

ជំពូកទី៧ ការជ្រើសរើសគំរូអ្វីបទីម៉ាល់

Model Optimization

រៀបរៀងនិងបង្រៀនដោយបណ្ឌិត ង៉ាន់ ស៊ុនដេត
សាស្ត្រាចារ្យនៃសាកលវិទ្យាល័យភូមិន្ទនីតិសាស្ត្រនិងវិទ្យាសាស្ត្រ
សេដ្ឋកិច្ច

ការជ្រើសរើសយកគំរូប្រសើរមំផុត Optimal selection Model

នៅក្នុងការអនុវត្ត ជារឿយៗសេដ្ឋកិច្ចមាត្រសាស្ត្រត្រូវបានជ្រើសរើសអថេរ
ពន្យល់ជាច្រើន $X_1, X_2, \dots, X_k$ ដើម្បីពន្យល់អថេរ Y ដែលមានជាប់ទាក់ទងគ្នា។
លក្ខណៈនៃវិធីដំណើរការស្ថិតិនាំឲ្យយើងដកចេញអថេរណាខ្លះ ឬបញ្ចូលអថេរ
ណាខ្លះទៅក្នុងគំរូវិភាគសេដ្ឋកិច្ចមួយ។ គំរូវិភាគអថេរពន្យល់នីមួយៗត្រូវមានអត្ថ
ន័យស្ថិតិ ប៉ុន្តែមិនដូចគ្នា។

គំរូអតិប្បរមា R^2 តម្រូវពិនិត្យមើលចំពោះគំរូណាដែលមានតម្លៃ R^2 ខ្ពស់ជាងគេ។

ការជ្រើសរើសយកគំរូប្រសើរបំផុត Optimal selection Model

- យើងនឹងពិនិត្យមើលវិធី៥យ៉ាងដែលនាំឲ្យយើងរក្សាបានគំរូវិភាគល្អប្រសើរហើយវិធីនេះត្រូវផ្សំអថេរនានាដែល ៖

- ❖ មានការពាក់ព័ន្ធខ្លាំងទៅនឹងអថេរត្រូវពន្យល់(ឧទាហរណ៍ជា Y)
- ❖ មានការពាក់ព័ន្ធតិចតួចទៅនឹងអថេរត្រូវពន្យល់

1) ចំនួនគំរូវិភាគដែលអាច

យើងមានរូបមន្តដែលអាចផ្សំអថេរនានាបានជាគំរូវិភាគគឺ $M = 2^K - 1$ ដែល k ជាចំនួនអថេរពន្យល់សរុបនៅក្នុងគំរូវិភាគ M ។

ឧទាហរណ៍ បើយើងយក $k = 10$ ចំនួនអថេរពន្យល់មាន $M = 2^{10} - 1 = 1023$ គំរូវិភាគ។

ការជ្រើសរើសយកគំរូប្រសើរមំផុត Optimal selection Model

2) Backward elimination (ការបំបាត់ត្រឡប់)

នៅក្នុងគំរូវិភាគដែលមាន k អថេរពន្យល់ វិធីនេះមានការលុបបំបាត់ម្តងមួយៗ $t - stat = t^*$ ដែលតូចជាងគេនៅខាងក្រោម $t - stat$ សម្រាប់វាយតម្លៃ(ហានិភ័យ)។

3) Forward Regression (ជ្រើសរើសមេគុណជម្រៀន)

វិធីនេះគឺជ្រើសរើសមេគុណកំណត់ទំនាក់ទំនង $R^2_{yX_j, X_i}, \forall j \neq i$ ដែលខ្ពស់ជាងគេ ហើយគិតដល់អថេរពន្យល់នៅក្នុងគំរូណាដែលមានមេគុណខ្ពស់ជាងគេ។ ការជ្រើសរើសត្រូវបញ្ចប់នៅពេលណាអស់ $t - stat = t^*$ របស់អថេរពន្យល់ដែលទាបជាង $t - stat$ សម្រាប់វាយតម្លៃ(ហានិភ័យ)។

ការជ្រើសរើសយកគំរូប្រសើរមំផុត Optimal selection Model

4) Stepwise regression(ជម្រៀនតាមជំហាន)

