

\*

Ordinary Least )

(Squares Method

( Multiple Liner Regression Model& Regression Quantile)

.Mse

### **Using Information of population in estimating parameters of multiple regression Models based on Quantile regression with Application**

#### **Abstract**

Is in this research the ordinary least squares method is reviewed and a brief account on the multiple linear regression model and regression quintile is given and then study the characteristics of the society and extracting information from them to add to the information data to get the multiple regression model which has less error than original after the addition of information of population, using regression quantile and comparison between the results using criterion Mse.

---

\* مدرس مساعد/ قسم الاحصاء والمعلوماتية/كلية علوم الحاسوب والرياضيات

[Mahmood81\\_tahr@yahoo.com](mailto:Mahmood81_tahr@yahoo.com)

تاريخ التسلم: 2011/ 1/ 31 — تاريخ القبول : 2011/ 3/13

( Regression Analysis )

)

Response variable (

[(1987) ] explanatory Variables

**Quantile )**

. **(Regression**

( )

**(Regression Quantile)**

( )

**Simple Linear Regression**

**-1**

**( Linear and non-linear regression)**

.( )

**( Simple Liner Regression Model )**

[Samprit &amp; Ali ,(2006)] [(1987) ]:

$$Y_i = \beta_0 + \beta_1 X + \varepsilon_i \quad \dots\dots\dots(1)$$

:  
:  $Y_i$   
:  $X$   
:  $\beta_0, \beta_1$   
:  $\varepsilon_i$

**(Multiple Liner Regression Model)**

[David,(2009)]:

$$\underline{Y} = X \underline{\beta} + \underline{U} \quad \dots\dots\dots(2)$$

:  $\underline{Y}$   
:  $X$   
:  $\underline{\beta}$   
:  $\underline{U}$

[David A,(2009)]:

$$\underline{\hat{y}} = x \underline{\hat{\beta}} \quad \dots\dots\dots(3)$$

 $\hat{y}$ **Ordinary Least Square****-2****Method (OLS)**

] [(1987) ]

: [(2002)

$$\min \sum_{i=1}^n (Y_i - \hat{Y}_i)^2 \dots\dots\dots(4)$$

(OLS) (2)

: [Samprit & Ali ,(2006)] [(1987) ]

$$\underline{Y} = \underline{X}\underline{\beta} + \underline{U}$$

$$\underline{U} = \underline{Y} - \underline{X}\underline{\beta}$$

$$\underline{U}'\underline{U} = (\underline{Y} - \underline{X}\hat{\underline{\beta}})'(\underline{Y} - \underline{X}\hat{\underline{\beta}})$$

$$\frac{\partial \underline{U}'\underline{U}}{\partial \underline{\beta}} = \frac{\partial}{\partial \underline{\beta}} [(\underline{Y} - \underline{X}\hat{\underline{\beta}})'(\underline{Y} - \underline{X}\hat{\underline{\beta}})] = 0 \dots\dots\dots(5)$$

$$\underline{\beta} \quad (5)$$

$$- 2\underline{X}'\underline{Y} + 2(\underline{X}'\underline{X})\hat{\underline{\beta}} = 0$$

:

$$(\underline{X}'\underline{X})\hat{\underline{\beta}} = (\underline{X}'\underline{Y})$$

$$(\underline{X}'\underline{X})^{-1}$$

[David et al,(2003)] [Sanford,(2005)]:

$$\hat{\underline{\beta}}_{ols} = (\underline{X}'\underline{X})^{-1} \underline{X}'\underline{Y} \dots\dots\dots( 6)$$

$$\begin{matrix} (m+1)(m+1) & (\underline{X}'\underline{X}) \\ \underline{X} & \end{matrix}$$

**The quantile function & quantile regression**

**The quantile function**

**Quantile Regression**

( )

1-3

The cumulative distribution function (CDF)

x

F(x)

x

[ Michael,( 2005) ]:

