

統計方法期末報告

-以每年參觀 Lake Keepit 的人數為例



指導老師：蔡碧紋

學生：林柏佐

系級：數學系 98 級

班別：四年乙班

學號：494402345

```

>
lake=read.table("D:lake.txt",header=T)
> lake
#Dist=Distance, Inc=Family Income,
Size=Family Members,
Y=Numbers of Vistors
  Dist  Inc Size  Y
1    27  4.45   5  1
2    72  7.69   4  1
3    44 10.04   4  5
4    23  8.97   4  4
5    89  9.15   5  2
6    58  5.80   5  5
7    84  4.94   4  2
8    62  5.05   1  1
9    10  2.89   4  8
10   77  4.21   3  0
11   91  2.93   3  0
12   72  5.46   5  0
13   54  7.98   2  3
14   34  7.39   1  6
15   94  4.71   4  1
16   98  6.55   4  0
17   90  3.61   3  0
18   63 10.42   4  0
19   19  7.36   2  5
20   34  2.67   3  1
21   45  3.76   4  5
22   91  8.71   5  1
23   40  6.32   3  3
24   78  5.64   2  4
25   76  7.51   2  0
26   64  7.24   5  5
27   62  9.13   1  1
28   20  7.98   4  3
29  119  3.61   5  0
30   50  3.21   3  2
31   81  6.55   1  1
32   38  2.66   2  3
33   29  6.68   4  3
34  100  5.60   1  1
35   78  9.69   5  0
36   46  7.82   5  4
37   56  7.34   2  0
38   34  5.39   1  2
39   63  8.36   5  4
40   34  5.95   1  1
41  120  6.86   1  1
42   85  9.50   2  1
43   25  8.00   2  2
44   37  8.29   5  8
45   21  8.57   4  7
46   88  6.29   3  0
47   66  5.08   1  2
48   48  5.82   2  1
49   90  7.39   2  1
50   57  6.61   4  2
51  109  9.88   2  0
52   57  7.96   3  2
53   19  3.86   2  1
54   79  3.15   3  0
55   30  9.64   3  5
56   48  3.56   2  3
57   41  6.18   2  0
58   49  5.79   2  3
59   59  6.89   2  0
60   72  2.74   2  2
61   76  8.85   5  3
62  110  9.83   4  0
63   47  6.38   1  0
64   88  7.47   3  0
65   10  5.39   4  3
66   21 10.54   5  7
67   64 10.39   3  2
68   46  7.89   4  4
69   44  5.40   2  4

```

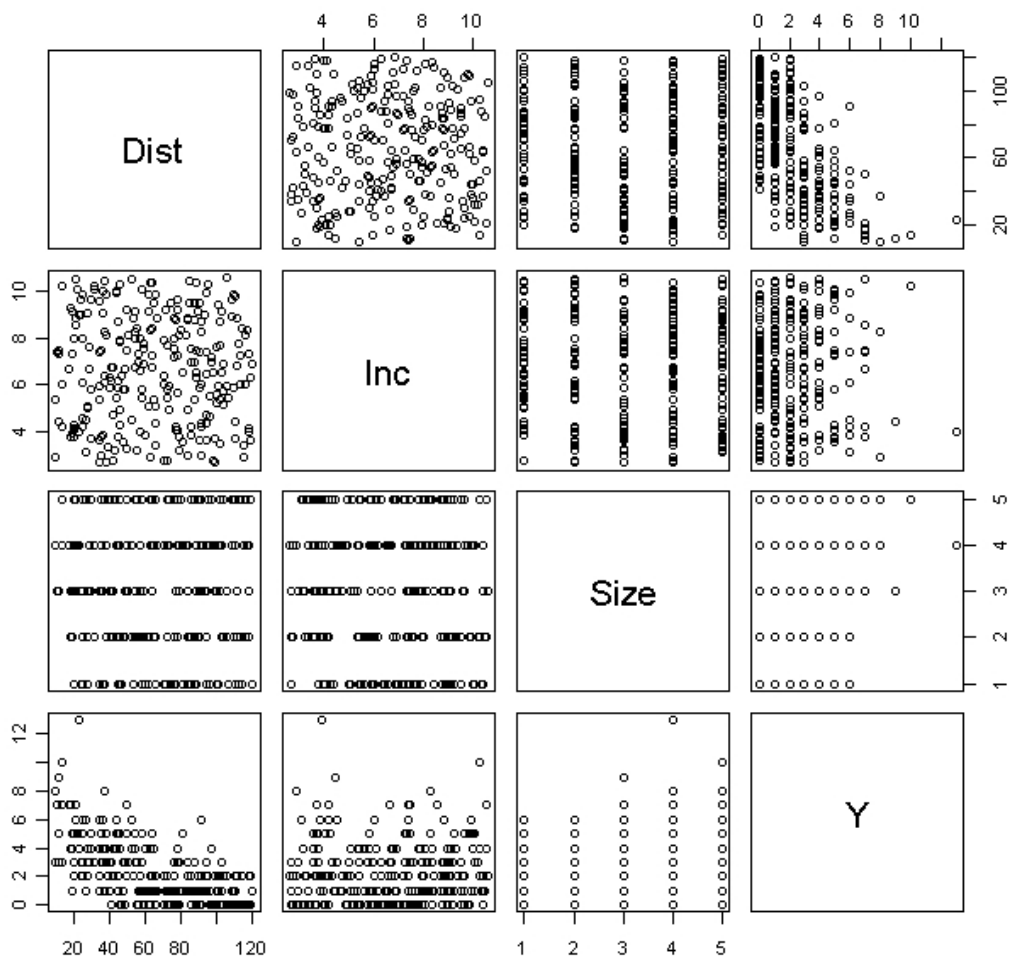
70	79	3.74	4	3	108	37	6.09	5	4
71	116	6.02	2	0	109	65	9.55	2	1
72	84	2.93	3	1	110	50	7.46	4	7
73	95	10.29	5	2	111	34	3.60	3	4
74	17	6.68	4	7	112	108	9.68	4	1
75	98	2.76	4	0	113	20	3.78	5	5
76	91	4.21	5	6	114	77	8.77	3	0
77	86	3.81	5	2	115	86	7.53	2	1
78	118	3.95	3	0	116	106	5.45	1	0
79	85	10.50	2	1	117	20	4.08	5	2
80	87	6.93	1	1	118	29	7.61	3	3
81	53	8.89	1	2	119	102	8.68	4	2
82	70	2.69	1	0	120	45	8.73	2	2
83	44	9.94	2	6	121	87	9.52	1	3
84	54	6.77	5	4	122	18	3.72	3	7
85	42	6.73	3	4	123	96	4.63	4	1
86	27	3.07	5	6	124	38	8.10	4	4
87	62	9.90	2	0	125	54	9.13	3	4
88	57	3.96	2	1	126	42	2.75	2	2
89	117	8.33	4	0	127	105	10.59	3	2
90	101	5.97	1	1	128	12	4.44	3	9
91	94	7.42	3	2	129	74	9.53	1	1
92	46	4.25	1	5	130	108	8.96	5	0
93	84	8.84	2	2	131	12	7.33	3	5
94	20	4.02	1	5	132	90	5.88	3	0
95	102	5.33	3	0	133	28	9.08	2	2
96	50	9.24	5	5	134	102	3.42	5	0
97	52	10.53	2	4	135	45	5.87	1	1
98	52	4.51	4	6	136	45	6.15	4	1
99	100	5.18	3	1	137	112	7.77	1	0
100	84	8.77	1	3	138	83	8.10	4	1
101	35	7.45	5	5	139	35	2.96	3	3
102	78	9.20	4	0	140	80	8.34	5	1
103	90	9.50	5	0	141	104	5.13	5	0
104	97	7.19	5	0	142	78	7.24	4	3
105	49	6.69	3	0	143	55	3.19	2	0
106	43	3.18	5	4	144	25	4.20	5	2
107	96	5.48	1	0	145	78	4.02	4	0

