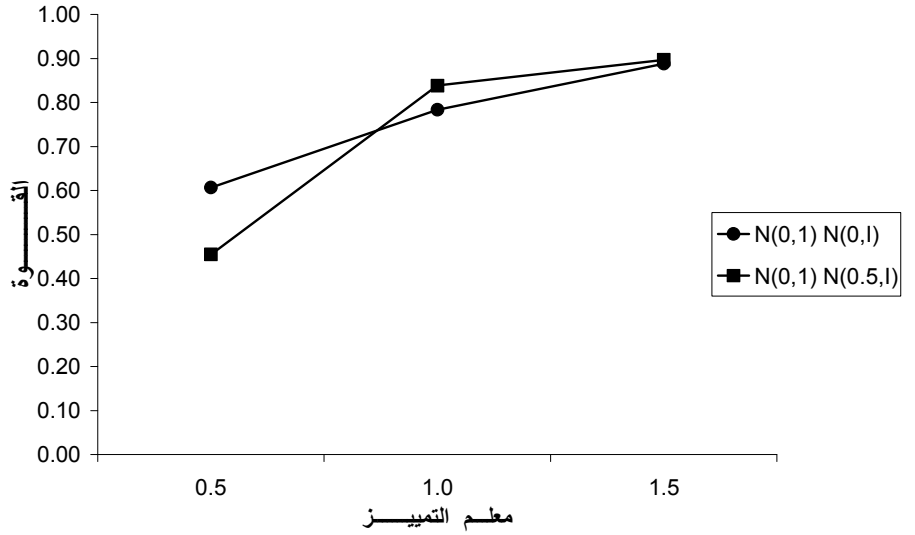
All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

المنارة للاستشارات
www.manaraa.com



الشكل ٣١. تأثير التفاعل بين حجم العينة ونسبتها على القوة لطريقة الانحدار اللوجستي عند وجود دالة فقرة تفاضلية مقدارها ٠,٥

يتضح من الرسم البياني أن الفروق الناتجة عن اختلاف نسبة العينة بين المجموعة المرجعية والمستهدفة كانت لصالح الحالة التي كانت فيها نسبة العينة ١—١ عندما كان حجم العينة صغيرا، ولصالح الحالة التي كانت فيها نسبة العينة ٢—١عندما كان حجم العينة متوسطا وكبيرا .



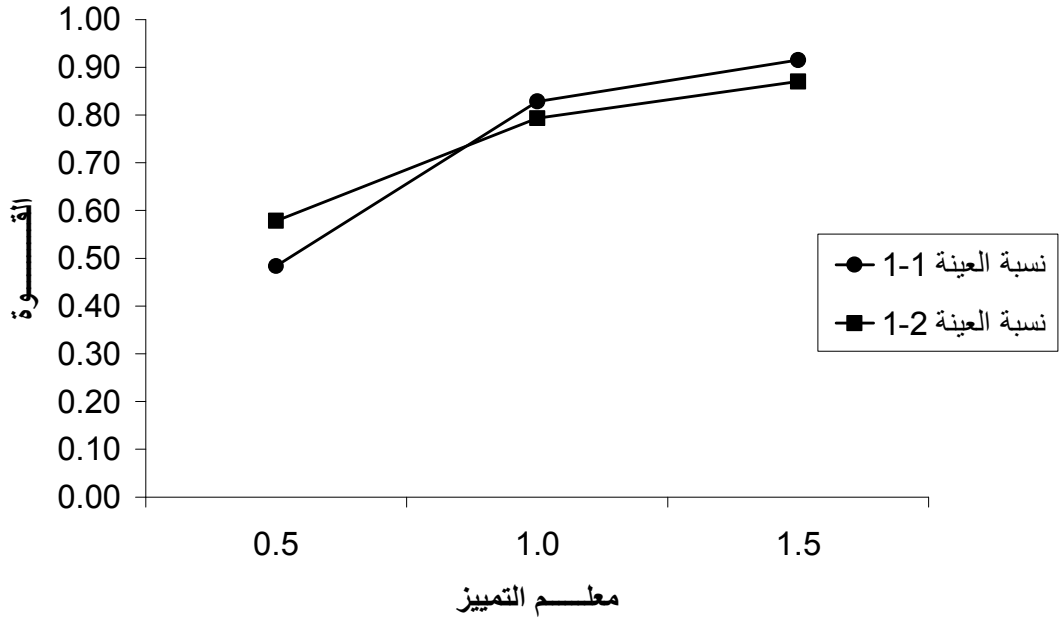
الشكل ٣٢. تأثير التفاعل بين معلم التمييز وتوزيع القدرة على القوة لطريقة الانحدار اللوجستي عند وجود دالة فقرة تفاضلية مقدارها ٠,٥

All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

المنارة للاستشارات
www.manaraa.com

يتضح من الرسم البياني أن قيمة القوة كانت أعلى للحالة التي كان فيها للمجموعتين نفس متوسط توزيع القدرة عندما كان معلم التمييز منخفضا، ولصالح الحالة التي كان فيها متوسط توزيع القدرة للمجموعتين مختلفا عندما كان معلم التمييز متوسطا ومرتفعا .

شكل رقم (٣٣)

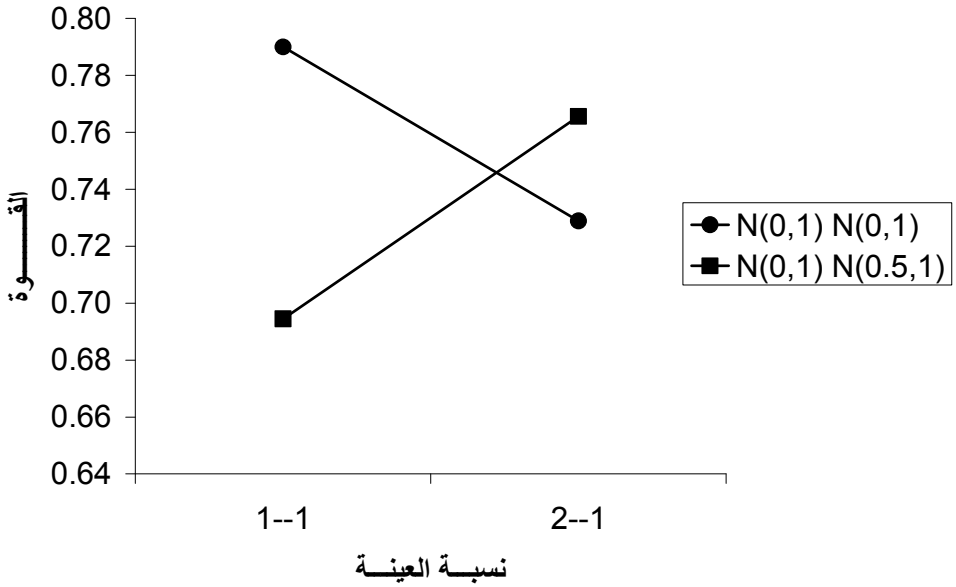


الشكل ٣٣. تأثير التفاعل بين معلم التمييز ونسبة العينة على القوة لطريقة الانحدار اللوجستي عند وجود دالة فقرة تفاضلية مقدارها ٠,٥

يتضح من الرسم البياني السابق أن قيمة القوة كانت أعلى للحالة التي كان فيها نسبة العينة ٢-١ عندما كان معلم التمييز منخفضا، ولصالح الحالة التي كان فيها نسبة العينة ١-١ عندما كان معلم التمييز متوسطا ومرتفعا .

All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

المنارة للاستشارات
www.manaraa.com



الشكل ٣٤. تأثير التفاعل بين توزيع القدرة ونسبة العينة بين المجموعة المرجعية والمجموعة المستهدفة على القوة لطريقة الانحدار اللوجستي عند وجود دالة فقرة تفاضلية مقدارها ٠,٥

يتضح من الرسم البياني أن قيمة القوة كانت أعلى للحالة التي كانت فيها نسبة العينة ١–١ عند تساوي متوسط توزيع القدرة للمجموعتين المرجعية والمستهدفة ، وأعلى للحالة التي كانت فيها نسبة العينة ٢—١ عند اختلاف متوسط توزيع القدرة بين المجموعتين.

All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

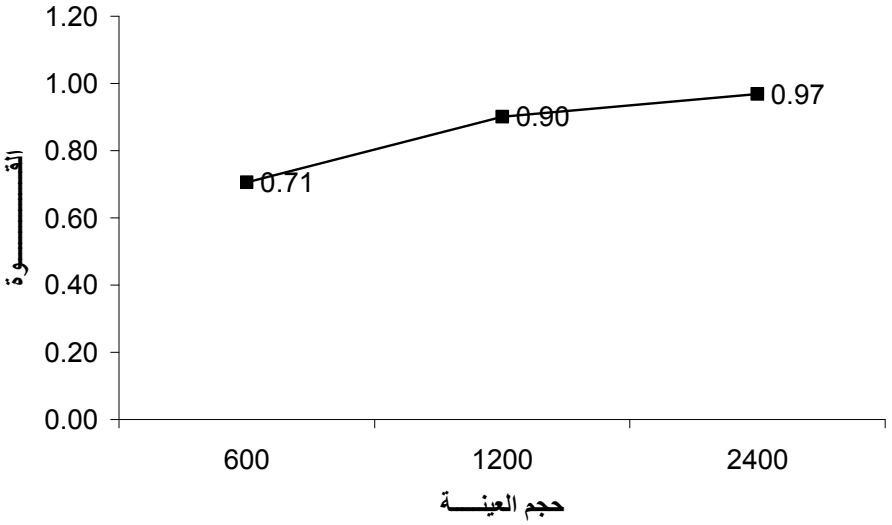
المنارة للاستشارات
www.manaraa.com

ج- طـــريقـــة اختبار تحيز الفقرة المتزامن SIBTEST

الجدول ١٩. نتائج تحليل الانحدار اللوجستي لتأثير متغيرات توزيع القدرة وحجم العينة ونسبتها ومعلم التمييز والتفاعل بينها على القوة لطريقة SIBTEST عند وجود دالة فقرة تفاضلية منتظمة مقدارها ٠,٥

