المفاضلة بين طرق تقدير الدوال الاقتصادية ذات المتغيرات التابعة النوعية

أ.د. عباس ناجي جواد

المستلخص

ان استخدام المتغيرات النوعية كمتغيرات مستقلة في معادلة الانحدار لم يعد يشكل مشكلة حيث اصبح ترميزها وطرق تقديرها وطرق تفسيرها شائعة باستخدام المتغيرات الوهمية.

غير ان استخدام المتغيرات النوعية كمتغيرات تابعة في معادلة الانحدار هو المشكلة بعينها بسبب عدم معرفة او عدم شيوع التكنيك اللازم للتقدير ذلك لان تقدير هذا النوع من المعادلات يتطلب نماذجا خاصة للحصول على تقديرات دقيقة.

ان اختيار النموذج الصحيح لتقدير المعادلات ذات المتغيرات التابعة النوعيةليس بالعملية الروتينية البسيطة ذلك لانها تتطلب الايفاء بعدة شروط منها ما يتعلق بعدد المتغيرات المستقلة ومنها ما يتعلق بطبيعة تلك المتغيرات .

سيحاول الباحث في هذا البحث مناقشة وتطبيق طرق تقدير المعادلات ذات المتغيرات التابعة النوعية من اجل اختيار الطريقة الافضل من بينها.

Abstract

The use of qualitative variables, as independent or explanatory variables in the regression equation, no longer a problem, where the methods of encoded, estimation and interpretation of results, are known, by using dummy variables.

But the use of these variables as dependent variables in the regression equation, it is the problem, due to the unknown of the necessary technique to perform this operation, because the use of qualitative variables as dependent variables, requires a special regression models in order to get accurate results.

The selection of the correct model is not a routine process because it requires several special conditions, about the number ,the composition and the nature of independent variables.

I try to explain, discuss and practice that analytical technique, of a limited use now, to enrich the quantitative analytical tools are currently known in the Arab economic literature, especially after the increasing need to adopt qualitative variables as dependent variables, which deepens the need to use appropriate analytical tools.

أولا : المقدمة:

يضطر الاقتصاديون إلى إدخال المتغيرات النوعية أو المتغيرات غير الكمية في نماذج الانحدار التي يقومون بتقديرها ، والمتغيرات النوعية أو الصفات أو المتغيرات المتقطعة او الفئوية، تشمل طائفة واسعة من المتغيرات المهمة المؤثرة على قيمة معلمات الدوال المقدرة ، وبدون استخدام هذه المتغيرات لا يمكن التوصل إلى تقدير معلمات انحدار دقيقة ، فمثلا حالة الحرب والسلم تؤثر تأثيرا بالغا على اغلب المتغيرات الاقتصادية الكلية كالدخل القومي والناتج القومي والاستهلاك والاستثمار والادخار والتقدم التكنولوجي .....الخ. وكذا الحال بالنسبة للمتغيرات الجندرية ،حيث تم اعتماد متغير الجنسانية أو الجنسوية في العديد من الدراسات الحديثة وخاصة ما يتعلق منها بمستويات البطالة والأجور والتعليم وتوقع الحياة ....الخ. وهكذا بالنسبة لبقية المتغيرات النوعية ذات الطبيعة المشابهة.

وقد جرت العادة على تمثيل المتغيرات النوعية بمتغيرات وهمية او صماء Dummy variables ، حيث يضم المتغير الوهمي قيمتين ، هما قيمة الصفر، وترمز للحالة الأساس وهي حالة السلم مثلا ، وقيمة 1 التي ترمز للحالة الاستثنائية وهي حالة الحرب مثلا.

مشكلة البحث:

إن إدخال المتغيرات النوعية ، كمتغيرات مستقلة أو توضيحية في معادلة الانحدار، لم يعد مشكلة، حيث أصبحت طرق ترميز وتقدير وتفسير النتائج شائعة ، باستخدام المتغيرات الوهمية ، غير أن استخدام تلك المتغيرات كمتغيرات تابعة في معادلة الانحدار، هو المشكلة بعينها، بسبب عدم انتشار التكنيك اللازم لإجراء هذه العملية، لان إدخال المتغيرات النوعية كمتغيرات تابعة في معادلة الانحدار يتطلب استخدام نماذج انحدار خاصة من اجل الحصول على نتائج منطقية. وان اختيار النموذج الصحيح ليس بالعملية الروتينية لأنه يتطلب عدة شروط خاصة بالمتغيرات المستقلة ، منها ما يتعلق بعددها ومنها ما يتعلق بطبيعتها المختلفة ، ومنها ما يتعلق بتفردها بنوع واحد أو المزيج فيما بين نوعين مختلفين منها او من البيانات.

فرضية البحث:

يفترض البحث أنه لا يمكن استخدام طريقة المربعات الصغرى الاعتيادية في تقدير نماذج الانحدار ذات المتغيرات التابعة النوعية لأنها تؤدي إلى تقدير معلمات انحدار غير منطقية وان استخدام نموذج الاحتمال الخطي في الانحدار لا يحل المشكلة إنما يقوم بتعميقها ، ومن هنا يأتي الافتراض الثاني للبحث وهو أن استخدام التحليل التمييزي أو نموذج الانحدار اللوجستي أو نموذج الانحدار اللوجيتي Logit (وهو غير اللوجستي Logistic) أو نموذج الانحدار البروبيتي Probit ، يؤدي إلى تقدير معلمات دقيقة ومنطقية.

هدف البحث: يهدف البحث الى :

١- التعريف بالطرق الممكنة لتقدير الدوال ذات المتغيرات التابعة النوعية.
٢- اختبار تلك الطرق عمليا من اجل المفاضلة فيما بينها.
٣- إثبات فرضيات البحث المذكورة أعلاه.

أهمية البحث:

تتأتى أهمية البحث من محاولته التعريف بأدوات تحليل كمية معدومة او قليلة الاستخدام حاليا، ترفد الأدوات التحليلية الكمية المتعارف عليها حاليا في الأدب الاقتصادي العربي ، خاصة بعد تزايد الحاجة الى اعتماد المتغيرات النوعية كمتغيرات تابعة، الأمر الذي يعمق الحاجة الى استخدام أدوات تحليل مناسبة.

منهجية البحث:

سيتبع البحث أسلوب التحليل الوصفي تارة وأسلوب التحليل القياسي الكمي تارة أخرى ، وحيث يتطلب الأمر، هذا من جهة، أما من جهة أخرى فسوف يتم انجاز هذا البحث بتناوله

لهذه المقدمة أولا ، ثم الاستعراض النظري لنماذج الانحدار المختلفة التي يمكن استخدامها في هذا الموضوع،ثانيا، أما ثالثا فسوف تتضمن تطبيقا عمليا باستخدام بيانات واقعية أحيانا وافتراضية في أحيان أخرى ، وبمساعدة الحزمة الإحصائية الجاهزة **SPSS17** ، وسوف يختم البحث بالاستنتاجات والمقترحات.

