"المقارنة بين بعض دوال طريقة M الحصينة في التوزيع اللوجستي في حالة أنموذج الانحدار الخطي المتعدد"

م.م. احمد ذياب احمد
كلية الادارة والاقتصاد
جامعة بغداد

Abstract:

This research discusses to compares between some of functions the M- robust method in the logistic distribution in state the multiply linear regression model by using the simulation on the size of samples different and we give theoretical M - Estimators.

1- المقدمة:

تعتبر طريقة M من أهم الطرائق الحصينة وذات كفاءة عالية مقارنة مع طريقة المربعات الصغرى عندما يتوزع الخطأ $N(0,\sigma^2)$ [2] وقد أطلق Huber في عام 1972 اسم مقدر M على صنف من المقدرات الحصينة وإن تطبيق هذه الطريقة في الدراسات المتعلقة بعلوم الحياة والعلوم الزراعية والطبية حيث يستخدم التوزيع السوقي (اللوجستي) في مثل هذا النوع من الدراسات[3] يكون بسبب حدوث خلل في إحدى فرضيات التوزيع الطبيعي أو تلوث البيانات بقيم شاذة.

وعلى فرض لدينا معادلة الانحدار آلاتية

$$\gamma = b_0 + b_1 x_1 + b_2 x_2 \quad \ldots\ldots\ldots\ldots\ldots (1)$$

حيث إن

γ : المتغير المعتمد

x_1, x_2 : المتغيرات المستقلة

b_0, b_1, b_2 : معلمات معادلة الانحدار

إن دالة الكثافة الاحتمالية للتوزيع هي

$$f(\chi) = \frac{1}{\beta}\left[e^{-\left(\frac{\chi-\alpha}{\beta}\right)}\right]\left[1+e^{-\left(\frac{\chi-\alpha}{\beta}\right)}\right]^{-2} \quad \ldots\ldots\ldots\ldots\ldots (2)$$

$$-\infty \prec \chi \prec \infty$$

حيث إن α, β تمثلان معلمتي التوزيع

$$\beta \succ 0 \quad , \quad -\infty \prec \alpha \prec \infty$$

أما الدالة التوزيعية لهذا التوزيع هي

$$F(\chi) = \left[1+e^{-\left(\frac{\chi-\alpha}{\beta}\right)}\right]^{-1} \qquad \ldots\ldots\ldots\ldots (3)$$

2- طريقة M[1]

تعتبر طريقة M من الطرائق الحصينة الواسعة الانتشار وذات الكفاءة العالية والمطلوب من هذه الطريقة هو تصغير المقدار

$$\sum_{i=1}^{r}\rho\left(\mathrm{Y}_i-\chi'_i\mathrm{B}\right) \quad \ldots\ldots\ldots\ldots (4)$$

حيث إن ρ دالة محدبة ومتماثلة ولتحقيق خاصية Scale Invariant تكون الصيغة (4) كلاتي

$$\sum_{i=1}^{r}\rho\left(\frac{\mathrm{Y}_i-\chi'_i\mathrm{B}}{\hat{\sigma}}\right) \quad \ldots\ldots\ldots\ldots (5)$$

وبأخذ المشتقة للدالة ρ بالنسبة إلى متجه المعلمات B ومساواتها إلى الصفر نحصل على

$$\sum_{i=1}^{r}\Psi\left[\frac{\left(\mathrm{Y}_i-\chi'_i\mathrm{B}\right)}{\hat{\sigma}}\right]\chi_{sj}=0 \quad \ldots\ldots\ldots\ldots (6)$$

$$j=1, 2, \ldots\ldots\ldots k$$

وبإعادة كتابة المعادلة (6) كلاتي

$$\sum_{i=1}^{r}w_i\chi_{ij}\left(\mathrm{Y}_i-\chi'_i\mathrm{B}\right)\Big/\hat{\sigma}=0 \quad \ldots\ldots\ldots (7)$$

حيث إن (w_i) دالة الوزن وتحسب كلاتي

$$w_i=\frac{\Psi\left(\frac{\left(\mathrm{Y}_i-\chi'_i\mathrm{B}_h\right)}{\hat{\sigma}}\right)}{\left[\frac{\left(\mathrm{Y}_i-\chi'_i\mathrm{B}_h\right)}{\hat{\sigma}}\right]} \quad \ldots\ldots\ldots\ldots (8)$$