146	87	8.78	5	3	184	94	6.45	1	1
147	111	8.85	3	2	185	24	9.64	3	2
148	36	10.01	1	4	186	73	5.92	4	1
149	116	8.12	5	0	187	25	6.06	1	6
150	36	3.84	3	6	188	91	4.31	5	2
151	104	6.88	5	2	189	81	3.54	4	5
152	60	3.79	1	4	190	118	7.32	4	0
153	114	8.42	1	2	191	86	4.20	2	0
154	56	6.95	2	1	192	86	6.98	1	0
155	66	4.33	5	2	193	97	6.93	5	4
156	65	5.34	4	4	194	41	6.72	3	3
157	105	5.95	2	0	195	77	4.01	1	0
158	66	9.90	2	0	196	111	6.81	2	0
159	94	9.96	4	1	197	70	7.94	1	0
160	113	6.03	2	0	198	66	3.36	2	1
161	93	9.04	3	1	199	30	3.69	5	5
162	51	8.13	3	3	200	115	3.38	4	0
163	87	5.35	2	1	201	12	7.46	3	3
164	22	10.09	4	5	202	100	7.61	1	0
165	46	4.88	4	0	203	71	9.22	1	1
166	99	2.63	4	2	204	37	10.43	1	1
167	14	5.99	4	3	205	35	3.78	4	3
168	14	10.25	5	10	206	110	3.68	5	2
169	89	9.14	2	1	207	76	8.01	5	1
170	103	6.55	4	1	208	119	6.30	5	2
171	20	4.24	3	5	209	70	6.39	1	0
172	32	10.21	2	3	210	104	5.74	2	0
173	109	5.90	5	1	211	38	9.91	2	5
174	55	8.39	5	3	212	70	6.04	4	2
175	94	3.61	2	2	213	88	6.51	4	1
176	64	8.40	3	1	214	30	5.82	5	4
177	110	4.41	1	1	215	76	7.18	1	2
178	99	3.69	3	0	216	73	4.91	5	0
179	36	9.05	1	3	217	54	8.27	5	5
180	46	10.31	2	4	218	28	5.09	3	4
181	115	2.87	4	0	219	64	6.74	4	2
182	25	9.90	2	5	220	28	4.99	1	3
183	63	2.87	4	2	221	37	6.41	1	1

222	19	9.14	3	4
223	59	7.78	3	1
224	102	7.71	4	0
225	34	9.62	3	2
226	14	4.20	5	7
227	88	10.37	1	1
228	113	7.22	5	0
229	72	5.81	4	1
230	42	4.72	3	2
231	18	9.30	2	4
232	58	5.40	5	3
233	91	5.98	4	0
234	67	8.86	4	1
235	94	4.35	4	3
236	39	3.30	5	2
237	79	8.16	4	1
238	103	9.16	2	3
239	91	8.24	5	1
240	101	9.41	4	2
241	59	7.58	2	0
242	25	4.52	3	3
243	11	7.37	3	7
244	21	8.85	3	6
245	40	5.86	5	1
246	23	3.94	4	13
247	58	10.07	1	1
248	24	8.99	1	3
249	32	10.10	4	3
250	118	4.14	2	2

> attach(lake)

> pairs(lake)



從圖形來看，距離與人數有很高的 correlation

```
> cor(lake)
```

	Dist	Inc	Size	Y
Dist	1.0000000	-0.00382533	0.02006100	-0.62618753
Inc	-0.00382533	1.0000000	-0.08055719	0.01823499
Size	0.02006100	-0.08055719	1.0000000	0.18961580
Y	-0.62618753	0.01823499	0.18961580	1.0000000

```
> fm<-lm(Y~Dist+Inc+Size)
```

```
> summary(fm)
```

Call:

```
lm(formula = Y ~ Dist + Inc + Size)
```

Residuals:

Min	1Q	Median	3Q	Max
-3.4949	-1.0841	-0.1381	1.0107	8.6605

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	3.973939	0.468093	8.490	1.98e-15 ***
Dist	-0.044974	0.003424	-13.136	< 2e-16 ***

```
Inc          0.031041  0.046206  0.672  0.502#無法 reject
Size         0.319426  0.075047  4.256 2.96e-05 ***
```

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.652 on 246 degrees of freedom

Multiple R-squared: 0.434, Adjusted R-squared: 0.4271#太低了

F-statistic: 62.89 on 3 and 246 DF, p-value: < 2.2e-16

> anova(fm)

Analysis of Variance Table

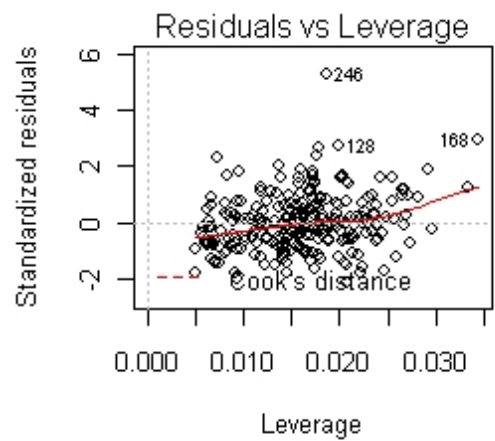
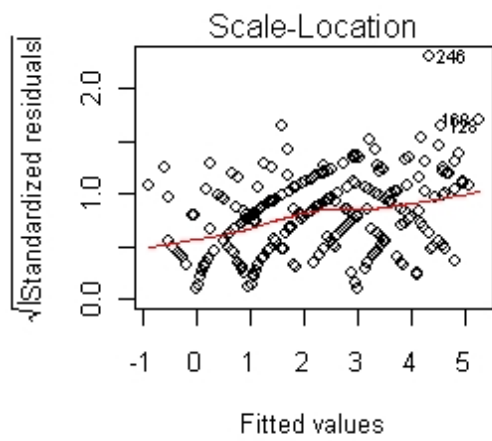
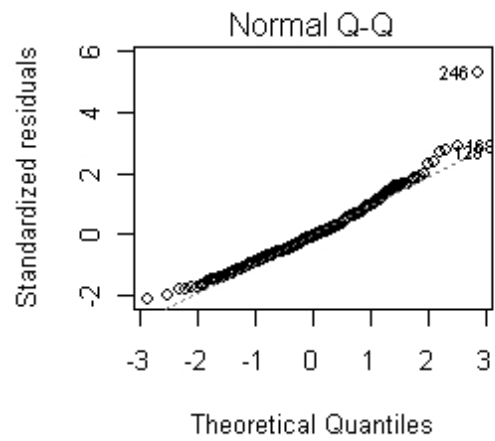
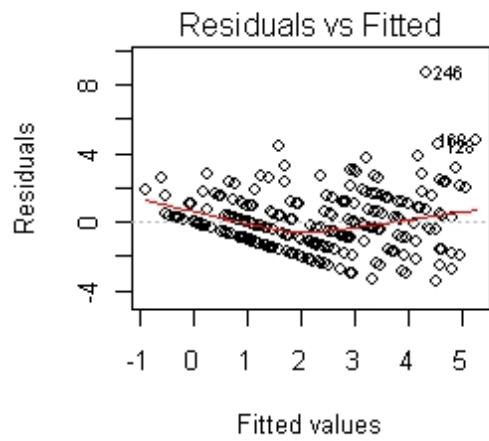
Response: Y

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Dist	1	465.29	465.29	170.4352	< 2.2e-16 ***
Inc	1	0.30	0.30	0.1091	0.7415#還是收入有問題
Size	1	49.46	49.46	18.1165	2.957e-05 ***
Residuals	246	671.58	2.73		

