استخدام النماذج الإحصائية العشوائية لتحديد العوامل المؤثرة على فقد الأجنة
(دراسة تطبيقية)

تحت إشراف

أ.د. / محمد توفيق اسماعيل البلقيني

أستاذ الرياضيات والإحصاء الإكتواري

كلية التجارة –جامعة المنصورة

الباحثة

حنان عبد السلام الفاهم

طالبة ماجستير كلية التجارة –جامعة المنصورة

المستخلص :-

هدفت الدراسة إلى تحديد العوامل المؤثرة على فقد الأجنة، عن طريق استخدام نماذج عشوائية، وهي نموذج الانحدار اللوجيستي الثنائي، وتحليل البقاء، وكذلك نموذج الانحدار المتعدد والمقارنة بينهم من خلال معامل التحديد وتحليل العلاقة ما بين المتغيرات والوصول الي المتغيرات الأكثر تأثيرًا على هذا الفقد، وشملت عينة الدراسة (٤٠٠) حالة منها (٣٠٠) حالة إجهاض وهي تمثل (٧٥%) من الحالات، و(١٠٠) حالة مولود حي وهي تمثل (٢٥%) من الحالات وذلك نظرًا للظروف التي كانت تعاني منها البلاد. ويتضح من النتائج أن قيم مربع كاي لجميع المتغيرات كانت ذات دلالة إحصائية عند مستوي المعنوية ٠.٠٥ فأقل ماعدا عمل الزوج وطبيعة الاقامة. أظهرت نتائج الانحدار الخطي المقدر بنموذج الانحدار المتعدد وجود علاقة انحدار معنويه بين نواتج حالات الحمل ومتغيرات الدراسة المستقلة وفقا لاختبار(**F**) عند مستوي معنوية (٠.٠٥)، وبفحص قيمة معامل التحديد R^2، نجد أن القدرة التفسيرية للنموذج المقدر تبلغ (٧٧.٧%)، معني ذلك أن التغيرات التي تحدث لحالات نواتج الحمل ترجع الي التغيرات التي تحدث للمتغيرات المستقلة، والباقي (٢٢.٣%) ترجع الي عوامل أخري، أما نتائج الانحدار اللوجيستي بلغت قيمه معامل التحديد الكلية (٨٩%)، وتحليل البقاء بلغت قيمة معامل التحديد(٩٠%). وهذا يدل على أن المتغيرات التفسيرية الموضحة تؤثر تأثير مباشر على فقد الأجنة.

ABSTRACT: -

THE STUDY AIMS TO DETERMINE THE FACTORS INFLUENCING THE LOSS OF EMBRYOS BY USING RANDOM MODELS: THE LEAST SQUARES REGRESSION MODEL, THE LOGISTIC REGRESSION AND THE SURVIVAL ANALYSIS,

COMPARING THEM THROUGH THE SELECTION FACTOR AND ANALYZING THE RELATION BETWEEN THE VARIABLES AND REACHING THE MOST IMPORTANT VARIABLES ON THIS LOSS. 400 CASES OF ABORTION, REPRESENTING 75% OF THE CASES, AND 100 CASES OF LIVE BIRTHS REPRESENTING 25% OF THE CASES DUE TO THE CIRCUMSTANCES EXPERIENCED BY THE COUNTRY. IT IS CLEAR FROM THE RESULTS THAT THE VALUES OF THE SQUARE OF KAY FOR ALL VARIABLES ARE STATISTICALLY SIGNIFICANT AT THE LEVEL OF 0.05 OR LESS WORK DONE HUSBAND AND NATURE OF RESIDENCE. THE RESULTS SHOW ALSO A SIGNIFICANT OF REGRESSION IN THE LOWER SQUARES METHOD. THERE WAS A SIGNIFICANT REGRESSION RELATIONSHIP BETWEEN PREGNANCY OUTCOMES AND INDEPENDENT STUDY VARIABLES ACCORDING TO F TEST AT A SIGNIFICANT LEVEL OF 0.05. IN EXAMINING THE VALUE OF THE R^2 FACTOR, THE EXPLANATORY CAPACITY OF THE ESTIMATED MODEL WAS 77.7% (22.3%). THE RESULTS OF LOGISTIC REGRESSION (R89%) AND SURVIVAL (R90) WERE SIGNIFICANT. THIS INDICATES THAT THE SUGGESTED VARIABLES HAVE A DIRECT EFFECT ON THE LOSS OF EMBRYOS.

١. المقدمة:

إن علم الإحصاء يُعد من العلوم الأساسية التي تعتمد في التخطيط والتنمية المستقبلية التي تحتاج إلى تحليل عملي دقيق للمتغيرات الأساسية المستخدمة فيها التي تساعد على وضع البرامج المعتمدة على البيانات الإحصائية المتوفرة عن هذه المتغيرات، ومن هنا لابد أن يكون هناك دور فعال لعلم الإحصاء في توضيح وتحليل ومعالجة المشكلة كونه العلم الذي يهتم بالتحليل والتنبؤ في جميع مجالات العلوم الأخرى وتبرز أهمية الإحصاء في تحديد أهم العوامل المؤثرة على فقد الأجنة لدى الأمهات الليبيات.

الإجهاض قضية حساسة اجتماعية وثقافية ودينية وقانونية وصحية وتعتبر مشكلة الإجهاض أهم المشكلات الشائعة التي قد تعانيها المرأة في سن الإنجاب وحدوث الاجهاضات المتكررة قد تسبب لها معاناة كبيرة صحية ونفسية وتكون حادثة مأساوية على حياة أي زوجة ويعد من أكبر التحديات الإكلينيكية التي تواجه الطبيب المعالج وفي غالب الحالات قد يعجز الطب عن إيجاد حلول ناجحة لمثل هذه المشكلة التي تؤدى إلى عدم الإنجاب والتي تؤثر على العلاقات الاجتماعية للأسرة. (مي حامد ،٢٠١٦)

تعتبر مرحلة الانجاب من المراحل العمرية الهامة التي تمر بها المرأة حيث إنها تتعرض خلالها لحالات متعاقبة من النشاط الفسيولوجي والبدني أثناء الحمل والولادة وهذه المرحلة لها انعكاسات كبيرة على الأوضاع الصحية والاجتماعية والاقتصادية كما تؤثر على قطاع واسع من السكان وبصفة خاصة في سن الإنجاب. (أحمد عبد المنعم، ماجدة محمد،٢٠٠٧)

كما يعتبر الإجهاض مشكلة شائعة يمكن أن يصيب أي امرأة. (Michael chapman 2011,)

ولكن المشكلة الأساسية في كثير من الأحيان لا تزال غير مكتشفة وغير معروفة في كثير من الحالات.(Milja kaare, 2009)

ستتطرق الباحثة في دراستها التطبيقية إلى استخدام نماذج عشوائية وهي ونموذج الانحدار اللوجيستي نموذج الانحدار المتعدد وتحليل البقاء وعمل مقارنة بينهم ومعرفة أيهم أفضل من خلال الدراسة التطبيقية على بيانات الأمهات الليبيات اللاتي تعرضن لفقد أجنتهن.