B_h : متجه المعلمات الابتدائية

$\hat{\sigma}$: تمثل معلمة القياس وتقدر من الصيغة آلاتية

$$\hat{\sigma}=1.483\left[Median\left|u_i-Median(u_i)\right|\right] \quad \ldots\ldots\ldots (9)$$

وباستخدام طريقة المربعات الصغرى يمكن حل المعادلة (7) كلاتي

$$\hat{\mathrm{B}}=(\chi' w\chi)^{-1}\chi' wy \quad \ldots\ldots\ldots (10)$$

وان

W: مصفوفة الأوزان القطرية عناصر القطر فيها الأوزان (w_i)

وبما إن حصانة المقدر تعتمد على الدالة ρ فان هناك صيغ عديدة لهذه الدالة منها

1- دالة Huber [4]

$$\rho(u)=\begin{cases}\frac{u^2}{2} & |u|\le c\\ c|u|-\frac{c^2}{2} & |u|\succ c\end{cases}$$

$$\Psi(u)=\begin{cases}u & |u|\le c\\ csign(u) & |u|\succ c\end{cases}$$

وان c=1.345

2- دالة ثنائية الأس Douple exponential

$$\Psi(u)=\begin{cases}-1 & u\prec 0\\ 1 & u\succ 0\end{cases}$$

3- دالة اندراوس Andrews

$$\Psi(u)=\begin{cases}\sin\left(\frac{u}{h}\right) & |u|\le h\pi\\ 0 & |u|\succ h\pi\end{cases}$$

وان h=1.5

4- دالة رمزي Ramsy

$$\Psi(u)=\begin{cases}u\exp\left(\frac{-a}{|u|}\right) & |u|\le\frac{1}{a}\\ 0 & o.w\end{cases}$$

وان a=0.3

3- المحاكاة:

وللمقارنة بين صيغ الدالة ρ في حالة التوزيع اللوجستي تم تكرار التجربة 1000 مرة ولإحجام عينات مختلفة (100,50,10) وبافتراض قيم معلمات التوزيع $\beta=1,3,5$, $\alpha=1,3,5$ والتي تستخدم لتوليد عينة عشوائية وفق الصيغة

$$\chi=-\beta\log\left(\frac{1}{u}-1\right)+\alpha$$

كما انه تم اخذ قيم معلمات معادلة الانحدار $b_0=0.5$, $b_1=0.9$, $b_2=0.5$

جدول (1) إيجاد MSE للمعلمات

حجم العينة	المعلمات	$\alpha=1,\beta=1$			
		Huber	Exponential	Andrews	Ramsy
n=10	b_0	0.10882	0.00985	0.08057	0.00174
	b_1	0.03729	0.00197	0.04203	0.00125
	b_2	0.00207	0.00062	0.00247	0.00177

n=50	b_0	0.00655	0.00128	0.00396	0.00046
	b_1	0.00231	0.0001	0.00107	0.0001
	b_2	0.00018	0.0004	0.00025	0.00069
n=100	b_0	0.00009	0.00062	0.01146	0.0008
	b_1	0.00273	0.0072	0.01991	0.00438
	b_2	0.00097	0.00166	0.00201	0.00153

جدول (2) إيجاد MSE للمعلمات

حجم العينة	المعلمات	$\alpha = 3 , \beta = 3$			
		Huber	Exponential	Andrews	Ramsy
n=10	b_0	0.07241	0.00664	0.07312	0.01579
	b_1	0.04965	0.004	0.05859	0.00509
	b_2	0.00262	0.00044	0.00303	0.00568
n=50	b_0	0.02864	0.01026	0.02039	0.01404
	b_1	0.00416	0.00075	0.00282	0.00119
	b_2	0.00009	0.00017	0.00011	0.00015
n=100	b_0	0.00487	0.00067	0.00082	0.0002
	b_1	0.00008	0.01662	0.00017	0.01826
	b_2	0.00029	0.00035	0.00021	0.00048

جدول (3) إيجاد MSE للمعلمات

حجم العينة	المعلمات	$\alpha = 5 , \beta = 5$			
		Huber	Exponential	Andrews	Ramsy
n=10	b_0	0.01092	0.0105	0.00001	0.00755
	b_1	0.00285	0.00909	0.03248	0.04074
	b_2	0.00167	0.00246	0.0037	0.00652
n=50	b_0	0.18361	0.01147	0.0836	0.00089
	b_1	0.03707	0.00059	0.02204	0.00228
	b_2	0.00003	0.00017	0.00001	0.00089

n=100	b_0	0.08425	0.0017	0.03588	0.01215
	b_1	0.10968	0.00578	0.07611	0.01452
	b_2	0.00596	0.00063	0.00469	0.0008