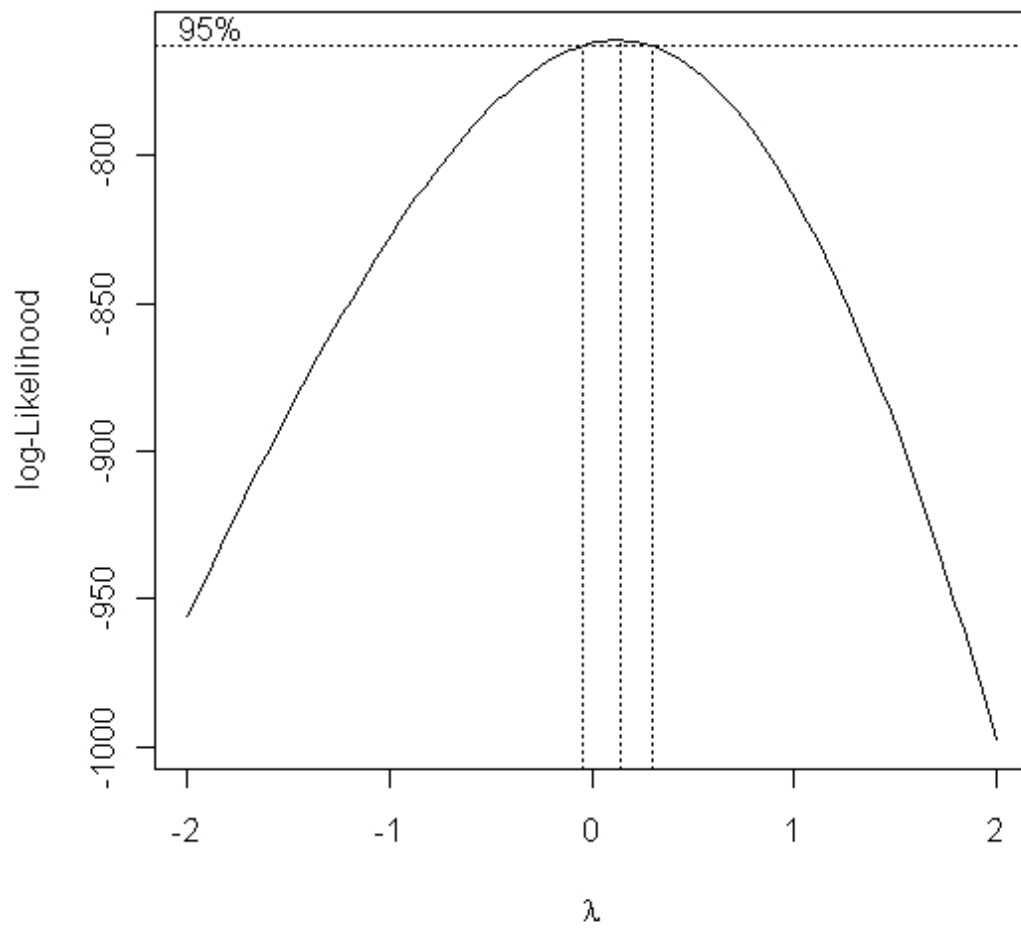
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

> par(mfrow=c(2,2))

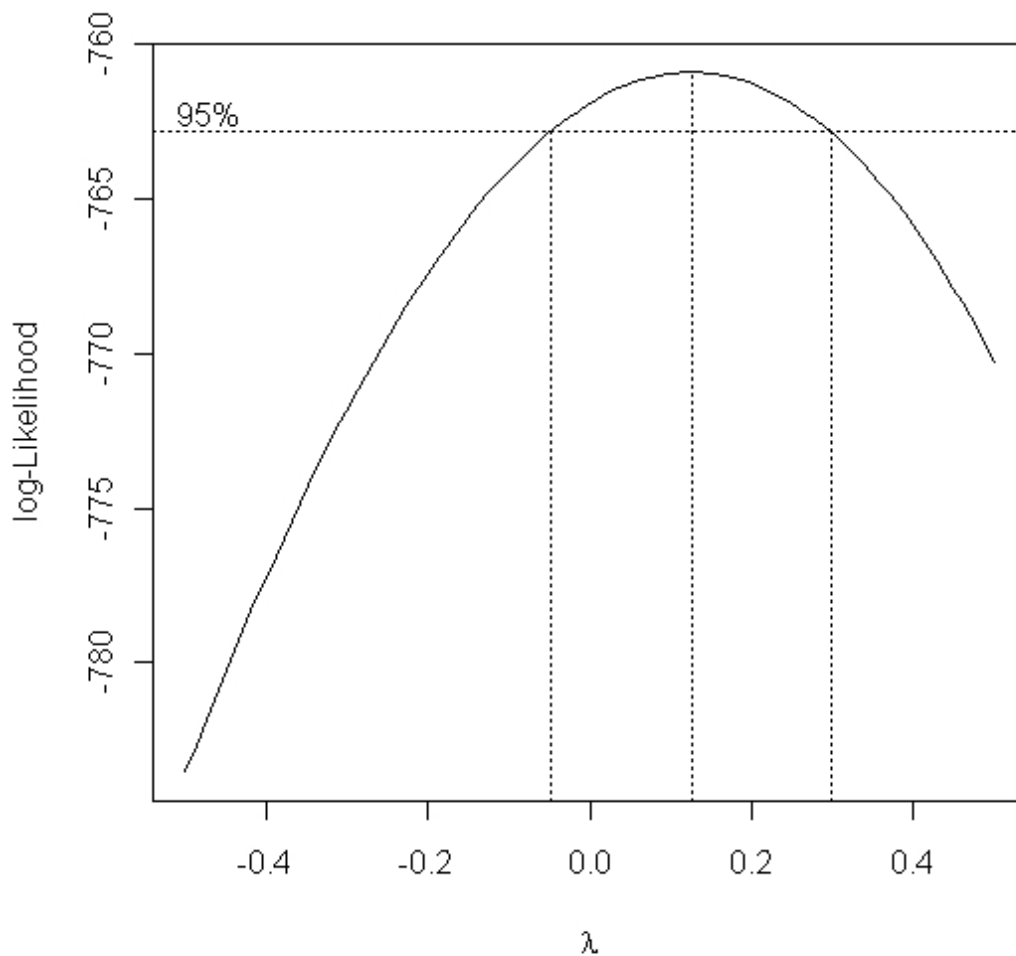
> plot(fm,ask=F)



由圖形來看，residuals 有一個 pattern 在，因此考慮是否要做 transformation
`> boxcox(fm)`



```
> boxcox(fm,lambda=seq(-0.5,0.5,0.1))
```



```
> fm1<-lm(log(Y+1)~Dist+Inc+Size)#在做轉換時，要注意各係數都必須是正數，
因為 Y 有 0，所以我讓其加 1，來做 regression
> summary(fm1)
```

Call:

```
lm(formula = log(Y + 1) ~ Dist + Inc + Size)
```

Residuals:

Min	1Q	Median	3Q	Max
-1.27011	-0.34697	0.04571	0.36514	1.23818

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	1.47987	0.14623	10.120	< 2e-16 ***
Dist	-0.01415	0.00107	-13.234	< 2e-16 ***
Inc	0.01892	0.01443	1.310	0.191285 #還是 Income 有問題
Size	0.08727	0.02345	3.722	0.000245 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.5162 on 246 degrees of freedom
 Multiple R-squared: 0.4335, Adjusted R-squared: 0.4266
 F-statistic: 62.75 on 3 and 246 DF, p-value: < 2.2e-16

```
> anova(fm1)
Analysis of Variance Table
```

Response: log(Y + 1)

