دراسة مقارنة للانحدار الحصين " دراسة باستخدام المحاكاة "

أ.م.د. خلود يوسف حمو ساندي قيس
كلية الإدارة والاقتصاد/جامعة بغداد

Estimation Study "By using Simulation Study " Robust Regression

Assis. Proof. Dr. Khlood Yousef H. Sandy Qais
Admin.&Econ. College/University of Baghdad

تاريخ استلام البحث 2016/11/29 تاريخ قبول النشر 2018/5/9

المستخلص:

فكرة البحث تتلخص في إيجاد مقدرات حصينة لأنموذج الانحدار الخطي في حالة وجود القيم الشاذة والتي تجعل من استخدام الطرائق التقليدية كطريقة OLS حساسة تجاه الشواذ وغير مقاومة تجاه الشواذ، ومن طرائق التقدير التي استخدمت الطرائق التقليدية OLS والطرائق الحصينة M وMM و R و S وLTS و LMS ، GM وقد وجد ان طرائق OLS و M و MM تملك نقطة انهيار صفر على عكس بقية الطرائق الحصينة وهي R و S و LTS و LMS و GM التي تملك نقطة انهيار 0.5 وهي بذلك مقاومة تجاه الشواذ وتملك تبايناً غير عالٍ وأقل MSE لذلك يفضل استخدامها في حالة الشواذ ولأحجام العينات كافة.

الكلمات المفتاحية: انحدار حصين.

ABSTRACT:

The idea of the research is to find robust estimator for linear regression model in the case of presence of outliers value that make traditional methods like OLS sensitive towards outliers and non resistance, an estimation methods are used traditional method OLS and robust method M,MM,S,R,LTS,LMS ,GM we find that OLS,M,MM methods has breakdown point zero were other robust methods has 0.5 breakdown point which is R,S,LTS,LMS,GM and there variance not high and minimum MSE were the best to use them in the stat of outliers for all size of samples.

Keywords: M method, MM, R method ,S Method, robust regression methods.

المقدمة:

طرائق تقديرات معلمات الانحدار تعول على افتراضات تكون شرعية لكن عند عدم تحقق الافتراضات بسبب وجود القيم الشاذة Outliers المقدر أو الإجراء الإحصائي يكون غير حصين تجاه الافتراضات ولتحقيق التقدير تستخدم طرائق حصينة مقاومة تجاه الشواذ ضد الطرائق التقليدية التي تكون حساسة تجاه الشواذ ومن مقدرات التقدير للمعلمات التي تستخدم على نطاق واسع مقدرات M حيث تعالج الخلل في متجه الأخطاء العشوائية بطريقة المربعات الصغرى الاعتيادية دون معالجة الخلل في المتغيرات التوضيحية وطورت الى مقدرات S ومقدرات R.

الجانب النظري:

مقدر MM-Estimation[1],[7],[6]

تقديرات المربعات الصغرى تعمل بسوء عند توزيع الخطأ فلا يكون طبيعياً وبخاصة عندما تكون الاخطاء ثقيلة ولإزالة تأثير المشاهدات الشاذة في مطابقة المربعات الصغرى يستخدم الانحدار الحصين والطريقة الأعم والأكثر شيوعاً للانحدار الحصين هي تقدير M، يفترض الأنموذج الخطي

$$y_i = x_i'\beta + \epsilon_i \quad (1)$$

وإلى المشاهدة i وبإعطاء المقدر b مطابقة الأنموذج هو

$$\hat{y}_i = a + b_1 x_{i1} + b_2 x_{i2} + \dots + b_k x_{ik} + e_i = x' b_i$$

البواقي تعطى بواسطة $e_i = y_i - \hat{y}_i$ مع تقدير M المقدرات b يمكن أن تحدد بواساطة تقليل دالة الهدف حل كل b

$$\sum_{i=1}^{n} \rho(e_i) = \sum_{i=1}^{n} \rho(y_i - x_i' b)$$

دالة الهدف ρ يجب أن تمتلك الخصائص الآتية:

