جمهورية العراق
وزارة التعليم العالي والبحث العلمي
جامعة القادسية
كلية التربية / قسم الرياضيات
الدراسات الصباحية

تحليل الانحدار اللوجستي

بحث مقدم من قبل الطالبان

سارة ناصر علي

احسان سعد طارش

قدم كجزء من متطلبات لنيل شهادة البكالوريوس

بأشراف الاستاذة

م. كوركيس شهيد

١٤٤٠ هـ ٢٠١٩ م

الاهـــــداء

الى من بلغ الرسالة وادى الامانة ونصح الامة .. الى نبي الرحمة .. ونور العالمين

سيدنا محمد (صلى الله علية واله وسلم)

****

الى من اعطاه الله بالهيبة والوقار .. الى من علمني العطاء دون انتظار

الى من احمل اسمه بكل افتخار .. الى من حصد الاشواك عند دربي

ليمهد لي طريق العلم .. والدي العزيز

****

الى ملاكي في الحياة .. الى معنى الحب والحنان والتفاني

الى بسمة الحياة وسر الوجود .. الى من كان دعاءها سر نجاحي وحنانها بلسم جراحي

الى القلب الناصع بالبياض .. والدتي العزيزة

****

الى القلوب الطاهرة الرقيقة والنفوس البريئة الى رياحين حياتنا .. اخوتي واصدقائي

****

الى الذين بذلوا كل جهد وعطاء لكي اصل الى هذه اللحظة .. اساتذتي الكرام

الى كل من ساعدني في انجاز هذا العمل شكري الجزيل وامتناني

الى الذين يضحون بأنفسهم لأجلنا

الى جيشنا الباسل والحشد الشعبي المقدس

الشـــكر والتقـــــدير

لابد لنا ونحن نخطو خطواتنا الأخيرة في الحياة الجامعية من وقفة نعود الى اعوام قضيناها في رحاب الجامعة مع اساتذتنا الكرام الذين قدموا لنا الكثير باذلين بذلك جهوداً كبيرة في بناء جيل الغد لتبعث الامة من جديد وقبل ان نمضي نتقدم بأسمى آيات الشكر والامتنان والتقدير والمحبة الى الذين حملوا أقدس رسالة في الحياة ... الى الذين مهدوا لنا طريق العلم والمعرفة

الى جميع اساتذتنا الافاضل .... واخص بالشكر والتقدير :

د. كوركيس شهيد

بسم الله الرحمن الرحيم

((قُلْ اعملوا فسيرى الله عملكم ورسوله والمؤمنين))

صدق الله العظيم

الفهرست

المستخلص

ان اسلوب الانحدار اللوجستي هو اوسع الاساليب الاحصائية المستخدمة كأنموذج في مجال تحليل البينات حيث لا يفترض على الانحدار اللوجستي اي شروط تخص توزيع المتغيرات المستقلة فهو اسهل واكثر حصانه من غيره ويتسم بالمرونة والبساطة حيث تفسيره للبيانات واضحاً وذات دلالة . ومعنى الوصف العلاقة ما بين متغيرين هو التابع والمستقل (التفسيرية) حيث يكون المتغير التابع من النوع الاسمي (بمستويين او اكثر) اما المتغيرات المستقلة من الممكن ان تكون وصفية او كمية والهدف من استخدام هذا النموذج هو التنبأ لوجود صفة معينة او ظاهرة بالاعتماد على قيم متغير مستقل اخر التي لها علاقة بالمتغير التابع

المقدمة

من المألوف في الدراسات الانسانية والاجتماعية والاقتصادية ان يكون المتغير التابع متغيراً منفصلاً (نوعياً) ، بحيث يأخذ قيمة ثنائية Dichotomous او اكثر ، وان هذا يشكل تحدياً كبيراً للباحثين عند محاولتهم توظيف تحليل الانحدار الخطي (البسيط او المتعدد) ، الذي يكون مقيداً نوعاً ما باشتراط ان يكون المتغير التابع متغيراً بدلاً من ان يكون وصفياً منفصلاً .

ان هدف دراسة علم الانحدار هو معرفة وتفسير العلاقات المختلفة بين الظواهر من خلال شيئين أساسيين تحديد العلاقات بين المتغيرات الاستفادة من هذه العلاقات في مجال البحث العلمي

لذا يرى (Lea 1997) انه يجب استخدام تقنية الانحدار اللوجستي Logistic Analysis Technique في مثل هذه الحالات ،وانه وان كانت هنالك العديد من الاساليب الاحصائية التي طورت لتحليل البيانات ذات المتغيرات الوصفية (النوعية) مثل تحليل الدوال التمييزية Discriminant Functions Analysis ، الا ان الانحدار اللوجستي يتمتع بالعديد من المميزات التي تجعله ملائما للاستخدام في مثل تلك الحالات (بابطين ، ٢٠٠٩ ،٤) .

وتمكن اهمية تحليل الانحدار اللوجستي عند مقارنته بالأساليب الاحصائية الاخرى (الانحدار الخطي والتحليل التمييزي) ، في ان الانحدار اللوجستي هو اداة اكثر قوه لأنه يقدم اختباراً لمعنوية المعاملات ، كما انه يعطي الباحث فكرة عن مقدار تأثير المتغير المستقل في المتغير التابع النوعي ثنائي القيمة ، بالإضافة الى ذلك ، فأن الانحدار اللوجستي يرتب تأثير المتغيرات المستقلة ، مما يسمح الباحث بالاستنتاج بأن متغيراً ما يعتبر اقوى من المتغير الاخر في فهم ظهور النتيجة المطلوبة ، كما ان تحليل الانحدار اللوجستي يمكنه ان يتضمن المتغيرات المستقلة النوعية وكذلك تأثير التفاعل بين المتغيرات المستقلة في المتغير التابع ثنائي القيمة ، كما ان من مزايا استخدام الانحدار اللوجستي هو انه اقل حساسية تجاه الانحرافات عن التوزيع الطبيعي لمتغيرات الدراسة ، وذلك مقارنة بأساليب احصائية اخرى مثل التحليل التمييزي والانحدار الخطي ،، كما ان الانحدار اللوجستي يستطيع ان يتجاوز العديد من الافتراضيات المقيدة لاستخدام طريقة المربعات الصغرى OLS في الانحدار الخطي ، الامر الذي يجعل في نهاية المطاف تحليل الانحدار اللوجستي الاسلوب الافضل في حالة المتغير التابع الفئوي ثنائي القيمة .