المتغير	B	wald	درجات الحرية	القيمة الاحتمالية
توزيع القدرة	٠,١٣٨-	٠,١٣٠	١	٠,٧١٩
حجم العينة	١,٩٢١	٤٩,٨٨٢	١	٠,٠٠٠
معلم التمييز	١,٥٢٠	٣٦,٣٢٥	١	٠,٠٠٠
نسبة العينة	٠,٠٨١-	٠,٠٥٠	١	٠,٨٢٣
حجم العينة * معلم التمييز	٠,٣٦٧-	٢٩,٨٠٩	١	٠,٠٠٠
حجم العينة * نسبة العينة	٠,٠٧٨	٠,٥٢١	١	٠,٤٧٠
حجم العينة*توزيع القدرة	٠,٢١٣-	٣,٧٢٤	١	٠,٥٤
معلم التمييز*توزيع القدرة	٠,٠٩٢-	٠,٧٧٤	١	٠,٣٧٩
معلم التمييز*نسبة العينة	٠,١٠٣-	١,٠١١	١	٠,٣١٥
توزيع القدرة * نسبة العينة	٠,٠١١	٠,٠٠٥	١	٠,٩٤٦
الثابت	٢,٩٩٤-	١٦,٤١٧	١	٠,٠٠٠

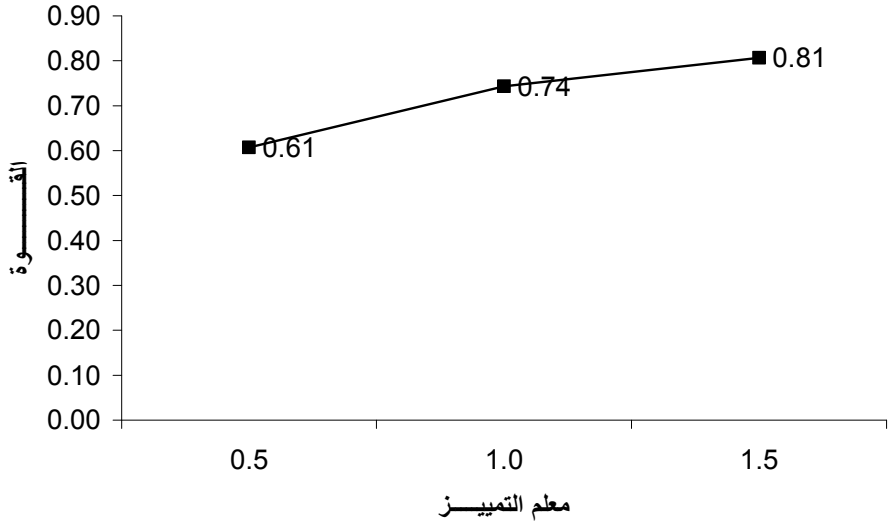
أظهـــــرت النتـــائج وجـــود تأثيـــر ذي دلالة إحصـــائية لمتغيـــر حجم العينـــة (Wald = 49.882 P. =0.000) ومتغير معلم التمييز (Wald = 36.325 P. =0.000) ، حيث أن القوة ازدادت مع ازدياد كل من حجم العينة ومعلم التمييز كما يتـضح مـن الرسـوم البيانية التالية :



الشكل ٣٥. تأثير حجم العينة على القوة لطريقة SIBTEST عند وجود دالة فقرة تفاضلية مقدارها ٠,٥

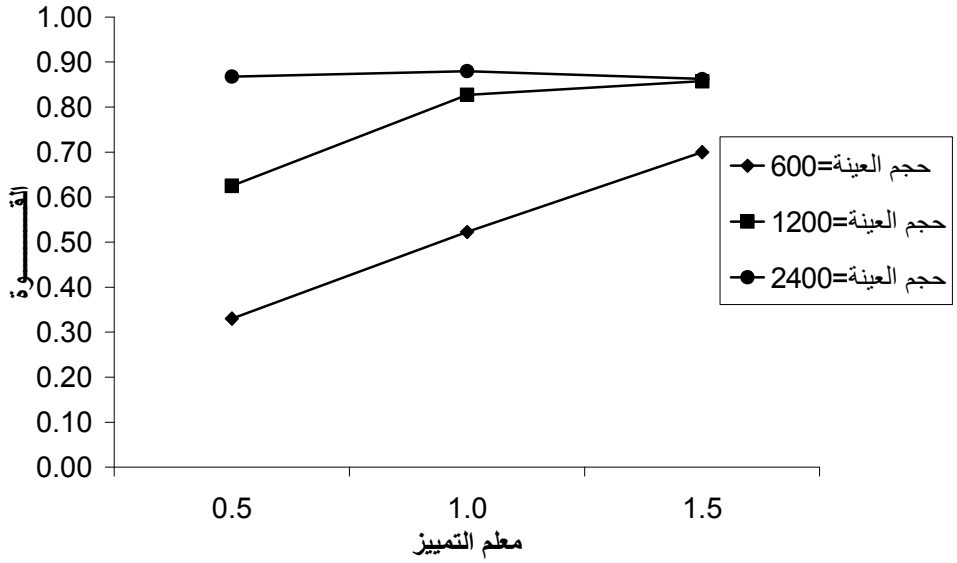
All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

المنارة للاستشارات
www.manaraa.com



الشكل ٣٦. تأثير معلم التمييز على القوة لطريقة SIBTEST عند وجود دالة فقرة تفاضلية مقدارها ٠,٥

كما أظهرت النتائج وجود تأثير للتفاعل بين حجم العينة ومعلم التمييز يتضح من الرسم البياني التالي :-



الشكل ٣٧. تأثير التفاعل بين معلم التمييز وحجم العينة على القوة لطريقة SIBTEST عند وجود دالة فقرة تفاضلية مقدارها ٠,٥

يتضح من الرسم البياني أن الفروق في القوة الناتجة عن اختلاف حجم العينة قلت مع ازدياد معلم التمييز

All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

المنارة للاستشارات
www.manaraa.com

٣-قوة الاختبار الإحصائي في حالة وجود تحيز غير منتظم

عند الحديث عن تأثير متغيرات الدراسة في قوة الاختبار الإحصائي سيتم عرض النتائج الخاصة بطريقة الانحدار اللوجستي فقط ، لأنها الطريقة الوحيدة التي تصلح للكشف عن دالة الفقرة التفاضلية ، وفيما يلي عرض لهذه النتائج

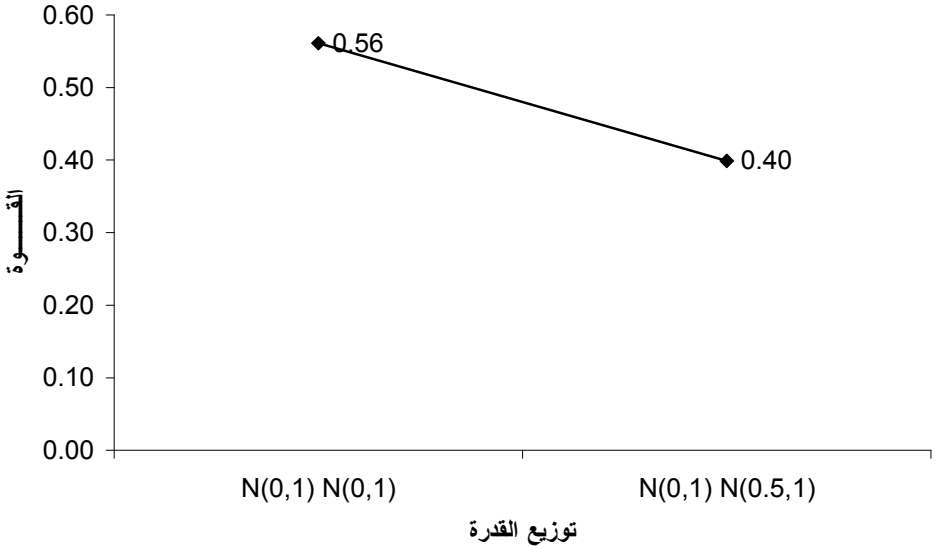
الجدول ٢٠. نتائج تحليل الانحدار اللوجستي لتأثير متغيرات توزيع القدرة والحجم العينة ونسبتها ومعلم التمييز والتفاعل بينها على القوة لطريقة الانحدار اللوجستي عند وجود دالة فقرة تفاضلية غير منتظمة.