- $\rho(e) \geq 0$ **غير سالب دائماً**
- $\rho(0) = 0$
- $\rho(e) = \rho(-e)$ **متماثلة**
- $\rho(e_i) \geq \rho(e_i')$ $|e_i|$ **متزايدة والى** $|e_i| > |e_i'|$

بفرض $\Psi = \rho'$ قابلة للاشتقاق حول ρ وتدعى منحنى التأثير وباشتقاق دالة الهدف نسبة الى المعاملات b وبوضع المشتقة الجزئية مساوية الى الصفر فان تقدير المعاملات

$$\sum_{i=1}^{n} \psi(y_i - x_i' b) x_i' = 0$$

بتعريف دالة الوزن

$$w(e) = w(e)/e$$
$$w_i = w(e_i)$$

$$\sum_{i=1}^{n} w_i (y_i - x_i' b) x_i' = 0 \qquad \dots(2)$$

التي تحل بطريقة تكرارية وتسمى المربعات الصغرى ذات الاوزان التكرارية المعادة **IRLS**

1. نختار التقديرات الاولية $b^{(0)}$ مثل تقديرات المربعات الصغرى.

2. عند كل تكرار t تحسب البواقي $e^{(t-1)}{}_i$ والوزن المرافق $w_i^{(t-1)} = w[e^{(t-1)}{}_i]$

3. تحسب تقديرات المربعات الصغرى الموزونة الجديدة $\mathbf{b^{(t)} = [x'w^{(t-1)}x]^{-1}x'w^{(t-1)}y}$
حيث x_i' هو السطر ith وان $\mathbf{w^{(t-1)} = diag[w^{(t-1)}{}_i]}$ الخطوات 2 و 3 تكرر حتى يحصل تقارب المعاملات المقدرة ، مصفوفة التباين المشترك التقاربية b هي

$$\psi(\mathbf{b}) = \frac{\mathbf{E}(\psi^2)}{[\mathbf{E}(\psi')]^2}(\mathbf{x'x})^{-1}$$

باستخدام $\sum[\psi(e_i)]^2$ لتقدير $E[\psi^2]$ و $[\sum\psi'(e_i)/n]^2$ لتقدير $[E(\psi')]^2$.

دوال الهدف

فيما يأتي شكل يمثل دوال الهدف ودوال الوزن لمقدرات M، **مقدر Huber ومقدر bisquare Tukey** او **biweight**

Huber M-estimater[5]

$$\psi(\mathbf{t}) = \begin{cases} \mathbf{t}(1-(\frac{\mathbf{t}}{\mathbf{c}})^2)^2 & \mathbf{if}\ |\mathbf{t}| < \mathbf{c} \\ \mathbf{0} & \mathbf{ow.} \end{cases} \qquad \psi(t) = \begin{cases} t & if\ \ t < b \\ b\,\mathrm{sgn}(t) & if\ t \geq b \end{cases}$$

b ثابت ، c ثابت

Andrew M-Estimator

$$\psi(t) = \begin{cases} \sin(t) & if\ \ -\pi \leq |t| < \pi \\ 0 & ow. \end{cases}$$

Hampel M-Estimator

$$\psi(t)\begin{cases} t & \text{if } |t| \\ a\,\text{sgn}(t) & \text{if } a \le |t| < b \\ \dfrac{c-|t|}{e-b}\text{sgn}(t) & \text{if } b < (t) < c \\ 0 & \text{ow.} \end{cases}$$

حيث a,b,c ثوابت

$$\hat{\beta}^{(M)} = \underset{\beta\in\Re^p}{\text{argmin}} \sum_{i=1}^{n} \rho\left(\frac{r_i(\beta)}{\hat{\sigma}}\right) \quad \dots \quad (3)$$

$$\hat{\sigma} = (median(\left|r_i - \underset{j}{median}(r_j)\right|)$$

حيث c عامل التصحيح والذي يعتمد على التوزيع c=1.4826

مقدر المربعات الصغرى المشذبة LTS[4],[5],[6],[7]

$$\hat{\beta}^{(LTS)} = \underset{\beta\in\Re^p}{\text{argmin}} = \sum_{i=1}^{h} r_{(i)}^{2}(\beta) \quad \dots(4)$$

