

# The Use of Regression, Log, and Top Models to Determine the Influencing Factors On Having Diabetes

Ahmed Hamza Abood<sup>1</sup>

Enas Abid alhafidh Mohammed<sup>2</sup>

1. Department of Anatomy, Medical Collage, Kerbala University, Iraq  
[dr.ahmedalbaghdadi@gmail.com](mailto:dr.ahmedalbaghdadi@gmail.com)

2. Statistics Department, Administration And Economics College, Kerbala University, Iraq  
[Enas.albasri@uokerbala.edu.iq](mailto:Enas.albasri@uokerbala.edu.iq)

## Article Information

**Submission date:** 1/ 11/ 2019

**Acceptance date:** 19 /11/ 2020

**Publication date:** 31/ 12/ 2020

## Abstract:

Statistical models have a major role in the health field, especially when dealing with diabetes and it is called diabetes sometimes (Diabetes Mellitus) is one of the chronic diseases and universally common, and is usually produced when the body is unable to produce sufficient quantities of the hormone insulin, or when the body is unable To benefit from insulin, which leads to high levels of sugar or blood glucose, and there are factors that lead to this disease, including age, blood sugar, social status, and educational qualification. Given that high sugar is a serious disease that affects a person without feeling it and that leads to sudden death so In this research, the traditional regression model was used and compared with the regression model of Topit, Probet, and Logit, which aims to know the factors affecting diabetes, so the researcher concluded that both the traditional least squares method and the method of repentance in the morale of the variables and the moral value of the model as a whole for both methods were converging while they were compatible Logit and Probet method in the results of the variables significance, Pseudo R<sup>2</sup> value and model significance

**Keyword:** logit model, probit model, tobit model, ordinary least square (OLS).

## الخلاصة

ان النماذج الاحصائية لها دور كبير في المجال الصحي وخاصة عند التعامل في داء مرض السكري ويسمى بمرض البول السكري أحيانا (Diabetes Mellitus) هو احد الأمراض المزمنة والشائعة عالميا، وينتج عادة عندما يعجز الجسم عن إنتاج كميات كافية من هورمون الإنسولين، أو عند عدم قدرة الجسم على الاستفادة من الإنسولين مما يؤدي إلى ارتفاع في مستويات سكر أو جلوكوز الدم وهناك عوامل تؤدي للإصابة بهذا المرض منها العمر ونسبة السكر والحالة الاجتماعية والمؤهل التعليمي وبالنظر لكون ارتفاع السكر من الأمراض الخطيرة التي تصيب الإنسان دون ان يشعر بها والتي تؤدي الى الموت المفاجئ لذا اقتضى في هذا البحث تم استخدام نموذج الانحدار التقليدي ومقارنته مع نموذج انحدار توبت وبروبت ولوجت والذي يهدف إلى معرفة العوامل المؤثرة على مرض السكري فتوصل الباحث ان كل من طريقة المربعات الصغرى التقليدية وطريقة توبت في معنوية المتغيرات وكانت قيمة المعنوية للنموذج ككل لكلا الطريقتين متقاربة في حين توافقت طريقة لوجت وبروبت في نتائج معنوية المتغيرات وقيمة Pseudo R<sup>2</sup> ومعنوية النموذج.

**الكلمات المفتاحية:** النموذج اللوجت، النموذج البروبت، النموذج التوبت، طريقة الانحدار Ols

## المقدمة

السكري أو الداء السكري أو مرض البوال السكري وغيرها (باللاتينية: Diabetes mellitus) هو من أكثر الأمراض انتشارا في جميع أنحاء العالم حيث يتسبب المرض بصورة كبيرة في وفاة الإنسان. وتتوقع منظمة الصحة العالمية بأن داء السكري سيصبح سابع عامل للوفاة في عام 2030 في ظل التقدم العلمي والتكنولوجي الذي يحصل الآن كان لا بد من استخدام النماذج والتطبيقات الإحصائية في المجالات الطبية ومن هنا جاء الاهتمام باستخدام نموذج توبت ويعتبر نموذج توبت تطبيق يقتضي تحقيق هدف يتم من خلاله دراسة العوامل المؤثرة على الإصابة بداء السكري وهو من الأمراض الخطيرة جدا والمنتشرة على مستوى العالم ومن هنا جاءت أهمية دراسة هذا المرض بصورة جدية وإجراء الأبحاث العلمية حول هذا المرض.

## أهمية البحث:

تكمن أهمية البحث في دراسة العوامل المؤثرة على الإصابة بداء السكري ومعرفة الأسباب الحقيقية للمرض للحد من انتشار المرض و النظر بالاعتبار من الباحثين و الأطباء تجاه هذا المرض من خلال الاستنتاجات والتوصيات التي يخرج بها الباحث , و توضيح أهمية المعلومات الإحصائية باستخدام نموذج توبت وبروبت ولوجت واستخدامها للتنبؤ بالقيم المستقبلية.

## بيانات البحث:

تم جمع البيانات من مركز الرعاية الصحية الأولى لناحية الجدول الغربي في مدينة كربلاء, وتم اخذ عينة بحجم (120) من المصابين بداء السكري بطريقة عشوائية و لفئات عمرية مختلفة لإجراء البحث حولها بتطبيق نموذج الانحدار التقليدي ومقارنته مع نموذج انحدار توبت وبروبت ولوجت , ومن ثم تعميم النتائج على المجتمع.

