The Use of Logistic Regression in studying the Causes of Death in Preterm babies in Babylon

استعمال الانحدار اللوجستي لدراسة اسباب وفيات الاطفال الخدج في محافظة بابل

أ.د عدنان كريم نجم الدين ميثم عبد الوهاب صالح
جامعة كربلاء / كلية الادارة والاقتصاد / قسم الاحصاء
بحث مستل من رسالة ماجستير في الاحصاء

ملخص:

تتركز الفكرة الأساسية في هذه الدراسة حول استخدام طرق إحصائية متقدمة لدراسة أهم العوامل المؤثرة على وفيات الأطفال الخدج في محافظة بابل من خلال بيانات حقيقية اخذت من دائرة صحة بابل ، وتناول هذا البحث اسلوب الانحدار اللوجيستي ثنائي الاستجابة لتمثيل الدالة المقدرة .وهو اسلوب انحدار وصفي في حال متغير الاستجابة الوصفي وتقدير معالم الانموذج بطريق الامكان الاعظم M.L.E ، ومن ثم تحسين الانموذج عن طريق التحليل العاملي وبطريقة تحليل المركبات الرئيسية وأشارت النتائج الى ان أهم العوامل المؤثرة على حياة الاطفال الخدج هي : مدة الحمل ,وزن الطفل ,ناتج الولادة (توأم او مفرد) ,سبب الرقود ,عمر الام ,معقم او غير معقم .

Abstract:

The main idea in this study is the use of advanced statistical methods to study the most important factors affecting the mortality of preterm infants in Babylon Governorate through the real data taken from the Babylon Health Directorate . This study dealt with the Logistic regression method of the binary response to represent estimated function. It is a descriptive regression method in the case of a descriptive response variable the parameters of the study will be estimated by the Maximum Likelihood Estemator , Thus improving the model by means of a Factor analysis and by Principle Components analysis . One of the main finding of the study is determining many factors which leads to infants death in Babylon Governorate such as:, The pregnancy period, child weight, Birth process outcome(Single or Twin), Pathological case of baby, mother's age, aseptic neonate or Septic neonate

المقدمة :

ان تطوير وتوسيع واقع الخدمات الصحية يعد ركناً اساسياً لخطط التنمية الاقتصادية والاجتماعية والتي تهدف الى انقاذ المجتمع من الاوبئة والامراض والاهتمام بتطوير الموارد البشرية. كما انها تاتي استجابة لكثير من المشكلات الاجتماعية والبيئية التي يواجهها الفرد.

وإن المعالجة الطبية الفعالة للكثير من الامراض من شأنها ان ترفد المجتمع بعناصر منتجة تساهم في تطور الناتج القومي والحركة التنموية في البلد، فضلا عن توفير افضل الخدمات للافراد كافة. وتحتل دراسة الوفيات وخاصة وفيات الاطفال مكانة

خاصة في مجال الابحاث الصحية والسكانية اذ انها تعكس الصورة الحقيقية للمستوى الصحي العام والجانب التنموي اذ تمثل الجانب السلبي للنمو السكاني .

ونظراً للتقدم العلمي والتكنولوجي الذي حصل في الكثير من المجالات البحثية وبالأخص في مجال الدراسات والبحوث الاحصائية فقد اصبح الإحصاء علم له اثره الواضح في خدمة الانسان وتطويره في شتى الميادين وذلك من خلال استخدامه في الاستدلال على سائر العلوم وتطويرها، وقد بدأت تطبيقات الاحصاء ومجالات استخدامه تنمو وتتقدم عملياً في كل حقول المعرفة وبشكل واضح باعتباره من العلوم التطبيقية الصرفة ذات المجال التطبيقي الواسع والتي تساعد على إتخاذ القرارات المبنية على أسس علمية صحيحة. وبات واضحاً استخدامه في المؤسسات الصحية لتحليل البيانات الخاصة بالامراض والوفيات والعوامل المسببة وتصميم نماذج الوفيات.

مشكلة الدراسة :

لا يزال الاطفال الخدج عرضة للعديد من المضاعفات، والتي تقلل من فرص بقائهم على قيد الحياة وتزيد من معدلات الوفيات في هذه الفئة العمرية لذلك اعتمدت هذه الدراسة دالة الانحدار اللوجستي للوقوف على اهم العوامل المؤثرة على حياة الاطفال الخدج.

أهمية الدراسة :

الأطفال هم مستقبل البلدان والاساس في بناء المجتمع وهم القاعدة التي ينطلق منها لبناء مستقبله السليم، فلابد من ان يكونوا بصحة وسلامة لاداء الدور المناط بهم في المستقبل، لذا تكمن أهمية الدراسة في اختيار اسلوب الانحدار اللوجستي ثنائي الاستجابة لتمثيل الدالة المقدرة وهو اسلوب انحدار وصفي في حال متغير الاستجابة الوصفي واعتماد طريقة الامكان الاعظم لتقدير معالم النموذج لدراسة المشاكل الصحية التي يعاني منها أطفال محافظة بابل والتي تؤدي الى ارتفاع نسب الوفاة ومحاولة وضع الحلول المستقبلية التي تقلل من اثر هذه الظاهرة وامكانية تعميمها على عموم المحافظات وذلك لتشابه الظروف المحيطة .

هدف الدراسة :

تهدف هذه الدراسة الى تحليل اهم العوامل المؤثرة في وفيات الاطفال الخدج في محافظة بابل عن طريق تحليل الانحدار اللوجستي. وصياغة توصيات تهدف إلى تخفيض نسب الوفيات عند الأطفال للارتقاء بواقع الخدمات الصحية المقدمة للاطفال الخدج .

حدود الدراسة :

اخذت بيانات الدراسة من خلال قسم الاحصاء في دائرة صحة بابل لعام 2016 وتشمل 12041 حالة دخلت وحدة العناية المركزة للاطفال الخدج خلال العام المذكور .

المبحث الاول : الجانب النظري لمفهوم تحليل الانحدار والانموذج اللوجستي

نتناول في هذا المبحث الدراسات السابقة واسباب وفيات الاطفال الخدج ومفهوم تحليل الانحدار والانموذج اللوجستي وطريقة تقدير معالم الانموذج .