الفصل الاول

الانحدار Regression

الفصل الاول

الانحدار Regression

يعتبر العالم الانكليزي francis galton اول من استخدم مفهوم الانحدار في التطبيقات البيولوجية بهدف اكتشاف بعض العلاقات بين بعض المتغيرات البيولوجية.

١-١ تعاريف اساسية

١- المصفوفة Matrix

المصفوفة:A هي عباره عن عناصر مرتبة في قائمة مستطيلة الشكل مثل

$$A=\begin{pmatrix} 1 & 2 & 4 \\ 5 & 9 & 8 \end{pmatrix}$$

العنصر (i,j) للمصفوفة A يرمز له a_{ij}

بعد dimension او حجم المصفوفة هو n (عدد الاسطر) وm(عدد الاعمدة) .

اذا كان $n = m$ فتسمى A مصفوفة مربعة .لكي نبين ابعاد المصفوفة نكتبها $A_{n\times m}$

٢- المتجه Vector

المتجه: هو مصفوفة تتكون من عمود واحد او سطر واحد . يكتب متجه عمود على الشكل

a_{n*1} ومتجه سطر على الشكل $a_{1\times n}$ سوف نفترض عند التكلم عن متجه انه متجه سطر مالم يذكر غير ذلك

$$A=(a1\ a2\cdots an)$$

٣- منقول الصفوفة Transpose

اذا كانت $A=(a_{ij})$ مصفوفة ذات ابعاد $n*m$ فان منقول A ويرمز له بالرمز A' هو المصفوفة ذات الابعاد $m*n$ والتي عناصرها a_{ji}

٤- المصفوفة المتناظرة Symmetric

اذا كان $A' = A$ فانه يقال ان المصفوفة متناظره

٥- مصفوفة الوحدة Idenitity Matrix

مصفوفة الوحدة I ذات الابعاد $n*n$ تكتب

$$I = I_n = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \ddots & 0 \\ 0 & 0 & \cdots & 1 \end{bmatrix}$$

٦- مصفوفة الاحاد Ones Matrix

مصفوفة الاحاد J ذات الابعاد $n*n$ تكتب

$$J = j_n \begin{bmatrix} 1 & 1 & \cdots & 1 \\ 1 & 1 & \cdots & 1 \\ 1 & 1 & \cdots & 1 \end{bmatrix}$$

٧- المصفوفة الصفرية Null Matrix

مصفوفة الاصفار O ذات الابعاد $n*n$ تكتب

$$O=O_n\begin{bmatrix}0 & \cdots & 0\\ \vdots & \ddots & \vdots\\ 0 & \cdots & 0\end{bmatrix}$$

٨- المقلوب Invers

اذا كانت A مصفوفة ذات بعد $n*n$ وكان يوجد مصفوفة C بحيث $AC = CA= I$

فان A مصفوفة غير شاذه والمصفوفة C تسمى مقلوب او معكوس A ويرمز له A^{-1}

اذا كانت A غير شاذه فان المقلوب يكون وحيد

$$A^{-1} = \frac{1}{|a|}\ \text{Adj } a$$

مثال/ جد معكوس المصفوفة $a=\begin{bmatrix}1 & 0 & -3\\ 2 & -2 & 1\\ 0 & -1 & 3\end{bmatrix}$

الحل/ $A^{-1} = \frac{1}{|a|}\ \text{Adj } a$

$$|a|=\begin{bmatrix}1 & 0 & -3\\ 2 & -2 & 1\\ 0 & -1 & 3\end{bmatrix}\begin{matrix}1 & 0\\ 2 & -2\\ 0 & 3\end{matrix}$$

=-6+6-(-1)

=1

$$\text{Adj } a=\begin{bmatrix}-5 & 3 & -6\\ -6 & 3 & -7\\ -2 & -1 & -2\end{bmatrix}$$

$$A^{-1}=\frac{1}{1}\begin{bmatrix}-5 & 3 & -6\\-6 & 3 & -7\\-2 & -1 & -2\end{bmatrix}$$

$$A^{-1}=\begin{bmatrix}-5 & 3 & -6\\-6 & 3 & -7\\-2 & -1 & -2\end{bmatrix}$$

٢-١ تعريف

يعرف الانحدار بشكل عام بانه مقياس رياضي لمتوسط العلاقة بين متغيرين او اكثر بدلالة وحدات قياس المتغيرات المعتمدة في العلاقات وغالبا ماتسمى العلاقات من هذا النوع بنماذج الانحدار

٣-١ انواع الانحدار

١- الانحدار الخطي
٢- الانحدار اللاخطي

لا يكتفي الجغرافيون عند دراسة العلاقة بين متغيرين بتحديد درجة الارتباط بينهما سواء كان ذلك الارتباط موجبا ام سالبا بل يفضلون ان يضعوا تلك العلاقة على شكل معادلة رياضية محددة تلخص خصائص وسمات شكل الانتشار ويعرف المتغير الذي يستخدم في عملية التنبؤ بالمتغير المستقل(Indepndent varible) ويعرف المتغير الاخر الذي يتم التنبؤ به بالمتغير التابع (Dependdent varible)

المتغير العشوائي :هو دالة قيمتها اعداد حقيقية وتعتمد على الحظ بصوره اكثر دقة .فان المتغير العشوائي هو الدالة التي ترافق التجربة العشوائية وتكون معرفة على فضاء العينة للتجربة وقيمتها عدد حقيقي

٤-١ انواع الانحدار

١-الانحدار الخطي البسيط Simple Linear Regression

٢-الانحدار الخطي المتعدد Multiple Linear Regression

٣- الانحدار الاسي Exponential Regression

٤-الانحدار اللاخطي Nonlinear Regression

٥-الانحدار الضبابي Fozzy Regression

٦-الانحدار اللوجستي Logistics Regression

الفصل الثاني

انواع الانحدار

الفصل الثاني

انواع الانحدار

انواع الانحدار

١- الانحدار الخطي البسيط Simple Linear Regression

هو عبارة عن دالة رياضية يمكن بها حساب احد المتغيرين بدلالة المتغير الاخر وابسط حالة لهذه الدالة عندما يكون لدينا متغير مستقل واحد فقط يرتبط مع المتغير التابع بعلاقة خط مستقيم كما يلي :

$$\hat{y} = a + bx + e$$

حيث $\hat{y}$ التابع

a ثابت الانحدار

b معامل الانحدار

e الباقي

حيث b يمكن استخراجها من القانون بطريقة المربعات الصغرى

$$b = \frac{n\sum xy - \sum x \sum y}{n\sum x^2 - (\sum x)^2}$$

$$a = \frac{\sum y}{n} - b\frac{\sum x}{n}$$

البرهان//

$$\hat{y} = a + bx$$

نجد a,b بواسطة طريقة المربعات

$Q=\sum_{i=1}^{n} ei^2$

$$= \sum(y_i - a - bx_i)^2 \cdots\cdots\cdots\cdots\cdots (1)$$

نشتق Q بالنسبة الى a,b ومن ثم نضع الناتج مساويا للصفر نحصل على

$$\partial Q/\ \partial a = -2\sum(y_i - a - bx_i) = 0 \cdots\cdots\cdots (3)$$

$$\frac{\partial Q}{\partial b} = -2\sum x_i(y_i - a - bx_i) = 0 \cdots\cdots\cdots (4)$$