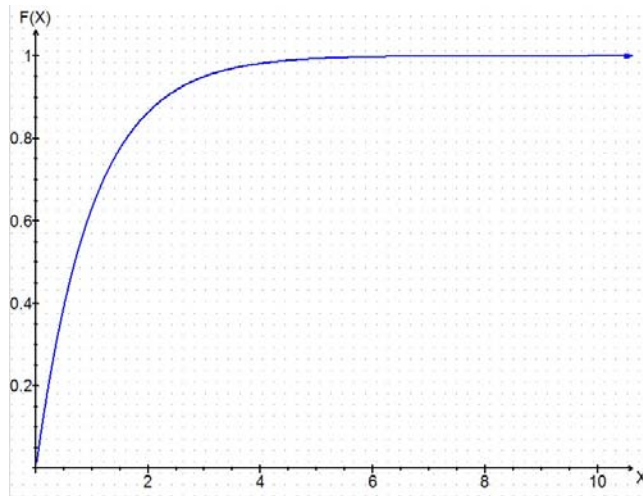
$$F(x) = p (X \leq x)$$

$$F(x) = p$$

p

F(x)

[Jean & Subhabrata,(2003)] :



F(x)

(1)

(1)

( )

**The quantile function (QF)**

**2-3**

$Q(p)$

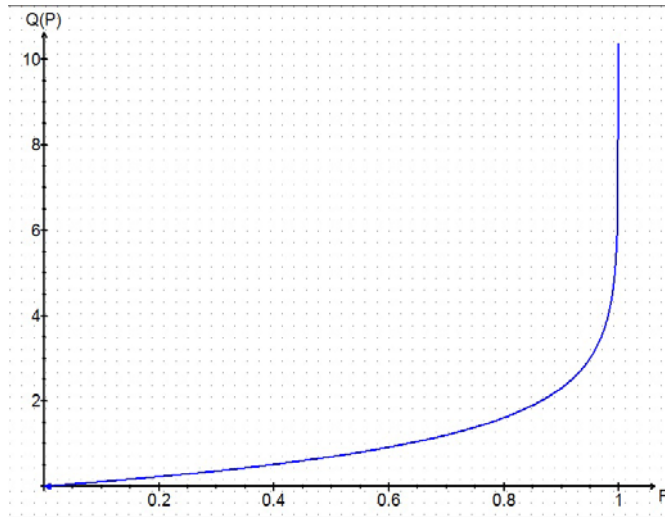
:

$p = p (X \leq x_p)$  عند احتمال  $x$  قيمة  $= x_p$

(p-quantile)  $x_p$

P      p-quantile       $x_p = Q(p)$

[Jean & Subhabrata ,(2003)]:



$QF$       (2)

:

$$F(x) = Q^{-1}(x) \qquad Q(p) = F^{-1}(p)$$

**Quantile Regression**

**3-3**

(2)

$y$

**(Multiple Linear Regression Model)**

.  $x$

[ Alexander and Christof ,(1998)]

$$C = \min \sum (y_i - \hat{y}_i)^2$$

ولإيجاد حل لمسألتنا أعلاه بعملية اشتقاق الدالة C بدلالة β ثم بمساواتها

[

. [ Warren ,(2000)] Samprit and Ali ,(2006)]

(2)

$$Q_y(p/x) = x\beta + \eta S(p) \dots\dots\dots(7)$$

(7)

y ( ) regression quantile function

x

x

(7)

S(p)

[Warren ,(2000)]

$$Q_y(p/x) = X\beta + \sigma N(p)$$

$\hat{y}$

(3)

[ David ,(2009)]:

$$e = y - x\hat{\beta}$$

$$e = y - \hat{y} \dots\dots\dots(8)$$

rank

[Warren ,(2000)]

:

$$p^* = BETAINV(0.5, n+1-r_i, r_i) \dots\dots(9)$$

(9)

$$N(p_i^*) = (p_i^*, 0, 1) \dots\dots\dots(10)$$

Excel (10) (9)

[ Warren ,(2000)] : 2003

$$\hat{y}_{M_i} = x\hat{\beta} + \hat{\sigma}N(p_i^*) \dots\dots\dots(11)$$

:  
:  $\hat{y}_{M_i}$   
:  $x\hat{\beta}$   
:  $\hat{\sigma}$   
:  $N(p_i^*)$

( )

( )

