4.4 The general manager of a large engineering firm wants to know whether the experience of technical artists influences their work quality. A random sample of 50 artists is selected. Using years of work experience (*EXPER*) and a performance rating (*RATING*, on a 100-point scale), two models are estimated by least squares. The estimates and standard errors are as follows:

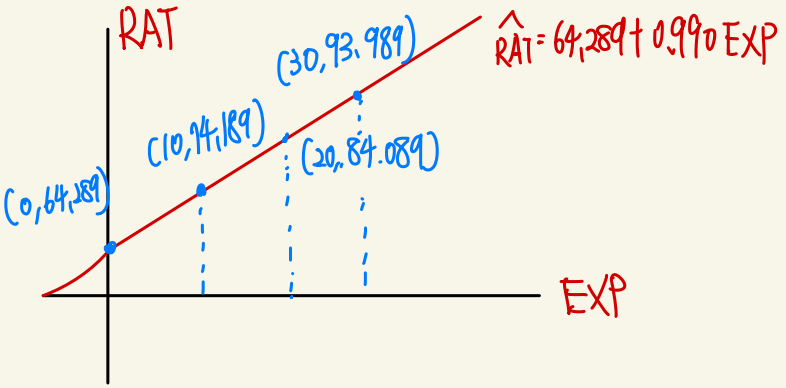
Model 1:

$$\widehat{RATING} = 64.289 + 0.990EXPER \quad N = 50 \quad R^2 = 0.3793$$
$$\text{(se)} \qquad (2.422) \quad (0.183)$$

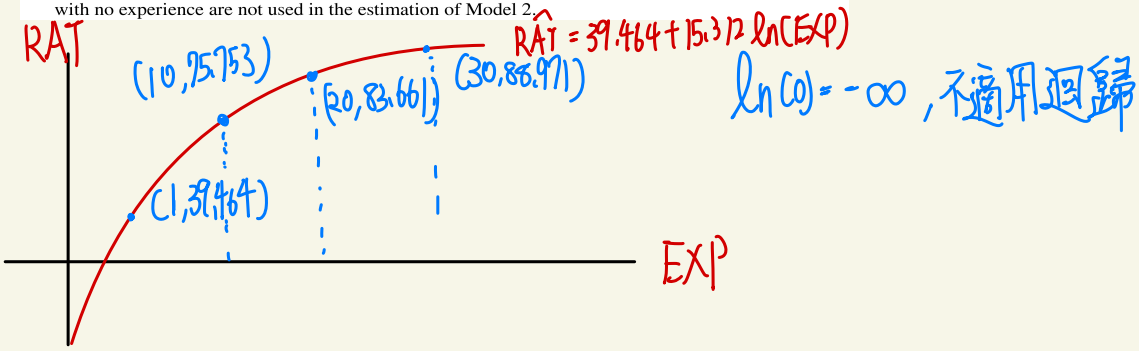
Model 2:

$$\widehat{RATING} = 39.464 + 15.312\ln(EXPER) \quad N = 46 \quad R^2 = 0.6414$$
$$\text{(se)} \qquad (4.198) \quad (1.727)$$

a. Sketch the fitted values from Model 1 for *EXPER* = 0 to 30 years.



b. Sketch the fitted values from Model 2 against *EXPER* = 1 to 30 years. Explain why the four artists with no experience are not used in the estimation of Model 2.



c. Using Model 1, compute the marginal effect on *RATING* of another year of experience for (i) an artist with 10 years of experience and (ii) an artist with 20 years of experience.

d. Using Model 2, compute the marginal effect on *RATING* of another year of experience for (i) an artist with 10 years of experience and (ii) an artist with 20 years of experience.

c. 0.990，邊際效果固定

d. (i) $15.312/10 = 1.5312$ (ii) $15.312/20 = 0.7656$

e. Which of the two models fits the data better? Estimation of Model 1 using just the technical artists with some experience yields $R^2 = 0.4858$.

f. Do you find Model 1 or Model 2 more reasonable, or plausible, based on economic reasoning? Explain.

e.