من المعادلات اعلاه نحصل على

$\sum y_i\text{-}na\text{-}b\sum x_i=0\cdots\cdots\cdots\cdots (5)$

$\sum x_i y_i\text{-}a\sum x_i\text{-}b\sum x_i^2 = 0 \cdots\cdots\cdots (6)$

من معادلة رقم 5 نجد ان

$a=\frac{1}{n}\sum y_i - \frac{b}{n}\sum x_i$

بالتعويض عن قيمة a بالمعادله (6) نحصل على

$$\sum x_i y_i - \frac{(\sum x_i)(\sum y_i)}{n} + \frac{b}{n}(\sum x_i)^2 - b\sum x_i^2$$

$b_1=\frac{n\sum x_i y_i - (\sum x_i)(\sum y_i)}{n\sum x_i^2 - (\sum x_i)^2}$

$a=\hat{y}\text{-}b_1\hat{x}_i$

الانحدار الخطي البسيط

الغرض من استخدام اسلوب تحليل الانحدار الخطي البسيط هو دراسة وتحليل اثر متغير كمي على متغير كمي اخر

لنعتبر النموذج الذي يكون فيه متغير الاستجابة y متصل خطيا بمتغير مستقل x بالعلاقة

$y_i=\beta 0+\beta 1x_i+\varepsilon_i$

$i=1,\cdots n$

حيث ε_i الخطأ العشوائي يكون توزيع متعدل وتعتبر دوما متغيرات عشوائية غير مترابطة

var = كميه ثابتة

$$\mu = 0$$

مثال// البيانات التالية تبين الكميات المعروضة من سعلة معينة

13	3	4	5	15	10	الكميهyالمعروضه
9	2	3	4	10	7	السعرx

المطلوب ١-تقدير معادلة الكمية المعروضة على السعر

٢-تقدير الكمية المعروضة عند السعر 20

الحل//

X^2	Xy	X	Y
49	70	7	10
100	150	10	15
16	20	4	5
9	12	3	4
4	6	2	3
18	117	9	13
∑X^2=259	∑Xy=375	∑X=35	∑Y=50

$b_1=\frac{n\sum xy-\sum x\sum y}{n\sum x^2-(\sum x)^2}$

$=\frac{6.357-35.50}{6.259-1.225}$

$=1.52$

$b_0=\frac{\sum y}{n}-b1\frac{\sum x}{n}$

$=\frac{50}{6}-1.52\frac{35}{6}$

$= -0.53$

معادلة الانحدار المقدرة

$$\hat{y}=\mathrm{b0}+\mathrm{b1x}$$

$=-0.53+1.52x$

٢- نقوم بالتعويض في معادلة الانحدار عن $x=20$

$$\hat{y}=-0.53+1.52(20)$$

$= 29.87$

طريقة تحليل الانحدار الخطي البسيط

طريقة المربعات الصغرى Least Square Method

تسمى ايضا بالمربعات الصغرى الكلاسيكية وتعد هذه الطريقة من الطرائق الواسعة الاستخدام في التطبيقات الاحصائية وتعتمد وجود علاقة بين متغيرين او اكثر ويستند مبدأ هذه الطريقة الى ايجاد ذلك الخط المستقيم الذي يتخلل نقاط الشكل الانتشاري بالشكل الذي يجعل مجموع مربعات ابعاد النقاط عنه اقل ما يمكن .

اي تحديد قيمه (β) التي تجعل هذا المجموع اقل ما يمكن في حاله الانحدار الخطي المتعدد وعملية تقدير العلاقة الخطية بين عدة متغيرات احداهما متغير الاستجابة والباقي متغيرات توضيحية .اما تقدير المعلمات بشكل عام يكون معادلة الانحدار في حالة (k) من المتغيرات التوضيحية بالشكل الاتي:

$$Y_{i=}b_0+b_1x_{i1}+b_2x_{i2}+\ldots+b_k\ x_{ik}+u_i$$

وفي حالة متغير توضيحي واحد فان

$$u_{i=}y_i\_b^{\wedge}{}_0\_b^{\wedge}{}_1\ x_{i\ldots.}$$

الشكل العام

$$Y=\beta 1+\beta 2x+u$$

$$Y_i=b1+b2xi+ui$$

$$\hat{y}=\hat{b}1+\hat{b}2xi$$

$$\hat{y}=\hat{b}_1+\hat{b}_2x$$

الغاية من المربعات الصغرى هي تقليل المسافة

Y الجزء الفعلي ، $\hat{y}$الجزء المقدر

$U=y-\hat{y}$مقدار الخطأ

$\sum U^2=\sum(y-\hat{y})^2$ مجموع مربعات الخطأ

$$\sum U^2=\sum(y-\hat{b}_1-\hat{b}_2x)^2$$

لكي تقدر المعادلة الى اقرب مايكون الى الواقع نقلل مجموع مربع هذه الاخطاء وذلك بأخذ المشتقة الاولى ونساويها الى الصفر

$$\frac{\partial u^2}{\partial b1}=0$$

$$\frac{\partial u^2}{\partial b2}=0$$

بأجراء العميات نحصل على القيم

$b_1=\bar{y} - b_0\bar{x}$

$$b2=\frac{\sum(x-\bar{x})(y-\bar{y})}{\sum(x-\bar{x})\ ^\wedge 2}$$

من هذه القوانين نستطيع الحصول على افضل مقدرات للمربعات الصغرى

مثال/ ماهي معادلة الاتجاه العام للبيانات المبينة في الجدول التالي

xالسنه	X	Y	xy	X^2
1985	1	40	40	1
1986	2	33	66	4
1987	3	29	87	9
1988	4	25	100	16
1989	5	21	105	25
1990	6	32	192	36
1991	7	40	280	49
1992	8	45	360	64
1993	9	41	369	81
1994	10	40	400	100
	$\bar{X}$=55	$\bar{Y}$=346	$\sum$Xy=1999	$\sum$X^2=385