٢-مشكلة البحث:

منذ الأحداث الأخيرة في ليبيا واستخدام كافة أنواع القنابل والغازات مما أدى إلى تفشي الكثير من الأجسام والغازات التي تسبب الإجهاضات العفوية بصوره ملحوظة بين السكان وخصوصًا القريبة من المناطق الموبوءة بتلك الانفجارات، لذا تمثلت مشكلة البحث في خطورة ارتفاع الإجهاضات لدى النساء، لذلك سيتم التطرق إلى ثلاثة نماذج عشوائية وهي نموذج الانحدار اللوجيستي ونموذج الانحدار المتعدد وتحليل البقاء لتحديد أهم العوامل التي تؤثر على المتغير التابع وفي هذه الدراسة نبين التالي:

١) تحديد المتغيرات المستقلة التي تؤثر على المتغير التابع .

٢) اختيار أهم المتغيرات المستقلة ذات التأثير المعنوي على المتغير التابع وذلك من خلال مقاييس خاصة بكل نموذج.

٣-أهمية البحث:

١) الوصول إلى أسلوب تنبؤ عام بعدد الأجنة المفقودة أثناء الولادة وتأثير ذلك على التنبؤ بحجم السكان.
٢) معرفة أي الأساليب أفضل وأدق في اظهار النتائج، وأيهم يظهر بأقل نسبة خطأ.

<u>٤- هدف البحث :-</u>

١) استخدام كل من نموذج الانحدار اللوجيستي ونموذج الانحدار المتعدد وتحليل البقاء والمقارنة بينهم لتحديد المتغيرات المؤثرة على هذا الفقد.
٢) اختيار المتغيرات المستقلة ذات التأثير المعنوي على فقد الأجنة وذلك من خلال مقاييس خاصة بكل نموذج.

<u>٥- النماذج المستخدمة في الدراسة :-</u>

أولًا: نموذج الانحدار المتعدد: *MULTIPLE REGRESSION MODEL*

يعتمد نموذج الانحدار المتعدد على استخدام طريقة المربعات الصغرى(LSM) في تقدير معالم معادلة الانحدار خطأ راجع للفرق بين القيم الحقيقية والقيم المقدّرة والناتجة عن التعويض المباشر في معادلة التقدير. وبالاشتقاق الجزئي للدالة لكل من **(μ, λ)** ومساواة ناتج كل اشتقاق بالصفر يتم الحصول على معادلتين، وبإجراء تبسيط رياضي يتم الوصول إلى قيمة مقدرات المربعات الصغرى $(\hat{\mu}, \hat{\lambda})$ للمعلمتين.(رشا ابراهيم ،٢٠١٢)

ومن خصائص هذه الطريقة أنها غير متحيزة فعند استخدام أسلوب المربعات الصغرى والتي تعتمد على الخطية في المعادلات يتم اللجوء إلى اعتماد التحويل الخاص بالتوزيع المفترض وهو التوزيع الطبيعي (***Normal Distribution***) وبمعلمتين (μ, λ) وذلك لغرض الحصول على المقدرات الخاصة بنموذج الانحدار المتعدد أولًا ثم التعويض بما يساويها للحصول على المعلمات المقدرة بصيغة **(α,β)** وكالآتي:(Syed E. Ahmed, et al, 2008)

$$\lambda/\sqrt{x} = Y + \mu\sqrt{x}$$

بأخذ الدالة التربيعية لأقل مجموع مربعات الخطأ الناتج من القيم المقدّرة وكالآتي:

$$g = (\frac{\lambda}{\sqrt{x}} - \mu\sqrt{x_i} - y_i)$$

$$g = \sum\left(\frac{\lambda}{\sqrt{x_i}} - \mu\sqrt{x_i} - y_i\right)^2 = 0 \qquad \text{.......... (1-1)}$$

وتؤخذ مشتقة العلاقة أعلاه مرة إلى المعلمة الاولى (**λ**) ومرة اخرى إلى المعلمة الثانية (μ) ومساواة المشتقة إلى الصفر ينتج:

$$\frac{\partial g}{\partial \mu} = \sum\left(\frac{\lambda}{\sqrt{x_i}} - \mu\sqrt{x_i} - y_i\right)^2$$

$$\Rightarrow 2\sum\left(\frac{\hat{\lambda}}{\sqrt{x_i}} - \hat{\mu}\sqrt{x_i} - y_i\right)\left(\sqrt{x_i}\right) = 0 \qquad \text{.......(1-2)}$$

$$\frac{\partial g}{\partial \hat{\lambda}} = 2\sum\left(\frac{\hat{\lambda}}{\sqrt{x_i}} - \hat{\mu}\sqrt{x_i} - y_i\right)\left(\frac{1}{\sqrt{x_i}}\right) = 0 \qquad \text{......(1-3)}$$

وبقسمة المعادلة رقم (**1-2**) والمعادلة رقم (**1-3**) على (**2**) ينتج لدينا:

$$\sum\left(\frac{\hat{\lambda}}{\sqrt{x_i}} - \hat{\mu}\sqrt{x_i} - y_i\right)\left(\sqrt{x_i}\right) = 0$$

$$\sum\frac{\hat{\lambda}}{\sqrt{x_i}} \cdot \sqrt{x_i} - \hat{\mu}\sum(\sqrt{x_i})(\sqrt{x_i}) - \sum y_i\sqrt{x_i} = 0$$

$$n\hat{\lambda} - \hat{\mu}\sum x_i - \sum y_i\sqrt{x_i} = 0$$

$$n\hat{\lambda} - \hat{\mu}\sum x_i = \sum y_i\sqrt{x_i} \qquad \text{........(1-4)}$$