ثانيا: الاستعراض النظري لنماذج انحدار المتغيرات التابعة النوعية

لا يمكن لنموذج الانحدار الاعتيادي المقدر بطريقة المربعات الصغرى الاعتيادية OLSوالذي يتخذ الصيغة $Y = a + bX_i$ أن يؤدي دوره بشكل جيد إن كانت المتغيرات التابعة متغيرات وهمية أي متغيرات مرمزة بقيم متقطعة كالصفر والواحد ، فهو لن يستطيع أن يقدر معلمات الانحدار بكفاءة، وبالتالي لن تكون مفيدة في توقع النتائج أو في التحليل والتنبوء. ذلك أن هذا النوع من الانحدار وبسبب طبيعة المتغير التابع سيؤدي إلى ظهور مشكلة عدم تجانس تباين الخطأ وارتباط قيم الخطأ بقيم المتغيرات المستقلة ، لذا فان القيمة المتوقعة للمتغير التابع سوف لن تقع بالضرورة ضمن المدى المنطقي والفعلي لذلك المتغير ،أي أنها سوف لن تقع بين الصفر والواحد، أي بين لا ونعم إن كان الصفر يمثل "لا" والواحد يمثل "نعم" مثلا، أو إن القيمة المتوقعة للمتغير التابع سوف لن تقع بين " أنثى" و"ذكر" إن كان الصفر يمثل الإناث والواحد يمثل الذكور مثلا، فان كانت القيم المتوقعة بهذا الشكل فمن تمثل إذن حيث لا يوجد لدينا غير النعم أو لا في عملية التصويت وغير الأنثى والذكر كمتغير جندري. وتبقى هذه المشكلة قائمة بغض النظر عن شكل العلاقة المقدرة أي في النماذج الخطية وفي النماذج اللاخطية أيضا.

ولمعالجة المشاكل أعلاه تم اللجوء إلى نموذج الاحتمال الخطي في الانحدار The linear probability model (***عطية، ٢٠٠٤، ٣٧٢-٣٧٥***)،وهو الذي يتخذ شكل المعادلة :

$$P = E(Y = 1/ X) = a + bX$$

وهو نموذج يحاول بيان اثر المتغير الكمي المستقل في المتغير التابع النوعي،والمتغير التابع هنا هو متغير احتمالي ذو طبيعة خاصة لأنه يتضمن قيمتي الصفر والواحد كما أسلفنا، لذا فان القيمة المتوقعة له ستكون بين الصفر والواحد

$$0 < E(Y / X) < 1$$

وهنا أيضا سوف لن يكون هناك توزيع معتدل لقيم الخطأ او الحد العشوائي ، وهذا يعقد علينا اختبار معنوية المعلمات المقدرة ، الذي يفترض عدالة التوزيع، هذا مع استمرار وجود مشكلة عدم تجانس تباين حد الخطأ بسبب ارتباط قيم هذا الحد مع قيم المتغيرات المستقلة ، وللتخلص من هذه المشكلة يتم قسمة طرفي النموذج على $K\wedge$ والتي تساوي :

$$K\wedge = \sqrt{Y(1-Y\wedge)}$$

هذا علاوة على أن قيمة معامل التحديد سوف لن تعبر بدقة عن جودة التوفيق او عن القوة التفسيرية للنموذج المقدر ذلك انه طالما يمر خط الانحدار بإحداثيات مشاهدات متطرفة فقط فان قيمته ستكون منخفضة لا محالة، هذا عدا افتراض النموذج أن العلاقة بين المتغير التابع والمتغير المستقل هي علاقة خطية.

ومن اجل الحصول على نتائج تقدير منطقية للدوال ذات المتغيرات التابعة النوعية ، لابد من اعتماد احد النماذج الآتية:

1- يستخدم **التحليل التمييزي او بالأحرى تحليل الدالة التمييزية** discriminant function analysis عادة في حالة المتغير التابع النوعي او الفئوي إن كانت جميع المتغيرات المستقلة متغيرات مستمرة وبشرط أن تكون موزعة توزيعا اعتداليا(***http://bear.warrington.ufl.edu***).

فالتحليل التمييزي تكنيك يستخدم لبناء نموذج تنبؤي للانتماء الى مجموعة معينة كمتغير تابع، من خلال خصائص المشاهدات الخاصة بالمتغير او المتغيرات المستقلة . فمثلا يمكن وضع الدول في مجموعة "فقيرة جدا" او "فقيرة" ، حسب بعض المؤشرات الاقتصادية كمعدل دخل الفرد مثلا .فالتحليل التمييزي ينشيء دوالا مستقاة من حالات العينة المدروسة ، تتعلق بتصنيف عضوية الدول في إحدى المجموعتين المذكورة. يمكن بعدها تطبيق تلك الدوال على حالات او مشاهدات جديدة ، أي انه إذا عرفنا معدل دخل الفرد لأي دولة خارج العينة المدروسة ، يمكننا عندئذ تطبيق الدالة التمييزية عليها لتحديد انتمائها، أي هل تنتمي الى مجوعة " فقيرة جدا" او الى مجموعة " فقيرة ".

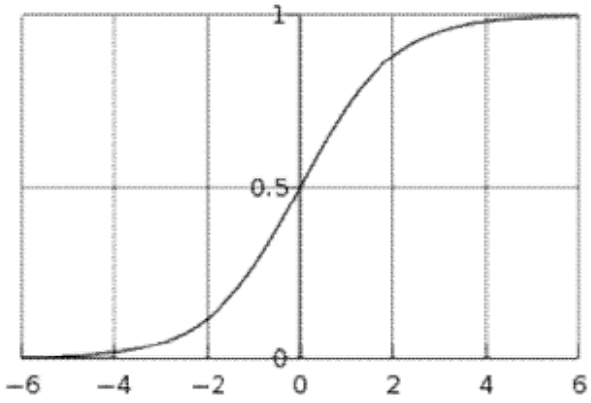
ومن الجدير بالذكر انه إذا كان لدينا مجموعتين(كحالة الفقيرة جدا والفقيرة في أعلاه) كمتغير تابع، فان التحليل التمييزي سينشيء دالة تمييزية واحدة ،وإذا كان لدينا أكثر من مجموعتين فستكون لدينا أكثر من دالة تمييزية ،غير أن الدوال المستخدمة في هذه الحالة هي الدوال الثلاث الأولى فقط عادة.

2- يستخدم **الانحدار اللوجستي** Logistic regression عادة في التنبوء بقيم المتغيرات النوعية او الفئوية categorical variable اعتمادا على مجموعة متغيرات مستقلة مختلطة mix of continuous and categorical variables ، كان يكون قسم منها متغيرات مستمرة او قياسات، أما القسم الآخر فيكون على شكل متغيرات متقطعة أي نوعية او فئوية (***http://core.ecu.edu***).