វិធីនេះត្រូវបានបង្ហាញនៅពីដំបូងម្តងហើយ លើកលែងតែថាបន្ទាប់ពីលាយបញ្ចូលអថេរពន្យល់ថ្មីបន្ថែម។ យើងពិនិត្យ $t - stat = t^*$ នៃអថេរពន្យល់នីមួយៗនៅក្នុងគំរូវិភាគដែលអាចជ្រើសរើសមកពិនិត្យ និងលុបបំបាត់ចោលពីគំរូវិភាគនៅពេលដែលវាមាន $t - stat = t^*$ នៅខាងក្នុងកម្រិតយកមែន(ហានិភ័យ)។

5) Stagewise regression(ជម្រៀនតាមដំណាក់)

នេះជាដំណើរការមួយនៃការជ្រើសរើសអថេរពន្យល់ដែលអនុញ្ញាតឲ្យអប្បបរមាអន្តរទំនាក់ទំនងរវាងសេរីអថេរពន្យល់តាមការសិក្សាតម្លៃលំអៀង/សំណល់ ៖

- ជំហានទី១. អថេរពន្យល់ដែលមានមេគុណទំនាក់ទំនងខ្ពស់ជាងគេជាមួយអថេរ Y ត្រូវបានជ្រើសរើសយកគឺ X_i ។

ការជ្រើសរើសយកគំរូបស៊េរីបំផុត Optimal selection Model

- ជំហានទី២. គណនាតួលំអៀងពីសមីការព្យាករណ៍ $Y = f(X_i)$: $e_1 = Y - \hat{\beta}_0 - \hat{\beta}_1 X_i$ ។ មេគុណទំនាក់ទំនាក់ទំនងរវាង e_1 និង អថេរពន្យល់ត្រូវបានគណនា យើងត្រឡប់ទៅអថេរពន្យល់ដែលមានមេគុណខ្ពស់ជាងគេ អថេរពន្យល់នោះគឺ X_j ។
- ជំហានទី៣. គណនាតួលំអៀងថ្មីពីសមីការព្យាករណ៍ $e_2 = Y - \hat{\beta}_0 - \hat{\beta}_1 X_i - \hat{\beta}_2 X_j$ ។ មេគុណទំនាក់ទំនាក់ទំនងរវាង e_2 និង អថេរពន្យល់ត្រូវបានគណនា យើងត្រឡប់ទៅអថេរពន្យល់ដែលមានមេគុណខ្ពស់ជាងគេ នេះនាំឲ្យបង្កើតបានជាតួលំអៀងជាថ្មីទៀត។
 - ❑ វិធីនេះត្រូវបញ្ចប់នៅពេលដែលមេគុណទំនាក់ទំនងគ្មានអត្ថន័យខុសពី ០ ។

ការជ្រើសរើសយកគំរូប្រសើរមំផុត

Optimal selection Model

- ឧទាហរណ៍. ឧបមាថាបញ្ហាសេដ្ឋកិច្ច Y មានជាប់ពាន់ព័ន្ធនឹងសេរីអថេរពន្យល់៤ X_1, X_2, X_3, X_4 . តារាងខាងក្រោមនេះជាទិន្នន័យដែលមានការរៀបចំរួចរាល់៖

Y	X1	X2	X3	X4
8.40	82.90	17.10	92.00	94.00
9.60	88.00	21.30	93.00	96.00
10.40	99.90	25.10	96.00	97.00
11.40	105.30	29.00	94.00	97.00
12.20	117.70	34.00	100.00	100.00
14.20	131.00	40.00	101.00	101.00
15.80	148.20	44.00	105.00	104.00
17.90	161.80	49.00	112.00	109.00
19.30	174.20	51.00	112.00	111.00
20.80	184.70	53.00	112.00	111.00

សំណួរ

ដោយប្រើវិធីទាំង៥ដែលបានសិក្សា
ចូរជ្រើសរើសក្នុងចំណោមអថេរ
ពន្យល់មួយ ឬច្រើនដែលអាចពន្យល់
នៅក្នុងគំរូវិភាគនៃសារៈសំខាន់ណា
មួយជាងគេ។