باستخدام الصيغ السابقة في طريقة المربعات الصغرى

$$b=\frac{n\sum xy-\sum x\sum y}{n\sum x^2-(\sum x)^\wedge 2}$$

$$=\frac{10(1999)-55\times 346}{10(385)-(55)^\wedge 2}$$

$$b=\frac{960}{825}$$

b=1.16

a=$y` - b\grave{x}$

a=34.6-6.38

a=28.22

$\hat{y} = a + bx$

=28.22+1.16x

هذه معادلة الاتجاه العام لقيم الاجور للفتره المبينة في الجدول

الانحدار غير الخطي البسيط Non Linear Regression

سبق دراسة الانحدار الخطي البسيط y=a+bx وحسبنا a,b

ويختلف غير الخطي البسيط عن الخطي البسيط كمايلي

المعامل a الثابت ليس حدا كما في y=a+bx

فظهر بالصوره y=ax^b يمكن تحويلها لخطية (لوغاريتم)

او y=ab^x يمكن تحويلها لخطية (لوغاريتم) او كمعادلة لوغاريتمي y=a+ln x

كما يمكن تحويلها لخطية بتغير تعريف المتغيرات فمثلا y=ax^b باخذ لوغاريتم الطرفين نحصل على

$$\ln y = \ln a + b \ln x$$

وذلك حسب قوانين الاسس وبوضع ln=a,lny=y,lnx=x

نحصل على المعادلة المرادفة

Y=a+bx

وبالتالي نحسب قيم a,b

مثال/ الجدول التالي يبين استهلاك 10 عائلات من اللحوم بالكيلوجرام (y) ومعدل دخلها الشهري بالدينار x والمطلوب تقدير معادلات الانحدار لأستهلاك اللحوم بدلالة الدخل الشهري ثم قدر استهلاك اللحم لعائلة معدل دخلها 250. استخدم المعادلة الاسية $y=ax^b$

الاستهلاك بالكيلوجرام	الدخل الشهري بالدينار	العائلة
45	250	1
42	300	2
44	290	3
52	255	4
50	260	5
55	250	6
40	200	7
40	220	8
50	270	9
60	300	10

الحل/

$Y=ax^b$ نحولها لمعادلة لوغارتيمية باخذ اللوغاريتم للطرفين فنحصل على

$$\ln y = \ln(ax)^b$$

$$=\ln a + \ln(x)^b$$

$$=\ln a + b\ln x$$

$$\ln y = \ln a + b\ln x$$

العائلة	الدخل الشهري x	الاستهلاك بالكيلوجرام y	X^2	$\ln x^2$	lnx	Lny	lnxlny
1	250	45	62500	11.0429	5.5215	3.8067	21.0183
2	300	42	90000	11.4076	5.7038	3.7377	21.3189
3	290	44	84100	11.3398	5.6699	3.7842	21.4559
4	255	52	65025	11.0825	5.5413	3.9512	21.8949
5	260	50	67600	11.1214	5.5607	3.9210	21.7535
6	250	55	62500	11.0429	5.5215	4.0073	22.1263
7	200	40	40000	10.5966	5.2983	3.6889	19.5449
8	220	40	48400	10.7873	5.3936	3.6889	19.8964
9	270	50	72900	11.1968	5.5984	3.9120	21.9012
10	300	60	90000	11.4076	5.7038	4.0943	23.3533

نحسب a,b من الصيغ السابقة

$$b=\frac{n\sum \ln x \ln y-\sum \ln x \sum \ln y}{n\sum(\ln x^2)-(\sum(\ln x))\text{^}2}$$

$$\ln a=\frac{\sum \ln y-b\sum \ln x}{n}$$

$$b=\frac{10(214.2635)-(55-5127)(38.5832)}{10(111.0254)-(55.5127)^2}$$

b=-0.0004

$$\ln a=\frac{38.5832-(-0.0004)(55.5127)}{10}$$

ln a=3.8605

$a=e^{3.8605}$

$a=47.4891\cong 47.5$

$Y=47.5x-0.0004$

والاستهلاك المطلوب لعائلة دخلها *250* دينار يكون

$Y=47.5x^{-0.0004}$

$Y=47.5*0.9978$

$Y=47.3955 \cong 47$

الانحدار الاسي Exponential Regression

في بعض الاحيان يكون تمثيل او توافق البيانات بخط مستقيم لايعطينا الدقة المطلوبة لذلك من المستحسن تمثيل البيانات بواسطة منحني انحدار غير خطي ومن ثم ايجاد شكل هذا المنحني وتقدير معلماتة .

كما ذكرنا سابقا فان شكل الانحدار يعطينا بعض الملامح عند شكل منحني الانحدار اذا كان خطيا او غير خطي.

لذا فان شكل الانحدار في هذه الحالة يؤشر لنا بان المعدل $\mu\, y/x$ يكون الافضل تمثيلة في حالة الانحدار غير الخطي بالصورة التالية

$$\mu\, y/x = n\lambda^{x}$$

حيثn, λ هما معلمتا المنحني ويجب تخمينهما اذا فرضنا ان مخمن n هوc

ومخمن λ هو d

فأننا نخمن الوسط $\mu\, y/x$ بالقيمة y

$Y=cd^{x_i} \cdots\cdots (1)$

نأخذ اللوغاريتم الاعتيادي للمعادلة رقم *(1)*

$$\log y = \log c + (\log d)x_i$$

وكل زوج من المشاهدات الموجودة في العينة يحقق الاتي

$$\log y = \log c + (\log d)x_i + e_i$$

حيث e_i هي الخطأ

نستطيع ان نكتب

$$\log y = a + bx_i + e_i$$

حيث ان $b=\log d, a = \log c$

وهي تشبه معادلة خط الانحدار والفرق الوحيد بينهما هو ان في خط الانحدار كانت البيانات على شكل (x_i,y_i) بينما هنا على شكل $(x_i,\log y_i)$

فلهذا نستطيع ايجاد قيمة a,b من المعادلة السابقة في طريقة المربعات الصغرى وذلك بعد تغيير كل y بالقيمة $\log y_i$ ومن ثم ايجاد a,b .