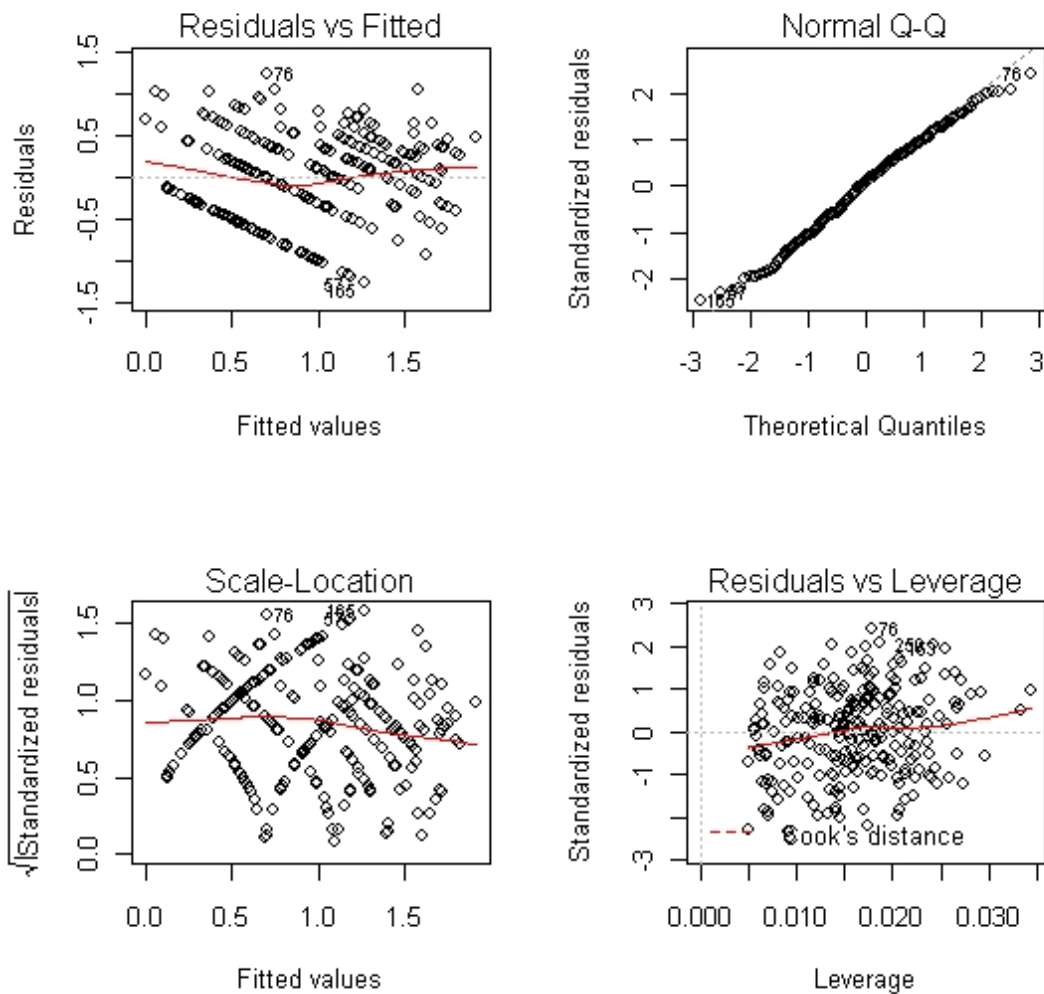
	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Dist	1	46.188	46.188	173.3584	< 2.2e-16 ***
Inc	1	0.274	0.274	1.0283	0.3115555
Size	1	3.692	3.692	13.8554	0.0002447 ***
Residuals	246	65.542	0.266		

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```
> par(mfrow=c(2,2))
```

```
>
```

```
plot(fm1,ask=F)
```



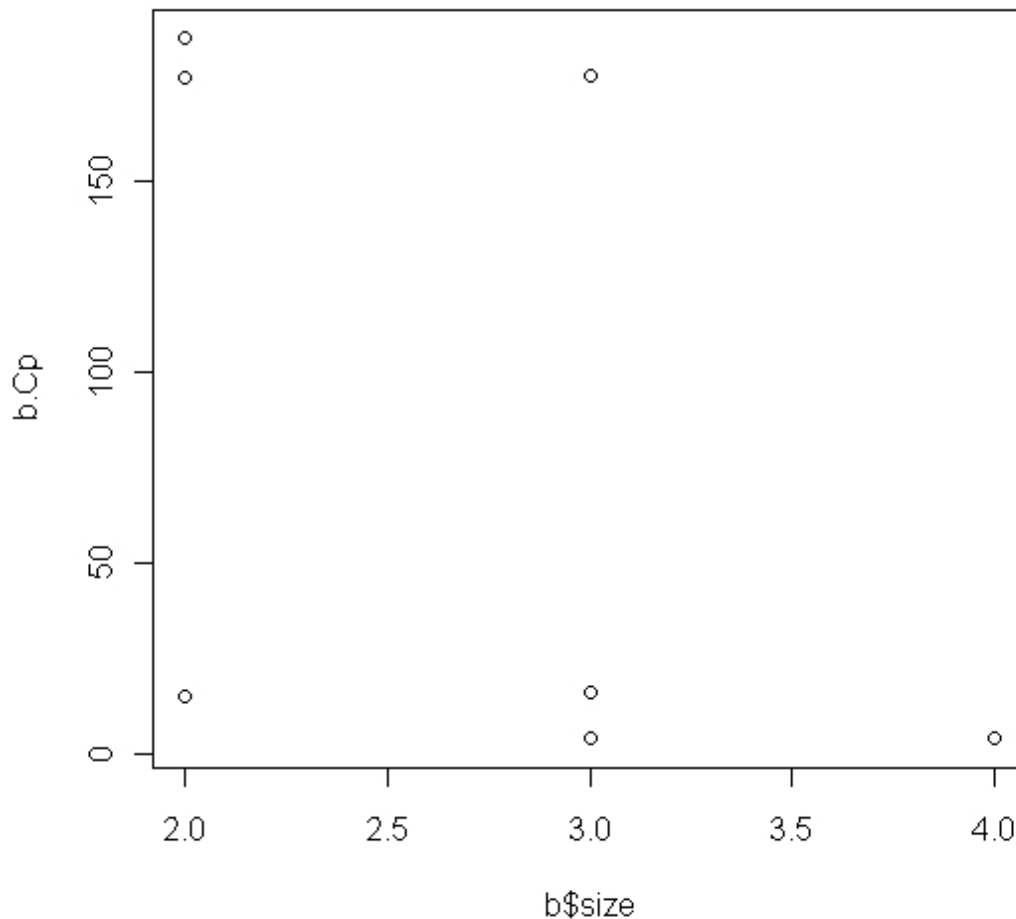
來做 model selection

```
> b=leaps(cbind(Dist,Inc,Size),log(Y+1),method="r2")
```

```

> b.r2=b$r2
> b=leaps(cbind(Dist,Inc,Size),log(Y+1),method="adjr2")
> b.adj2=b$adjr2
> b=leaps(cbind(Dist,Inc,Size),log(Y+1),method="Cp")
> b.Cp=b$Cp
> plot(b$size,b.Cp)

```



```

      1 2 3      b.r2      b.adj2      b.Cp
1 1 0 0 0.399220586 0.396798088 14.883715
1 0 0 1 0.026055915 0.022128721 176.927526
1 0 1 0 0.002609064 -0.001412674 187.109136
2 1 0 1 0.429541430 0.424922332 3.717128
2 1 1 0 0.401588611 0.396743175 15.855419
2 0 1 1 0.030189294 0.022336576 177.132639
3 1 1 1 0.433495741 0.426587152 4.000000
> bs=regsubsets(log(Y+1)~Dist+Inc+Size,data=lake)
> rs=summary(bs)
> names(rs)
[1] "which" "rsq" "rss" "adjr2" "cp" "bic" "outmat"
[8] "obj"
> print(cbind(rs$which,R2=rs$rsq,adjR2=rs$adjr2,Cp=rs$Cp,BIC=rs$bic))

```

```

      (Intercept) Dist Inc Size          R2      adjR2          Cp          BIC
1             1      1  0      0 0.3992206 0.3967981 14.883715 -116.3389
2             1      1  0      1 0.4295414 0.4249223  3.717128 -123.7643
3             1      1  1      1 0.4334957 0.4265872  4.000000 -119.9818
> fm2<-lm(log(Y+1)~Dist+Size)
> summary(fm2)

```

Call:

```
lm(formula = log(Y + 1) ~ Dist + Size)
```

Residuals:

```

      Min          1Q      Median          3Q          Max
-1.30145 -0.33937  0.04727  0.36763  1.19679

```

Coefficients:

```

              Estimate Std. Error t value Pr(>|t|)
(Intercept)  1.613546    0.104927   15.378 < 2e-16 ***
Dist         -0.014158    0.001071  -13.218 < 2e-16 ***
Size          0.084795    0.023403    3.623 0.000353 ***
---

```

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.5169 on 247 degrees of freedom

Multiple R-squared: 0.4295, Adjusted R-squared: 0.4249

F-statistic: 92.99 on 2 and 247 DF, p-value: < 2.2e-16

```
> anova(fm2)
```

Analysis of Variance Table

Response: log(Y + 1)