ការជ្រើសរើសយកគំរូបស៊េរីមធ្យម

Optimal selection Model

A. ខាងក្រោមនេះជាតារាងវិភាគមាន $M = 2^4 - 1 = 15$ គំរូវិភាគ

M1	series	X1	0.995	R-square		
	coefs	0.118				
	t-stat	41.9				
M2	series	X2	0.967			
	coefs	0.327				
	t-stat	15.3				
M3	series	X3	0.952			
	coefs	0.516				
	t-stat	12.5				
M4	series	X4	0.978			
	coefs	0.663				
	t-stat	18.6				
M5	series	X1	X2	0.996	R-square	
	coefs	0.132	-0.039			
	t-stat	8.42	0.53			
M6	series	X1	X3	0.996		
	coefs	0.126	-0.036			
	t-stat	8.42	0.53			
M7	series	X1	X4	0.996		
	coefs	0.102	0.09			
	t-stat	5.6	0.86			
M8	series	X2	X3	0.975		
	coefs	0.209	0.194			
	t-stat	2.55	1.48			
M9	series	X2	X4	0.988		
	coefs	0.137	0.395			
	t-stat	2.42	3.45			
M10	series	X3	X4	0.979		
	coefs	-0.166	0.871			
	t-stat	0.73	3.04			
M11	series	X1	X2	X3	0.996	R-square
	coefs	0.139	-0.038	-0.035		
	t-stat	5.58	0.66	0.49		
M12	series	X1	X2	X4	0.996	
	coefs	0.115	-0.028	0.074		
	t-stat	3.55	0.48	0.64		
M13	series	X1	X3	X4	0.998	
	coefs	0.104	-0.188	0.319		
	t-stat	7.46	2.47	2.62		
M14	series	X2	X3	X4	0.993	
	coefs	0.164	-0.3	0.72		
	t-stat	3.32	1.99	3.81		
M15	series	X1	X2	X3	X4	R-square
	coefs	0.097	0.015	-0.199	0.34	0.998
	t-stat	3.66	0.3	2.2	2.27	

ការជ្រើសរើសយកគំរូប្រសើរមំផុត Optimal selection Model

- ដូច្នេះយើងត្រូវពិនិត្យមើលតម្លៃពិន្ទុ (Ratio Student) $t - stat = t^*$ មួយ ឬ ច្រើន ដើម្បីបំបាត់គំរូវិភាគណាដែលអថេរពន្យល់មិនមានអត្ថន័យស្ថិតិគ្រប់គ្រាន់ក្នុងការបដិសេធន័ H_0 ។ ក្នុងករណី

ដែលយើងយកចំនុចគ្រីទិក $t_{(10,1)}^{5\%} = 2.306$:គំរូ M5, M6, M7, M8, M10, M11, M12, M14, M15 ត្រូវបំបាត់ចោល។

- ដោយឡែកគំរូវិភាគ ៦ ដែលនៅសល់, គេពិនិត្យមើលតម្លៃ R^2 ខ្ពស់ជាងគេដូចករណី M13ជ្រើសយក X_1, X_3, X_4 ។ ចេញពីឧទាហរណ៍នេះ យើងចាប់ផ្តើមនូវវិធី

B. វិធីបំបាត់តំរែតតំរង់(Backward elimination)

- មេគុណព្យាករណ៍មានគ្រប់គ្រាន់។ ចេញពីគំរូព្យាករណ៍មាន៤អថេរពន្យល់ X_2 មាន Ratio Student ទាបជាងគេ។
- នៅសល់គំរូមាន ៣អថេរ។ M13 ជាគំរូវិភាគដែលអថេរទាំង៣ X_1, X_3, X_4 មាន $t-stat = t^* > 2.306$ ។

