ជំពូកទី១០ គំរូមិនលីនេអ៊ែរ

Non Linear Modeling

រៀបរៀងនិងបង្រៀនដោយបណ្ឌិត ង៉ាន់ ស៊ុនដេត

សាស្ត្រាចារ្យនៃសាកលវិទ្យាល័យភូមិន្ទនីតិសាស្ត្រនិងវិទ្យាសាស្ត្រសេដ្ឋកិច្ច

គំរូមិនលីនេអ៊ែរ
Non Linear Modeling

- យើងបានសិក្សារួចមកហើយគំរូលីនេអ៊ែរដែលមានអថេរពន្យល់បង្ហាញថេរៈវេលាដូចគ្នា។ នៅពេលនេះទ្រឹស្តីសេដ្ឋកិច្ចតម្រូវឲ្យរៀបចំឲ្យបានល្អ។ ជារឿយៗទ្រឹស្តីសេដ្ឋកិច្ចតម្រូវឲ្យរៀបចំផ្លូវគណនាតាមវិធីទំនាក់ទំនងមិនលីនេអ៊ែរដែរ។
- ទំនាក់ទំនងមិនលីនេអ៊ែរមានច្រើនប្រភេទ និងអាចស្មុំញ៉ាំផង តែពេលនេះគឺលើកយកមកសិក្សាតែករណីពិសេសខ្លះតែប៉ុណ្ណោះ។
- សិក្សារបៀបបំលែងនិងវិធីគណនាមេគុណប៉ារ៉ាម៉ែត្រ។

គំរូមិនលីនេអ៊ែរ
Non Linear Modeling

1) អនុគមន៍ផលិតកម្ម Cobb-Douglas

អនុគមន៍ផលិតកម្ម Cobb-Douglas មានរាង$Q = f(K, L) = AK^{\alpha}L^{\beta}e^{u_i}, i = 1,2,\dots,n$ ។

Q= វាល់បរិមាណផលិតកម្ម

K= ដើមទុន/កត្តាផលិតកម្មនៃធនធានមូលធនហិរញ្ញវត្ថុ

L=កម្លាំងពលកម្ម/កត្តាផលិតកម្មនៃធនធានពលកម្មនិងកម្រិតសិក្សា

ឧទាហរណ៍ ការអង្កេត២៥សហគ្រាសៈ បរិមាណផលិតកម្ម Q, កត្តាផលិតកម្មនៃធនធានមូលធនហិរញ្ញវត្ថុ K, កត្តាផលិតកម្មនៃធនធានពលកម្ម L, អថេរទាំងអស់នេះធ្វើការវាស់ជារាប់លានដុល្លា(1,000,000.00$)។អនុគមន៍ផលិតកម្ម Cobb-Douglas មានរាង$Q = f(K, L) = AK^{\alpha}L^{\beta}$។

Ob	Q	K	L
1	106.00	8.00	23.00
2	81.08	9.00	14.00
3	72.80	4.00	38.00
4	57.34	2.00	97.00
5	66.79	6.00	11.00
6	98.23	6.00	43.00
7	82.63	3.00	93.00
8	99.77	6.00	49.00
9	110.00	8.00	36.00
10	118.93	8.00	43.00
11	95.05	4.00	61.00
12	112.83	8.00	31.00
13	64.54	3.00	57.00
14	137.22	6.00	97.00
15	86.17	4.00	93.00
16	56.25	2.00	72.00
17	81.10	3.00	61.00
18	65.23	3.00	97.00
19	149.56	9.00	89.00
20	65.43	3.00	25.00
21	36.06	1.00	81.00
22	56.92	4.00	11.00
23	49.59	2.00	64.00
24	43.21	3.00	10.00
25	121.24	6.00	71.00

គំរូមិនលីនេអ៊ែរ
Non Linear Modeling

តារាងទិន្នន័យអង្កេត២៥សហគ្រាស

1) ផ្តល់ការបកស្រាយអត្ថន័យសេដ្ឋកិច្ចមេគុណ α *and* β

2) គណនាតម្លៃមេគុណ α *and* β និងអត្ថាធិប្បាយលទ្ធផល

3) គណនាហានិភ័យ $\alpha\%$ នៅក្នុងលំហារទិន្នផលមាត្រដ្ឋាន។

ស្រាយបញ្ជាក់

1. មេគុណ α *and* β បកស្រាយរៀងគ្នាពីភាពយឺតនៃផលិតកម្មរបស់កត្តាផលិតកម្មដើមទុន និងកត្តាផលិតកម្មធនធានពលកម្ម។

   គណនា α *and* β ចេញពីសមីការ $Q = f(K, L) = AK^{\alpha}L^{\beta}$។

2. បំលែងជា លោការីតទស្សភាគគោល១០ យើងបាន៖

   $$Log(Q) = Log(A) + \alpha Log(K) + \beta Log(L) + u_i, i = 1,2, \dots, 25$$

   យើងបានតារាងខាងក្រោម៖

គំរូមិនលីនេអ៊ែរ Non Linear Modeling

លទ្ធផលវិភាគតាមExcel

Ob	Q	K	L	LOGQ	LOGK	LOGL
1	106.00	8.00	23.00	2.025306	0.90309	1.361728
2	81.08	9.00	14.00	1.908914	0.954243	1.146128
3	72.80	4.00	38.00	1.862131	0.60206	1.579784
4	57.34	2.00	97.00	1.758458	0.30103	1.986772
5	66.79	6.00	11.00	1.824711	0.778151	1.041393
6	98.23	6.00	43.00	1.992244	0.778151	1.633468
7	82.63	3.00	93.00	1.917138	0.477121	1.968483
8	99.77	6.00	49.00	1.999	0.778151	1.690196
9	110.00	8.00	36.00	2.041393	0.90309	1.556303
10	118.93	8.00	43.00	2.075291	0.90309	1.633468
11	95.05	4.00	61.00	1.977952	0.60206	1.78533
12	112.83	8.00	31.00	2.052425	0.90309	1.491362
13	64.54	3.00	57.00	1.809829	0.477121	1.755875
14	137.22	6.00	97.00	2.137417	0.778151	1.986772
15	86.17	4.00	93.00	1.935356	0.60206	1.968483
16	56.25	2.00	72.00	1.750123	0.30103	1.857332
17	81.10	3.00	61.00	1.909021	0.477121	1.78533
18	65.23	3.00	97.00	1.814447	0.477121	1.986772
19	149.56	9.00	89.00	2.174815	0.954243	1.94939
20	65.43	3.00	25.00	1.815777	0.477121	1.39794
21	36.06	1.00	81.00	1.557026	0	1.908485
22	56.92	4.00	11.00	1.755265	0.60206	1.041393
23	49.59	2.00	64.00	1.695394	0.30103	1.80618
24	43.21	3.00	10.00	1.635584	0.477121	1
25	121.24	6.00	71.00	2.083646	0.778151	1.851258

SUMMARY OUTPUT						
Regression Statistics						
Multiple R	0.970351					
R Square	0.941581					
Adjusted R Square	0.93627					
Standard Error	0.040044					
Observations	25					
ANOVA						
	df	SS	MS	F	Significance F	
Regression	2	0.568584	0.284292	**177.2947**	**2.7046E-14**	
Residual	22	0.035277	0.001603			
Total	24	0.603861				
	Coefficients	Standard Error	**t Stat**	**P-value**	Lower 95%	Upper 95%
Intercept	**1.077552**	0.055847	19.29474	2.82E-15	0.96173215	1.193371
Log(K)	**0.640124**	0.034724	**18.43482**	**7.26E-15**	0.5681115	0.71213635
Log(L)	**0.257304**	0.026954	**9.546226**	**2.79E-09**	0.20140614	0.31320246

គំរូមិនលីនេអ៊ែរ
Non Linear Modeling

យើបានតម្លៃគណនា ratio-student លើតារាងវិភាគរបស់មេគុណអថេរដើមទុន និង អថេរ

កម្លាំងពលកម្មគឺធំជាងតម្លៃ ratio-student ដែលអានក្នុងតារាងច្បាប់ណ័រម៉ល $t^{\frac{5\%}{2}}_{(25,2)} = 2.07$, ដូច្នេះមេគុណទាំងអស់មានអត្ថន័យស្ថិតិគ្រប់គ្រាន់ខុសពី 0 ត្រង់ហានិភ័យ ៥%។