## منهجية البحث:

استخدم المنهج الاستقرائي لوصف متغيرات الدراسة والتحليل باستخدام اختباري  $f$ ,  $t$  للتوصل للنتائج وفرضيات الدراسة و ذلك لتحليل المتغيرات المختلفة في بيانات الأشخاص المصابين بداء السكري , من خلال سحب عينة وتعميم نتائجها على المجتمع المدروس .

## الجانب النظري

### التمهيد حول داء السكري :

مرض السكري معروف منذ القدم, فقد وجد وصف المرض على صفحة من صفحات البردي منذ 300 سنة قبل الميلاد عند قدماء المصريين وعرفه الإغريق . أطلق الرومان على مرض السكر ديابيتس ميليتس وهي تسمية لاتينية ترجع إلى ما قبل الميلاد حيث تعني كلمة ديابيتس الماء الجاري إشارة إلى كثرة التبول بكميات كبيرة ومعنى كلمة ميليتس العسل إشارة إلى مرور السكر مع البول.

### أنواع داء السكري:

يصنف مرض داء السكري إلى أربعة أنواع حسب تصنيف منظمة الصحة العالمية وهذه الأنواع هي [ 1 ]

### النوع الأول Type 1 Diabetes Mellitus: المعتمد على الانسولين :

وهذا النوع يسمى قديما (IDDM) أي مرضى السكر الذين يعتمدون على الانسولين في علاجهم . ومعظم مرضى هذا النوع من الأطفال والشباب صغار السن , ويصيب الشخص بين عمر (11-13) سنة ويستمر طيلة الحياة ولا يمكن الشفاء منه بشكل نهائي إذ يحتاج المصاب إلى حقن الانسولين باستمرار .

### - أسباب الإصابة بمرض السكر:

نقص في كفاءة الجهاز المناعي للجسم : تشير الدراسات إلى أن الاستعداد الوراثي والإصابة ببعض الفيروسات تؤدي إلى الإصابة بهذا النوع من المرض .

الفيروسات: يعتقد الباحثون أن سبب حدوث هذا النوع من مرض السكري هو بعض الفيروسات مثل الحصبة الألمانية والنكاف (mumps) وخاصة عند الأشخاص الذين لديهم استعداد وراثي .

اختلاف الأجناس أو السلالات: إن أكثر الأشخاص المصابين بهذا النوع هم المنحدرون من شمال أوروبا وبعض مناطق البحر المتوسط مثل (منطقة سردينيا) ويكون أقل شيوعا بين الأمريكيين والآسيويين .

تناول الأطفال حليب البقر: تشير الدراسات إلى أن الأطفال الذين تم تغذيتهم بشرب حليب البقر خلال الأيام الثمانية الأولى للولادة يكونوا أكثر عرضة للإصابة بهذا النوع من السكري بما يعادل مرة ونصف مقارنة بأولئك الذين يعتمدون على الرضاعة الطبيعية.

### النوع الثاني Type2 Diabetes Mellitus :

إن معظم المصابين بهذا النوع من السكر هم من البالغين , فأجسامهم تقاوم الانسولين , أي أنهم لا يتمكنون من الاستفادة من الإنسولين بشكل صحيح . ويجب عليهم الانتباه إلى نوعية طعامهم وممارسة التمرينات الرياضية من أجل السيطرة على مرض السكري ويمثل هذا النوع 90% من مرضى السكري والذي يميزه إن المرضى لديهم نسبة عالية من الإنسولين في دمهم وإن خلايا الجسم قد فقدت الإحساس بهذا الهرمون, ويكون أكثر المصابين في هذا النوع هم فوق سن الخامسة والخمسين ومن لديهم أحد الأقارب من الدرجة الأولى (أحد الوالدين أو الأشقاء) مصاب بالسكري وأيضا من لديهم تاريخ مرضي في ارتفاع ضغط الدم أو ارتفاع نسبة الكوليسترول في الدم, ولا يوجد فرق في الإصابة بمرض داء السكري



بين الجنسين خلال عمر 25 سنة الأولى من حياة المصاب إلا أن الميزان يتجه للانحراف باتجاه جنس الإناث بعد هذا السن, أي إن الإناث أكثر عرضة للإصابة بهذا النوع من الذكور [ 1 ]

### النوع الثالث "مرض السكري أثناء الحمل" Gestational Diabetes Mellitus :

يعتبره البعض تابعا للنوع الثاني في بعض النواحي , ويحصل بنسبة 2\_5% بين الحوامل وتزداد النسبة مع

تقدم العمر ويمكن أن يستمر أو يختفي بعد الولادة . ويحتاج إلى إشراف طبي خلال فترة الحمل . [ 1 ]  
أنواع أخرى من مرض السكري:

رغم أن معظم الحالات تصنف ضمن النوع الأول أو الثاني, إلا إن مرض السكري يصنف بصفة ثانوية نتيجة لوجود علة مرضية محددة مسبقاً ومحدثة للمريض عن طريق شل نشاط الخلايا التي تفرز الإنسولين في البنكرياس أو إتلاف تلك الخلايا . , ومن هذه العلة أو الأسباب : [ 2 ]

