استعمال طريقة الإمكان الأعظم المعدلة لتقدير انموذج الانحدار المتعدد الحدود

الباحث مهدي يونس محمد أ.م.د. احمد ذياب احمد

جامعة بغداد / كلية الادارة والاقتصاد / قسم الاحصاء

الملخص

هذا البحث تم الحصول على تقديرات انموذج ومعلمات الانحدار المتعدد الحدود من الدرجة الثانية في حالة ان الخطأ العشوائي يتوزع Long Tailed Symmetric وذلك باستعمال طريقة الإمكان الأعظم المعدلة, ولقد اثبتت هذه الطريقة كفاءة اكثر من طريقة المربعات الصغرى الاعتيادية من خلال المقارنة بينها باستعمال متوسط مربعات الخطأ وباسلوب المحاكاة ولثلاثة احجام للعينات (٦٠, ٩٠, ١٢٠).

ABSTRACT

In this research, estimates of the model and parameters of second-degree polynomial regression were obtained in the event that the random error is distributed Long Tailed Symmetric, using the modified maximum Likelihood method, This method has proven more efficient than the ordinary least squares method by comparing it using mean squares error, simulation method, and three sample sizes (60,90,120)

مفاتيح الكلمات: الانحدار المتعدد الحدود, توزيع طويل الذيل المتماثل, طريقة الامكان الاعظم المعدلة.

١- المقدمة

يستعمل الانحدار[1] لتفسير العلاقة بين متغيرين او اكثر يسمى احد هذه المتغيرات بالمتغير التابع والبقية تسمى بالمتغيرات التوضيحية, يعمل الانحدار على نمذجه البيانات لكي نتمكن من تفسير العلاقة بين المتغيرات. ويقسم الانحدار الى الانحدار الخطي و الانحدار اللاخطي ويستعمل الانحدار الخطي عندما يكون اس المتغيرات واحد اي من الدرجة الاولى, وعندما يكون في معادلة الانحدار متغير توضيحي واحد يسمى بانموذج الانحدار الخطي البسيط ويسمى بالانحدار الخطي المتعدد اذا كان هناك اكثر من متغير توضيحي واحد. اما الانحدار اللاخطي فيستعمل عندما تكون المتغيرات ذات اس اكبر من واحد او متعدد الحدود او ذات صيغ لوغاريتمية.

٢- الانحدار المتعدد الحدود

يتم في الانحدار المتعدد الحدود[3]] نمذجة العلاقة بين المتغيرات التوضيحية (Xs) والمتغير المعتمد (Y) بشرط ان تكون X من الدرجة n حيث يعد الانحدار المتعدد الحدود نوع من أنواع الانحدار اللاخطي وله مجالات تطبيق مختلفة مثل المجالات الاقتصادية والمجالات الطبية ومجالات أخرى. وصيغة انموذج الانحدار المتعدد الحدود من الدرجة (k) هي[7] [13]:

$$Y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x^2_{i1} + \ldots + \beta_k x^k_{i1} + e_i \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots \qquad (1)$$

ويدعى بأنموذج الانحدار البسيط اذا كان $k=1$ وبأنموذج الانحدار المتعدد الحدود من الدرجة الثانية اذا كان $k=2$ ومن الدرجة الثالثة اذا كان $k=3$.

ان مقدرات طريقة المربعات الصغرى (OLS) ليست فعالة في تقدير معلمات أنموذج الانحدار المتعدد الحدود بسبب وجد القيم الشاذة [5] [14] او المتطرفة التي يمكن ان يكون مصدرها خطأ في قراءة البيانات او في تسجيلها او قد يكون مصدر القيم الشاذة هو المجتمع الاحصائي وهذا عندما تكون بعض قيم البيانات تختلف بشكل كبير عن البقية وان الاختلاف غير

ناتج عن خطأ وانما حالة حقيقية موجودة في الواقع, لذا قام الباحثان (Akkaya and Tiku) في عام ٢٠٠٨ باعادة نمذجة انموذج الانحدار الخطي المتعدد وكما يلي[1]

$$y_i = \theta_0 + \sum_{j=1}^{q} \theta_j\, u_{ij} + e_i \qquad 1 \le j \le q \quad ,1 \le i \le n \qquad \dots\dots\dots\dots \quad (2)$$

$$u_{ij} = \frac{x_{ij} - \bar{x}_j}{s_i} \quad , \quad \bar{x}_j = \frac{\sum_{i=1}^{n} x_{ij}}{n} \quad , \quad s_j^2 = \sum_{i=1}^{n} \frac{1}{n}(x_{ij} - \bar{x}_j)^2$$