យើងបានភាពយឺតរបស់ដើមទុនស្មើនឹង $\alpha = 0.64$ មានន័យថាបើដើមទុនកើន ១0% នាំឲ្យផលិតកម្មកើនបាន ៦.៤% និងភាពយឺតរបស់កម្លាំងពលកម្មស្មើនឹង $\beta = 0.25$ មានន័យថាបើកម្លាំងពលកម្មកើន ១0% នាំឲ្យផលិតកម្មកើនបាន ២.៥%។

សមីការផលិតកម្មកំណត់ដោយ $\hat{Q}_i = 10^{1.08384} K^{0.64} L^{0.25} \Rightarrow \boldsymbol{\hat{Q}_i = 12.13 K^{0.64} L^{0.25}}$ ។

គំរូម៉ូឌែលមិនលីនេអ៊ែរ
Non Linear Modeling

3. តើអនុគមន៍ផលិតកម្មមានទិន្នផលមាត្រដ្ឋានថេរ ឬទេ?

តាមគោលការទិន្នផលមត្រដ្ឋានមានបីករណី៖

បើ $\alpha+\beta\ >1$: គឺជាទិន្នផលមាត្រដ្ឋានកើន(Increasing return to scale).

បើ $\alpha+\ \beta<1$: គឺជាទិន្នផលមាត្រដ្ឋានថយចុះ(decreasing return to scale).

បើ $\alpha+\ \beta=1$: គឺជាទិន្នផលមាត្រដ្ឋានថេរ(constant return to scale).

ឥឡូវនេះយើងមាន $\alpha+\ \beta=0.64+0.25=0.89,$

តើមានតម្លៃប្រូបាប៊ីលីតេដែលនាំឲ្យតម្លៃផលបូកនេះមានអត្ថន័យស្ថិតិខុសពី 1 ឬទេ?

សម្មតិកម្មគឺ $H0$: $\alpha+\beta=1\ \ and\ \ H1$: $\alpha+\beta<1.$

តម្លៃ ratio-student ដើម្បីបដិសេធសម្មតិកម្ម Ho កំណត់ដោយ : $t^*=\frac{\hat{\alpha}+\hat{\beta}-1}{\widehat{sd}_{\hat{\alpha}+\hat{\beta}}}$

គំរូមិនលីនេអ៊ែរ
Non Linear Modeling

ម៉ាទ្រីសវ៉ារ្យង់កូវ៉ារ្យង់របស់មេគុណព្យាករណ៍កំណត់ដោយ

$$\Omega_{\hat{\beta}} = \delta^2_{\hat{\beta}}(X'X)^{-1} = \frac{\sum_{i=1}^{25} e_i^2}{n-k-1}(X'X)^{-1} = 0.0016(X'X)^{-1} = \begin{array}{cccc} C & LogK & LogL & \\ 0.003112 & -0.00124 & -0.00138 & C \\ -0.00124 & 0.001203 & 0.000299 & LogK \\ -0.00138 & 0.000299 & 0.000725 & LogL \end{array}$$

យើងបានវ៉ារ្យង់ $\hat{\sigma}^2_{\hat{\alpha}+\hat{\beta}} = \widehat{sd}^2_{\hat{\alpha}} + \widehat{sd}^2_{\hat{\beta}} + 2cov(\hat{\alpha}, \hat{\beta}) = 0.0012 + 0.00072 + 2 \times 0.0003 = 0.002638$

$$\Rightarrow \text{standard diviation: } \hat{\sigma}_{\hat{\alpha}+\hat{\beta}} = \sqrt{\hat{\sigma}^2_{\hat{\alpha}+\hat{\beta}}} = \sqrt{0.002638} = 0.05136$$

$$t^* = \frac{\hat{\alpha} + \hat{\beta} - 1}{\widehat{sd}_{\hat{\alpha}+\hat{\beta}}} = \frac{0.64 + 0.25 - 1}{0.05136} = -2.14 = |-2.14| = 2.14 > t^{\frac{5\%}{2}}_{(25,2)} = 2.074$$

ដូច្នេះយើងមាន ៥%នៃហានិភ័យក្នុងការបដិសេធសម្មតិកម្ម $H0$, យើងអាចសន្និដ្ឋានថា ផលិតកម្មមានអត្ថន័យគ្រប់គ្រាន់ជាទិន្នផលមាត្រដ្ឋានថយចុះ។

គំរូមិនលីនេអ៊ែរ
Non Linear Modeling

2) អនុគមន៍លូស៊ីស្ទិក(Logistic Curve)

គំរូ Logistic Curve ត្រូវបានកំណត់ជាអនុគមន៍ $Y_t = \frac{Y_{max}}{1-Cr^t}$. ដែល c និង r ជាម៉ាប៉ាម៉ែត្រពីររបស់គំរូ (0 < r < +1), r : កំណត់លក្ខណៈល្បឿននៃលំហូរសេវាផ្គត់ផ្គង់តម្រូវការនានា។ C ជាលក្ខណៈដែលបានបញ្ជាទិញនៅប្រភពដើម ឬ យើងអាចនិយាយថាជាចំនួនថេរដែរ។

Y_{max} ជាចំនុចឆ្អែត។ បើ $t \to +\infty \Rightarrow Y_t \to Y_{max}$, $and\ if\ t \to -\infty \Rightarrow Y_t \to 0$ ។

ចំនុចរបត់របស់គំរូ Logistic Curve គឺថេរ និងនៅត្រង់ចំនុច $Y_t = \frac{Y_{max}}{2}$ ។

ក្នុងការអនុវត្តន៍គឺត្រូវបំលែងកន្សោមតាមអនុគមន៍លោការីតជា *លោការីតនេពែរ ឬ លោការីតទសភាគគោល១០ក៏បាន លទ្ធផលដូចគ្នាតែមួយ* ។

គំរូមិនលីនេអ៊ែរ
Non Linear Modeling

បំលែងកន្សោមយើងបាន $Y_t = \frac{Y_{max}}{1+Cr^t} \Rightarrow 1 + Cr^t = \frac{Y_{max}}{Y_t} \Rightarrow Cr^t = \frac{Y_{max}}{Y_t} - 1$ បំលែងជាអនុគមន៍

លោការីយើងបាន $Log(C) + tLog(r) = Log(\frac{Y_{max}}{Y_t} - 1)$

តាងកន្សោម $\alpha = Log(C), \beta = Log(r), and\ Q_t = Log(\frac{Y_{max}}{Y_t} - 1)$

យើងបានសមីការលីនេអ៊ែរ $Q_t = \alpha + \beta t$ ។

ឧទាហរណ៍ ការលក់ CD ចម្រៀងមានឈ្មោះល្បីមួយនៅប្រទេសបារាំងក្នុងទីប្រជុំជនមួយមានប្រជាជនរស់នៅ១០០០នាក់មានការកើនឡើងជាលំដាប់យោងតាមទិន្នន័យអង្កេត១៩សប្តាហ៍។

ក. សង់ក្រាភិកតាងកំណើនតម្រូវការ CD ចម្រៀងនេះជាអនុគមន៍នៃពេល ។ ខ. គណនាមេគុណកំណើនករណីប្រទេសបារាំង។ ទិន្នន័យនេះមានលក្ខណៈស្រដៀងគ្នានឹងក្រុងតូចមួយនៃប្រទេសស្វីសជាប់នឹងបារាំង។ គណនាមេគុណកំណើនករណីប្រជាជនស្វ៊ីសមាន៨០០នាក់ រួចគណនា *dw*។

គំរូមិនលីនេអ៊ែរ
Non Linear Modeling

- ទិន្នន័យអង្កេត

Week	CD audio	week	CD audio
1	40.8	11	108.8
2	44	12	118.3
3	48.5	13	128.6
4	52.8	14	140.8
5	59.7	15	156.4
6	66	16	170.9
7	74.3	17	183.3
8	81.9	18	197.6
9	90.7	19	211
10	100.9		