وبعد ذلك نستطيع ايجاد قيمة d,c من العلاقة

$C=10^{a}$

$d=10^{b}$

وعند التعويض في معادلة رقم(1) نحصل على معادلة الانحدار الأسي.

مثال/

18	15	12	9	6	5	2	x
2.6	5.3	7.9	8.4	9.8	10.7	12.5	y

وافق الجدول اعلاه لمنحني انحدار أسي على شكل $\mu\, y/x = cd^{x}$ ثم جد y عندما x=10

الحل/

x	Y	$\log y$	x^2	$x\log y$
2	12.5	1.097	4	2.194
5	10.7	1.029	25	5.145
6	6.8	0.991	36	5.96
9	8.4	0.924	81	8.316
12	7.9	0.898	144	10.776
15	5.3	0.724	255	10.86
18	2.6	0.415	324	7.47
$\sum X=67$	$\sum Y=57.2$	$\sum \log y = 6.078$	$\sum x^2 = 839$	$\sum=50.707$

الان نجد a,b من المعادلات الطبيعية فنحصل على

$$b=\frac{(7)(50.707)-(67)(6.087)}{(7)(839)-(67)^2}$$

$$=-0.038$$

$$a=\overline{\log y}+b\bar{x}$$

$$\bar{x}=\frac{67}{7}=9.571$$

$$\overline{\log y}=\frac{6.078}{7}$$

$$=0.868$$

$$=0.868+(0.038)(9.571)$$

$$=1.232$$

$$C=10^{a}$$

$$=(10)^{1.232}=17.061$$

$$d=10^{b}$$

$$=(10)^{-0.038}$$

$$=0.916$$

$$Y=(17.061)(0.916)^{x}$$

$$Y_{10}=(17.061)(0.916)^{10}$$

$$=7.1$$

الانحدار الخطي المتعددMultiple Linear Regressio

ان نموذج الانحدار الخطي المتعدد هو عباره عن انحدار المتغير التابع (y) على العديد من المتغيرات المستقلة $x1, x2 \cdots xn$

ويسمى هذا بنموذج الانحدار الخطي المتعدد multiple linear regression

ويهدف هذا المقال الى توضيح كيفية تقدير نموذج الانحدار الخطي المتعدد ثم تحديد اهم افتراضات النموذج ،يضاف الى ذلك بيان عدم وجود علاقة خطية تامة بين المتغيرات المستقلة وكيف ان المصفوفة (x) تكون مصفوفة غير شاذة (non-singlar) اذا كان محددها لايساوي صفرا. ثم يتم بعد ذلك تقدير معلومات النموذج ، تقدير التباين والتباين المشترك والانحراف المعياري لها للوصول الى اختيار معلمات النموذج.

نموذج الانحدار الخطي المتعدد:

يستند النموذج الخطي المتعدد على افتراض وجود علاقة خطية بين متغير تابع Y_i وعدد من المتغيرات المستقلة $x1, x2 \cdots xn$ وحد عشوائي U_i . ويعبر عن هذه العلاقة بالنسبة لـn من المشاهدات وk من المتغيرات المستقلة بالشكل الاتي

$$Y_i=b_0+b_1x_{i1} +b_2x_{i2} +\cdots + b_kx_{ik}+u_i \cdots\cdots (1)$$

وفي واقع الامر فأن هذه المعادلة هي واحده من جملة معادلات يبلغ عددها (n) تكون نظام المعادلات الاتي :

$$Y_1=b_0+b_1x_{11} +b_2x_{12} +\cdots + b_kx_{1k} +u_1$$

$$Y_2=b_0+b_1x_{21}+b_2x_{22}+\cdots + b_kx_{2k}+u_2$$

$$\vdots \quad \vdots \qquad \vdots \qquad \vdots$$

$$Y_n=b_0+b_1x_{n1}+b_2x_{n2}+\cdots + b_kx_{nk} +u_n$$

وهذه المعادلات تتضمن (k+1) من المعادلات المطلوب تقديرها علما ان الحد الاول منها (b_0) يمثل الحد الثابت ، الامر الذي يتطلب اللجوء الى المصفوفات والمتجهات لتقدير تلك المعلمات.

عليه يمكن صياغة هذه المعادلات في صوره مصفوفات وكالاتي:

$$Y=XB+U$$

Y=تحوي (n+1)مشاهدات المتغير متجه عمودي

X= تحوي مشاهدات (1+nxk) مصفوفة ابعادها المتغيرات المستقلة يحتوي عمودها الاول على قيم الواحد الصحيح ليمثل الحد الثابت

B= يحتوي على المعالم (11x+k) متجه عمودي ابعاده المطلوب تقديرها

U= يحتوي على الاخطاء (1xn) متجه عمودي ابعاده عشوائية

وبما ان المعادلة (1) هي العلاقة الحقيقة المجهولة والمراد تقديرها بأستخدام الاحصاءات المتوفرة عن المتغير التابع y والمتغيرات المستقلة $x1, x2 \cdots x3$

مثال/ في دراسة على تذوق جبنة الشدر حللت $n=20$ عينة لمحتواها الكيميائي واخضعت لأختبارات تذوق وجمعت بيانات عن المتغيرات التالية

$y=$ نتيجة اختبار التذوق (*TASTE*)

$X_1=$تركيز حامض الخليك (*ACETIC*)

$X_2=$ تركيز سولفيت الهيدروجين (H_2S)

$X_3=$ تركيز الحامض اللبني(LACTIC)

وضعت بالجدول التالي

TASTE	ACETIC	H_2S	LACTIC	TASTE	ACETIC	H_2S	LACTIC
11.3	4.543	3.135	0.86	40.9	6.365	9.588	1.74
19.9	4.159	5.043	1.53	15.9	4.787	3.912	1.16
38.0	5.366	5.438	1.57	6.4	5.412	4.700	1.49
48.9	6.759	7.496	1.81	18.0	5.247	6.174	1.63
5.7	5.663	3.807	0.99	38.9	5.438	9.064	1.99
25.9	5.697	7.601	1.29	14.0	4.564	4.949	1.51
26.9	6.082	8.72	1.30	15.2	5.298	5.22	1.33

