تقدير معاملات ومعلمة القياس لأنموذج انحدار القيمة المتطرفة الخطي (النوع الاول) وللقيم الكبرى مع جانب تطبيقي[1]

م . د. فاطمة جاسم محمد العزاوي
قسم الاقتصاد الزراعي / كلية الزراعة
جامعة بغداد

أ. م .د.عادل احمد هدو الربيعي
قسم الاحصاء / كلية الادارة والاقتصاد
الجامعة المستنصرية

المستخلص

تم تقدير معاملات أنموذج الانحدار الخطي في حالة الاخطاء التي تسلك على وفق توزيع القيمة المتطرفة (النوع الاول) وللقيم الكبرى كتطوير الى الدراسات حول القيم الصغرى، واتبعت طريقتان في التقدير وهي طريقة المربعات الصغرى التي استخدمت كتقدير اولي يعتمد عليه في الطريقة الثانية (طريقة الامكان الاعظم)، وتم اللجوء الى طريقة نيوتن رافسن واسلوب فشر وهما من طرائق التعويض المتتالي للحصول على تقديرات الامكان الاعظم وذلك لصعوبة حلها بالاسلوب الاعتيادي .ولمعرفة اي الطريقتين السابقتين اكفا في التقدير تم استخدام الكفاءة التقريبية.كما تم التطرق الى اجراءات الاستدلال الاحصائي لانموذج انحدار القيمة المتطرفة (النوع الاول) وللقيم الكبرى وذلك من خلال تقدير حدود الثقة لكل من (معلمة القياس ، معاملات الانحدار) وكذلك اختبار الفرضيات من خلال استخدام احد اختبارات حسن المطابقة الخاصة بهذا التوزيع وهو اختبار الرسم الاحتمالي تم التوصل الى انه ليس هناك اي دليل ضد التوزيع المفترض للاخطاء.الجانب التطبيقي للبحث تناول جانبان مهمان هما ازمنة البقاء (مرض ابيضاض الدم) والثاني درجات الحرارة العظمى

المصطلحات الأساسية: انحدار القيمة المتطرفة، معاملات الانحدار، معلمة القياس، المربعات الصغرى، الإمكان الأعظم، اختبار الرسم الاحتمالي، أزمنة البقاء، مرض ابيضاض الدم، درجات الحرارة العظمى.

ESTIMATION OF COEFFICIENTS AND SCALE PARAMETER FOR LINEAR (TYPE 1) EXTREME VALUE REGRESSION MODEL FOR LARGEST VALUES WITH APPLICATIONS

ABSTRACT

In this paper we estimate the coefficients and scale parameter in linear regression model depending on the residuals are of type 1 of extreme value distribution for the largest values . This can be regard as an improvement for the studies with the smallest values . We study two estimation methods (OLS & MLE) where we resort to Newton – Raphson (NR) and Fisher Scoring methods to get MLE estimate because the difficulty of using the usual approach with MLE . The relative efficiency criterion is considered beside to the statistical inference procedures for the extreme value regression model of type 1 for largest values . Confidence interval , hypothesis testing for both scale parameter and regression coefficients , goodness of fit statistics based on the observed residuals are considered . As a conclusion and through the probability plot test we get no evidence against using the assumed residuals distribution.

Keywords: Extreme value regression, The regression coefficient, Scale parameter, Least squares, Maximum likelihood, Probability plot test, Leukemia, Maximum grades temperature.

مجلة العلوم
الاقتصادية والإدارية
المجلد 18
العدد 68
الصفحات 373- 402

1- البحث مستل من رسالة ماجستير للباحث الأول .[1]

1- المقدمة

تقدير معاملات ومعلمة القياس لأنموذج انحدار القيمة المتطرفة الخطي (النوع الاول) وللقيم الكبرى مع جانب تطبيقي

تعد الدراسات المتعلقة بنظرية القيمة المتطرفة (xtreme value theory) من الدراسات المهمة في الإحصاء ولها تطبيقات عملية كثيرة وتعنى بدراسة الخصائص الإحصائية لأكبر او اصغر المشاهدات التي يشار اليها $X_{(n)}$ او $X_{(1)}$ على التوالي، وتعد حالة خاصة من النظرية العامة للاحصاءات المرتبة (General theory of order statistic)، وقد ركزت غالبية الدراسات السابقة على التعرف على خواص التوزيع (وسط، تباين) وبيان اهم تطبيقاته: كأزمنة البقاء Survival times , فيضان الانهر Flood flows ، الحالات الارصادية المتطرفة General meteorological data، مشاكل الانكسار Fracture problems ، وفي الثمانينيات تم تناول المسائل المتعلقة بوجود انحدار خطي مقترح في دالة الكثافة الاحتمالية لتوزيع القيمة المتطرفة عندما تسلك الاخطاء وفق توزيع القيمة المتطرفة (النوع الاول) وللقيم الصغرى .[12],[9].

2- هدف البحث

ركزت اغلب الدراسات في مجال تحليل الانحدار حول افتراض كون الخطأ العشوائي في الانموذج يتبع التوزيع الطبيعي في النماذج الخطية فعند تقدير معاملات أنموذج الانحدار فان مقدرات المربعات الصغرى ستكون متطابقة مع مقدرات الامكان الاعظم ولهذا تعد مقدرات كفوءة وجيدة ولكن ماذا سيحدث لخصائص مقدرات معاملات أنموذج الانحدار عندما تتوزع الاخطاء توزيعا غير التوزيع الطبيعي هذا السؤال وغيره دفع الباحثين في فترة الثمانيات الى تركيز دراساتهم و بحوثهم على ان الاخطاء في النماذج الخطية لها توزيعات غير التوزيع الطبيعي اذ اشارت تلك البحوث الى امكانية احتواء دالة الكثافة الاحتمالية لمتغير الاستجابة لتوزيع القيمة المتطرفة على أنموذج انحدار. يهدف هذا البحث الى تقدير معاملات أنموذج الانحدار الخطي في حالة الاخطاء التي تسلك وفق توزيع القيمة المتطرفة (النوع الاول) وللقيم الكبرى كتطوير الى الدراسات السابقة حول القيم الصغرى .

3- الجانب النظري

1.3 توزيعات القيمة المتطرفة Extreme value distributions

توزيعات القيمة المتطرفة تؤخذ لتدرس بصورة عامة للعوائل الثلاث الاتية :[12]

1- النوع الاول :

$$P_r[\,X \le x\,] = \exp[\,-\exp\,(\,-\,(\,x - \xi\,)\,/\,\theta\,)\,]\,,\quad -\infty < x < \infty \qquad (1)$$

2- النوع الثاني

$$P_r[\,X \le x\,] = \begin{cases} 0 & , \; x < \xi \\ \exp\,[(-(x-\xi)/\theta)^{-k}\,] & , \; x \ge \xi \end{cases} \qquad (2)$$

النوع الثالث :

$$P_r[\,X \le x\,] = \begin{cases} \exp\,[(-(\xi-x)/\theta)^{k}\,] & , \; x \le \xi \\ 1 & , \; x > \xi \end{cases} \qquad (3)$$

حيث ان (ξ , θ , k(>0)) وتمثل المعلمات، توزيع النوع الاول يشار اليه بتوزيع القيمة المتطرفة النوع الاول للقيم الكبرى (Type 1E.V. dist.for greatest values) وينشا كتوزيع محدد للمشاهدة الكبرى عندما تكون العينة مسحوبة من التوزيع الطبيعي (Normal dist.) و التوزيع اللوغارتمي الطبيعي(Logarithmic normal dist.) والتوزيع الاسي (Exponential dist.) ، وعند اخذ المتغير العشوائي $y = -x$ واستبدال ξ ب $-\xi$ سوف نستخرج توزيع القيمة المتطرفة النوع الاول للقيم الصغرى . (Least values)دالة الكثافة الاحتمالية لتوزيع القيمة المتطرفة النوع الاول للقيم الكبرى هي :

$$f_x(x) = \frac{1}{\theta} exp\left[\left(-\frac{x-\xi}{\theta}\right)\left(-\exp\left(-\frac{x-\xi}{\theta}\right)\right)\right] \quad , -\infty < x < \infty \qquad (4)$$

حيث ان ξ تمثل معلمة الموقع و θ معلمة القياس ($\theta > 0$).وعندما $\xi=0$ و $\theta=1$ فيدعى توزيع القيمة المتطرفة النوع الاول القياسي وله دالة الكثافة الاحتمالية التالية :

$$f_x(x) = \exp(-x - \exp(-x)) \quad , -\infty < x < \infty \qquad (5)$$

وبافتراض ان , $\xi=0$, $\theta=1$, $y = (x-\xi) / \theta$ ، نحصل على الصيغة القياسية الاتية :

$$f_y(y) = \exp(-y - \exp(-y)) \qquad (6)$$

وتمثل توزيع القيمة المتطرفة القياسي للقيم الكبرى ، فالدالة المولدة للعزوم للمتغير y هي :

$$\mu_y(t) = \Gamma(1-t) \quad , (t < 1) \qquad (7)$$

حيث $\Gamma(x)$ تمثل دالة كاما (Gamma function) التي تعطي وسط مساوي الى $\Gamma'_{(1)} = -\gamma$ حيث $\gamma = 0.57722$ ويمثل (Eulers constant) وتباين مساوي الى $\frac{1}{6}\pi^2\theta^2$ ومنه فان الدالة المولدة للعزوم للمتغير X ($X = Y\theta + \xi$) ستكون

$$\mu_x(t) = e^{t\xi}\,\Gamma(1-\theta) \quad , (|t| < \frac{1}{\theta}) \qquad (8)$$

حيث ان $\Gamma(\alpha)$ تمثل دالة كاما ومن خلالها فان :

$$E(x) = \xi + \gamma\theta = \xi + 0.57722\,\theta \qquad (9)$$

$$V(x) = \frac{1}{6}\pi^2\theta^2 \qquad (10)$$

علما بان التوزيع غير متماثل (Non symmetrical) مع معاملات التواء وتفلطح (Skewness and Kurtosis) مساوية الى 1.29857 و 5.4 على التوالي . علما بان توزيع القيمة المتطرفة النوع الاول يرتبط مباشرة بتوزيع وايبل باستخدام العلاقات : $\xi = -\log\alpha$, $\theta = \beta^{-1}$, $x=-\log T$ حيث ان T يسلك على وفق توزيع وايبل .

2.3- نموذج انحدار القيمة المتطرفة (An extreme value regression model)

نموذج الانحدار العام الاعتيادي :

$$Y_i = \underset{\sim}{x}'_i \ \underset{\sim}{\beta} + \varepsilon_i \ , \ i = 1,2,\ldots\ldots,n \qquad (11)$$

حيث ان :

$$\underset{\sim}{x}'_i = (1, x_{i1}, x_{i2}, , \quad , x_{ik}) \ , \ \underset{\sim}{\beta}' = (\beta_0, \beta_1, \ldots., \beta_k) \qquad (12)$$

اذ تشير القيم x_{i1} , x_{i2} , …, x_{ik} كي تمثل المشاهدات ل **k** من المتغيرات المستقلة (المتغيرات التفسيرية غير العشوائية) للمفردة **(i)** . اما $\underset{\sim}{\beta}'$ فتمثل موجه معاملات الانحدار المجهولة ، ε_i تمثل الاخطاء العشوائية على افتراض انها تسلك وفق توزيع القيمة المتطرفة (النوع الاول) وللقيم الكبرى مع **p.d.f.** :

$$E(\varepsilon_I) = 0 \ , \ Var(\varepsilon_I) = \frac{1}{6} \qquad f_{(\varepsilon_I)} = \frac{1}{\theta} \exp\left[-\left(\frac{\varepsilon_i}{\theta} + \gamma\right) - \exp\left(-\left(\frac{\varepsilon_i}{\theta} + \gamma\right)\right)\right] \qquad (13)$$

$\pi^2\theta^2$ حيث ان γ وضح سابقا وان و $\theta > 0$ (معلمة القياس العامة التي لها علاقة في حساب التباين للتوزيع السابق) . فان دالة الكثافة الاحتمالية ل (y_i) ستكون

$$f_{Y_i}(y_i) = \frac{1}{\theta} \exp\left[-\left(\frac{Y_i - x'_i \ \beta}{\theta} + \gamma\right) - \exp\left(-\left(\frac{Y_i - x'_i \ \beta}{\theta} + \gamma\right)\right)\right], \ -\infty < y < \infty \qquad (14)$$

حيث $\gamma = 0.57722$ و $\theta > 0$ اما الوسط والتباين للتوزيع :

$$E(y_i) = \underset{\sim}{x}'_i \ \underset{\sim}{\beta} \ , \ Var(y_i) = \frac{1}{6} \pi^2 \theta^2 \qquad (15)$$

التوزيع غير متماثل مع معامل التواء (1.29857) ومعامل تفلطح (2.4) .

3-3 تقدير معاملات ومعلمة القياس لنموذج انحدار القيمة المتطرفة الخطي (النوع الاول وللقيم الكبرى باستخدام طريقة الامكان الاعظم

اساس الطريقة يعتمد على ايجاد دالة الامكان للمتغيرات العشوائية ومن ثم ايجاد النقطة التي تجعلها اعظم ما يمكن (باخذ التفاضلات الجزئية) التي تمثل مقدر الامكان الاعظم وتتطلب طريقة (ML) تخصيص او تحديد التوزيع للاخطاء وذلك للحصول على دالة الامكان الاعظم ولكن (OLS) لا تحتاج الى تخصيص التوزيع للاخطاء .