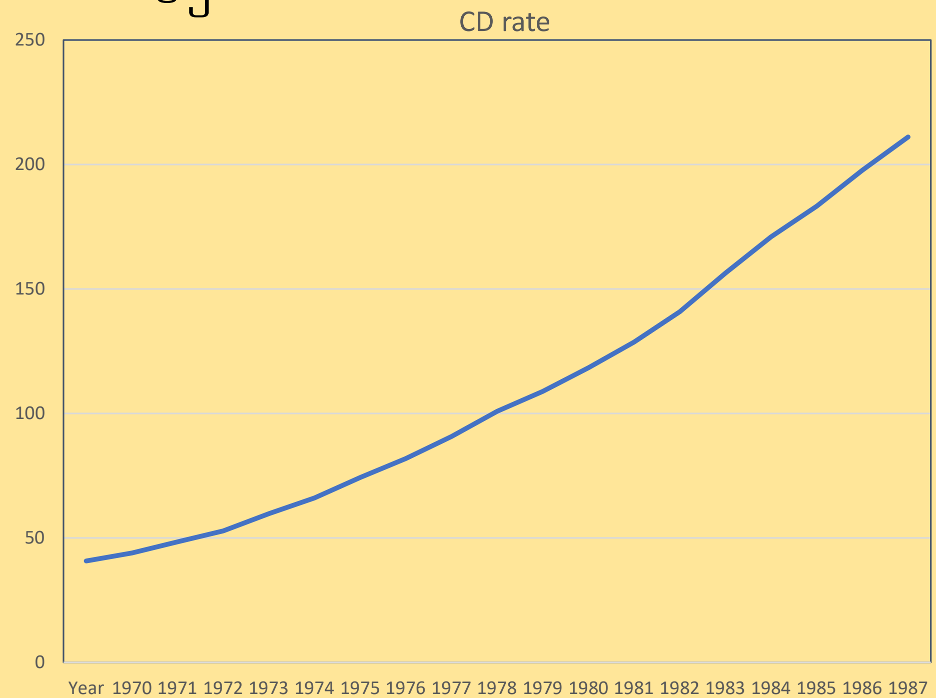
ទិន្នន័យបំលែងដោយសន្មត់សប្តាហ៍គោលចេញពី 1៖

week	Qi=Ln(B/Yi-1), B=1000	Qi=Ln(B/Yi-1), B=400	week	Qi=Ln(B/Yi-1), B=1000	Qi=Ln(B/Yi-1), B=400
1	3.157418	2.175197	11	2.103058	0.984499
2	3.078568	2.090741	12	2.008628	0.867619
3	2.976476	1.980646	13	1.913394	0.746887
4	2.886999	1.88339	14	1.808661	0.61026
5	2.756867	1.740496	15	1.685262	0.443111
6	2.649822	1.621486	16	1.579262	0.29308
7	2.522439	1.477866	17	1.494148	0.16739
8	2.416807	1.356867	18	1.401362	0.024001
9	2.305118	1.226754	19	1.318908	-0.11011
10	2.187264	1.086648			

គំរូមិនលីនេអ៊ែរ
Non Linear Modeling

- បំណាកស្រាយក្រភិកទិន្នន័យរបស់តម្រូវការថាសចម្រៀង

ក្រភិកបង្ហាញការប្រែប្រួលតម្រូវការ មានកើនឡើងខ្លាំងនាំឲ្យបរិមាណលក់ បូកបន្តកើនឡើងជាអនុគមន៍នៃឆ្នាំ។ ខ្សែកោងនេះអាចផ្ដល់យោបល់ការ វិវត្តទៅជាខ្សែកោងកើននិងមានលក្ខណៈ ជាលំហូរ *ឡូជីស្ទិក* ដែលយើងអាចនឹង ស្មានចំនុចថេររបស់ចំនុចរបត់បាន។



គំរូមិនលីនេអ៊ែរ
Non Linear Modeling

- បំលែងជាលោការីតនេពែរ និងប្រើឧបករណ៍វិភាគ យើងបានលទ្ធផលខាងក្រោះ

មេគុណ $\alpha = 3.2720,\ \beta = -0.1048$

$Ln(C) = \alpha = 3.2720$

$\Rightarrow C = e^{3.2720} = 26.3652$

$Ln(r) = \beta = -0.1048$

$\Rightarrow r = e^{-0.1048} = 0.900$

ដូច្នេះយើងបានសមីការប៉ាន់ស្មាន ឡូស៊ីស្ទិកនៅប្រទេសបារាំងគឺ

$Y_{t\_France} = \frac{1000}{1+26.741\times(0.900)^t}$ ។

ករណីប្រទេសស្វីស យើងបាន

$Y_{t\_Swiss} = \frac{40}{1+10.73355\times(0.8793)^t}$ ។

SUMMARY OUTPUT For FRANCE						
Regression Statistics						
Multiple R	0.9994					
R Square	0.9988					
Adjusted R Square	0.9988					
Standard Error	0.0202					
Observations	19					
ANOVA						
	Df	SS	MS	F	Significance F	
Regression	1	6.264352	6.264352	15302.039	1.46183E-26	
Residual	17	0.006959	0.000409			
Total	18	6.271311				
	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%
Intercept	3.2720	0.009663	338.627	5.41206E-34	3.251	3.292
T = week	-0.1048	0.000847	-123.701	1.46183E-26	-0.106	-0.103

គំរូមិនលីនេអ៊ែរ
Non Linear Modeling

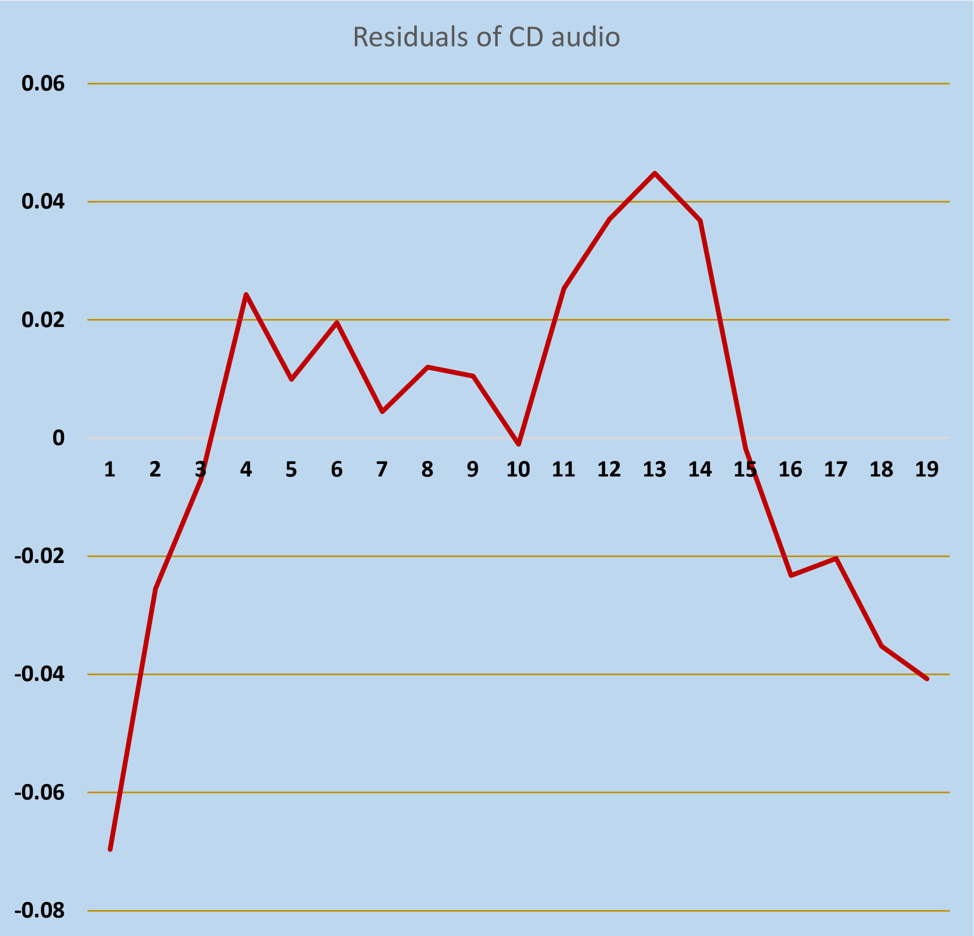
គណនាស្ថិតិ *dw* និងការសន្និដ្ឋានស្វ័យលំអៀង

$$H_0 : \rho = 0$$
$$H_1 : \rho \neq 0$$

$$dw = \frac{\sum_{t=2}^{19}(e_t - e_{t-1})^2}{\sum_{t=1}^{19}(e_t^2)^2} = \frac{0.007152059}{0.016127181} = 0.443$$