حيث $r^2_{(1)} \le \dots \le r^2_{(n)}$ بواقي المربعات المرتبة فدائما يوجد مقدر LTS

$$h = [n/2] + [(p+1)/2]p > 1$$

نقطة الانهيار لمقدر LTS هي $\xi^* = ([n-p]/\xi+1)/n$ LTS هو حساس جداً لتغير صغير جداً بالبيانات او لحذف نقطة واحدة في البيانات يسبب تغيراً كبيراً في التقدير.

مقدر المربعات الصغرى الموزونة LWS

لاي $\beta \in \Re^p$ يعرف بواقي الرتبة ith حيث $r_i(\beta) = Y_i - X_i'\beta$

$$\hat{\beta}^{(LWS,W,n)} = \underset{\beta\in\Re^p}{\text{argmin}} \sum_{i=1}^{n} w_i r^2_i(\beta)$$

$$= \underset{\beta\in\Re^p}{\text{argmin}} \sum_{i=1}^{n} W\left(\frac{i=1}{n}\right) r^2(i)(\beta) \quad \dots(5)$$

فالاوزان w_i تعرف دالة الوزن w والتي هي مستمرة بإطلاق $w(0) = 1$.

مقدر انحدار اقل وسيط المربعاتLMS

افترض من قبل Rousseeuw عام (1984) ويسمى اختصارا LMS حيث تبدل مجموع بواقي المربعات التي تتميز بـ OLS مع بواقي وسيط المربعات

$$\min M(y_i - \sum x_{ij}\beta_j)^2 = \min M(e^2{}_i) \;\; \dots \; (6)$$

الفكرة تبديل المجموع مع الوسيط الحصين المقدر الناتج مقاوم للشواذ.

مقدر R

افترض من قبل Jaeckel(1972) والذي يعتمد على التوفيق الخطي إذ R_i تمثل رتب البواقي e_i التي تقلل مجموع الدرجات للبواقي المرتبة

$$\min \sum_{i=1}^{n} a_n(R_i)e_i \;\dots(7)$$

إذ $a_n(i)$ دالة الدرجات التصاعدية والتي تحقق

$$\sum_{i=1}^{n} a_n(i) = 0$$

تحسب رتبة المشاهدات من الوسيط $a_n(i) = i - (\frac{n+1}{2})$

درجات wilcoxcon $a_n(i) = \sin[i - (\frac{n+1}{2})]$

مقدر MM[5]

افترضت أولاً من قبل Yahai عام (1987) وهي الأكثر شيوعاً لتقنية الانحدار الحصين إجراء الطريقة هو كالآتي :

1. التقدير الابتدائي للمعاملات $\hat{\beta}^{(1)}$ وتتبعها البواقي $e^{(1)}{}_i$ من الانحدار المقاوم (ويعني الانحدار مع نقطة الانهيار 50%.
2. البواقي $e^{(1)}{}_i$ من التقدير الابتدائي في الخطوة 1 تستخدم لحساب M لمقياس البواقي $\hat{\sigma}_e$.
3. التقدير الابتدائي للبواقي $e^{(1)}{}_i$ من الخطوة 1 ومن قياس البواقي $\hat{\sigma}_e$ في الخطوة 2 يستخدم في التكرار الاول للمربعات الصغرى الموزونة لتحديد تقديرات M معامل الانحدار هو

$$\sum_{i=1}^{n} w_i(e^{(1)}{}_i/\hat{\sigma}_e)x_i = 0 \;\;\dots\; (8)$$

إذ wi هو Huber او اوزان bisquare.

4. تحسب الأوزان الجديدة $w_i^{(2)}$ باستخدام البواقي من WLS الابتدائي.

5. بالاحتفاظ بثبات مقياس البواقي من الخطوة 2 وبإستمرار تعاد الخطوات حتى التقارب.