ويستخدم هذا النوع من الانحدار أيضا في حالة إذا لم تكن تلك المتغيرات المستقلة موزعة توزيعا معتدلا، حيث أن الانحدار اللوجستي لا يضع أية شروط تتعلق بطبيعة توزيع المتغيرات المستقلة.وقد شاع استخدام الانحدار اللوجستي في الدراسات الطبية حيث أن الكثير من الدراسات الطبيعة ذات متغيرات تابعة متقطعة او نوعية فإما أن يكون المرء فيها مريضا او سليما. (***http://en.wikipedia.org***) .

الانحدار اللوجستي هو نموذج يستخدم في التنبوء باحتمال حدوث حدث معين بتوفيق البيانات بشكل منحنى لوجستي، فهو إذن نموذج خطي عام generalized linear model يتخذ شكل الدالة الاتية :

شكل 1 يوضح الدالة اللوجستية حيث يقاس *z* على المحور الأفقي أما *f(z)* فيقاس على المحور العمودي



حيث أن : $f(z) = \frac{e^{z}}{e^{z}+1} = \frac{1}{1+e^{-z}}$

فالشكل أعلاه يوضح أن المتغير المستقل هو Z بينما المتغير التابع فهو $F(Z)$ ، وهي دالة مفيدة حيث أن المتغير المستقل Z يمكن أن يكون أية قيمة سالبة او موجبة ، بينما سيكون المتغير التابع بقيمة تقع بين الصفر والواحد الصحيح، والمتغير المستقل هنا يمثل مجموعة متغيرات مسببة مختلطة بين المتغيرات المستمرة والمتقطعة بينما المتغير التابع يمثل احتمالية النتيجة.

لذا يعرف المتغير المستقل Z على انه:

$$Z = a + b_1X_1 + b_2X_2 + b_3X_3 + \ldots + b_kX_k$$

أما إذا كانت المتغيرات المستقلة من نوع المتغيرات المتقطعة او الفئوية ، مع استمرار المتغير التابع بهيئته الثنائية ، فلابد من استخدام احد نماذج الانحدار الآتية:

3- **نموذج الانحدار اللوجيتي** Logit Regression Model الذي يتجاوز أو يتحايل على مشكلة عدم تجانس تباين الحد العشوائي لأنه يفترض توزيعا مختلفا لهذا الحد ***(عطية،٣٨٢، ٢٠٠٤-٣٨٥)*** و ***(http://en.wikipedia.org)***،وذلك باعتماد هذا النموذج على دالة التوزيع اللوجستي المتراكم Cumulative Logistic Probability Function والذي يعتمد الصيغة :

$$Ln(P/1-P) = a + bX$$

أما P هنا فهو عبارة عن نسبة احتمال حدوث الحدث المدروس إلى حجم العينة ، لذا فان $1-P$ ستمثل نسبة احتمال عدم حدوث الحدث، فلو كان حجم العينة 20 مشاهدة منها 12 حالة "نعم" مثلا ، فهنا سيكون:

$$P = 12/20 = 0.60 \quad , \; 1-P = 1-0.60 = 0.40$$

4- او استخدام **نموذج الانحدار البروبيتي** Probit Regression Model الذي يؤدي نفس مهمة التخلص من مشكلة عدم تجانس تباين الحد العشوائي باعتماده دالة التوزيع الطبيعي الاحتمالي المتراكم Cumulative Normal Probability Function والذي يعتمد الصيغة:

$$Ln(P) = a + bX$$

ثالثا- التطبيق العملي لنماذج انحدار المتغيرات التابعة النوعية

من اجل استكمال خطة البحث لابد لنا من تطبيق النماذج النظرية الخاصة بطرق تقدير الدوال ذات المتغيرات التابعة النوعية الواردة في ثانيا أعلاه ، تطبيقا عمليا ، ومن اجل ذلك تم اختيار 21 دولة نامية مدرجة في الجدول أدناه، وهي دول افريقية وآسيوية عموما، تتميز

بحصة واطئة للفرد من الناتج المحلي الإجمالي ، وهذه الحصة لم تبلغ الألف دولار سنويا، معتمدين في ذلك على بيانات الأمم المتحدة، وتم استكمال البيانات من نفس المصدر بإضافة مؤشر معدل السنوات الدراسية للفرد ومعدل وفيات الأطفال دون سن الخامسة، والمؤشرات الثلاث مجتمعة تعطينا صورة واضحة عن حالة الفقر والتخلف التي تتسم بها البلدان المدروسة، وتم تصنيف هذه الدول إلى صنفين هما الدول الفقيرة جدا والتي لا تتجاوز حصة الفرد من الناتج المحلي فيها 500 دولار أمريكي وهي 10 دول ، تم إعطائها الرمز صفر، أما الصنف الثاني فهو الدول الفقيرة وهي التي تقع حصة الفرد من الناتج المحلي فيها بين

جدول 1 بعض المؤشرات المختارة لدول العينة

المؤشرات المدروسة / دول العينة	حصة الفرد بالدولار من الناتج المحلي الإجمالي 2008	معدل سنوات الدراسة للفرد لسنة 2006	معدل وفيات الأطفال دون الخامسة (بالألف) 2006	الموقع القاري للدولة =0آسيا =1أفريقيا =2أمريكا الجنوبية	تصنيف الباحث للدول =0فقيرة جدا =1فقيرة
	X1	X2	X3	CONT	Y
أفغانستان	466.00	8.00	285.00	0.00	.00
بنغلاديش	494.00	8.00	57.00	0.00	.00
بنين	767.00	8.00	121.00	1.00	1.00
بوركينا فاسو	522.00	6.00	157.00	1.00	1.00
بوروندي	138.00	9.00	166.00	1.00	.00
كمبوديا	769.00	10.00	89.00	0.00	1.00
تشاد	765.00	6.00	211.00	1.00	1.00
الكونغو الديموقراطية	181.00	8.00	198.00	1.00	.00
أثيوبيا	319.00	8.00	131.00	1.00	.00
غامبيا	636.00	9.00	116.00	1.00	1.00
غانا	709.00	10.00	117.00	1.00	1.00
غويانا	505.00	8.00	148.00	2.00	1.00
العراق	788.00	10.00	41.00	0.00	1.00
كينيا	788.00	10.00	104.00	1.00	1.00
قرغيزستان	934.00	13.00	46.00	0.00	1.00
ليسوتو	788.00	10.00	104.00	1.00	1.00
ليبيريا	219.00	11.00	140.00	1.00	.00
ملاوي	278.00	9.00	121.00	1.00	.00
النيجر	354.00	4.00	172.00	1.00	.00
طاجيكستان	364.00	11.00	78.00	0.00	.00
زيمبابوي	314.00	4.00	94.00	1.00	.00

المصدر:إحصاءات الأمم المتحدة المنشورة على

http://unstats.un.org/unsd/demographic/products/socind/statistics.htm

500–1000 دولار أمريكي وهي 11 دولة ، تم إعطائها الرمز 1، وبذلك أصبح لدينا متغير تابع Y وهو صنف الدولة وهو متغير نوعي او فئوي وثلاث متغيرات مستقلة مؤثرة فيه وهي X_1, X_2, X_3 وكما مؤشر في الجدول رقم 1 أعلاه.