ស្ថិតិ *dw* អានក្នុងតារាង៥% $d_1 = 1.18$ & $d_2 = 1.40$

$\widehat{dw} = 0.443 < d_1 = 1.18 \Rightarrow$ គ្មានអត្ថន័យស្ថិតិគ្រប់គ្រាន់ក្នុងការទទួលយក $H_0 \Rightarrow \rho > 0$ ដូច្នេះគំរូព្យាករណ៍មានស្វ័យលំអៀងវិជ្ជាមានក្នុងការប៉ាន់ស្មានតម្លៃ។ ប៉ុន្តែតម្លៃស្វ័យលំអៀងមានកម្រិតទាប។



គំរូមិនលីនេអ៊ែរផ្សេងៗទៀត
Non Linear Modeling

- គំរូមិនលីនេអ៊ែរផ្សេងទៀតករណីអនុគមន៍ $E(y) = \beta_0 + \beta_1 X$ អាចបំបែកជាករណីពិសេស៖

  ឬជា $E(y) = \beta_0 + \beta_1\left(\frac{1}{X}\right) + u_i$

  ឬជា $E(y) = \beta_0 + \beta_1 \log(X) + u_i$

  ឬជា $E(y) = \beta_0 X^{\beta_1} u_i$

- គំរូមិនលីនេអ៊ែរមានច្រើនទម្រង់ទៅតាមស្ថានភាពនៃទ្រឹស្តីនិងភាពជាក់ស្តែង។ មានដូចជា៖

  $E(y) = \beta_0 X_1^{\beta_1} X_2^{\beta_2} \dots X_k^{\beta_k} u_i$

  $E(Y) = \beta_0 + \beta_1 X_1^1 + \beta_2 X_2^2 + \cdots + \beta_k X_k^k + u_i$

- ក្នុងការបង្កើតគំរូវិភាគជាការចាំបាច់គឺត្រូវពិនិត្យទ្រឹស្តីនិងពិសោធន៍ទិន្នន័យរបស់សំណាក។

គំរូមិនលីនេអ៊ែរផ្សេងៗទៀត
Non Linear Modeling

- ករណីអនុគមន៍ពហុធាគេអាចប្រើវិធី OLSដើម្បីគណនាមេគុណ Y_i ជាមួយអថេរ $X, X^2, X^3, \ldots, . X^k$.
- ក្នុងការធ្វើតេស្តសម្មតិកម្មគឺអនុវត្តលើសម្មតិកម្មដំបូងនិងសម្មតិកម្មឆ្លាស់៖

  Null hypothesis H_0: $\beta_1 = 0,\ \beta_2 = 0, \ldots, \beta_k = 0\ \ vs.$

  $Alternative\ hypothessis\ \ H_1\!:\ \ at\ least\ one\ \beta_k \neq 0,\ \ k = 1,2, \ldots, n.$
- ការធ្វើតេស្តមេគុណ β_k គឺអនុវត្តតេស្តម្តងមួយៗរហូតដល់វាខុសពី0គឺអាចឈប់នោះ។
- **ឧទាហរណ៍** សិក្សាអនុគមន៍ត្រីធាដឺក្រេទីពបង្ហាញពីទំនាក់ទំនងរវាងប្រាក់ចំណូលក្នុងស្រុកដើម្បីធ្វើតេស្តពិន្ទុកំណត់ដោយសមីការ $\widehat{TestScore} = 600.1 + 5.02(Income) - 0.096(Income)^2 + 0.00069(Income)^3$

  (.)=sd, (5.1) (0.71) (0.029) (0.00035)

t-statistic របស់មេគុណ $(Income)^3$ ស្មើនឹង1.97 មានអត្ថន័យស្ថិតិខុសពី 0 ត្រង់ហានិភ័យ 5% ។ F-statistic របស់ Fisher ជាប់ពាក់ព័ន្ធ $(Income)^2$ និង $(Income)^3$ ស្មើនឹង 37.7 មានហានិភ័យតិចជាង 0.01% អាចបដិសេធន៍សម្មតិកម្ម H_0។ ដូច្នេះសមីការនេះវាជារាងការ៉េ(quadratic) ឬជាដឺក្រេទីព(cubic)។

គំរូមិនលីនេអ៊ែរផ្សេងៗទៀត
Non Linear Modeling

- ឧទាហរណ៍ ការវិវត្តរបស់ចំនួនបាក់តេរី(Y) ក្នុងពេលវេលានិងសីតុណ្ហភាពអំណោយផលត្រូវបានគេធ្វើអង្កេតនៅក្នុងបន្ទប់ពិសោធន៍ពីមួយម៉ោងទៅមួយម៉ោង(T)លទ្ធផលបង្ហាញក្នុងតារាងខាងក្រោមនេះ ៖

T	0	1	2	3	4	5	6	7	8	9	10
Y	2	3	4	5	6	8	10	13	15	22	28

H1: ចំនួនបាក់តេរីមានការកើនឡើងជានិច្ចពីមួយម៉ោងទៅមួយម៉ោង។

H2: វាមានការកើនឡើងពីមួយម៉ោងទៅមួយម៉ោងក្នុងអត្រាថេរ។

ក. តើគំរូវិភាគណាមួយដែលលាក់កំបាំងនៅក្នុងសម្មតិកម្មទាំងពីរខាងលើនេះ?

ខ. ចូរវិនិច្ឆ័យតាមគ្រប់មធ្យោបាយដែលអាចរបស់សម្មតិកម្មទាំងពីរនេះ។ តើសម្មតិកម្មណាមួយត្រូវពិចារណាឲ្យប្រហាក់ប្រហែលនឹងស្ថានភាពអាចទទួលយកបាន?

គ. តើអ្នកអាចធ្វើការព្យាករណ៍ចំនួនបាក់តេរីបានទេនៅម៉ោងទី២៥នៃការពិសោធន៍?

គំរូមិនលីនេអ៊ែរផ្សេងៗទៀត
Non Linear Modeling

ក. តើគំរូណាមួយដែលលាក់កំបាំងនៅក្នុងសម្មតិកម្មទាំងពីរខាងលើនេះ?

H1: ចំនួនបាក់តេរីមានការកើនឡើងជានិច្ចពីមួយម៉ោងទៅមួយម៉ោងមានន័យថាជាគំរូលីនេអ៊ែរ ពីព្រោះវាមានចំនួនបាក់តេរីប្រែប្រួលកើនឡើងគឺសមមាត្រទៅនឹងរយៈពេលកើនឡើងដូចគ្នាមានន័យថា $dY = adT$. ដូច្នេះយើងបានសមីការលីនេអ៊ែរ $Y = aT + b, b\colon constant$.

H2: វាមានការកើនឡើងពីមួយម៉ោងទៅមួយម៉ោងក្នុងអត្រាថេរ។ សម្មតិកម្មនេះគឺជាគំរូអិចស្ប៉ូណង់ស្យែល (កំណើនក្នុងអត្រាថេរ)។ អត្រាប្រែប្រួលចំនួនបាក់តេរី Y គឺសមមាត្រនឹងពេលវេលាកន្លងទៅ $\frac{dY}{Y} = adT$.

ដូច្នេះយើងបានសមីការអិចស្ប៉ូណង់ស្យែល $Y = e^{aT+b}, b\colon constant$. យើងអាចធ្វើទៅជាលីនេអ៊ែរបានដោយបំលែងជាអនុគមន៍លោការីត $Ln(Y) = aT + b, b\colon constant$.