اما الاساس الذي اتبع لطريقة الامكان فهو استخدام طريقة نيوتن رافسن واسلوب فشر **(Newton – Raphson and Fisher' s scoring approach)** وهما من طرائق اساليب التعويض المتتالي **(Iterative procedures)** التي يتم اللجوء اليها في حالة الحصول على تقديرات يصعب حلها بالاسلوب الاعتيادي ، وعلى افتراض ان Y_1 , Y_2 ,….,Y_n متغيرات عشوائية مستقلة حيث ان Y_i يسلك وفق توزيع القيمة المتطرفة النوع الاول وللقيم الكبرى وله **p.d.f** المعطاة في الصيغة رقم (14)، وبوضع

$$Z_i = \frac{Y_i - x'_i \ \beta}{\theta} + \gamma \ , \quad i = 1, 2, \ldots.., n \qquad (16)$$

تقدير معاملات ومعلمة القياس لأنموذج انحدار القيمة المتطرفة الخطي (النوع الاول) وللقيم الكبرى مع جانب تطبيقي

فدالة الكثافة الاحتمالية ل Y_i ستكون :

$$f_i(y_i;\boldsymbol{\beta},\theta)=\theta^{-1}\exp(-z_i-\exp(-z_i)) \qquad (17)$$

نفترض $f_i(y_i;\boldsymbol{\beta},\theta)=f_{Y_i}(y_i)$ وذلك لسهولة العمل ، فتفاضلات المرتبة الاولى ل β_r و θ :

$$\frac{\partial z_i}{\partial \beta_r}=\frac{-X_{ir}}{\theta}\ ,\quad \frac{\partial z_i}{\partial \theta}=-(\ )=-\theta^{-1}\left(\frac{Y_i-x'_i\beta}{\theta^2}\right)=-\frac{Y_i-x'_i\beta}{\theta}\frac{(z_i-\gamma)}{\theta}$$

وباخذ اللوغارتم لجعل المعادلة (17) خطية لا مكانية اشتقاقها :

$$\log f_{Y_i}(y_i)=-\log\theta+(-z_i-\exp(-z_i))$$

نفترض :

$$U_r^{(i)}=\frac{\partial \operatorname{Log} f_{yi}(yi)}{\partial\beta_r}\ ,\quad U_\theta^{(i)}=\frac{\partial \operatorname{Log} f_{yi}(yi)}{\partial\theta}\ ,\quad V_{rs}^{(i)}=\frac{\partial^2 \operatorname{Log} f_{yi}(yi)}{\partial\beta_r\,\partial\beta_s}$$

$$V_{r\theta}^{(i)}=\frac{\partial^2 \operatorname{Log} f_{yi}(yi)}{\partial\beta_r\,\partial\theta}\ ,\quad V_{\theta\theta}^{(i)}=\frac{\partial^2 \operatorname{Log} f_{yi}(yi)}{\partial\theta^2}\ ,\quad r\ s=0,1,\ldots,k \qquad (18)$$

فالتفاضلات الجزئية المطلوبة في الاعلى ستكون :

$$-U_r^{(i)}\ \frac{\partial z_i}{\partial\beta_r}=-(1-e^{-z_i})=\frac{X_{ir}}{\theta}(1-e^{-z_i}),\ r=0,1,\ldots,k \qquad (19)$$

$$-U_\theta^{(i)}=-\frac{1}{\theta}-\frac{\partial z_i}{\partial\theta}-\left(-\frac{\partial z_i}{\partial\theta}\right)e^{-z_i}$$

$$-\frac{1}{\theta}=-\left(-\frac{(z_i-\gamma)}{\theta}\right)+\ )\ e^{-z_i}\frac{(z_i-\gamma)}{\theta}(-$$

$$=\theta^{-1}((z_i-\boldsymbol{\gamma})(1-e^{-z_i})-1) \qquad (20)$$

حيث : $\dfrac{\partial z_i}{\partial\beta_r}=-\dfrac{X_{ir}}{\theta}\ ,\ \dfrac{\partial z_i}{\partial\theta}=-\dfrac{(z_i-\gamma)}{\theta}$

$$-V_{rs}^{(i)}=\frac{\partial}{\partial\beta_s}U_r^{(i)}=\left(\frac{\partial}{\partial\beta_s}(1-e^{-z_i}\frac{X_{ir}}{\theta}\ )\right)$$

$$=-\frac{X_{ir}X_{is}}{\theta^2}e^{-z_i} \qquad (21)$$

حيث : $\dfrac{\partial}{\partial\beta_s}\dfrac{X_{ir}}{\theta}=\text{zero}\ ,\ \dfrac{\partial z_i}{\partial\beta_s}=-\dfrac{X_{is}}{\theta}$

$$-V_{r\theta}^{(i)}=\frac{\partial}{\partial\theta}U_r^{(i)}=\frac{\partial}{\partial\theta}\frac{X_{ir}}{\theta}(1-e^{-z_i})$$

$$=-\frac{X_{ir}}{\theta^2}[(1-e^{-z_i})+e^{-z_i}(z_i-\boldsymbol{\gamma})] \qquad (22)$$

$$-V_{\theta\theta}^{(i)}=\frac{\partial}{\partial\theta}U_\theta^{(i)}=\frac{\partial}{\partial\theta}\left[\frac{(z_i-\gamma)(1-e^{-z_i})-1}{\theta}\right]$$

$$=-\frac{1}{\theta^2}[(z_i-\boldsymbol{\gamma})^2e^{-z_i}+2(z_i-\boldsymbol{\gamma})(1-e^{-z_i})-1] \qquad (23)$$

اذا كان :

$$L(\beta,\theta)=\sum_i log\ f_{Yi}(y_i)$$

حيث ان $L(\beta,\theta)$ تشير الى لوغارتم الامكان (Log – likelihood) وعلى فرض :

$$\frac{\partial L(\beta,\theta)}{\partial\beta_r} = \sum_i U_r^{(i)},\quad \frac{\partial L(\beta,\theta)}{\partial\theta}=\sum_i U_\theta^{(i)} \quad (24)$$

وهكذا فان تقديرات الامكان الاعظم ML تعطى بواسطة حل K + 2 من المعادلات لذلك :

$$=\sum_i U_r^{(i)} \quad \sum_i x_{ir}\ \theta^{-1}\ (1-e^{-zi})$$

وبمساوتها بالصفر تصبح :

$$\sum_i x_{ir}\ (1-e^{-\hat{z}i}) = 0 \quad (25)$$

كما ان :

$$=\sum_i U_\theta^{(i)} \sum_i \theta^{-1}((z_i-\gamma)(1-e^{-zi})-1)$$

وبمساواتها بالصفر تصبح :

$$\sum_i (\hat{Z}_i-\gamma)(1-e^{-\hat{z}i})=n \quad (26)$$

ل : $r = 0, 1, \ldots\ldots, k$ وحيث ان :

$$\hat{z}_i=\frac{(Yi-x'i\ \hat{\beta})}{\hat{\theta}}+\gamma \qquad i=1,2,\ldots..,n \quad (27)$$

ولاستخراج تفاضلات المرتبة الثانية للوغارتم الامكان الاعظم نفترض :

$$\frac{\partial^2 L(\beta,\theta)}{\partial\beta r\,\partial\beta s}=\sum_i V_{rs}^{(i)},\quad \frac{\partial^2 L(\beta,\theta)}{\partial\beta r\,\partial\theta}=\sum_i V_{r\theta}^{(i)},\quad \frac{\partial^2 L(\beta,\theta)}{\partial\theta^2}=\sum_i V_{\theta\theta}^{(i)} \quad (28)$$

فان :

$$-\sum_i V_{rs}^{(i)} = -\sum_i x_{ir}x_{is}\quad \theta^{-2}\ e^{-zi}$$

وبمساواتها بالصفر تصبح :

$$\sum_i x_{ir}x_{is}\quad e^{-\hat{z}i}=0 \quad (29)$$

$$-\sum_i V_{r\theta}^{(i)} = -\sum_i x_{ir}\ \hat{\theta}^{-2}\ [(1-e^{-zi})+e^{-zi}(z_i-\gamma)]$$

بمساواتها بالصفر تصبح :

$$\sum_i x_{ir}\quad e^{-\hat{z}i}(1-(\hat{z}_i-\gamma))=\sum_i x_{ir} \quad (30)$$

$$-\sum_i V_{\theta\theta}^{(i)}=\frac{1}{\theta^2}\ \sum_i[(z_i-\gamma)^2\ e^{-zi}+2(z_i-\gamma)(1-e^{-zi})-1]$$

وبمساواتها بالصفر :

$$\sum_i [\,(\hat{z}_i - (1-e^{-\hat{z}_i}\ \gamma)^2 e^{-\hat{z}_i} + 2(\hat{z}_i - \gamma)\,) = n \qquad (31)$$

أ – طريقة نيوتن رافسن Newton – Raphson Method : وهي احدى الطرائق العددية لحل معادلات الامكان اعلاه عن طريق ايجاد جذور تقريبية من خلال الحل المتتالي لها وذلك باستخدام طريقة المربعات الصغرى كتقديرات اولية لمعاملات الانحدار ومعلمة القياس والتي تستخدم للحصول على تقديرات معاملات الانحدار ومعلمة القياس لطريقة الامكان الاعظم، ويتم ذلك بوضع :

$$\underset{\sim}{D}{}^{(\ell)}_{1(k+2)*1} = \begin{bmatrix} \partial L(\beta,\theta)/\partial\beta_j \\ \partial L(\beta,\theta)/\partial\theta \end{bmatrix}_{\beta=\hat{\beta}^{(\ell)},\ \theta=\hat{\theta}^{(\ell)}} ,\quad J = 0, 1, \ldots, k \qquad (32)$$

$$\underset{\sim}{D}{}^{(\ell)}_{2(k+2)*(k+2)} = \begin{bmatrix} \frac{\partial^2 L(\beta,\theta)}{\partial\beta_0^2} & \frac{\partial^2 L(\beta,\theta)}{\partial\beta_0\partial\beta_1} & \cdots & \frac{\partial^2 L(\beta,\theta)}{\partial\beta_0\partial\beta_k} & \frac{\partial^2 L(\beta,\theta)}{\partial\beta_0\partial\theta} \\ - & - & & - & - \\ \frac{\partial^2 L(\beta,\theta)}{\partial\beta_0\partial\beta_k} & \frac{\partial^2 L(\beta,\theta)}{\partial\beta_1\partial\beta_k} & \cdots & \frac{\partial^2 L(\beta,\theta)}{\partial\beta_k^2} & \frac{\partial^2 L(\beta,\theta)}{\partial\beta_k\partial\theta} \\ \frac{\partial^2 L(\beta,\theta)}{\partial\beta_0\partial\theta} & \frac{\partial^2 L(\beta,\theta)}{\partial\beta_1\partial\theta} & \cdots & \frac{\partial^2 L(\beta,\theta)}{\partial\beta_k\partial\theta} & \frac{\partial^2 L(\beta,\theta)}{\partial\theta^2} \end{bmatrix}_{\beta=\hat{\beta}^{(\ell)},\theta=\hat{\theta}^{(\ell)}}$$

............(33)

- حيث ان $\hat{\beta}^{(\ell)}$, $\hat{\theta}^{(\ell)}$ تشير الى التقريبات ل $\hat{\beta}$ و $\hat{\theta}$ الى درجة (ℓ) من التكرار اما التقريبات الجديدة فتعطى بواسطة السلسلة التالية

$$\begin{bmatrix} \hat{\beta}^{(\ell+1)} \\ \hat{\theta}^{(\ell+1)} \end{bmatrix} = \begin{bmatrix} \hat{\beta}^{(\ell)} \\ \hat{\theta}^{(\ell)} \end{bmatrix} - \left[D_2^{(\ell)}\right]^{-1} D_1^{(\ell)} \quad (34)$$

التكرار المستمر باستخدام السلسلة المعطاة في (34) حتى يستخرج التقريب الملائم .