كما وتوصل الباحث (waebe)[16] في عام 2008 الى ان غالبا ما تتوزع البيانات التي تمثل الخسائر والعوائد المالية توزيعا ملتويا, وفي عام (2019) وضح الباحث (kitic)[9] ان العديد من التطبيقات البيولوجية لا يمكن تحليلها باستعمال الإحصائية الخطية, وتوصل الباحثان (Puthenpura and sinha)[15] في عام 1968 الى ان طريقة المربعات الصغرى تكون غير فعالة بوجود الشواذ لذا اقترحا استعمال طريقة الإمكان الأعظم المعدلة, وفي عام 2008 قام الباحثان (AKKaya and Tiku)[1] بتقدير معلمات انموذج الانحدار المتعدد باستعمال طريقة الامكان الاعظم المعدلة و طريقة المربعات الصغرى والمقارنة بين كل من المتوسط والتباين وتبين ان طريقة الامكان الأعظم المعدلة افضل بكثير لانها كانت اقل تحيز, ان انموذج الانحدار المتعدد الحدود من الدرجة الثانية (Second order polynomial regression model) يكون بالصيغة الالتية:

$$y_i = \theta_0 + \sum_{j=1}^{q} \theta_j\, u_{ij} + \sum_{j=1}^{q} \theta_{jj}\, u_{ij}^2 + \sum_{j=1}^{q-1} \sum_{k=j+1}^{q} \theta_{jk}\, u_{ij}\, u_{ik} + e_i \qquad \dots\dots\dots\dots\dots (3)$$

$$1 \le j \le q \quad , \quad 1 \le i \le n$$

q : عدد المتغيرات المستقلة, ويمكن كتابة النموذج (٣) بصيغة المصفوفات وكالاتي:

$$Y = U\theta + e \qquad \dots\dots\dots\dots (4)$$

اذ ان :

$$Y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ \vdots \\ y_n \end{bmatrix} \quad , \quad e = \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ \vdots \\ e_n \end{bmatrix} \quad , \quad \theta = \begin{bmatrix} \theta_0 \\ \vdots \\ \theta_q \\ \theta_{11} \\ \vdots \\ \theta_{qq} \\ \theta_{12} \\ \vdots \\ \theta_{q-1,q} \end{bmatrix}$$

$$U = \begin{pmatrix} 1 & u_{11} & \dots & u_{1q} & u_{11}^2 & \dots & u_{1q}^2 & u_{11}u_{12} & \dots & u_{1q-1}u_{1q} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 1 & u_{n1} & \dots & u_{nq} & u_{n1}^2 & \dots & u_{nq}^2 & u_{n1}u_{n2} & \dots & u_{n,q-1}u_{nq} \end{pmatrix}$$

3- طرائق التقدير

لغرض تقدير معلمات انموذج الانحدار المتعدد الحدود من الدرجة الثانية سيتم استعمال ثلاثة طرائق هي طريقة المربعات الصغرى الاعتيادية (Ordinary Least Square Method) وطريقة الإمكان الأعظم المعدلة (Modified Maximum Likelihood Method) وطريقة (M) الحصينة, وان صيغة طريقة المربعات الصغرى الاعتيادية هي[2]:

$$\tilde{\theta} = (\grave{U}U)^{-1}(\grave{U}Y) \qquad \dots\dots\dots\dots\dots (4)$$

$$\sigma^2 = \frac{Ee^2}{n-c} = \frac{(Y-U\theta)'(Y-U\theta)}{n-c} \quad , \quad c = 1 + 2q + \frac{q(q-1)}{2}$$

σ^2: التباين
c: درجة الحرية

3-1- طريقة الإمكان الأعظم المعدلة (Modified maximum Likelihood method)

ان صيغة اشتقاق طريقة الامكان الاعظم المعدلة يكون كالاتي[12] وبافتراض ان الخطأ يتوزع طويل الذيل المتماثل (Log Tailed Sympatric distribution) [11]: درجة الحرية

$$f(e)=\frac{\Gamma(p)}{\sigma\sqrt{k}\,\Gamma\left(\frac{1}{2}\right)\Gamma\left(\frac{p-1}{2}\right)}\left\{1+\frac{e^2}{k\sigma^2}\right\}^{-p}.-\infty<e<+\infty \quad \text{.................} (5)$$

$$E(e)=0.\ \ V(e)=\sigma^2\ ,\quad k=2p-3\quad ,\quad t=\sqrt{\frac{v}{k}}\frac{e}{\sigma}$$