المتغير	B	wald	درجات الحرية	القيمة الاحتمالية
توزيع القدرة	١,٢٠٣	٩,٠١٣	١	٠,٠٠٣
حجم العينة	٠,٥٥٨	٥,٢٥٧	١	٠,٠٢٢
معلم التمييز	٠,٦١١	٦,٣٨٤	١	٠,٠١٢
نسبة العينة	-٠,٣٠٩	٠,٧١٠	١	٠,٣٩٩
حجم العينة * معلم التمييز	٠,٤٠٥	٤٠,٤٨٠	١	٠,٠٠٠
حجم العينة * نسبة العينة	٠,٣٣٠	١٠,٧٧٧	١	٠,٠٠١
حجم العينة*توزيع القدرة	-٠,٥٥٦	٢٩,٣٢٩	١	٠,٠٠٠
معلم التمييز*توزيع القدرة	-٠,٣٢٤	١٠,٠٨٨	١	٠,٠٠١
معلم التمييز*نسبة العينة	٠,٠٣١	٠,٠٩٨	١	٠,٧٥٤
توزيع القدرة * نسبة العينة	٠,٢٢٢	٢,٠١٤	١	٠,١٥٦
الثابت	-٣,٢٥٧	١٧,٥٤٤	١	٠,٠٠٠

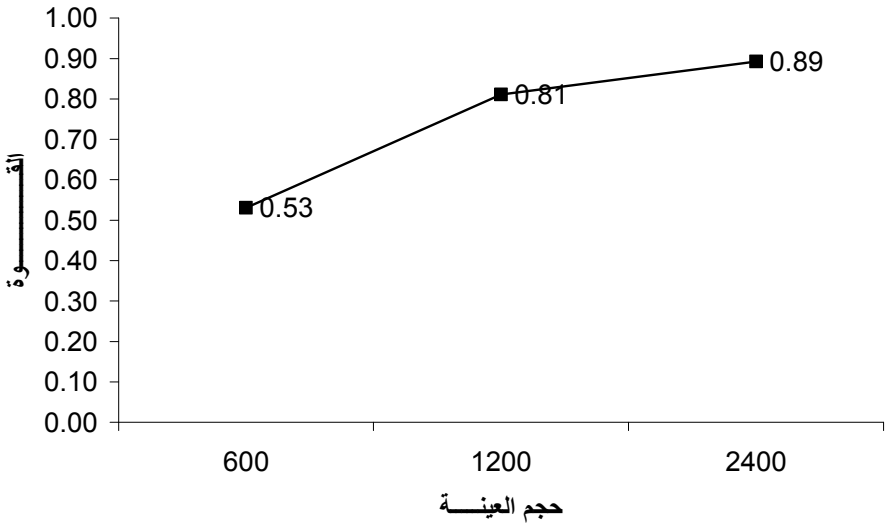
أظهرت النتائج وجود تأثير لمتغيرات توزيع القدرة (Wald =9.013, P. =0.003) وحجم العينة (Wald = 5.257 P. =0.022) ومعلم التمييز (Wald = 6.384, P. =0.012) على القوة حيث قلت القوة عند وجود اختلاف في القدرة بين المجموعة التجريبية والمجموعة الضابطة، كما أن القوة ازدادت مع زيادة كل من حجم العينة ومعلم التمييز كما يتضح من الرسوم البيانية التالية :-

All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

المنارة للاستشارات
www.manaraa.com



الشكل ٣٨. تأثير توزيع القدرة على القوة لطريقة الانحدار اللوجستي عند وجود دالة فقرة تفاضلية غير منتظمة

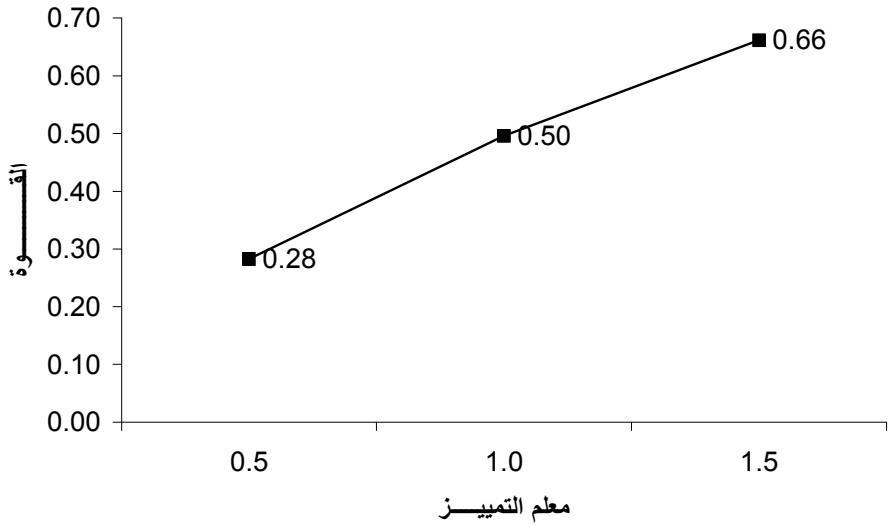


الشكل ٣٩. تأثير حجم العينة على القوة لطريقة الانحدار اللوجستي عند وجود دالة فقرة تفاضلية غير منتظمة

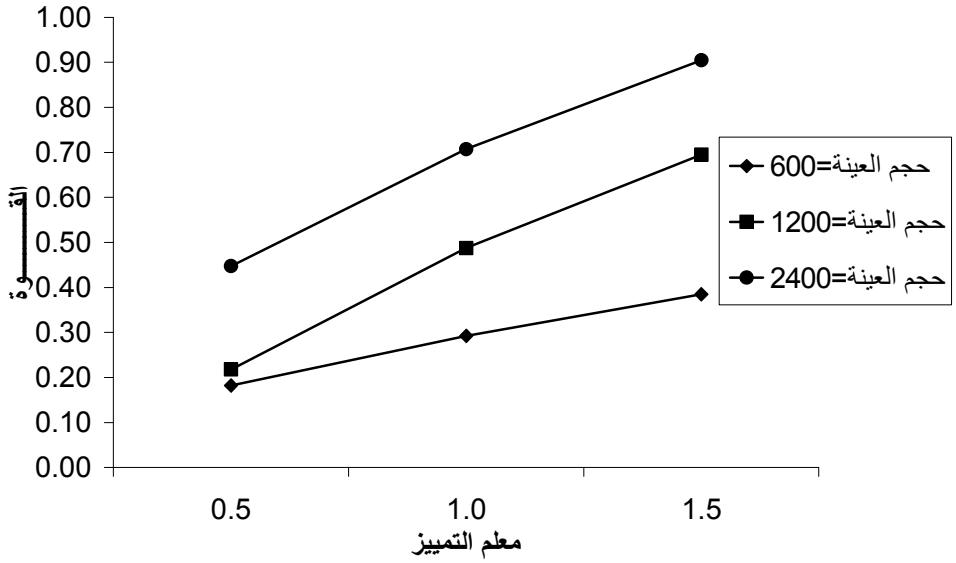
الشكل ٤٠. تأثير معلم التمييز على القوة لطريقة الانحدار اللوجستي عند وجود دالة فقرة تفاضلية غير منتظمة

All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

المنارة للاستشارات
www.manaraa.com



كما أظهرت النتائج وجود تأثير للتفاعل بين حجم العينة ومعلم التمييـــز (Wald = 40.48 P. =0.000)، والتفاعل بين حجم العينة ونسبتها (Wald = 10.77 P. =0.001)، والتفاعل بـــين حجم العينة وتوزيع القدرة (Wald = 29.32 P. =0.001) ، والتفاعل بـــين معلـــم التمييـــز وتوزيع القدرة (Wald = 10.088 P. =0.000)، على القـــــوة كما يتـــضح مـــن الرســـوم البيانية التالية :-

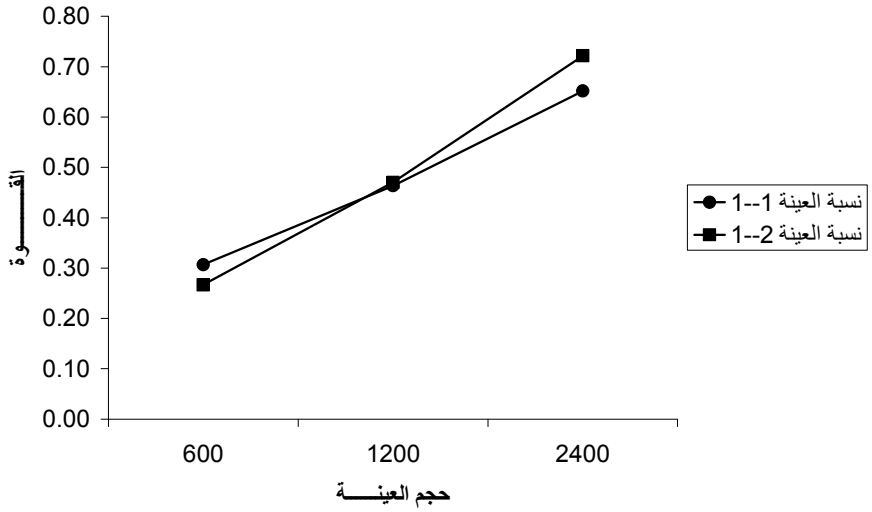


الشكل ٤١. تأثير التفاعل بين حجم العينة ومعلم التمييز على القوة لطريقة الانحدار اللوجستي عند وجود دالة فقرة تفاضلية غير منتظمة

يتضح من الرسم البياني أن الفروق الناتجة عن ازدياد حجم العينة ازدادت مع ازدياد معلم التمييز .

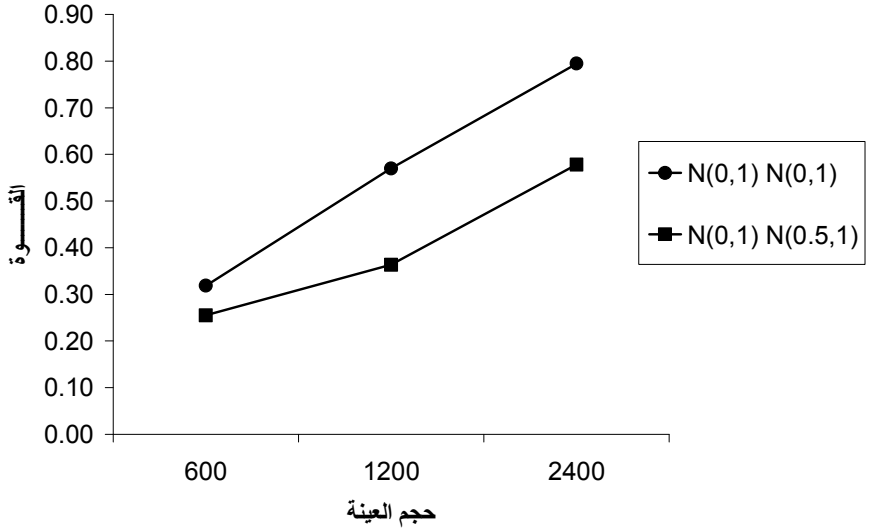
All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

المنارة للاستشارات
www.manaraa.com



الشكل ٤٢. تأثير التفاعل بين حجم العينة ونسبتها بين المجموعة المرجعية والمجموعة المستهدفة على القوة لطريقة الانحدار اللوجستي عند وجود دالة فقرة تفاضلية غير منتظمة

يتضح من الرسم البياني أن الفروق الناتجة عن اختلاف نسبة العينة بين المجموعة المرجعية والمستهدفة كانت لصالح الحالة التي كانت فيها نسبة العينة ١—١ عندما كان حجم العينة صغيرا، ولصالح الحالة التي كانت فيها نسبة العينة ٢—١عندما كان حجم العينة متوسطا وكبيرا.