គំរូមិនលីនេអ៊ែរផ្សេងៗទៀត
Non Linear Modeling

ខ. ធ្វើការវិនិច្ឆ័យតាមមធ្យោបាយដែលអាចរបស់សម្មតិកម្មទាំងពីរនេះ៖

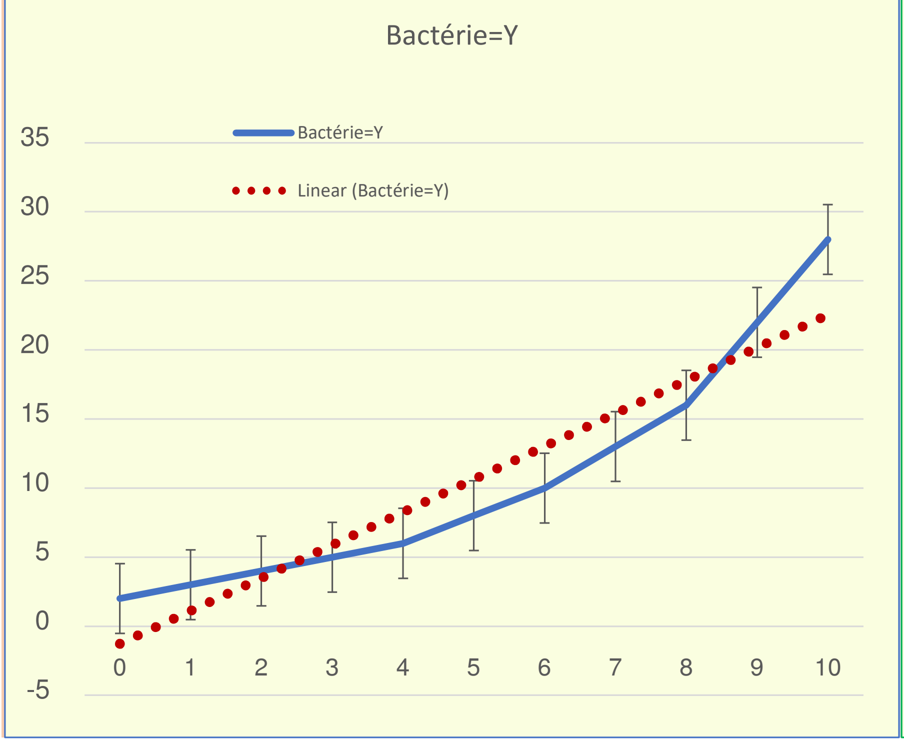
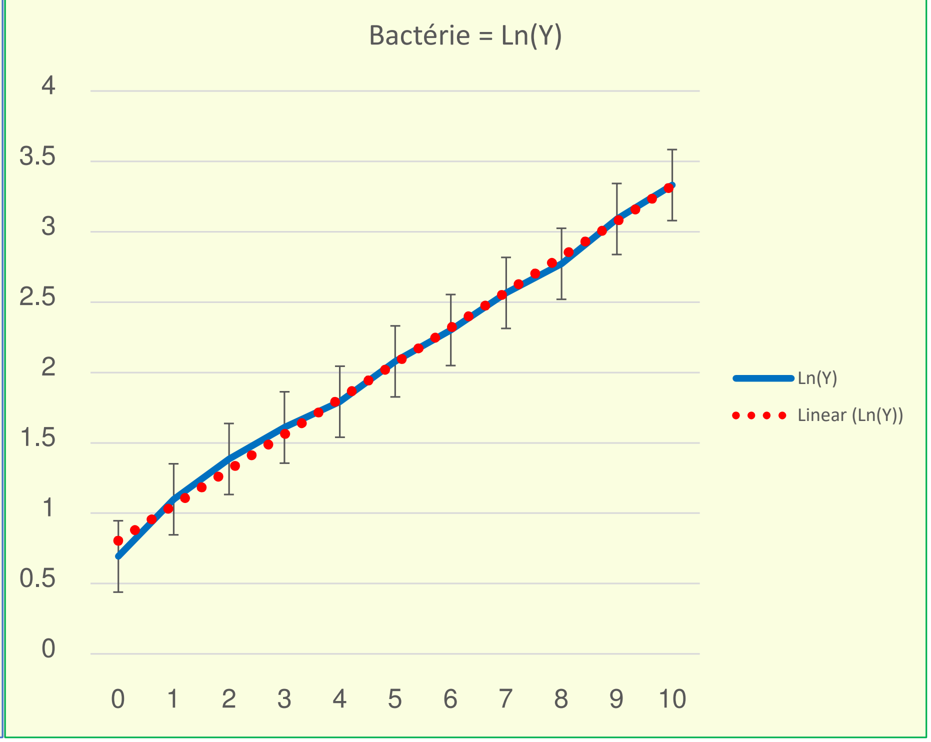
យើងមានមធ្យោបាយពីរដើម្បីវិនិច្ឆ័យសម្មតិកម្មទាំងពីរនេះ(ក្នុងការចាប់ផ្តើម)៖

ទី១ ប្រើវិធីក្រាគិក៖ គឺយើងសង់ក្រាភិកដើម្បីពិនិត្យមើលទម្រង់នៃការវិវត្តន៍របស់ទិន្នន័យជាពិសេសគឺចំនួនបាក់តេរីមានការប្រែប្រួលកើនឡើងបន្ទាប់មកយើងធ្វើការប្រៀបធៀបគ្នាមើលតើគំរូមួយណាដែលខិតជិតភាពជាក់ស្តែងជាង។ លទ្ធផលនោះគឺជាអំណះអំណាងនៃការសម្រេចចិត្តបង្កើតគំរូវិភាគបានល្អជាងគេ។

ទី២ មេគុណរបស់ការកំណត់ទំនាក់ទំនង(R-square)៖ គឺត្រូវគណនារកតម្លៃមេគុណកំណត់ទំនាក់ទំនងដើម្បីធ្វើការប្រៀបធៀបគ្នាមើល តើគំរូវិភាគណាមួយដែលបង្ហាញពីភាពមានការជាប់ពាក់ព័ន្ធខ្លាំងជាងគេសម្រាប់ធ្វើការបកស្រាយកម្រិតទំនាក់ទំនងរវាងអថេរពន្យល់ និងអថេរត្រូវពន្យល់(IV & DV)។

ក្នុងករណីដែលកំណត់បានគំរូវិភាគដែលមានកម្រិតទំនាក់ទំនងល្អជាងគេបានហើយ នៅទីបញ្ចប់គឺគេធ្វើសេចក្តីសន្និដ្ឋានយកគំរូចុងក្រោយនោះសម្រាប់ធ្វើការវាយតម្លៃស្ថានភាពទិន្នន័យនោះតែងម្តង។

គំរូមិនលីនេអ៊ែរផ្សេងៗទៀត
Non Linear Modeling

គំរូលីនេអ៊ែរ	គំរូអិចស្ប៉ូណង់ស្យែល
	

គំរូមិនលីនេអ៊ែរផ្សេងៗទៀត
Non Linear Modeling

គំរូលីនេអ៊ែរ	គំរូអិចស្ប៉ូណង់ស្យែល
សមីការវិភាគគំរូលីនេអ៊ែរ $Y = aT + b, b: constant.$ កំណត់ដោយសមីការវិភាគ៖ $$\underset{(8.457)}{\mathrm{Y} = 2.382T} \underset{(-0.763)}{- 1.27}$$ $$n = 11, R - square = 0.8882, (.) = t - statics$$ គំរូព្យាករណ៍ត្រូវពន្យល់ដោយអថេរអង្កេត 88.82%	សមីការវិភាគគំរូអិចស្ប៉ូណង់ស្យែល $Ln(Y) = aT + b, b: constant$ កំណត់ដោយ៖ $$\underset{(49.070)}{Ln(Y) = 0.252T} \underset{(26.454)}{+ 0.804}$$ $$n = 11, R - square = 0.9962, (.) = t - statics$$ $$\Rightarrow Y = e^{0.252Y+0.804} = 2.234e^{0.252Y}$$ គំរូព្យាករណ៍ត្រូវពន្យល់ដោយអថេរអង្កេត 99.62%