وبالعودة إلى معادلة رقم (**1-4**) بعد قسمتها على (**2**):

$$\sum(\frac{\hat{\lambda}}{\sqrt{x_i}} - \hat{\mu}\sqrt{x_i} - y_i)\left(\frac{1}{\sqrt{x_i}}\right) = 0$$

$$\sum\frac{\hat{\lambda}}{\sqrt{x_i}} \times \frac{1}{\sqrt{x_i}} - \hat{\mu}\sum\sqrt{x_i} \cdot \frac{1}{\sqrt{x_i}} - \sum y_i\frac{1}{\sqrt{x_i}} = 0$$

$$\hat{\lambda}\sum\frac{1}{x_i}-n\hat{\mu}=\sum\frac{y_i}{\sqrt{x_i}} \qquad .....(1\text{-}5)$$

بحل المعادلات **(1-4)** **(1−5)** آنيا بالحذف والتعويض ينتج لدينا:

$$\left(n\hat{\lambda}-\hat{\mu}\sum x_i=\sum y_i\sqrt{x_i}\right)(\mathbf{n})$$

$$\left[\hat{\lambda}\sum\frac{1}{x_i}-n\hat{\mu}=\sum\frac{y_i}{\sqrt{x_i}}\right]\left[\sum x_i\right]$$

$$n^2\hat{\lambda}-n\hat{\mu}\sum x_i=n\sum y_i\sqrt{x_i}$$

بالطرح $$-\hat{\lambda}\sum\frac{1}{x_i}\cdot\sum x_i\pm n\hat{\mu}\sum x_i=-\sum\frac{y_i}{\sqrt{x_i}}\cdot\sum x_i$$

$$\hat{\lambda}\left[n^2-\sum\frac{1}{x_i}\right]=n\sum y_i\sqrt{x_i}-\sum\frac{y_i}{\sqrt{x_i}}\cdot\sum x_i$$

$$\therefore\ \hat{\lambda}=\frac{n\sum y_i\sqrt{x_i}-\sum\frac{y_i}{\sqrt{x_i}}\cdot\sum x_i}{n^2-\sum\frac{1}{x_i}}$$

وبقسمة جميع الحدود على (n^2) للطرف الأيمن ينتج لدينا:

$$\hat{\lambda}=\frac{\frac{n}{n^2}\sum y_i\sqrt{x_i}-\left(\sum\frac{y_i}{\sqrt{x_i}}\ .\sum x_i\right)/n^2}{1-\frac{1}{n^2}\sum\frac{1}{x_i}}$$

$$\therefore\ \hat{\lambda}=\frac{\sum\frac{y_i\sqrt{x_i}}{n}-(\sum\frac{y_i}{\sqrt{x_i}}\cdot\sum x_i)/n^2}{1-\frac{1}{n^2}\sum\frac{1}{x_i}} \qquad ......(1\text{-}6)$$

وللحصول على **($\hat{\mu}$)** نحتاج إلى عمليات التعويض بالصيغة رقم **(1-6)** أعلاه وكالآتي:

وبالعودة إلى معادلة رقم **(1-5)** والتعويض فيها بما يساويها في معادلة رقم **(1-6)** ينتج لدينا:

$$\hat{\lambda}\sum\frac{1}{x_i}-n\hat{\mu}=\sum\frac{y_i}{\sqrt{x_i}} \qquad \text{......(1-7)}$$

$$\left(\frac{\sum\frac{y_i\sqrt{x_i}}{n}-\left(\sum\frac{y_i}{\sqrt{x_i}}\cdot\sum x_i\right)/n^2}{1-\frac{1}{n^2}\sum\frac{1}{x_i}}\right)\left(\sum\frac{1}{x_i}\right)-n\hat{\mu}=\sum\frac{y_i}{\sqrt{x_i}}$$

$$\left(\frac{\sum\frac{y_i\sqrt{x_i}}{n}-\left(\sum\frac{y_i}{\sqrt{x_i}}\cdot\sum x_i\right)/n^2}{1-\frac{1}{n^2}\sum\frac{1}{x_i}}\right)\left(\sum\frac{1}{x_i}\right)-\left(\sum\frac{y_i}{\sqrt{x_i}}\right)=n\hat{\mu}$$

$$\hat{\mu}=\frac{\left[\frac{\left(\sum\frac{y_i\sqrt{x_i}}{n}-\left(\sum\frac{y_i}{\sqrt{x_i}}\cdot\sum x_i\right)/n^2\right)}{1-\frac{1}{n^2}\sum\frac{1}{x_i}}\left(\sum\frac{1}{x_i}\right)-\left(\sum\frac{y_i}{\sqrt{x_i}}\right)\right]}{n}$$

$$\therefore \hat{\mu} = \frac{\dfrac{\sum \dfrac{y_i\sqrt{x_i}}{n} \cdot \sum \dfrac{1}{x_i} - \left(\sum \dfrac{y_i}{\sqrt{x_i}} \cdot \sum x_i \sum \dfrac{1}{x_i}\right)/n^2}{1 - \dfrac{1}{n^2}\sum \dfrac{1}{x_i}} - \left(\sum \dfrac{y_i}{\sqrt{x_i}}\right)}{n} \quad \text{........(1-8)}$$

وللحصول على مقدرات طريقة (***OLS***) بصيغ معلمات التوزيع (***BISA***) والتي معلماتها الواجبة التقدير تساوي (**β,α**) يمكن اللجوء إلى المقدرات التالية:

بما إنه لدينا التوزيع (***Birnbaum-Saunders***) بدلالة المعلمات (**β,α**) وحسب العلاقات الآتية والتي تساوي:

$$\hat{\alpha} = \frac{1}{\sqrt{\hat{\mu}\hat{\lambda}}} \quad , \quad \hat{\beta} = \frac{\hat{\lambda}}{\hat{\mu}}$$