- وراثي, نتيجة خلل في الصبغة الوراثية يؤدي إلى نقص في تكوين وإفراز الإنسولين .
- مرض يحطم البنكرياس .
- أمراض الغدد الصماء , مثل متلازمة كوشينج, Cushing's syndrome فرط إفراز الغدة الدرقية Hyperthyroidism وضخامة النهايات (الأطراف . Acromegaly )
- نتيجة أخذ العقاقير مثل هرمون الغدة الدرقية Thyroid Hormone وحمض النيكوتينيك Nicotinic Acid والكورتيزونات steroids .
- نتيجة الالتهابات الفيروسية التي تؤثر على البنكرياس مثل التهاب فيروس سايتوميغالو cytomegalo virus والحصبة الحلقية Congenital Rubella أي طفل ولد مصاباً من الأم أثناء الحمل
- الحالة النفسية: كثير من المصابين يزعمون إنهم أصيبوا بالمرض بعد تعرضهم (لهزة نفسية) إلا إن الإصابة بالمرض كانت لديهم (كامنة) أي قبل تعرضهم (لهزة النفسية) فالهزة النفسية زادت أعراض المرض تفاقماً إلى حد كبير ولكنها لا تكون السبب وراء الإصابة بالمرض [ 2 ] .

### تعريف العوامل المؤثرة على الإصابة بداء السكري: [ 3 ], [ 4 ]

- 1- الجنس: لا يوجد فرق في الإصابة بمرض السكري بين الجنسين خلال عمر 25 سنة الأولى من حياة المصاب, إلا إن الميزان ينحرف باتجاه جنس الإناث بعد هذا السن, أي إن الإناث أكثر عرضة للإصابة بهذا المرض من الذكور .
- 2- العمر : تزداد خطورة الإصابة بالنوع الثاني من مرض السكر , كلما زاد عمر الشخص, خاصة بعد سن 30 عام مع قلة النشاط و زيادة الوزن .
- 3- الوزن: زيادة وزن الجسم هو احد أهم العوامل المؤثرة في الإصابة \_معظم المصابين بداء السكر من النوع الثاني يكون وزنهم أعلى من المعدل الطبيعي بكثير (حوالي 8 أشخاص من كل 10 أشخاص مصابين بالمرض يعانون من زيادة الوزن) . [ 3 ]

- 4- نسبة السكر: يمثل هذا المتغير نسبة السكري الموجود في الدم وقد تم قياسها بشكل دقيق لجميع مرضى العينة، وتتراوح نسبة السكر الطبيعي عند الانسان العادي من 60 الى 150ملغ، وان اي زيادة نقصان في هذه النسبة يسبب خلل في انظمة الجسم.
- 5- الحالة الاجتماعية: يمثل هذا المتغير حالة مريض السكري الاجتماعية، هل هو متزوج أم أعزب وغيرها، حيث إن مرضى السكري يصابون بنوع من الحالة النفسية و عدم القدرة على تحمل الضغوط الكبيرة، كذلك فإن ألعازبين يتخوفون من عدم قدرتهم على تكوين أسرة ناجحة وممارسة حياتهم بشكل طبيعي.
- 6- المؤهل العلمي: للدلالة على المستوى العلمي الحاصل عليه مريض السكري، إذ إن لذلك أهمية كبيرة لان مستوى المريض المعرفي يساعده في معرفة ما يعود على صحته بالفائدة.
- 7- الوظيفة: مرضى السكري بحاجة إلى رعاية وعناية جيدتين، كذلك تخفيف الضغوط الحياتية قدر الإمكان، لهذا فإن نوع الوظيفة التي يعمل بها مريض السكري حتما تختلف ضغوطها ونوعيتها من مريض إلى آخر، فبعض الوظائف تحتاج إلى مجهود بدني و نفسي كبيرين مما قد لا يعود على صحة المريض بالفائدة.
- 8- وجود وعدم وجود أمراض أخرى: يمثل هذا المتغير تأثير الأمراض الأخرى على مرضى السكري، وعلى وجه الخصوص ارتفاع ضغط الدم حيث إن العلاقة بين السكر وارتفاع ضغط الدم وثيقة للغاية، ومثبتة علميا في العديد من الأبحاث، حيث أشارت هذه الأبحاث إلى إن حوالي 50\_60% من مرضى السكر يصابون بارتفاع ضغط الدم حيث تتشابه الأسباب المؤدية إلى المرضين بشكل كبير.
- 9- السكن: تم اخذ هذا المتغير لمعرفة هل هناك علاقة بين الريف والمدينة في انتشار مرض السكري، حيث إن سكان المدن يحصلون على الوجبات السريعة والأغذية العالية بالسعرات الحرارية، بعكس الريف حيث يكون هناك حد اقل مما هو في المدينة

#### تمهيد عن نماذج الانحدار:

في هذا الفصل سيتم التطرق إلى الأساليب والنماذج الإحصائية و طرق التحليل الإحصائي التي تستخدم في تحليل بيانات البحث فيتم استخدام نموذج توبت لتحليل بيانات البحث باستخدام البرنامج الإحصائي Stata .

