دراسة مقارنة للانحدار الحصين "دراسة باستخدام المحاكاة"

أ. م. د. خلود يوسف

كلية الإدارة والاقتصاد/ جامعة بغداد

Robust Estimation Study "By using Simulation Study Regression"

Assist. Prof. Dr. Claud yosif

College Administration &Economic/ University of Baghdad

تاريخ استلام البحث 2017/11/6 تاريخ قبول النشر 2018/5/9

المستخلص:

فكرة البحث تتلخص في إيجاد مقدرات حصينة لأنموذج الانحدار الخطي في حالة وجود القيم الشاذة التي 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 أن طرائق OLS و M و MM تملك نقطة انهيار صفر على خلاف 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 ولكافة العينات.

الكلمات المفتاحية: **انحدار حصين.**

Abstract:

Robust Regression Estimators Study "By using Simulation Study"

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.**

1- 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 بخلاف الطرائق التقليدية التي تكون حساسة تجاه الشواذ ومن مقدرات التقدير للمعلمات التي تستخدم بشكل واسع مقدرات M إذ تعالج الخلل في متجه الأخطاء العشوائية بطريقة المربعات الصغرى الاعتيادية دون معالجة الخلل في المتغيرات التوضيحية وطورت إلى مقدرات S ومقدرات R.

2- الجانب النظري

2-1 مقدر MM-Estimation[1],[7],[6]

تقديرات المربعات الصغرى تعمل بشكل سيء حينما لا يكون توزيع الخطأ طبيعياً وبخاصةً حينما تكون الأخطاء ثقيلة الأثر ولإزالة تأثير المشاهدات الشاذة لمطابقة المربعات الصغرى يستخدم الانحدار الحصين والطريقة الأكثر عموماً وشيوعاً 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 بوساطة $e_i = y_i - \hat{y}_i$ مع تقدير M المقدرات b يمكن أن تحدد بوساطة 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 الاتية:

- $\rho(e) \geq 0$ 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 إلى الصفر فإن تقدير المعاملات:

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

بتعريف دالة الوزن

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

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

التي تحل بطريقة تكرارية وتسمى المربعات الصغرى ذات الأوزان التكرارية المعادة IRLS

1. نختار التقديرات الأولية $b^{(0)}$ 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 و $\mathbf{w^{(t-1)} = diag[w^{(t-1)}{}_i]}$ الخطوات 2 و 3 تكرر حتى يحصل تقارب المعاملات المقدرة ، مصفوفة التباين المشترك التقاربية b هي:

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

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

2-1-1 دوال الهدف

الشكل الاتي يمثل دوال الهدف ودوال الوزن لمقدرات M، مقدر Huber ومقدر bisquare أو Tukey biweightHuber M-estimater[5]

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

b ثابت، c ثابت

Andrew M-Estimator

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

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

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

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

حيث c عامل التصحيح الذي يعتمد على التوزيع c=1.4826

2-2 مقدر المربعات الصغرى المشذبة LTS [4], [5], [6], [7]

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

حيث $r^2{}_{(1)} \le \ldots \le r^2{}_{(n)}$ بواقي المربعات المرتبة ويوجد مقدر LTS دائماً

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

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

3-2 مقدر المربعات الصغرى الموزونة LWS

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

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

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

والأوزان w $_i$ تعرف دالة الوزن W وهي مستمرة مطلقاً $w(0) = 1$.

2-4 مقدر انحدار أقل وسيط المربعات LMS

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

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

والفكرة تبديل المجموع مع الوسيط الحصين المقدر الناتج مقاوم للشواذ.

2-5 مقدر R

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

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

فدالة $a_n(i)$ دالة الدرجات التصاعدية التي تحقق

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

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

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

2-6 مقدر MM[5]

افترضت أولا من قبل Yahai عام (1987) وهي الأكثر شيوعا لتقنية الانحدار الحصين إجراء الطريقة هو كالاتي:

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 الأول للمربعات الصغرى الموزونة لتحديد تقديرات M معامل الانحدار هو

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

حيث wi هو Huber أو أوزان bisquare.

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

5. بالاحتفاظ بثبات مقياس البواقي من الخطوة 2 وتعاد الخطوات بأستمرار حتى التقارب.