σ: تمثل معلمة القياس (Scale Parameter) و p: تمثل معلمة الشكل (Shape Parameter), وان دالة الامكن الاعظم هي:

$$L=\prod_{i=1}^{n}f(e)\qquad ,\qquad z_i=\frac{e_i}{\sigma}$$

$$\ln L=n\ln d-n\ln\sigma-p\sum_{i=1}^{n}\ln\left(1+\frac{{z_i}^2}{k}\right) \quad \dots\dots\dots\dots \quad (6)$$

$$e_i=y_i-\left(\theta_0+\sum_{j=1}^{q}\theta_j\,u_{ij}+\sum_{j=1}^{q}\theta_{jj}\,u_{ij}^2+\sum_{j=1}^{q-1}\sum_{k=j+1}^{q}\theta_{jk}\,u_{ij}\,u_{ik}\right)$$

$$\frac{\partial L}{\partial\theta_0}=\frac{2p}{k\sigma}\sum_{i=1}^{n}\frac{z_i}{1+\frac{z_i^2}{k}}=0 \quad \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots (7)$$

$$\frac{\partial L}{\partial\theta_j}=\frac{2p}{k\sigma}\sum_{i=1}^{n}\sum_{j=1}^{q}\frac{u_{ij}z_i}{1+\frac{z_i^2}{k}}=0 \quad \dots\dots\dots\dots\dots (8)$$

$$\frac{\partial L}{\partial\theta_{jj}}=\frac{2p}{k\sigma}\sum_{i=1}^{n}\sum_{j=1}^{q}\frac{{u_{ij}}^2z_i}{1+\frac{z_i^2}{k}}=0 \quad \dots\dots\dots\dots\dots\dots (9)$$

$$\frac{\partial L}{\partial\theta_{jk}}=\frac{2p}{k\sigma}\sum_{i=1}^{n}\sum_{j=1}^{q-1}\sum_{k=j+1}^{q}\frac{u_{ij}u_{ik}z_i}{1+\frac{z_i^2}{k}}=0 \quad \dots\dots\dots\dots\dots (10)$$

$$\frac{\partial L}{\partial\sigma}=\frac{2p}{k\sigma}\sum_{i=1}^{n}\frac{{z_i}^2}{1+\frac{z_i^2}{k}}=0 \quad \dots\dots\dots\dots\dots (11)$$

$$g(z_i)=\frac{z_i}{1+\frac{{z_i}^2}{k}} \quad \text{..........................} (12)$$

وان المعادلات أعلاه تحتوي على دوال صعبة وبتطبيق خطوات الإمكان الأعظم المعدلة بوضع العادلات بمصطلح الترتيب الـ (Variables Order) بترتيب المتغيرات بشكل تصاعدي واستبدال (z_i) في العادلات اعلاه بـ $\left(z_{(i)}\right)$ وكما يلي[12]

$$z_{(1)}\le z_2\le z_{(3)}\le\cdots\le z_{(n)}$$

$$\frac{\partial L}{\partial\theta_0}=\frac{2p}{k\sigma}\sum_{i=1}^{n}g\left(z_{(i)}\right)=0 \quad \dots\dots\dots\dots\dots \quad (13)$$

$$\frac{\partial L}{\partial \theta_j}=\frac{2p}{k\sigma}\sum_{i=1}^{n}\sum_{j=1}^{q}u_{ij}g\left(z_{(i)}\right)\ =0 \qquad \dots\dots\dots\dots\dots\ (14)$$

$$\frac{\partial L}{\partial \theta_{jj}}=\frac{2p}{k\sigma}\sum_{i=1}^{n}\sum_{j=1}^{q}{u_{ij}}^2g\left(z_{(i)}\right)=0 \qquad \dots\dots\dots\dots\dots\ (15)$$

$$\frac{\partial L}{\partial \theta_{jk}}=-\frac{2p}{k\sigma}\sum_{i=1}^{n}\sum_{j=1}^{q-1}\sum_{k=j+1}^{q}u_{ij}u_{ik}g\left(z_{(i)}\right)\ =0 \qquad \dots\dots\dots\dots\dots\ (16)$$

$$\frac{\partial L}{\partial \sigma}=-\frac{n}{\sigma}+\frac{2p}{k\sigma}\sum_{i=1}^{n}z_ig\left(z_{(i)}\right)\ =0 \qquad \dots\dots\dots\dots\dots\ (17)$$