```

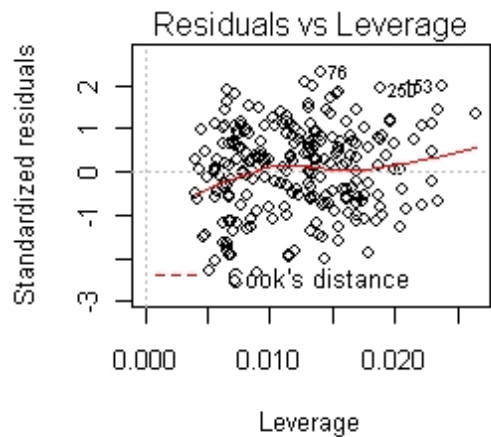
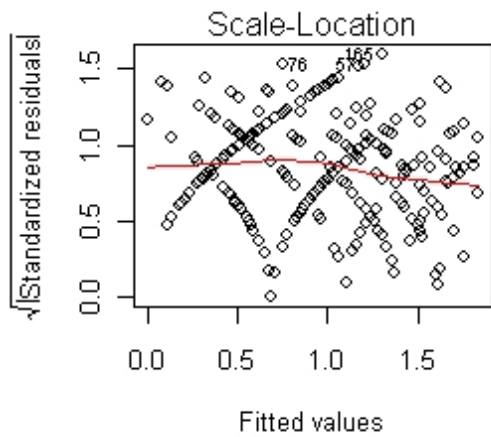
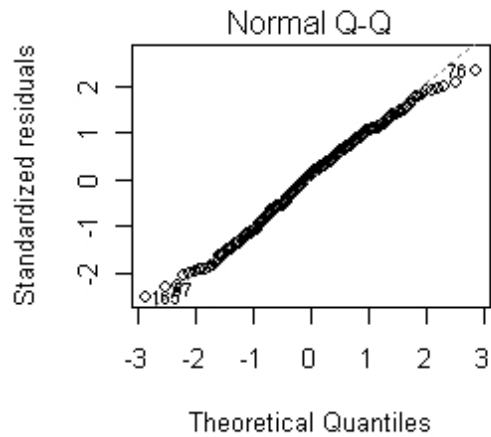
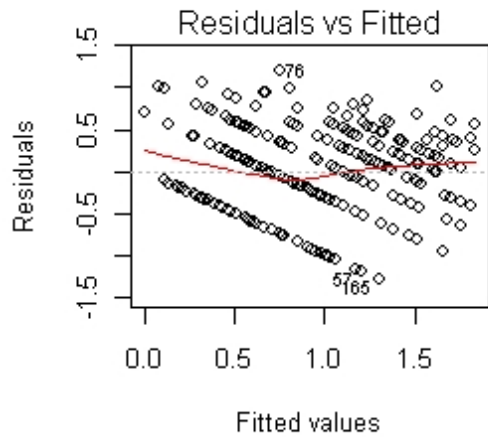
      Df Sum Sq Mean Sq F value    Pr(>F)
Dist   1 46.188  46.188 172.857 < 2.2e-16 ***
Size   1  3.508   3.508  13.129 0.000353 ***
Residuals 247 66.000   0.267
---

```

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```
> par(mfrow=c(2,2))
```

```
> plot(fm2,ask=F)
```



```
> fm3<-lm(sqrt(Y)~Dist+Inc+Size) #variance stable
> summary(fm3)
```

```
Call:
lm(formula = sqrt(Y) ~ Dist + Inc + Size)
```

```
Residuals:
    Min       1Q   Median       3Q      Max
-1.6186 -0.4273  0.0727  0.4570  1.6070
```

```
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  1.884998   0.184333  10.226 < 2e-16 ***
Dist         -0.017528   0.001348 -13.000 < 2e-16 ***
Inc           0.024612   0.018196  1.353  0.17742
Size          0.104933   0.029553  3.551  0.00046 ***
```

```
--
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
Residual standard error: 0.6507 on 246 degrees of freedom
```

Multiple R-squared: 0.424, Adjusted R-squared: 0.417
 F-statistic: 60.36 on 3 and 246 DF, p-value: < 2.2e-16

> anova(fm3)

Analysis of Variance Table

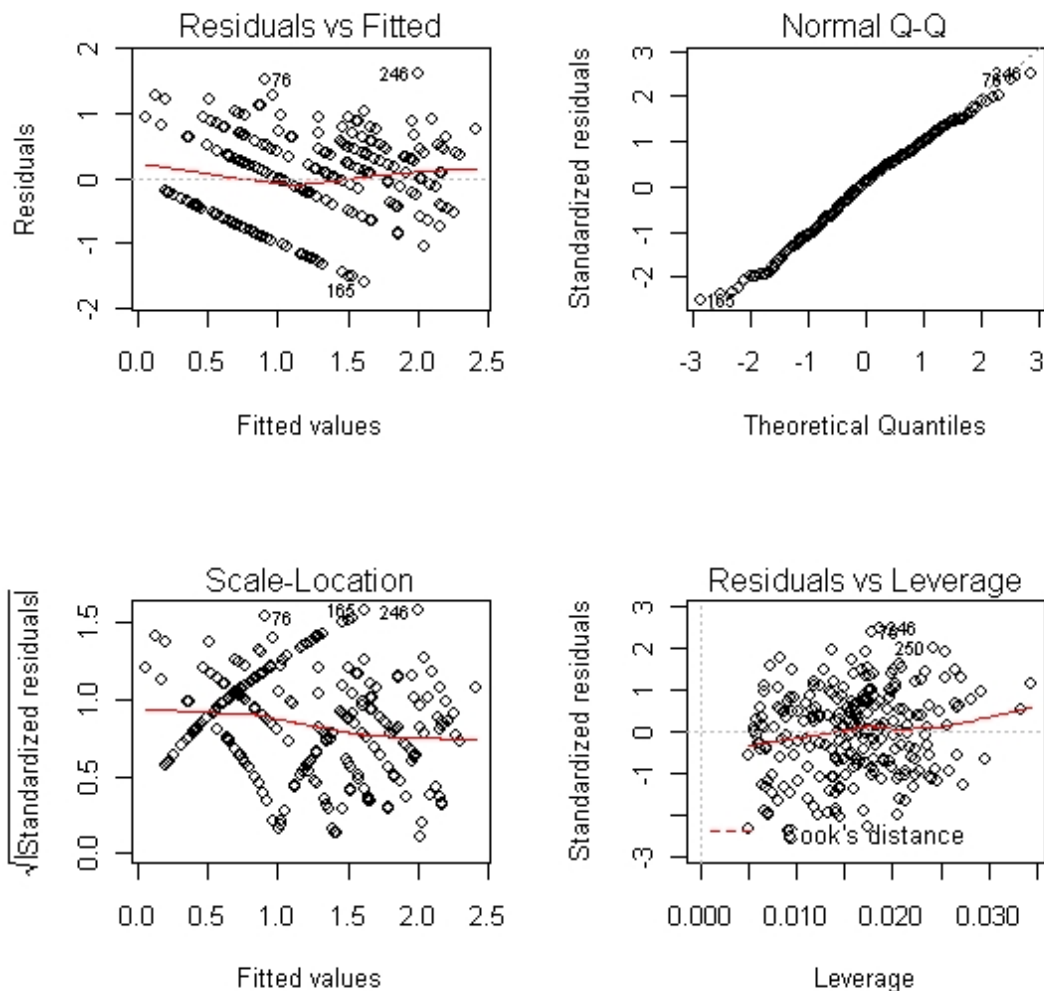
Response: sqrt(Y)

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Dist	1	70.844	70.844	167.3406	< 2.2e-16 ***
Inc	1	0.485	0.485	1.1455	0.2855408
Size	1	5.337	5.337	12.6072	0.0004603 ***
Residuals	246	104.145	0.423		

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

> par(mfrow=c(2,2))

> plot(fm3,ask=F)



> res=residuals(fm3)

> fm4=lm(abs(res)~Dist+Inc+Size)

> wi=1/((fitted(fm4))^2)

> print(cbind(Dist,Inc,Size,Y,res,"abs(rs)"=abs(res),weights=wi),digits=2)

Dist	Inc	Size	Y	res	abs(rs)	weights
------	-----	------	---	-----	---------	---------