وبالعودة إلى المعادلتين رقم (**1-7**) و (**1-8**) ينتج لدينا:

$$\hat{\beta} = \frac{\dfrac{\dfrac{\sum y_i\sqrt{x_i}}{n} - \left(\sum \dfrac{y_i}{\sqrt{x_i}} \cdot \sum x_i\right)/n^2}{1 - \dfrac{1}{n^2}\sum \dfrac{1}{x_i}}}{\dfrac{\dfrac{\dfrac{\sum y_i\sqrt{x_i}}{n} \cdot \sum \dfrac{1}{x_i} - \left(\sum \dfrac{y_i}{\sqrt{x_i}} \cdot \sum x_i \sum \dfrac{1}{x_i} - \dfrac{\sum y_i}{\sqrt{x_i}}\right)}{1 - \dfrac{1}{n^2}\sum \dfrac{1}{x_i}}}{n}}$$

$$\hat{\beta} = \left(\frac{\frac{\sum y_i\sqrt{x_i}}{n} - \left(\sum \frac{y_i}{\sqrt{x_i}} \cdot \sum x_i\right)/n^2}{1 - \frac{1}{n^2}\sum \frac{1}{x_i}} \frac{n}{\frac{\frac{\sum y_i\sqrt{x_i}}{n} \cdot \sum \frac{1}{x_i} - \left(\sum \frac{y_i}{\sqrt{x_i}} \cdot \sum x_i \sum \frac{1}{x_i} - \frac{\sum y_i}{\sqrt{x_i}}\right)}{1 - \frac{1}{n^2}\sum \frac{1}{x_i}}} \right)$$

لنفرض أن:

$$T_1 = \sum y_i\sqrt{x_i}$$

$$T_2 = \sum \frac{y_i}{\sqrt{x_i}}$$

$$T_3 = \sum x_i$$

$$T_4 = \sum \frac{1}{x_i}$$

$$\frac{\hat{\beta}}{x} = \left(\frac{\frac{T_1}{n} - (T_2T_3)/n^2}{1 - \frac{T_4}{n^2}} \right) \left(\frac{n}{\frac{\frac{T_1T4}{n} - (T_2T_3T_4 - T_2)}{1 - \frac{T_4}{n^2}}} \right) \left(\frac{n^2}{n^2 - T_4} \right)$$

$$\hat{\beta} = \left(\frac{\frac{nT_1 - T_2T_3}{n^2}}{\frac{n^2 - T_4}{n^2}} \right) \left(\frac{n}{\frac{\frac{T_1T4}{n} - (T_2T_3T_4 - T_2)}{1 - \frac{T_4}{n^2}}} \right) \left(\frac{n^2}{n^2 - T_4} \right)$$

$$\widehat{\beta} = \left(\frac{nT_1 - T_2T_3}{n^2 - T_4}\right)\left(\frac{n}{n^2T_1T_4 - T_2T_3T_4 - n^2T_2}\right)\left(\frac{n^2}{n^2 - T_4}\right)$$

$$\widehat{\beta} = \left(\frac{nT_1 - T_2T_3}{n^2 - T_4}\right)\left(\frac{n^3}{[n^2 - T_4][n^2T_1T_4 - T_2T_3T_4 - n^2T_2]}\right)$$

$$\widehat{\beta}_{OLS} = \left(\frac{n^4T_1 - n^3T_2T_3}{[n^2 - T_4]^2[n^2T_1T_4 - T_2T_3T_4 - n^2T_2]}\right) \qquad \text{……(1-9)}$$

$$\because\ \widehat{\alpha} = \frac{1}{\sqrt{\widehat{\mu}\widehat{\lambda}}}$$

$$\widehat{\alpha} = \frac{1}{\sqrt{\left[\dfrac{\dfrac{\sum\frac{y_i\sqrt{x_i}}{n}\cdot\sum\frac{1}{x_i} - \left(\sum\frac{y_i}{\sqrt{x_i}}\cdot\sum x_i\sum\frac{1}{x_i}\right)\Big/n^2 - \sum\frac{y_i}{\sqrt{x_i}}}{1 - \frac{1}{n^2}\sum\frac{1}{x_i}}}{n}\right]\left[\dfrac{\frac{\sum y_i\sqrt{x_i}}{n} - \left(\sum\frac{y_i}{\sqrt{x_i}}\cdot\sum x_i\right)/n^2}{1 - \frac{1}{n^2}\sum\frac{1}{x_i}}\right]}}$$

$$\widehat{\alpha} = \frac{1}{\sqrt{\left[\dfrac{\dfrac{\frac{T_1T_4}{n} - (T_2T_3T_4)/n^2 - T_2}{1 - \frac{T_4}{n^2}}}{n}\right]\left[\dfrac{\frac{T_1}{n} - \frac{(T_2T_3)}{n^2}}{1 - \frac{T_4}{n^2}}\right]}}$$

$$\therefore\ \widehat{\alpha}_{OLS} = \frac{1}{\sqrt{\dfrac{\left(\frac{T_1T_4}{n} - \frac{T_2T_3T_4}{n^2} - T_2\right)\left(\frac{T_1}{n} - \frac{T_2T_3}{n^2}\right)}{\left(n - \frac{T_4}{n}\right)\left(1 - \frac{T_4}{n^2}\right)}}} \qquad \text{……..(1-10)}$$

وبهذا يمكن إيجاد مقدر دالة المعولية التقريبية وكالآتي:

$$\left(\frac{1}{\sqrt{\frac{\left(\frac{T_1T_4}{n}-\frac{T_2T_3T_4}{n^2}-T_2\right)\left(\frac{T_1}{n}-\frac{T_2T_3}{n^2}\right)}{\left(n-\frac{T_4}{n}\right)\left(1-\frac{T_4}{n^2}\right)}}}\left[\left(\frac{t}{\frac{n^4T_1-n^3T_2T_3}{[n^2-T_4]^2[n^2T_1T_4-T_2T_3T_4-n^2T_2]}}\right)^{1/2}\right.\right.$$

$$\left.\left.-\left(\frac{t}{\frac{n^4T_1-n^3T_2T_3}{[n^2-T_4]^2[n^2T_1T_4-T_2T_3T_4-n^2T_2]}}\right)^{-1/2}\right]\right)$$

$$\widehat{R}_{OLS}(t)=\left(\frac{1}{\widehat{\alpha}_{OLS}}\left[\left(\frac{t}{\widehat{\beta}_{OLS}}\right)^{1/2}-\left(\frac{t}{\widehat{\beta}_{OLS}}\right)^{-1/2}\right]\right) \qquad \text{……….(1-11)}$$