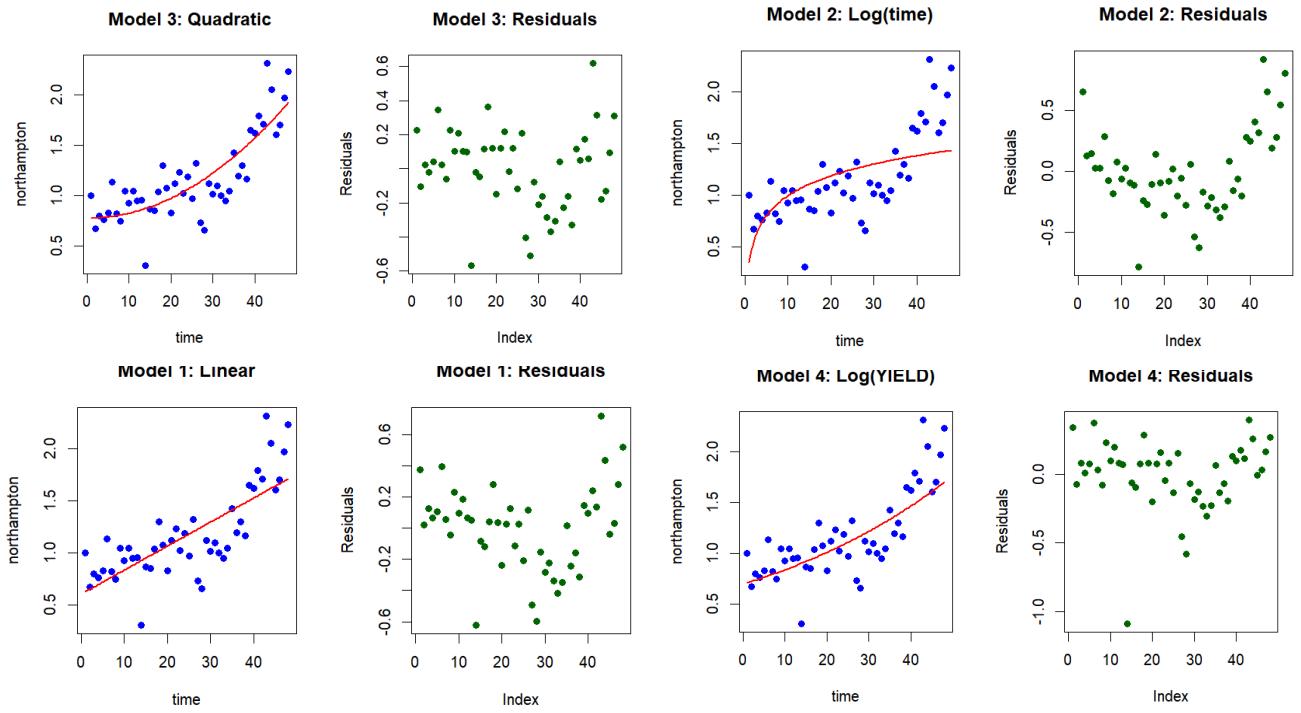
Model 2 之 R^2 較高，Model 2 效果較好

f.

Model 2 較合理，在吸取一定經驗，對 rating 的增量效果應逐漸降低

Q28

(a)應選擇model3(Quadratic Model)，因其殘差最近常態分配，且擁有最高的R平方(解釋力強)




```
> shapiro.test(model3$residuals)

        Shapiro-Wilk normality test

data:  model3$residuals
W = 0.98589, p-value = 0.8266
```

```
> # R平方比較
> summary(model1)$r.squared
[1] 0.5778369
> summary(model2)$r.squared
[1] 0.3385733
> summary(model3)$r.squared
[1] 0.6890101
> summary(model4)$r.squared
[1] 0.5073566
```

(b)

```
Residuals:
     Min       1Q   Median       3Q      Max
-0.56899 -0.14970  0.03119  0.12176  0.62049

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) 7.737e-01  5.222e-02   14.82  < 2e-16 ***
I(time^2)   4.986e-04  4.939e-05   10.10 3.01e-13 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.2396 on 46 degrees of freedom
Multiple R-squared:  0.689,     Adjusted R-squared:  0.6822
F-statistic: 101.9 on 1 and 46 DF,  p-value: 3.008e-13
```