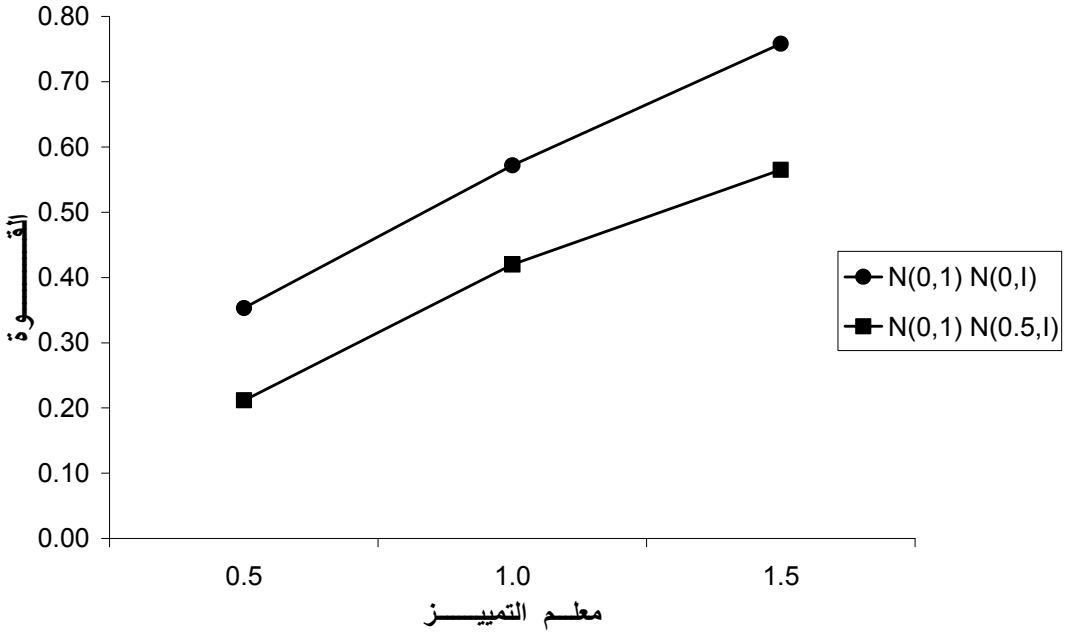


الشكل ٤٣. تأثير التفاعل بين حجم العينة وتوزيع القدرة على القوة لطريقة الانحدار اللوجستي عند وجود دالة فقرة تفاضلية غير منتظمة

All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

المنارة للاستشارات
www.manaraa.com

يتضح من الرسم البياني أن الفروق الناتجة عن الاختلاف في توزيع القدرة كانت أكبر عندما كانت حجم العينة كبيرا ومتوسطا منها عندما كان حجم العينة صغيرا .



الشكل ٤٤. تأثير التفاعل بين معلم التمييز وتوزيع القدرة للمجموعة المرجعية والمجموعة المستهدفة على القوة لطريقة الانحدار اللوجستي عند وجود دالة فقرة تفاضلية غير منتظمة

يتضح من الرسم البياني أن الفروق الناتجة عن اختلاف متوسط توزيع القدرة كانت أكبر عندما كان معلم التمييز مرتفعا منها عندما كان معلم التمييز متوسطا ومنخفضا .

All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

المنارة للاستشارات
www.manaraa.com

الفصل الخامس
منـــــــاقشة النتـــــــائج

All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

المنارة للاستشارات
www.manaraa.com

مناقشة النتائج

هدفت هذه الطريقة إلى التحقق من فاعلية أربعا من الطرائق الإحصائية المستخدمة في الكشف عن دالة الفقرة التفاضلية، اثنتين من هذه الطـــرائق تستند إلى النظـــرية الكلاسيكية في القيـــاس (طريقة مانتل هينزل ، وطريقة الانحدار اللوجستي)، واثنتين تستندان إلى النظرية الحديثة في القياس طريقة نسبة الأرجحية (Likelihood Ratio) وطريقة اختبار تحيز الفقرة المتزامن (SIBTEST)، كما هدفت إلى دراسة تأثير المتغيرات التالية (الاختلاف في القدرة بين المجموعتين المرجعية والمستهدفة وحجم العينة ونسبتها بين المجموعتين المرجعية والمستهدفة ، ومعلم التمييز ، ومقدار دالة الفقرة التفاضلية ونوعها، على الخطأ من النوع الأول والقوة الإحصائية للطرائق الأربع وفيما يلي عرض لمناقشة النتائج التي أظهرتها الدراسة :

أولا : مناقشة النتائج المتعلقة بفاعلية الطرائق الإحصائية (مانتل هينزل، والانحـــدار اللوجستي، واختبار التحيز المتزامن، ونسبة الأرجحية) في الكشف عن دالة الفقرة التفاضلية باستخدام الخطأ من النوع الأول، وتأثير المتغيرات المستقلة على معدل الخطأ من النوع الأول.

أظهرت نتائج الدراسة أن معدل الخطأ من النوع الأول كان قريبا من القيمة ٠,٠٥ لثلاث طرائق هي طريقة الانحدار اللوجستي وطريقة مانتل هينزل وطريقة اختبار التحيـــز المتـــزامن كمـــا أظهرت النتائج وجود تضخم واضح لمعدل الخطأ من النوع الأول لطريقة نسبة الأرجحية.
ولقد كانت أفضل الطرائق فاعلية في الكشف عن دالة الفقرة التفاضلية وفقا لمعيار الخطـــأ مـــن النوع الأول طريقة الانحدار اللوجستي التي تراوح معدل الخطأ مـــن النـــوع الأول لهـــا بـــين (٠,٠٠-٠,٠٨)، ففي حالة تساوي توزيع القدرة للمجموعتين المرجعية والمستهدفة كان معـــدل الخطأ من النوع الأول ≥ ٠,٠٥ في جميع الحالات باستثناء خليتين، أما في حالة عدم تـــساوي القدرة كان معدل الخطأ من النوع الأول <٠,٠٥ في سبع خلايا . كانت فـــي معظمهـــا ذات حجم الكبير وذات معلم تمييز متوسط أومرتفع .
وقد جاءت طريقة مانتل هينزل التي تراوح معدل الخطأ من النوع الأول لهـــا بـــين (٠,٠٠-٠,٠٩) في المرتبة الثانية ففي حالة تساوي توزيع القدرة للمجموعتين المرجعية والمستهدفة كان معدل الخطأ من النوع الأول ≥ ٠,٠٥ في جميع خلايا التصميم باستثناء ثلاث خلايا كان فيهـــا حجم الخلايا كبيرا ومعلم التمييز متوسطا ومرتفعا، أما عندما كان التوزيعان مختلفـــان، حيـــث

All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

المنارة للاستشارات
www.manaraa.com

متوسط القدرة للمجموعة المستهدفة أقل من المجموعة المرجعية، فقد كان معدل الخطأ من النوع الأول > ٠,٠٥ في تسع خلايا، ولم يكن هنالك نمط عام يدل على وجود تأثير لحجم العينة أو معلم التميز .

وقد جاءت طريقة اختبار التحيز المتزامن التي تراوحت قيم معدل الخطأ من النوع الأول لها بين (٠,٠٠٠- ٠,١٥) في المرتبة الثالثة، فعند تساوي القدرة للمجموعتين المرجعية والمستهدفة كان عدد الخلايا التي فيها معدل الخطأ من النوع الأول > ٠,٠٥ سبع خلايا، كانت في معظمها ذات حجم العينة كبير ومعلم تمييز مرتفع، وعندما كان مستوى قدرة المجموعة المستهدفة أقل من المجموعة المرجعية، فقد كانت معدلات الخطأ من النوع الأول > من ٠,٠٥ في عشر خلايا، تركزت بشكل أساسي في الخلايا التي كانت النسبة فيها بين المجموعة المرجعية والمستهدفة (١:١)، ولم يكن هنالك نمط عام يدل على تأثير لحجم العينة أو معلم التمييز.

وبشكل عام يمكننا القول أن ثلاث طرائق وهي (مانتل هينزل ، والانحدار اللوجستي، واختبار التحيز المتزامن) تمتعت بمعدلات مقبولة من الخطأ من النوع الأول، وقد كانت أفضل معدلات للخطأ من النوع الأول لطريقتي مانتل هينزل والانحدار اللوجستي في الخلايا التي كان فيها متوسط توزيع القدرة متساويا بين المجموعتين، وحجم العينة متوسطا أو صغيرا ومعلم التمييز ≥ ١، أما طريقة اختبار التحيز المتزامن، فقد كانت أفضل معدلات للخطأ من النوع الأول في الخلايا التي كان فيها متوسط توزيع القدرة متساويا بين المجموعتين، وحجم العينة متوسطا.