جدول (4) إيجاد MSE للمعلمات

حجم العينة	المعلمات	$\alpha=1, \beta=3$			
		Huber	Exponential	Andrews	Ramsy
n=10	b_0	0.30055	0.01267	0.17022	0.01719
	b_1	0.10871	0.00037	0.08403	0.00075
	b_2	0.00399	0.00111	0.0037	0.0005
n=50	b_0	0.00075	0.00144	0.00233	0.00164
	b_1	0.00022	0.00032	0.00005	0.0014
	b_2	0.00042	0.00005	0.00057	0.00005
n=100	b_0	0.00366	0.00561	0.03664	0.02887
	b_1	0.03157	0.06249	0.0622	0.05689
	b_2	0.00298	0.00597	0.00402	0.00861

جدول (5) إيجاد MSE للمعلمات

حجم العينة	المعلمات	$\alpha=1, \beta=5$			
		Huber	Exponential	Andrews	Ramsy
n=10	b_0	0.24088	0.00075	0.22139	0.00006
	b_1	0.1298	0.00109	0.14099	0.00015
	b_2	0.00484	0.00027	0.0054	0.00028
n=50	b_0	0.00475	0.00408	0.01188	0.00613
	b_1	0.00001	0.00004	0.00307	0.00035

	b_2	0.00054	0.00022	0.00101	0.00026
n=100	b_0	0.01117	0.01102	0.08496	0.02446
	b_1	0.07501	0.09112	0.14462	0.0652
	b_2	0.00524	0.0074	0.00732	0.00917

جدول (6) إيجاد MSE للمعلمات

حجم العينة	المعلمات	$\alpha = 3 , \beta = 5$			
		Huber	Exponential	Andrews	Ramsy
n=10	b_0	0.00312	0.00348	0.00401	0.00849
	b_1	0.003	0.01974	0.0459	0.00256
	b_2	0.00174	0.0038	0.00452	0.00056
n=50	b_0	0.02859	0.01265	0.04896	0.05809
	b_1	0.01172	0.00809	0.02034	0.02952
	b_2	0.00049	0.00023	0.0005	0.0007
n=100	b_0	0.15223	0.00634	0.08222	0.03071
	b_1	0.15315	0.00452	0.11252	0.01136
	b_2	0.00737	0.00041	0.00623	0.00061

جدول (7) إيجاد MSE للمعلمات

حجم العينة	المعلمات	$\alpha = 5 , \beta = 1$			
		Huber	Exponential	Andrews	Ramsy
n=10	b_0	0.05526	0.0142	0.05136	0.03178
	b_1	0.03716	0.01188	0.03757	0.02097
	b_2	0.00198	0.00101	0.00203	0.00139
n=50	b_0	0.01777	0.00247	0.04683	0.00271

	b_1	0.00448	0.00014	0.00441	0.00001
	b_2	0.00008	0.00062	0.00008	0.00033
n=100	b_0	0.04634	0.01858	0.0475	0.03404
	b_1	0.00308	0.00009	0.00299	0.0036
	b_2	0.00015	0.00031	0.00016	0.00008

جدول (8) إيجاد MSE للمعلمات

حجم العينة	المعلمات	$\alpha = 5 , \beta = 3$			
		Huber	Exponential	Andrews	Ramsy
n=10	b_0	0.01498	0.01048	0.00501	0.00232
	b_1	0.00168	0.0023	0.00035	0.01018
	b_2	0.00025	0.00011	0.00065	0.0051
n=50	b_0	0.28026	0.0016	0.20374	0.00014
	b_1	0.08285	0.00001	0.06337	0.00093
	b_2	0.00039	0.00037	0.00026	0.00057
n=100	b_0	0.00848	0.00023	0.00221	0.00012
	b_1	0.01918	0.05995	0.02823	0.0399
	b_2	0.00234	0.00552	0.00273	0.00404