ثانيًا: الانحدار اللوجيستي: –

٢–١ مفهوم الانحدار اللوجيستي:

يعتبر نموذج الانحدار اللوجيستي من النماذج الرياضية المستخدمة في وصف العلاقة بين بعض المتغيرات التفسيرية والمتغير التابع الثاني Binary lndendent variable سواء كانت المتغيرات التفسيرية كمية أو وصفية ويتم استخدام الانحدار اللوجيستي وبصفة خاصة في المجالات الاقتصادية والاجتماعية والطبية أن نماذج الانحدار تنقسم إلى نوعين:

النوع الأول ويعرف بنماذج الانحدار الخطي Linear Regression Model

النوع الثاني يسمى بنماذج الانحدار اللاخطي Non Linear Regression Model

وكما هو معروف فإن العديد من الظواهر عند دراسة سلوكها تجدها تسلك سلوكاً غير خطي، ولتحليل تلك الظواهر فإنه يستخدم النماذج اللاخطية لوصف وتحليل تلك

الظواهر وفي الواقع فإنه بسبب صعوبة واستخدام هذه النماذج فإنه غالباً ما تستخدم نماذج الانحدار الخطي لهذا الغرض، وأن نماذج الانحدار اللوجيستي يعد واحد من هذه النماذج الإحصائية التي تستخدم لوصف وتحليل تلك الظواهر.

لقد ازدادت أهمية استخدام التحليل اللوجيستي يوماً بعد آخر، لكونه يهتم بتحليل البيانات ذات الاستجابة الثنائية والتي عادة ما يكون فيها المتغيير التابع (Response variable) ثنائياً (Bimary)، ففي حالة النجاح (Success) يأخذ متغير الاستجابة القيمة (1) وحالة الفشل (Failure) يأخذ القيمة (0) ويستخدم نموذج الانحدار اللوجيستي (Logistic Model) لوصف العلاقة بين متغير الاستجابة (Y) ومتغير توضيحي (مستقل) واحد X أو عدة متغيرات مستقلة (تفسيرية).

X_n، .... ، X_1 ويتم التعبير عن تلك العلاقة بالصيغة الآتية:

$$P(x) = \frac{1}{1+e^{-\alpha - Bx}} \qquad \text{............(2-1)}$$

حيث $-\infty < \beta < \infty$

α، β : معلمتا النموذج المراد تقديرهما.

$P(X)$: احتمال الاستجابة

X : المتغير المستقل

والصيغة أعلاه تعرف بدالة الاستجابة اللوجستية وتمتاز بأنها $P(X)$ محددة بين (1, 0) وأن المعلمتين (α ، B) غير مقيدتين وهناك نوعان من نماذج الانحدار اللوجيستي، الأول يسمى بنموذج الانحدار اللوجيستي ثنائي الاستجابة، والثاني يسمى بنموذج الانحدار اللوجيستي متعدد المتغيرات.(كريم عزو ،٢٠١٦)

٢-٢ نموذج الانحدار اللوجيستي ثنائي الاستجابة:

يبني نموذج الانحدار اللوجيستي على فرض أساسي وهو أن المتغير التابع (Y) متغير الاستجابة الذي يهتم بدراسته هو متغير ثنائي يتبع توزيع بيرنولي (Bernoull) ويأخذ القيمة (1) باحتمال مقداره (π) والقيمة (0) باحتمال ($1-\pi$) أي حدوث الاستجابة وعدم حدوثها، وكما نعلم في الانحدار الخطي الذي تأخذ متغيراته

التوضيحية ومتغير الاستجابة قيماً مستمرة فإن النموذج الذي يربط بين المتغيرات هو على النحو الآتي (منتظر محمد ،٢٠١٥)

$$y_i = \beta_0 + \beta_1 X_{i1} + e_i \qquad \ldots\ldots\ldots\ldots(2\text{-}2)$$

إذ أن (Y) يمثل متغيراً مشاهداً مستمراً ويفرض أن متوسط قيم (Y) المشاهدة الفعلية عند قيمة معينة للمتغير (X) هي E(Y) فإنه يمكن كتابة النموذج على النحو التالي:

$$E\,(y \mid x) = \beta_0 + \beta_1 X_{i1} \qquad \ldots\ldots\ldots(2\text{-}3)$$

ومن المعروف أن الطرف الأيمن في نموذج الانحدار يأخذ قيماً من (∞-) إلى (∞+) ولكن عندما يكون لدينا متغيران أحدهما ثنائي (y) فإن نموذج الانحدار الخطي البسيط لا يكون ملائماً لأن:

$$E(y \mid x) = P_r\,(y = 1) = \pi \qquad \ldots\ldots\ldots(2\text{-}4)$$

وبذلك تكون قيمة الطرف الأيمن محصورة ما بين الرقمين (1 ، 0) وبذلك يكون النموذج غير قابل للتطبيق من وجه نظر الانحدار وأن إحدى طرق الحل هذه المشكلة هو إدخال تحويله رياضية مناسبة على المتغير التابع (y)، ومن المعروف أن ($0 \leq \pi \leq 1$) ومن ثم النسبة $\frac{\pi}{1-\pi}$ هي عبارة عن مقدار موجب محصور بين (0 ، ∞) أي $(0 \leq \frac{\pi}{1-\pi} \leq \infty)$ وبأخذ اللوغاريتم الطبيعي للأساس (e) للتحويل فإن مجال قيمة تصبح محصورة بين $(-\infty \leq \log e \frac{\pi}{1-\pi} \leq \infty)$ وعليه فإنه يمكن كتابة نموذج الانحدار في حالة المتغير المستقل الواحد على النحو التالي:

$$\log\left(\frac{\pi}{1-\pi}\right) = \beta_0 + \beta_1 x \qquad \ldots\ldots\ (2\text{-}5)$$

والصيغة (2-5) تمثل نموذج الانحدار اللوجيستي بصيغة خطية بدلالة ما يعرف بـ (اللوجيت) (Logit).