(C)打星號為可能影響迴歸線的資料

	dfb.1_	dfb.I..2	dffit	cov.r	cook.d	hat	inf
1	0.21672	-0.16229	0.2167	1.052	2.35e-02	0.0474	
2	-0.09773	0.07306	-0.0977	1.087	4.86e-03	0.0472	
3	0.02031	-0.01514	0.0203	1.096	2.11e-04	0.0469	
4	-0.02066	0.01534	-0.0207	1.095	2.18e-04	0.0464	
5	0.03759	-0.02777	0.0376	1.094	7.22e-04	0.0458	
6	0.32374	-0.23761	0.3238	0.994	5.11e-02	0.0451	
7	0.01981	-0.01443	0.0198	1.093	2.01e-04	0.0443	
8	-0.05344	0.03856	-0.0535	1.089	1.46e-03	0.0434	
9	0.20370	-0.14536	0.2040	1.047	2.08e-02	0.0423	
10	0.09085	-0.06401	0.0911	1.081	4.22e-03	0.0412	
11	0.17959	-0.12470	0.1802	1.052	1.63e-02	0.0400	
12	0.08610	-0.05878	0.0865	1.078	3.81e-03	0.0387	
13	0.07951	-0.05324	0.0801	1.078	3.27e-03	0.0373	
14	-0.48945	0.32052	-0.4944	0.826	1.09e-01	0.0359	*
15	-0.01477	0.00943	-0.0150	1.082	1.15e-04	0.0345	
16	-0.03636	0.02252	-0.0370	1.079	7.01e-04	0.0331	
17	0.08685	-0.05198	0.0891	1.067	4.03e-03	0.0316	
18	0.26559	-0.15266	0.2746	0.970	3.66e-02	0.0301	
19	0.08416	-0.04612	0.0880	1.063	3.93e-03	0.0287	
20	-0.09745	0.05045	-0.1032	1.056	5.40e-03	0.0274	
21	0.07761	-0.03750	0.0836	1.061	3.55e-03	0.0261	
22	0.13414	-0.05951	0.1477	1.032	1.09e-02	0.0249	
23	-0.00912	0.00363	-0.0103	1.070	5.45e-05	0.0238	
24	0.06641	-0.02295	0.0778	1.057	3.07e-03	0.0228	
25	-0.05955	0.01690	-0.0728	1.057	2.69e-03	0.0220	
26	0.10001	-0.02109	0.1289	1.033	8.35e-03	0.0214	
27	-0.18618	0.02301	-0.2561	0.936	3.14e-02	0.0210	
28	-0.21991	0.00382	-0.3278	0.863	4.94e-02	0.0208	*
29	-0.02836	-0.00324	-0.0466	1.062	1.11e-03	0.0209	
30	-0.06998	-0.01971	-0.1295	1.032	8.43e-03	0.0213	
31	-0.04707	-0.02357	-0.1006	1.048	5.12e-03	0.0220	
32	-0.07267	-0.05810	-0.1848	1.004	1.69e-02	0.0231	
33	-0.07996	-0.09838	-0.2520	0.961	3.07e-02	0.0246	
34	-0.05243	-0.09984	-0.2165	0.996	2.31e-02	0.0265	
35	0.00521	0.01617	0.0308	1.074	4.83e-04	0.0288	
36	-0.01739	-0.10166	-0.1741	1.036	1.52e-02	0.0316	
37	-0.00445	-0.08158	-0.1283	1.061	8.33e-03	0.0350	
38	0.00733	-0.19326	-0.2836	0.997	3.94e-02	0.0389	
39	-0.00851	0.07524	0.1043	1.081	5.53e-03	0.0434	
40	-0.00652	0.03719	0.0492	1.096	1.24e-03	0.0486	
41	-0.03218	0.14139	0.1800	1.078	1.64e-02	0.0544	
42	-0.01371	0.05039	0.0621	1.110	1.97e-03	0.0610	
43	-0.20253	0.65218	0.7823	0.798	2.64e-01	0.0683	*
44	-0.11635	0.33832	0.3967	1.042	7.72e-02	0.0764	
45	0.07730	-0.20715	-0.2382	1.112	2.86e-02	0.0854	
46	0.06520	-0.16340	-0.1848	1.139	1.73e-02	0.0953	*
47	-0.05361	0.12702	0.1417	1.160	1.02e-02	0.1061	*
48	-0.20393	0.46077	0.5078	1.089	1.26e-01	0.1180	