គំរូមិនលីនេអ៊ែរផ្សេងៗទៀត
Non Linear Modeling

តើសម្មតិកម្មណាមួយត្រូវពិចារណាឲ្យប្រហាក់ប្រហែលនឹងស្ថានភាពអាចទទួលយកបាន?
បើពិនិត្យលើក្រាភិកដែលបានបង្ហាញខាងលើ សម្មតិកម្មទី២ ដូចជាខិតទៅជិតស្ថានភាពជាងសម្មតិកម្មទី១។ នៅក្រាភិកទី១ ចំនុចនីមួយៗមិនស្ថិតនៅលើបន្ទាត់ត្រង់ជួរទេ។ គេឃើញចំនុចទាំងនោះរត់នៅលើបន្ទាត់របស់ខ្សែកោងផតបត់ទៅខាងលើ(ផ្នែក Y>0)ដែលអាចនិយាយថាជាខ្សែកោងអ៊ិចស្ប៉ូណង់ស្យែល។ ផ្ទុយមកវិញនៅលើខ្សែកោងទី២(ហៅថា «semi log»)ចំនុចជាច្រើនស្ទើរតែរត់ត្រង់ជួរ។ ដូច្នេះសម្មតិកម្មទី២គឺខិតទៅជិតទំនាក់ទំនងអ៊ិចស្ប៉ូណង់ស្យែលជាងគេបំផុត។
យើងធ្វើការប្រៀបធៀបរវាងមេគុណកំណត់ទំនាក់ទំនងរបស់អំណះអំណាងសម្មតិកម្មទាំងពីរ៖

$$\left.\begin{aligned} H1&: R^2 = 0.8882 \\ H2&: R^2 = 0.9962 \end{aligned}\right\} \Rightarrow$$ សម្មតិកម្មទី២លេចឡើងឲ្យយើងឃើញយ៉ាងច្បាស់គឺគំរូវិភាគខិតជិតបញ្ហា។ ដូច្នេះយើងហៅគំរូវិភាគរបស់សម្មតិម្មទី២ថាជាគំរូអ៊ិចស្ប៉ូណង់ស្យែល ពីព្រោះ R^2 របស់គំរូលីនេអ៊ែរគឺជាផ្នែកមួយនៃការពន្យល់វ៉ារ្យង់របស់ Y និង R^2 របស់គំរូអ៊ិចស្ប៉ូណង់ស្យែលគឺជាផ្នែកមួយនៃការពន្យល់វ៉ារ្យង់របស់លោការីត Y ដែរ។ ដូច្នេះយើងប្រើប្រាស់តម្លៃ R^2 ដែលធំជាងគេក្នុងការកំណត់គំរូវិភាគដែលជាប់នឹងបញ្ហា។

គំរូមិនលីនេអ៊ែរផ្សេងៗទៀត
Non Linear Modeling

គ. ព្យាករណ៍ចំនួនបាក់តេរីនៅម៉ោងទី២៥នៃការពិសោធន៍

យើងមានសមីការព្យាករណ៍៖ $Ln(Y) = 0.252T + 0.804$

បើ $T = 25 \Rightarrow \widehat{Ln(Y)} = 0.252 \times 25 + 0.804 = 7.104 \Rightarrow \hat{Y} = e^{7.104} \cong 1216$

ប៉ុន្តែតម្លៃ 1216 នេះគឺការប៉ាន់ស្មានត្រង់ចំនួនមធ្យមប៉ុណ្ណោះ។ ដូច្នេះដើម្បីបញ្ជាក់ភាពសមហេតុសមផល (Logical calculus)គឺត្រូវគណនាតាមវិធីចន្លោះអង្កត់ជឿជាក់របស់មេគុណរយៈពេល (T) វិញ។

យើងមាន $a \in \left(\hat{a} \pm \sigma_{\hat{a}} \times t_{(n;k)}^{\alpha\%/2}\right) \Rightarrow a \in \left(0.252 \pm 0.0051 \times t_{(9)}^{1\%/2}\right) = (0.252 \pm 0.0051 \times 3.25)$

$\Rightarrow a \in (0.236; 0.269) \Rightarrow \hat{Y}_{T=25} \in (e^{0.236\times25+0.804};\ e^{0.269\times25+0.804}) = (e^{6.692};\ e^{7.527}) = (806; 1858)$

ដូច្នេះចំនួនបាក់តេរីនៅម៉ោងទី២៥ក្នុងបន្ទប់ពិសោធន៍នៃប៉ាន់ស្មានត្រង់កម្រិតហានិភ័យ ១%

គឺនៅចន្លោះ $(806 < \hat{Y}_{T=25} < 1858)$។ ប៉ុន្តែបើប៉ាន់ស្មានត្រង់កម្រិតហានិភ័យ ៥%, តម្លៃពិន្ទុ $t_{(9)}^{5\%/2} = 2.262$

$\Rightarrow \hat{Y}_{T=25} \in (e^{0.241\times25+0.804};\ e^{0.264\times25+0.804}) = (e^{6.82};\ e^{7.401}) = (916; 1638).$

ដូច្នេះចំនួនបាក់តេរីប៉ាន់ស្មាននៅម៉ោងទី២៥គឺនៅចន្លោះ $(916 < \hat{Y}_{T=25} < 1638)$ ត្រង់កម្រិតហានិភ័យ៥%។

The three Logarithmic Regression Model
Y = f(X)

- មានបីករណីដែលត្រូវបានប្រើនៅក្នុងការវិភាគ៖ នៅពេលដែល X ត្រូវបានបំលែងដោយលោការីត, Y មិនបំលែង, នៅពេលដែល Y ត្រូវបានបំលែងដោយលោការីត, X មិនបំលែង និង (X,Y) ត្រូវបានបំលែងដោយលោការីតទាំងពីរ។

➢ ករណីទី១៖ X ត្រូវបានបំលែងដោយលោការីត, Y មិនបំលែង

ក្នុងករណីនេះយើងបានសមីការវិភាគ $Y_i = \beta_0 + \beta_1 Ln(X_i) + u_i, i = 1,2, \ldots, n.$

ពីព្រោះ Y មិនបំលែងដោយលោការីត តែ X បំលែងវិធីនេះហៅថា Linear-log model ។

នៅក្នុងគំរូវិភាគ Linear-log model, បើ X ប្រែប្រួល 1% នាំឲ្យមានជាប់ទាក់ទងក្នុង Y ប្រែប្រួល $0.01\beta_1$។ ការពិនិត្យករណីនេះគឺអាស្រ័យនឹងភាពខុសគ្នានៃការប្រែប្រួលអនុគមន៍នៅពេលដែល X ប្រែប្រួល ΔX ៖

$[\beta_0 + \beta_1 Ln(X + \Delta X)] - [\beta_0 + \beta_1 Ln(X)] = \beta_1[Ln(X + \Delta X) - Ln(X)] \cong \beta_1(\frac{\Delta X}{X})$ ដែលជាជំហ៊ាននៃការប្រើវិធីគណនាតម្លៃប្រហែល។ បើ X ប្រែប្រួល 1% នាំឲ្យ $\frac{\Delta X}{X} = 0.01$ ។ នេះជាការបកស្រាយរបស់គំរូវិភាគខាងលើដែលថា បើ X ប្រែប្រួល 1% នាំឲ្យមានជាប់ទាក់ទងក្នុង Y ប្រែប្រួល $0.01\beta_1$។

The three Logarithmic Regression Model
Y = f(X)

- ឧទាហរណ៍១ សិក្សាទំនាក់ទំនងរវាងការវាយតម្លៃប្រសិទ្ធភាពគ្រប់គ្រង(ដាក់ពិន្ទុTest score)របស់សង្កាត់មួយជាមួយនឹងចំណូលដែលសង្កាត់ប្រមូលបាន។ សមីការវិភាគកំណត់ដូចខាងក្រោម៖

$$\widehat{Testscore} = 557.8 + 36.42Ln(income),\ R - square\ adjusted = 0.561$$
$$(3.8) \quad (1.40) \quad (.) = standard\ error$$