وبما أن المتغير التابع هو متغير نوعي ، عندئذ يمكن استخدام احد النماذج النظرية التي تم استعراضها سابقا، وبما إن المتغيرات المستقلة هي متغيرات مستمرة فلابد إذن من استخدام دالة الانحدار الخطي باستخدام طريق المربعات الصغرى ، او دالة الاحتمال الخطي لتقدير معلمات الدالة او استخدام نموذج التحليل التمييزي لتكوين دالة تمييزية تنفع في التحليل

والتنبؤ بحيث يتم من خلال استخدامها تصنيف أية دولة أخرى خارج دول العينة تصنيفا يتعلق بانتمائها الى مجموعة الدول الفقيرة جدا أم الى مجموعة الدول الفقيرة.

1- نتائج تقدير نموذج المربعات الصغرى الاعتيادية

عند تقدير نموذج المربعات الصغرى الاعتيادية وهو نموذج خطي ،باستخدام البيانات الأصلية الواردة في الجدول أعلاه، تم الحصول على النتائج الآتية والتي تفيدنا في المقارنة مع النماذج الأخرى المقترحة في هذا البحث:

$$Y = -0.517 + 0.002X1 - 0.005X2 + 0.001X3$$

t = - 1.367 7.01* - 0.160 0.451

F = 18.638*

$Adj\ R^2 = 0.726$

$D.W. = 1.044$

VIF 1.121 1.296 1.341

ويبدو من النتائج أن النموذج المقدر ذو قوة تفسيرية جيدة حيث يستطيع تفسير ٧٢% من التغيرات في المتغير التابع بإرجاعها الى تغيرات في المتغيرات المستقلة،غير أن المعلمات المقدرة للنموذج لم تكن معنوية حسب اختبار t عدا معلمة المتغير X_1 ، أما المعنوية الإجمالية للنموذج فقد كانت عالية جدا حسب اختبار F ، غير أن النموذج المقدر يعاني من مشكلة الارتباط الذاتي حسب اختبار D.W. ، غير انه لا يعاني من مشكلة عدم تجانس التباين كما يتضح من الشكل البياني 2 في أدناه، ولا يعاني من مشكلة التعدد الخطي كما يتضح من اختبار VIF .

أما من الناحية الاقتصادية فتبدو قيم وإشارة المعلمات متسقة مع المنطق الاقتصادي ، ورغم ذلك لا يمكننا قبول نتائج التقدير بسبب افتراض خطية العلاقة بين المتغيرات المستقلة والمتغير التابع (لعدم إمكانية تقدير نماذج لوغاريتمية بسبب وجود الصفر في قيم المتغير التابع)، وعدم معنوية المعلمات المقدرة علاوة على رفض بعض القيم المتوقعة للمتغير التابع والتي تكون اكبر من الواحد الصحيح او تلك السالبة منها، وهي سبع مشاهدات فقط ، كما يتضح من الجدول 2 في أدناه.

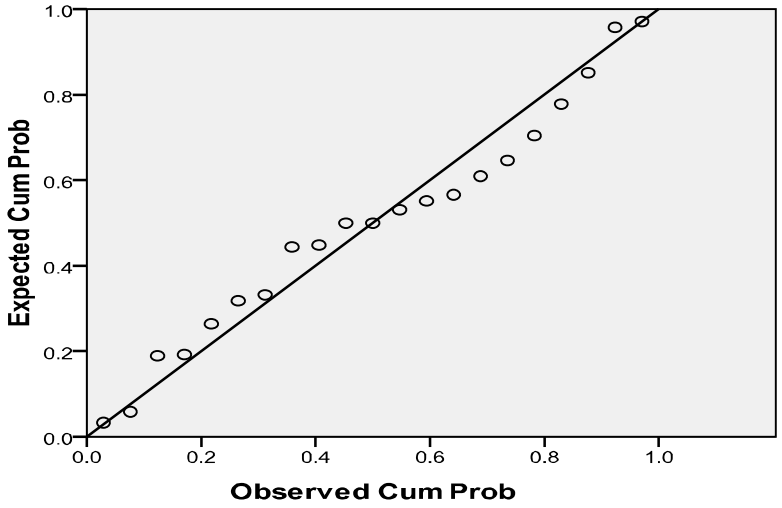
جدول 2 القيم المقدرة لـ Y باستخدام نموذج المربعات الصغرى الاعتيادية

رقم المشاهدة	القيم المقدرة لـY	رقم المشاهدة	القيم المقدرة لـY
1	0.49342	12	0.49248
2	0.42118	13	0.96556
3	0.97931	14	1.00034

4	**0.53981**	**15**	**1.23322**
5	**-0.20529**	**16**	**1.00034**
6	**0.95568**	**17**	**-0.07433**
7	**1.03496**	**18**	**0.03797**
8	**-0.10038**	**19**	**0.23617**
9	**0.12691**	**20**	**0.16912**
10	**0.72078**	**21**	**0.11651**
11	**0.85623**		

المصدر: من عمل الباحث

شكل 2 توزيع الحد العشوائي وفق نموذج المربعات الصغرى



2- نتائج تقدير نموذج الاحتمال الخطي:

عند تقدير نموذج الاحتمال الخطي The linear probability model باستخدام البيانات المعدلة سواء بالنسبة للمتغير التابع او المتغيرات المستقلة والتي تتطلب العديد من خطوات معالجة البيانات الأولية للوصول إليها(للاطلاع على تفاصيل خطوات معالجة البيانات الأولية ، ***راجع: عطية، 2004***، 280) تم الحصول على النتائج الاتية:

$$Y = 2.292 + 0.001\ X_1 - 0.010\ X_2 + 0.0\ X_3$$

t = 2.481 2.986* -2.423* -0.885*

*F = 3.926**

Adj R^2 = 0.305

D.W. = 1.368

VIF 26.577 32.462 11.914

ويبدو أن النتائج أعلاه جيدة في بعض الجوانب الإحصائية فعلى الرغم من قيمة معامل التحديد المتدنية جدا نجد أن اغلب معلمات المتغيرات المستقلة ذات تقدير معنوي حسب

اختبار t مع معنوية إجمالية عالية جدا للنموذج المقدر حسب اختبار F ، مع عدم خلو النموذج من مشكلتي التعدد الخطي والارتباط الذاتي، علاوة على ذلك فأن مشكلة عدم تجانس تباين الخطأ قد تفاقمت كما يبدو ذلك واضحا من توزيع حد الخطأ في الشكل 3 أدناه