ب – طريقة فشر Fisher's scoring method : اساس هذه الطريقة اجراء بعض التعديلات على طريقة نيوتن رافسن وتعتمد على ايجاد التوقع للتفاضلات الجزئية الثنائية لمصفوفة المعلومات ، وتحتاج هذه الطريقة اولا الى بعض نتائج التقدير البسيط للمتغيرات العشوائية:

$$Z_i = \theta^{-1}(y_i - \underset{\sim}{x}_i'\ \boldsymbol{\beta}) + \gamma \qquad i = 1, 2, \ldots, n$$

التي تمثل موجه الاخطاء الحقيقية التي تكون مستقلة ومتطابقة التوزيع مع دالة كثافة احتمالية p.d.f التالية:

$$f_z(z) = \exp(-z - e^{-z}), \quad (-\infty < z < \infty) \qquad (35)$$

حيث ان (35) تمثل دالة الكثافة الاحتمالية لتوزيع القيمة المتطرفة النوع الاول وللقيم الكبرى وان الدالة المولدة للعزوم لتوزيع E.V.D النوع الاول القياسي :

$$M_z(t) = \int_{-\infty}^{\infty} e^{zt} e^{-z} \exp(-\exp(-z))\,dz = \Gamma(1-t), \ (t<1) \qquad (36)$$

ومن خلالها فان :

$$E(z^r e^{zt}) = d^r M_z(t) / dt^r = d^r \Gamma(1-t)/dt^r \quad (37)$$

مع $M_z'(t) = -\Gamma'(1-t)$, $E(z e^{-z}) = -(1-\boldsymbol{\gamma}) = -0.422784$

$$M_z''(t) = \Gamma''(1-t) , E(z^2 e^{-z}) = \Gamma''(2) = 0.823680 , E(e^{-z}) = \Gamma(2) = 1 \quad (38)$$

حيث ان $\frac{d}{dx}$ **Log Γ (x)** و $\frac{d^2}{dx^2}$ **Log Γ (x)** تشير الى دوال **digamma** و **Trigamma** بالتعاقب وباستخدام المعادلات (36) ، (37) نحصل على :

$$E[e^{-z}(z-\boldsymbol{\gamma})] = -\Gamma'(2) - \boldsymbol{\gamma} = -1$$

$$E[e^{-z_i}((z_i-\boldsymbol{\gamma})^2 - 2(z_i-\boldsymbol{\gamma}) - 1] = 2.644934$$

نفرض : $I_{rs} = E(-\sum_i V_{rs}^{(i)}), \quad I_{r\theta} = E(-\sum_i V_{r\theta}^{(i)}) , I_{\theta\theta} = E(-\sum_i V_{\theta\theta}^{(i)})$

باستخدام (21) ,(22) ,(23) ,(38) فالعناصر في مصفوفة المعلومات يتم الحصول عليها عن طريق الصيغ التالية :

$$I_{rs} = \theta^{-2} \sum_i x_{ir}\, x_{is} \quad I_{r\theta} = -\theta^{-2} \sum_i x_{ir} , I_{\theta\theta} = 2.644934\, n\, \theta^{-2} \quad (39)$$

والتي تستبدل بتلك الموجودة في المصفوفة $\underset{\sim}{D}_2$ والتقدير لدرجة (ℓ) من التكرار . استخدام التقريب الحالي $\bar{\theta}^{(\ell)}$ الى $\bar{\theta}$ فالعناصر تصبح مستقلة عن $\underset{\sim}{\beta}$. التكرار المستمر باستخدام السلسلة المعطاة في (34) حتى يتم استخراج التقارب الملائم . فمصفوفة المعلومات (**Information matrix**) ستكون :

$$\underset{\sim}{I} = \frac{1}{\theta^2}\begin{bmatrix} 2.644934n & -n & -n\,\bar{x}_1 & -n\,\bar{x}_2 & \cdots & -n\,\bar{x}_k \\ -n & n & n\,\bar{x}_1 & n\,\bar{x}_2 & \cdots & n\,\bar{x}_k \\ -n\,\bar{x}_1 & n\,\bar{x}_1 & \Sigma x_{i1}^2 & \Sigma x_{i1}x_{i2} & \cdots & \Sigma x_{i1}x_{ik} \\ -n\,\bar{x}_2 & n\,\bar{x}_2 & \Sigma x_{i1}x_{i2} & \Sigma x_{i2}^2 & \cdots & \Sigma x_{i2}x_{ik} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ -n\,\bar{x}_k & n\,\bar{x}_k & \Sigma x_{i1}x_{ik} & \Sigma x_{i2}x_{ik} & \cdots & \Sigma x_{ik}^2 \end{bmatrix} \quad (40)$$

وحيث انه في الغالب يكون المتوسط العام (**General mean**) مهما في النموذج الخطي ، فاذا اخذنا بنظر الاعتبار العبارة التالية (**Without loss of generality**) اي من دون خسارة للعمومية نفترض :

اذا كانت قيم **x** تاخذ القيم المركزية اي بمعنى اخر :

$$\bar{x}_j = 0 \;, j = 1,2,\ldots,k \quad \text{او} \quad x_{i0} = 1 , i = 1, 2, \ldots, n , \sum_{i=1}^{n} x_{ij} = 0 , j = 1, 2, \ldots, k$$

فمصفوفة المعلومات سوف تجزا الى الشكل التالي :

$$\underset{\sim}{I} = \frac{1}{\theta^2}\begin{bmatrix} I_{1_{2*2}} & Q_{2*k} \\ Q_{2*k} & I_{2_{k*k}} \end{bmatrix} \qquad (41)$$

حيث ان I_1 تشير الى θ و β_0 ، I_2 تشير الى $\beta_1, \beta_2, \dots, \beta_k$ و

$$I_1 = \begin{bmatrix} 2.644934n & -n \\ -n & n \end{bmatrix} \quad , I_2 = \begin{bmatrix} \sum_i x_{i\,1}^2 & \sum_i x_{i\,1}x_{i\,2} & \cdots & \sum_i x_{i\,1}x_{i\,k} \\ \sum_i x_{i\,1}x_{i\,2} & \sum_i x_{i\,2}^2 & \cdots & \sum_i x_{i\,2}x_{i\,k} \\ \vdots & \vdots & \vdots & \vdots \\ \sum_i x_{i\,1}x_{i\,k} & \sum_i x_{i\,2}x_{i\,k} & \cdots & \sum_i x_{i\,k}^2 \end{bmatrix} \qquad (42)$$

وهكذا فالمعكوس لمصفوفة $\underset{\sim}{I}$: $\quad \underset{\sim}{I}^{-1} = \theta^2 \begin{bmatrix} I_1^{-1} & Q \\ Q & I_2^{-1} \end{bmatrix}$ **حيث ان :**

$$I_1^{-1} = \frac{1}{n}\begin{bmatrix} 0.607927 & 0.607927 \\ 0.607927 & 1.607927 \end{bmatrix} \;,\; I_2^{-1} = \left\|\sum_i x_{i\,r}x_{i\,s}\right\|^{-1}$$

فستكون التقريبات القياسية (تباينات العينة الكبيرة)

$$=(\hat{\beta}_0)\ \text{Var}\ \frac{1.607927\,\theta^2}{n}, \text{Var}\ (\hat{\theta}) \ =(\hat{\beta}_0, \hat{\theta}\,), \text{Cov}\ \frac{0.607927\,\theta^2}{n} = \ \frac{0.607927\,\theta^2}{n}$$

والتقريب لمصفوفة التباين المشترك ل) $\hat{\beta}_1, \cdots, \hat{\beta}_k$ = $\hat{\beta}_i'$) هو :

$$= \text{Cov}\ (\hat{\beta}_1, \hat{\beta}_2, \cdots, \hat{\beta}_k)\ (\hat{\beta}_i)\ \ \theta^2 \approx \left\|\sum_i x_{i\,r}x_{i\,s}\right\|^{-1} \qquad (43)$$

وفي حالة وجود متغير مستقل واحد (حالة الانحدار البسيط) يمكن ان نحصل على مقدرات الامكان الاعظم من خلال الاتي ، بما ان دالة الامكان الاعظم لانموذج انحدار القيمة المتطرفة (النوع الاول) وللقيم الكبرى التي لها دالة كثافة احتمالية :

$$L = -\,n\ \ \text{Log}\,\theta - n\boldsymbol{\gamma} - \ -\sum_{i=1}^{n}\left(\frac{\text{Yi} - \ \text{x' i}\,\beta}{\theta}\right)\sum_{i=1}^{n} exp\left(-\left(\frac{\text{Yi} - \ \text{x' i}\ \ \beta}{\theta} + \gamma\right)\right) \qquad (44)$$

وبافتراض ان :

$$\underset{\sim}{x}'_i\,\boldsymbol{\beta} = \boldsymbol{\beta}_0 + \boldsymbol{\beta}_1 + \ldots + \boldsymbol{\beta}_k = \ \sum_{r=0}^{k}\beta_r x_{i\,r} \ \ , x_{i\,0} = 1$$

فاستخراج التفاضلات الاولية وذات المرتبة الثانية لكل من $r = 0 , 1$ اي β_0 , β_1 و θ ومنها فان :

$$\underset{\sim}{D}^{(1)} = (\ \boldsymbol{\beta}_0\ ,\ \ \boldsymbol{\beta}_1\ \ ,\theta\ \ ) = \begin{bmatrix} \partial L/\partial\beta_0 \\ \partial L/\partial\beta_1 \\ \partial L/\partial\theta \end{bmatrix} \qquad (45)$$

وبمساواتها بالصفر نحصل على تقديرات الإمكان الاعظم اي ان :

$$\underset{\sim}{D}^{(1)} \ = (\ \hat{\beta}_0, \hat{\beta}_1, \hat{\theta}\ ) = \ \ Q \qquad (46)$$

اما مصفوفة المعلومات فستكون :

$$\mathbf{I} = \theta^{-2}\begin{bmatrix} n & 0 & -n \\ 0 & \sum x_{i\,1}^2 & 0 \\ -n & 0 & 2.644934n \end{bmatrix} \qquad (47)$$

فاذا كانت $\tilde{\theta}^{(1)}$ و $\tilde{\beta}_1^{(1)}$ و $\tilde{\beta}_0^{(1)}$ تشير الى التقريبات الاولية لتقديرات الامكان الاعظم (ML)، وباستخدام طريقة فشر فالتقديرات الجديدة تعطى عن طريق التكرار المستمر باستخدام السلسلة المعطاة في (34) حتى استخراج التقارب الملائم لمعاملات الانحدار ومعلمة القياس (التي تم استخدامها في الجانب التطبيقي بعد تبديل قيم المشتقات الثانية في مصفوفة المعلومات بقيم متوقعة للمشتقات الثانية التي تتطلبها طريقة فشر) .

4-3 خصائص العزوم لمقدرات المربعات الصغرى الاعتيادية

مقدر المربعات الصغرى الاعتيادية (O.L.S) ل $\boldsymbol{\beta}$ هو :

$$\tilde{\beta} = (\underset{\sim}{x}'\underset{\sim}{x})^{-1}\underset{\sim}{x}'\underset{\sim}{y} \qquad (48)$$

يوضح تقديرات المعلمة $\tilde{\beta}_r$ حيث ان $r = 0, 1, 2, \ldots, k$ وتعد افضل مقدر (اصغر تباين) غير متحيز (B.L.U.E.) بالمقارنة مع كل التقديرات غير المتحيزة التي تكون خطية في المشاهدات (Y_i) . علما بان :

$$E(\tilde{\beta}) = \boldsymbol{\beta}\ , \ V.C(\tilde{\beta}) = \boldsymbol{\sigma}^2 (\underset{\sim}{x}'\underset{\sim}{x})^{-1} \qquad (49)$$

وان مقدر المربعات الصغرى ل $\boldsymbol{\sigma}^2$ وهو تقدير غير متحيز يساوي :

$$(n-k-1)^{-1}(\underset{\sim}{Y} - \underset{\sim}{X}\tilde{\beta} = \tilde{\sigma}^2)'(\underset{\sim}{Y} - \underset{\sim}{X}\tilde{\beta}) \qquad (50)$$

وحيث ان :

$$\boldsymbol{\sigma}^2 = \frac{1}{6}\boldsymbol{\pi}^2\theta^2$$

وهكذا فان :

$$Var(\tilde{\beta}_0) = \frac{\pi^2\theta^2}{6\,n}\ , \ (\boldsymbol{\pi} = 3.1415903) \qquad (51)$$

اما مصفوفة التباين والتباين المشترك ل $\tilde{\beta}_1, \tilde{\beta}_2, \ldots, \tilde{\beta}_k$ فهي :

$$Cov(\tilde{\beta}_1, \tilde{\beta}_2, \ldots, \tilde{\beta}_k) = \frac{1}{6}, \pi^2\theta^2 \underset{\sim}{I}_2^{-1} = \underset{\sim}{I}_2^{-1} \ \|\textstyle\sum_i x_{ir} x_{is}\|^{-1}, \ r = s = 0,1,2,\ldots,k \qquad (52)$$

اما مقدر المربعات الصغرى الاعتيادية (OLS) $\tilde{\theta}$ ل θ

$$\tilde{\theta} = \theta\,[\,1 + (\ -\ \theta^2)/\theta^2\,]^{1/2}\tilde{\theta}^2 \qquad (53)$$

اما قيمة التوقع والتباين ل $\tilde{\theta}$ فهي على التوالي :

$$E(\tilde{\theta}) = \theta\,[\,1 - \frac{1}{4\,(n-k-1)}\,[\,1 + \frac{1.2}{n-k-1} \ \sum_{i=1}^{n}(1 - h_{ii})^2\,]] \qquad (54)$$

$$Var(\tilde{\theta}) = \frac{1.1\,\theta^2}{n} \qquad (55)$$

3-5 الكفاءة التقريبية Asymptotic Efficiency

اولا: الكفاءة التقريبية ل ($\tilde{\beta}_0$) بالمقارنة مع ($\hat{\beta}_0$):

$$\frac{Var(\hat{\beta}_0)}{Var(\tilde{\beta}_0)} = \left(\frac{1.607927\,\theta^2}{n}\right)/\left(\frac{\pi^2\theta^2}{6n}\right) = 0.978 \qquad (56)$$

ثانيا: الكفاءة التقريبية ل ($\tilde{\beta}_j$) مع ($\hat{\beta}_j$) ولكل مقدر على حدة :

$$\frac{Var(\hat{\beta}_j)}{Var(\tilde{\beta}_j)} = \left(\frac{1}{\theta^2}.\frac{1}{6}\,\pi^2\theta^2\right)^{-1} = 0.608 \quad , j = 1,2, \ldots,k \qquad (57)$$

ثالثا: الكفاءة التقريبية ل ($\tilde{\theta}$) بالمقارنة مع ($\hat{\theta}$) :

$$\frac{Var(\hat{\theta})}{Var(\tilde{\theta})} = \left(\frac{6\,\theta^2}{\pi^2 n}\right)/\left(\frac{1.1\,\theta^2}{n}\right) = 0.553 \qquad (58)$$

ومن خلال القيم الثلاث التي استخرجت وجميعها اقل من واحد صحيح ، اي ان تقديرات الامكان الاعظم (MLE) اكثر كفاية من تقديرات المربعات الصغرى الاعتيادية (OLSE) في تقدير معالم النموذج .