الدراسات السابقة :[8]

اعتمدت الدراسات والبحوث التي تناولت موضوع وفيات الاطفال الرضع على طريقة (Brass) لتقدير معدلات الوفيات العمرية (Age Specific Deaths Rates) للأطفال والأطفال الرضع وتعتمد على البيانات المتعلقة بالنساء في عمر ألإنجاب وعدد المولودين أحياء وعدد الباقين على قيد الحياة لسنة محددة . وقد تم اعداد دراسة عن هذا الموضوع عام (1992) من قبل الباحثين الامريكيين (Ronald Lee & Lawrence Carter) وذلك اعتماداً على اسلوب جديد في التقدير والتنبؤ بمعدلات الوفيات العمرية (ASDRs) والتي عرفت فيما بعد باسميهما (Lee-Carter Method)، وتعتمد هذه الطريقة على الاتجاهات العمرية والزمنية الماضية ألتي تستنبط من البيانات التاريخية المستخدمة في تقدير الأنموذج الاحصائي الامثل. لقد قام عدد من الدارسين والباحثين في المجالين الصحي والاحصائي بدراسات عديدة للتوصل الى اسباب الوفاة للاطفال ووضع نموذج احصائي ملائم ومنها :

دراسة شادي اسماعيل التلباني[4] . قسم الاحصاء / جامعة الازهر- غزة ، 2004 والتي تناولت (استخدام نموذجي الانحدار اللوجستي وانحدار كوكس لدراسة العوامل المؤثرة على وفيات الاطفال الرضع في فلسطين)، وتوصلت هذه الدراسة الى انه على الرغم من اختلاف النموذجين الا انهما توصلا الى نفس المتغيرات المؤثرة على الظاهرة ، كما توصلت الدراسة الى ان اهم عامل يؤثر في وفيات الاطفال الخدج هو جنس المولود وطبيعة المولود مفرد او توأم ، وقد اوصت بان المستوى التعليمي للام والحالة التغذوية للام والطفل اهمية كبيرة في تفسير حالة البقاء على قيد الحياة للطفل الرضيع.

وفي عام 2005 قامت الهاشمي[6] ، سما سعدي علي، بـ(دراسة احصائية لوفيات الاطفال الرضع لمحافظة نينوى للفترة 1987 – 2004) باستخدام النماذج التنبؤية والمقارنة بين هذه النماذج عبر خمسة مؤشرات وهي :MSE وهو متوسط مجموع مربعات الخطأ وMAE وهو متوسط البواقي المطلقة و MAPE وهو متوسط مجموع مربعات الخطأ و ME متوسط البواقي و MPE متوسط البواقي المئوية وانتقاء الأنموذج الأفضل في ضوء هذه المؤشرات التي تتباين عن بعضها البعض.

اما دراسة حسين، حسن محمد[9] ،(2006) فقد تناولت(وفيات الاطفال تحت تاثير العوامل الطبيعية والبشرية في محافظة ديالى) استخدم الباحث معامل الارتباط لبيرسون للفصل بين العوامل المؤثرة وغير المؤثرة واوصى في دراسته الى ضرورة الاهتمام

بالجانب الصحي للام الحامل وعدم الاقتراب من المناطق التي تعرضت لضربات عسكرية ملوثة وتوجيه الامهات بعدم استخدام العقاقير دون استشارة الطبيب وعدم تعرضهن للاجهاد .

وفي عام 2008 قامت مديرية الرعاية الصحية الأولية[10] في وزارة الصحة السورية بالتعاون مع منظمة اليونسيف – المكتب المركزي للاحصاء باعداد (دراسة أسباب وفيات الأطفال دون الخمس سنوات) واعتمدت الدراسة على معدلات الوفيات وبينت تاثير عامل التشوهات الخلقية على وفيات الاطفال الخدج ،وقد اوصت بتطبيق برنامج الفحص الطبي ما قبل الزواج و إعطاء حمض الفوليك للنساء الراغبات في الحمل بغية الوقاية من تشوهات الأنبوب العصبي .

وفي العام نفسه قدمت (باناجيوتا كيتسانتاس)(Panagiota Kitsantas)[15] دراسة (الاسباب الكامنة لوفيات الاطفال الخدج Underlying and Multiple Causes of Death in Preterm Infants) وباستخدام المنحنى اللوجستي وتحليل البيانات باستخدام برنامج spss اظهرت الدراسة ان قصر فترة الحمل وامراض الجهاز التنفسي للاطفال من اهم الاسباب المؤدية الى وفاة الاطفال الخدج .كما اوصت الدراسة احتمال الوفاة تزداد (خاصة بالنسبة لأولئك الذين ولدوا بين 24-30 أسبوع) والذين لديهم مشاكل في الجهاز التنفسي .

اما السباح، شروق عبد الرضا[5]،(2009) فقد تناولت (بناء أنموذج انحدار لوجستي معدل لحياة الأطفال الخدج في محافظة كربلاء) وتبين لديها أن حياة الطفل الخديج مرتبطة بالعوامل (مدة الحمل ،وزن الطفل ،نوع المستشفى ، سبب الرقود، عمر الأم) أكثر من غيرها .كما اوصت باعتماد النتائج في دراسة واقع المستشفيات ومحاولة تعزيزها بالمستلزمات الضرورية لتوفير العناية الصحية للمواطنين .

وفي نفس العام قامت البير، خولة علي[3] بدراسة (تقييم الواقع الصحي للطفولة في العراق بالعمر (4-0) سنة والخدمات المتاحة لهم ومقترحات تطويرها) واستخدمت في الدراسة المنهاج الوصفي التحليلي ، واكدت ان نسبة كبيرة من وفيات الاطفال هي بسبب الامراض التنفسية ونقص التغذية والاسهال . كما اكدت على نقص مراكز الرعاية الصحية الاولية التي تعنى بالام والطفل .

كما توصل الباحثان Saad J. Sulaiman*, Hanady J. Mahmood[17] في دراستهما عام 2014 (Assessment Of Factors Causing Mortality Rate Of Neonate In Al-Batool Teaching Hospital In Mosul City) تقييم اسباب وفيات الأطفال حديثي الولادة خلال عام واحد 2010 وتوصلا الى 38% من الوفيات تحدث في الايام الاولى من عمر الطفل. وكانت الاكثر شيوعا لاسباب الوفاة هي متلازمة الضائقة التنفسية بنسبة (42%) يتبعها (19%) الولادة المبكرة، والتشوهات الخلقية بنسبة (14%) و كذلك نقصان الوزن بنسبة 11% .

اما الدراسة التي اعدت من قبل زيد، افنان محمد محمود[7] ،(2014) والتي تناولت (اثر العوامل الديموغرافية والاقتصادية والاجتماعية في وفيات الاطفال الرضع في الاراضي الفلسطينية من واقع مسح الاسر الفلسطينية) فقد استخدمت اسلوب التحليل الوصفي وقد استنتجت ان هناك علاقة عكسية تربط بين عمر الام وصلة القرابة بين الزوجين وبين متوسط وفيات الاطفال واضافت اثر المستوى التعليمي والرضاعة الطبيعية ووزن الطفل على وفيات الاطفال.