#### نماذج الانحدار: [ 5 ], [ 6 ]

تعتمد العديد من الدراسات والبحوث على اساليب متطورة من اجل الحصول على نتائج تتصف بالفعالية والدقة العاليتين، وقد كان لعلم الاحصاء وفروعه المرتبطة به الاثر الكبير في بناء النماذج الرياضية وتحليل البيانات من خلالها وصولا للقرارات السليمة يعد تحليل الانحدار اهم فروع علم الاحصاء، والذي يهتم ببناء العلاقة الرياضية بين متغير الاستجابة والمتغيرات التوضيحية وتمثل هذه العلاقة كتركيبية خطية تدعى معادلة الانحدار اذ دقتها تعتمد على صحة تقدير معالمها والتي تشترط توفر شروط التحليل، وتعتبر طريقة المربعات الصغرى الاعتيادية احد اكثر طرق التحليل استخداما وتتصف هذه الطريقة بكفاءتها العالية بتقدير معالم نموذج الانحدار عند توفر فروض التحليل،.

#### نموذج توبت (Tobit Model): [ 1 ]، [ 7 ]

نحن نعلم إن تحليل الانحدار هو احد الوسائل الإحصائية التي تصف شكل العلاقة ما بين المتغيرات التفسيرية و المتغير المعتمد، فإذا كانت قيم المتغيرات التفسيرية معلومة فإن تحليل الانحدار يساعدنا في التوقع (التنبؤ) بقيمة المتغير التابع لذلك فإن اختيار النموذج الملائم للبيانات المتوفرة هي من ضروريات هذا التحليل، وبهدف الوصول إلى نتائج دقيقة





$b_i$ : هي معالم النموذج ( $i=0,1,2,\dots,k$ )

$x_i$ : هي المتغيرات التفسيرية ( $i=0,1,2,\dots,k$ )

$e_i$ : حد الخطأ العشوائي .

تؤخذ قيمة  $a$  (نقطة التقيد) مساوية إلى أي قيمة محددة حسب بيانات الدراسة، وفي الظاهرة قيد الدراسة كانت نقطة التقيد مساوية إلى الصفر لذا فإن شكل الدالة سيأخذ الشكل التالي:

$$y = 0 \quad \text{if } y^* \leq 0$$

$$y = y^* \quad \text{if } y^* > 0 \quad a = 0$$

تقدير دالة نموذج توبت: [1], [9], [10],

توجد عدة طرق لتقدير دالة نموذج توبت، منها طريقة المربعات الصغرى، و طريقة العزوم (Maximum likelihood)، ونستخدم هنا طريقة العزوم. ومن خلال صيغة دالة نموذج توبت المبينة أعلاه سوف يتم التعامل في هذا البحث مع جزأين من البيانات الجزئية الأولى من المشاهدات عندما تكون قيم المشاهدات مساوية إلى الصفر والجزء الثاني من تلك المشاهدات عندما تكون مساوية إلى كميات موجبة، لذا فلكل جزء من هذه المشاهدات دالة رياضية خاصة بها ومن ضرب هذه الدوال مع بعضها البعض سنحصل على دالة مختلطة لنموذج توبت .

**أولا الخطية: Linearity**

عندما  $y^* > 0$  أي عندما تكون المشاهدات مساوية إلى كميات موجبة :

$$\text{if } y_i = \beta_i x_i + e_i \dots \dots \dots (2)$$

$$e_i = y_i - \beta_i x_i \dots \dots \dots (3)$$

دالة العزوم (Likelihood Function) عندما تكون المشاهدات اكبر من الصفر تعطى من الصيغة الآتية :

$$Li = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{e_i^2}{2\sigma^2}}$$

$$Li = \frac{1}{\sigma\sqrt{2\pi}} \exp 2\sigma^{-2}(y_i - \beta_i x_i)$$

$$= \frac{1}{\sigma} \phi \left( \frac{y_i - \beta_i x_i}{\sigma} \right) \dots \dots \dots (4)$$

ثانيا : الاحتمالية عندما  $y_i \leq 0$  عندما تكون المشاهدات مساوية الى الصفر :

$$\Pr(y_i = 0) = \Pr(y^* \leq 0)$$

$$\Pr(e_i \leq x_i \beta) = \Pr\left(\frac{e_i}{\sigma} \leq -\frac{x_i \beta}{\sigma}\right) \dots \dots \dots (5)$$

$$= \phi\left(-\frac{x_i \beta}{\sigma}\right) = 1 - \phi\left(\frac{x_i \beta}{\sigma}\right) \quad , \quad \phi \text{ هي دالة تجميعية}$$

حيث أن  $\phi$  (p.d.f) هي دالة  $\phi$

ومن خلال ضرب دالة الكثافة الاحتمالية والدالة التجميعية نحصل على دالة ومن خلال ضرب دالة الكثافة الاحتمالية والدالة التجميعية نحصل على دالة مختلطة من نموذج توبت.

$$P(y) = \left[ \frac{1}{2\pi\sigma^2} e^{-\frac{(y_i - XB)^2}{2\sigma^2}} \right] \left[ 1 - \phi \left( \frac{\alpha - XB}{\sigma} \right) \right] \dots \dots \dots (6) \quad \text{فان المعادلة تصيح كالتالي :}$$

اذا  $a=0$

$$p(y) = \left[ \frac{1}{2\pi\sigma^2} e^{-\frac{(y_i - XB)^2}{2\sigma^2}} \right] \left[ 1 - \phi \frac{-XB}{\sigma} \right]$$