3-6 اجراءات الاستدلال الاحصائي لنموذج انحدار القيمة المتطرفة (النوع الاول) وللقيم الكبرى

نتائج التوزيع التقريبي (Asymptotic distribution) لتقديرات الامكان الاعظم ML تكون :

$$\hat{\beta}_0 \ \mathrm{N}(\overset{a}{\sim}\ \beta_0\ , 1.6079\, n^{-1}\,\theta^2 \qquad (59)$$

$$\hat{\beta}_r\ , \beta_r \mathrm{N}(\overset{a}{\sim}\ \theta^2 V^{r\,r}\ ) \quad , r = 1\ , 2\ , \ldots, k \qquad (60)$$

$$\hat{\theta}\ \overset{a}{\sim}\ \mathrm{N}(\ \theta\ , 0.6079\ \ n^{-1}\,\theta^2\ ) \qquad (61)$$

حيث ان $V^{r\,r}$ هي العنصر (r , r) لمعكوس مجموع مصفوفة الضرب المتتالي :

$$\underset{\sim}{V} = \begin{bmatrix} \sum_i x_{i1}^2 & \sum_i x_{i1}x_{i2} & \cdots & \sum_i x_{i1}x_{ik} \\ \sum_i x_{i1}x_{i2} & \sum_i x_{i2}^2 & \cdots & \sum_i x_{i2}x_{ik} \\ \vdots & \vdots & \vdots & \vdots \\ \sum_i x_{i1}x_{ik} & \sum_i x_{i2}x_{ik} & \cdots & \sum_i x_{ik}^2 \end{bmatrix} \qquad (62)$$

فاذا تحقق الشرط ($\sum x_{ir} = 0$, r = 1, 2, …,k) فان التباين المشترك التقريبي (Asymptotic covariance) بين كل من $\hat{\beta}_r$ و $\hat{\theta}$ وبين $\hat{\beta}_r$ و $\hat{\beta}_0$ يكون مساويا للصفر لكل من r = 1, 2 ,…,k وكذلك يكون :

$$cov_a\,(\ \hat{\beta}_0, \hat{\theta})\ \ 0.6079\, n^{-1}\theta^2, \ = \ (cov_a\, \hat{\beta}_r\ ,) = \ \hat{\beta}_s\theta^2 V^{r\,s}\ , r,s \neq 1,2,..,k \qquad (63)$$

تقدير معاملات ومعلمة القياس لأنموذج انحدار القيمة المتطرفة الخطي

(النوع الاول) وللقيم الكبرى مع جانب تطبيقي

ولتحديد حدود الثقة التقريبية لمعاملات الانحدار ومعلمة القياس بالاعتماد على تقديرات ML ويتم ذلك من خلال الاستناد الى اخذ التوزيعات التقريبية للمتغيرات العشوائية المدارية (Pivotal R.) التي تتطابق مع نظرية الحد المركزي وهي كالاتي :

$$\theta^* = \hat{\theta}/\theta = \beta_r^*, \ - \hat{\beta}_r(\ \beta_r \ ) / \theta , r = 0,1,\ldots,k \qquad (64)$$

التي تكون مستقلة عن β و θ ، ومن الملاحظ ان الخصائص الخاصة بهذه التوزيعات تمتلك العمومية لكل حجوم العينات ، لذلك يجب استخدام صفات التوزيع ل θ^* لصنع الاستدلالات الاحصائية حول θ وصفات التوزيع للمتغير العشوائي :

$$T_r \ - \hat{\beta}_r = (\ ) /\beta_r \ \hat{\theta} \qquad (65)$$

للاستدلال الاحصائي حول β_r . وحيث ان $T_r = \beta_r^*/\theta^*$ فان التوزيع ل T_r هو كذلك مستقل عن β و θ ومن الملاحظ احصائيا بان :

$$T_1 = \ ,\frac{\bar{X}-\mu}{S/\sqrt{n}} \quad \frac{(n-1)S^2}{\sigma^2} \backsim \chi^2_{n-1} \ \ T_2 \longrightarrow = \frac{S^2\,(n-1)}{\sigma^2} \qquad (66)$$

حيث ان المتغير العشوائي (R.V.) (T_1) يعتمد فقط على قيمة المعلمة المجهولة (μ) وتوزيعه مستقل ل μ و . اما (T_2) فهو متغير عشوائي (R.V.) يعتمد فقط على σ والتوزيع له مستقل عن μ و σ .و هذا ما ينطبق على المتغيرات العشوائية التي اعتمدت وهي : $=\beta_r^* \ \frac{\hat{\beta}_r - \beta_r}{\theta} \ , \ \theta^* = \frac{\hat{\theta}}{\theta}$

وعليه فحدود الثقة التقريبية ذات الجانبين ل θ والمستندة الى مقدر الامكان الاعظم $\hat{\theta}$ هي :

$$\frac{\hat{\theta}}{1+0.7797n^{-1/2}U_{1-\alpha/2}}(, \ \frac{\hat{\theta}}{1-0.7797n^{-1/2}U_{-\alpha/2}} \ ) \qquad (67)$$

حيث ان $U_{1-\alpha}$ تمثل المئوية ($1-\alpha$) العليا للتوزيع الطبيعي القياسي N(0,1) بينما تكون حدود الثقة التقريبية لمعاملات الانحدار والمستندة على تقديرات الامكان الاعظم β_r :

$$\hat{\beta}_r(- \hat{\theta} \ t_{r,1-\alpha/2} \ , \hat{\beta}_r \ \ -) \ , r = 1,2,\ldots,k \hat{\theta} \ t_{r,\alpha/2} \qquad (68)$$

حيث ان :

$$t_{r,1-\alpha} = - \ \ t_{r,\alpha} = \frac{(V^{rr})^{1/2}\, U_{1-\alpha}}{(1-U_{1-\alpha}^{\ 2}\ \frac{0.6079}{n})^{1/2}} \qquad (69)$$

و V^{rr} هي العنصر (r,r) لمعكوس مجموع مصفوفة الضرب المتتالي كما ذكرت سابقا .

كما تم التركيز في اختبار الفرضيات المتعلقة بمعاملات الانحدار التي تكون قيمها صفرا وقيمة θ (معلمة القياس) تساوي واحدا وذلك لتجاوز الافتراض المعتمد على افتراض قيمة مثل ($\beta = 3$) لان من المحتمل ان لاتكون ($\beta = 3$) وانما تاخذ قيمة اخرى اي بمعنى الحصول على نمط واحد لقبول الفرضية H_0 او رفضها عندما تكون $\beta = 0$ و $\theta = 1$. ومن الجدير بالذكر احصائيا ان نتيجة حدود الثقة لها هي علاقة متطابقة مع نتيجة اختبار الفرضيات التي يتم التركيز فيها.

7-3 حسن المطابقة لتوزيع القيمة المتطرفة: عند افتراض توزيع معين للاخطاء فان المتغير المعتمد يتبع توزيعه التوزيع المفترض للاخطاء نفسه ومن اجل فحص هذه الاخطاء واختبارها على انها تسلك وفق توزيع القيمة المتطرفة (النوع الاول) تم استخدام احد اختبارات حسن المطابقة (Goodness of fit tests) الخاصة لهذا التوزيع وهو اختبار الرسم الاحتمالي (Probability plot) والذي يتم من خلال رسم قيم المتغير $R_{(i)}$ (الاحصاءات المرتبة لقيم الاخطاء المشاهدة) مقابل قيم K_i حيث ان

$$R_i \; , \frac{Y_i - x'_i\,\hat{\beta}}{\hat{\theta}} = \; = -\log\left(\log\left(K_i\,\frac{n+1}{n+1-i}\right)\right),\; i=1,2,\ldots,n \qquad (70)$$

ويتم التقويم الاولي للرسم البياني الاحتمالي فيكون ذلك بقدر تراص وضع تلك النقاط باخذها تقريبا خطا مستقيما قدر الامكان وعكسه يعطي نقصا في مطابقة التوزيع المفترض ، اما صيغة الاحصاءة (S_R) الخاصة باختبار الرسم الاحتمالي فهي :

$$S_R = 1 - r^2(R_{(i)}\,, K_i\,)\,,\; i=1,2,\ldots,n \qquad (71)$$

حيث ان $r(R_i,K_i)$ تمثل معامل الارتباط بين K_i و R_i وبمقارنة قيمة S_R مع قيم جدولية ولمستوى معنوية معين والموجودة في جداول خاصة اعطيت من قبل Smith and Bain [9] بهذا الاختبار ولحجوم العينة (80,60,40,20,8) يمكن تحديد فيما اذا كان التوزيع المفترض للاخطاء هو التوزيع الملائم لبيانات تلك العينة .

3- الجانب التطبيقي

1-4 التطبيق الاول : ازمنة البقاء (Survival Times)

يعد مرض ابيضاض الدم (Leukaemias) مرض التكاثر او التوالد الشاذ للخلايا المنتجة للدم بسبب الزيادة التدريجية لترشيح نخاع العظم ، فهو ورم خبيث ينشا عن اضطراب للخلايا المنتجة للدم التي تخضع لتغير جوهري بسبب منبعث من قيود طبيعية تفرض على فاعلية التكاثر ويحدث المرض باشكال عديدة مختلفة ويصنف الى ابيضاض الدم الحاد وابيضاض الدم المزمن و ذلك طبقا الى الملاحظة السريرية للمريض والى الشوكاني (اي ما يتعلق بالنخاع الشوكي)واللمفاوي (متعلق بالنسيج اللمفاوي) طبقا الى حقل الخلية السائدة والمتضمنة لحالة اللوكيميا .حيث سيتم التركيز في ابيضاض الدم الحاد (Acute L.) ، حيث يقسم ابيضاض الدم النقياني الحاد AML (Acute Myelogeneous Leukaemia) الى سبعة انواع تاخذ الرموز من M_1 الى M_7 وابيضاض الدم اللمفاوي الحاد ALL (Acuye Lymphoblastic Leukaemia) الى ثلاثة انواع هي (L_3 , L_1 , L_2) , اما عن اهم العوامل التي تسبب المرض فجزءا منها مازال لحد الان لم يعرف لدى العلماء وخاصة الطبيعة الدقيقة للتغيرات الاساسية التي تودي على مستوى جزئي الى تطور مرض اللوكيميا وانشائه ولكن هناك عددا من العوامل المساعدة التي يشار اليها كعوامل تعمل على تطور التغير الورمي وتتضمن الاشعاع ، المواد الكيميائية ، الفايروسات ،اسباب تعود الى الاستعداد الوراثي للمرض والعوامل النفسية ويلاحظ ان اغلبية المرضى بهذا المرض يموتون كنتيجة للمرض ولو ان زمن البقاء للمريض يمكن ان يطول من خلال المعالجة فاجراءات العلاج يمكن ان تساعد المريض على مواصلة الحياة وتساعد على اطالتها. ويمكن ان يحدث المرض في اي عمر حيث انه في الاطفال يكون حدوثه اكثر في الست سنوات الاولى من العمر اما عند البالغين فيحدث في كل الاعمار وكذلك يحدث عند الاشخاص الكبار في السن .

4-1-1 اسلوب جمع البيانات

ان اسلوب جمع البيانات كان باخذ الفايلات الخاصة بالمرضى والذين انقطعوا عن المراجعة اي المتوفين) المصابين بمرض (1) ابيضاض الدم النقياني الحاد (2) ابيضاض الدم اللمفاوي الحاد للفترة من 1985/10/20 الى 1994/4/20. وذلك من عيادة امراض الدم (مستشفى مدينة الطب) فتم تدوين المعلومات المطلوبة عن كل مريض من خلال استمارة الطبلة الخاصة بالمصاب الذي يوضح من خلالها شرح تفصيلي لحالة المريض وتاريخ التشخيص للمرض وفصول العلاج التي اخذها وتواريخها . وقد كان عدد المشاهدات 29 مشاهدة (ابيضاض الدم النقياني الحاد AML) و 43 مشاهدة (ابيضاض الدم اللمفاوي الحاد ALL) ومن خلال هذه الاستمارات والمعلومات الاخرى المذكورة في الفايلات تم اخذ المعلومات التالية عن كل مريض وهي: الجنس، العمر، بداية التشخيص للمرض، تاريخ اخر مراجعة للعيادة، كمية كريات الدم البيضاء WBC (White biood cells) (عند تشخيص المرض)، كمية الخلايا المهلكة (Blast cells) عند تشخيص المرض. وكذلك تحديد العاملين الاخيرين في اخر فحص للمريض قبل انقطاعه عن المراجعة وقد لوحظ في اغلبها ارتفاع كبير في نسبة هذين العاملين في الدم مما كانا سببا في الوفاة) لذلك اعتمد زمن البقاء كمتغير معتمد وكمية كريات الدم البيضاء كمتغير مستقل رغم ان هناك 13 عامل يوثر على الدم ولكن بعضها غير مقاس لذلك اكتفينا باخذ عامل واحد وهو كمية كريات الدم البيضاء ولكن على الرغم من ذلك تم اختبار تاثير بعض تلك العوامل وهي كمية الخلايا المهلكة (Blast cells) والعمر كلا على انفراد كعوامل موثرة في زمن البقاء وظهرت النتائج ايجابية (تاثير سلبي على المرض) وقد اخذت بعض النقاط الاساسية عند اختيار المجتمعات الخاصة بالبحث وهي :

1- تم استبعاد الحالات التي لم تذكر معلومات كافية فيها عن تاريخ تشخيص المرض ولا توجد اي معلومات عن حالة المريض والفحوصات الخاصة به .
2- تم تحديد زمن البقاء لكل مريض من خلال حساب الفترة من تاريخ بداية تشخيص المرض الى تاريخ اخر مراجعة له للعيادة (وقد اخذ بنظر الاعتبار ان بداية تشخيص المرض لا تعني تاريخ دخوله الى المستشفى بل تاريخ اول فحص للدم شخص من خلاله اصابته بالمرض سواء كان قبل تاريخ دخوله الى المستشفى او بعده بالاعتماد على الفايلات الخاصة بالمرضى).[1]
3- تم اخذ عامل واحد من العوامل التي توثر في زمن البقاء وهو كمية كريات الدم البيضاء WBC .
4- تم سحب عدة عينات بصورة عشوائية وبحجوم مختلفة لكل حالة (ALL,AML) وبعد اختبارها تم اختيار افضلها وهي التي نتطرق اليها في بحثنا .
5- وحدة القياس الخاصة بزمن البقاء هي الاسابيع ، اما وحدة القياس الخاصة بكمية كريات الدم البيضاء فهي نسبة كريات الدم البيضاء في كل ميكرولتر من الدم وفي احيان اخرى تعطى بواسطة كل لتر وقد تم تحويلها كلها الى الحالة الاولى . وفي ضوء ذلك فقد شمل مجتمع البحث الاول الخاص بمرض (AML) على 19 مريضا مصابا بهذا المرض اما مجتمع البحث الثاني الخاص بمرض (ALL) فشمل على 22 مريضا مصابين بهذا المرض .