		6				0	
38.3	4.230	7.966	1.09	32.0	5.455	9.242	1.44
20.9	4.210	3.850	1.57	56.7	5.855	10.20	2.01
19.1	5.321	4.174	1.59	16.8	5.366	3.664	1.31
20.0	7.021	6.908	1.86	11.6	6.043	3.219	1.46
33.9	6.121	2.996	1.06	26.5	6.458	6.962	1.72
58.2	7.025	4.942	1.30	0.7	5.328	3.912	1.25
0.8	9.012	6.142	1.52	13.4	5.802	6.685	1.08
55.9	6.151	6.752	1.76	5.5	6.176	4.781	1.25

المتغيرات ACETIC و H_2S ممثلة بالتدريج اللوغارتمي الطبيعي ،المتغير LACTIC لم يتم تحويلة. لنفرض ان الباحثين وضعوا فرضية ان كل من الثلاثة المحتويات الكيميائية المتغيرة $x1, x2, x3$ مهمه في وصف التذوق . وفي هذه الحالة فقد اعتمدوا مبدئيا نموذج الانحدار التالي

$$y_i = \beta 0 + \beta 1 x11 + \beta 2 x12 + \beta 3 x13 + \varepsilon_i$$

لقيم i=1,2,…20

في الصيغة المصفوفية $Y=XB+\varepsilon$

يكتب

$$Y_{20*1}=\begin{pmatrix} 11.3 \\ 19.9 \\ \vdots \\ 55.9 \end{pmatrix}$$

$$X_{20*4}=\begin{bmatrix} 1 & 4.543 & 3.315 & 0.86 \\ 1 & 4.159 & 5.043 & 1.53 \\ 1 & 5.366 & 5.438 & 1.57 \\ \vdots & \vdots & \vdots & \vdots \\ 1 & 6.151 & 6.752 & 1.76 \end{bmatrix}$$

$$\beta_{4*1}=\begin{pmatrix} \beta 0 \\ \beta 1 \\ \beta 2 \\ \beta 3 \end{pmatrix}$$

$$\varepsilon_{20*1}=(\varepsilon 1, \varepsilon 2, \cdots \varepsilon 20)$$

يمكن استخدام الانحدار الخطي المتعدد في حالة توافر الشروط التالية

١- ان تكون العلاقة خطية بين المتغيرات المستقلة والمتغير التابع.

٢- ان تكون البيانات موزعه توزيعا طبيعيا للمتغيرات المستقلة والمتغير التابع.

٣- يجب ان تكون قيم المتغير التابع من المستوي الترتيبي على الاقل.

الفصل الثالث

الانحدار اللوجستي

Logistics Regression

الفصل الثالث

الانحدار اللوجستي Logistics Regression

التوزيع السوقي (اللوجستي)(Logistics Distribution)

تبرز استخدامات هذا التوزيع وبشكل خاص في الدراسات المتعلقة بعلوم الحياة والعلوم الزراعية والطبية وبشكل عام في الدراسات ذات الطابع التجريبي .

يعرف هذا التوزيع:

يقال ان المتغير العشوائي x هو ذو توزيع سوقي اذا كانت دالة الكثافة الاحتمالية لهذا المتغير هي:

$$f(x;\alpha,\beta)=\frac{1}{4\beta}.\sec h^2\left[\frac{1}{2}\left(\frac{x-\alpha}{\beta}\right)\right]$$

$$=\frac{1}{\beta}\left[e^{-\left(\frac{1-\alpha}{\beta}\right)}\right].\left[1+\left(\frac{x-\alpha}{\beta}\right)\right]^2 ;-\infty<x<\infty$$

حيث β,α معلمتي التوزيع وان $\beta>0,-\infty<\alpha<\infty$

وبالرموز فان $X\sim LOGISTIC(\alpha,\beta)$

وفيما يلي خصائص التوزيع

١- ان الدالة التوزيعية لهذا التوزيع هي:

$$f(x)=p_{\mathrm{r}}(x\le x)=\frac{1}{2}\left[1+\tan h\left[\frac{1}{2}\left(\frac{x-\alpha}{\beta}\right)\right]\right]$$

$$=\left[1+e^{-\left(\frac{x-\alpha}{\beta}\right)}\right]^{-1}$$

كذلك يمكن التعبير عن f(x) بدلالة F(x) من خلال العلاقة التالية

$$f(x) = \frac{1}{\beta}\mathrm{F}(x)(1 - f(x))$$

٢- ان الدالة المولدة للعزوم x حول نقطة الاصل هي :

$M_x = e^{xt}.\Gamma(1-\beta).\Gamma(1-\beta t)$

$= e^{xt}.(\pi\beta t).\csc(\pi\beta t)$

ومن خلال هذه الدالة يمكن بيان ان

$\mu_x = EX = M_{x'}(0) = \alpha, Ex^2$

$= M_x''(0) = \alpha^2 + \frac{\pi^2\beta^2}{3}$

وان

$$\sigma^2 = \frac{\pi^2\beta^2}{3}$$

٣- يتحقق المنوال في هذا التوزيع عندما $x=\alpha$ وهذا ناتج من حل المعادلة التفاضلية $f'(x)=0$

بشرط ان $f''(x) < 0$.وعندئذ فان

$\mathrm{Max}.f(x) = f(x)]x = \alpha = \frac{1}{4\beta}$

٤- كذلك يتحقق الوسيط في هذا التوزيع عندما $x=\alpha$ وهذا ناتج من حل الصيغة $F(x)=\frac{1}{2}$. نلاحظ في هذا التوزيع تساوي الاوساط الثلاثة (الوسط،الوسيط،المنوال) وهذا يعني ان منحنى دالة هذا التوزيع متماثل حول المحور $x=\alpha$.

٥- اذا كانت $\alpha = 0, \beta = 1$ عندئذ نحصل على الشكل المعياري لهذا التوزيع وفي هذه الحالة تكون:

$f(x) = e^{-x}(1 + e^{-x})^{-2} = \frac{1}{4}\mathrm{sech}^2\frac{x}{2}; -\infty < x < \infty$

وان

$$F(x)=[1+e^{-x}]^{-1}=\frac{1}{2}\left[1+\tanh\frac{x}{2}\right]$$

كذلك فان

$$M_x(t)=\Gamma(1-t).\Gamma(1+t)=\pi t.\csc \pi t$$

وان

$$\mu_x=0, \sigma^2=\frac{\pi^2}{3}$$

التحليل اللوجستي Logistics Analysis

في السنوات الاخيرة ازدادت اهمية التحليل اللوجستي في مجال تحليل البيانات المصنفة وخاصة في مجالات البحوث الطبية والاجتماعية والزراعية وغيرها. وذلك لكونه يهتم بتحليل البيانات ذات الاستجابة الثنائية والتي يكون فيها متغير الاستجابة (Response Variable) ثنائيا (Binary) ،حيث حيث حالة النجاح يكون فيها متغير الاستجابة يأخذ القيمة (1) وحالة الفشل ياخذ القيمة (0)