وإذا كان لدينا أكثر من متغير مستقل واحد فإن نموذج الانحدار اللوجيستي المتعدد يصبح كالآتي:

$$\log\left(\frac{\pi}{1-\pi}\right)=\beta_0+\sum_{i=1}^{P}\beta_i X_{ip} \qquad ....(2-6)$$

إذ أن $i = 1 , 2 .... , P$ ، $j = 1 , 2 , ...n$

تسمى النسبة $\frac{\pi}{1-\pi}$ بنسبة الأفضلية أو أفضلية النجاح (odds of success) أو بنسبة الأفضلية للحدث المرغوب ونسبة $\frac{\pi}{1-\pi}$ يمكن ان تسمى أيضًا نسبة افضلية الفشل (Odds of failure) وأن المقدار ($\frac{\pi}{1-\pi}$ loge) يسمى لوغاريتم نسبة الأفضلية (log odds rating) أو اللوجيت ان نموذج الانحدار اللوجيستي لا يشترط توافر الافتراضات الآتية:

١. وجود علاقة خطية ما بين المتغيرات المستقلة والمتغير التابع.
٢. أن يكون توزيع المتغيرات توزيعًا طبيعيًا.
٣. تحقق خاصية ثبات التباين (تجانس البيانات).

ثالثًا: نماذج البقاء:

٣-١ توزيعات البقاء

٣-١-١ دلالات

إذا افترضنا أن T هي متغير عشوائي مستمر غير سالب يمثل وقت البقاء، مع دالة كثافة الاحتمال الدالة f (t) (pdf) ودالة توزيع تراكمي (cdf) F(t) = Pr{T < t} . وبالتركيز على دالة البقاء S(t) = Pr{T > t}، يكون احتمال البقاء على قيد الحياة عند وقت t ، ودالة المخاطرة A(t) = f(t)/S(t).

اذا كان $\Lambda(t) = \int_0^t \lambda(u)du$ تشير إلى الخطر تراكمي (أو تكاملي)، يمكن القول بأن

$$S(t) = \exp\{-A(t)\}. \qquad ....(3-1)$$

أي توزيع محدد لـ $t \in [0,\infty)$ يمكن أن يكون بمثابة توزيع للبقاء. يمكننا أيضًا أن نسحب إلى توزيعات الخدمة المحددة لـ $y \in (-\infty,\infty)$ من خلال النظر في t = exp{y}، حيث y = log t. وبشكل عام ، يمكننا أن نبدأ من r.v. W مع توزيع قياسي في (−∞،∞) لإنتاج مجموعات من توزيعات البقاء عن طريق إدخال تغييرات الموقع والنطاق على النموذج

$$\log T = Y = a + \sigma W. \qquad \ldots\ldots\ldots\ldots(3\text{-}2)$$

٣-١-٢ اللوغاريتم المنطقي

يكون لديT توزيع لوغاريتمي منطقي "لوجيستي" إذا كان:

$$Y = \log T = a + \sigma W \qquad \ldots\ldots\ldots\ldots(3\text{-}3)$$

إذا كان W لديها توزيع لوجستي "منطقي" قياسي، مع دالة كثافة الاحتمال ل pdf

$$f_W(w) = \frac{e^w}{(1+e^w)^2}, \qquad \ldots\ldots\ldots\ldots(3\text{-}4)$$

ومع دالة التوزيع التراكمي

$$F_W(w) = \frac{e^w}{1+e^w}. \qquad \ldots\ldots\ldots\ldots(3\text{-}5)$$

فتكون دالة البقاء المتممة

$$S_W(w) = \frac{1}{1+e^w}. \qquad \ldots\ldots\ldots\ldots(3\text{-}6)$$

بتغيير المتغيرات الي T نجد ان دالة اللوغاريتم اللوجيستي أو المنطقي للباقين هي

$$S(t) = \frac{1}{1+(\lambda t)^p}, \qquad \ldots\ldots\ldots\ldots(3\text{-}7)$$

حيث كالمعتاد a = — log A and p = 1/ a . وبأخذ اللوغاريتم log يمكننا الحصول على سالب تكامل المخاطر، و تفاضل w.r.t. t يمكننا إيجاد دالة المخاطر

$$\lambda(t) = \frac{\lambda p(\lambda t)^{p-1}}{1 + (\lambda t)^p}. \qquad \ldots\ldots\ldots\ldots(3\text{-}8)$$

يلاحظ أن لوغاريتمية دالة البقاء تكون خطية في لوغاريتم تي log T يقدم هذا منحني تفسيري: إذا كان لديك تقدير غير حدودي لدالة الباقين أو الناجين حيث يمكنك رسم لوغاريتمية الاحتمالية مقابل لوغاريتم الوقت ؛ إذا كان الرسم البياني يبدو وكأنه خط مستقيم فان دالة الباقين تكون لوغاريتمية لوجستية "log-logistic ".

وتكون المخاطر نفسها:

- تناقصية رتيبة من ∞ إذا كانت $p < 1$ ،
- تناقصية رتيبة من λ إذا $p = 1$ ، و
- تشبه لوغاريتمية طبيعية log-normal إذا كان $p > 1$.

٣-٢ النماذج مع متغيرات مصاحبة

يوجد أربع طرق لنمذجة بينات البقاء في وجود متغيرات مصاحبة

مجموعات حدودية" بارا مترية"

حياة معجلة

مخاطر تناسبية

أرجحية تناسبية

ونقوم بوصف كل طريقة فيما يلي

٣-٢-١ المجموعات البارا مترية Parametric Families

وهو نهج عام يتمثل في اختيار أحد التوزيعات المعلمية التي ناقشناها وترك معلمات ذلك التوزيع تعتمد على المتغيرات المشتركة. فمثلا،

في التوزيع الأسي، يمكننا الافتراض أن المعلمة λ تابعة لمتجه من المتغيرات المشتركة أو المصاحبة x ، على سبيل المثال باستخدام نموذج لوغاريتمي –خطي

$$\log \lambda = x'\beta$$

في توزيع Weibull ، يمكننا استخدام نموذج مماثل لـ ʎ مع الاحتفاظ بـ p ثابتة ، أو يمكننا أن ندع p تابعة للمتغيرات المصاحبة أيضًا ، على سبيل المثال

log p = X' ʎ

في نموذج كولي–ماك نيل " Coale-McNeil" باستخدام المعلمة رودي جوي – تروسيل Rodriguez-Trussell ، يمكن استخدام نموذج خطي للمتوسط

$$\mu = x'\beta$$

بينما يظل الانحراف المعياري σ ثابتًا (كما يحدث عادة في النماذج الخطية)، على العكس من ذلك، يمكننا السماح للتشتت ان يكن تابع لمتغيرات مصاحبة أيضًا، باستخدام

$$\log\sigma = x'\gamma,$$

مع المعلمات γ . في الحالة الأكثر شيوعًا، يمكننا ترك التناسب الذي في النهاية سيكون تابعا لمجموعة أخرى من المعلمات.