2-1-4 التحليل الاحصائي

اولا : ابيضاض الدم النقيائي الحاد (AML)

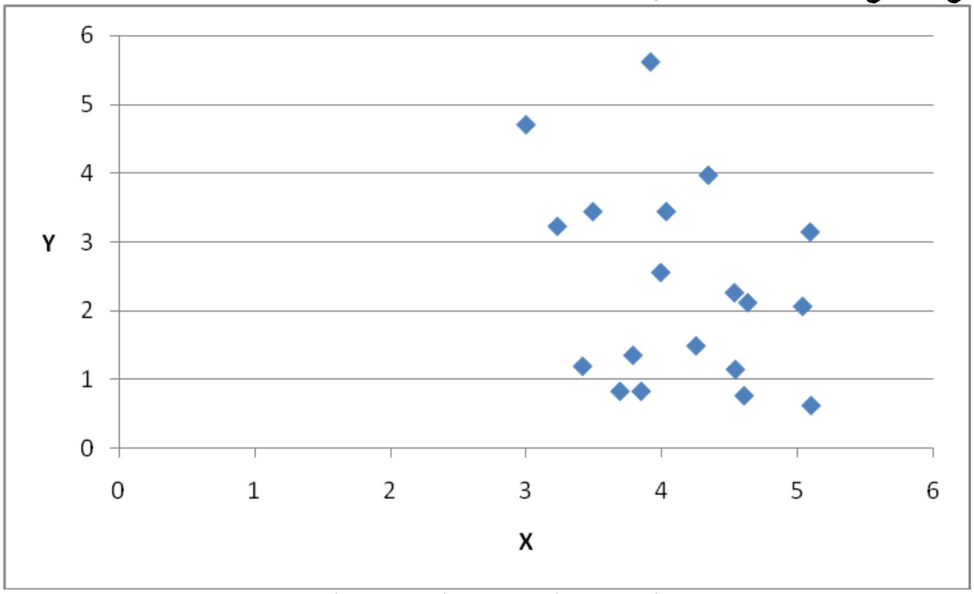
-1 **مجتمع البحث : لقد شمل مجتمع البحث على 19 مريضا مصابا بمرض (AML) الذي تتمثل قراءته ببعدين اولهما هو المتغير التابع (Y) الذي يمثل عدد الاسابيع التي يعيشها المريض حتى الوفاة ومن خلال استخدام عدة تحويلات وجدنا ان افضل تحويل كان باستخدام التحويل اللوغارتمي والذي حقق الحالة الخاصة بنا وهي توزيع القيمة المتطرفة (حيث ان Y=Log (T) اي ان توزيع القيمة المتطرفة يمثل لوغارتم توزيع زمن البقاء) والمتغير المستقل (X) ويمثل لوغارتم كريات الدم البيضاء للمريض والمبينة في الجدول رقم (1) وباتخاذ العلاقة مابين المتغيرين لشكلا خطيا بسيطا على وفق نموذج الانحدار الاتي :**

$$Y_i = \beta_0 + \beta_1 (\; X_i \;\; \theta Z_i + \; - \bar{X} \;) \quad , i = 1,2,\ldots,19 \qquad (72)$$

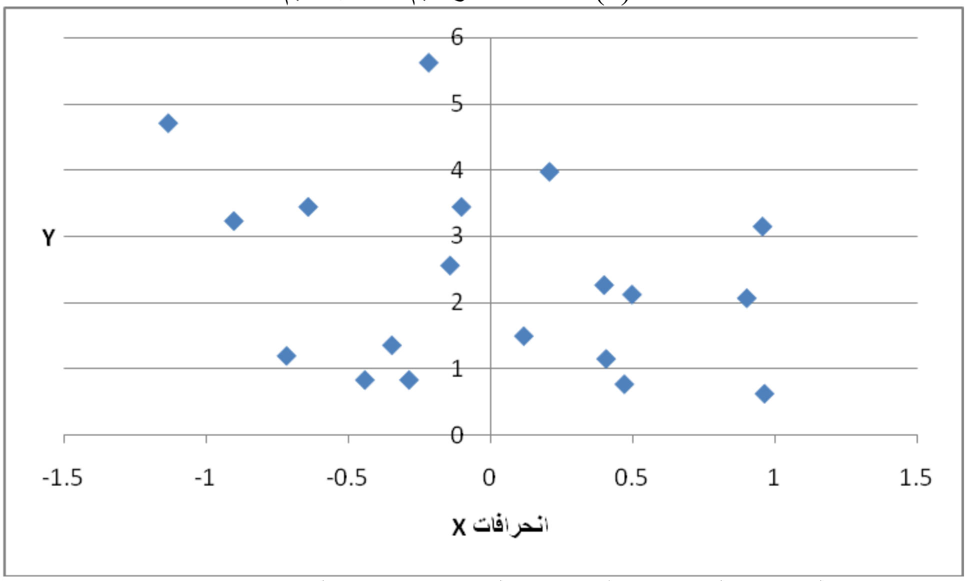
جدول رقم (1) بيانات ازمنة البقاء للمرضى المصابين بمرض AML
(Y يمثل لوغارتم توزيع زمن البقاء (t) و X يمثل لوغارتم كريات الدم البيضاء للمريض)

t_i	Y_i	X_i	R_i	$\hat{\mu}_i$
52.857	3.968	4.338	1.55057	2.29063
12.857	2.554	3.989	0.424	2.09533
3.286	1.1897	3.415	-1.27366	2.56753
9.571	2.259	4.53	0.02401	2.23303
1.857	0.619	5.093	-1.33588	2.06413
3.857	1.3499	3.785	-1.02297	2.45653
31.143	3.439	4.029	0.97587	2.38333
23.143	3.142	5.086	0.99444	2.06623
25.143	3.225	3.23	0.55647	2.62303
110.143	4.702	3	1.85802	2.69202
274.286	5.614	3.914	2.95455	2.41783
4.429	1.488	4.248	-0.76691	2.31763
2.286	0.827	3.69	-1.53268	2.48503
8.286	2.115	4.628	-0.08193	2.20363
7.856	2.061	5.031	-0.02009	2.08273
3.143	1.145	4.537	-1.00383	2.23093
2.286	0.827	3.845	-1.4897	2.43853
2.143	0.762	4.601	-1.34013	2.21173
31.143	3.439	3.491	0.82667	2.54473

2- لأجل التحقق من مدى صحة النموذج المفترض تم رسم الشكل البياني لمعرفة نوع العلاقة واتضح من الشكل المرقم (1) بان العلاقة بين قيم Y مقابل قيم X علاقة خطية ويتضح من الشكل المرقم (2) ان العلاقة خطية بين قيم (Y) مقابل قيم الانحرافات للمتغير المستقل (X) ومن خلالها يتضح بان استخدام الانحرافات لا يوثر في نتائج النموذج ولكن فقط لغرض السهولة توخذ .



شكل (1) شكل الانتشار لقيم Y مقابل قيم X



شكل (2) شكل الانتشار لقيم Y مقابل قيم الانحرافات ل X

3- بعد ان تم افتراض والتحقق من ان الاخطاء في النموذج المفترض تسلك على وفق توزيع القيمة المتطرفة (النوع الاول) وللقيم الكبرى والمتمثلة بدالة الكثافة الاحتمالية التالية :

$$f(z)= \exp\{ - (z+ \gamma) - \exp(- (z+ \gamma)) \} \quad , \quad -\infty < Z < \infty$$

3-1 طريقة المربعات الصغرى: تقديرات OLS هي كالاتي :

$$= - 0.632 , \tilde{\beta}_1 = 2.35 , \tilde{\beta}_0 \tilde{\theta} = 1.10913$$

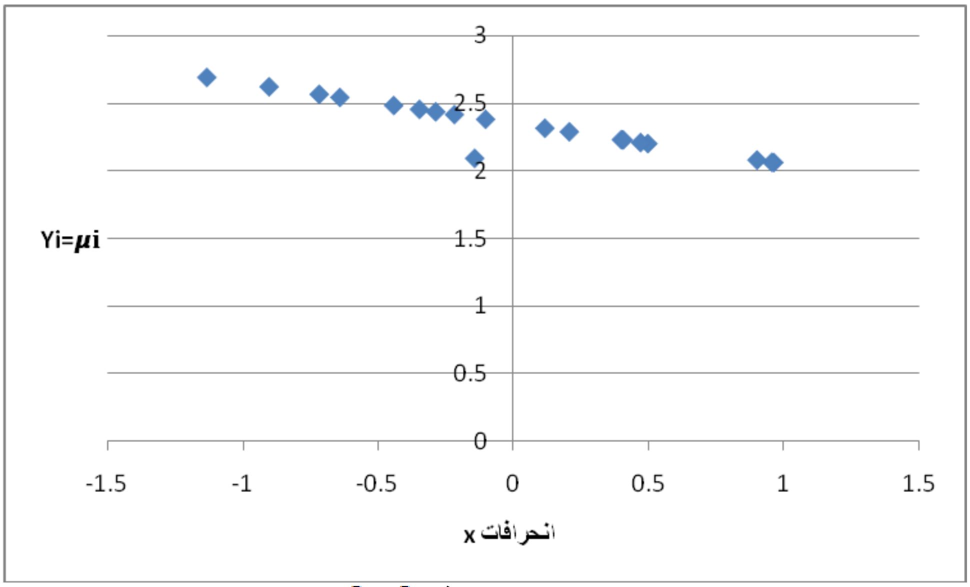
حيث ان $\tilde{\beta}_1$ و $\tilde{\beta}_0$ تمثل معاملات النموذج المفترض بينما $\tilde{\theta}$ تمثل معلمة القياس .

3-2 طريقة الامكان الاعظم:على اعتبار ان طريقة (OLS) كتقدير اولي يستخدم في تقديرات الامكان الاعظم (MLE) والمعتمدة على طريقة فشر والموضحة في الجانب النظري للبحث فتقديرات الامكان الاعظم هي :

$$= 2.33708 , \hat{\beta}_0 \hat{\beta}_1 = - 0.299997 , \hat{\theta} = 1.08178$$

3-3 معادلة الانحدار المقدرة: $- X_i = 2.33708 - 0.299997 (\hat{Y}_i \bar{X})$

ويتضح من خلالها بان اضافة وحدة واحدة من x (كمية كريات الدم البيضاء) تقلل بمقدار (0.299997) من Y (زمن البقاء) للمريض ويتضح من الشكل (6) خط انحدار العينة لقيم $\hat{Y}_i = \hat{\mu}_i$ جدول رقم (1) مقابل قيم X_i (الانحرافات).



شكل (6) خط الانحدار مع القيم التقديرية $\tilde{\beta}_1 , \tilde{\beta}_0$

4- **اختبار الرسم الاحتمالي (P.P)** للقيم المرتبة للاخطاء المشاهدة وذلك لمعرفة فيما اذا كان هناك دليل ضد توزيع القيمة المتطرفة للاخطاء وذلك عن طريق استخدام الصيغة الاتية للاخطاء المشاهدة.

$$R_i = \left(\hat{\theta}^{-1} Y_i \;\; \hat{\beta}_0 - - \hat{\beta}_1 \, x_i \;\right))$$

حيث ان R_i تمثل القيم المرتبة للاخطاء المشاهدة (جدول رقم 1) مقابل قيم K_i حيث ان :

$$K_i = - \mathrm{Log}(\;\; \mathrm{Log}[\, 20 / (\, 20 - i \,)])$$

وتطبيقا لاختبار الرسم الاحتمالي فان:

$$r\,(\, R_{(i)} \,, K_i \,) = - 0.921 \,, \; S_R = \; 0.151759$$

وبمقارنة هذه القيمة مع القيم الجدولية ولمستوى معنوية (0.05,0.01)= α ، n=19 والموجودة في جداول خاصة [9] بهذا الاختبار فان القيم الجدولية هي (0.189,0.271) وعليه فان التوزيع المفترض للاخطاء يعد التوزيع الملائم لبيانات هذه الدراسة وذلك لان القيمة المستخرجة هي اصغر من القيم الجدولية.