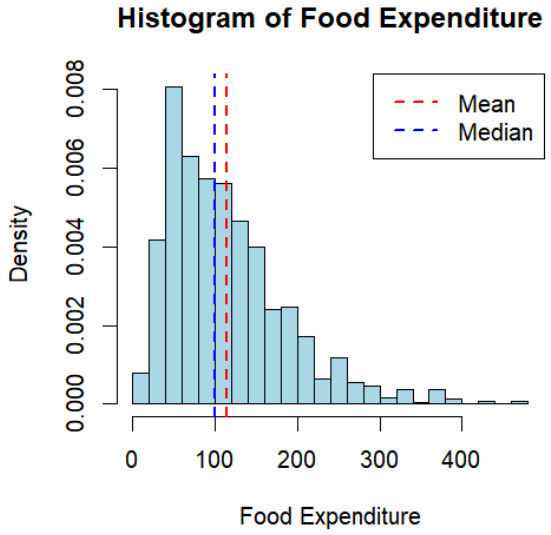
(d)實際值為2.2318，落在95%預測區間內

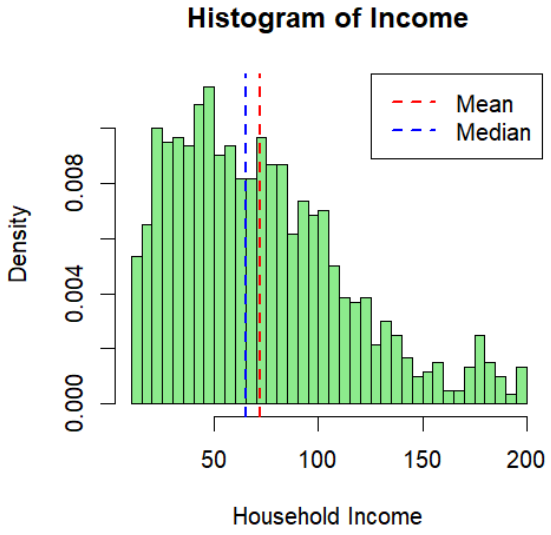
	fit	lwr	upr
1	1.922482	1.412563	2.432401

Q29.

(a)food、income之敘述統計量與直方圖，兩者都不是bell-shaped，且平均數都大於中位數，進行Jarque-bera Test後，皆無法拒絕分布非常態分配，皆可能為正偏。

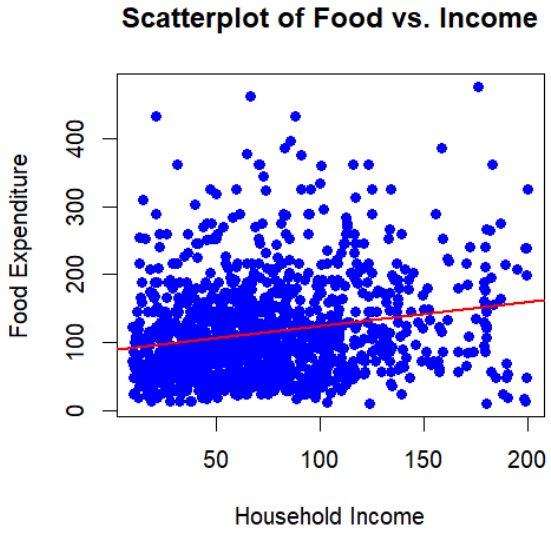
	food	income
Mean	114.4431	72.14264
Median	99.8000	65.29000
Min	9.6300	10.00000
Max	476.6700	200.00000
SD	72.6575	41.65228





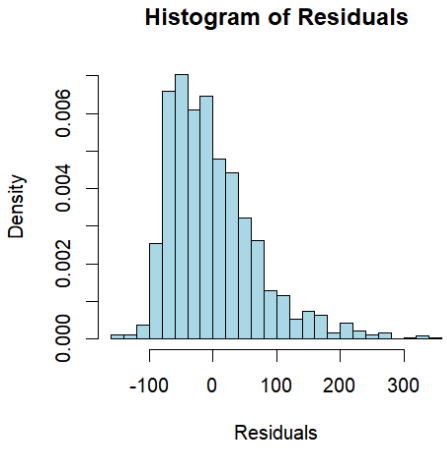
```
data:  cex5_small$food
JB = 648.65, p-value < 2.2e-16
alternative hypothesis: greater

>
> jb_income <- jarque.test(cex5_small$income)
> print(jb_income)

        Jarque-Bera Normality Test

data:  cex5_small$income
JB = 148.21, p-value < 2.2e-16
alternative hypothesis: greater
```