យោងលើសមីការដែលមានស្រាប់នេះ បើចំណូលកើន 1% នាំឲ្យពិន្ទុកើនបាន $0.01 \times 36.42 = 0.36$ point ម្យ៉ាងទៀតយើងក៏អាចគណនាដោយផ្ទាល់តាមរយៈចំណូលកើនបានដែរ។ ឧបមាថាសង្កាត់រកចំណូលបានកើនពី $10,000 ទៅ $11,000។ ចូរគណនាពិន្ទុដែលទទួលបាន។ យើងយក ΔY ជាបម្រែបម្រួលរវាងតម្លៃព្យាករណ៍គឺ $\Delta\hat{Y} = [557.8 + 36.42Ln(11)] - [557.8 + 36.42Ln(10)] = 36.42[Ln(11) - Ln(10)] \rightarrow$ $= 36.42 \times Ln(11/10) = 36.42 \times 0.095310 = 3.47119$ តម្លៃនេះគឺប្រហាក់ប្រហែលនឹងការគណនាខាងលើដែរ(ករណីនេះចំណូលកើន10%)។ បើករណីចំណូលកើនពី $40,000 ទៅ $41,000 យើងបានចំនួនពិន្ទុថយចុះស្មើនឹង $36.42 \times Ln(41/40) = 36.42 \times 0.024692 = 0.90$ point (ករណីនេះចំណូលកើន 2.5%) ។

វិបាក៖ វិធី Linear-log model មានតម្លៃពិន្ទុថយចុះជាបន្តបន្ទាប់កាលណាចំណូលកាន់តែកើនច្រើនឡើង។

The three Logarithmic Regression Model
Y = f(X)

- ករណីទី២៖ Y ត្រូវបានបំលែងដោយលោការីត, X មិនបំលែង

  ក្នុងករណីនេះយើងបានសមីការវិភាគ $\mathrm{Ln}(Y_i) = \beta_0 + \beta_1(X_i) + u_i, i = 1,2,\dots,n.$
  ពីព្រោះ Y បំលែងដោយលោការីត តែ X មិនបំលែង, វិធីនេះហៅថា Log-linear model ។

- វិធី Log-linear model កាលណាកើន X ១ឯកតា($\Delta X = 1$) នាំឲ្យមានកំណើន $100 \times \beta_1\%$ ក្នុង Y ។ ដើម្បីប្រៀបធៀបតម្លៃប៉ាន់ស្មាន $Ln(Y)$ នៅពេល X ប្រែប្រួល ΔX យើងបានកន្សោមបម្រែបម្រួលគឺ $Ln(Y + \Delta Y) - Ln(Y) = [\beta_0 + \beta_1(X + \Delta X)] - [\beta_0 + \beta_1(X)] = \beta_1 \Delta X.$ ដោយសារលក្ខណៈបម្រែបម្រួលក្នុងកន្សោម $Ln(Y + \Delta Y) - Ln(Y) = \Delta Y/Y$. ដូច្នេះយើងអាចទាញបាន $\Delta Y/Y = \beta_1 \Delta X$ បើករណី $\Delta X = 1$ ដូច្នេះបើ កើន X ១ឯកតា នាំឲ្យផលធៀប $\Delta Y/Y$ កើន β_1 ឯកតា។ បកស្រាយជាភាគរយ មានន័យថា បើ X ១ឯកតា នាំឲ្យមានកំណើន $100 \times \beta_1\%$ ក្នុង Y ។

  ឧទាហរណ៍២ សិក្សាពិនិត្យមើលទំនាក់ទំនងរវាងអាយុរបស់បុគ្គលិកនិងការទទួលបានប្រាក់ចំណូល។ លទ្ធផលសិក្សាបច្ចុប្បន្នភាពមួយនៅឆ្នាំ២០០៩ លើចំនួនអង្កេត ១៩៩០៧នាក់ ទាក់ទងជាមួយអាយុនិង

The three Logarithmic Regression Model
Y = f(X)

ចំណូលដែលទទួលបានកំណត់ដោយសមីការវិភាគខាងក្រោមនេះ៖

$$Ln(\widehat{Earning}) = \underset{(0.018)}{2.805} + \underset{(0.0004)}{0.0087}Age, \ \bar{R}^2 = 0.027$$

យោងតាមសមីការវិភាគនេះ ចំណូលត្រូវបានព្យាករណ៍ថាមានកំណើន 0.87% [(100 × 0.0087)%] សម្រាប់ឆ្នាំបន្ថែមនីមួយៗនៃអាយុ។

- ករណីទី៣៖ ទាំង X និង Y ត្រូវបានបំលែងដោយលោការីត

ក្នុងករណីនេះយើងបានសមីការវិភាគកំណត់ដោយ $\text{Ln}(Y_i) = \beta_0 + \beta_1 Ln(X_i) + u_i, i = 1,2, \dots, n.$

ពីព្រោះ ទាំង X និង Y ត្រូវបានបំលែងដោយលោការីត, វិធីនេះហៅថា Log-log linear model ។

វិធី Log-log linear model កាលណាកើន X 1% នាំឲ្យមានកំណើន β_1% ក្នុង Y ។ ដូច្នេះ នៅក្នុងករណីពិសេសនេះ មេគុណជាភាពយឺតនៃ Y ទាក់ទងនឹង X ។

ដើម្បីបកស្រាយករណីនេះឲ្យបានឃើញជាក់ស្តែង យើងប្រើគោលការណ៍បម្រែបម្រួលក្នុង X ដោយ ΔX ដែលនាំឲ្យមានបម្រែបម្រួល Y ក្នុង ΔY ដោយ ។

The three Logarithmic Regression Model
Y = f(X)

យើងបាន $Ln(Y+\Delta Y)-Ln(Y)=[\beta_0+\beta_1 Ln(X+\Delta X)]-[\beta_0+\beta_1 Ln(X)]=\beta_1[Ln(X+\Delta X)-Ln(X)]$ អនុវត្តន៍គោលការណ៍តម្លៃប្រហែលករណីអនុគមន៍លោការីតរបស់អង្គទាំងពីរនៃសមិការ យើងបាន៖

$$\frac{\Delta Y}{Y}\cong\beta_1\frac{\Delta X}{X}\Rightarrow\ \beta_1=\frac{\frac{\Delta Y}{Y}}{\frac{\Delta X}{X}}=\frac{100\times\left(\frac{\Delta Y}{Y}\right)}{100\times\left(\frac{\Delta X}{X}\right)}=\frac{percentage\ chang\ \ in\ Y}{percentage\ chang\ \ in\ X}$$ ។

ដូច្នេះ នៅក្នុងវិធី Log-log linear model, β_1បញ្ជាក់ពី ratio របស់បម្រែបម្រួលជាភាគរយក្នុង Y ធៀបនឹងបម្រែបម្រួលជាភាគរយក្នុង X ។ ប្រសិនបើបម្រែបម្រួលក្នុង X មាន 1% $(i.e\ \Delta X=0.01X)$ នោះ β_1ជាបម្រែបម្រួលភាគរយនៅក្នុង Y ទាក់ទងនឹងបម្រែបម្រួលក្នុង X មាន 1%។ ដូច្នេះ β_1ជាភាពយឺតនៃ Yទាក់ទងនឹង X។ ឧទាហរណ៍៣ ត្រឡប់មកពិនិត្យទំនាក់ទំនងរវាងការវាយតម្លៃប្រសិទ្ធភាពគ្រប់គ្រង(ដាក់ពិន្ទុTest score)របស់សង្កាត់មួយជាមួយនឹងចំណូលដែលសង្កាត់ប្រមូលបានជាថ្មីម្តងទៀត។ សមីការវិភាគកំណត់ដូចខាងក្រោម៖

$$Ln(\widehat{Testscore})=6.336+0.0554Ln(income),\ \ R-square\ adjusted=0.557$$
$$(0.006)\quad(0.0021)\qquad(.)=standard\ deviation$$

The three Logarithmic Regression Model
Y = f(X)

យោងតាមសមីការវិភាគខាងលើ បើមានកំណើន ១% នៅក្នុងចំណូលនាំឲ្យមានកំណើន ០.០៥៥៨%កើនក្នុងពិន្ទុវាយតម្លៃ(test score)។