من ثم يتم ابدال $g(z_i)$ بالدالة الخطية الاتية

$$g(z_i)=\alpha_i+\beta_iz_{(i)} \qquad \dots\dots\dots\dots\dots\ (18)$$

وللحصول على تقديرات β,α_i يكون باستعمال اول حدين من سلسلة تايلر لل $g(z_i)$ حول $t_{(i)}$

$$g\left(z_{(i)}\right)=g\left(t_{(i)}\right)+\left(z_i-t_{(i)}\right)g'\left(t_{(i)}\right) \qquad \dots\dots\dots\dots\dots\ (19)$$

$$\alpha_i=\frac{2\frac{{t_{(i)}}^3}{k}}{\left(1+\frac{{t_{(i)}}^2}{k}\right)^2}\quad . \qquad \beta_i=\frac{1-\frac{{t_{(i)}}^2}{k}}{\left(1+\frac{{t_{(i)}}^2}{k}\right)^2}$$

وعلى فرض ان (q_i) تمثل تقدير للدالة التراكمية $F(t_i)$ وياخذ الصيغة الاتية

$$q_i=\frac{i}{n+1}$$

وبما ان (t_i) تمثل معكوس الدالة التراكمية فلايجاد تقديرها يكون كل الاتي:

$$F(t_i)=\frac{\Gamma(p)}{\sigma\sqrt{k}\,\Gamma\left(\frac{1}{2}\right)\Gamma\left(\frac{p-1}{2}\right)}\int_{-\infty}^{t_{(i)}}\left(1+\frac{e^2}{k\sigma^2}\right)^{-p}dz \quad \dots\dots\ (20)$$

$$q_i=\frac{\Gamma(p)}{\sigma\sqrt{k}\,\Gamma\left(\frac{1}{2}\right)\Gamma\left(\frac{p-1}{2}\right)}\int_{-\infty}^{t_{(i)}}\left(1+\frac{e^2}{k\sigma^2}\right)^{-p}dz \quad \dots\dots\ (21)$$

بأجراء بعض التكاملات نحصل عل تقدير قيمة (t), وبتعويض $\alpha_i+\beta_iz_{(n)}$ في المعادلات ١٣, ١٤, ١٥, ١٦, ١٧ و كالاتي[12]:

$$\frac{\partial L}{\partial \theta_0}=\frac{2p}{k\sigma}\sum_{i=1}^{n}\left(\alpha_i+\beta_iz_{((i))}\right)\ =0 \qquad \dots\dots\ (22)$$

$$\frac{\partial L}{\partial \theta_j}=\frac{2p}{k\sigma}\sum_{i=1}^{n}\sum_{j=1}^{q}u_{ij}\ \left(\alpha_i+\beta_iz_{(i)}\right)=0 \qquad \dots\dots\ (23)$$

$$\frac{\partial L}{\partial \theta_{jj}} = \frac{2p}{k\sigma}\sum_{i=1}^{n} {u_{ij}}^{2}\left(\alpha_i + \beta_i z_{(i)}\right) = 0 \qquad \ldots\ldots (24)$$

$$\frac{\partial L}{\partial \theta_{jk}} = \frac{2p}{k\sigma}\sum_{i=1}^{n}\sum_{j=1}^{q-1}\sum_{k=j+1}^{q} u_{ij}u_{ik}\left(\alpha_i + \beta_i z_{(i)}\right) \qquad \ldots\ldots (25)$$

$$\frac{\partial L}{\partial \sigma} = -\frac{n}{\sigma} + \frac{2p}{k\sigma}\sum_{i=1}^{n} z_i\left(\alpha_i + \beta_i z_{(i)}\right) \qquad \ldots\ldots (26)$$