جدول (9) إيجاد MSE للمعلمات

حجم العينة	المعلمات	$\alpha = 3 , \beta = 1$			
		Huber	Exponential	Andrews	Ramsy
n=10	b_0	0.00123	0.00001	0.00693	0.00002
	b_1	0.00229	0.00011	0.01024	0.00018
	b_2	0.00071	0.0004	0.00125	0.00021

n=50	b_0	0.01211	0.00404	0.01479	0.00396
	b_1	0.00169	0.00008	0.00297	0.00005
	b_2	0.00019	0.00034	0.00013	0.00035
n=100	b_0	0.00241	0.00226	0.00546	0.00005
	b_1	0.00155	0.00351	0.00305	0.00335
	b_2	0.00016	0.00003	0.00011	0.00001

4- الاستنتاجات:

1- في حالة $\alpha = 1 , \beta = 1$

عندما يكون حجم العينة (n=10) تكون دالة Douple Exponential أفضل من دالة Ramsy ودالة Huber ودالة Andrews في تقدير أنموذج الانحدار الخطي المتعدد, أما عندما تكون (n=50,100) فان دالة Huber تكون أفضل من الدوال الأخرى.

2- في حالة $\alpha = 3 , \beta = 3$

عندما يكون حجم العينة (n=10) تكون دالة Douple Exponential أفضل من الدوال الأخرى في تقدير أنموذج الانحدار الخطي المتعدد أما عندما تكون (n=50) فان دالة Huber تكون أفضل من الدوال الأخرى وعندما تصبح (n=100) تكون دالة Andrews هي الأفضل.

3- في حالة $\alpha = 5 , \beta = 5$

عندما يكون حجم العينة (n=10) تكون دالة Huber أفضل من الدوال الأخرى في تقدير أنموذج الانحدار الخطي المتعدد أما عندما تكون (n=50) فان دالة Andrews تكون أفضل من الدوال الأخرى وعندما تصبح (n=100) تكون دالة Ramsy هي الأفضل.

4- في حالة $\alpha = 1 , \beta = 3$

عندما (n=10,50) فان دالة Ramsy أفضل من الدوال الأخرى وعند n=100 فان دالة Huber هي الأفضل .

5- في حالة $\alpha = 1 , \beta = 5$

عندما (n=10,50) فان دالة Douple Exponential أفضل من الدوال الأخرى وعند n=100 فان دالة Huber هي الأفضل .

6- في حالة $\alpha = 3 , \beta = 5$

عندما (n=10) فان دالة Ramsy أفضل من الدوال الأخرى وعند (n=50,100) فان دالة Douple Exponential تكون أفضل .

7- في حالة $\alpha = 5 , \beta = 1$

عندما (n=10) فان دالة Douple Exponential أفضل من الدوال الأخرى وعند n=50 فان دالة Andrews تكون أفضل وعند n=100 فان دالة Ramsy هي الأفضل .

8- في حالة$\alpha = 5 , \beta = 3$

عندما (n=10) فان دالة Douple Exponential أفضل من الدوال الأخرى وعند n=50 فان دالة Andrews تكون أفضل وعند n=100 فان دالة Huber هي الأفضل .

9- في حالة $\alpha = 3 , \beta = 1$

عندما (n=10,100) فان دالة Ramsy أفضل من الدوال الأخرى وعند n=50 فان دالة Andrews تكون أفضل .

5- التوصيات:

1- نوصي باستخدام طريقة M الحصينة في حالة التوزيع اللوجستي.
2- استخدام دالة Huber عند تطبيق طريقة M الحصينة عندما يكون حجم العينة كبير.
3- استخدام دالة Douple exponential عند تطبيق طريقة M الحصينة عندما يكون حجم العينة صغير.

6- المصادر:

1- العزاوي, احمد ذياب (2005)"المقارنة بين بعض طرائق تقدير أنموذج انحدار اللوجستك والطرائق الحصينة للتجارب الحياتية ذات الاستجابة الثنائية باستخدام أسلوب المحاكاة"رسالة ماجستير, كلية الإدارة والاقتصاد, جامعة بغداد.
2- عبد الجبار, انعام عبد الوهاب (2002)"دراسة مقارنة مقدرات طريقتي R,M الحصينتين مع مقدرات المربعات الصغرى" المجلة العراقية للعلوم الاحصاءية, العدد 4, المجلد 2 .
3- هرمز, أمير حنا (1990)"الإحصاء الرياضي"دار الكتب للطباعة والنشر, الموصل.
4-Huber, P.J. (1981) "Robust Statistics" New York, John Wiley & Sons.