مقدر LWS[5],[6],[7]

اقترحت من قبل visek عام (2000)

$$\hat{\beta}^{(LWS,w,n)} = \underset{\beta\in\Re^p}{\arg\min} \sum_{i=1}^{n} w_i r_{(i)}^{2}(\beta)$$
$$= \underset{\beta\in\Re^p}{\arg\min} \sum_{i=1}^{n} w\left(\frac{i-1}{n}\right) r_{(i)}^{2}\beta \quad ...(9)$$

$w(0)=1$ مع المشتقة $w'(t)$

$$\hat{\beta}^{(LWS,w,n)} = \underset{\beta\in\Re^p}{\arg\min} \sum_{i=1}^{n} w\left(\frac{\pi(\beta_{1i})-1}{n}\right) r_i^{2}(\beta)$$
$$\sum w\left(\frac{\pi(\beta_{1i})-1}{n}\right) X_i(Y_i - X_i'\beta) = 0$$

مقدر GM

مقدرات M العمومية مقدمة لغرض تحديد دالة التأثير بمعنى دالة الوزن w

$$\hat{\beta}^{(GM)} = \underset{\beta\in\Re^p}{\arg\min} \sum_{i=1}^{n} w(X_i)\frac{\rho(r_i(\beta\beta))}{\hat{\sigma}} \quad ...(10)$$

التعريف يمكن إعادة كتابته

$$\sum w(x)\psi\left(\frac{r_i}{\sigma}\right)X_i = 0 \quad ...(11)$$

حيث ρ عدد معاملات الانحدار وتحسب كالآتي :

1. تقدير $\hat{\beta}^{(OLS)}$ لـ β^{0}

2. تحسب البواقي $i=1,...,n$ $\quad r_i(\hat{\beta}) = (Y_i - \hat{Y}_i) = Y_i - X_i'\hat{\beta}$

3. حساب تقدير $\hat{\sigma}$ $\quad$ MAD $\hat{\sigma} = 1.4826\, median(|r_i - median(r_i)|)$

4. حساب الوزن wi مثال

$$Andrew's\,\psi = w_i = \frac{\psi\left(\frac{r_i}{\hat{\sigma}}\right)}{\frac{r_i}{\hat{\sigma}}}$$

5. تحديد التقدير $\hat{\beta}$ بوساطة انجاز أوزان المربعات الصغرى مع اوزان w_i.

حساب $\hat{\beta}^{(WLS)} = (X'WX)^{-1} X'WY$

6. الرجوع للخطوة 2 ونستمر حتى التقارب.

مقدر S

لمعالجة نقطة الانهيار الصغرى لمقدر M افترض Rousseews&Leroy(1987) مقدر S الذي يوجد التشتت الأقل للبواقي

$$\min \hat{\sigma}(e_1(\hat{\beta}),\ldots,e_n(\hat{\beta}))$$

مقدر S الحصين لمقياس البواقي

$$\frac{1}{n}\sum_{i=1}^{n}\rho\left(\frac{e_i}{\hat{\sigma}_e}\right) = b \quad \ldots(12)$$

حيث b هو ثابت يعرف S

$$b = E\phi[\rho(e)] \quad \ldots \quad (13)$$

و ϕ تمثل التوزيع الطبيعي القياسي باشتقاق المعادلة 12 وحل النتائج

$$\frac{1}{n}\sum_{i=1}^{n}\psi\left(\frac{e_i}{\hat{\sigma}_e}\right) = b \quad \ldots (14)$$

حيث ψ تبدل بدالة وزن ملائم و توظف دالة الوزن عادة biweight مع أنه مقدر S تملك نقطة انهيار 0.5.