وبحل المعادلات اعلاه نحصل على مقدرات الإمكان الأعظم المعدلة وكالاتي:

$$\theta = K + D\tilde{\sigma} \qquad \ldots\ldots (27)$$
$$K = (W'\mathfrak{B}W)^{-1}(W'\mathfrak{B}Y) = (K_\ell) \qquad \ldots\ldots (28)$$
$$D = (W'\mathfrak{B}W)^{-1}(W'\alpha I) \qquad \ldots\ldots (29)$$

$$\tilde{\sigma} = \frac{B + \sqrt{B^2 + 4nc}}{2\sqrt{n(n-c)}} \qquad \ldots\ldots\ldots\ldots\ldots (30)$$

$$\alpha = \mathrm{diag}(\alpha_i) \quad , \quad \mathfrak{B} = \mathrm{diag}(\beta_i \quad , \quad I' = [1.1.\ldots\ldots.1]$$

٣- المحاكاة

ان القيم الافتراضية التي تم استعمالها في المحاكاة لمعلمتي الشكل والقياس هي(P=3,5,7) و $(\sigma^2 = 1)$ اما القيم الافتراضية لمعلمات انموذج الانحدار فتكون كالاتي:

cof	θ_0	θ_1	θ_2	θ_3	θ_{11}	θ_{22}	θ_{33}	θ_{12}	θ_{13}	θ_{23}
value	77.21	-8.79	-7.43	-0.05	-3.06	**-3.52**	**-1.73**	**-4.68**	**-2.08**	**-1.17**

وصيغة توليد المتغير العشوائي هي:

$$e = \frac{t\sigma}{\sqrt{\frac{v}{k}}}$$

و(t) البيانات العشوائية التي يتم توليدها وفق برنامج (MATLAB) $t = trad(2p-1)$ وتم اختيار ثلاثة احجام للعينات (60,90,120) وتكرار التجربة (1000) , وللمقارنة بين طرائق التقدير تم استعمال المقياس الاحصائي متوسط مربعات الأخطاء (MSE) وكالاتي:

١- لمعلمات الانموذج وفق الصيغة الأتية:

$$MSR(\hat{\theta}) = \frac{\sum_{i=1}^{r}(E(\hat{\theta}) - \theta)^2}{R}$$

٢- للأنموذج وفق الصيغة الأتية:

$$MSR(\hat{\theta}) = \frac{\sum_{i=1}^{n}(y_i - \hat{y})^2}{n-c}$$

R: عدد مرات تكرار التجربة, وتم كتابة البرنامج بلغة (MATLA) وفيما يلي نتائج المحاكاة:

جدول (1) (MSE) لمعلمات الانموذج لجميع طرائق التقدير وعندما (P=3)

parameters	**60**		**90**		**120**	
	ols	**MMLE**	**ols**	**MMLE**	**ols**	**MMLE**
θ_0	0.085391	322.5215	0.057637	295.2409	0.040806	284.3584
θ_1	0.018077	149.1451	0.012634	128.7969	0.009336	123.0894

θ_2	0.019532	59.17818	0.011726	49.9603	0.009996	47.09173
θ_3	0.019172	31.17377	0.012874	27.35147	0.009287	26.12682
θ_{11}	0.025042	8.438129	0.015988	5.337014	0.011281	4.849357
θ_{22}	0.026141	9.723491	0.015613	7.85317	0.011065	7.3537
θ_{33}	0.025014	9.277775	0.015902	6.999886	0.011597	6.511795
θ_{12}	0.021078	17.96225	0.012657	13.48186	0.008233	12.29989
θ_{13}	0.022571	32.6825	0.013362	27.42217	0.008764	26.11237
θ_{23}	0.020516	12.88293	0.013378	10.11803	0.009412	9.588027

جدول (2) (MSE) لمعلمات الانموذج لجميع طرائق التقدير وعندما (P=5)

parameters	60		90		120	
	ols	MMLE	ols	MMLE	ols	MMLE
θ_0	0.086411	0.0000043	0.057641	0.0000042	0.041381	0.000004
θ_1	0.020628	0.0000059	0.011939	0.0000056	0.009276	0.0000055
θ_2	0.020502	0.0000034	0.011668	0.0000032	0.008546	0.0000031
θ_3	0.018804	0.00000023	0.01308	0.00000019	0.008984	0.00000018
θ_{11}	0.023547	0.00000025	0.015179	0.00000021	0.011124	0.00000019
θ_{22}	0.026713	0.0000011	0.015229	0.0000011	0.011657	0.000001
θ_{33}	0.026161	0.00000039	0.015638	0.00000034	0.011665	0.00000031
θ_{12}	0.020664	0.0000006	0.013248	0.00000057	0.008805	0.00000054
θ_{13}	0.023107	0.00000082	0.011968	0.00000072	0.009459	0.0000007
θ_{23}	0.020105	0.00000027	0.01385	0.00000023	0.009347	0.00000022