أما فيما يتعلق بتأثير متغيرات الدراسة المستقلة (توزيع القدرة، وحجم العينة ونسبتها بين المجموعتين المرجعية والمستهدفة، ومعلم التمييز) على الخطأ من النوع الأول فقد أظهرت نتائج الدراسة عدم وجود تأثير ذو دلالة إحصائية ($\alpha \geq$ ٠,٠٥) لأي من المتغيرات المستقلة أو التفاعلات الثنائية بينها على الخطأ من النوع الأول لطريقتي مانتل هينزل، وكذالك عدم وجود تأثير وجود تأثير لأي من متغيرات الدراسة المستقلة على الخطأ من النوع الأول لطريقة اختبار التحيز المتزامن، لكن كان هنالك تأثير للتفاعل بين معلم التمييز ونسبة حجم العينة بين المجموعة المرجعية والمستهدفة فقط على الخطأ من النوع الأول.

كما أظهرت النتائج المتعلقة بطريقة الانحدار اللوجستي وجود تأثير ذو دلالة إحصائية ($\alpha \geq$ ٠,٠٥) لمتغير حجم العينة ومعلم التمييز والتفاعل بين حجم العينة ونسبتها على الخطأ من النوع الأول، حيث أن الخطأ من النوع الأول كان يزداد بزيادة حجم العينة ومعلم التمييز، كما أن نسبة الخطأ من النوع الأول كانت أكبر عندما كانت النسبة بين

All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

المنارة للاستشارات
www.manaraa.com

المجموعة المرجعية والمستهدفة (٢: ١) وحجم العينة صغيرا ومتوسطا، وأعلى للنسبة (١:١) عندما كان حجم العينة كبيرا .

ومما لا شك فيه أن عدم وجود تأثير للمتغيرات على الخطأ من النوع الأول لطريقتي مانتل و اختبار التحيز المتزامن يعتبر أمرا ايجابيا فهو يعني أن الطريقتين سوف تكونا فعالتين في الكشف عن الخطأ من النوع الأول في ظروف متعددة ومختلفة، أما طريقة الانحدار اللوجستي فعلى الرغم من ازدياد معدل الخطأ من النوع الأول مع ازدياد حجم العينة ومعلم التمييز إلا أنه بقي مقبولا وقريبا من القيمة ٠,٠٥ ، وعلى الرغم من ازدياد الخطأ من النوع الأول بشكل طفيف عند وجود فروق معتدلة في توزيع القدرة بين المجموعتين المرجعية والمستهدفة في الطرائق الأربعة، إلا أنه لم يكن هنالك تأثير دال إحصائيا لمتغير توزيع القدرة على الخطأ من النوع الأول لأي من الطرائق الأربعة، وعدم تأثر هذه الطرائق بالفروق المعتدلة في توزيع القدرة أمر هام فالفروق في الأداء بين المجموعتين يجب أن تعزى إلى الفقرة حتى تعد أداء تفاضليا وليس إلى فروق حقيقية في القدرة بين المجموعتين والتي ترتبط بمفهوم تأثير الفقرة (Item Impact) ، ويمكن أن يعزي عدم تأثر طريقة مانتل هينزل بالفروق المعتدلة في توزيع القدرة إلا أن الأفراد في هذه الطريقة يتم تقسيمهم إلى فئات وفقا للدرجة الكلية المشاهدة، حيث تتم مقارنة كل فئة في المجموعة المرجعية مع الفئة التي تناظرها في المجموعة المستهدفة، أما طريقة اختبار التحيز المتزامن فتعتمد بشكل كبير على عملية تصحيح انحدار الدرجة الحقيقية على الدرجة المشاهدة للحد من التضخم في الخطأ من النوع الأول ، أما طريقة نسبة الأرجحية فتعمل على تقدير معالم الفقرات في كل مجموعة استنادا إلى النظرية الحديثة في القياس التي من مميزاتها عدم تغير تقديرات معالم الفقرات وفقا للعينة المستخدمة في تقدير المعالم عند تحقق افتراضات وشروط معينة .

وتتفق هذه النتائج مع دراسة روز وستاوت (Rousses & Stout, 1996) التي أشارت إلى أن معدل الخطأ من النوع الأول لطريقتي مانتل هينزل واختبار التحيز المتزامن كان جيدا مع ازدياد معدل الخطأ من النوع الأول بشكل طفيف مع زيادة حجم العينة وزيادة الفرق بين متوسط توزيع القدرة للمجموعتين المرجعية والمستهدفة، كما أن الدراسات التي أجريت على الفقرات ذات الاستجابة المتعددة مثل دراسة (Kristjansson. , Aylesworth , Mcdowell & Zumbo , 2005) أظهرت أن طريقتي الانحدار اللوجستي ومانتل هينزل المخصصة للكشف عن دالة الفقرة التفاضلية للفقرات متعددة الاستجابة كان لها معدلات جيدة للخطأ من النوع الأول وقريبة أيضا من القيمة ٠,٠٥ ، وكذلك دراسة (Tian ,1999) التي أجريت على

All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

المنارة للاستشارات
www.manaraa.com

فقرات متعددة الاستجابة والتي أشارت إلى تمتع طريقتي مانتل هينزل وطريقة اختبار التحيز المتزامن بمعدلات جيدة للخطأ من النوع الأول عندما تكون المجموعتين المرجعية والمستهدفة متساويتين في القدرة أو عند وجود فروق صغير في القدرة بين المجموعتين .

أما طريقة نسبة الأرجحية(Log Likelihood Ratio)فقد أظهرت النتائج وجود تضخم واضح في الخطأ من النوع الأول لهذه الطريقة حيث تراوح معدل الخطأ من النوع الأول بين (٠,١٢ – ٠,٢٩)، وهي معدلات مرتفعة وغير مقبولة نظر للاحتمال المرتفع للكشف عن وجود دالة الفقرة التفاضلية بشكل خاطئ في الفقرات التي يتم التحق من وجود دالة الفقرة التفاضلية فيها باستخدام هذه الطريقة ، وتتفق نتائج هذه الدراسة مع دراسة (Cohen, Kim & Wallack ,1996) التي أشارت إلى تضخم نسبة الخطأ من النوع الأول لطريقة نسبة الأرجحية (log likelihood ratio) عند استخدام النموذج اللوجستي ثلاثي المعالم مقارنة مع النموذج اللوجستي ثنائي المعالم (Ankenmann, Witt & Dumber, 1999) .

ثانيا : مناقشة النتائج المتعلقة بفاعلية الطرائق الإحصائية (مانتل هينزل ، والانحدار اللوجستي ، واختبار تحيز الفقرة المتزامن، ونسبة الأرجحية) في الكشف عن دالة الفقرة التفاضلية من حيث القوة ، وتأثير المتغيرات المستقلة على القوة الإحصائية لهذه الطرائق .

أولا القوة عند وجود دالة فقرة تفاضلية منتظمة مقدارها ٠,٢٥

على الرغم من أن القوة الإحصائية لطريقة نسبة الأرجحية كانت الأعلى بين الطرائق الأربع وبخاصة عند وجود دالة فقرة تفاضلية منتظمة مقدارها ٠,٥، أو عند وجود دالة فقرة تفاضلية غير منتظمة حيث كانت القوة الإحصائية لها ≤ ٠,٨٠ في معظم الحالات باستثناء الخلايا التي كان حجم العينة فيها صغيرا ومعلم التمييز منخفضا، إلا هذه النتائج لم يتم التعليق عليها في فصل النتائج، نظرا للمعدلات المرتفعة للخطأ من النوع الأول لهذه الطريقة والتي تجعل من مناقشة النتائج الخاصة بالقوة لهذه الطريقة غير ذات معنى، بسبب الاحتمال المرتفع للكشف عن دالة الفقرة التفاضلية بشكل خاطئ.

وقد أظهرت النتائج أن القوة الإحصائية للطرائق الإحصائية المتبقية عند وجود دالة تفاضلية منتظمة مقدارها ٠,٢٥ كانت ضعيفة بشكل عام، فقد كانت القوة الإحصائية للطرائق الثلاث في معظم الخلايا ≤ ٠,٦٠ ولم تتجاوز القوة الإحصائية لهذه الطرائق ٠,٨٠ إلا في عدد قليل من

All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

المنارة للاستشارات
www.manaraa.com

الخلايا، ففي طريقة الانحدار اللوجستي كانت القوة الإحصائية ≥ ٠,٨٠ في جميع الخلايا باستثناء أربع خلايا كان معلم التمييز لها مرتفعا وحجم العينة لثلاث منها كبيرا وواحدة متوسطا. كما أن القوة الإحصائية كانت بين ٠,٦٠-٠,٨٠ في سبع خلايا ، أما بقية الخلايا فقد كانت القوة الإحصائية فيها أقل من ٠,٦٠

أما طريقتي مانتل هينزل و اختبار التحيز المتزامن فقد كانت قيمة القوة الإحصائية ≥ ٠,٨٠ ، فقط في خلية واحدة، وهي الخلية التي كان فيها للمجموعتين نفس المتوسط لتوزيع القدرة بين المجموعتين المرجعية والمستهدفة، وحجم عينة كبير ومتساو بين المجموعتين، ومعلم تمييز مرتفعا، كما كانت قيمة القوة بين ٠,٦٠-٠,٨٠ في ثمان خلايا لطريقة مانتل هينزل وخمس خلايا لطريقة اختبار تحيز الفقرة المتزامن، أما بقية الخلايا فقد كانت قيمة القوة الإحصائية لها أقل من ٠,٦٠ .