3. الجانب التجريبي

تم استخدام أسلوب المحاكاة لتوليد بيانات لمتغيرات عشوائية إذ تم توليد متغير توضيحي وفق التوزيع الطبيعي بمتوسط صفر وتباين واحد ثم اجراء عملية التلويث لقيم xi و ei

لحجمي العينة n=50 و n=100 وبنسبة تلويث 10% الحالة الثانية تلويث قيم xi و ei بنسبة تلويث 20% ولحجمي العينة 50 و100 ولمتغير واحد ومتغيرين توضيحيين، وقد تم استخدام البرنامج الإحصائي Matlab إذ يُعد من البرمجيات (Software) القابلة للبرمجة وذو إمكانيات عالية في الجوانب الرياضية والإحصائية والهندسية إذ يوظف الأدوات بإستقامة لبرمجة متقدمة جداً.

وتم الاعتماد على الإمكانية العالية لبرنامج MATLAB إذ تم إجراء خطوات عديدة للحصول على البيانات وكانت كما يأتي:

1. تم توليد البيانات بشكل غير الملوث بالاستعانة بالدالة المكتبية Randn (التي تعمل على توليد خلايا من الارقام العشوائية التي تتوزع توزيعاً طبيعياً بمتوسط وتباين $\sigma^2 = 1$)

ومن ثم تحويل البيانات من التوزيع N(0,1) الى التوزيع $N(\mu,1)$ من خلال العلاقة $X = Z + \mu$ وبهذه الطريقة تم توليد بيانات بتوزيع N(-1,1).

2. توليد المتغيرات بالشكل الملوث حيث تم اتباع أسلوب توليد متغيرات التوزيع الطبيعي نفسه لتوليد البيانات النظيفة بمتوسط μ وتباين 1 ولكن تم استقطاع جزء من البيانات بنسبة 10% و 20% والتي تمثل نسب التلوث ولكل حجم عينة 50 و100 باستخدام نفس دالة التوليد Randn لكن يكون توزيع الشواذ N(9,1).

جدول رقم (1)

نقطة الانهيار لطرائق التقدير

Estimator	Breakdown point
$\hat{\beta}_{OLS}$	0
$\hat{\beta}_{M}$	0
$\hat{\beta}_{MM}$	0
$\hat{\beta}_{LTS}$	0.50
$\hat{\beta}_{LMS}$	0.50
$\hat{\beta}_{R}$	0.50
$\hat{\beta}_{GM}$	0.50
$\hat{\beta}_{S}$	0.50

جدول رقم (2)

تقديرات المعلمات بالطرائق الحصينة لحجم العينة 50 وبنسبة تلويث 10% ولمتغير توضيحي واحد اي انموذج بسيط

المعلمات	$\hat{\beta}_{OLS}$	$\hat{\beta}_{M}$	$\hat{\beta}_{MM}$	$\hat{\beta}_{LTS}$
intercept	-0.322	0.325	-4.254	-3.256
$\hat{\beta}_1$	6.324	4.322	3.267	6.345
MSE	34.26	22.18	23.66	11.66
المعلمات	$\hat{\beta}_{LMS}$	$\hat{\beta}_{R}$	$\hat{\beta}_{GM}$	$\hat{\beta}_{S}$
intercept	-0.122	-0.076	2.345	-3.567
$\hat{\beta}_1$	2.671	1.879	3.456	2.667
MSE	16.87	4.52	8.96	0.89

جدول رقم (3)
نقطة الانهيار لطرائق التقدير

Breakdown point	Estimator
0	$\hat{\beta}_{OLS}$
0.2	$\hat{\beta}_{M}$
0.23	$\hat{\beta}_{MM}$
0.50	$\hat{\beta}_{LTS}$
0.50	$\hat{\beta}_{LMS}$
0.50	$\hat{\beta}_{R}$
0.50	$\hat{\beta}_{GM}$
0.50	$\hat{\beta}_{S}$

جدول رقم (4)
تقديرات المعلمات بالطرائق الحصينة لحجم عينة 50 وبنسبة تلويث 10% ولمتغيرين توضيحيين اي انموذج متعدد