جدول (3) (MSE) لمعلمات الانموذج لجميع طرائق التقدير وعندما (P=7)

parameters	60		90		120	
	ols	MMLE	ols	MMLE	ols	MMLE
θ_0	0.093566	0.000017	0.057695	0.000016	0.043648	0.000015
θ_1	0.019611	0.00000024	0.011803	0.00000022	0.00929	0.00000021
θ_2	0.019475	0.00000014	0.01223	0.00000013	0.009222	0.00000012
θ_3	0.020782	0.000000005	0.012491	0.00000000043	0.00855	0.000000004
θ_{11}	0.025861	0.00000001	0.016656	0.00000000095	0.011016	0.0000000085
θ_{22}	0.026257	0.000000024	0.015856	0.0000000023	0.011888	0.000000022
θ_{33}	0.026164	0.000000015	0.014241	0.0000000014	0.011402	0.000000013
θ_{12}	0.021792	0.000000013	0.012719	0.0000000011	0.009505	0.00000001
θ_{13}	0.020102	0.000000016	0.01245	0.0000000014	0.00952	0.000000014

			8		5	
θ_{23}	0.022106	0.000000005 1	0.01235 4	0.0000000004 1	0.00967 1	0.000000003 7

جدول (4) (MSE) للأنموذج ولجميع طرائق تقدير

	n	ols	MMLE	الافضل
P=3	60	1.202293	609.4715	ols
	90	1.130553	510.5617	ols
	120	1.084538	476.5779	ols
P=5	60	1.195609	0.994181	MMLE
	90	1.128175	1.003263	MMLE
	120	1.092455	0.996361	MMLE
P=7	60	1.20235	0.999345	MMLE
	90	1.123508	1.002293	MMLE
	120	1.083188	0.992137	MMLE

تحليل النتائج

- **متوسط مربعات الخطأ (MSE) لمعلمات الانموذج**

١. عند P=3 ولجميع حجوم العينة تكون طريقة المربعات الصغرى افضل طريقة لانها تعطي اقل MSE
٢. عند P=5 ولكل حجوم العينة تعطي طريقة MMLE اقل MSE
٣. عند P=7 ولكل حجوم العينة تعطي طريقة MMLE اقل MSE

- **متوسط مربعات الخطأ (MSE) للانموذج**

١. عند P=3 ولكل حجوم العينة تعطي طريقة ols اقل متوسط لمربعات الخطأ MSE
٢. عند P=7,5 ولكل حجوم العينة تعطي طريقة MMLE اقل متوسط لمربعات الخطأ MSE

6 الجانب التطبيقي

تم جمع البيانات الخاصة بمستوى السكر في الجسم من خلال احدى المختبرات المرخصة من وزارة الصحة العراقية (مختبر كلكامش) لسنة 2022 ولعينة حجم مشاهداتها (90) فرد وتم حساب المتغيرات التالية :

c-peptide: يقيس هذا الاختبار مستوى c-peptide في الدم و هو مادة مصنعة في البنكرياس مع الانسولين يعمل على التحكم في مستوى الجلوكوز(السكر في الدم)ويعد الجلوكوز المصدر الرئيسي للطاقة في الجسم .يتم انتاج الانسولين و c-peptid c من البنكرياس في نفس الوقت وبكميات متساوية تقريباً.لذلك من الممكن ان يكون هذا الاختبار طريقة جيدة لقياس الانسولين لانه يميل الى البقاء في الجسم فترة اطول واذا لن ينتج الجسم كمية كافية من الانسولين سوف يكون هذا علامة على الاصابة بمرض السكر

HBa1c: أو ما يُعرف بفحص الهيموغلوبين السكري ((Hemoglobin A1c Test، أو فحص السكر التراكمي ويقيس هذا الفحص كمية السكر في الدم التي تكون مرتبطة بالهيموغلوبين .وتأتي اهمية هذا الفحص كونه يوضح متوسط كمية الغلوكوز في الدم التي تكون مرتبطة بالهيموغلوبين خلال الاشهر الثلاث الماضية ويعود ذللك كون حياة خلايا الدم الحمر في مجرى الدم ثلاث اشهر

نتائج فحص الهيموغلوبين HBa1c

تفسير النتائج

أقل من ٤٢ ملليمول / مول

(٥,٦٪)