ويمكن التعبير عن الدالة اعلاه بشكل اخر  $P(y) = \left[ \frac{1}{\sigma} \phi \left( \frac{y_i - XB}{\sigma} \right) \right] \left[ 1 - \phi \left( \frac{-XB}{\sigma} \right) \right]$

ومن خلال ما ذكر اعلاه يمكن تقدير معالم نموذج تويت بالاعتماد على طريقة الامكان الاعظم :

وللحصول على لوغاريتم دالة العزوم ندخل دالة الترجيح الاعظم لطرفي المعادلة:

$$L = \prod_{i=1}^n p(y_i) = \frac{1}{\sigma^n} \prod_{i=1}^n \left[ \phi \left( \frac{y_i - X\beta}{\sigma} \right) \right] [1 - \phi]$$

بإدخال اللوغاريتم الطبيعي لجميع أطراف المعادلة نحصل على الآتي:

$$\ln p(y) = \sum_{i=1}^n \left[ -\ln \sigma + \ln \phi \left( \frac{y_i - XB}{\sigma} \right) \right] + \ln \left[ 1 - \phi \left( \frac{-XB}{\sigma} \right) \right] \dots \dots \dots (7)$$

ويمكن استخدام الطرق العددية لحساب معاملات نموذج تويت

### نموذج LOGIT: [8], [9], [10]

يتم استخدام هذا النموذج اذا كان المتغير التابع بمثابة متغير صوري يأخذ قيمتين فقط هما 0,1 يطلق عليه Variable Binary Dummy Dependent وهو متغير وصفي Qualitative وتكون المتغيرات المستقلة وصفية او كمية Quantitative , ان نموذج logit يعالج تباين المتغير التابع الذي يتغير بتغير المتغير المستقل وان تباين الخطأ لا يتوزع توزيعاً طبيعياً وان القيم المقدرة لا يمكن تفسيرها بوصفها احتمالات ذلك لأن قيمتها لا تتراوح بين الصفر والواحد ويعتمد على دالة التوزيع التراكمي Cumulative Logistic وان العلاقة بين الاحتمال والمتغير التفسيري تعد علاقة غير خطية وتأخذ العلاقة الصيغة التالية :

$$p_i = E \left( y = \frac{1}{x_i} \right) = b_0 + b_1 x_1 + \dots \dots \dots (8)$$

$$p_i = \frac{1}{1 + e^{(b_0 + b_1 x_1)}}$$

$$p_i = \frac{1}{1 + e^{-z_i}} \dots \dots \dots (9)$$



والمعادلة (9) تشير الى دالة التوزيع التراكمي :

$$1 - p = \frac{1}{1 + e^{-zi}}$$

اذ ان  $p$ : احتمال النجاح

$1-p$ : احتمال الاخفاق

$e \cdot \frac{p}{1-p}$ : 2.7 تمثل نسبة الترجيح للحدث محل الاهتمام *odds*

وبذلك يمكن كتابة معادلة الانحدار للعدد من المدخلات:

$$\ln\left(\frac{p}{1-p}\right) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k \dots \dots \dots (10)$$

تستخدم طريقة الامكان الاعظم لحساب معاملات *logit* وتهدف هذه الى تعظيم لوغاريتم الاحتمال *log likelihood* الذي يعكس مدى امكانية او احتمال ان تكون تلك القيم المشاهدة للمتغير التابع في الامكان توقعها او تنبؤها من خلال المتغير او المتغيرات المستقلة, ويلاحظ ان تقديرات الامكان الاعظم انها طريقة تكرارية تبدأ بقيمة اولية لما ينبغي ان تكون معاملات *logit* ثم تحدد هذه الطريقة اتجاه ومقدار التغير في معاملات *logit* الذي يزيد عن لوغاريتم الاحتمال. كما يمكن تفسير معاملات الانحدار في هذا الانموذج بدلالة الاحتمالات حيث ان زيادة المتغير المستقل بوحدة واحدة ستزيد  $L$  او لوغاريتم معامل الترجيح *Log odds* بأن يكون المتغير التابع يساوي واحد بمقدار احتمال معين هو معامل المتغير المستقل. كما ان نموذج *logit* يتحايل على مشكلة عدم ثبات تجانس تباين الخطأ العشوائي لانه يفترض توزيعا مختلفا لهذا الحد , ويكون معامل التحديد لأنموذج *logit* على النحو الاتي :

$$R^2 : R^2_{Mf} = 1 - (LLF_{\beta} / LLF_0)$$

$$LLF_{(\beta)} = \sum_{t=1}^T \{ [Y_t \ln \hat{Y}_t] + [(1 - Y_t) \ln(1 - \hat{Y}_t)] \}$$

نموذج **PROBIT**: [12], [10], [11],

هذا الانموذج يشبه انموذج *Logit* في طبيعة المتغير التابع اذ يكون متغير نوعي يأخذ صفتين هما الصفر والواحد ويعتمد على دالة التكتيف الاحتمالي  $f(X_t B)$  ودالة التوزيع التراكمي  $F(X_t B)$  واذا كان توزيع طبيعي *Normal Distribution* بمتوسط  $\mu=0$  وتباين  $\sigma^2=1$  فان دالتي *PDF, CDF* تكون على النحو الآتي :

$$f_t = f(x_t B) = \exp\left[\frac{-0.5(X_t B)^2}{\sqrt{2\pi}}\right]$$

$$F_t = F(X_t B) = \int_{-\infty}^{X_t B} \exp\left[\frac{-0.5(X_t B)^2}{\sqrt{2\pi}}\right] dB$$