1	27	4.4	5	1	-1.0459	1.0459	4.2
2	72	7.7	4	1	-0.2320	0.2320	3.6
3	44	10.0	4	5	0.4554	0.4554	5.4
4	23	9.0	4	4	-0.1224	0.1224	6.0
5	89	9.2	5	2	0.3393	0.3393	3.5
6	58	5.8	5	5	0.7002	0.7002	3.6
7	84	4.9	4	2	0.4602	0.4602	2.9
8	62	5.0	1	1	-0.0275	0.0275	3.3
9	10	2.9	4	8	0.6278	0.6278	4.4
10	77	4.2	3	0	-0.9538	0.9538	2.9
11	91	2.9	3	0	-0.6769	0.6769	2.5
12	72	5.5	5	0	-1.2821	1.2821	3.2
13	54	8.0	2	3	0.3873	0.3873	4.2
14	34	7.4	1	6	0.8736	0.8736	4.8
15	94	4.7	4	1	0.2269	0.2269	2.6
16	98	6.5	4	0	-0.7482	0.7482	2.8
17	90	3.6	3	0	-0.7112	0.7112	2.6
18	63	10.4	4	0	-1.4570	1.4570	4.7
19	19	7.4	2	5	0.2931	0.2931	5.5
20	34	2.7	3	1	-0.6696	0.6696	3.6
21	45	3.8	4	5	0.6275	0.6275	3.5
22	91	8.7	5	1	-0.0290	0.0290	3.4
23	40	6.3	3	3	0.0778	0.0778	4.3
24	78	5.6	2	4	1.1335	1.1335	3.1
25	76	7.5	2	0	-0.9476	0.9476	3.5
26	64	7.2	5	5	0.7700	0.7700	3.8
27	62	9.1	1	1	-0.1279	0.1279	4.3
28	20	8.0	4	3	-0.4185	0.4185	5.7
29	119	3.6	5	0	-0.4127	0.4127	2.2
30	50	3.2	3	2	0.0118	0.0118	3.3
31	81	6.5	1	1	0.2686	0.2686	3.2
32	38	2.7	2	3	0.2378	0.2378	3.5
33	29	6.7	4	3	-0.2288	0.2288	4.8
34	100	5.6	1	1	0.6250	0.6250	2.7
35	78	9.7	5	0	-1.2810	1.2810	3.9
36	46	7.8	5	4	0.2041	0.2041	4.5
37	56	7.3	2	0	-1.2940	1.2940	4.0
38	34	5.4	1	2	-0.1124	0.1124	4.2
39	63	8.4	5	4	0.4888	0.4888	4.1
40	34	6.0	1	1	-0.5404	0.5404	4.4
41	120	6.9	1	1	0.9445	0.9445	2.5
42	85	9.5	2	1	0.1612	0.1612	3.7
43	25	8.0	2	2	-0.4394	0.4394	5.5
44	37	8.3	5	8	0.8632	0.8632	5.0
45	21	8.6	4	7	0.4982	0.4982	5.9
46	88	6.3	3	0	-0.8122	0.8122	3.0
47	66	5.1	1	2	0.4561	0.4561	3.2
48	48	5.8	2	1	-0.3968	0.3968	3.9
49	90	7.4	2	1	0.3007	0.3007	3.1
50	57	6.6	4	2	-0.0541	0.0541	3.8

51	109	9.9	2	0	-0.4275	0.4275	3.2
52	57	8.0	3	2	0.0176	0.0176	4.1
53	19	3.9	2	1	-0.8568	0.8568	4.3
54	79	3.1	3	0	-0.8927	0.8927	2.7
55	30	9.6	3	5	0.3248	0.3248	5.9
56	48	3.6	2	3	0.3909	0.3909	3.4
57	41	6.2	2	0	-1.5283	1.5283	4.2
58	49	5.8	2	3	0.3535	0.3535	3.8
59	59	6.9	2	0	-1.2303	1.2303	3.8
60	72	2.7	2	2	0.5139	0.5139	2.7
61	76	8.8	5	3	0.4367	0.4367	3.8
62	110	9.8	4	0	-0.6186	0.6186	3.1
63	47	6.4	1	0	-1.3232	1.3232	4.0
64	88	7.5	3	0	-0.8412	0.8412	3.2
65	10	5.4	4	3	-0.5301	0.5301	5.2
66	21	10.5	5	7	0.3448	0.3448	7.0
67	64	10.4	3	2	0.0805	0.0805	4.6
68	46	7.9	4	4	0.3073	0.3073	4.5
69	44	5.4	2	4	0.5434	0.5434	3.9
70	79	3.7	4	3	0.7199	0.7199	2.8
71	116	6.0	2	0	-0.2098	0.2098	2.5
72	84	2.9	3	1	0.2004	0.2004	2.6
73	95	10.3	5	2	0.4164	0.4164	3.6
74	17	6.7	4	7	0.4746	0.4746	5.3
75	98	2.8	4	0	-0.6550	0.6550	2.3
76	91	4.2	5	6	1.5312	1.5312	2.6
77	86	3.8	5	2	0.4181	0.4181	2.7
78	118	4.0	3	0	-0.2288	0.2288	2.2
79	85	10.5	2	1	0.1365	0.1365	3.9
80	87	6.9	1	1	0.3644	0.3644	3.1
81	53	8.9	1	2	0.1344	0.1344	4.5
82	70	2.7	1	0	-0.8292	0.8292	2.8
83	44	9.9	2	6	0.8812	0.8812	5.3
84	54	6.8	5	4	0.3702	0.3702	3.9
85	42	6.7	3	4	0.3707	0.3707	4.3
86	27	3.1	5	6	0.4375	0.4375	3.9
87	62	9.9	2	0	-1.2518	1.2518	4.5
88	57	4.0	2	1	-0.1933	0.1933	3.2
89	117	8.3	4	0	-0.4590	0.4590	2.8
90	101	6.0	1	1	0.6334	0.6334	2.7
91	94	7.4	3	2	0.6794	0.6794	3.1
92	46	4.2	1	5	0.9478	0.9478	3.6
93	84	8.8	2	2	0.5741	0.5741	3.6
94	20	4.0	1	5	0.4977	0.4977	4.3
95	102	5.3	3	0	-0.5432	0.5432	2.6
96	50	9.2	5	5	0.4754	0.4754	4.8
97	52	10.5	2	4	0.5574	0.5574	5.1
98	52	4.5	4	6	0.9452	0.9452	3.5
99	100	5.2	3	1	0.4255	0.4255	2.6
100	84	8.8	1	3	0.9986	0.9986	3.5

101	35	7.4	5	5	0.2565	0.2565	4.8
102	78	9.2	4	0	-1.1640	1.1640	3.8
103	90	9.5	5	0	-1.0660	1.0660	3.6
104	97	7.2	5	0	-0.8865	0.8865	3.0
105	49	6.7	3	0	-1.5056	1.5056	4.1
106	43	3.2	5	4	0.2658	0.2658	3.4
107	96	5.5	1	0	-0.4422	0.4422	2.7
108	37	6.1	5	4	0.0890	0.0890	4.3
109	65	9.6	2	1	-0.1906	0.1906	4.3
110	50	7.5	4	7	1.0338	1.0338	4.3
111	34	3.6	3	4	0.3075	0.3075	3.8
112	108	9.7	4	1	0.3500	0.3500	3.2
113	20	3.8	5	5	0.0839	0.0839	4.3
114	77	8.8	3	0	-1.0660	1.0660	3.7
115	86	7.5	2	1	0.2272	0.2272	3.2
116	106	5.4	1	0	-0.2662	0.2662	2.5
117	20	4.1	5	2	-0.7453	0.7453	4.4
118	29	7.6	3	3	-0.1467	0.1467	5.1
119	102	8.7	4	2	0.6837	0.6837	3.1
120	45	8.7	2	2	-0.1068	0.1068	4.8
121	87	9.5	1	3	1.0327	1.0327	3.6
122	18	3.7	3	7	0.6699	0.6699	4.3
123	96	4.6	4	1	0.2640	0.2640	2.6
124	38	8.1	4	4	0.1620	0.1620	4.9
125	54	9.1	3	4	0.5220	0.5220	4.6
126	42	2.8	2	2	-0.0122	0.0122	3.4
127	105	10.6	3	2	0.7942	0.7942	3.4
128	12	4.4	3	9	0.9013	0.9013	4.8
129	74	9.5	1	1	0.0726	0.0726	4.0
130	108	9.0	5	0	-0.7372	0.7372	3.0
131	12	7.3	3	5	0.0662	0.0662	5.9
132	90	5.9	3	0	-0.7670	0.7670	2.9
133	28	9.1	2	2	-0.4134	0.4134	5.8
134	102	3.4	5	0	-0.7060	0.7060	2.4
135	45	5.9	1	1	-0.3457	0.3457	4.0
136	45	6.2	4	1	-0.6674	0.6674	4.1
137	112	7.8	1	0	-0.2181	0.2181	2.8
138	83	8.1	4	1	-0.0493	0.0493	3.4
139	35	3.0	3	3	0.0729	0.0729	3.6
140	80	8.3	5	1	-0.2127	0.2127	3.6
141	104	5.1	5	0	-0.7131	0.7131	2.5
142	78	7.2	4	3	0.6163	0.6163	3.4
143	55	3.2	2	0	-1.2094	1.2094	3.2
144	25	4.2	5	2	-0.6606	0.6606	4.2
145	78	4.0	4	0	-1.0365	1.0365	2.8
146	87	8.8	5	3	0.6312	0.6312	3.5
147	111	8.8	3	2	0.9422	0.9422	2.9
148	36	10.0	1	4	0.3947	0.3947	5.7
149	116	8.1	5	0	-0.5763	0.5763	2.7
150	36	3.8	3	6	0.7862	0.7862	3.8