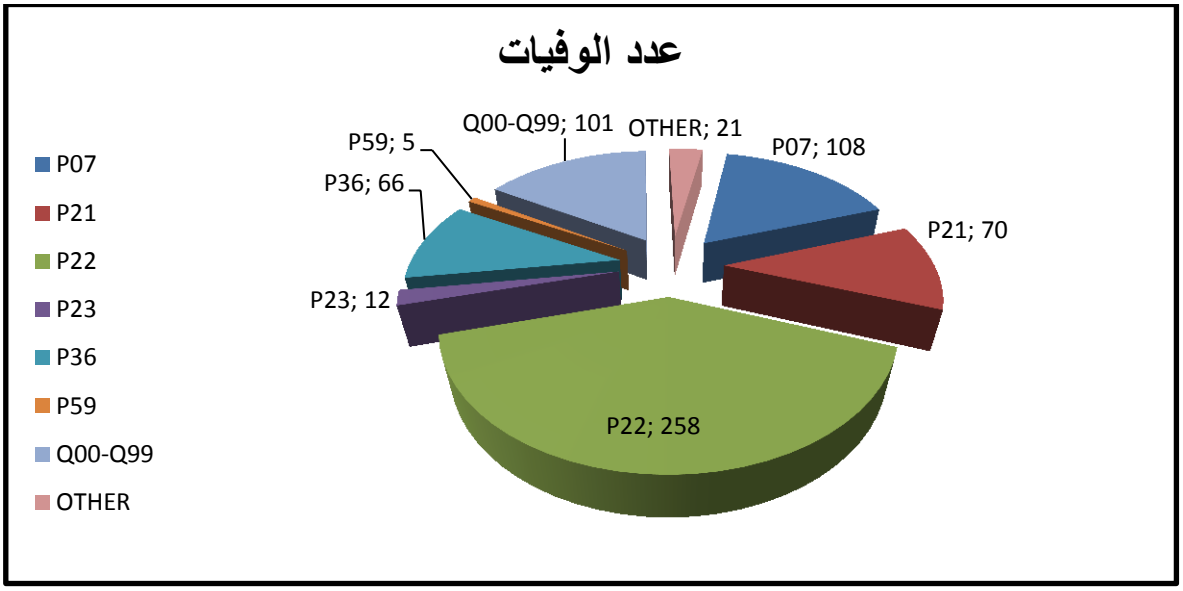
اسباب الوفاة :

تكتسب تقديرات وفيات الأطفال حديثي الولادة او الخدج (وهو احتمال الوفاة خلال الشهر الاول) ووفيات ما بعد حديثي الولادة (وهو احتمال الوفاة قبل اتمام السنة من العمر) ووفيات الطفولة (اقل من خمس سنوات) اهميتها من خلال امرين هما : رصد وتقييم البرامج الصحية الراهنة، ولصياغة السياسات وتوزيع الاولويات للبرامج المستقبلية . كما ينظر اليها كمؤشرات أساسية للوضع الاجتماعي والاقتصادي ونوعية الحياة ومستوى المعيشة العام في المجتمع .

ويعرف حديث الولادة (new born) بانه الطفل الذي يصل عمره لأقل من شهر واحد بعد الولادة. والخديج (Preterm): مَن وُلد مبكرا اي لفترة حمل أقل من 37 أسبوعا.

وفقا لدراسة أجرتها منظمة الأمم المتحدة للطفولة (يونيسيف) وُجد أن معظم أسباب وفيات الأطفال حديثي الولادة يعود إلى تشوهات خلقية بنسبة 2.27 % و الولادات المتعددة بنسبة 26 % و عدم النضوج بنسبة 7.21 %، ومرض الأم بنسبة 7.6 % وأوضاع معينة للوليد 4.6 % والاختناق بنسبة 9.4 %. وأوضحت الدراسة أن أكثر التشوهات الخلقية شيوعًا، كانت للأمراض القلبية الخلقية والتشوهات الخلقية المتعددة، التي شكلت نسبتها 8.25 % و 1.19 % على التوالي [11].

وتشير البيانات الاحصائية التي تم الحصول عليها من دائرة صحة بابل الى ان عدد الاطفال المتوفين لعام 2016 بلغ (641) وكانت بنسبة 40.25% من اسباب الوفاة تعود لعسر التنفس (P22) ،تليها اسباب نقص الوزن ونقص فترة الحمل (P07) بنسبة 16.84% ، ومن ثم التشوهات الخلقية (Q00-Q99) بنسبة 15.75% و الاختناق الولادي (P21) بنسبة 10.92%، و تسمم الدم الجرثومي(P36) بنسبة 10.29% وتاتي اسباب الوفاة (ذات الرئة الولادي (P23) واليرقان الولادي (P59)واسباب اخرى بنسبة 5.92%.) والشكل (1) يوضح ذلك .



شكل (1) عدد وفيات الخدج حسب اسباب الوفاة في محافظة بابل لعام 2016

والولادة المبكرة (Premature birth) هي اكبر سبب لوفيات المواليد، وهي ثاني أكبر سبب لوفاة الأطفال دون سن الخامسة ، اذ أكثر من مليون طفل يموت سنويا بسبب مضاعفات الولادة المبكرة، ومعظمها في العالم النامي.
وعلى الصعيد العالمي فان أكثر من 1 من كل 10 حالات الحمل تنتهي بالولادة المبكرة. وان الأطفال المولودين مبكرا في أقل من 23 أسبوعا تكون فرصة بقائهم على قيد الحياة قريبة من الصفر، وفي 23 أسبوعا 15%، وفي 24 أسبوعا 55% وترتفع هذه النسبة الى 80% للاطفال المولودين بعمر 25 أسبوعا.

تحليل الانحدار[16] Regression Analysis

يعد تحليل الانحدار واحداً من الادوات الاحصائية الاكثر استخداما لتحليل البيانات متعددة العوامل بالاضافة الى انه يوفر مفهوما وبطرق بسيطة لطبيعة العلاقات الدالية بين المتغيرات .ويعرف تحليل الانحدار على انه مجموعة من التقنيات التحليلية التي تدرس العلاقات المتبادلة بين مجموعة معينة من المتغيرات .
ويكتسب الانحدار أهمية بالغة في تحليل العلاقات الدالية كونه يقدم اربعة وظائف رئيسية هي[1] :
1- الوصف (description): إذ يمكن وصف وتلخيص البيانات بمعادلة واحدة او اكثر لوصف طبيعة العلاقة بين تلك المتغيرات .
2- السيطرة (control): السيطرة على نتائج المتغير التابع عند تغير قيم المتغيرات المفسرة .
3- التنبؤ (prediction): تقدير قيم متغير الاستجابة والتنبؤ بالمقادير التي يمكن ان يحصل عليها في للمستقبل.
4- تقدير المعلمات (Estimation of parameters):للاستدلال على اهمية وقوة العلاقة بين المتغيرات واتجاه العلاقة .
وتنقسم نماذج الانحدار الى قسمين هما نماذج الانحدار الخطية (linear regression) ونماذج الانحدار اللاخطية (nonlinear regression)