وتأخذ دالة الاحتمال اللوغاريتمي  $LLF(B)$  لأنموذج بروبوت الشكل الآتي:

$$LLF(\beta) = \sum_{t=1}^T (X_t \ln F(X_t \beta) + (1 - X_t) \ln[1 - F(X_t \beta)])$$

ولو فرضنا العلاقة الآتية:

$$I = b_0 + b_1 X \dots \dots \dots (11)$$

إذا افترضنا ان  $I$  تمثل 1 في حال حدوث الحدث و 0 في حال عدم حدوثه و  $X$  متغير مستقل . وعلى فرض ان مستوى من  $I$  يطلق عليه  $I^*$  فإذا زاد  $I$  أعلى من  $I^*$  فإنه سيحدث الحدث المرغوب أي ان  $I=1$  , وإذا كان التوزيع طبيعي مع نفس المتوسط والتباين نستطيع ان نقدر معاملات المعادلة (11) وتحسب على النحو الآتي

$$P\left(Y = \frac{1}{Y}\right) = P(I^* \leq I) = P(Z \leq B_0 + B_1 X) \\ = (F(B_0 + B_1 X)) \dots \dots \dots (12)$$

اذ ان :  $P(Y = \frac{1}{X}) =$  متوسط احتمال حدوث الحدث عندما تكون  $X$  قيمة معطاة.  
 $Z =$  متغير طبيعي معير Standard Normal.

$$F(I_i) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{I_i} e^{-z^2/2} dz \\ = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\beta_0 + \beta_1 X_i} e^{-z^2/2} dz$$

وهنا احتمالية حدوث الحدث تقاس بمساحة المنحنى الطبيعي القياسي من  $-\infty$  الى  $I$  وللحصول على قيمة او كذلك المعلمات  $\beta_0, \beta_1$  نأخذ معكوس الدالة (11)  
 $I = F^{-1}(I) = F^{-1}(pi)$

إذا  $F^{-1} =$  معكوس الدالة CDF.

عند اجراء مقارنة بسيطة بين أنموذج Probit و Logit نرى ان معظم التطبيقات تشير الى ان النتائج في الأنموذجين تكون مقارنة ومتماثلة نوعيا لكن الفرق الرئيس هو ان Logit لديه زيول وامتدادات تضخمية في التوزيع وان  $Pi$  (الاحتمالية الشرطية) فيه تقترب من الصفر والواحد بمعدل ابطأ من Probit, لكن كثيرا من الباحثين يفضلون Logit لأنه ابسط رياضيا نسبيا. وقيم المقدرات غير قابلة للمقارنة والسبب Logistic قياسي , والتوزيع الطبيعي للأنموذجين هو متوسط صفر وتباين يختلف عن الواحد والمعيار الطبيعي هو  $\frac{\pi^2}{3}$  لتوزيع Logistic لذلك اذا ضربنا مقدر Probit ب 1.81 نحصل على مقدر Logit وبالعكس اذا ضربنا مقدر Logit ب 0.55 فسوف نحصل على مقدر Probit لكن Amemiya اقترح ضرب مقدر Logit ب 0.625 للحصول على تقدير مناظر اكثر لمقدر Probit ب 1.6 للحصول على مقدر Logit , وعليه اقترح Amemiya ان معاملات Logit & Probit بينهما العلاقة الآتية :

$$B_{LPM} = 0.025 B_{logit} + 0.5 \quad (\text{مع الثابت})$$

علما ان أنموذج Probit يتخلص هو الاخر من مشكلة عدم تجانس تباين الخطأ العشوائي باعتماده دالة التوزيع الاحتمالي التراكمي .

## الجانب العملي التطبيقي

### تمهيد

في هذا الفصل سيتم التطرق الى الجانب التطبيقي للبحث و ذلك باستخدام البرنامج الاحصائي STATA والذي يمتاز بسهولة التعامل لاحتوائه على العديد من الادوات التي تساعد كثير من الباحثين في عملية التحليل والتوصل الى النتائج الدقيقة

نتائج البيانات باستخدام نموذج توبت :

جدول (1) يبين نتائج معنوية المتغيرات حسب نموذج توبت :

Tobit regression	Number of obs	=	120
	LR chi2(8)	=	16.59
	Prob > chi2	=	0.0347
Log likelihood = -705.19812	Pseudo R2	=	0.0116

D	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
A	-1.245115	19.69028	-0.06	0.950	-40.25889 37.76866
B	-.3384686	.5788674	-0.58	0.560	-1.48542 .808483
C	.9389909	.6369426	1.47	0.143	-.3230292 2.201011
E	12.58961	14.2996	0.88	0.381	-15.74322 40.92243
H	-1.440701	3.57178	-0.40	0.687	-8.517725 5.636323
I	-63.20714	19.42376	-3.25	0.002	-101.6928 -24.72145
F	-12.34191	4.600684	-2.68	0.008	-21.45758 -3.226247
G	17.28174	12.92907	1.34	0.184	-8.335547 42.89903
_cons	255.0497	73.89572	3.45	0.001	108.6348 401.4646
/sigma	89.61611	5.822931			78.07872 101.1535