- ក្នុងការប្រៀបធៀបវិធី Log-log model និងវិធី Log-linear model ករណីឧទាហរណ៍១ដដែល យើងបាន៖

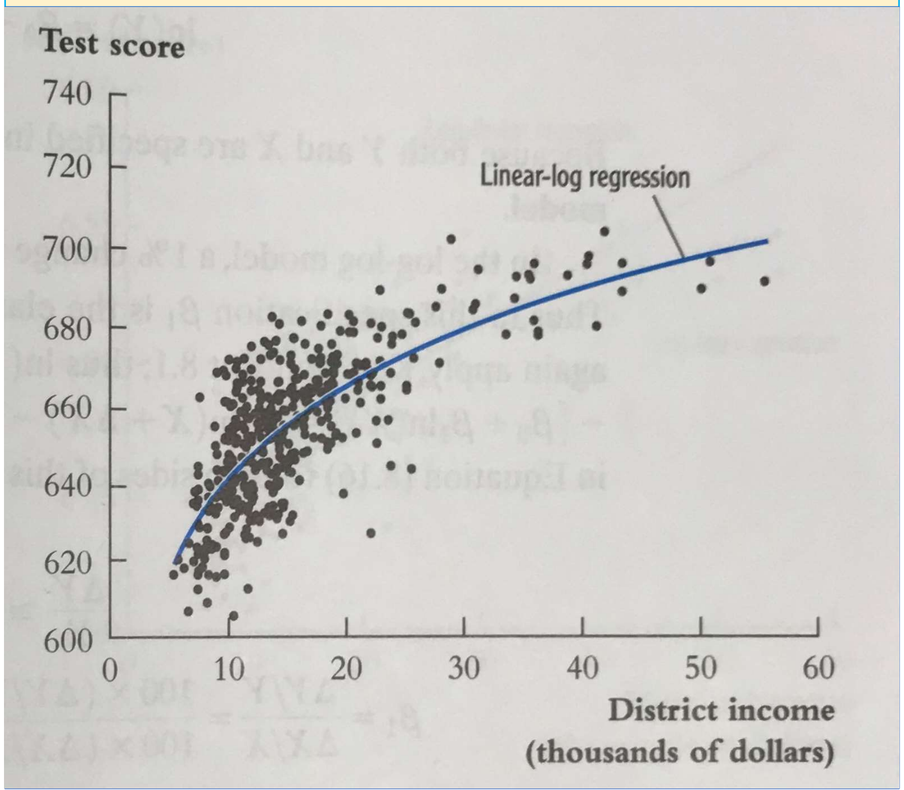
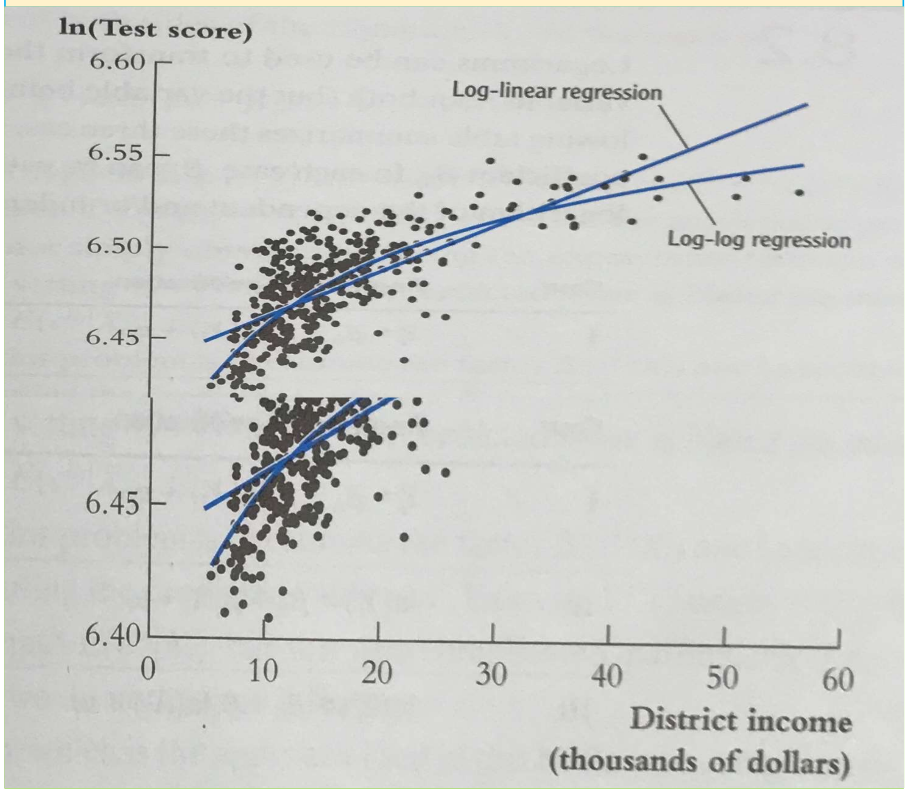
$$Ln(\widehat{Testscore}) = 6.439 + 0.00284(income),\ R - square\ adjusted = 0.497$$
$$(0.003)\quad (0.00018)\qquad (.) = standard\ deviation$$

យើងសង្កេតឃើញថា មេគុណកំណត់ទំនាក់ទំនងកែតម្រូវ($\bar{R}^2$) របស់វិធី Log-log regression ធំជាងវិធី Log-linear regression គឺ $\bar{R}^2_{log-log\ model} = 0.557 > \bar{R}^2_{log-linear\ model} = 0.497$.

❖ ទោះយ៉ាងនេះក្ដី ការបញ្ជាក់កំណត់ហេតុមិនសមនឹងទិន្នន័យជាពិសេសទេ៖ នៅត្រង់តម្លៃទាបរបស់ប្រាក់ចំណូលនៃការអង្កេតភាគច្រើនធ្លាក់ក្រោមខ្សែកោង log-log ខណ:ដែលនៅក្នុងជួរប្រាក់ចំណូលកណ្ដាលភាគច្រើនរបស់ការសង្កេតបានធ្លាក់ចុះលើសតម្លៃអនុគមន៍តំរែតំរង់ដែលបានប៉ាន់ស្មាន។

ជាបន្តយើងពិនិត្យមើលក្រាភិកដែលបកស្រាយប្រៀបធៀបរវាងវិធី log-linear model & log-log model ៖

The three Logarithmic Regression Model
Y = f(X)

វិធី Linear-log regression model	ប្រៀបធៀបវិធី log-linear & log-log regression model
	

The three Logarithmic Regression Model
Y = f(X)

- ជាសង្ខេបយើងបានតារាងបែងចែកភាពផ្សេងគ្នារបស់វិធីលោការីទាំងបីដូចមានខាងក្រោម៖

ករណី	ការបញ្ជាក់ប្រភេទតំរែតំរង់	អត្ថន័យបកស្រាយរបស់មេគុណ β_1
ទី ១	$Y_i = \beta_0 + \beta_1 Ln(X_i) + u_i$ Linear-log regression model	ប្រែប្រួល 1% ក្នុង X គឺជាប់ពាក់ព័ន្ធជាមួយការប្រែប្រួលស្មើនឹង $0.01\beta_1$. នៅក្នុងតម្លៃ Y ។
ទី ២	$Ln(Y_i) = \beta_0 + \beta_1 (X_i) + u_i$ Log-linear regression model	ប្រែប្រួលក្នុង X ១ឯកតា($\Delta X = 1$) គឺជាប់ពាក់ព័ន្ធប្រែប្រួល $100\beta_1$% នៅក្នុងតម្លៃ Y ។
ទី ៣	$Ln(Y_i) = \beta_0 + \beta_1 Ln(X_i) + u_i$ Log-log regression model	ប្រែប្រួល 1% ក្នុង X គឺជាប់ពាក់ព័ន្ធប្រែប្រួលβ_1% នៅក្នុងY ។ ដូច្នេះ β_1គឺជាភាពយឺតរបស់ Y ទាក់ទងនឹង X។