5 - **حدود الثقة :حدود الثقة التقريبية %95 ل β_1 هي :**

$$- 1.118761 \;\; \leq \beta_1 \leq 0.5187672$$

ويتضح من هذه النتيجة بانه لا يوجد دليل واضح ضد الفرضية القائلة بان $\beta_1 = 0$. اما حدود الثقة %95 ل θ :

$$0.800965 \;\; \leq \theta \;\; \leq \;\; 1.665804$$

ويتضح من خلالها بان لا يوجد دليل ضد الفرضية القائلة بان $\theta = 1$.

6 – **الكفاية التقريبية :** ظهرت ل $\hat{\beta}_1$ مساوية ل (0.57889) ول $\hat{\theta}$ مساوية ل (0.5057405) ومن خلال هذه القيم وهي اقل من الواحد الصحيح اي ان تقديرات الامكان الاعظم اكثر كفاية من تقديرات المربعات الصغرى الاعتيادية في تقدير معالم النموذج .

ثانيا: ابيضاض الدم اللمفاوي الحاد (ALL)

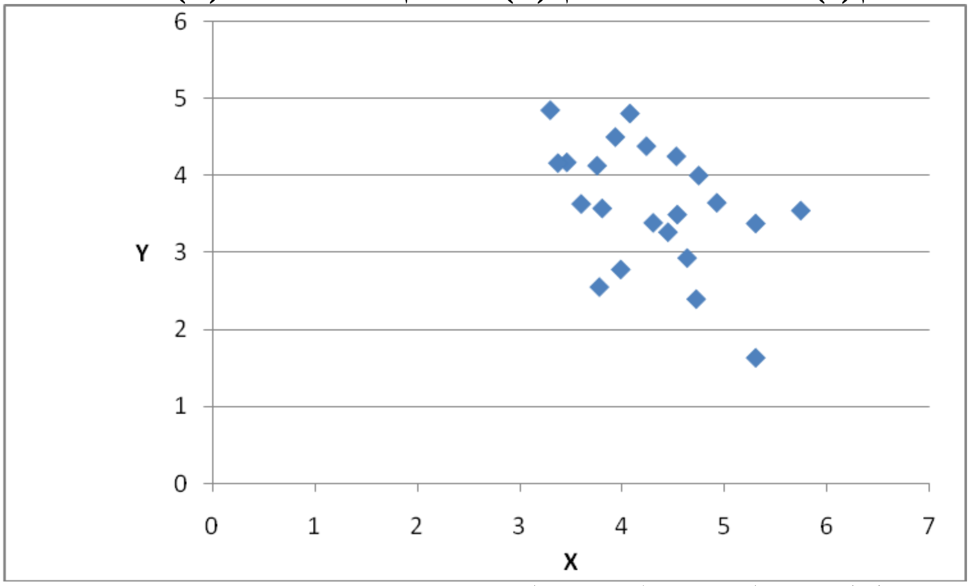
1- مجتمع البحث : لقد شمل مجتمع البحث على 22 مريضا مصابا بمرض (ALL) الذي تتمثل قراءته ببعدين اولهما هو المتغير التابع (Y) الذي يمثل عدد الاسابيع التي يعيشها المريض حتى الوفاة ومن خلال استخدام عدة تحويلات للبيانات وجدنا ان افضل تحويل كان باستخدام التحويل اللوغارتمي والذي حقق الحالة الخاصة بنا وهي توزيع القيمة المتطرفة (حيث ان Y=Log (T) اي ان توزيع القيمة المتطرفة يمثل لوغارتم توزيع زمن البقاء) والمتغير المستقل (X) ويمثل لوغارتم كريات الدم البيضاء للمريض والمبينة في الجدول رقم (2) وباتخاذ العلاقة مابين المتغيرين لشكلا خطيا بسيطا على وفق نموذج الانحدار الاتي :

$$Y_i = +\beta_0 \;\; \beta_1 \;\; (\;\, + \;\;\, - \bar{X} \,) \, X_i \;\; \theta Z_i \quad , \mathrm{i} = 1,2,\ldots,22 \qquad (73)$$

جدول رقم (2) بيانات ازمنة البقاء للمرضى المصابين بمرض ALL
(Y يمثل لوغارتم توزيع زمن البقاء ، X يمثل لوغارتم كريات الدم البيضاء)

t_i	Y_i	X_i	$\hat{\mu}_i$	R_i
32.857	3.492	4.538	3.43777	0.07356
18.714	2.929	4.633	3.3728	-0.60201
79.571	4.377	4.238	3.64294	0.99575
16.143	2.781	3.987	3.81459	-1.40207
64.571	4.168	3.462	4.17364	-0.00765
69.857	4.246	4.528	3.44461	1.08709
34.571	3.543	5.74	2.61573	1.25784
5.143	1.638	5.301	2.91596	-1.73355
26.143	3.264	4.447	3.5	-0.32014
54.429	3.9969	4.747	3.29484	0.95235
12.857	2.554	3.778	3.95753	-1.90388
10.857	3.385	4.303	3.59848	-1.64609
121.571	4.8	4.076	3.75373	1.41927
29.286	3.377	5.301	2.91596	0.6254
37.714	3.63	3.602	4.07789	-0.60757
64	4.159	3.375	4.23314	-0.10057
38.286	3.645	4.923	3.17447	0.63827
61.857	4.125	3.756	3.97257	0.20676
89.571	4.495	3.934	3.85084	0.8738
11	2.398	4.722	3.31193	-1.23975
126.857	4.843	3.301	4.28375	0.75863
35.571	3.572	3.806	3.93838	-0.49699

2- لاجل التحقق من مدى صحة النموذج المفترض تم رسم الشكل البياني لمعرفة نوع العلاقة واتضح من الشكل المرقم (2) ان العلاقة خطية بين قيم (Y) مقابل قيم المتغير المستقل (X) .



شكل (3) شكل الانتشار لقيم Y مقابل قيم X

3- بعد ان تم افتراض ان الاخطاء في النموذج المفترض تسلك على وفق توزيع القيمة المتطرفة (النوع الاول) وللقيم الكبرى والمتمثلة بدالة الكثافة الاحتمالية التالية :

$$f(z)= \exp\{ - (z+ \gamma) - \exp(- (z+ \gamma)) \} \quad , \quad -\infty < Z < \infty$$

3-**1طريقة المربعات الصغرى:تقديرات OLSهي كالاتي** :

$$\tilde{\beta}_0 = 3.56 \text{ , } = - 0.553 \text{ , } \tilde{\beta}_1 \ \tilde{\theta} = 0.608861$$

حيث ان $\tilde{\beta}_0$ و $\tilde{\beta}_1$ تمثل معاملات النموذج المفترض بينما $\tilde{\theta}$ تمثل معلمة القياس .

2-3 طريقة الامكان الاعظم:على اعتبار ان طريقة (OLS) كتقدير اولي يستخدم في تقديرات الامكان الاعظم (MLE) والمعتمدة على طريقة فشر والموضحة في الجانب النظري للبحث فتقديرات الامكان الاعظم هي :

$$\hat{\beta}_0 = 3.60371 \text{ , } \hat{\beta}_1 = - 0.683894 \text{ , } \hat{\theta} = 0.737193$$

3-3 معادلة الانحدار المقدرة: $= 3.60371 - 0.683894 \ (\hat{Y}_i X_i - \bar{X})$

ويتضح من خلالها بان اضافة وحدة واحدة من x (كمية كريات الدم البيضاء) تقلل بمقدار (0.683894) من Y (زمن البقاء) للمريض .

4- اختبار الرسم الاحتمالي (P.P) للقيم المرتبة للاخطاء المشاهدة وذلك لمعرفة فيما اذا كان هناك دليل ضد توزيع القيمة المتطرفة للاخطاء وذلك عن طريق استخدام الصيغة التالية للاخطاء المشاهدة.

$$R_i = \ (\hat{\theta}^{-1} Y_i \ \hat{\beta}_0 - - \hat{\beta}_1 \ x_i \ ))$$

حيث ان R_i تمثل القيم المرتبة للاخطاء المشاهدة (جدول رقم2) مقابل قيم K_i حيث ان :

$$K_i = - \text{Log}(\ \text{Log}[\ 23 / (\ 23 - i\ )])$$

وتطبيقا لاختبار الرسم الاحتمالي فان:

$$r\ (,R_{(i)} K_i) = - 0.982 \ \ ,S_R = \ 0.035676$$

وبمقارنة هذه القيمة مع القيم الجدولية ولمستوى معنوية α =(0.05,0.01) ، n=22 والموجودة في جداول خاص[9] بهذا الاختبار فان القيم الجدولية هي (0.189,0.271) وعليه فان التوزيع المفترض للاخطاء يعد التوزيع الملائم لبيانات هذه الدراسة وذلك لان القيمة المستخرجة هي اصغر من القيم الجدولية.

5 - **حدود الثقة** :حدود الثقة التقريبية 95% ل β_1 هي :

$$- 1.1870473 \ \leq \beta_1 \leq - 0.1807407$$

ويتضح من هذه النتيجة بانه لايوجد دليل واضح ضد الفرضية القائلة بان 0 = β_1. اما حدود الثقة 95% ل θ :

$$0.5560297 \ \leq \theta \ \leq \ 1.0934594$$

ويتضح من خلالها بان لايوجد دليل ضد الفرضية القائلة بان 1= θ .

6 – **الكفاية التقريبية** : ظهرت ل $\hat{\beta}_1$ مساوية ل (0.891) ول $\hat{\theta}$ مساوية ل (0.810)ومن خلال هذه القيم وهي اقل من الواحد الصحيح اي ان تقديرات الامكان الاعظم اكثر كفاية من تقديرات المربعات الصغرى الاعتيادية في تقدير معالم النموذج .

4-2 **التطبيق الثاني** : درجات الحرارة العظمىٔ : للطقس تاثير عميق في مختلف اوجه النشاط الانساني، وقد تعرض العراق لموجات من الحر الشديد بالغة الاثر خلال السنوات الماضية في مختلف محافظات القطر وعلى الاخص في محافظة بغداد وتعزى هذه الزيادة في درجات الحرارة الى تظافر جملة عوامل منها الظروف التضاريسية المحلية وتاثيراتها في اعاقة حركة الكتل الهوائية وعلى توزيعات درجه الحرارة وتاثيرات العوارض الجبلية والوديان وطبيعة تكوين سطح التربة ووجود المسطحات المائية ومصادر المياه الموزعة في مختلف المناطق اضافة الى بعض التاثيرات الناتجة عن الفعاليات البشرية . ان صفة الشمول لهذا التاثيرات كانت حافزا لاعطاء هذا الموضوع اهتماما كبيرا وذلك للاحتمالية الكبيرة في تكرار مثل هذه الاحداث المميزة والخطيرة في المستقبل والتي قد تعرض العراق لاضرار مختلفة في الامكان تلافيها وذلك من خلال تهيئة الوسائل اللازمة والسبل المستخدمة لاعداد التنبؤات الجوية الصائبة لدرء الاضرار والاخطار ومن هنا جاءت اهمية هذا التطبيق .

4-2-1 درجة الحرارة العظمى: تعرف الحرارة بانها نوع من انواع الطاقة وهي كمية الطاقة المنتقلة الى جسم لترفع من درجة حرارته متجهة من الدرجات الحرارية العالية الى الدرجات الحرارية المنخفضة والمصدر الرئيسي لها في الغلاف الجوي هو الشمس وهي احد عناصر الطقس واهمها (عناصر الطقس هي الاشعاع الشمسي والحرارة والضغط والرياح والرطوبة) ويوجد نظامان لقياسها الاول النظام المئوي (السيليسيوس) المتبع في الهيئة العامة للانواء الجوية وفيه نقطة الغليان هي درجة المئة مئوي ونقطة الجليد درجة الصفر المئوي اما النظام الثاني فهو الفهرنهايت. اما درجة الحرارة العظمى فهي اعلى درجة حرارة يبلغها الهواء السطحي خلال اليوم الواحد وتحدث بين الساعة الثانية والثالثة بعد الظهر بينما درجة الحرارة الصغرى فهي اوطأ درجة حرارة يبلغها الهواء السطحي خلال اليوم الواحد وتحدث قبل شروق الشمس بدقائق وهناك عوامل كثيرة توثر في درجة الحرارة يتعلق البعض منها بطبيعة مناخ المنطقة من حيث خطوط العرض، توزيع اليابسة والتضاريس والتيارات المحيطة والبعد عن الساحل والبعض الاخر خاضع الى تاثيرات عملية متعلقة بنشاطات الانسان التي لاحصر لها. وقد تم التاكيد على عنصرين مهمين من عناصر الطقس التي تتاثر بتلك العوامل وتوثر بدورها في درجة الحرارة وهي شدة الاشعاع الشمسي والغطاء الغيمي، حيث ان كمية الاشعاع الشمسي الواصل الى كوكبنا يتوقف على فصول السنة فتختلف في فصل الصيف عنها في فصل الشتاء بالاضافة الى الظروف الجوية المتغيرة دائما كوجود الغيوم او الغبار والشوائب الاخرى اخرى وتقدر كمية الاشعاع بالسعرات لكل سم^2 وبالملي واط لكل سم^2 (1 سعرة /سم^2 =1.163 ملي واط /سم^2) اي ان كل سنتمتر مربع من سطح الارض يكتسب طاقة اشعاع تساوي 1 سعرة او ما مقداره 1.163 ملي واط . اما كمية التغييم (Clodness) فتقاس اما ب (OKTA) او (8) والمقصود بها ان السماء تقسم الى ثمان اثمان فاذا كانت مغطاة بالغيوم تكون (8 / 4) وهكذا والحساب الاخر هو ان اسماء اما صافية او غائمة او غائمة جزئيا ويتم تفسير ذلك على وفق الاتي : صافية يعني كمية الغيوم اقل من 1.6 ، غائمة يعني كمية الغيوم اكبر من 6.5 وغائمة جزئيا اي كمية الغيوم بين 1.6 الى 6.5 .