غير مصاب بالسكري

بين ٤٢ و٤٧ ملليمول / مول

(٥,٧% - ٦,٤٪)

مقدمات السكري Prediabetes))
٤٨ ملليمول / مول
(٦,٥٪) أو أكثر
مصاب بداء السكري من النوع ٢
(random blood sugar) RBS
:هو اختبار لسكر الدم بشكل عشوائي في اي وقت من اليوم ويقوم هذا لتحليل بقياس نسبة الجلوكوز في الدموذلك بغض النظر عن اخر مرة تناول الطعام ويمكن ايطاً اخذ اكثر من قياس على مدار اليوم وبالنسبة للا شخاص الطبيعين فأن مستويات السكر العشوائي لا يتغير على مدار اليوم اما وجود مستويات مختلفة بدرجات كبيرة على مدار اليوم هذا يعني وجود مشكلة
UREA:تعتبراليويا فضلات طبيعية ينتجها جسم الإنسان بعد تناول الطعام، حيث يقوم الكبد بتحليل البروتين الموجود في الطعام منتجًا بذلك مادة اليوريا، وتكون نسبة اليوريا تختلف من فرد الى اخر بختلاق العمر او الجنس وعوامل أخرى وتكون النسبة الطبيعية لليوريا في الدم كما يلي
الرجال (24 – 8ملليغرام/ ديسيلتر)
النساء (21 – 6ملليغرام/ ديسيلتر)
الأطفال حتى عمر ١٧ عام (20 – 7ملليغرام/ ديسيلتر)
c-peptide: y
HBa1c:X1
RBS :X2
UREA :X3
تم استعمل اختبار كولموكوروف- سميرنوف(Kolmogorov-smirnov Tast) وتبين ان البينات لا تتبع توزيع طويل الذيل المتماثل لذلك تم معالجة البيانات بأخذ الدرجة المعيارية للمتغير المعتمدY وتم اختبارها مرة ثانية وتبين انها تتبع توزيع طويل الذيل المتماثل وكانت القيمة المحسوبة للاختبار D=0.1411 اقل من القيم الجدولية مستوى معنوية $K_{0.05}$ = 0.1434 وعند$K_{0.01} = 0.1718$
واثبت ان القيمة الجدولية اكبر من المحسوبة وهذا يعني قبول فرضية العدم H_0 أي ان البينات تتبع توزيع طويل الذيل المتماثل

الطرائق المستعملة في تحليل النتائج

تم استعمال طريقة الإمكان الأعظم المعدلة (MMLE) كونها تعطي اقلMSE للمعلمات وللأنموذج عن الطراق الأخرى ولجميع حجوم العينة ,وتم الحصول على متوسط مربعات الأخطاء عن طريق المحاكاة, كما ان هذه الطريقة تحتاج الى قيم أولية لذلك تم الحصول على هذه القيم باستعمال طريقة المربعات الصغرى

جدول(٤-٢) القيمة المقدرة لمعلمات طريقة الامكان الاعظم المعدلة عندما (p=7)

parameters.	θ_0	θ_1	θ_2	θ_3	θ_{11}
Value	-0.02223	0.003171	-0.24895	0.277254	0.013605
Parameters	θ_{22}	θ_{33}	θ_{12}	θ_{13}	θ_{23}
Value	0.05333	0.07001	-0.03021	-0.45654	0.335203

$$\hat{y}=-0.02223+0.003171u_{i1}-0.24895u_{i2}+0.277254u_{i3}+0.013605u_{i1}^2+0.05333u_{i2}^2+0.07001u_{i3}^2-0.03021u_{i1}u_{i2}-0.45654u_{i1}u_{i3}+0.335203u_{i2}u_{i3}$$

الاستنتاجات
1- ان العلاقة بين المتغير المستقل RBC مع المتغير المعتمد C-PeP تكون طردية .اي ان نسبة زيادة RBC يودي الى زيادة C-PeP الاستنتاجات
2- ان العلاقة بين المتغير المستقل HBAIC مع المتغير المعتمد C-PeP تكون عكسية أي ان أي زيادة في HBAIC يؤدي الى نقصان في C-PeP
3- ان العلاقة بين المتغير المستقل UREA مع المتغير المعتمد C-PeP تكون طردية .اي ان نسبة زيادة UREA يودي الى زيادة C-PeP