$\hat{\beta}_{LTS}$	$\hat{\beta}_{MM}$	$\hat{\beta}_{M}$	$\hat{\beta}_{OLS}$	***المعلمات***
-0.321	-0.732	-0.629	-0.573	***intercept***
7.921	3.325	3.781	10.327	$\hat{\beta}_1$
0.089	6.733	3.716	0.356	$\hat{\beta}_2$
6.77	21.78	23.66	26.98	MSE

$\hat{\beta}_{S}$	$\hat{\beta}_{GM}$	$\hat{\beta}_{R}$	$\hat{\beta}_{LMS}$	***المعلمات***
-0.736	-2.627	-3.671	-0.923	***intercept***
5.327	3.271	4.325	3.267	$\hat{\beta}_1$
6.725	8.571	5.121	5.627	$\hat{\beta}_2$
3.22	11.23	2.34	5.86	MSE

جدول رقم (5)
نقطة الانهيار لطرائق التقدير

Breakdown point	Estimator
0	$\hat{\beta}_{OLS}$
0.2	$\hat{\beta}_{M}$
0.3	$\hat{\beta}_{MM}$
0.50	$\hat{\beta}_{LTS}$
0.50	$\hat{\beta}_{LMS}$
0.50	$\hat{\beta}_{R}$
0.50	$\hat{\beta}_{GM}$
0.50	$\hat{\beta}_{S}$

جدول رقم (6)
تقديرات المعلمات بالطرائق الحصينة لحجم عينة 50 وبنسبة تلويث 20% ولمتغير توضيحي اي انموذج بسيط

المعلمات	$\hat{\beta}_{OLS}$	$\hat{\beta}_{M}$	$\hat{\beta}_{MM}$	$\hat{\beta}_{LTS}$
intercept	-0.345	-0.225	-0.347	-0.678
$\hat{\beta}_1$	8.965	7.623	7.009	5.236
MSE	17.324	16.234	15.225	9.876

المعلمات	$\hat{\beta}_{LMS}$	$\hat{\beta}_{R}$	$\hat{\beta}_{GM}$	$\hat{\beta}_{S}$
intercept	-0.075	-0.256	-0.674	-0.256
$\hat{\beta}_1$	4.325	6.234	2.456	3.227
MSE	3.267	2.987	7.235	4.336

جدول رقم (7)
تقدير المعلمات بالطرائق الحصينة لحجم عينة 50 وبنسبة تلويث 20% ولمتغيرين توضيحيين اي انموذج متعدد

المعلمات	$\hat{\beta}_{OLS}$	$\hat{\beta}_{M}$	$\hat{\beta}_{MM}$	$\hat{\beta}_{LTS}$
intercept	0.876	0.654	0.886	0.324
$\hat{\beta}_1$	32.892	27.567	33.879	18.965
$\hat{\beta}_2$	17.654	15.236	11.256	10.456
MSE	77.23	70.34	66.78	11.67

المعلمات	$\hat{\beta}_{LMS}$	$\hat{\beta}_{R}$	$\hat{\beta}_{GM}$	$\hat{\beta}_{S}$
intercept	0.224	0.256	0.334	0.456
$\hat{\beta}_1$	22.665	14.234	11.345	10.235
$\hat{\beta}_2$	8.457	6.235	5.897	12.234
MSE	10.45	6.34	23.45	8.34

جدول رقم (8)
نقطة الانهيار لطرائق التقدير

Estimator	Breakdown point
$\hat{\beta}_{OLS}$	0
$\hat{\beta}_{M}$	0
$\hat{\beta}_{MM}$	0
$\hat{\beta}_{LTS}$	0.50
$\hat{\beta}_{LMS}$	0.50
$\hat{\beta}_{R}$	0.50
$\hat{\beta}_{GM}$	0.50
$\hat{\beta}_{S}$	0.50

جدول رقم (9)

تقديرات المعلمات بالطرائق الحصينة بحجم عينة 100 وبنسبة تلويث 10% ولمتغير توضيحي اي انموذج بسيط