4-2-2 اسلوب جمع البيانات : تم جمع البيانات الخاصة بالمعدلات الشهرية لكل من درجات الحرارة العظمى والاشعاع الشمسي وكمية الغيوم وخلال السنة للفترة من (1971 - 1980)(تطبيق احدث للحالة [2]) والخاصة بمدينة بغداد (تقع في المنطقة الوسطى من القطر على خط طول '41 °44 شرقا وخط عرض '14 °32 شمالا ارتفاعها عن مستوى سطح البحر بمقدار 31.7 متر ولقد تم الحصول على هذه البيانات من الهيئة العامة للانواء الجوية– قسم المناخ ولقد اخذت بعض النقاط الاساسية عند اخذ العينة الخاصة بالبحث وهي :

- ممكن اخذ البيانات الخاصة بدرجات الحرارة العظمى بشكل يومي او شهري او فصلي او سنوي وذلك على وفق مايلائم النموذج المعطى وقد تم اخذ البيانات بشكل فصلي اي تقسيم السنة الى اربعة فصول (فصل$_1$، فصل$_2$ ، فصل$_3$ ، فصل$_4$) وعلى هذا الاساس فان العينة التي اخذت للفترة من (1971-1980) ستكون مؤلفة من 40 مشاهدة والجدول رقم (3) يوضح الصيغة النهائية للبيانات التي استخدمت في التحليل .
- تم تحويل المتغيرات التوضيحية (المستقلة) الى قيمها المعيارية وذلك نتيجة لاختلاف وحدات القياس الخاصة بها .

جدول رقم (3) الصيغة النهائية لبيانات درجات الحرارة العظمى، شدة الاشعاع الشمسي والغطاء الغيمي التي استخدمت في التحليل

السنوات	الفصول	درجات الحرارة العظمى	شدة الاشعاع الشمسي	الغطاء الغيمي
1971	1	60.9	1311.9	9.9
	2	104.4	2065.6	7.1
	3	127.3	2120.2	0.4
	4	70.8	1154.4	7.9
1972	1	50.1	1175.2	10.6
	2	104.2	1987	8.8
	3	126.6	2074.9	1.6
	4	73.4	1148.6	8
1973	1	61.2	1370.7	7.6
	2	105.4	2189	4.7
	3	127.8	2048	0.6
	4	74.3	1190.7	6.6
1974	1	49.3	1120.3	12.8
	2	104	2144	5.9
	3	124.2	1971.8	0.9
	4	75.9	1118.6	7.5
1975	1	54	1239.5	8.5
	2	107.7	1887.4	7.2
	3	127.1	2004.9	1.3
	4	71.1	1153.2	7.2
1976	1	52.8	1041.9	10.2
	2	104.1	1764.4	7.5
	3	122	1822.7	0.6
	4	78.4	953.1	9.1
1977	1	59.4	1059.5	10
	2	106.5	2079.8	6.5
	3	128.2	2091.4	0.1
	4	70.1	1090.2	8.1
1978	1	62.5	1261.5	8.8
	2	107.3	2094.9	5.7
	3	125.5	1982.9	0.3
	4	73.9	1077	7.8
1979	1	62.7	1154.8	10
	2	108.4	1947.9	7.9
	3	129	1921	2
	4	77.3	1038.1	8.9
1980	1	55.5	1147.9	11.6
	2	107.6	2027.2	6.1
	3	125.7	1994.7	0.9
	4	73.2	1061.8	8.9

3-2-4 التحليل الاحصائي

تقدير معاملات ومعلمة القياس لأنموذج انحدار القيمة المتطرفة الخطي

(النوع الاول) وللقيم الكبرى مع جانب تطبيقي

1- **عينة البحث : لقد شملت عينة البحث 40 مشاهدة تمثل المجاميع الفصلية لثلاثة متغيرات اولها هو المتغير التابع Y_i الذي يمثل المجاميع الفصلية لدرجات الحرارة العظمى والمتغير المستقل الاول X_{i1} ويمثل المجاميع الفصلية لكمية الاشعاع الشمسي والمتغير المستقل الثاني X_{i2} ويمثل المجاميع الفصلية لكمية الغيوم .وحيث ان تحليل الانحدار يتطلب بعض الاجراءات وخاصة في اختيار النموذج المقترح ومن اجل ذلك تم اقتراح عدة نماذج واجراء الاختبار عليها وتبين بان افضل نموذج يلائم بياناتنا هو النموذج الاسي التالي :**

$$Y_i = \exp(\beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2}) \,.\, \theta Z_i \quad (i=1,2,\dots,40)$$

وذلك من خلال معامل التحديد ومتوسط مربع الخطا حيث ان : $R^2 = 87.9\%$ $S^2 = 0.0130$

ويتم تحويل النموذج الى الصيغة الخطية باخذ اللوغارتم الطبيعي لطرفيه فينتج :

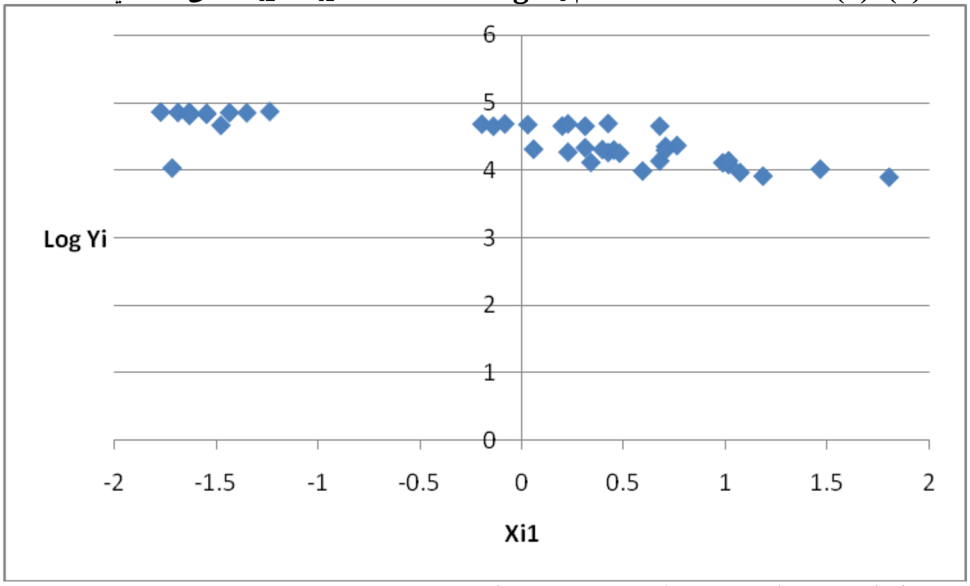
$$\text{Log } Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \text{Log}(\theta)\, Z_i$$

والجدول رقم (4) يوضح بيانات درجات الحرارة العظمى ل40 فصل مع القيم المعيارية للمتغيرات المستقلة وبعض المقاييس المهمة التي يشار اليها خلال الفقرات القادمة.

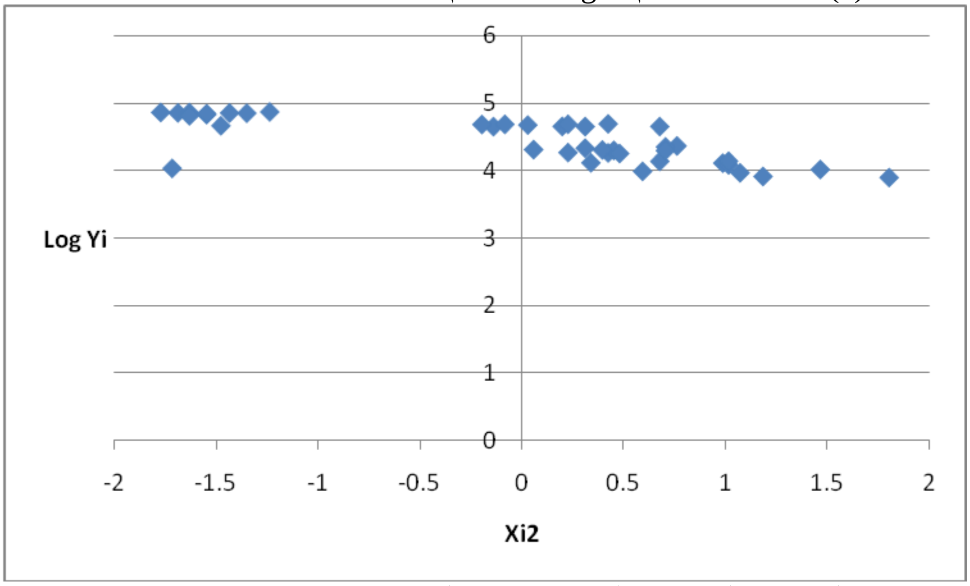
جدول رقم (4) درجات الحرارة ل40 فصل مع القيم المعيارية للمتغيرات المستقلة وبعض المقاييس المهمة

Log Y_i	X_{i1}	X_{i2}	R_i	$\hat{\mu}_i$
4.10923	-0.59595	0.98495	-0.71389	4.18593
4.64823	1.09702	0.19643	-0.00302	4.64855
4.84655	1.21966	-1.69039	-1.31821	4.98817
4.25986	-0.94973	0.42172	0.45432	4.21105
3.91402	-0.90301	1.18208	-1.66564	4.09297
4.64631	0.92046	0.67517	1.04563	4.53397
4.84103	1.11791	-1.35246	-0.6583	4.91176
4.29592	-0.96276	0.44988	0.85754	4.20379
4.11415	-0.46388	0.33723	-1.9169	4.32009
4.65776	1.3742	-1.47945	-1.47072	4.81577
4.85047	1.05748	-1.63407	-0.89915	4.94707
4.30811	-0.86819	0.05562	0.18511	4.28822
3.89792	-1.02633	1.80163	-0.6265	3.96523
4.64439	1.27312	-0.14151	-0.88517	4.73949
4.82189	0.88632	-1.54959	-0.72234	4.8995
4.32942	-1.03015	0.30907	1.07255	4.21419
3.98898	-0.75858	0.59069	-2.1514	4.22012
4.67935	0.69674	0.22459	1.05827	4.56565
4.84497	0.96067	-1.43694	-0.46728	4.89518
4.26409	-0.95243	0.22459	0.1916	4.2435
3.96651	-1.20243	1.06943	-0.80807	4.05333
4.64535	0.42046	0.30907	1.37573	4.49755
4.80402	0.55141	-1.63407	-0.41131	4.84821
4.36182	-1.40189	0.75966	2.75175	4.06619
4.08429	-1.1629	1.01311	0.12866	4.07047
4.66815	1.12891	0.02746	-0.13876	4.68305
4.85359	1.15497	-1.77488	-1.26657	4.98967
4.24992	-1.09394	0.47804	0.71173	4.17346
4.13517	-0.70916	0.67517	-0.74904	4.21564
4.67563	1.16283	-0.19783	-0.48159	4.79737
4.03231	0.91126	-1.71855	-0.93387	4.93264
4.30271	-1.12359	0.39356	1.12545	4.1818
4.13836	-0.94883	1.01311	0.2427	4.11229
4.68583	0.83264	0.42172	1.17846	4.55922
4.85981	0.77221	-1.23981	0.32045	4.82538
4.34769	-1.21097	0.70333	2.18537	4.11291
4.01638	-0.96433	1.46369	-0.16285	4.03388
4.67842	1.01076	-0.08519	-0.0037	4.67822
4.8339	0.93776	-1.54959	-0.70412	4.90955
4.2932	-1.15773	0.70333	1.58131	4.12331

2- لاجل التحقق من مدى صحة النموذج المفترض فقد تم رسم الشكل البياني لمعرفة نوع العلاقة ويتضح من الشكل (3)و(4) بان العلاقة خطية بين قيم $Log\ Y_i$ مقابل كل من X_{i1} وX_{i2} على التوالي.



شكل(4) شكل الانتشار لقيم LogYi مقابل قيم Xi1



شكل (5) شكل الانتشار لقيم LogYi مقابل قيم Xi2)

3- **اختبار فرضية تجانس التباين** :تم استخدام اختبار بارتليت لاختبار فرضية تجانس التباين وظهرت القيمة المحتسبة (0.0180972) وبمقارنتها بقيمة χ^2 الجدولية عند مستوى معنوية (0.05) ودرجة حرية 1 والتي تساوي (3.841) نجد ان القيمة المحتسبة اصغر من قيمتها الجدولية وعليه لا يوجد دليل ضد فرضية العدم التي تنص على ثبات تجانس تباين الخطا .

4- بعد ان تم افتراض ان الاخطاء في النموذج المفترض تسلك على وفق توزيع القيمة المتطرفة (النوع الاول) وللقيم الكبرى والمتمثلة بدالة الكثافة الاحتمالية التالية:

$$f(z)= \exp\{ - (z+ \gamma) - \exp(- (z+ \gamma)) \} \quad , \quad -\infty < Z < \infty$$

4-1 طريقة المربعات الصغرى: تقديرات OLS هي كالاتي :

$$\tilde{\beta}_0 =0.169 \tilde{\beta}_1= 4.46 \ , \ \tilde{\beta}_2 = -0.149,, \tilde{\theta} = 0.0888498$$

حيث ان $\tilde{\beta}_0$ ، $\tilde{\beta}_1$ و $\tilde{\beta}_2$ تمثل معاملات النموذج المفترض بينما $\tilde{\theta}$ تمثل معلمة القياس .