151	104	6.9	5	2	0.6581	0.6581	2.8
152	60	3.8	1	4	0.9684	0.9684	3.1
153	114	8.4	1	2	1.2152	1.2152	2.8
154	56	7.0	2	1	-0.2844	0.2844	3.9
155	66	4.3	5	2	0.0548	0.0548	3.1
156	65	5.3	4	4	0.7031	0.7031	3.3
157	105	6.0	2	0	-0.4009	0.4009	2.6
158	66	9.9	2	0	-1.1817	1.1817	4.4
159	94	10.0	4	1	0.0977	0.0977	3.5
160	113	6.0	2	0	-0.2627	0.2627	2.5
161	93	9.0	3	1	0.2078	0.2078	3.4
162	51	8.1	3	3	0.2261	0.2261	4.4
163	87	5.3	2	1	0.2984	0.2984	2.9
164	22	10.1	4	5	0.0686	0.0686	6.6
165	46	4.9	4	0	-1.6186	1.6186	3.7
166	99	2.6	4	2	0.7800	0.7800	2.3
167	14	6.0	4	3	-0.4747	0.4747	5.2
168	14	10.2	5	10	0.7457	0.7457	7.3
169	89	9.1	2	1	0.2401	0.2401	3.5
170	103	6.5	4	1	0.3394	0.3394	2.7
171	20	4.2	3	5	0.2825	0.2825	4.4
172	32	10.2	2	3	-0.0532	0.0532	6.0
173	109	5.9	5	1	0.3556	0.3556	2.6
174	55	8.4	5	3	0.0799	0.0799	4.3
175	94	3.6	2	2	0.8781	0.8781	2.5
176	64	8.4	3	1	-0.2848	0.2848	4.0
177	110	4.4	1	1	0.8296	0.8296	2.4
178	99	3.7	3	0	-0.5554	0.5554	2.4
179	36	9.1	1	3	0.1504	0.1504	5.3
180	46	10.3	2	4	0.4576	0.4576	5.3
181	115	2.9	4	0	-0.3597	0.3597	2.1
182	25	9.9	2	5	0.3357	0.3357	6.3
183	63	2.9	4	2	0.1431	0.1431	2.9
184	94	6.4	1	1	0.4989	0.4989	2.9
185	24	9.6	3	2	-0.6022	0.6022	6.3
186	73	5.9	4	1	-0.1709	0.1709	3.3
187	25	6.1	1	6	0.7486	0.7486	4.7
188	91	4.3	5	2	0.4935	0.4935	2.6
189	81	3.5	4	5	1.2639	1.2639	2.7
190	118	7.3	4	0	-0.4166	0.4166	2.6
191	86	4.2	2	0	-0.6909	0.6909	2.7
192	86	7.0	1	0	-0.6544	0.6544	3.1
193	97	6.9	5	4	1.1199	1.1199	2.9
194	41	6.7	3	3	0.0855	0.0855	4.4
195	77	4.0	1	0	-0.7390	0.7390	2.8
196	111	6.8	2	0	-0.3169	0.3169	2.6
197	70	7.9	1	0	-0.9584	0.9584	3.7
198	66	3.4	2	1	-0.0207	0.0207	3.0
199	30	3.7	5	5	0.2614	0.2614	3.9
200	115	3.4	4	0	-0.3723	0.3723	2.2

201	12	7.5	3	3	-0.4410	0.4410	5.9
202	100	7.6	1	0	-0.4245	0.4245	3.0
203	71	9.2	1	1	0.0276	0.0276	4.0
204	37	10.4	1	1	-0.5981	0.5981	5.9
205	35	3.8	4	3	-0.0523	0.0523	3.8
206	110	3.7	5	2	0.8420	0.8420	2.3
207	76	8.0	5	1	-0.2747	0.2747	3.6
208	119	6.3	5	2	0.9353	0.9353	2.5
209	70	6.4	1	0	-0.9203	0.9203	3.4
210	104	5.7	2	0	-0.4133	0.4133	2.6
211	38	9.9	2	5	0.5633	0.5633	5.6
212	70	6.0	4	2	0.1878	0.1878	3.3
213	88	6.5	4	1	0.0775	0.0775	3.0
214	30	5.8	5	4	-0.0271	0.0271	4.5
215	76	7.2	1	2	0.5797	0.5797	3.4
216	73	4.9	5	0	-1.2510	1.2510	3.1
217	54	8.3	5	5	0.5693	0.5693	4.4
218	28	5.1	3	4	0.1657	0.1657	4.3
219	64	6.7	4	2	0.0654	0.0654	3.6
220	28	5.0	1	3	0.1101	0.1101	4.3
221	37	6.4	1	1	-0.4992	0.4992	4.4
222	19	9.1	3	4	-0.0917	0.0917	6.3
223	59	7.8	3	1	-0.3572	0.3572	4.0
224	102	7.7	4	0	-0.7067	0.7067	2.9
225	34	9.6	3	2	-0.4264	0.4264	5.7
226	14	4.2	5	7	0.3781	0.3781	4.6
227	88	10.4	1	1	0.2973	0.2973	3.8
228	113	7.2	5	0	-0.6068	0.6068	2.7
229	72	5.8	4	1	-0.1857	0.1857	3.3
230	42	4.7	3	2	-0.1656	0.1656	3.8
231	18	9.3	2	4	-0.0083	0.0083	6.5
232	58	5.4	5	3	0.2061	0.2061	3.5
233	91	6.0	4	0	-0.8569	0.8569	2.9
234	67	8.9	4	1	-0.3485	0.3485	4.1
235	94	4.3	4	3	0.9678	0.9678	2.6
236	39	3.3	5	2	-0.3931	0.3931	3.6
237	79	8.2	4	1	-0.1209	0.1209	3.6
238	103	9.2	2	3	1.2171	1.2171	3.2
239	91	8.2	5	1	-0.0175	0.0175	3.3
240	101	9.4	4	2	0.6482	0.6482	3.3
241	59	7.6	2	0	-1.2473	1.2473	4.0
242	25	4.5	3	3	-0.1408	0.1408	4.3
243	11	7.4	3	7	0.4574	0.4574	5.9
244	21	8.8	3	6	0.4000	0.4000	6.1
245	40	5.9	5	1	-0.8528	0.8528	4.2
246	23	3.9	4	13	1.6070	1.6070	4.2
247	58	10.1	1	1	-0.2212	0.2212	4.7
248	24	9.0	1	3	-0.0585	0.0585	5.9
249	32	10.1	4	3	-0.2604	0.2604	6.0
250	118	4.1	2	2	1.2857	1.2857	2.2

```
> summary(fm5)
```

```
Call:
```

```
lm(formula = sqrt(Y) ~ Dist + Inc + Size, weights = wi)
```

```
Residuals:
```

Min	1Q	Median	3Q	Max
-3.1818	-0.8134	0.1118	0.9051	3.2484

```
Coefficients:
```

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	1.874293	0.181835	10.308	< 2e-16 ***
Dist	-0.018063	0.001328	-13.599	< 2e-16 ***
Inc	0.027104	0.017669	1.534	0.126323#收入太低決定將其拿掉
Size	0.113891	0.029023	3.924	0.000113 ***

```
---
```

```
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
Residual standard error: 1.233 on 246 degrees of freedom
```

```
Multiple R-squared: 0.4511, Adjusted R-squared: 0.4444
```

```
F-statistic: 67.38 on 3 and 246 DF, p-value: < 2.2e-16
```

```
> anova(fm5)
```

```
Analysis of Variance Table
```

```
Response: sqrt(Y)
```