(b)




```
> confint(lm_model, level = 0.95)
                 2.5 %    97.5 %
(Intercept) 80.5064570 96.626543
income       0.2619215  0.455452
>
```

(c)無法拒絕殘差不是常態分配(可能為正偏)

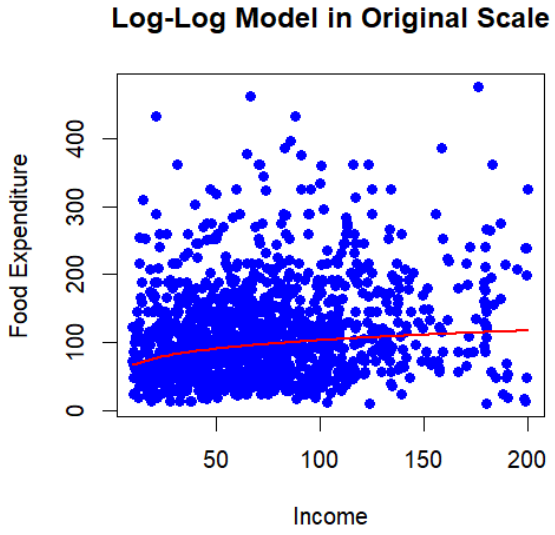



```
> jarque.test(residuals_lm)

        Jarque-Bera Normality Test

data:  residuals_lm
JB = 624.19, p-value < 2.2e-16
alternative hypothesis: greater
```

(d)彈性隨收入增加而單調遞增，且信賴區間不重疊。收入增加代表家庭可能花更多錢在食物上(更為精緻的飲食，而非單純滿足生理需求)。

```
> print(result)
  income_values  food_hat elasticity lower_bound upper_bound
1            19  95.38155 0.07145038  0.05219387  0.09070689
2            65 111.88114 0.20838756  0.15222527  0.26454986
3           160 145.95638 0.39319883  0.28722827  0.49916940
```

(e)log-log Model的表現無較佳

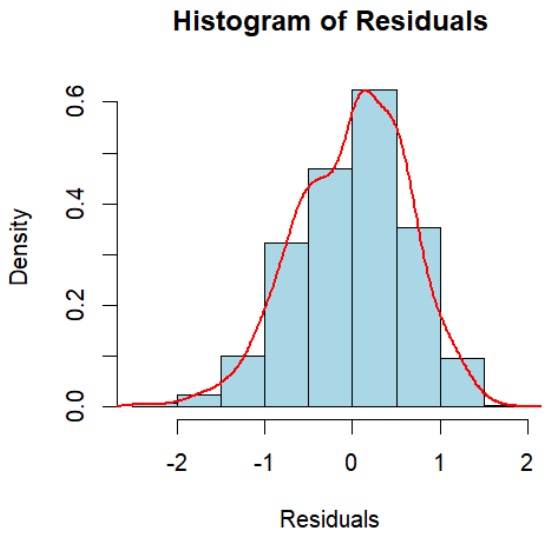



```
> cat("Linear Model R²: ", r2_linear, "\n")
Linear Model R²:  0.0422812
> cat("Log-Log Model R²: ", r2_log_log, "\n")
Log-Log Model R²:  0.03322915
> cat("Generalized R² for Log-Log Model: ", ge
Generalized R² for Log-Log Model:  0.0332564
```