(11)  
.(  $e^*$ ) (8)

$$\underline{e}^* = \underline{y} - \hat{y}_{M_i} \dots\dots\dots(12)$$

Mean Square Error : - 4

[ David,( 2009)]

$$MSE = \hat{\sigma}^2 = \frac{\sum_{i=1}^n e^2_i}{n-p-1}$$



[ Samprit and Ali,(2006)]:

$$MSE = \hat{\sigma}^2 = \frac{e'e}{n - m - 1} \dots\dots\dots(13)$$

$$e'e = y'y - \hat{\beta}x'y$$

:  $y'y$

:  $\hat{\beta}$

:  $x'y$

: n

: m

-4

Excel 2007

Mse

Minitab15

-  
2008

-

:  
:  
:  
:  
:

100

(1)

(1)

$Y_i$	$X_1$	$X_2$	$X_3$	$X_4$
2.21	58	85	59	891
2.92	56	88	71	1018
2.8	54	87	66	1042
2.81	60	87	63	918
3.43	59	87	84	983
3.3	57	88	78	998
1.95	57	82	61	845
2.73	55	84	70	914
2.58	52	85	70	937

(6)

$$\hat{\beta}_{ols} = (X'X)^{-1} X'Y$$

$$\hat{\beta} = \begin{bmatrix} 9 & 508 & 773 & 622 & 8546 \\ 508 & 28724 & 43639 & 35111 & 482076 \\ 773 & 43639 & 66425 & 53490 & 734895 \\ 622 & 35111 & 53490 & 62681 & 592984 \\ 8546 & 482076 & 734895 & 592984 & 8148376 \end{bmatrix}^{-1} \begin{bmatrix} 24.73 \\ 1397.43 \\ 2130.21 \\ 1735.71 \\ 23664.37 \end{bmatrix} = \begin{bmatrix} -9.67 \\ 0.0161 \\ 0.101 \\ 0.0365 \\ 0.00032 \end{bmatrix}$$

(3)

 $\hat{y}$

$$\hat{y} = \begin{bmatrix} 2.28742 \\ 3.03686 \\ 2.72884 \\ 2.67626 \\ 3.44746 \\ 3.30206 \\ 2.0266 \\ 2.54698 \\ 2.60704 \end{bmatrix}$$

Mean Square

 $\hat{y}$ 

( 13 )

Error

$$MSE = 0.02061$$

 $\hat{y}$ 

-

(12)

$$e = \begin{bmatrix} -0.07742 \\ -0.11686 \\ 0.07116 \\ 0.13374 \\ -0.01746 \\ -0.00206 \\ -0.0766 \\ 0.18302 \\ -0.02704 \end{bmatrix}$$

 $p^*$ 

(9)

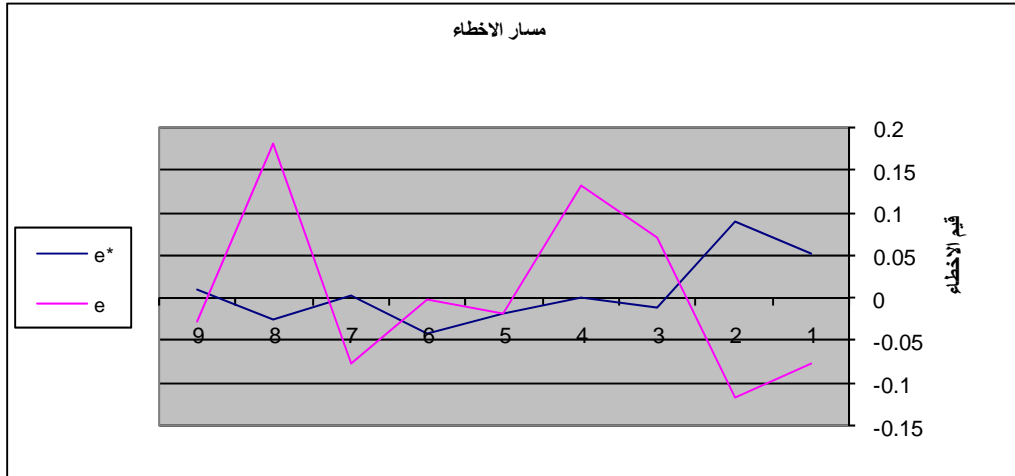
$$p^* = \begin{bmatrix} 0.17961955 \\ 0.07412529 \\ 0.71376324 \\ 0.82038045 \\ 0.50000000 \\ 0.60691524 \\ 0.28623676 \\ 0.92587471 \\ 0.39308476 \end{bmatrix} N(P^*) \quad (10)$$