أما فيما يتعلق بتأثير المتغيرات المستقلة للدراسة على التفاعل للطرائق الثلاث فقد أظهرت النتائج وجود تأثير لتوزيع القدرة على القوة لطريقتي الانحدار اللوجستي و اختبار تحيز الفقرة المتزامن، حيث أدى اختلاف متوسط توزيع القدرة للمجموعتين إلى ارتفاع القوة الإحصائية لطريقة الانحدار اللوجستي وانخفاضها لطريقة اختبار تحيز الفقرة المتزامن، كما أظهرت النتائج وجود تأثير لحجم العينة ومعلم التمييز على القوة لطريقتي مانتل هينزل و اختبار تحيز الفقرة المتزامن حيث كانت القوة تزداد بزيادة كل من حجم العينة ومعلم التمييز، وقد كان هنالك أثر لنسبة العينة على القوة لطريقة اختبار تحيز الفقرة المتزامن، حيث انخفضت القوة عند اختلاف نسبة العينة، وفيما يتعلق بالتفاعلات الثنائية بين المتغيرات المستقلة، فقد كان هنالك تأثير للتفاعل بين حجم العينة ومعلم التمييز على القوة للطرائق الثلاث حيث كانت الفروق الناتجة عن حجم العينة تزداد عندما يكون معلم التمييز متوسطا ومرتفعا، كما كان هنالك تأثير للتفاعل بين حجم العينة وتوزيع القدرة على القوة لطريقة اختبار تحيز الفقرة المتزامن فعند تساوي متوسط توزيع القدرة للمجموعتين المرجعية والمستهدفة كانت القوة أعلى لحجم العينة الكبير والمتوسط، وقد كان هنالك تأثير للتفاعل بين معلم التمييز توزيع القدرة لطريقة اختبار التحيز المتزامن حيث كانت قيمة القوة أعلى عندما اختلاف متوسط توزيع القدرة بين المجموعتين، وعندما كانت قيمة معلم التمييز مرتفعة، كما كان هنالك تأثير للتفاعل بين معلم التمييز ونسبة العينة على القوة لطريقة مانتل و اختبار تحيز الفقرة المتزامن حيث كانت أعلى قيمة للقوة عند ما كانت نسبة العينة (١:١) بين المجموعتين وعندما كانت قيمة معلم التمييز متوسطة أو مرتفعة.

All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

المنارة للاستشارات
www.manaraa.com

ثانيا القوة عند وجود دالة فقرة تفاضلية مقدراها ٠,٥٠

أظهرت النتائج أن الطرائق الثلاث تتمتع بقدرة عالية في الكشف عن الفقرات التي يوجد فيها دالة فقرة تفاضلية مقدارها ٠,٥ ، مع كون نتائج طريقتي الانحدار اللوجستي أفضل قليلا من طريقة اختبار تحيز الفقرة المتزامن، فقد كانت قيمة القوة للطرائق الثلاث(≤ ٠,٨٠) في جميع الخلايا التي كان فيها حجم العينة كبيرا وجميع الخلايا التي كان فيها حجم العينة متوسطا ومعلم التمييز مرتفعا أو متوسطا (باستثناء خليتين لهما متوسط غير متساوي لتوزيع القدرة للمجموعتين وحجم متوسط وغير متساوي لطريقة اختبار التحيز المتزامن)، كما أن قيم القوة كانت في معظمها بين ٠,٦٠-٠,٧٩ للطرائق الثلاث عندما كان حجم العينة صغيرا ومعلم التمييز كبيرا ، وقد كانت قيم القوة ≤ ٠,٦٠ فقط في الخلايا التي كان فيها حجم العينة صغير ومعلم التمييز منخفضا والخلية التي كان فيها حجم العينة متوسطا ومتساويا بين المجموعتين ومتوسط توزيع القدرة مختلف بين المجموعتين ومعلم التمييز منخفض للطرائق الثلاث .

وفيما يتعلق بتأثير المتغيرات المستقلة على القوة فقد أظهرت النتائج وجود تأثير لتوزيع القدرة على القوة لطريقة الانحدار اللوجستي، حيث أن قيمة القوة كانت أعلى عند تساوي متوسط توزيع القدرة بين المجموعتين، كما أظهرت النتائج وجود تأثير لحجم العينة ومعلم التمييز على القوة للطرائق الثلاث، حيث كانت القوة تزداد مع حيث كانت الزيادة في القوة الناتجة عن الزيادة حجم العينة تقل مع ازدياد معلم التمييز وهذه النتائج تتعارض مع النتائج السابقة التي ظهرت عند وجود دالة فقرة تفاضلية مقدارها ٠,٢٥ حيث كانت الزيادة الناتجة عن زيادة حجم العينة تزداد مع ازدياد معلم التمييز ويمكن تفسير ذلك بأن قيمة القوة الإحصائية للطرائق الثلاث عند وجود دالة فقرة تفاضلية مقدارها ٠,٥ كانت مرتفعة وقريبة من القيمة القصوى التي يمكن أن تصل إليها القوة الإحصائية، وعند ازدياد حجم العينة ومعلم التمييز كانت قيم القوة تقترب من القيمة القصوى وتقل الفرصة الممكنة للزيادة في قيم القوة الإحصائية للطرائق، كما أظهرت النتائج أن القوة الإحصائية لطريقة الانحدار اللوجستي تأثرت بالتفاعل بين حجم العينة ونسبتها فقد كانت قيمة القوة أعلى عند اختلاف نسبة العينة(٢ : ١) وعندما كان حجم العينة متوسطا وكبيرا، كما تأثرت القوة الإحصائية لطريقة الانحدار اللوجستي بالتفاعل بين معلم التمييز وتوزيع القدرة وكذالك بالتفاعل بين معلم التمييز ونسبة العينة حيث كانت أعلى قيم للقوة عندما كان معلم التمييز متوسطا ومرتفعا وعند اختلاف نسبة العينة (٢ : ١) أو عند اختلاف متوسط توزيع القدرة للمجموعتين المرجعية والمستهدفة وتأثرت القوة الإحصائية لطريقة الانحدار اللوجستي كذلك

All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

المنارة للاستشارات
www.manaraa.com

بالتفاعل بين توزيع القوة ونسبة العينة فقد كانت أعلى قيمة القوة عندما كان حجم المجموعتين متساويا (١ : ١) ومتوسط توزيع القدرة متساويا .

ثالثا القوة عند وجود تحيز غير منتظم

أظهرت نتائج الدراسة أن اثنتين من الطرائق لا تصلح للكشف عن التحيز غير المنتظم وهما طريقة مانتل هينزل وطريقة اختبار تحيز الفقرة المتزامن، فقد كانت قيم القوة لطريقة مانتل هينزل في جميع الخلايا (< ٠,٦٠) باستثناء أربعة خلايا، أما طريقة اختبار تحيزالفقرة المتزامن فقد كانت قيم القوة (< ٠,٦٠) في جميع خلايا باستثناء أربعة خلايا كان فيها حجم العينة كبيرا، وهذه النتائج كانت متوقعة حيث أن هذه الطرائق في الأصل غير مصممة للكشف عن دالة الفقرة التفاضلية غير المنتظمة وهذا ما أكدته العديد ممن الدراسات، فقد أشارت دراسة (Kristjansson. , Aylesworth , McDowell & Zumbo (2005 التي أجريت على الفقرات متعددة الاستجابة أن طريقة مانتل هينزل لا تصلح للكشف عن دالة الفقرة التفاضلي غير المنتظمة، كما أشارت دراسة Tian, (1999)، ودراسة Chang And Rouses (1996) أن طريقة مانتل هينزل وطريقة اختبار التحيز المتزامن لا تصلحان للكشف عن دالة الفقرة التفاضلية، الطريقة الوحيدة التي كانت ملائمة للكشف عن دالة الفقرة التفاضلية غير المنتظمة هي طريقة الانحدار االوجستي وذلك في الخلايا التي كان فيها حجم الخلية كبيرا أو متوسطا ومعلم التمييز مرتفعا أو متوسطا عند تساوي متوسط توزيع القدرة وفي الخلايا التي كان فيها حجم العينة كبيرا ومعلم التمييز مرتفعا أو متوسطا عند عدم تساوي توزيع القدرة .

وفيما يتعلق بتأثير المتغيرات المستقلة على القوة لطريقة الانحدار اللوجستي فقد أظهرت النتائج وجود تأثير لتوزيع القدرة ونسبة العينة بين المجموعتين المرجعية والمستهدفة على القوة لطريقة الانحدار اللوجستي، حيث أن قيمة القوة كانت أعلى عند تساوي متوسط توزيع القدرة بين المجموعتين وعندما كانت نسبة العينة (١ : ١)، كما أظهرت النتائج وجود تأثير لحجم العينة ومعلم التمييز على القوة، حيث كانت القوة تزداد بزيادة كل من حجم العينة ومعلم التمييز، وفيما يتعلق بالتفاعلات الثنائية بين المتغيرات المستقلة، فقد كان هنالك تأثير للتفاعل بين حجم العينة ومعلم التمييز على القوة الإحصائية لطريقة الانحدار اللوجستي فقد كانت الفروق الناتجة عن حجم العينة أكبر عندما معلم التمييز متوسطا ومرتفعا، وقد كان هنالك تأثير للتفاعل بين حجم العينة ونسبة العينة حيث كانت قيمة القوة أعلى عندما كانت نسبة حجم العينة بين المجموعتين (١ : ١) وعندما كان حجم العينة كبيرا ومتوسطا بينما كانت قيمة القوة أعلى عند اختلاف حجم

All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

المنارة للاستشارات
www.manaraa.com

المجموعتين (٢ : ١) عندما كان حجم العينة صغيرا، كما كان هنالك تأثير للتفاعل بين معلم التمييز وتوزيع القدرة على القوة، حيث كانت الفروق الناتجة عن اختلاف متوسط توزيع القدرة بين المجموعتين أكبر عندما كان حجم العينة متوسطا وكبيرا، وكذلك الأمر بالنسبة للتفاعل بين معلم التمييز وتوزيع القدرة.