ការជ្រើសរើសយកគំរូប្រសើរបំផុត Optimal selection Model

C. ជ្រើសរើសយកតំរែតំរង់(Forward Regression)

- ❖ គណនាមេគុណទំនាក់ទំនងរវាង Y និង X_1, X_2, X_3, X_4 ។ ក្នុងតារាងវិភាគយើងមាន $R_{yX1} = 0.997$, $R_{yX2} = 0.983$, $R_{yX3} = 0.975$, $R_{yX4} = 0.988$ ។ អថេរ X_1 ត្រូវបានជ្រើសរើសយក។
- ❖ ព្យាករណ៍គំរូដែលមាន២អថេរពន្យល់ X_1 និង X_2 , X_1 និង X_3 , X_1 និង X_4 ។ គំរូ M5, M6, M7 ក្នុងគំរូទាំងបីនេះគ្មានអថេរពន្យល់ថ្មីណាមួយដែលមានអត្ថន័យស្ថិតិអាចបដិសេធន៍សម្មតិកម្ម H_0 បានទេ លើកលែងតែ X_1 ហេតុនេះ X_1 ត្រូវជ្រើសរើសយក។

D. តំរែតំរង់មួយជំហានទៅមួយជំហាន(Stepwise Regression)

- ❖ គណនាមេគុណទំនាក់ទំនងរវាង Y និង X_1, X_2, X_3, X_4 ។ អថេរពន្យល់ដែលត្រូវបានជ្រើសរើសយកគឺ X_1 ។
- ❖ ព្យាករណ៍គំរូដែលមាន២អថេរពន្យល់ X_1 និង X_2 , X_1 និង X_3 , X_1 និង X_4 ដែលក្នុងគំរូនីមួយៗគ្មានអថេរណាមួយមានអត្ថន័យស្ថិតិ ដូច្នេះនីតិវិធីត្រូវបញ្ចប់ ហើយត្រឡប់ជ្រើសរើស X_1 ។

ការជ្រើសរើសយកគំរូប្រសើរមំផុត Optimal selection Model

E. តំរែតំរែតាមដំណាក់(Stagewise Regression)

- ❖ គណនាមេគុណទំនាក់ទំនងរវាង Y និង X_1, X_2, X_3, X_4 ។ អថេរទី១ដែលត្រូវជ្រើសរើសគឺ X_1 ។
- ❖ តំរែតំរែនិងការគណនា $e_1 = Y + 1.24 - 0.11X_1$ (គំរូវិភាគ M1)

Ob	e1	X1	X2	X3	X4
1	-0.123762	82.90	17.10	92.00	94.00
2	0.47523064	88.00	21.30	93.00	96.00
3	-0.12712	99.90	25.10	96.00	97.00
4	0.23651925	105.30	29.00	94.00	97.00
5	-0.4247536	117.70	34.00	100.00	100.00
6	0.00791332	131.00	40.00	101.00	101.00
7	-0.4190136	148.20	44.00	105.00	104.00
8	0.07829998	161.80	49.00	112.00	109.00
9	0.01702708	174.20	51.00	112.00	111.00
10	0.2796589	184.70	53.00	112.00	111.00

- ❖ គណនាមេគុណទំនាក់ទំនងរវាង e_1 និង X_1, X_2, X_3, X_4 យើងបាន $R_{e_1X_1} = 0.000$, $R_{e_1X_2} = 0.0016$, $R_{e_1X_3} = 0.0015$, $R_{e_1X_4} = 0.0023$ ។ មេគុណទំនាក់ទំនងរវាង e_1 និង X_1 គឺស្របតាមសំណង់សម្មតិកម្ម ០ ពីព្រោះថាព័ត៌មានទំនាក់ទំនងពី X_1 គឺគ្មានសោះនៅក្នុងសំណល់(ឬថាវិវិទ្យា e_1 និង X_1 គឺកែងគ្នា)។
- ❖ មេគុណទំនាក់ទំនងផ្សេងទៀតគ្មានអត្ថន័យស្ថិតិខុសពី ០ ។ ដូច្នេះដំណើរការនីតិវិធីជ្រើសរើសគំរូវិភាគត្រូវបញ្ចប់។ អថេរប្រសើរដែលត្រូវជ្រើសរើសគឺ X_1 ។

សម្គាល់: ក្នុងការអនុវត្តវិធីប្រើញឹកញាប់គឺ **Stepwise Regression** ពីព្រោះវាមានជាស្រេចនៅក្នុងកម្មវិធីវិភាគទិន្នន័យ។