جدول 3 لقيم المقدرة لـ Y باستخدام نموذج الاحتمال الخطي

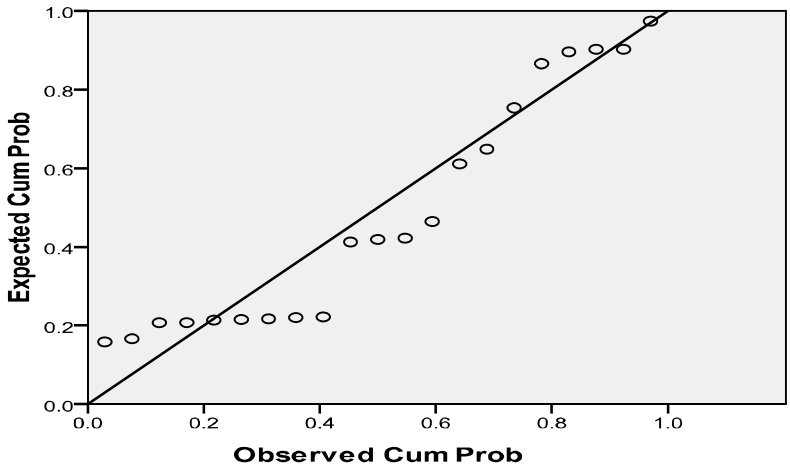
رقم المشاهدة	القيم المقدرة لـY	رقم المشاهدة	القيم المقدرة لـY
1	2.60719	12	2.68476
2	2.69872	13	4.22703
3	4.75534	14	5.76267
4	2.74248	15	6.38591
5	-6.39719	16	5.76267
6	3.92545	17	3.21943
7	5.89196	18	2.59931
8	3.31932	19	2.63383
9	2.56530	20	2.54308
10	2.88500	21	2.70132
11	3.14218		

المصدر: من عمل الباحث

هذا علاوة على افتراض النموذج أن العلاقة بين المتغير التابع وهو تصنيف الدول حسب حالة الفقر، هي علاقة خطية مع المتغيرات المستقلة ، وهو افتراض لا يمكن القبول به. هذا علاوة على أن القيم المقدرة للمتغير التابع Y يجب أن تقع بين الصفر والواحد الصحيح ، فكلما اقتربت من الصفر يعني ذلك أن الدولة فقيرة جدا وكلما اقتربت من الواحد يعني ذلك أن الدولة فقيرة ، وبالتالي لا يمكن القبول بالقيم السالبة او القيم التي تكون قيمتها اكبر من الواحد الصحيح ، كما هو الحال في جميع المشاهدات في الجدول 3 أعلاه، حيث أن جميع القيم المذكورة ليست بذات معنى.

شكل 3

توزيع الحد العشوائي وفق نموذج الاحتمال الخطي



3- نتائج تقدير نموذج الانحدار التمييزي:

وباستخدام التحليل التمييزي حصلنا على النتائج الآتية:

$$Y = -4.688 + 0.009\,X_1 - 0.022\,X_2 + 0.002\,X_3$$

Wilks' Lambda = 0.238 0.937 0.929*

Canonical Correlation = 0.876

$R^2 = 0.767$

ويبدو جليا أن النموذج التمييزي المقدر يتفق مع المنطق الاقتصادي من حيث قيمة وإشارة معلماته الاقتصادية فكلما كبرت قيمة Y أي كلما اقتربت من الواحد الصحيح كلما كانت الدولة ضمن مجموعة الدول الفقيرة وكلما صغرت قيمة Y أي كلما اقتربت من الصفر كلما انتمت الدولة الى مجموعة الدول الفقيرة جدا. وان المعلمة الأساس في النموذج هي معلمة X_1 أي معلمة حصة الفرد من الناتج المحلي الإجمالي، حيث إنها قد اجتازت اختبار Wilks بينما فشلت بذلك بقية المعلمات ، لذا فان الدالة التمييزية قد اعتمدت على مؤشر حصة الفرد من الناتج لتصنيف الدول الى فقيرة وأخرى فقيرة جدا كما فعل البحث تماما ، مع الأخذ بنظر الاعتبار بقية المؤشرات المدروسة، لذا اعترض النموذج

جدول 4 القيم المقدرة للمتغير التابع وفق نموذج الانحدار التمييزي

احتمال الانتماء الى المجموعة	احتمال الانتماء الى المجموعة	القيم المقدرة للمتغير التابع	القيم الأصلية للمتغير التابع	اسم الدولة	المشاهدة

				الفقيرة	الفقيرة جدا
1	أفغانستان	0.00	0	0.18442	0.81558
2	بنغلاديش	0.00	0	0.33263	0.66737
3	بنين	1.00	1	0.99910	0.00090
4	بوركينا فاسو	1.00	1	0.52350	0.47650
5	بوروندي	0.00	0	0.00002	0.99998
6	كمبوديا	1.00	1	0.99915	0.00085
7	تشاد	1.00	1	0.99905	0.00095
8	الكونغو الديموقراطية	0.00	0	0.00007	0.99993
9	أثيوبيا	0.00	0	0.00355	0.99645
10	غامبيا	1.00	1	0.96484	0.03516
11	غانا	1.00	1	0.99538	0.00462
12	غويانا	1.00	0	0.40472	0.59528
13	العراق	1.00	1	0.99950	0.00050
14	كينيا	1.00	1	0.99950	0.00050
15	قرغيزستان	1.00	1	0.99999	0.00001
16	ليسوتو	1.00	1	0.99950	0.00050
17	ليبيريا	0.00	0	0.00021	0.99979
18	ملاوي	0.00	0	0.00112	0.99888
19	النيجر	0.00	0	0.00949	0.99051
20	طاجيكستان	0.00	0	0.01254	0.98746
21	زيمبابوي	0.00	0	0.00309	0.99691

المصدر: من عمل الباحث

التمييزي على تصنيف دولة واحدة فقط من بين دول العينة الواحدة والعشرين وهي دولة غويانا حيث وجد أن احتمال انتمائها الى الدول الفقيرة جدا 0.59528 ، وهو اكبر من احتمال انتمائها الى مجموعة الدول الفقيرة البالغ 0.40472 ، بذلك يكون البحث قد صنف 95.2% من دول العينة تصنيفا صحيحا.