ويلاحظ مما سبق أن حجم العينة ومعلم التمييز من أهم العوامل التي أثرت على القوة لجميع الطرائق سواء عند وجود دالة فقرة تفاضلية منتظمة أو غير منتظمة، وقد كان هذا متوقعا وبشكل خاص لحجم العينة نظرا للعلاقة بين حجم العينة والقوة، فالقوة تزداد عند زيادة حجم العينة، وتتفق نتائج هذه الدراسة مع نتائج دراسة Jiang & Stout (1998) والتي أشارت إلى أن القوة لطريقتي مانتل هينزل واختبار تحيز الفقرة المتزامن تزداد بزيادة حجم العينة، فقد كانت قيم القوة للطريقتين ضعيفة عندما كان حجم العينة صغيرا في المجموعتين المرجعية والمستهدفة (٢٥٠،٢٥٠) ومرتفعة عندما كان حجم العينة كبير في المجموعتين المرجعية والمستهدفة (٣٠٠٠،٣٠٠٠) .

كما أنها اتفقت مع هذه النتائج دراسة Kristjansson., Aylesworth, Mcdowell & Zumbo (2005) التي أجريت على الفقرات متعددة الاستجابة والتي أظهرت أن القوة لطريقتي مانتل هينزل والانحدار اللوجستي تزداد بزيادة حجم العينة، كما أشارت نتائج دراسة Tian 1999 أن طريقة مانتل هينزل واختبار التحيز المتزامن تتمتعان بقوة عالية في الكشف عن دالة الفقرة التفاضلية المنتظمة ولم يكن أدائها ضعيفا إلا عند وجود معلم تمييز منخفض وحجم عينة صغير، حيث أن ازدياد حجم العينة ومعلم التمييز أدى إلى زيادة القوة . كما يلاحظ من النتائج اختلاف في تأثير متغير توزيع القدرة على القوة للطرائق الثلاث حيث كان هنالك تأثير لتوزيع القدرة على طريقة الانحدار عند وجود دالة فقرة تفاضلية منتظمة وغير منتظمة، و تأثير لمتغير توزيع القدرة على طريقة اختبار تحيز الفقرة المتزامن عند وجود دالة فقرة تفاضلية مقدراها منتظم مقداره ٠,٢٥، ولم يكن هنالك تأثير لتوزيع القدرة على القوة لطريقة مانتل، وقد اختلفت نتائج الدراسات السابقة حول تأثير توزيع القدرة أيضا فقد أشارت دراسة Tian (1999) إلى وجود تأثير لتوزيع القدرة حيث ازدادت القوة عند وجود اختلاف في توزيع القدرة، بينما أشارت دراسة Kristjansson., Aylesworth, Mcdowell & Zumbo (2005) إلى عدم وجود تأثير لمتغير توزيع القدرة على القوة، كما أظهرت النتائج أن القوة لطريقة اختبار التحيز المتزامن عند وجود دالة فقرة تفاضلية منتظمة مقدارها ٠,٢٥ ،ولطريقة الانحدار اللوجستي عند وجود دالة فقرة تفاضلية غير منتظمة كانت أعلى عندما كانت نسبة العينة متساوية(١ : ١)،

All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

المنارة للاستشارات
www.manaraa.com

وهذه النتائج تختلف مع دراسة Kristjansson., Aylesworth, Mcdowell & Zumbo (2005) التي أشارت إلا أن القوة ازدادت قليلا عندما كانت نسبة العينة بين المجموعة المرجعية والمستهدفة مختلفة (٤-١).

وإذا ما أخذنا بعين الاعتبار كل من الخطأ الأول والقوة فإن الطرق الأربع لا يفضل استخدامها عندما يكون حجم العينة صغيرا حيث أن قدرتها في الكشف عن الأداء التفاضلي بين المجموعات تكون ضعيفة. كما يمكننا القول بشكل عام أن طريقة الانحدار اللوجستي هي أكثر الطرائق فاعلية في الكشف عن دالة الفقرة التفاضلية، نظرا لكون معدل الخطأ من النوع الأول الأقل من بين الطرائق الأربع، وكون القوة الإحصائية لها أعلى من طريقتي مانتل هينزل واختبار تحيز الفقرة المتزامن، وهي الطريقة الوحيدة التي أظهرت النتائج أنها تصلح للكشف عن دالة الفقرة التفاضلية.

All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

المنارة للاستشارات
www.manaraa.com

المراجع العربية

البستنجي، محمود (٢٠٠٤)، **مقارنة أربع طرائق للكشف عن الأداء التفاضلي لمتغير الجنس في فقرات اختبار قدرات عقلية خاصة للفئة العمرية من ١٥-١٦ سنة في الأردن**، رسالة دكتوراه غير منشورة، جامعة عمان العربية للدراسات العليا، عمان، الأردن.

جرادات، عبد الله (٢٠٠٣)، **مقارنة بين طريقة مانتل هينزل وطريقة الصعوبة المحولة في الكشف عن تحيز الفقرات**، رسالة ماجستير غير منشورة ، جامعة مؤتة، الكرك، الأردن.

العطيوي، إيمان (٢٠٠٤)، **الكشف عن الأداء التفاضلي لمتغير الجنس لفقرات اختبار تحصيلي في العلوم العامة للصف الثامن الأساسي في مديرية تربية وتعليم عمان الرابعة**، رسالة ماجستير غير منشورة، جامعة عمان العربية للدراسات العليا، عمان، الأردن.

القرعان، محمود (١٩٩٧)، **مقارنة فعالية ثلاث طرائق للكشف عن تحيز الفقرة**، رسالة ماجستير غير منشورة، جامعة اليرموك ، اربد ، الأردن.

ظاظا، حيدر (٢٠٠٦)، **فاعلية طريقة الكشف عن التحيز في الفقرة المبنية على الإجراءات التحكيمية مقارنة بالطرائق التجريبية المبنية على إجراءات النظرية الكلاسيكية في القياس ونظرية استجابة الفقرة**، أطروحة دكتوراه غير منشورة، الجامعة الأردنية، عمان، الأردن .

بدر، سالم (١٩٩٣)، **التحيز الاجتماعي الاقتصادي في الاختبارات العقلية الجمعية للأفراد الأردنيين في الفئة العمرية(٩-١٢) سنة**، أطروحة دكتوراه غير منشورة، الجامعة الأردنية، عمان، الأردن.

All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

المنارة للاستشارات
www.manaraa.com

المراجع الأجنبية

Ankenmann, R. , Witt, E. & Dunbar, S. (1999) , An Investigation Of The Power Of The Likelihood Ratio Goodness-Of-Fit Statistic In Detecting Differential Item, **Journal Of Educational Measurement**, 36(4), 277-300.

Beer M. D. (2004), Use Of Differential Item Functioning (Dif) Analysis For Bias Analysis In Test Construction**, Journal Of Industrial Psychology,** 30 (4), 52-58.

Berk Ronald A . (1982), **Hand Book Of Methods For Detecting Test Bias**, Baltimore: Johns Hopkins University Press, Maryland.

Bert, D.& Stoneberg, J. (2004), **Study Of Gender-Based And Ethnic-Based Differential Item Functioning (DIF) In The Spring 2003 Idaho Standard Achievement Tests Applying The Simultaneous Bias Test (SIBTEST) And The Mantel-Hanszel Chi Square Test,** University Of Maryland, Statistics and evaluation department and national center for education statistics, (Nces), Assessment Division.

Bertrand, R. & Boiteau, N. (2003), Comparing The Stability Of IRT-Based And Non IRT-Based Dif Methods In Different Cultural Contexts Using TIMMS Data Methods**, (Eric D. No. 476924).**

Camilli G. & Shepard, L. (1994)**, Methods For Identifying Biased Test Items**, New York: Sage Publications.

All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

www.manaraa.com

Chaimongkol, S. (2005), **Modeling Differential Item Functioning (Dif) Using Multilevel Logistic Regression Models: A Bayesian Perspective**, Unpublished Doctoral Dissertation, University Of Florida, Florida.

Chang H., Mazzeo J. , And Rouses, L. (1996), Detecting Dif For Polytomously Scored Items: An Adaptation Of The SIBTEST Procedure, **Journal of Education Measurement**, 33(3), 333-353.

Cohen, A. S. And Kim, S. (1993), A Comparison Of Lord's Χ And Raju's Detection Of DIF, **Applied Psychological Measurement**, 17(1), 39-52.

Conoley, A. C. (2003), **Differential Item Function In The Peabody Picture Vocabulary Test-Third Editions: Partial Correlation Versus Expert Judgment**, Unpublished Doctoral Dissertation, Texas A&M University, Texas.

Crocker L. & Algina J. (1986), **Introduction To Classical And Modern Test Theory,** New York : CBS College Publishing .

Daniel M. B. (2002), A Monte Carlo Comparison Of Parametric And Nonparametric Polytomous DIF Detection Methods , **Applied Measurement In Education**, 15(2), 113–141.

Finch, W. H. & French, B. F. (2007), Detection of Crossing Differential Item Functioning A Comparison of Four Methods, **Educational and Psychological Measurement**, 67(4), 565-582.

All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

المنارة للاستشارات
www.manaraa.com

Gao, L. & Wang C. (2005), Using Five Procedure To Detect DIF With Passage-Based Testlets, **Paper Prepared For The Poster Presentation At The Graduate Student Poster Session At The Annual Meeting Of The National Council Of Measurement In Education**, Montreal, Quebec.

Gibson S.G. (1998), Gender And Ethnicity-based differential item functioning on armed services vocational aptitude battery, blacks burg, va . **http://scholar.lib.vt.edu/theses/available/etd-93098-11430/unrestricted /Sggps2.pdf. 12-6-2007.**

Gierl M. J., Gotzman, N. A. & Boughton, K. A. (2004), Performance Of SIBTEST When The Percentage Of Dif Items Is Large, **Applied Measurement In Education**, 17(3), 241-264.