ការជ្រើសរើសយកគំរូបស៊េរីបំផុត Optimal selection Model

- សិក្សាឧទាហរណ៍ខាងក្រោមនេះនិងធ្វើការសន្និដ្ឋានវាយតម្លៃគំរូវិភាគថ្មី៖

$$1. \log P_t = 9.213 + 0.148WL - 0.33DA - 0.0057(D75) - 0.203\log A$$

$$(0.204) \quad (0.099) \quad (0.011) \quad (0.0142) \quad (0.0345)$$

$n = 50, R^2 = 0.559, (.) = \text{standard error}$

ចូរធ្វើតេស្តមេគុណនៃអថេរទាំងនេះ និងសន្និដ្ឋានសមីការថ្មីត្រង់ហានិភ័យ ៥% ។

$$2. \hat{Y} = 0.171 + 7.275X_1 + 0.547X_2 + 0.00073X_3 + 0.0638X_4 - 0.0043X_5 + 0.857X_6.$$

$$(2.97) \quad (2.32) \quad (1.34) \quad (3.26) \quad (2.24) \quad (1.80)$$

$n = 64, R\text{-square} = 0.897, (.) = t\text{-statistics}$

Y =change in post-proposition 13 mean house prices,

X_1 =post-proposition 13 decrease in property tax bill on mean house, X_4 =mean age of house,

X_2 =mean square footage of house, X_5 =transportation time to San Francisco,

X_3 =median income of the families in the area, X_6 =housing-quality index as computed by real estate appraisers.

ចូរធ្វើតេស្តមេគុណនៃអថេរទាំងនេះ និងសន្និដ្ឋានសមីការថ្មីត្រង់ហានិភ័យ ៥% ។

ការជ្រើសរើសយកគំរូបស៊េរីមធ្យម Optimal selection Model

- នៅសហរដ្ឋអាមេរិកគេធ្វើការអង្កេតមួយតាក់ទងនឹងទិន្នន័យប្រើប្រាស់អាហាររបស់មនុស្សម្នាក់ៗ ដែលមានជាប់ពាក់ព័ន្ធជាមួយថ្លៃអាហារ និងចំណូលក្នុងបុគ្គលម្នាក់ៗ។ គេមានគំរូវិភាគពីរដែលអាច៖

$$(1). \log Q = \alpha + \beta_1 \log P + \beta_2 \log Y$$

$$(2). \log Q = \alpha + \beta_1 \log P + \beta_2 \log Y + \beta_3 \log P \cdot \log Y$$

where:

Q = food consumption per capita,

P = food price,

Y = consumer income,

ចូរសរសេរសមីការវិភាគទាំងពីរនេះនៅចន្លោះឆ្នាំ1927-1941 មួយ និង 1948-1962 មួយ និងចំនួនអង្កេតទាំងអស់មួយ។ បន្ទាប់មកគណនាគណនាភាពយឺតសមីការ(១)ធៀបនឹងថ្លៃនិងចំណូល និងសមីការ(២)ធៀបនឹងថ្លៃនិងចំណូលតែវាមានលក្ខណៈជាអន្តរកម្មផលគុណថ្លៃនិងចំណូល។

Year	Food Consumption per Capita, Q	Food Price, P	Consumer Income, Y
1927	88.9	91.7	57.7
1928	88.9	92.0	59.3
1929	89.1	93.1	62.0
1930	88.7	90.9	56.3
1931	88.0	82.3	52.7
1932	85.9	76.3	44.4
1933	86.0	78.3	43.8
1934	87.1	84.3	47.8
1935	85.4	88.1	52.1
1936	88.5	88.0	58.0
1937	88.4	88.4	59.8
1938	88.6	83.5	55.9
1939	91.7	82.4	60.3
1940	93.3	83.0	64.1
1941	95.1	86.2	73.7

(World War II year excluded)			
1948	96.7	105.3	82.1
1949	96.7	102.0	83.1
1950	98.0	102.4	88.6
1951	96.1	105.4	88.3
1952	98.1	105.0	89.1
1953	99.1	102.6	92.1
1954	99.1	101.9	91.7
1955	99.8	100.8	96.5
1956	101.5	100.0	99.8
1957	99.9	99.8	99.9
1958	99.1	101.2	98.4
1959	101.0	98.8	101.8
1960	100.7	98.4	101.8
1961	100.8	98.8	103.1
1962	101.0	98.4	105.5