المعلمات	$\hat{\beta}_{OLS}$	$\hat{\beta}_{M}$	$\hat{\beta}_{MM}$	$\hat{\beta}_{LTS}$
intercept	-0.826	-0.326	0.426	-0.533
$\hat{\beta}_1$	8.321	6.215	5.351	5.726
MSE	36.34	22.67	22.33	21.34

المعلمات	$\hat{\beta}_{LMS}$	$\hat{\beta}_{R}$	$\hat{\beta}_{GM}$	$\hat{\beta}_{S}$
intercept	-0.627	0.267	0.627	0.626
$\hat{\beta}_1$	3.711	3.255	2.701	3.261
MSE	21.45	5.35	12.34	5.27

جدول رقم (10)

نقطة الانهيار لطرائق التقدير

Estimator	Breakdown point
$\hat{\beta}_{OLS}$	0
$\hat{\beta}_{M}$	0
$\hat{\beta}_{MM}$	0
$\hat{\beta}_{LTS}$	0.50
$\hat{\beta}_{LMS}$	0.50
$\hat{\beta}_{R}$	0.50
$\hat{\beta}_{GM}$	0.50
$\hat{\beta}_{S}$	0.50

جدول رقم (11)

تقديرات المعلمات بالطرائق الحصينة لحجم عينة 100 وبنسبة تلويث 10% ولمتغيرين توضيحيين اي انموذج متعدد

intercept	$\hat{\beta}_{OLS}$	$\hat{\beta}_{M}$	$\hat{\beta}_{MM}$	$\hat{\beta}_{LTS}$
$\hat{\beta}_1$	-0.345	-0.634	-0.989	-0.567
$\hat{\beta}_2$	6.232	4.324	3.245	3.456
MSE	88.35	56.34	36.24	18.97

intercept	$\hat{\beta}_{LMS}$	$\hat{\beta}_{R}$	$\hat{\beta}_{GM}$	$\hat{\beta}_{S}$
$\hat{\beta}_1$	-0.732	-0.534	-0.734	-0.534
$\hat{\beta}_2$	7.345	3.456	3.456	5.350
MSE	22.87	4.25	13.35	8.56

جدول رقم (12)

نقطة الانهيار لطرائق التقدير

Estimator	Breakdown point
$\hat{\beta}_{OLS}$	0
$\hat{\beta}_{M}$	0.20
$\hat{\beta}_{MM}$	0.20
$\hat{\beta}_{LTS}$	0.50
$\hat{\beta}_{LMS}$	0.50
$\hat{\beta}_{R}$	0.50
$\hat{\beta}_{GM}$	0.50
$\hat{\beta}_{S}$	0.50

جدول رقم (13)

تقدير المعلمات بالطرائق الحصينة ولحجم العينة 100 وبنسبة تلويث 20% ولمتغير توضيحي اي انموذج بسيط

المعلمات	$\hat{\beta}_{OLS}$	$\hat{\beta}_{M}$	$\hat{\beta}_{MM}$	$\hat{\beta}_{LTS}$
intercept	-0.965	-0.912	-0.888	-0.061
$\hat{\beta}_1$	-23.783	-33.451	-18.304	-15.327
MSE	12.89	11.87	12.22	8.77

المعلمات	$\hat{\beta}_{LMS}$	$\hat{\beta}_{R}$	$\hat{\beta}_{GM}$	$\hat{\beta}_{S}$
intercept	0.345	-0.234	-0.256	0.006
$\hat{\beta}_1$	17.729	-15.364	-12.404	-13.444
MSE	7.66	4.23	13.45	2.36

جدول رقم (14)

نقطة الانهيار لطرائق التقدير

Estimator	Breakdown point
$\hat{\beta}_{OLS}$	0
$\hat{\beta}_{M}$	0
$\hat{\beta}_{MM}$	0
$\hat{\beta}_{LTS}$	0.50
$\hat{\beta}_{LMS}$	0.50
$\hat{\beta}_{R}$	0.50
$\hat{\beta}_{GM}$	0.50
$\hat{\beta}_{S}$	0.50

جدول رقم (15)

تقدير المعلمات بالطرائق الحصينة ولحجم العينة 100 وبنسبة تلويث 20% ولمتغيرين توضيحيين اي انموذج متعدد