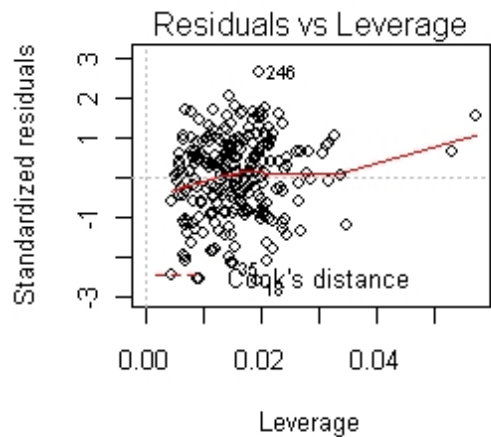
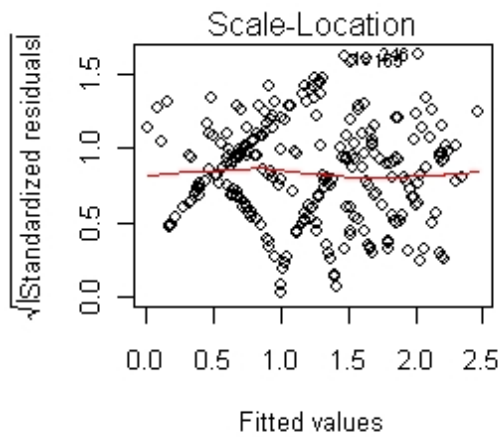
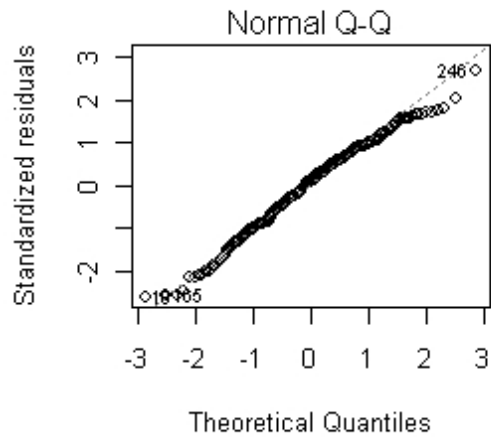
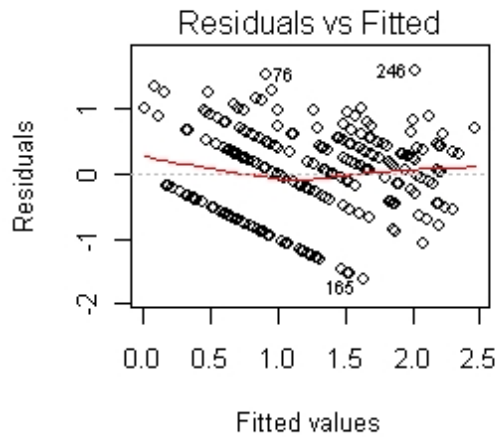
	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Dist	1	281.61	281.61	185.3227	< 2.2e-16 ***
Inc	1	2.14	2.14	1.4075	0.236623
Size	1	23.40	23.40	15.3988	0.000113 ***
Residuals	246	373.81	1.52		

```
---
```

```
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
> par(mfrow=c(2,2))
```

```
> plot(fm5,ask=F)
```



```
> fm6=lm(sqrt(Y)~Dist+Size,weights=wi)
> summary(fm6)
```

```
Call:
lm(formula = sqrt(Y) ~ Dist + Size, weights = wi)
```

```
Residuals:
    Min       1Q   Median       3Q      Max
-3.2532 -0.8942  0.1028  0.9767  3.0845
```

```
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  2.077797   0.124694  16.663 < 2e-16 ***
Dist         -0.018099   0.001332 -13.590 < 2e-16 ***
Size          0.109893   0.028985   3.791 0.000188 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
Residual standard error: 1.236 on 247 degrees of freedom
Multiple R-squared: 0.4458, Adjusted R-squared: 0.4413
```

F-statistic: 99.34 on 2 and 247 DF, p-value: < 2.2e-16

```
> anova(fm6)
```

Analysis of Variance Table

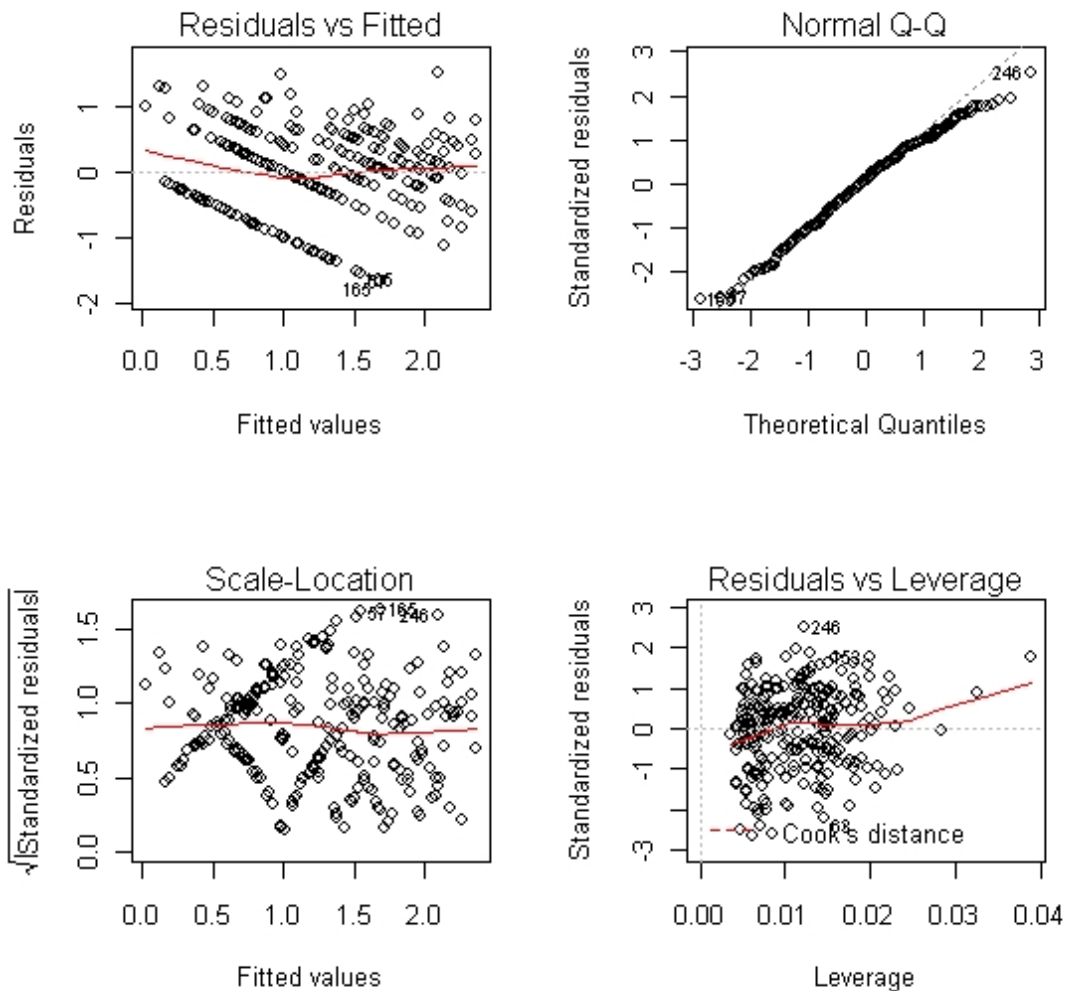
Response: sqrt(Y)

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Dist	1	281.61	281.61	184.313	< 2.2e-16 ***
Size	1	21.96	21.96	14.374	0.0001884 ***
Residuals	247	377.38	1.53		

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```
> par(mfrow=c(2,2))
```

```
> plot(fm6,ask=F)
```



最後我選擇 $\sqrt{Y} = 2.077797 - 0.018099\text{Dist} + 0.109893\text{Size}$ 這個 model, 但是其實還是有很多要改進, 它的 R-squared 太低, 解釋力不夠。或許利用 generalized least square 可以解決這個問題。