وقد بلغت القوة التفسيرية المتمثلة بقيمة R^2 نسبة 77% تقريبا وهي قوة تفسيرية جيدة ، هذا علاوة على خلو النموذج المقدر من المشاكل القياسية وذلك بسبب انتفاء قيمة حد الخطأ تقريبا بسبب عدم وجود انحرافات بين قيم Y الأصلية و Y المقدرة في 20 مشاهدة من مجموع 21 كما يتضح مفصلا من الجدول 4 في أعلاه ، هذا عدا عن منطقية القيم المقدرة للمتغير التابع التي وقعت جميعا بين الصفر والواحد الصحيح

4- نتائج تقدير النموذج اللوجستي

عند الانتقال الى استخدام النموذج اللوجستي، لابد لنا من إضافة متغير مستقل نوعي الى المتغيرات المستقلة الموجودة أصلا من اجل الاستيفاء بأحد شروط استخدام هذا النموذج الذي يتطلب أن تكون المتغيرات المستقلة مزيجا من المتغيرات المستمرة والمتغيرات المتقطعة، لذا تم اختيار متغير الموقع القاري لدول العينة وهي دول افريقية وآسيوية ، وتم تكوين المتغير الوهمي d لتمثيل القارة بحيث أعطينا الرمز 0 للقارة الآسيوية والرمز 1 للقارة الإفريقية والرمز 2 لقارة أمريكا الجنوبية، ثم أجرينا عملية تقدير المتغير التابع على أربعة متغيرات مستقلة بدلا من ثلاث متغيرات، وقد حصلنا على نتائج التقدير الآتية:

$$Y = -110.156 + 0.191 X_1 + 2.517 X_2 + 0.101 X_3 - 48.264\ cont1 - 2.754\ cont2$$

Wald 0.0 0.0 0.0 0.0 0.0 0.0

$Cox \& Snell\ R^2 = 0.749$

ويبدو من النتائج المقدرة منطقية المعلمات المقدرة من حيث إشارة معلمات المتغيرات المستقلة عدا إشارة معلمة X_2 حيث يفترض بها أن تكون سالبة.
ويلاحظ هنا أن معادلة القارة الآسيوية ستكون:

$$Y = -110.156 + 0.191 X_1 + 2.517 X_2 + 0.101 X_3$$

بينما معادلة القارة الإفريقية ستكون:

$$Y = -110.156 + 0.191 X_1 + 2.517 X_2 + 0.101 X_3 - 48.264\ cont_1$$

بينما ستكون معادلة القارة الأمريكية الجنوبية :

$$Y = -110.156 + 0.191 X_1 + 2.517 X_2 + 0.101 X_3 - 2.754\ cont_2$$

أي أن تأثير القارة الإفريقية هو تأثير كبير جدا ، فهو عامل حاسم في تحديد انتماء الدول الى صنف الدول الفقيرة جدا .أما اختبار Wald الذي يعاني من انتقادات مهمة ، فقد اثبت عدم معنوية المعلمات المقدرة جميعا.
ومع ذلك جاءت قيمة معامل التحديد $Cox \& Snell\ R^2$ جيدة حيث بلغت 0.749، والاهم من كل ذلك أن قيم Y المقدرة قد جاءت مطابقة تماما لقيمه الأصلية ، لذا انعدم حد الخطأ تماما ، وهذا يؤثر طبعا على دقة اختبار Wald.
إن انعدام حد الخطأ تماما يعني التخلص من مشكلة عدم تجانس تباين حد الخطأ ومشكلة الارتباط الذاتي بين قيم الحد العشوائي، ويتبقى لدينا مشكلة التعدد الخطي والتي يمكن أن تدلنا مصفوفة معاملات الارتباط بين المتغيرات المستقلة عليها نوعا ما ، حيث لم نجد ارتباطا معنويا سوى بين X_2 و X_1 أي بين عدد سنوات الدراسة وبين معدل وفيات الأطفال دون الخامسة.

جدول 5 القيم المقدرة للمتغير التابع وفق النموذج اللوجستي

القيم المقدرة للمتغير التابع	القيم الأصلية للمتغير التابع	اسم الدولة	المشاهدة	القيم المقدرة للمتغير التابع	القيم الأصلية للمتغير التابع	اسم الدولة	المشاهدة
1.00	1.00	غويانا	12	0.00	0.00	أفغانستان	1
1.00	1.00	العراق	13	0.00	0.00	بنغلاديش	2
1.00	1.00	كينيا	14	1.00	1.00	بنين	3
1.00	1.00	قرغيزستان	15	1.00	1.00	بوركينا فاسو	4
1.00	1.00	ليسوتو	16	0.00	0.00	بوروندي	5
0.00	0.00	ليبيريا	17	1.00	1.00	كمبوديا	6
0.00	0.00	ملاوي	18	1.00	1.00	تشاد	7
0.00	0.00	النيجر	19	0.00	0.00	كونغو ديموقراطية	8
0.00	0.00	طاجيكستان	20	0.00	0.00	أثيوبيا	9
0.00	0.00	زيمبابوي	21	1.00	1.00	غامبيا	10
				1.00	1.00	غانا	11

المصدر: من عمل الباحث

5- نتائج تقدير نموذج الانحدار اللوجيتي

لا يمكننا استخدام النموذج اللوجيتي ولا النموذج البروبيتي لتقدير معادلة انحدار وفقا للبيانات التي استخدمت في تقدير النماذج السابقة ، ذلك أن الانحدار اللوجيتي والبروبيتي يتطلب نوعا محددا من
البيانات، فهو يتطلب علاوة على المتغير التابع النوعي كتملك او عدم تملك الأسرة لسيارة خاصة ،

جدول 6 بيانات استخدام نماذج الانحدار اللوجيتي والبروبيتي

رقم المشاهدة	مستويات دخل الأسرة الشهري بالمليون دينار income	حجم العينة المدروسة S	عدد الأسر التي تمتلك سيارة C
1	0.5	12.00	2.00
2	1	13.00	3.00
3	1.5	15.00	2.00
4	2	12.00	5.00
5	2.5	42.00	20.00
6	3	25.00	15.00
7	3.5	63.00	45.00
8	4	78.00	60.00
9	4.5	41.00	38.00
10	5	52.00	50.00

المصدر : بيانات افتراضية

$$Y = -2.73 + 0.528\ INCOME$$
$$SE = 0.436 \quad 0.064$$
$$Z = -6.261^* + 8.211^*$$

جدول 7 الخطأ والقيم المتوقعة واحتمالية المتغير التابع في النموذج اللوجيتي

	Number	INCOME	Number of Subjects	Observed Responses	Expected Responses	Residual	Probability
LOGIT	1	.5	12	2	1.194	.806	.100
	2	1	13	3	2.051	.949	.158
	3	1.5	15	2	3.616	-1.616	.241
	4	2	12	5	4.200	.800	.350
	5	2.5	42	20	20.041	-.041	.477
	6	3	25	15	15.185	-.185	.607
	7	3.5	63	45	45.611	-.611	.724
	8	4	78	60	63.678	-3.678	.816
	9	4.5	41	38	36.198	1.802	.883
	10	5	52	50	48.226	1.774	.927