2-7 مقدر LWS[5],[6],[7]

اقترحت من قبل visek عام (2000)

$$\hat{\beta}^{(LWS,w,n)} = \underset{\beta \in \Re^p}{\operatorname{argmin}} \sum_{i=1}^{n} w_i r_{(i)}{}^2(\beta)$$

$$= \underset{\beta \in \Re^p}{\operatorname{argmin}} \sum_{i=1}^{n} w\left(\frac{i-1}{n}\right) r_{(i)}{}^2 \beta \quad \dots (9)$$

$w'(t)$ مع المشتقة $w(0)=1$

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

$$\sum w\left(\frac{\pi(\beta_{1i})-1}{n}\right) X_i (Y_i - X_i'\beta) = 0$$

2-8 مقدر GM

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

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

التعريف يمكن إعادة كتابته

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

حيث ρ عدد معاملات الانحدار وتحسب كالآتي:

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

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

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

4. حساب الوزن wi مثال

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

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

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

6. الرجوع للخطوة 2 ونستمر حتى التقارب.

2-9 مقدر S

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

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

مقدر S الحصين لمقياس البواقي

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

حيث b هو ثابت يعرف S

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

و ϕ تمثل التوزيع الطبيعي القياسي باشتقاق المعادلة 12 وحل النتائج

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

حيث ψ تبدل بدالة وزن ملائم وعادة توظف دالة الوزن biweight مع انه 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 إجراء 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 يعد من البرمجيات (Software) القابلة للبرمجة وذو إمكانات عالية في الجوانب الرياضية والإحصائية والهندسية فيوظف الأدوات بإستقامة لبرمجة متقدمة جداً.

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

1. تم توليد البيانات بالشكل غير الملوث بالاستعانة بالدالة المكتبية Randn (التي تعمل على توليد خلايا من الأرقام العشوائية التي تتوزع توزيعا طبيعيا بمتوسط 0 وتباين $\sigma^2=1$) ومن ثم تحويل البيانات من التوزيع N(0,1) إلى التوزيع $N(\mu,1)$ من خلال العلاقة $X=Z+\mu$ وبهذه الطريقة تم توليد بيانات بتوزيع N(−1,1).
2. توليد المتغيرات بالشكل الملوث فتم اتباع أسلوب توليد متغيرات التوزيع الطبيعي نفسه لتوليد البيانات النظيفة بمتوسط μ وتباين 1 ولكن تم استقطاع جزء من البيانات بنسبة 10% و 20% التي 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 أي أنموذج 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

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

جدول رقم (4)

تقديرات المعلمات بالطرائق الحصينة لحجم عينة 50 وبنسبة تلويث 10% ولمتغيرين توضيحيين أي أنموذج متعدد

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

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

جدول رقم (5)

نقطة الانهيار لطرائق التقدير

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

جدول رقم (6)

تقديرات المعلمات بالطرائق الحصينة لحجم عينة 50 وبنسبة تلويث 20% ولمتغير توضيحي أي أنموذج 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 تلوث 20% ولمتغيرين توضيحيين أي أنموذج 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 أي أنموذج 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 أي أنموذج 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 أي أنموذج 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 أي أنموذج متعدد

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

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

من الجداول المذكور انفاً يلاحظ أن المقدرات M و MM وOLS لا يعملان جيداً لأنهما يملكان نقطة انهيار صفر أو قريبة من الصفر في حين أن المقدرات R و S و LTS و LMS تعملان جيداً وتمتلكان 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 فإن مقدر R وS و GM وLMS يملكان أقل MSEفي حين أن المقدرات OLS و M وMM وLTS تملكان MSE عالياً.

4- الاستنتاجات

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

3. يلاحظ أنه في حالة وجود القيم الشاذة وعدم تحقق الفرضيات نجدإلى أن مقدر Rو S و GM وLMS ينتجان أقل MSE بخلاف المقدرات OLS وM و MM و LTS ولحجوم العينات كافة.

5- 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 ولحجوم العينات كافة.
2. لا نوصي باعتماد المقدرات OLSوM و MM كونها غير مقاومة تجاه الشواذ وتملك نقطة انهيار صفر أو قريبة من الصفر أي تفشل بإعطاء مقدر جيد في حالة وجود الشواذ وتملك MSE عالياً ولحجوم العينات كافة.

المصادر

أولا: العربية

1. شاكر، صالح مويد " تحسين أسلوب M الحصين في تقدير معلمات أنموذج الانحدار الخطي، المجلة العراقية للعلوم الإحصائية 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.
3. Fox,John& ,Weisberg, Sanford "Robust Regression in R, December 2010.
4. Fox,john& Weisberg, Sanford "Robust Regression" ,2013.
5. France,Jiri"Robust regression : Robust Estimation of Regression Coefficient in Linear Regression Model when orthogonality condition
6.https://www.sagepub.com/sites/default/files/upmbinaries/17839_Chapter_4.pdf
7. http://www.mathworks.com/help/stats/robustfit.html