Gierl, M. J., Bisanz, J., Bisanz, G. L., Boughton K. A., Khaliq S. N., (2001), Illustrating the Utility of Differential Bundle Functioning Analyses to Identify and Interpret Group Differences on Achievement Tests, **Paper Presented at the Annual Meeting of the American Educational Research Association (AERA)**, Seattle, Washington, USA.

Hambleton, R. & Swaminathan, H. (1985), **Item Response Theory principles and application**, Boston : Kluwer Nijhoff Publishers Group.

Hambleton R. & Swaminthan, H .& Rogers, H. (1991), **Fundamental of Item Response Theory,** London : Sage Publications.

All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

المنارة للاستشارات
www.manaraa.com

Hambleton, R. & Rodgers, J. (1995), Item Bias Review, **http://www.eric facility.net/databases/eric digests/ed398241.html. 12-6-2007 .**

Hidalgo, M. D. & López-Pina, J. A. (2004), Differential Item Functioning Detection And Effect Size: A Comparison Between Logistic Regression And Mantel-Haenszel Procedures, **Educational And Psychological Measurement**, 64(6), 903-915.

Jensen R. (1980), **Bias In Mental Testing**, New York: The Free Press A Division Of Macmillan Publishing Co.

Jiang, H. & Stout, W. (1998), Improved type I error control and reduced Estimation Bias For DIF Detection Using SIBTEST, **Journal Of educational And Behavioral Statistics**, 23, 291- 322.

Kim, S. & Cohen, A. S. (1995), A Comparison Of Lord'S Chi Square Raju,S Area Measures, And Likelihood Ratio Test On Detection Of Differential Item Function, **Applied Measurement In Education** , V4(3), 291-312 .

Kristjansson, E. , Aylesworth, R. , Mcdowell, I. & Zumbo, B. (2005), Comparison of four methods for detecting differential item functioning in ordered response items. **Educational And Psychological Measurement**, 65(6).

Li, H. & Stout, W. (1996), Anew Procedure For Detection Of Crossing Dif**,** **Psychometrika**, 61, 647-677.

All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

المنارة للاستشارات

www.manaraa.com

Luppescu S. (2002), Dif Detection In Hlm **, Paper Presented At The Aera Annual Meeting,** New Orleans, U.S.A.

Morales L. S. (2004), Assessing Item And Scale Differential Functioning Using The DFIT Methodology, **Http://Outcomes.Cancer.Gov/Conference/Irt/Morales.Pdf 12-6-2007.**

O'neal, M. R. (1991), A Comparison Of Methods For Detecting Item Bias. **Paper Presented At The Annual Meeting Of The Mid-South Educational Research Association (20Th, Lextington, Ky, November 12-15).**

Pastor D. A. & Richarde R. G. (2003), Investigation of Differential Tem Functioning For African-American And Anglo-American College Students In The Modified Achievement Goal Questionnaire, **Poster Presented at the National Council of Measurement in Education's Graduate Student Poster Session,** Chicago.

Raju, N. S. (1988), The Area Between Two Item Characteristic Curves, **Psychometrika** , 53(4), 495-502.

Raju, N. S. (1990), Determining the Significance of Estimated Signed and Unsigned Areas Between Two Item Response Functions, **Applied Psychological Measuremen**t, 14, 197-207.

All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

www.manaraa.com

Raju, N. S. , Drasgow, F . & Slind, J. (1993), An Empirical Comparison Of The Area Methods, Lord's Chi-Square Test, And The Mantel-Haenszel Technique For Assessment Differential Item Function, **Educational And Psychological Measurement**, 53(1) 301-314.

Roever C. (2005), That's not fair Fairness, bias, and differential item functioning in language testing, **http://www2.hawaii.edu/~roever/brownbag.pdf 25-6-2007.**

Rousses, Louis A., Stout, William F. (1996), Simulation Studies Of The Effects Of Small Sample Size And Studied Item Parameters On SIBTEST And Mantel-Hanszel Type I Error Performance, **Journal Of Educational Measurement** 33 (2).

Samuelsen K. (2005), **Examining Differential Item Functioning From A Latent Class Perspective.** Unpublished Doctoral Dissertation,. University Of Maryland, U.S.A.

Shealy, R. & Stout, W. (1993), A Model-Based Standardization Approach That Separate True Bias/ Dif From Group Ability Differences And Detects Test Bias/DTF As Well As Item Bias/Dif, **Psychometrika**, 58, 159-194.

Sheilai J., Teresi J., Gershon R. (2005), Procedures For The Analysis Of Differential Item Function (DIF) For Small Sample Size**. Evaluation & The Health Professions**, 28(3), 283-294.

All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

المنارة للاستشارات
www.manaraa.com

Skaggs G. & Lissitz, R. (1992), The Consistency Of Detecting Item Bias Across-Different Test Administration: Implication Of Another Failure, **Journal Of Education Measurement**, 29(3), 227-242

Su, Y. & Wang W. (2005), Efficiency of the Mantel, Generalized Mantel–Haenszel, and Logistic Discriminant Function Analysis Methods in Detecting Differential Item Functioning for Polytomous Items **Applied Measurement In Education**, 18(4), 313–350 .

Teresi, J. (2004), Differential Item Functioning And Health Assessment, Columbia University Stroud Center And Faculty Of Medicine. NewYork State Psychiametric Institute, **Http://Outcomes.Cancer.Gov/Conference/Irt/Teresi.Pdf . 12-6-2007.**

Teresi, J. A., Klenman M. & Ocepek-Weliksonm K. (2000), Modern Psychometric Methods For Detection Of Differential Item Functioning: Application To Cognitive Assessment Measures, **Statistics In Medicine** ,19,1651-1683.

The William Stout Institute For Measurement (2002), **Dimensionality –Based Dif Analysis Package : SIBTEST, Poly SIBTEST, Crossing SIBTEST, DIFSIM, DIFCOMP**, USA .

Thissen D. (2001), **Software For The Computation Of The Statistics Involved In Item Response Theory Likelihood-Ratio Tests For Differential Item Functioning**, *L.L.* Thurstone Psychometric Laboratory University Of North Carolina .

All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

المنارة للاستشارات
www.manaraa.com

Tian, F. (1999), **Detecting DIF In Polytomous Item Responses,** Unpublished Doctorial Dissertation, University Of Ottawa, Ottawa.

Wang, W. & Su, Y. (2004), Effects Of Average Signed Area Between Two Item Characteristic Curves And Test Purification Procedures, **Applied Measurement Education**, 17(2), 113-144.

Zumbo, B. D. (1999**), A Handbook on the Theory and Methods of Differential Item Function (Dif)**, Ottawa, On Directory of Human Resources research And Evaluation Department of National Defense.

Zumbo B. D. (2001), Investigation DIF By Statistical Modeling Of The Probability Of Endorsing An Item: Logistic Regression And Extensions Thereof. **Paper Presented At The National Council For Measurement In Education Meeting** , University Of South Carolina.

All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

المنارة للاستشارات
www.manaraa.com

A COMPARISON BETWEEN FOUR METHODS FOR DETECTING ITEM FUNCTION

(ASSIMILATION STUDY)

By
Nezar Allabadi

Supervisor
Dr. Khalil Ellian, Prof.

ABSTRACT

The study aimed at evaluating the effectiveness and comparing the four differential item function (DIF) methods. Two of these method are based on the classical theory; they are Mantel Haenzsel (MH) and the logistic regression (LR). The other two are based on the item response theory: they are item response theory likelihood-ratio test (IRTLR) and simultaneous item bias (SIBTEST). The four methods were evaluated by using the simulation technique in generating data based on two criteria: type I error and the power. The data were generated for a test consisted of forty items based on the three parameter logistic model. all the items had the same parameters in the two groups: reference and focal group except the last item. The study examined the effectiveness of those methods under different conditions of the ability distribution, the sample size, its ratio among the two groups the referential group and the foal group, the item discrimination, the type differential item function (DIF) and its size.

The study revealed the fowling results:

- The percentage of the first types error was acceptable and near the theoretical value 0.05 for three methods (LR, MH, and SIBTEST). The results also showed a clear inflation of type I error for IRTLR. The most effective method based on the criteria of the first types error was LR then MH then SIBTEST, and finally IRTLR.

All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

المنارة للاستشارات
www.manaraa.com

- There was an effect of sample size, item discrimination parameter and the interaction between the sample size and its ratio on the type I error of IRTLR. The result also revealed that there was an effect by the interaction between the item discrimination parameter and the ratio of sample on the type I error.
- The statistical power of the three methods that has an acceptable level of the fist type error (LR, MH, & SIBTEST) when used to reveal the uniform DIF of the value 0.25 was in general week , because the statistical power of the three functions in most of the cells was ≤.60. But when used to reveal the uniform DIF of the value 0. 5 the three method's (LR, MH, SIBTEST) had high ability to examine the items that have uniform DIF, however, the result of the two LR and MH were a little bit better than SIBTEST. Most of the power value for the three method was ≥0.80. The result also showed that the only method that can reveal the differential item was the LR in the cells that had average or large sample size and average or high discrimination parameter when there was an equity of the means of ability distribution between the two groups, And in the cells that had large sample size and average or high discrimination parameter when the means of ability distribution differed .
- Regarding the effect of the independent variable on the power of the three methods that has an acceptable level of the fist type error (LR, MH, & SIBTEST), the sample size and the discrimination parameter was the most important variable that had effect on the power, because there was an increase of power due to the increase of sample size and discrimination parameter.

All Rights Reserved - Library of University of Jordan - Center of Thesis Deposit

المنارة للاستشارات
www.manaraa.com