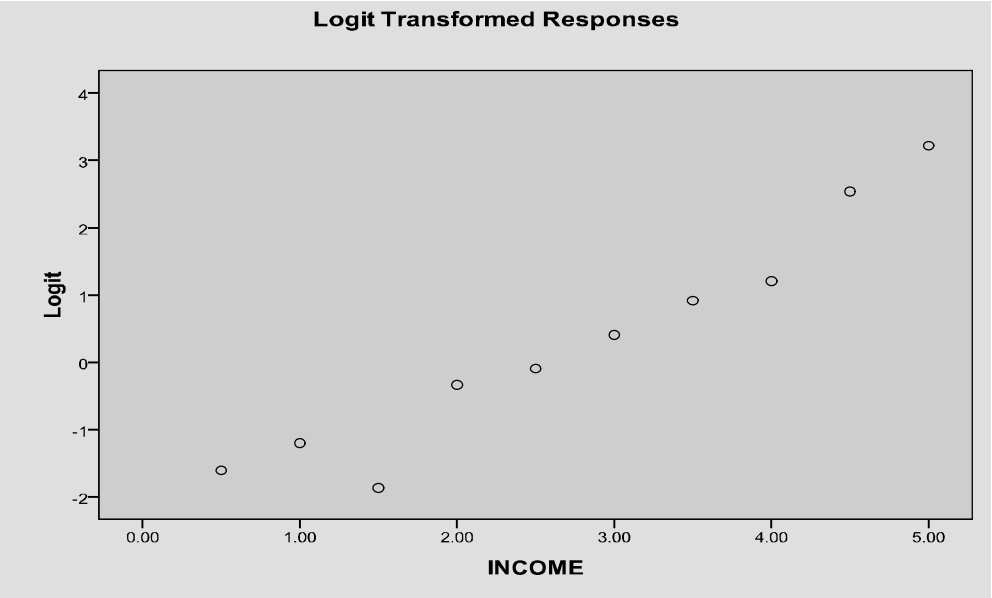
المصدر: من عمل الباحث

متغيرا مستقلا على شكل احتمال او نسبة بين متغيرين كعدد لأسر التي تمتلك سيارة خاصة،الى عدد الأسر في العينة المدروسة ،عند كل مستوى من مستويات الدخل، فمثلا باستخدام البيانات المفترضة في الجدول 6 أعلاه ، تم تقدير دالة الانحدار اللوجيتي الاتية:

ويتضح أن المعادلة اللوجيتية المقدرة تتفق مع المنطق الاقتصادي فكلما ارتفع مستوى دخل الأسرة كلما زاد احتمال امتلاكها لسيارة خاصة ، وقد اجتازت معلمات الدالية اختبارات المعنوية حسب اختبار Z ، علما أن احتمالات حدوث الحدث أي امتلاك سيارة والذي يتمثل باحتمالات القيم المقدرة للمتغير التابع قد جاءت ضمن الاحتمالات المنطقية فقد تراوحت بين حد أدنى بلغ 0.10 وبين حد أعلى بلغ 0.927 ، فكلما اقترب الاحتمال من الصفر دل ذلك على ضعف فرصة امتلاك الأسرة لسيارة ، كما في المشاهدات الخمس الأولى في الجدول 7

أعلاه، وعلى العكس كلما اقتربت قيمة الاحتمال من الواحد د ذلك على زيادة فرصة الأسرة في امتلاك سيارة ، كما في المشاهدات الخمس الأخيرة.

شكل 4 القيم المحولة للمتغير التابع حسب النموذج اللوجيتي



6- نتائج تقدير نموذج الانحدار البروبيتي

باستخدام نفس البيانات التي استخدمت في تقدير الدالة اللوجيتية ، حصلنا على نتائج تقدير الدالة البروبيتية ، وكما يأتي:

$$Y = -1.614 + 0.314\ INCOME$$
$$SE = 0.244 \quad 0.035$$
$$Z = -6.613^{*} + 8.885^{*}$$

ويتضح أن نتائج تقدير النموذج البروبيتي تقترب جدا من نتائج تقدير النموذج اللوجيتي، فالمعادلة البروبيتية تتفق مع المنطق الاقتصادي أيضا، فكلما ارتفع مستوى دخل الأسرة كلما زاد احتمال امتلاكها لسيارة خاصة ، وقد اجتازت معلمات الدالية اختبارات المعنوية حسب اختبار Z أيضا، علما أن احتمالات حدوث الحدث أي امتلاك سيارة والذي يتمثل باحتمالات القيم المقدرة للمتغير التابع قد جاءت ضمن الاحتمالات المنطقية أيضا، فقد تراوحت بين حد أدنى بلغ 0.097 وبين حد أعلى بلغ 0.936 ، فكلما اقترب الاحتمال من الصفر دل ذلك على ضعف فرصة امتلاك الأسرة لسيارة ، كما في المشاهدات الخمس الأولى في الجدول 8 أدناه ، وعلى العكس كلما اقتربت قيمة الاحتمال من الواحد د ذلك على زيادة فرصة الأسرة في امتلاك سيارة ، كما في المشاهدات الخمس الأخيرة.

جدول 8 قيم الخطأ والقيم المتوقعة واحتمالية المتغير التابع في النموذج البروبيتي

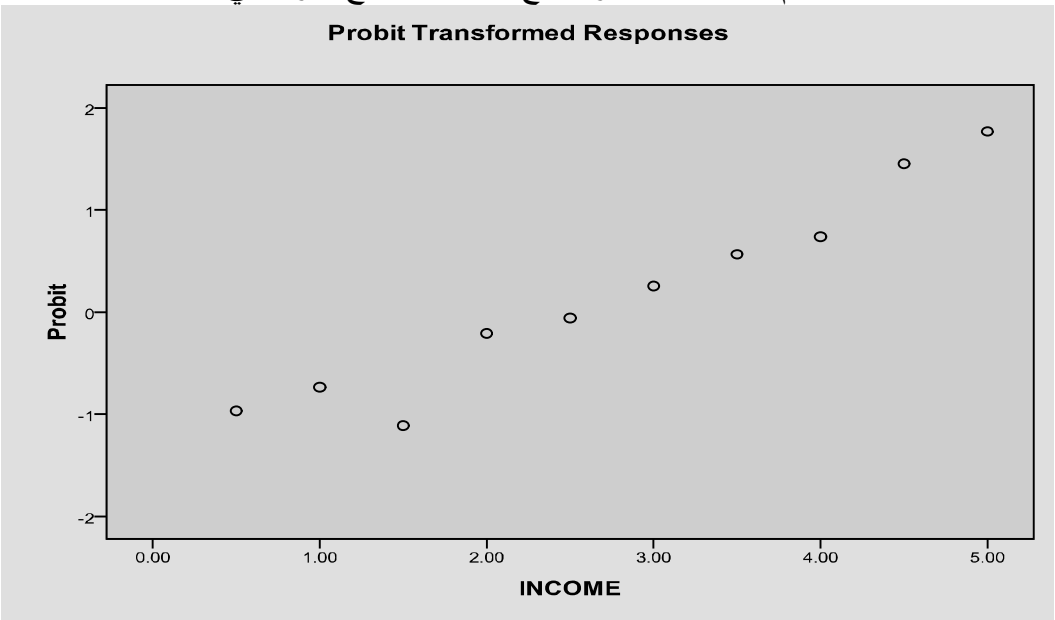
	Number	INCOME	Number of Subjects	Observed Responses	Expected Responses	Residual	Probability
PROBIT	1	.5	12	2	1.161	.839	.097
	2	1	13	3	2.105	.895	.162
	3	1.5	15	2	3.756	-1.756	.250