الانموذج اللوجستي : Logistic Model[12]

الانحدار الخطي غالبا ما يعمل بشكل جيد عندما يكون متغير الاستجابة كمي،أما الحالة التي يأخذ فيها متغير الاستجابة احد اثنين من القيم المحتملة (1,0) فهي في الغالب ناجمة عن استجابة نوعية . يستخدم تحليل الإنحدار اللوجستي في الدراسات الوبائية والطبية اذ يتم تحديد المتغيرات المستقلة الكمية والنوعية التي تؤثر في إحتمال حدوث المتغير التابع .
يستند النموذج الى دالة الاحتمال اللوجستي ويتم تحديده بصورة عامة :

$$E(Yi) = \frac{\exp(B_0+\beta 1X)}{1+\exp(B_0+\beta 1X)}$$

ويمكن ان تؤول الى الصيغة :

$$E(Yi) = \frac{1}{1+\exp-(B_0+\beta 1X)}$$

Yi : وهو متغير عشوائي ثنائي الاستجابة binary response يأخذ القيم اما(0) او (1) اذ يكون توزيعه الاحتمالي كالاتي :

$$Yi \sim B(1,P)$$
$$Yi = 0 \Rightarrow pr(Yi = 1) = pi$$
$$Yi = 1 \Rightarrow pr(Yi = 0) = 1 - pi$$

pi : هو قيمة الاحتمال عندما $Yi = 0$.
$1 - pi$: هو قيمة الاحتمال عندما $Yi = 1$.

B_0: الحد الثابت .
$\beta 1$: الميل (slop) وهو تعبير عن احتمال لوغارتمي ناتج عن زيادة وحدة واحدة في X .
هذا يعني ان نسبة الرجحان e^{B1} تتغير عند زيادة X بمقدار وحدة واحدة .
Xi : المتغير المفسر .
e: هو اساس اللوغاريتم الطبيعي ≈ 2.718
ei : هو حد الخطأ العشوائي ، اذ ان E(ei)=0 .
وان التوقع لمتغير الاستجابة والتباين يكونان بالشكل الاتي :

$$E(Yi) = (B_0 + \beta 1Xi) = pi$$
$$var(Yi) = E[(Yi - p)]^2 = p(1-p)$$

ويمكن تحويل النموذج اللوجستي الى دالة خطية :

$$zi = lin(\frac{pi}{1-pi})$$
$$zi = B_0 + \beta x \; or \; zi = \beta 0 + \sum \beta ixi$$

وفي حالة وجود اكثر من متغير تفسيري يكون:

$$lin\left(\frac{pi}{1-pi}\right) = B_0 + \sum_{i=1}^{k} BjXij$$
i=1,2,…,n ; j=1,2,…,k

نلاحظ ان اللوغارتم الطبيعي لنسبة الرجحان هو الدالة الخطية للمتغير المستقل.
ويمكن التعبير عن دالة التوزيع الاحتمالي بالشكل الاتي[14] :

$$fi(Yi) = pi^{Yi}(1-pi)^{1-Yi}$$
Yi=0.1 , i=1,….,n

عادة في انموذج الانحدار اللوجستي يستخدم اكثر من متغير تفسيري للحصول على وصف كافٍ للدالة المقدرة .
ببساطة نستبدل (B_0+B_1X) بـــ$(B_0+B_1X_1+\ldots..+B_{P-1}X_{P-1})$ ولتبسيط الصيغة نستبدلها بالشكل الاتي :

$$X'B = B_0 + B_1X_1 + \cdots\cdots + B_{p-1}X_{p-1}$$
$$Xi'B = B_0 + B_1X_{i1} + \cdots\cdots + B_{p-1}X_{i.p-1}$$
$$E(Yi) = \frac{\exp(X'B)}{1+\exp(Xi'B)}$$

ومن خصائص عنصر الخطأ :
1- حد الخطأ لا يتوزع توزيعاً طبيعياً للمتغير ثنائي الاستجابة وان:

$$ei = yi - (B_0 + \beta Xi)$$

ممكن ان تاخذ قيمتين فقط وهي :

1- If Yi=1 : ei=1- $B_0 - \beta Xi$
2- If yi=0 : ei = - $B_0 - \beta Xi$

2- تباين حد الخطا غير ثابت : حد الخطا(ei) ليس له تباينات متساوية في حالة المتغير ثنائي الاستجابة :

$$\delta^2(Yi) = E[Yi - E(Yi)]^2$$
$$= (1\text{-pi})^2\text{pi} + (0\text{-pi})^2\,(1\text{-pi})$$

ان تباين ei هو نفس تباين yi بسبب ei=yi-pi و pi ثابت

$$\delta^2(ei) = pi(1-pi) = \big(E(Yi)\big)(1-E(Yi)$$

$\delta^2(ei)$ تعتمد على xi لذا فان تباينات عنصر الخطا تختلف باختلاف قيم xi .
3- القيود على الدالة ثنائية الاستجابة : بما ان دالة الاستجابة تمثل احتمالات المتغير المفسر عندما يكون قيمة متغير الاستجابة هو 0,1 يجب ان تكون متوسط الاجابة مقيدة على النحو التالي . $0 \le E(Yi) = pi \le 1$

اختبار هوزمر – ليمشو[13] :- The Hosmer- Lemeshow Test

وهذا الاختبار يستعمل لتقيم حسن المطابقة للأنموذج وهو يسمح بأي عدد من المتغيرات التوضيحية سواء أكانت مستمرة أم متقطعة ، و تتوزع توزيع x^2 بدرجة حرية (8) واذا كانت قيمتها اكبر من مستوى المعنوية فان هذا يؤكد جودة التوفيق للأنموذج بالكامل .ونرمزلها (C) ويتم احتساب الاختبار وفقا للصيغة الأتية :-

$$C = \sum_{k=1}^{10} \frac{(O_k - n'_k \overline{\pi_k})^2}{n'_k \overline{\pi_k}(1-\overline{\pi_k})}$$

اذ ان :

O_k : التكرار المشاهد للمجموعة (k) عند الاستجابة y=1 .

$$O_k = \sum_{k=1}^{n'_i} Y_j$$

$\overline{\pi_k}$: متوسط الاحتمالات المتوقعة للمجموعة (k)

$$\overline{\pi_i} = \sum_{j=1}^{m'_i} \frac{n_j \widehat{\pi_j}}{m'_i}$$

n'_k : مجموع التكرار المشاهد في المجموعة (k) .