مفهوم الانحدار اللوجستي ثنائي الاستجابة

يبنى نموذج الانحدار اللوجستي على فرض اساسي هو ان المتغير التابع (y) متغير الاستجابة تهتم بدراسته هو متغير ثنائي يتبع توزيع برينولي (Bernoulli) ،يأخذ القيمة (1) باحتمال مقداره (π) والقيمة (0) باحتمال ($1-\pi$) اي الى حدوث الاستجابة وعدم حدوثها وكما نعلم في الانحدار الخطي الذي نأخذ متغيراتة المستقلة والمتغير التابع قيما مستمرة فأن النموذج الذي يربط بيم المتغيرات هو على النحو الاتي :

$$Y=B_0+B_1x+\epsilon \cdots\cdots (22)$$

اذان (y) يمثل متغيرا مشاهدا مستمرا وبفرض ان متوسط قيم (y) المشاهدة او الفعلية عند قيمة معينه للمتغير (x) هي E(y) فانه يمكن كتابة النموذج على النحو التالي:

$$E(y/x) = \beta 0 + \beta 1x \cdots\cdots (23)$$

ومن المعروف في الانحدار ان الطرف الايمن لهذا النموذج يأخذ قيما من $(-\infty)$الى (∞)ولكن عندما يكون لدينا متغيران احدهما ثنائي (y) فان نموذج الانحدار الخطي البسيط لايكون ملائما لان

$$E(y/x) = p_r(y=1)\pi \cdots\cdots (24)$$

وبذلك تكون قيمة الطرف الايمن محصورة مابين الرقمين (0,1) وبذلك يكون النموذج غير قابل للتطبيق من وجهة نظر الانحدار. وان احدى طرائق حل هذه المشكلة هو ادخال تحويلة رياضية مناسبة على المتغير التابع.

(y) ومن المعروف ان $(1 \le \pi \le 0)$ ومن ثم النسبه$(\frac{\pi}{1-\pi})$ هي عبارة عن مقدار موجب محصور بين $(\infty, 0)$ اي$(\infty \le \frac{\pi}{1-\pi} \le 0)$ وبأخذ اللوغاريتم الطبيعي للاساس (e) للتحويل $(\frac{\pi}{1-\pi})$ فان مجال قيمه تصبح محصورة $(\infty \le \log_e(\frac{\pi}{1-\pi}) \le -\infty)$ وعلية يمكن كتابة نموذج الانحدار في حالة متغير مستقل واحد على النحو الاتي:

$$\log_e(\frac{\pi}{1-\pi}) = \beta 0 + \beta 1x \cdots\cdots\cdots (25)$$

واذا كان لدينا p من المتغيرات المستقلة فان النموذج يكون

$$\log_e\left(\frac{\pi}{1-\pi}\right) = \beta 0 + \sum_{j=1}^{p} \beta j x i j \cdots\cdots\cdots (26)$$

اذان i=1,2,….n

J=1,2,…p

ويسمى هذاالنموذج بنموذج الانحدار اللوجستي ثنائي الاستجابة وتسمى التحويلة $\log_e(\frac{\pi}{1-\pi})$

بتحويلة لوجيت |(logit transformation) وان الدالة اللوجستية هي دالة مستمرة تأخذ القيم ما بين (1,0) وتقترب y من الصفر كلما اقترب الطرف الايمن للدالة اللوجستية من $(-\infty)$

وتقترب y من الواحد كلما اقترب الطرف الايمن لهذه الدالة من (∞) وهي دالة متماثلة عندما يكون الطرف الايمن لهذه الدالة مساويا للصفر .

تسمى النسبه ($\frac{\pi}{1-\pi}$) بنسبه الاقضليه او افضليه النجاح (odds of success) اونسبة الافضلية للحدث المرغوب والنسبة ($\frac{\pi}{1-\pi}$) يمكن ان تسمى ايضا نسبة افضلية الفشل (odds of failure)

وان المقدار $\log_e(\frac{\pi}{1-\pi})$ يسمى لوغاريتم نسبة الافضلية (log odds ratio) اواللوجيت

-الصيغة الاحتمالية لنموذج الانحدار اللوجستي ثنائي الاستجابة

يمكن كتابة الانحدار اللوجستي ثنائي الاستجابة بالصيغة الاحتمالية وذلك برفع طرفي المعادلة(26) للأساس (e) ونحصل على

$$\left(\frac{\pi}{1-\pi}\right) = e^{\beta 0+\sum_{j=1}^{p}\beta jxij} \cdots\cdots (27)$$

$$\frac{1}{\frac{1}{\pi}-1} = \frac{1}{e^{(\beta 0+\sum_{j=1}^{p}\beta jxij)-1}} \cdots\cdots\cdots (28)$$

وبأستخدام الطرائق الجبرية فان المعادلة (28) يمكن ان تكتب كالاتي :

$$\pi = \frac{1}{1+e^{(\beta 0+\sum_{j=1}^{p}\beta jxij)-1}} \cdots\cdots (29)$$

وبالتالي فان احتمال متغير الاستجابة y ياخذ القيمة (1)

$$P(y=1/x) = \frac{1}{1+e^{(\beta 0+\sum_{j=1}^{p}\beta jxij)-1}} \cdots\cdots\cdots\cdots (30)$$

واحتمال ان متغير الاستجابة y ياخذ قيمة (0) يكون

$$P(y=0/x) = 1-p(y=1/x) \cdots\cdots\cdots (31)$$

وان الصيغ التقديرية للمعادلتين (31,30) هي

$$\hat{p}(y=1/x)=\frac{1}{1+e^{(\hat{\beta}0+\sum_{j=1}^{p}\hat{\beta}jxij)-1}}\cdots\cdots\cdots(32)$$

وان

$$\hat{p}(y=0/x)=1-\hat{p}(y=1/x)\cdots\cdots\cdots(33)$$

حيث ان $\widehat{\beta p},\cdots,\cdots,\widehat{\beta 1},\widehat{\beta 0}$ تمثل مقدرات الامكان الاعظم للمعالم المجهولة $\beta p,\cdots,\cdots,\beta 1,\beta 0$ وان قاعدة التصنيف وفق الصيغة الاحتمالية لنموذج الانحدار اللوجستي ثنائي الاستجابة تكون:

ان المشاهده $\underline{x}$ تعود للمجموعة G_0 اذا كان

$$\hat{p}(y=0/\underline{x})\geq\hat{p}(y=1/\underline{x})$$

وتعود المشاهده $\underline{x}$ للمجموعة G_1 اذا كان

$$\hat{p}(y=0/\underline{x})<\hat{p}(y\,1/\underline{x})$$

وتجد الاشارة الى ان نموذج الانحدار اللوجستي لايشترط توفر الافتراضات الاتية:

١- وجود علاقة خطية ما بين المتغيرات التوضيحية والمتغير المعتمد .