تم الاعتماد على متغير نسبة السكري باعتباره المتغير المعتمد واعتبار المتغيرات الاخرى هي المستقلة. وقد تم اختيار قيمة  $\alpha=0.05$  كمستوى معنوية معتمد وكان التفسير كالاتي :

وجد ان متغير السكن (i) والمؤهل العلمي (f) لهما تاثير معنوي للاصابة بداء السكري خلافا لبقية المتغيرات المتمثلة بالوزن (c) والعمر (b) والجنس (a) و الوظيفة (g) والحالة الاجتماعية (e) و وجود امراض اخرى (h) كانت نتائجهم غير معنوية لانها اكبر من مستوى الثقة المساوي ل 0.05.

وقد كان النموذج معنوي كون قيمة P-value للنموذج مساوية الى 0.0347 وهي اقل من 0.05 و قيمة

. Pseudo R<sup>2</sup>=0.0116



نتائج البيانات باستخدام نموذج لوجت :

الجدول (2) يبين نتائج معنوية المتغيرات حسب نموذج لوجت :

Logistic regression for grouped data		Number of obs = 840	
Log likelihood = -346.44219		LR chi2(6) = 107.49	Prob > chi2 = 0.0000
		Pseudo R2 = 0.1343	

  

_outcome	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]
A	1.132838	.5740124	1.97	0.048	.0077943 2.257882
B	-.1444288	.0183342	-7.88	0.000	-.1803632 -.1084944
C	.4420372	.1030326	4.29	0.000	.240097 .6439774
E	0 (omitted)				
F	-.1403763	.0815084	-1.72	0.085	-.3001298 .0193773
G	-2.431331	.4153537	-5.85	0.000	-3.245409 -1.617253
H	-.0237471	.5073807	-0.05	0.963	-1.018195 .9707007
I	0 (omitted)				
_cons	-22.43751	7.974514	-2.81	0.005	-38.06727 -6.807754

من النتائج الموضحة في الجدول اعلاه والخاصة بنموذج لوجت و بالاعتماد على متغير نسبة السكري كمتغير معتمد و بقية المتغيرات الاخرى مستقلة وباخذ مستوى معنوية قدره  $\alpha=0.05$  كان التفسير كالاتي : وكان متغير الجنس (a) ومتغير العمر (b) و متغير الوزن (c) و متغير الوظيفة (g) كانت ذات تاثير معنوي على الاصابة بداء السكري بخلاف المتغيرات الاخرى المذكورة وهي (المؤهل العلمي, و وجود امراض اخري) اما معنوية النموذج الكلية فكانت مساوية الى 0.000 وهي اقل من 0.05 فضلا عن قيمة  $Pseudo R^2$  والتي تساوي 0.1343 .



نتائج البيانات باستخدام طريقة المربعات الصغرى :

الجدول (4) يمثل نتائج معنوية المتغيرات حسب طريقة المربعات الصغرى التقليدية :

. regress D A B C E F G H I

Source	SS	df	MS	Number of obs		
Model	143475.319	8	17934.4149	F( 8, 111)	-	2.10
Residual	949784.148	111	8556.61394	Prob > F	-	0.0420
Total	1093259.47	119	9187.05434	R-squared	-	0.1312
				Adj R-squared	-	0.0686
				Root MSE	-	92.502

  

D	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
A	-2.030894	20.3116	-0.10	0.921	-42.27969	38.2179
B	-.3384729	.5975267	-0.57	0.572	-1.522512	.8455662
C	-.9549477	.6573033	1.45	0.149	-.3475427	2.257438
E	12.82079	14.75894	0.87	0.387	-16.42503	42.06661
F	-12.28784	4.748711	-2.59	0.011	-21.69773	-2.877948
G	17.79339	13.33718	1.33	0.185	-8.635134	44.22191
H	-1.472236	3.686794	-0.40	0.690	-8.777865	5.833393
I	-63.2609	20.04981	-3.16	0.002	-102.9909	-23.53088
_cons	253.73	76.26763	3.33	0.001	102.6006	404.8594

من النتائج اعلاه والتي تمثل نتائج البيانات باستخدام طريقة المربعات الصغرى و بالاعتماد على متغير نسبة السكري كمتغير معتمد و المتغيرات الاخرى مستقلة وبمستوى معنوية  $\alpha=0.05$  تبين مايلي :

ان جميع المتغيرات لا تؤثر معنويا في الاصابة بداء السكري باستثناء متغير السكن (i) والحالة الاجتماعية كان تشيرهما معنويا .