نسبة الارجحية : Add Ratio

وهي تشير الى مقدار التغير الحاصل في نسبة ارجحية وقوع الحدث عند حدوث تغير في قيمة المتغير المستقل (x_i) المرتبط بالمعلمة (B) . وتزداد نسبة ارجحية وقوع الحدث اذا كانت قيمة ((exp(B) اكبر من الواحد ، بينما تنخفض هذه النسبة مع كل زيادة في قيمة المتغير المستقل (x_i) عندما تكون قيمة ((exp(B) اقل من الواحد .

$$\exp(B) = 2.718^{(B)}$$

التحليل العاملي [2] : Factor Analysis

وهو اسلوب احصائي يستخدم في دراسة الظواهر بهدف ارجاعها الى العوامل المؤثرة فيها ويهدف الى تفسير معاملات الارتباط الموجبة بين مختلف المتغيرات . وان العوامل الناتجة من عملية التحليل العاملي هي عبارة عن متغيرات كامنة مشتقة من مجموعة متغيرات .ويهدف الى :

أ- يلخص المتغيرات في عدد اقل من العوامل الرئيسة التي يمكن ان تفسر الظاهرة وذلك عن طريق الحصول على مجموعة جديدة من المتغيرات (العوامل) وبعدد اقل لتحل جزئيا او كليا محل المجموعة الاصلية من المتغيرات .

ب- التعرف على المتغيرات التي لها دلالة احصائية مهمة والتي تتطلب مزيدا من عمليات التحليل الاخرى .

جـ - يحل مشكلة الارتباطات العالية بين المتغيرات المستقلة والتي تؤدي الى تقديرات خاطئة لمعاملاتها في حال بقاء الارتباط بينها

وهناك عدة طرائق للتحليل العاملي اهمها واكثرها شيوعاً هي طريقة المكونات الرئيسية (Principle Components) ولها عدة مزايا منها تؤدي الى تشبعات دقيقة وكل عامل يستخرج اقصى كمية من التباين وانها تؤدي الى اقل قدر ممكن من البواقي كما ان مصفوفة الارتباطات تختزل الى اقل عدد ممكن من العوامل المتعامدة غير المرتبطة .

يعد تحليل المركبات الرئيسية Principle Components Analysis (PCA) احد الاساليب المهمة المستعملة في تحليل البيانات متعدد المتغيرات و يستعمل عندما تكون المتغيرات التوضيحية ذات ارتباط عالٍ اي وجود مشكلة التعدد الخطي . كما يستعمل في تحليل البيانات ذات العدد الكبير للمتغيرات التوضيحية نسبة الى عدد المشاهدات والذي يعد أحد التقنيات المستعملة لتلخيص البيانات واختصارها. ويستند الى ايجاد الجذور المميزة والمتجهات المميزة لمصفوفة الارتباطات .

تقدير معالم الانموذج

يمكن تقدير معلمات انموذج الانحدار اللوجستي بطريقة المربعات الصغرى وطريقة الامكان الاعظم وقد تم اعتماد طريقة الإمكان الأعظم (MLM) باعتبارها طريقة كفوءة علما بانها تتطلب طرائق تكرارية لحسابها لذلك تحتاج إلى قيم أولية لـ (B0 , B1) .

طريقة الامكان الاعظم Maximum Likelihood Method [18]

تعتمد هذه الطريقة على ايجاد قيم $\hat{B}$ وهي عبارة عن تقديرات لمتجه المعلمات B التي تجعل الدالة في نهايتها العظمى وعلى فرض ان لدينا n من المتغيرات العشوائية المستقلة $Y_1, Y_2, \ldots \ldots Y_n$ وتتوزع وفق توزيع ثنائي الحدين بالمعلمتين (ni, pi) وان Yi يمثل مجموع حالات النجاح في كل محاولة من (ni) وان هناك (k) من المتغيرات المستقلة في كل مجموعة من المجاميع .

لتحويل الانموذج اللوجستي الى الانموذج الخطي باخذ اللوغارتم الطبيعي نحصل على :

$$\log\left(\frac{pi}{1-pi}\right) = \sum_{k=0}^{K} x_{ik}B_K \qquad i=1,2,\ldots,N \qquad \ldots\ldots (1)$$

وان دالة الكثافة الاحتمالية ل (Yi) هي :

$$f(y|B) = \prod_{i=1}^{N} \frac{ni!}{yi!(ni-yi)!} P^{yi}(1-Pi)^{ni-yi} \qquad \ldots\ldots (2)$$

هناك C_{yi}^{ni} طرق مختلفة لترتيب yi حالة نجاح من ni من المحاولات .

وان احتمال النجاح لاي وحدة محاولة من ni هي P ، والاحتمال لعدة نجاحات هي P^{yi} . وان احتمال (ni-yi) من الفشل هي $(1-Pi)^{ni-yi}$.

ويعبر عن دالة الامكان الاعظم لقيم B بالصيغة :

$$L(B|y) = \prod_{i=1}^{N} \frac{ni!}{yi!(ni-yi)!} P^{yi}(1-Pi)^{ni-yi} \qquad \ldots\ldots (3)$$

تقديرات الامكان الاعظم هي قيم B التي تعظم دالة الاحتمال في المعادلة (3) .

كذلك المعادلة دون المفكوك يمكن ان تعطي نفس النتيجة كما لو كانت متضمنة المفكوك . اذ يمكن كتابة المعادلة المراد تعظيمها على النحو التالي :

$$\prod_{i=1}^{N} \left(\frac{Pi}{1-Pi}\right)^{yi} (1-Pi)^{ni} \qquad \ldots\ldots (4)$$

وبعد أخذ e للطرفين في المعادلة (1) نحصل على :

$$\left(\frac{Pi}{1-Pi}\right) = e^{\sum_{k=0}^{K} x_{ik}B_K} \qquad \ldots\ldots (5)$$

وبعد ايجاد Pi تصبح

$$Pi = \left(\frac{e^{\sum_{k=0}^{K} x_{ik}B_K}}{1+e^{\sum_{k=0}^{K} x_{ik}B_K}}\right) \qquad \ldots\ldots (6)$$

وبتعويض المعادلة (5,4) في المعادلة 6 نحصل على:

$$\prod_{i=1}^{N} \left(e^{\sum_{k=0}^{K} x_{ik}B_K}\right)^{yi} \left(1 - \frac{e^{\sum_{k=0}^{K} x_{ik}B_K}}{1+e^{\sum_{k=0}^{K} x_{ik}B_K}}\right)^{ni} \qquad \ldots\ldots (7)$$

ويمكن كتابة المعادلة (7) بالشكل :

$$\prod_{i=1}^{N} \left(e^{\sum_{k=0}^{K} x_{ik}B_K}\right)^{yi} \left(1 + e^{\sum_{k=0}^{K} x_{ik}B_K}\right)^{-ni} \qquad \ldots\ldots (8)$$