ការជ្រើសរើសយកគំរូបស្ទើរដ៏ល្អបំផុត

Optimal selection Model

- បំរែបំរួលជាលោការីតទិន្នន័យឆ្នាំ១៩២៧ដល់១៩៤១

Year	$\ln(Q_i)$	$\ln(P_i)$	$\ln(Y_i)$	$\ln(P_i)\ln(Y_i)$
1927	4.487512	4.518522	4.055257	18.32377
1928	4.487512	4.521789	4.082609	18.460696
1929	4.489759	4.533674	4.127134	18.711083
1930	4.485260	4.509760	4.030695	18.177465
1931	4.477337	4.410371	3.964615	17.485425
1932	4.453184	4.334673	3.793239	16.442452
1933	4.454347	4.360548	3.779634	16.481273
1934	4.467057	4.434382	3.867026	17.147868
1935	4.447346	4.478473	3.953165	17.704141
1936	4.483003	4.477337	4.060443	18.179971
1937	4.481872	4.481872	4.091006	18.335364
1938	4.484132	4.424847	4.023564	17.803655
1939	4.518522	4.411585	4.099332	18.084554
1940	4.535820	4.418841	4.160444	18.384341
1941	4.554929	4.456670	4.300003	19.163694

- បំរែបំរួលជាលោការីតទិន្នន័យឆ្នាំ១៩៤៨ដល់១៩៦២

1948	4.571613	4.656813	4.407938	20.526945
1949	4.571613	4.624973	4.420045	20.442587
1950	4.584967	4.628887	4.484132	20.756538
1951	4.565389	4.657763	4.480740	20.870224
1952	4.585987	4.653960	4.489759	20.895162
1953	4.596129	4.630838	4.522875	20.944701
1954	4.596129	4.623992	4.518522	20.893611
1955	4.603168	4.613138	4.569543	21.079934
1956	4.620059	4.605170	4.603168	21.198373
1957	4.604170	4.603168	4.604170	21.193767
1958	4.596129	4.617099	4.589041	21.188055
1959	4.615121	4.593098	4.623010	21.233937
1960	4.612146	4.589041	4.623010	21.215182
1961	4.613138	4.593098	4.635699	21.29222
1962	4.615121	4.589041	4.658711	21.379015

ការជ្រើសរើសយកគំរូបស្ទើរដុត Optimal selection Model

- បន្ទាប់ពីបំប្លែងទិន្នន័យជាអនុគមន៍លោការីតរួចជាបន្ទាប់ត្រូវគណនាមេគុណ β_0 , β_1 , β_2 និង β_3 សម្រាប់ឆ្នាំ 1927-1942 និង 1948-1962 តាមលំដាប់លំដោយ និងបញ្ចូលគ្នាទាំងពីររយៈពេលដែលមានបង្ហាញនៅក្នុងតារាងវិភាគលទ្ធផលណែនាំបន្តបន្ទាប់ខាងមុខ។
- នៅក្នុងតារាងមានបង្ហាញចំនុចសំខាន់សម្រាប់វាយតម្លៃនិងវិភាគតេស្តបានដូចជា parameters estimation, R-square, F-fisher, t-ratio student and p-values. ដែលធាតុទាំងអស់នេះអាចឲ្យយើងវិភាគប្រៀបធៀបបំប្លែងលក្ខណៈតម្រូវការពីកំឡុងពេលមួយទៅកំឡុងពេលមួយទៀតបានជាពិសេសរវាងឆ្នាំ 1927-1942 និង 1948-1962។
- ការគណនានិងបកស្រាយភាពយឺតអាចធ្វើបានតាមទ្រឹស្តីនិងអាចប្រៀបធៀបជាមួយលទ្ធផលជាក់ស្តែងដែលបានរកឃើញតាមរយៈទិន្នន័យអង្កេត។
- ភាពយឺតរបស់តម្រូវការធៀបនឹងចំណូលជាចំនួនវិជ្ជមាន។
- ភាពយឺតរបស់តម្រូវការធៀបនឹងថ្លៃទំនិញជាចំនួនអវិជ្ជមាន។