جدول يوضح نتائج المفاضلة بين نماذج الانحدار اعلاه :

جدول (5) يبين نتائج المفاضلة بين نماذج الانحدار المستخدمة في هذا التحليل

المتغيرات	نموذج توبت	نموذج لوجت	نموذج بروبوت	النموذج التقليدي
معنوي (السكن، المؤهل العلمي)	معنوي (العمر، الجنس، الوزن، الوظيفة)	معنوي (العمر، الجنس، الوزن، الوظيفة)	معنوي (العمر، الجنس، الوزن، الوظيفة)	معنوي (السكن، المؤهل العلمي)
النموذج	0.0347	0.000	0.000	0.0420
Pseudo R <sup>2</sup>	0.0116	0.1343	0.1343	-



## الاستنتاجات:

- ان اهم الاستنتاجات التي تخص الاسلوب الاحصائي ومؤشراته التي تم تطبيقها على عينة الدراسة هي:-
- 1- ان متغير السكن (i) والمؤهل العلمي (f) لهما تأثير معنوي للاصابة بداء السكري خلافا لبقية المتغيرات المتمثلة بالوزن (c) والعمر (b) والجنس (a) و الوظيفة (g) والحالة الاجتماعية (e) و وجود امراض اخرى (h) كانت نتائجهم غير معنوية لانها اكبر من مستوى الثقة المساوي ل 0.05 من النتائج المبينه اعلاه التي تخص نموذج توبت.
  - 2- استنتجنا من تحليل البيانات باستخدام نموذج لوجت ان متغير الجنس (a) والعمر (b) والوزن (c) والوظيفة (g) لهم تأثير معنوي بخلاف المتغيرات الاخرى, وهي المؤهل العلمي (f) و وجود امراض اخرى (h) وكانت معنوية النموذج مساوية الى 0.000 وهي اقل من قيمة  $\alpha=0.05$  فضلا عن قيمة  $Pseudo R^2=0.1343$ .
  - 3- كانت النتائج المستحصلة من نموذج بروبوت توضح ان متغير الجنس (a) و العمر (b) و الوزن (c) و الوظيفة (g) ذات تأثير معنوي خلافا للمتغيرات الاخرى وهي المؤهل العلمي (f) و وجود امراض اخرى (h) تأثيرهما غير معنوي, ونلاحظ ان نموذجي لوجت و بروبوت ذات نتائج متشابهة .
  - 4- اشتركت كل من طريقة المربعات الصغرى التقليدية وطريقة توبت في معنوية المتغيرات وكانت قيمة المعنوية للنموذج ككل لكلا الطريقتين متقاربة في حين توافقت طريقة لوجت وبروبوت في نتائج معنوية المتغيرات وقيمة  $Pseudo R^2$  ومعنوية النموذج .

## التوصيات:

- 1- نوصي بمواصلة اجراء البحوث الاحصائية وغير الاحصائية حول مرض داء السكري نظرا لكثرة انتشاره وخطورته من اجل الاحاطة به اكثر للوصول الى ناجعة للحد منه وعلاجه .
- 2- كما نوصي بإجراء بحوث احصائية حول بقية العوامل المسببة مثل (العامل الوراثي , الصدمات المفاجئة , الحالة النفسية , وغيرها من العوامل التي لم نتناولها).

## Conflict of Interests.

There are non-conflicts of interest .

## References

- 1- بحث تخرج منشور ,استخدام نموذج توبت في تحليل العوامل المؤثرة على الاصابة بضغط الدم لمرض الفشل الكلوي ,اليمين ابكر عبد الله ,رحاب يحيى عبد الكريم ,سارة احمد عيسى , سبتمبر 2016م .
- 2- الاقدام السكرية , أ د حسن بن علي الزهراني 1427هـ - 2006م مدينة الملك عبد العزيز للعلوم والتقنية ,المملكة العربية السعودية .
- 3- الامراض النفسية وعلاقتها بمرض العصر (السكر) الدكتور جاسم محمد عبد الله محمد المرزوقي دار العلم والايامن للنشر والتوزيع , الطبعة الاولى , تاريخ النشر 2008 .
- 4- داء السكري هل السكري مرض يمكن منعه وعلاجه؟ اعداد الدكتور موسى بن تسمحان العنزي , الطبعة الاولى 1431هـ - 2010م , الرياض ,المملكة السعودية العربية .
- 5- حسين ,مجيد علي ,سعيد ,عفاف عبد الجبار , (1998م) , "الاقتصاد القياسي بين النظرية و التطبيق " , دار وائل للنشر , عمان , الأردن .
- 6- بحث تخرج منشور , دراسة بعض العوامل المؤثرة في حدوث مرض داء السكري في محافظة الديوانية - دراسة احصائية- أورهان مؤيد محمد يونس ,حسين هلال عبد الحسين , 1438هـ-2017م .
- 7- السكري / <https://ar.m.wikipedia.org/wiki>
- 8-مجلة العلوم الزراعية , قياس كفاءة حقول الابقار باستخدام نماذج الاستجابة النوعية - اسكندر حسين علي ,قسم الاقتصاد الزراعي- كلية الزراعة - جامعة بغداد-2016 .
- 9-Gramer ,J.S (1986)."Econometric applications of Maximum Likelihood methods".  
Cambridge University press .
- 10.S.Madheswaran , "Probit, Logit and Tobit Models", Ineitude for Social and Economic Change , **Bangalor (2001)** .
- 11- LM Tierney ,MA Papadkis,"Current medical diagnosis& treatment", New York: Lang Medical Books L McGraw-Hill, International Edition (2002).
- 12- Rajulton Fernando; "Logit, Probit and Tobit Models for Categorical and Limited Dependent Variables ", March 23 Statistics and Data Series at Western (2012).