بشكل عام، مع مجموعات k يمكن إعطاء كل مجموعة توزيعها الخاص في العائلة. هذا نهج قابل للتطبيق، لكنه لا يفسح المجال للتفسيرات السهلة

٣-٢-٢ نماذج الحياة المعجلة

إذا اعتبارنا نموذج الانحدار العادي للوغاريتم وقت بقاء، من النموذج

$$Y = \log T = -x'\beta + \sigma W,$$

حيث يكون معامل الخطأ W توزيع مناسب، على سبيل المثال، القيمة القصوى، القيمة القصوى المعمم، طبيعية أو لوجستية، هذا يؤدي إلى Weibull، جاما معممة، نماذج لوغاريتمية عادية أو لوغاريتمية –لوجستية لـ T.

٣-٢-٣ مخاطر تناسبية Proportional Hazards

هناك نهج بديل لنمذجة بيانات البقاء وهو افتراض أن تأثير المتغيرات المصاحبة هو زيادة أو تقليل الخطر بمقدار متناسب في جميع الفترات. وهكذا

$$\lambda(t,x) = \lambda_0(t)e^{x'\beta},$$

حيث $\lambda_0(t)$ هو خط الأساس للمخاطر، أو الخطر بالنسبة لفرد مرجعي مع قيم المتغير صفر، و $\exp\{x'\beta\}$ هو الخطر النسبي المرتبط بقيم المتغير X.

من الواضح أن الأخطار التراكمية سوف تتبع نفس العلاقة، كما يمكن رؤيتها من خلال

$$S(t,x) = S_0(t)^{e^{x'\beta}}$$

دمج جانبي المعادلة السابقة. نكشف عن الخطر المتكامل الذي نعثر عليه من الناجين

٣-٣ تقدير أقصى احتمال Maximum Likelihood Estimation

جميع النماذج البارمترية قد تكون مناسبة من خلال تعظيم دالة الاحتمال المناسبة.

تتألف البيانات من أزواج $\{t_i, d_i\}$ حيث

- t_i هو البقاء أو وقت المراقبة، و
- d_i هو مؤشر الوفاة ، مع أخذ القيمة ١ للوفيات وصفر للحالات الخاضعة للمراقبة

دالة الاحتمال تحت المراقبة العامة غير المعلوماتية لها الشكل

$$L(\theta) = \prod_{i=1}^{n} \lambda(t_i|x_i)^{d_i} S(t_i|x_i),$$

وبشكل عام، يجب تعظيمه رقميًا باستخدام إجراء مثل Newton-Raphson.

Kalbfleisch و prentice لدينا مناقشة لطيفة للإجراءات التي يجب اتباعها في توضيب النماذج البارمترية، بما في ذلك المشتقات الأولى والثانية لنماذج الحياة المتسارعة باستخدام التوزيعات المعملية التي نوقشت هنا.

يمكن أن يصلح ترتيب Stata لعدد من النماذج البارا مترية، بما في ذلك النماذج الأسية، و Weibull و Gompertz في إطار المخاطر النسبية، وتسجيل الدخول

العادي، و log logistic، و جاما المعمم (وكذلك التوزيع الأسي وتوزيع Weibull) في إطار وقت الفشل المتسارع. الآن أنت تعرف لماذا يتم تضمين توزيع Weibull. (Rodriguez, G, 2010)

٦-**الجانب التطبيقي :**

٦-١ التطبيق العملي باستخدام الانحدار اللوجيستي:

لا يوجد تأثير ذات دلالة احصائية للمتغيرات المستقلة والمتمثلة في إحدى عشر متغير على مستوي صحة الحالة الحمل.

ولاختبار هذا الفرض قد استخدم الباحث الانحدار اللوجيستي والذي يدرس إثر المتغيرات المستقلة على المتغير التابع حيث يعتبر هذا النموذج من أشهر وأهم النماذج الإحصائية لصياغة دالة التمييز والتقسيم في الانحدار اللوجيستي ليس المهم تقدير المعالم بقدر استخدامها في حساب احتمال ناتج الحمل (إجهاض / مولود حي).

جدول ١: نتائج تقدير معالم نموذج الانحدار اللوجيستي

Variables in the Equation						
Sig.	Df	Wald	S.E.	B		
0.001	1	4.682	5.251	-8.065	**العمر**	Logistic
0.005	1	4.268	1.2963	-6.6	**ترتيب الحمل**	
0	1	2.521	2.9852	7.526	**عدد اسابيع الحمل**	
0	1	6.157	0.1254	-0.115	**عدد حالات الإجهاض السابقة**	
0.001	1	5.691	0.1047	-9.704	**الفترة بين الحملين**	
0.056	1	2.681	0.2652	-4.515	**العمل**	
0.216	1	1.768-	1.9	3.36	**طبيعة الاقامة**	
0.197	1	2.985	0.2271	14.111	**نوع الجنين**	
0.009	1	4.638	0.1365	-7.136	**الامراض المزمنة**	
0.001	1	7.352	0.1674	-12.490	**امراض أثناء الحمل**	

0.568	1	1.639	0.1833	-4.033	**عمل الزوج**	
.0.003	1	6.352	83.43	26.6754	**الثابت**	
0.895	معامل التحديد		٠.٠٠٠	المعنوية	328.13	مربع كاي

ويعتبر جدول رقم (١) أحد مخرجات تحليل الانحدار اللوجيستي وتوضح هذه القيم أو المؤشرات مدي أهمية المتغيرات المستقلة ومدي تفسيرها للتغيرات التي تحدث للمتغير التابع حيث قيمة مربع كاي (٣٢٨.١٣) بمعنوية (٠.٠٠٠) وهي أقل من مستوي الخطأ المسموح به ومن ثم نرفض الفرض العدمي ونقبل الفرض البديل القائل بأنه على الأقل يوجد متغير ذات معامل لا يساوي صفر، كما بلغت قيمة معامل التحديد الكلية الصحيحة (٨٩.٥%) ويعتبر نموذج بشكل عام جيد حيث ان معنوية النموذج أقل من مستوي الخطأ المسموح به.