٤- تبين ان طريقة المربعات الصغرى العتيادية اقل كفاءة في تقدير انموذج الانحدار المتعدد الحدود من الدرجة الثانية
٤- ان قيم متوسط مربعات الخطأ (MSR) للمعلمات وللنموذج يتناقص بزيادة حجم العينة
٥- اثبتت طريقة الإمكان الأعظم المعدلة كفاءتها قي تقدير انموذج اللانحدار المتعدد الحدود من الدرجة الثانية مقارنتا بطريقة المربعات الصغرى الاعتيادية

التوصيات

- اوصي باستعمال طريقة الإمكان الأعظم المعدلة لتقدير معلمات انموذج الانحدار المتعدد الحدود عندما يتوزع الخطاء طويل الذيل المتماثل
- ضرورة استعمال انموذج الانحدار المتعدد العوامل المتعدد الحدود من الدرجة الثانية في دراسة العوامل المؤثرة على c-peptide
- اعتماد على c-peptide فيما اذا كان الفرد مصاب بالسكر ام لا
- استعمال متغيرات أخرى مختلفة عن الذي تم استعمالها
- استعمال طرائق تقدير أخرى لتقدير انموذج الانحدار المتعدد العوامل المتعدد الحدود من الدرجة الثانية مثل طريقة SوM الحصينتين

المصادر

References

1. Akkaya, A.D., and Tiku, M.L. (2008) " Robust estimation in multiple linear regression model with non-Gaussian noise". Automatica 44, 407-417
2. Akkaya, A.D., and Tiku, M.L. (2010)." Estimation in multifactor polynomial regression under non- normality" . Pak. J. Statist. Vol. 26(1), 49-68
3. Al-Mashhadani, M. H., Hormuz, A. H. (1989). Statistics. College of Administration and Economics, University of Baghdad.
4. Daniel P. Normolle .(2003)."An Algorithm for Robust nonlinear Analysis of Radioimmunoassays and Othre Bioassays "Ann Arbor,MI 48109-2029 , 313-936-1013
5. David M. Rocke and David L.Woodruff, (1998) ," Some Statistical Tools for data Mining Applications", University of California. ,Davis.
6. Filzmoser, P.(2004) , " Amultivariate Outlier detection Method", Department of statistics and Probability Theory. Vienna, Austria, Volume 1,PP. 18-22
7. Islam,T., Shaibur,M.R. and Hossain,S.S.(2009)." Effectivity of Modified Maximum Likelihood Estimators Using Selected Ranked Set Sampling Data". AUSTRIAN JOURNAL OF Statistics Volume 38, Number 2, 109–120
8. Jajo, N. K. (1989). Robust estimator of the linear regression model. Master's thesis in Statistics, Second College of Education, Ibn Al-Haytham, University of Baghdad.
9. Kiliç , Muhammet (2020) , ―Using Genetic Algorithms For Parameter Estimation Of A Two-
10. M.L. Tiku, M.Q. Islam, and A.S. YILDIRIM,(2001), NONNORMAL REGRESSION. I. SKEW DISTRIBUTION, Commun. Stat. Theory Methods 30, no. 6 , pp. 1021–1045
11. M.L. Tiku, R.P. Suresh, (1992) , A new method of estimation for location and scale parameters, J. Stat. Plann. Inference 30 (2) 281–29
12. M.L. Tiku, S. Kumra, (1985) , Expected values and variances and covariances of order statistics for a family of symmetric distributions (Student's t), selected tables in mathematical statistics 141–270
13. Ostertagova, E.(2012)"Modelling using polynomial regression" Procedia Engineering 48 500 – 506

14. Qumsiyeh, S.B. (2007)."Non-normal bivariate distribution estimation and hypothesis testing" . A thesis submitted of the graduate school of natural and applied sciences of middle east technical 1-4
15. Ungwon Yu, Soyoung Yang, Jinhong Kim, Youngjae Lee, Kil-Taek Lim, Seiki Kim, Sung-Soo Ryu, Hyeondeok Jeong.(2001)" A Confidence Interval-Based Process Optimization Method Using Second-Order Polynomial Regression Analys". Processes
16. Waeber , Rolf , Embrechts , Paul , Roy , Parthanil And Lysenko , Natalia (2008) , ―Multivariate Skew-Normal Distributions And Their Extremal Properties ― , Master Thesis Swiss Federal Institute Of Technology, Eth Zurich