وهذه هي نواة دالة الامكان الاعظم . وباخذ اللوغارتم الطبيعي للمعادلة (8) نحصل على :

$$L(B) = \sum_{i=1}^{K} yi\left(\sum_{k=0}^{K} x_{ik}B_K\right) - ni.\log\left(1 + e^{\sum_{k=0}^{K} x_{ik}B_K}\right) \qquad \ldots\ldots (9)$$

لايجاد النقطة الحرجة للوغارتم دالة الامكان الاعظم نشتق لكل قيم B ونساويها بالصفر .

$$\frac{\partial}{\partial B_K}\sum_{k=0}^{K} x_{ik}B_K = xik \quad \ldots\ldots\ldots (10)$$

$$\frac{\partial(B)}{\partial B_K} = \sum_{i=1}^{N} y_i x_{ik} - ni.\frac{1}{1+e^{\sum_{k=0}^{K} x_{ik}B_K}}.\frac{\partial}{\partial B_K}(1+e^{\sum_{k=0}^{K} x_{ik}B_K})$$
$$= \sum_{i=1}^{N} y_i x_{ik} - ni.\frac{1}{1+e^{\sum_{k=0}^{K} x_{ik}B_K}}.\frac{\partial}{\partial B_K}\sum_{k=0}^{K} x_{ik}B_K$$
$$= \sum_{i=1}^{N} y_i x_{ik} - ni.\frac{1}{1+e^{\sum_{k=0}^{K} x_{ik}B_K}}.e^{\sum_{k=0}^{K} x_{ik}B_K}.xik$$
$$= \sum_{i=1}^{N} y_i x_{ik} - ni\ Pixik \quad \ldots\ldots\ldots (11)$$

تقدير الامكان الاعظم لB يمكن ان نحصل عليه بواسطة وضع K+1 من المعادلات في المعادلة (11) ومساواتها للصفر والحل لكل قيم B_K ، عندها نحصل على نظام من K+1 من المعادلات اللاخطية يحوي على K+1 من المتغيرات . ولحل النظام لمتجه B_K نستخدم طريقة نيوتن رافسون (The Newton – Raphson Method) لايجاد تقديرات الامكان الاعظم وبالصيغة التالية :

$$B^{(1)} = B^{(0)} + [X^T W X]^{-1}.X^T(y-\mu) \quad \ldots\ldots\ldots (12)$$

اذ ان :

μ : متجه عمودي بطول N للعناصر $n_i P_i$

W: مصفوفة مربعة من الرتبة N بالعناصر القطرية $n_i P_i(1-P_i)$

X: مصفوفة المتغيرات المستقلة .

$B^{(1)}$: تمثل متجه المعلمات B المراد تقديرها .

$B^{(0)}$: تمثل متجه القيم الاولية للمعلمات

اذ يتم التوقف عن أيجاد التقديرات لمتجه المعلمات (B) عندما يكون الفرق بين الدورة السابقة والدورة اللاحقة صغير جدا ويقترب من الصفر .

المبحث الثاني : الجانب التطبيقي وتحليل النتائج

تم تحليل البيانات بواسطة حزمة البرامج spss من خلال طريقة الحذف الخلفي backward ، للحصول على انموذج احصائي يمثل طبيعة هذه البيانات وذلك للتحقق من أدائها في الواقع العملي . ولدراسة العلاقة بين المتغيرات اعتمدت بيانات من سجلات دائرة صحة بابل شملت12041 من مرضى الأطفال الخدج خلال عام (2016).

تمثلت البيانات بكل من المتغير المعتمد (متغير الاستجابة) وفاة او عدم وفاة الطفل الخديج . بينما كانت المتغيرات التوضحية تتمثل بمدة الحمل و وزن الطفل ونتيجة الولادة (مفرد ،توأم) وسبب الرقود (نقص الوزن و قصر مدة الحمل ، ذات الرئة الولادي ، ، يرقان ولادي ، تشوهات خلقية ، متلازمة ضيق التنفس (عسر التنفس) ، اختناق ولادي ، تسمم الدم الجرثومي ، أخرى) وعمر الأم ،معقم اوغير معقم ، كمتغيرات مستقلة تؤثر في حياة الطفل .

ان تأثير المتغيرات التوضيحية في المتغير التابع يقاس عن طريق معاملات الانحدار B ، ولمعرفة فيما اذا كان للمتغيرات التوضيحية تأثير في المتغير التابع نفترض الفرضيات الآتية :-

فرضية العدم :- لا يوجد تأثير للمتغيرات التوضيحية في المتغير التابع $\boldsymbol{H_0: B_i = 0}$

الفرضية البديلة :- يوجد تأثير للمتغيرات التوضيحية في المتغير التابع $\boldsymbol{H_1: B_i \neq 0}$

$i = 1,2,3, \ldots\ldots$

والنتائج مبينة في **الجدول (1)** والذي يبين معالم النموذج التي تم الحصول عليها وفيه قيم المعالم المقدرة للمتغيرات الداخلة اضافة الى قيمة الثابت والخطا المعياري ومعنوية كل معلمة ومنه يمكن ملاحظة المتغيرات المعنوية وغير المعنوية .

جدول (1) المتغيرات الداخلة في المعادلة (Variables in the Equation)

			B	S.E.	Df	Sig.	Exp(B)
Step 3[a]	X1	مدة الحمل	-0.170	0.018	1	0.000	0.844
	X2	وزن الطفل	-0.491	0.086	1	0.000	0.612
	X6	تؤام او مفرد			2	0.022	
	X6.(1)	تؤام او مفرد(1)	-1.012	1.294	1	0.434	0.363
	X6,(2)	تؤام او مفرد(2)	-1.259	1.295	1	0.331	0.284
	X7	سبب الرقود			7	0.000	
	X7,(1)	سبب الرقود(1)	0.470	0.208	1	0.024	1.600
	X7,(2)	سبب الرقود(2)	0.434	0.200	1	0.030	1.543
	X7,(3)	سبب الرقود(3)	0.046	0.168	1	0.786	1.047
	X7,(4)	سبب الرقود(4)	-0.166	0.334	1	0.620	0.847
	X7,(5)	سبب الرقود(5)	0.455	0.209	1	0.030	1.577
	X7,(6)	سبب الرقود(6)	-2.488	0.300	1	0.000	0.083
	X7,(7)	سبب الرقود(7)	2.269	0.200	1	0.000	9.671
	X5,(2)	نوع الولادة(2)	-0.238	0.118	1	0.043	0.788
		Constant	5.783	1.413	1	0.000	324.728

اظهرت نتائج التحليل دخول المتغيرات (مدة الحمل ، وزن الطفل ، سبب الرقود ، نوع الولادة) واستبعاد العوامل الأخرى لأنها غير معنوية . **والجدول (1)** يمثل المعالم المقدرة للمتغيرات الداخلة فضلاً عن قيمة الثابت والخطأ المعياري . والقيمة الاحتمالية المقابلة (sig.) ويستعمل لاختبار الدلالة لكل متغير فان كانت ذات دلالة احصائية فتعني ان قيمة ($\hat{b}$) لذلك المتغير لا تساوي صفراً . وكذلك يبين الجدول معنوية المعالم (sig.) ومنه يمكن ملاحظة المتغيرات المعنوية وغير المعنوية فان كانت ($sig. < 0.05$) يدل على ان المتغير ذو تاثيرمعنوي. ويظهر ايضا في الجدول قيم (exp(B)) التي تمثل نسبة الارجحية (Adds Ratio) وهي تشير الى مقدار التغير الحاصل في نسبة ارجحية وقوع الحدث (الوفاة) عند حدوث تغير في قيمة المتغير المستقل المرتبط بالمعلمة (B) ، وتزداد نسبة الارجحية لوقوع الوفاة كلما زادت عن الواحد .