- $\log y = 26.6754 - 8.065x_1 - 6.6x_2 + 7.526x_3 - 0.115x_4 - 9.704x_5 - 4.515x_6 + 3.36x_7 + 14.111x_8 - 7.136x_9 - 12.490x_{10} - 4.033x_{11}$

جدول ٢: التوفق لاختبار Hosmer and Lemeshow

Contingency Table for Hosmer and Lemeshow Test						
Total	y = مولود حي		y = إجهاض			
	Expected	Observed	Expected	Observed		
40	0	0	40	40	1	Logistic
40	0	0	40	40	2	
40	0	0	40	40	3	
40	0	0	40	40	4	
40	0	0	40	40	5	
40	0	0	40	40	6	
40	0	0	40	40	7	
40	20	20	20	20	8	
8	8	8	0	0	9	
72	72	72	0	0	10	

ويتضح من الجدول رقم (٢) وهو يمثل اختبار لا معلمي لجودة توفيق النموذج إذا يعتمد على حساب إحصاءه كاي للفرق بين القيم المشاهدة والقيم المتوقعة وقد

أقترح Hosmer and Lemeshow باستخدام توزيع مربع كاي للكشف عن انحرافات النموذج الانحدار اللوجيستي حيث يوجد جزء من الملاحظات لا يستند الي النموذج نظريًا والآخر متوقع محسوب من تقديرات النموذج اللوجيستي ومن ثم تحسب مربع كاي كمقياس لجودة التوفيق

مربع كاي	المعنوية
٠.٠١٩١	١.٠٠٠

نلاحظ أن المعاملات المشتركة المتغيرات تظل إيجابية، والقيم p-value صغيرة أقل من ٠.٠٠٥، ومعايير Akaike (AIC) أصغر بشكل هامشي، وعلى العموم ربما لا يوجد سبب واضح لتفضيل نموذج واحد على الآخر. ومع ذلك، وهذه هي النقطة الرئيسية، بالنسبة لمعظم مستخدمي البرامج الإحصائية في العلوم الاجتماعية، لا تبقى هناك خيارات أخرى لنمذجة حقيقية فيما يتعلق بالنمذجة اللوجستية لهذه البيانات.

٨- **التوصيات :-**

استنادًا إلى النتائج التي تم التوصل إليها، يمكن استخلاص عدة توصيات بهدف التخفيف من المشكلات الناتجة عن الإجهاض لدي السيدات، والتي تؤثر تأثيرًا مباشرًا على صحتهن:

١. أن تكون هناك فترة زمنية بين الإنجاب والإنجاب الأخر، وهذه الفترة يمكن أن تكون ٢-٣ سنوات لكي تعطي درجة من الراحة للمرأة الوالدة.
٢. يتطلب الحمل والإنجاب وجود بيئة سكنية هادئة وآمنة يمكن أن تحافظ على مقومات الحمل وتحافظ على نجاحه لكي ينشئ الطفل في كنف الرعاية والصحة والحنان التي تمكنه من العيش السليم وسط أجواء إيجابية.
٣. تؤكد الدراسة على أهمية توفير بيانات تمكننا من دراسة تأثير أبعاد العوامل الصحية المختلفة سواء من حيث دور الأمراض، وسلوك السيدة الصحي وكذلك مستوي الخدمة والرعاية المقدمة.

٤. الحث على تجنب الممارسات الضارة التي يأتي من ضمنها الزواج المبكر ودعوة على الحرص من الإنجاب في سن مبكر.
٥. استخدام أساليب إحصائية متقدمة أخري لدراسة العوامل التي تؤثر على فقد الأجنة ومقارنة نتائجها مع الأساليب المستخدمة في الدراسة.
٦. إدخال متغيرات أخري للدراسة مثل الوزن، التدخين، صلة القرابة بين الزوجين وغيرها من المتغيرات التي تؤثر على هذا الفقد.

٩-المراجع:

أولًا: المراجع العربية: -

١. أحمد عبد المنعم، ماجدة محمد عبد الحميد (٢٠٠٧)،" **أثر العوامل الاقتصادية والاجتماعية على استخدام خدمات رعاية الأمومة في بعض الدول العربية،"** المؤتمر السنوي الثاني والأربعون ٢-٥ ديسمبر ص ٤٦ جامعة القاهرة.

٢. رشا أبراهيم محمد (٢٠١٢)،"**مقارنة بعض طرائق تقدير المعلمات ودالة المعولية لتوزيع Birnbau-Saunders ذو معلمتين باستخدام المحاكاة**"، رسالة ماجستير، كلية الاقتصاد والإدارة – جامعة بغداد.

٣. كريم خلف عزر (٢٠١٦)، "**استخدام الانحدار اللوجيستي والتحليل التميزي لدراسة حالات الإصابة بمرض الإسهال لدي الأطفال في العراق،** "رسالة ماجستير، كلية التجارة – جامعة المنصورة.

٤. منتظر محمد عبد الحميد (٢٠١٥٠)،"**دراسة إحصائية لأثر التلوث البيئي على الصحة العامة في العراق**"، رسالة ماجستير غير منشورة، كلية التجارة – جامعة المنصورة.

٥. مي حامد محمود (٢٠١٦)،"**المشكلات الاجتماعية الناتجة عن الإجهاض المتكرر لدى السيدات وتصوير لدور الخدمة الاجتماعية في التخفيف منها**"، رسالة ماجستير غير منشورة، كلية الخدمة الاجتماعية – جامعة الفيوم.

ثانيا: المراجع الأجنبية:

6. Michael Ghapman (2011)." **Recurrent Miserriage**, Australia National in Ferrility Network ", p1.
7. Milja Kaare (2009).''**Genetic Studies on Reurrent Miscarriage Medical Genetics Universit**", Helsinki, Finland, p11.
8. Syed E. Ahmed, Kamon Budsaba & Supranee Lisawadi ,(2008).''**Parametric Estimation For The Birnpaum-Saunders Lifetime Distrib Ution Based On New Parametric Zation**" Thailand stat istician ,6(2), Vol . 213-2

9. Rodrıguez, G. (2010)." **Parametric Survival Models**", Technical report, Princeton: Princeton University.