4	2	12	5	4.315	.685	.360
5	2.5	42	20	20.227	-.227	.482
6	3	25	15	15.135	-.135	.605
7	3.5	63	45	45.318	-.318	.719
8	4	78	60	63.526	-3.526	.814
9	4.5	41	38	36.344	1.656	.886
10	5	52	50	48.668	1.332	.936

المصدر: من عمل الباحث

شكل 5

القيم المحولة للمتغير التابع حسب النموذج البروبيتي



رابعا - الاستنتاجات والمقترحات

أ- الاستنتاجات

من خلال استعراض وتطبيق جميع النماذج التي يمكن من خلالها تقدير معادلات الانحدار التي تتضمن متغيرات تابعة نوعية، اتضحت الاستنتاجات الآتية:

1- اتضح أن استخدام نموذج المربعات الصغرى الاعتيادية OLS يؤدي إلى الحصول على نتائج غير دقيقة لان بعض القيم المقدرة للمتغير التابع كانت اكبر من الواحد أو اصغر من الصفر وهي قيم لا معنى لها لان القيم الأصلية للمتغير التابع تأخذ قيم الصفر وتعني انعدام الحدث او الصفر ويعنى حدوث الحدث.

وهذا الاستنتاج ينطبق أيضا على استخدام نموذج الاحتمال الخطي، بل أن القيم المتوقعة للمتغير التابع قد ازدادت سوءا عن ما كانت عليه في نموذج المربعات الصغرى الاعتيادية.

2- وان بقية النماذج التي تم استخدامها كنموذج الانحدار التمييزي ونموذج الانحدار اللوجستي ونموذج الانحدار اللوجيتي وأخيرا نموذج الانحدار البروبيتي، قد أدت إلى الحصول على نتائج منطقية في حالة تقدير المعادلات ذات المتغيرات التابعة النوعية.

3- يختلف استخدام النماذج الأربع المذكورة في ثانيا ، حسب طبيعة البيانات وعلى الوجه الآتي:

أ- إن كانت جميع المتغيرات المستقلة من نوع البيانات المستمرة يمكن حينئذ استخدام نموذج التحليل التمييزي.

ب- إن كانت المتغيرات المستقلة خليط بين المتغيرات المستمرة والمتغيرات المتقطعة او الفئوية، عندئذ يمكن استخدام النموذج اللوجستي.

ت- إن كان احتمال حدوث الحدث في المتغير التابع يستخرج من العلاقة بين بيانات حجم العينة المدروسة وبين عدد مرات حدوث الحدث فيها، عند كل مستوى من مستويات المتغير المستقل ، يمكن حينئذ استخدام النموذج اللوجيتي أو النموذج البروبيتي.

4 – وتختلف النماذج الأربع أيضا في الاختبارات الخاصة باختبار المعلمات واختبار النموذج ككل فالتحليل التمييزي يستخدم اختبار Wilks' Lambda لاختبار معنوية المعلمات المقدرة بينما يستخدم الانحدار اللوجستي اختبار Wald في ذلك في حين تستخدم نماذج الانحدار اللوجيتي والبروبيتي اختبار Z .

5- تتجاوز النماذج الأربع جميعا اختبارات المشاكل القياسية، كاختبارات الارتباط الذاتي والتعدد الخطي وعدم تجانس التباين ، ذلك أن بعضها يتخلص تماما من حد الخطأ أو أن يجعله صغيرا جدا وبذلك تنتهي مشاكل الارتباط الذاتي بين قيم الحد العشوائي ومشاكل عدم تجانس تباين الخطأ لعدم وجود خطأ أصلا .

6- علاوة على ذلك فان متطلبات نموذج الانحدار التمييزي ونموذج الانحدار اللوجستي فيما يتعلق بالمتغير التابع، متطلبات بسيطة جدا ، وهي أن يكون المتغير التابع متغيرا نوعيا يأخذ قيمتي الصفر والواحد.

7- غير أن نموذجي الانحدار اللوجيتي والبروبيتي ، يشترطان تشكيلة من المتغيرات، تتم معالجتها من اجل التوصل إلى قيم المتغير التابع الاحتمالي، وهي لا تتوفر دائما في الواقع العملي. فهذه النماذج تتطلب متغير مستقلا يمثل مستويات مختلفة لأخذ العينات التي تمثل المتغير الثاني وهو عدد الحالات في العينة لكل مشاهدة ، ثم متغيرا ثالثا يمثل حدوث الحدث ضمن العدد الذي يمثله المتغير الثاني، وبالتالي يتم احتساب احتمالية المتغير التابع بمعالجة المتغير الثاني والثالث وبمساعدة المتغير الأول.

8- طبقا لما جاء في رابعا وخامسا أعلاه يصبح استخدام نموذجي الانحدار التمييزي والانحدار اللوجستي أكثر واقعية وانسجاما مع طبيعة البيانات المتوفرة بشكل عام ، وأسهل بكثير من استخدام نموذجي الانحدار اللوجيتي والانحدار البروبيتي.

ب- المقترحات

1- نقترح على الباحثين الذين يقومون بتقدير دوال ذات متغيرات تابعة نوعية ، استبعاد استخدام نموذج المربعات الصغرى الاعتيادية ونموذج الاحتمال الخطي والتركيز على استخدام بقية النماذج التي تمت دراستها في هذا البحث مع مراعاة طبيعة البيانات المستخدمة وشروط ومتطلبات استخدام كل نموذج.

2- نقترح على الباحثين وطلبة الدراسات العليا القيام بمزيد من البحث في مجال طرق تقدير نماذج الانحدار ذات المتغيرات التابعة النوعية.

المصادر

1- عطية، عبد القادر محمد، الحديث في الاقتصاد القياسي بين النظرية والتطبيق ،2004 ، .www.tahasoft.com

2- http://bear.warrington.ufl.edu/shugan/7626/discriminant-analysis-output-lecture-note.pdf.

3- http://core.ecu.edu/psyc/wuenschk/MV/Multreg/Logistic-SPSS.doc.

4- http://en.wikipedia.org/wiki/Logistic_regression.

5- http://en.wikipedia.org/wiki/Logit.

6- http://unstats.un.org/unsd/demographic/products/socind/statistics.htm.

This document was created with Win2PDF available at http://www.win2pdf.com.
The unregistered version of Win2PDF is for evaluation or non-commercial use only.
This page will not be added after purchasing Win2PDF.