$\hat{\beta}_{LTS}$	$\hat{\beta}_{MM}$	$\hat{\beta}_{M}$	$\hat{\beta}_{OLS}$	المعلمات
78.458	101.245	112.245	121.243	***intercept***
-23.246	-22.009	-22.678	-24.456	$\hat{\beta}_1$
19.245	16.245	-19.24	21.562	$\hat{\beta}_2$
24.66	24.78	30.78	33.45	MSE

$\hat{\beta}_{S}$	$\hat{\beta}_{GM}$	$\hat{\beta}_{R}$	$\hat{\beta}_{LMS}$	المعلمات
78.224	77.234	66.457	68.345	***intercept***
-16.234	17.234	18.905	26.231	$\hat{\beta}_1$
15.987	16.987	15.237	15.236	$\hat{\beta}_2$
6.22	28.55	7.24	15.24	MSE

من الجداول المذكورة آنفاً يلاحظ أن المقدرات M و MM و OLS لا تعمل جيداً إذ انها تملك نقطة انهيار صفر او قريبة من الصفر في حين أن المقدرات R و S و LTS و LMS تعمل جيداً وتملك نقطة انهيار جيدة هي 0.5 أو قريبة من 0.5 وتملك أقل تباين.

كما أن في حالة وجود القيم الشاذة وفي حالة تلويث البيانات وعدم تحقق الفرضيات فان مقدر R و S و GM و LMS تملك اقل MSE في حين أن المقدرات OLS و M و MM و LTS تملك MSE عالي.

الاستنتاجات والتوصيات

أولا: الاستنتاجات:

1. يلاحظ أن نقطة انهيار مقدر M ومقدر MM هو قريب من مقدر OLS اي انهما ليسا أفضل من مقدر OLS ولكافة حجوم العينات.
2. مقدر R ومقدر S ومقدر LTS ومقدر LMS تعمل جيداً وتملك نقطة انهيار جيدة 0.5 لذلك يفضل استخدامها في حالة احتواء البيانات على الشواذ وتملك تبايناً قليلاً و حجوم العينات كافة.

3. يلاحظ أنه في حالة وجود القيم الشاذة وعدم تحقق الفرضيات نجد الى أن مقدر R و S و GM وLMS تنتج أقل MSE على عكس المقدرات OLS وM و MM و LTS ولكافة حجوم العينات.

ثانيا: التوصيات

1. عند انتهاك الفرضيات وعدم تحقق شرط التوزيع الطبيعي نوصي باعتماد المقدرات R,S, LMS، LTS، GM كونها تملك نقطة انهيار 0.5 وهي مقاومة تجاه الشواذ واقل MSE ولكافة حجوم العينات.
2. لا نوصي باعتماد المقدرات OLS وM و MM كونها غير مقاومة تجاه الشواذ وتملك نقطة انهيار صفر أو قريبة من الصفر اي تفشل بإعطاء مقدر جيد في حالة وجود الشواذ وتملك MSE عالياً ولكافة حجوم العينات.

المصادر:

أولا: العربية

1. شاكر، صالح مويد "تحسين اسلوب M الحصين في تقدير معلمات انموذج الانحدار الخطيً ، المجلة العراقية للعلوم الاحصائية 16، 2009.

ثانيا: الاجنبية

1. Brown,Lawrence,Cai,T.Tony&Zhou,H.H,"Robust nonparametric estimation via wavelet median regression, 2008, Vol. 36 No.5 2055-**2084.**
2. Fox,John& ,Weisberg, Sanford "Robust Regression in R ,December 2010.
3. Fox,john& Weisberg, Sanford "Robust Regression" ,2013.
4. France,Jiri"Robust regression : Robust Estimation of Regression Coefficient in Linear Regression Model when orthogonality condition
5. https://www.sagepub.com/sites/default/files/upm-binaries/17839_Chapter_4.pdf
6. http://www.mathworks.com/help/stats/robustfit.html