ومن كل ما سبق يمكن ان تكون معادلة الانحدار اللوجستي بالصيغة الاتية :

$$\text{Log}\left(\left(\frac{\rho}{1-\rho}\right)\right) = 5.783 - 0.17X_1 - 0.491X_2 + 0.47X_{7,1} + 0.434X_{7,2} + 0.455X_{7,5} - 2.488X_{7,6} + 2.269X_{7,7} - 0.238X_{5,1}$$

جدول (2)اختبار الدلالة الاحصائية للانموذج

Omnibus Tests of Model Coefficients				
		Chi-square	df	Sig.
Step 3a	Model	1133.853	12	0.000

من **جدول (2)** نلاحظ اختبار كاي تربيع (χ^2) والذي يعتمد على المقارنة بين القيم المشاهدة والقيم المتوقعة للانموذج الذي يحتوي الحد الثابت فقط والانموذج الذي يحتوي كافة المتغيرات .نجد قيمة مربع كاي وتساوي (1133.853) عند درجة حرية (12) ومستوى معنوية (P-Value=0.00 < 0.05) وهذا يعني ان الانموذج ذو دلالة احصائية ويوفق البيانات بشكل جيد اي ان المتغيرات المتضمنة في الانموذج لها تاثير ودلالة ومساهمة في التصنيف .

ولاختبار فيما اذا كان الانموذج يمثل البيانات بشكل جيد ام لا نلاحظ الجدول الاتي :

جدول (3) اختبار هوزمر- ليمشو

Hosmer and Lemeshow Test			
Step	Chi-square	df	Sig.
1	15.373	8	0.052
2	12.700	8	0.123
3	12.603	8	0.126

من **جدول (3)** اختبار هوزمر- ليمشو وهو احد الاختبارات المهمة لقياس جودة توفيق انموذج الانحدار اللوجستي ويعتمد على مدى قرب الاحتمالات المشاهدة من الاحتمالات المتوقعة . ويختبر الفرضية الاتية :

H_0 : الانموذج المقدر يوافق البيانات بشكل جيد

H_1 : الانموذج المقدر لا يوافق البيانات بشكل جيد

وتتم مقارنة قيم ($P - Value$) لكاي تربيع (χ^2) مع مستوى المعنوية ونلاحظ ان قيمة ($P - Value = 0.126 > 0.05$) وهي معنوية لجودة التوفيق للانموذج .وهذا يعني لا يوجد دليل كاف لرفض فرضية العدم اي ان الانموذج ملائم ويوافق البيانات بشكل جيد .

والجدول(4) يمثل التحقق من صحة النموذج الذي تم توفيقه للبيانات وفي هذا الاختبار يتم تقسيم البيانات الى عشر مجموعات ويتم تعريف هذه المجموعات حسب الترتيب التصاعدي لتقدير المخاطرة فالمجموعة تقابل مجموعة المفردات الاقل مخاطرة(حياة) والمجوعة الثانية تقابل مجموعة مفردات الاكثر مخاطرة (وفاة) .

جدول (4)

Contingency Table for Hosmer and Lemeshow Test						
		حياة = حياة او وفاة		وفاة = حياة او وفاة		
		Observed	Expected	Observed	Expected	Total
Step 1	1	1200	1202.012	4	1.988	1204
	2	1196	1200.657	8	3.343	1204
	3	1199	1199.026	5	4.974	1204
	4	1178	1188.190	26	15.810	1204
	5	1183	1174.368	21	29.632	1204
	6	1171	1165.316	33	38.684	1204
	7	1158	1154.674	46	49.326	1204
	8	1141	1138.458	63	65.542	1204
	9	1109	1101.772	95	102.228	1204
	10	866	876.528	339	328.472	1205

والخطوة الاخيرة لتحسين الانموذج ومعرفة الاهمية النسبية لكل متغير نقوم بتحليل المركبات الرئيسية وكما ياتي :

جدول (5) جدول تفسير التباين

Total Variance Explained									
Component	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	2.393	29.910	29.910	2.393	29.910	29.910	2.131	26.642	26.642
2	1.767	22.085	51.995	1.767	22.085	51.995	1.920	24.004	50.646
3	1.037	12.963	64.958	1.037	12.963	64.958	1.145	14.312	64.958
4	.853	10.665	75.623						
5	.679	8.487	84.111						
6	.504	6.303	90.413						
7	.403	5.037	95.450						
8	.364	4.550	100.000						

ويفسر **الجدول(5)** العوامل وايها اكثر تاثيرا من خلال القيمة eigen values إذ كلما كان القيمة اكبر من واحد دل على ان المكون ساهم في تفسير التباين الحادث بين المتغيرات ونلاحظ تاثير التباين 29.9% للمكون الاول و 22% للمكون الثاني و12% للمكون الثالث . وتم استبعاد العوامل من تسلسل 4 الى تسلسل 8 إذ ان قيمها اقل من واحد .