3-2 **طريقة الامكان الاعظم:**على اعتبار ان طريقة (OLS) كتقدير اولي يستخدم في تقديرات الامكان الاعظم (MLE) والمعتمدة على طريقة فشر والموضحة في الجانب النظري للبحث فتقديرات الامكان الاعظم هي :

$$\hat{\beta}_0 =4.46712 \ , \hat{\beta}_1 =0.195341 \ , \hat{\beta}_2 = -0.167296 \ , \hat{\theta} = 0.107436$$

3-3 **معادلة الانحدار المقدرة:**

$$\hat{Y}_i = 4.46712+ 0.195341 \ X_{i1}- 0.167296 \ X_{i2}$$

ويتضح من خلالها بان اضافة وحدة واحدة من X_{i1} (كمية كريات الاشعاع الشمسي) تزيد بمقدار (0.195) من Y (درجة الحرارة العظمى) وان اضافة وحدة واحدة من X_{i2} (كمية الغيوم) تقلل بمقدار (0.167) من قيمة Y (درجة الحرارة العظمى) .

5- **اختبار الرسم الاحتمالي** (P.P) للقيم المرتبة للاخطاء المشاهدة وذلك لمعرفة فيما اذا كان هناك دليل ضد توزيع القيمة المتطرفة للاخطاء وذلك عن طريق استخدام الصيغة التالية للاخطاء المشاهدة.

$$R_i = (\hat{\theta}^{-1} \ Y_i - \hat{\beta}_0 - \hat{\beta}_1 \ x_i - \hat{\beta}_2 X_{i2} \ ))$$

حيث ان R_i تمثل القيم المرتبة للاخطاء المشاهدة (جدول رقم) مقابل قيم K_i حيث ان :

$$K_i = - \text{Log}(\ \text{Log}[\ 41 \ / \ (\ 41 - i \ )])$$

وتطبيقا لاختبار الرسم الاحتمالي فان:

$$r \ (\ R_{(i)} \ , K_i \ ) = - 0.952, \ \ S_R = \ 0.0940737$$

وبمقارنة هذه القيمة مع القيم الجدولية ولمستوى معنوية α =(0.05,0.01) ، n=40 والموجودة في جداول خاصة بهذا الاختبار فان القيم الجدولية هي (0.147,0.225) وعليه فان التوزيع المفترض للاخطاء يعد التوزيع الملائم لبيانات هذه الدراسة وذلك لان القيمة المستخرجة هي اصغر من القيم الجدولية.

5 - **حدود الثقة :** حدود الثقة التقريبية 95% ل β_1 هي :

$$0.161027 \leq \beta_1 \leq 0.229655$$

ويتضح من هذه النتيجة بان هناك دليل واضح ضد الفرضية القائلة بان $\beta_1 = 0$. اما حدود الثقة التقريبية 95% ل β_2 باعتبارها تمثل مقدار التغير في المتغير المعتمد المتاتي من التغير بوحدة واحدة مثلا في المتغير المستقل الثاني الذي يمثل كمية الغيوم هي :

$$-0.2016099 \leq \beta_2 \leq -0.132982$$

ويتضح من هذه النتيجة بان هناك دليل واضح ضد الفرضية القائلة بان $\beta_2 = 0$ اما حدود الثقة 95% ل θ :

$$0.086528 \leq \theta \leq 0.1416672$$

ويتضح من خلالها بانه يوجد دليل ضد الفرضية القائلة بان $\theta = 1$.

6 – الكفاية التقريبية: ظهرت ل $\hat{\beta}_j$ حيث j=1,2 مساوية ل (0.888683) ول $\hat{\theta}$ مساوية ل (0.0.8082949) ومن خلال هذه القيم وهي اقل من الواحد الصحيح اي ان تقديرات الامكان الاعظم اكثر كفاية من تقديرات المربعات الصغرى الاعتيادية في تقدير معالم النموذج .

5- الاستنتاجات والتوصيات

1-5 الاستنتاجات : من خلال سلوك الاخطاء على وفق توزيع القيمة المتطرفة (النوع الاول) وللقيم الكبرى ومن خلال الجانب النظري والتطبيقي لابد من الاشارة الى بعض النقاط التالية:

اولا:الجانب الطبي : ازمنة البقاء – مرض ابيضاض الدم

1- ان نموذج الانحدار البسيط المقترح كان هو النموذج المناسب وكذلك من خلال رسم قيم Y مقابل قيم X تبين ان هناك علاقة خطية بين المتغيرين .

2- من خلال استخدام اختبار الرسم الاحتمالي تبين انه ليس هناك اي دليل ضد التوزيع المفترض للاخطاء وهو توزيع القيمة المتطرفة (النوع الاول) وللقيم الكبرى .

3- من خلال استخدام صيغ حدود الثقة ل β_1 و θ اتضح :

- ابيضاض الدم النقيائي الحاد (AML) :انه لا يوجد دليل واضح ضد الفرضية القائلة بان $\beta_1 = 0$ وهذا يعني ان المتغير المستقل (كمية كريات الدم البيضاء) لا توثر في زمن البقاء للمريض (وقد اكد الاطباء هذه الحالة بانه كمية كريات الدم البيضاء صحيح يعد عاملا من العوامل التي تؤثر في مرض اللوكيميا ولكن تاثيرها يختلف حسب انواع المرض وكذلك اكدوا ان زيادة كمية كريات الدم البيضاء لايعني دائما بان تاثيرها يكون سيئا في المرض لانه في بعض الحالات تكون اصلا كمية هذه الكريات في دم المرضى هي 400 لكل ميكرولتر من الدم فقط وهذه النسبة قليلة جدا بالمقارنة مع النسبة الطبيعية التي تتراوح بين (4000 - 11000) كرية بيضاء ولكن في حالة زيادة كمية كريات الدم البيضاء وفي الوقت نفسه كمية الكريات المهلكة فان الحالة ستكون هنا خطرة جدا وتودي الى الوفاة وهذا ما لوحظ من خلال اخر فحص للدم اجري للمرضى الذين توفوا .

ولθ اتضح انه لا يوجد دليل ضد الفرضية القائلة بان $\theta=1$.

-ابيضاض الدم اللمفاوي الحاد (ALL) : ان هناك دليلا واضحا ضد الفرضية القائلة بان $\beta_1 = 0$ اي ان المتغير X الذي يمثل كمية كريات الدم البيضاء له تاثير عكسي في المتغير Y الذي يمثل زمن البقاء للمريض وهذا ما يؤكد التوضيح الذي ذكر في حالة ابيضاض الدم النيقائي الحاد AML ولθ اتضح بانه لا يوجد دليل واضح ضد الفرضية القائلة بان $\theta=1$.

4- من خلال قيم الكفاءة التقريبية لكل من β_1 ، θ اتضح بان تقديرات الامكان الاعظم اكثر كفاءة من تقديرات المربعات الصغرى الاعتيادية في تقدير معالم النموذج .

ثانيا: جانب المناخ (درجات الحرارة العظمى)

1- ان النموذج الاسي المقترح كان هو النموذج المناسب (من خلال معامل التحديد ومتوسط مربع الخطا) وكذلك من خلال رسم $Log\ Y_i$ مقابل X_i(i=1,2) تبين ان هناك علاقة خطية بين المتغيرين .

2- من خلال اختبار بارتليت لاختبار فرضية تجانس التباين ظهر بانه لا يوجد دليل ضد الفرضية التي تنص على ثبات تجانس تباين الخطا.

3- من خلال استخدام اختبار الرسم الاحتمالي تبين بانه لا يوجد اي دليل ضد التوزيع المفترض للاخطاء وهو توزيع القيمة المتطرفة (النوع الاول) وللقيم الكبرى.

4- من خلال معادلة الانحدار المقدرة اتضح بان هناك علاقة طردية بين درجة الحرارة العظمى وشدة الاشعاع الشمسي وعلاقة عكسية بين درجة الحرارة العظمى وكمية الغيوم ومن خلال استخراج القيم المقدرة $\widehat{\mu}_{i}$ ومقارنتها مع القيم الحقيقية اتضح بان هناك فرقا قليلا جدا مما يدل على صحة النموذج المقترح.

5- من خلال استخدام حدود الثقة ل β_1, β_2، θ اتضح بان هناك دليلا واضحا ضد الفرضية القائلة بان $\beta_1 = \beta_2 = 0$ و$\theta=1$.

6- من خلال ملاحظة نتائج تطبيق الكفاءة التقريبية اتضح بان تقديرات الامكان الاعظم اكثر كفاءة في تقدير معالم النموذج من طريقة المربعات الصغرى .

2-5 التوصيات :

اولا: الجانب النظري : من خلال ما توصل اليه الباحث عادل احمد هدو[9] [10] من تحديد الصيغ المتعلقة بالتحيز لمعاملات ومعلمة القياس لنموذج انحدار القيمة المتطرفة (النوع الاول) وللقيم الصغرى نقترح بالاخذ بهذه الصيغ للوصول الى تحديد الصيغ في حالة القيم الكبرى للتوزيع نفسه .

ثانيا : الجانب التطبيقي:

1- الجانب الطبي :ان تعمل دراسة مماثلة تاخذ العوامل الاساسية جميعها التي توثر في المرض كنموذج انحدار متعدد وهذه العوامل ذكرت اغلبها ضمن فقرة (جمع البيانات الخاصة بالبحث) او ان تعمل الدراسة بشكل مباشر على المرضى وان تراعى مسالة الوقت اي مثلا تحدد ثلاث او اربع سنوات وهذه الحالة مفضلة اكثر للحصول على اغلبية المعلومات المتعلقة بالمرض من المريض مباشرة لتفادي اي خطا ممكن ان يقع فيه الباحث .

2- جانب المناخ (درجات الحرارة العظمى): ان تعمل دراسة مماثلة تدرس اولا العوامل التي توثر في الاشعاع الشمسي مثل الغبار والمواد العالقة الاخرى والتي تؤثر بالتالي في درجة الحرارة وكل هذه العوامل مقاسة بالانواء ، ومن جانب اخر ومن ضمن العوامل التي توثر في درجة الحرارة ان يؤخذ عامل انعكاسية السطح (Albedo) فطبيعة الارض سواء كانت مزروعة ام لا رطبة ام جافة رملية ام صخرية وكذلك لونها كل هذا يؤثر في درجة الحرارة وفي هذه الحالة تعمل دراسة ميدانية لتحديد هذا العامل للمنطقة المراد وضع الدراسة حولها وهذه العوامل يمكن ان تقاس من قبل الانواء الجوية.

المصادر

1- العزاوي، فاطمة جاسم محمد (تقدير معاملات ومعلمة القياس لنموذج انحدار القيمة المتطرفة الخطي (النوع الأول) وللقيم الكبرى)، رسالة ماجستير، كلية الإدارة والاقتصاد، جامعة بغداد، 1995.

2-العزاوي، فاطمة جاسم محمد (أهمية استخدام نموذج انحدار القيمة المتطرفة مع جانب تطبيقي)، بحث منشور، (مجلة وقائع المؤتمر السنوي الخامس للتخطيط والتنمية)، هيئة التخطيط، 2001.

3-الطواش ، ميسون سالم مجيد "دراسة احصائية حول توزيع ويبل وتوزيع القيمة المتطرفة مع تطبيقات عملية "- رسالة ماجستير – كلية الادارة والاقتصاد –جامعة بغداد -1990 .

4-العباسي، جمال ناجي صالح "احصاء القيم القصوى مع تطبيق لتقدير مخاطر الزلازل في العراق"-رسالة ماجستير- كلية الادارة والاقتصاد-جامعة بغداد-1984.

5-صالح، معتز محمد"موجة الحر التي اثرت على القطر العراقي في شهر تموز1987"،20،Aت 1982، الهيئة العامة للانواء الجوية العراقية-بغداد.

6-هدو، عادل احمد "مقارنة طريقة المربعات الصغرى (LSE) وطريقة الامكان الاعظم (MLS) باستخدام الكفاءة التقريبية (Asymptotic Efficiency) ،مؤتمر جامعة القادسية 1989 ، مجلة كلية الادارة والاقتصاد-الجامعة المستنصرية.

7-Galambos ,J.(1978)."The asymptotic theory of extreme order statistics",john wiley &Sons,N.Y.

8-Gumbel ,E.J.(1954),"statistical theory of extreme values and some practical applications",Nati.Bureau of stand.,Appl.Math.ser.No.33.

9-Haddow ,A.A.(1986)."some statistical problems connected with extreme value regression ",ph.D.thesis,university of Brunel,U.K.

10-Haddow , A.A.and Young,D.H.(1986)."moment properties of estimators for a type 1 extreme-value regression model",communications in statistics,vol.15,No.8,2527-2539.

11-Johnson,N.L.and Kotz,S.(1970)."distribution in statistics,continuous univariate distributions-1",Boston,Houghton Mifflin company.

12-Lawless,J.F.(1982),"Statistical models and methods for lifetime data ".New York,wiley.

13-Leblanc,Michael and Moon,James (2006)."Extreme Regression ",Biostatistics,vol.7,issue 1,pp71-84.

14-Mann,N.R.(1968),"point and interval estimation procedures for the twe-parameter weibull or extreme value distributions",Technometrics,10,231-256.

15-Silvey,S.D.(1978)."statistical inference",John wiley & Sons i,New York

16-Stephens,M.A.(1977),"Goodness of fit for the extreme value distribution ",Biometrika,64,583-588.

17-Thomas , Michael (2011),"Regression class : Linear and non-linear regression",Jave scientific library ,13 november.