$$N(P^*) = \begin{bmatrix} -0.91681583 \\ -1.445738733 \\ 0.564412252 \\ 0.91681583 \\ 5.47142E-10 \\ 0.271288163 \\ -0.564412252 \\ 1.445738733 \\ -0.271288163 \end{bmatrix} \hat{y} \quad (11)$$

$$\hat{y}_{M_i} = \begin{bmatrix} 2.28742 \\ 3.03686 \\ 2.72884 \\ 2.67626 \\ 3.44746 \\ 3.30206 \\ 2.0266 \\ 2.54698 \\ 2.60704 \end{bmatrix} + 0.1435618334 \begin{bmatrix} -0.916815934 \\ -1.445738435 \\ 0.564412365 \\ 0.916815934 \\ -1.39214E-16 \\ 0.271288008 \\ -0.564412365 \\ 1.445738435 \\ -0.271288008 \end{bmatrix} = \begin{bmatrix} 2,155800224 \\ 2,82930714 \\ 2,809868074 \\ 2,807879776 \\ 3,44746 \\ 3,341006604 \\ 1,945571926 \\ 2,75453286 \\ 2,568093396 \end{bmatrix}$$

$$MSE = 0.00308 \quad (13)$$

Mean Square Error



( 3 )

$$e^* \quad (3)$$

$e$

-0.11       $e$       0.09      -0.04       $e^*$

0.18

MSe

...

-5

-1

( 0.00308 )

( 0.02061)

-2

-3

-6

"

(1987)

-1

(2002)

-2

"

- 1-Alexander von Eye & Christof Schuster,(1998), "Regression Analysis for Social Sciences ",Academic Press.
- 2- David A. Freedman,(2009)," Statistical Models Theory and Practice", Cambridge, New York.
- 3-David Ruppert &M. P. Wand & R. J. Carroll,(2003)," Semiparametric Regression ",Cambridge, New York.
- 4-Jean Dickinson Gibbons & Subhabrata Chakraborti, (2003), "Nonparametric Statistical Inference" Fourth Edition, Marcel Dekker, INC, New York.
- 5- Michael W. Trosset,(2005),"An Introduction to Statistical Inference and Its Applications ",Chapman & Hall/CRC Texts in Statistical Science.
- 6-Samprit Chatterjee & Ali S.Hadi,(2006)," Regression Analysis by Example ",Fourth Edition , John Wiley & Sons.
- 7- Sanford Weisberg,(2005)," Applied Linear Regression ",Third Edition, John Wiley & Sons.
- 8-Warren G,(2000),"Statistical Modelling with Quantile Functions",Chapman& Hall/CRT.



## (الملحق (2)

## تحليل بيانات التجربة الأصلية باستخدام برنامج أل Minitab 15

**Regression Analysis: Yi versus X1; X2; X3; X4**

The regression equation is

$$Y_i = -9.67 + 0.0161 X_1 + 0.101 X_2 + 0.0365 X_3 + 0.00032 X_4$$

Predictor	Coef	SE Coef	T	P
Constant	-9.670	3.142	-3.08	0.037
X1	0.01608	0.03100	0.52	0.631
X2	0.10114	0.06913	1.46	0.217
X3	0.036504	0.007781	4.69	0.009
X4	0.000317	0.002307	0.14	0.897

$$S = 0.143573 \quad R\text{-Sq} = 95.3\% \quad R\text{-Sq}(\text{adj}) = 90.6\%$$

## Analysis of Variance

Source	DF	SS	MS	F	P
Regression	4	1.67830	0.41958	20.35	0.006
Residual Error	4	0.08245	0.02061		
Total	8	1.76076			

Source	DF	Seq SS
X1	1	0.04839
X2	1	1.11256
X3	1	0.51696
X4	1	0.00039