٢- ان يكون توزيع المتغيرات توزيعا طبيعيا .

٣- تحقق خصيصة ثبات التباين .

يستخدم الانحدار اللوجستي في التسويق لحساب توقعات ميل المستهلك الى شراء منتج ما او امتناعه عن الشراء.

تقدير وتفسير معاملات الانحدار اللوجستي :

من اجل تقدير معاملات الانحدار اللوجستي يتم اللجوء الى طريقة الاحتمال الاعظم والتي تعتبر الطريقة الاكثر ملائمة لكافة النماذج الخطية وغير الخطية وتعرف طريقة الاحتمال الاعظم بأنها طريقة تكرارية تعتمد على تكرار العمليات الحسابية عدة مرات حتى يتم الوصول الى افضل تقدير للمعاملات.

وتستخدم طريقة الاحتمال الاعظم لحساب معاملات اللوجت في الانحدار اللوجستي وتهدف هذه الطريقة الى تعظيم لوغاريتم الاحتمال الذي يعكس مدى امكانية او احتمال ان تكون تلك القيم المشاهده للمتغير التابع في الامكان توقعها او التنبوء بها.

طريقة الامكان(الترجيح)الاعظم Method of Maximum Likelihood

دالة التراجيح Likelihood Function

دالة الترجيح الى n من المتغيرات العشوائية $x1, x2, \cdots xn$ تعرف بانها تساوي دالة الكثافة المشتركة لهذه المتغيرات وهي دالة الى المعلمة θ

بصوره اخرى اذا كانت $x1, x2, \cdots xn$ عينة عشوائية من دالة الكثافة الاحتمالية $f(x,\theta)$ فان دالة الترجيح $L(\theta: x1, \cdots xn)$ تعرف كالاتي

في الحالات الاعتيادية تكون دالة الترجيح محدوده من الاعلى ومستمره بالنسبة الى θ وان معادلة دالة الترجيح لها حل واحد فقط هو الذي يعمل على ان تكون $L(\theta: x1, x2 \cdots xn)$ اعظم مايمكن

بعد ان استعرضنا مفهوم دالة الترجيح فان طريقة الترجيح الاعظم يمكن تلخيصها بالخطوات التالية

١- نجد دالة الترجيح في التعريف اعلاه

٢- نأخذ اللوغاريتم الطبيعي او الاعتيادي (حسب ظروف المسألة) الى داله الترجيح

٣- نجد المشتقة بالنسبة الى θ ثم نضع الناتج مساويا للصفر

٤- نحل المعادلة الناتجة ونجد مخمن الترجيح الاعظم ويرمز لهذا المخمن بالرمز $\hat{\theta}$

مثال/ لتكن $x1, x2, \cdots xn$ عينة عشوائية سعتها n من التوزيع الاسي بمعلمة λ جد المخمن الى λ بطريقة الترجيح الاعظم

الحل//

$$f(x,\lambda) = \lambda\, e^{-\lambda x} \quad ، x > 0$$

$$\mathrm{L}(\lambda: x1, \cdots xn) = \prod_{i=1}^{n} f(xi, \lambda)$$

$$= \lambda^{n} \; e^{-\lambda \sum xi}$$

$$\mathrm{Ln\ L}(\lambda: x1, \cdots xn) = n\; ln\lambda - \lambda \sum_{i=1}^{n} xi$$

$$\frac{\mathrm{d} \ln l}{\mathrm{d}\lambda} = \frac{n}{\lambda} - \sum xi = 0$$

$$\frac{n}{\lambda} = \sum xi$$

$$\hat{\lambda} = \frac{n}{\sum xi} = \frac{1}{\frac{\sum xi}{n}}$$

$$= \frac{1}{\bar{\mathrm{x}}}$$

مثال/ لتكن $x1, x2 \cdots xn$ عينه عشوائية من توزيع بواسون جد المخمن الى λ بطريقة الترجيح الاعظم

الحل//

$$\mathrm{p}\left(x: \lambda\right) = \frac{\lambda^{x} e^{-x}}{x!}, \quad \mathrm{x}=0,1,2\cdots n$$

$$\mathrm{L}(\lambda: x1, \cdots xn) = \prod_{i=1}^{n} p(xi, \lambda)$$

$$= \frac{\lambda \sum xi \; e^{-n\lambda}}{\prod_{i=1}^{n} xi!}$$

$$ln\, l(\lambda: x1, x1, \cdots xn) = (\sum xi) \ln \lambda - n\lambda - \sum_{i=1}^{n} \ln(xi)!$$

$$\frac{\mathrm{d} \ln l}{\mathrm{d}\lambda} = \frac{\sum xi}{\lambda} - n - 0 = 0$$

$$\lambda = \frac{\sum xi}{n} = \bar{x}$$

المصادر

المصادر العربية

١- باطين ، عادل احمد حسن (٢٠٠٩) ، انحدار اللوجستي وكيفية استخدامه في بناء نماذج التنبؤ للبيانات ذات المتغيرات المتابعة ثنائية القيمة ، اطروحة دكتوراه غير منشورة ، اختصاص احصاء وبحوث ، جامعه ام القرى – كلية التربية – قسم علم النفس ، السعودية

٢- بطارسة ، صالح رشيد ، ٢٠٠٩ ، كتاب الاحصاء والاحتمالات ، دار اسامة للنشر والتوزيع ، الاردن – عمان

٣- عطية ، عبد القادر محمد عبد القادر (٢٠٠٤) ، الحديث في الاختصاص القياسي بين النظرية والتطبيق ، مكة المكرمة ، السعودية

المصادر الانكليزية

4- Binay Logistic Regression with PASW/SPSS, available

5- Garson , David (2006) , logistic Regression available

6- Draper,N,R.and smith , H,(1981); Applied Regression analysis , new york , P.413