الجدول (6) جدول المعاملات للعوامل الرئيسية

		B	S.E.	df	Sig.	Exp(B)	95% C.I.for EXP(B)	
							Lower	Upper
Step 1[a]	factor V1	-.884-	.032	1	.000	.413	.388	.440
	factor V2	-.423-	.046	1	.000	.655	.599	.717
	factor V3	-.115-	.043	1	.008	.891	.819	.970
	Constant	-3.397-	.056	1	.000	.033		

ومن **جدول (6)** نلاحظ معنوية العوامل الرئيسية ($sig < 0.05$) وكذلك قيم المعاملات (B) وبذلك تكون معادلة الانحدار بالشكل الاتي :

$$\mathrm{Log}\left(\frac{\rho}{1-\rho}\right) = -3.397 - 0.884V_1 - 0.423V_2 - 0.115V_3$$

الاستنتاجات:

من خلال تحليل البيانات والتطبيق العملي تم التوصل الى الاستنتاجات الآتية :-

1- يعتبرالانموذج اللوجستي انموذجاً جيدا لتفسير العلاقة بين وفيات الاطفال الخدج (كمتغير تابع) والمتغيرات التي تفسر سلوك هذا المتغير مثل (مدة الحمل، وزن الطفل، تؤام او مفرد، سبب الرقود، عمر الام، معقم او لا) وان جميع المعايير التي استخدمت اعطت دلالات جيدة على جودة الانموذج اللوجستي ومنها اختبار هوزمر- ليمشو اذ كانت ($P-Value = 0.126 > 0.05$)

2- بينت نتائج التحليل أن (مدة الحمل ، وزن الطفل ، سبب الرقود(P07 ، P36 ، P59)) ذات تاثير بدرجة معنوية عالية في حياة الطفل الخديج . لذلك تعتبر هذه العوامل من اهم مسببات الوفاة في الاطفال الخدج .

3- ان نسبة الوفيات الخدج قد بلغت 5.3% وهي نسبة عالية وكان سبب عسر التنفس (P22) و اسباب نقص الوزن ونقص فترة الحمل (P07) و التشوهات الخلقية (Q00-Q99) الحصة الاكبر من اسباب الوفاة .

4- ساهم التحليل العاملي في شرح التباين الحادث بين المتغيرات ونلاحظ ذلك في جدول (5) ، كمااستطاع التحليل العاملي في اختزال عدد المتغيرات الى ثلاث مركبات رئيسية ونلاحظ معنوية هذه المركبات من خلال قيمة .sig وكما مبين في جدول (6) .

التوصيات:

في ضوء الاستنتاجات التي تم التوصل إليها ، نوصي بما يلي:-

1- ضرورة استخدام الأنموذج اللوجستي في حالة البيانات الثنائية (Binary Responses) لدقة نتائجه .

2- نوصي الاهتمام بالعوامل المؤثرة في حياة الطفل المولود حديثا واعطائها اولويات في التداخلات الصحية والطبية وحسب درجة معنوية تلك العوامل .

3- دراسة إمكانية إدخال متغيرات جديدة قد تكون لها تأثير ومرتبطة بوفاة أو بقاء الطفل الخديج مع توخي الدقة في اختيار المتغيرات التي لها تأثير فعلي على احتمال الاستجابة، لان الدقة في الحساب تؤثر في دقة النتائج التي نحصل عليها

4- ضرورة استخدام هذه المعايير في حالة أنموذج الانحدار اللوجستي متعدد الاستجابة وملاحظة مدى كفاءة أداءهم .

5- اعتماد النتائج لغرض دراسة الواقع الصحي للمستشفيات لتوفير عناية صحية افضل للاطفال الخدج من خلال توفير الاجهزة الطبية الحديثة والمتطورة والكادر الطبي المناسب للتعامل مع هذه الحالات التي تتطلب رعاية خاصة .

المصادر

[1] التميمي، زهرة حسن عباس وآخرون ، 2014 ، تحليل الانحدار. مديرية دار الكتب للطباعة والنشر / جامعة البصرة .

[2] ابو فايدة ، احمد ،(التحليل العاملي :مفهومه، أهدافه، شروطه،انواعه، خطواته، مثال تطبيقي لكيفية استخراج التحليل العاملي بنظامSPSS)، جامعة الازهر ، غزة .2016

[3] البير، خولة علي،2009 (تقييم الواقع الصحي للطفولة في العراق بالعمر 4-0 سنة والخدمات المتاحة لهم ومقترحات تطويرها) ،وزارة التخطيط والتعاون الانمائي العراقية- دائرة التنمية البشرية

[4] التلباني ،شادي اسماعيل، 2004 (استخدام نموذجي الانحدار اللوجستي وانحدار كوكس لدراسة العوامل المؤثرة على وفيات الاطفال الرضع في فلسطين)، جامعة الازهر- غزة .

[5] السباح، شروق عبد الرضا،(2009) (بناء أنموذج انحدار لوجستي معدل لحياة الأطفال الخدج في محافظة كربلاء)،رسالة دكتوراه- جامعة بغداد .

[6] الهاشمي ، سما سعدي علي ،2005 (دراسة إحصائية لوفيات الأطفال الرضع لمحافظة نينوى للفترة (2004 – 1987) .رسالة ماجستير- جامعة بغداد

[7] زيد، افنان محمد محمود ،2014 ،(اثر العوامل الديموغرافية والاقتصادية والاجتماعية في وفيات الاطفال الرضع في الاراضي الفلسطينية من واقع مسح الاسر الفيسطينية 2010) ،رسالة ماجستير – جامعة النجاح الوطنية –نابلس

[8] فتاح، احمد فاضل ،2006 (التنبؤ بمعدلات الوفيات وبناء جداول الحياة لدولة فرنسا) رسالة ماجستير-الجامعة المستنصرية .

[9] حسين، حسن محمد ،(2006) (وفيات الاطفال تحت تاثير العوامل الطبيعية والبشرية في محافظة ديالى – دراسة ميدانية)،مجلة الفتح- العدد 26 .

[10] المكتب المركزي للاحصاء ،2008 (دراسة أسباب وفيات الأطفال دون الخمس سنوات) ، مديرية الرعاية الصحية الأولية في وزارة الصحة السورية بالتعاون مع منظمة اليونسيف .

[11] خوجة ، عبد الحفيظ ،2016، وفيات الاطفال الأطفال حديثو الولادة وعوامل الخطورة التي تهدد حياتهم، جدة، العدد 13548 ، http://aawsat.com/home/article/532821

[12] Douglas C. Montgomery & George C. Runger,2003, Applied statistics and probability for engineers , Printed in the United States of America. Third Edition
[13] HOSMER,DAVID W, LEMESHOW,STANLEY,(2000),Applied Logistic Regression, Second edition, John Wiley &Sons.INC,New York
[14] John.O.Rawlings & others,1998,Applied regression analysis A research Tool North Carolina State University , USA , Second Edition
[15] Panagiota Kitsantas,2008, Underlying and Multiple Causes of Death in Preterm Infants, George Mason University. PDF
[16] Chatterjee & Alis .Hadi ,2006 regression analysis by Example , by John Wiley & Sons, Inc . Fourth Edition.
[17] Mahmood, Hanady J. & Saad J. Sulaiman,2013, Assessment Of Factors Causing Mortality Rate Of Neonate In Al-Batool Teaching Hospital In Mosul City(2010). Kufa Journal for Nursing Sciences.
http://www.uokufa.edu.iq/journals/index.php/kjns/article/view/2212
[18] Czepie, Scott A. l Maximum Likelihood Estimation of Logistic Regression Models: Theory and Implementation , , PDF
https://czep.net/stat/mlelr.pdf