(f)與linear-Model不同，因log-log彈性為常數

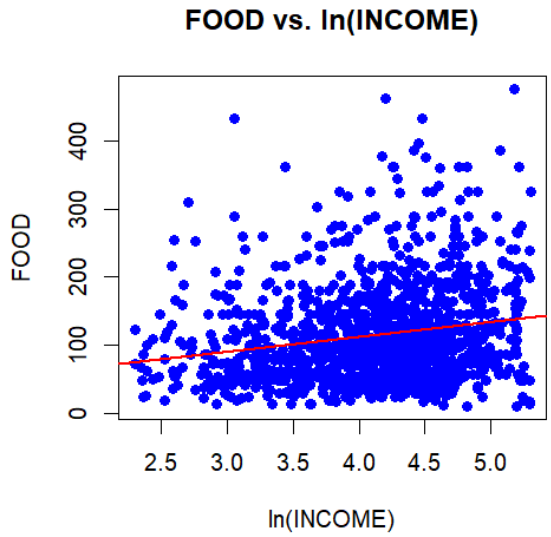
```
> cat("Point Estimate of Elasticity: ", beta2_hat, "\n")
Point Estimate of Elasticity:  0.1863054
> cat("95% Confidence Interval: (", CI_lower, ",", CI_upper, ")\n")
95% Confidence Interval: ( 0.1293997 , 0.243211 )
```

(g)log-log model之殘差並非常態分配




```
Jarque-Bera Test Statistic:  25.84998
> cat("p-value: ", jb_test$p.value, "\n")
p-value:  2.436404e-06
>
> if (jb_test$p.value < 0.05) {
+   cat("Conclusion: Residuals are NOT normally dist
+ } else {
+   cat("Conclusion: Residuals are normally distribu
+ }
Conclusion: Residuals are NOT normally distributed.
```

(h) linear-log Model的表現大同小異

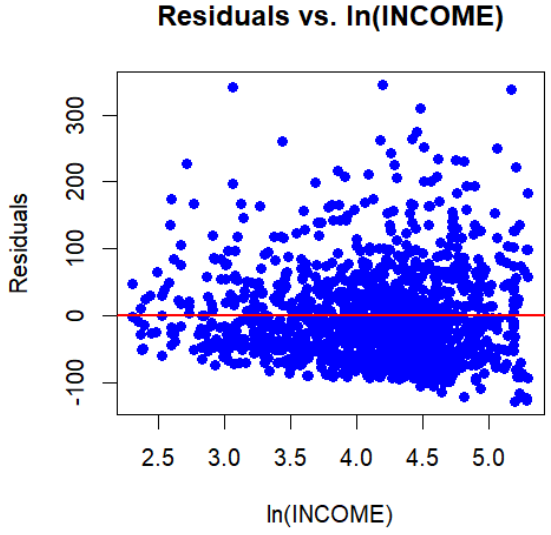


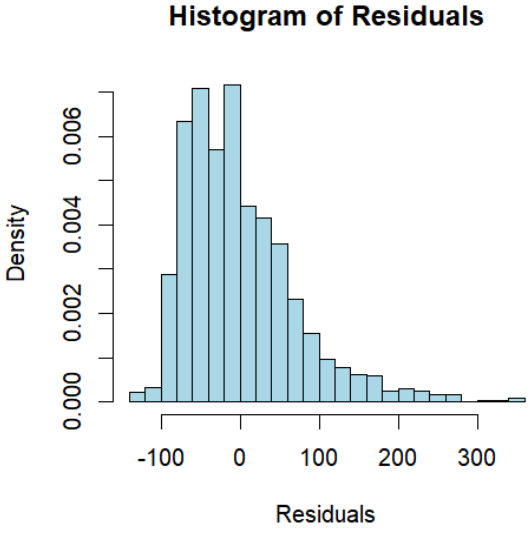

```
cat("R-squared for Linear Model: ", r2_line
-squared for Linear Model:  0.0422812
cat("R-squared for Log-Log Model: ", r2_log
-squared for Log-Log Model:  0.03322915
cat("R-squared for Linear-Log Model: ", r2_
-squared for Linear-Log Model:  0.03799984
```

(i)linear-log之彈性與上述兩個模型完全不同，因其主要取決於1/Income。

Income	Elasticity_Estimate	Lower_95_CI	Upper_95_CI
19	1.1677571	0.83503867	1.5004756
65	0.3413444	0.24408823	0.4386005
160	0.1386712	0.09916084	0.1781815

(j)linear-log之殘差檢定=>可能也不是常態分配






```
        Jarque-Bera Normality Test

data:  residuals_linear_log
JB = 628.07, p-value < 2.2e-16
alternative hypothesis: greater
```

(h)在R平方差不多(0.03~0.04)的情況下，linear log Model的殘差更接近常態分配，可能是較好的估計模型。