ការជ្រើសរើសយកគំរូបស្ទើរដុត Optimal selection Model

- តារាងលទ្ធផល

<i>Equatyon1: $\ln(Q_i) = \beta_0 + \beta_1 \ln P_i + \beta_2 \ln Y_i$</i>									
	1927-1941			1948-1962			All obs.(1927-1962)		
	Coeff.	t-stat	p-valu.	Coeff.	t-stat	p-valu.	Coeff.	t-stat	p-valu.
Intercept	4.555	22.674	0.000	5.052	5.605	0.000	4.047	29.762	0.000
Price	-0.235	-4.406	0.001	-0.237	-1.540	0.149	-0.119	-2.946	0.007
Incomes	0.243	10.626	0.000	0.141	3.035	0.010	0.241	17.949	0.000
n	15			15			30		
d.f.	12			12			27		
R-square	0.907			0.874			0.973		
F-fisher	58.235		0.000	41.675		0.000	488.315		0.000

ការជ្រើសរើសយកគំរូបស្នើរបំផុត Optimal selection Model

- តារាងលទ្ធផល

<i>Equation2: $\ln(Q_i) = \beta_0 + \beta_1 \ln P_i + \beta_2 \ln Y_i + \beta_3 \ln P_i \ln Y_i$</i>									
	1927-1941			1948-1962			All obs. (1927-1962)		
	Coeff.	t-stat	p-valu.	Coeff.	t-stat	p-valu.	Coeff.	t-stat	p-valu.
Intercep	4.058	0.539	0.600	16.632	0.611	0.554	8.029	4.465	0.000
Prices	-0.123	-0.072	0.944	-2.745	-0.465	0.651	-0.996	-2.509	0.019
Incomes	0.368	0.195	0.849	-2.416	-0.402	0.695	-0.718	-1.661	0.109
Interaction	-0.028	-0.066	0.948	0.554	0.425	0.679	0.211	2.220	0.035
n	15			15			30		
d.f.	11			11			26		
R-square	0.907			0.876			0.977		
F-fisher	35.603		0.000	25.947		0.000	374.547		0.000

ការជ្រើសរើសយកគំរូប្រសើរមំផុត

Optimal selection Model

- Income elasticity: $\eta_{QY} = \frac{dLn(Q)}{dLn(Y)} = \beta_2 + \beta_3 LnP = -0.718 + 0.211LnP$. កន្សោមភាពយឺតជាអនុគមន៍នៃ $Ln(P)$ មានន័យថាបើលំដាប់ $Ln(P)$ កើនពី $Ln(P) = 4.33$ (1932) ទៅ $Ln(P) = 4.66$ (1951) នាំឲ្យភាពយឺតនៃតម្រូវការធៀបនឹងចំណូលកើនពី $\eta_{QY} = 0.195$ ទៅ $\eta_{QY} = 0.265$ ។
- Price elasticity: $\eta_{QP} = \frac{dLn(Q)}{dLn(P)} = \beta_1 + \beta_3 LnY = -0.996 + 0.211LnY$. កន្សោមភាពយឺតជាអនុគមន៍នៃ $Ln(Y)$ មានន័យថាបើចំណូលកើន នាំឲ្យតម្រូវការអាហារមានផ្លែទាបជាងមុន។ បើលំដាប់ $Ln(Y)$ កើនពី $Ln(Y) = 3.78$ (1933) ទៅ $Ln(Y) = 4.66$ (1962) នាំឲ្យភាពយឺតនៃតម្រូវការធៀបនឹងចំណូលកើនពី $\eta_{QY} = -0.198$ ទៅ $\eta_{QY} = -0.013$ ។

ជីវិតមានន័យ



គឺពោរពេញទៅដោយសេរី
ភាព និងឯករាជ្យភាព។



Table 4.3 Indexes of food consumption, food price, and consumer income^a

Year	Food Consumption per Capita, q	Food Price, p^b	Consumer Income, y^c
1927	88.9	91.7	57.7
1928	88.9	92.0	59.3
1929	89.1	93.1	62.0
1930	88.7	90.9	56.3
1931	88.0	82.3	52.7
1932	85.9	76.3	44.4
1933	86.0	78.3	43.8
1934	87.1	84.3	47.8
1935	85.4	88.1	52.1
1936	88.5	88.0	58.0
1937	88.4	88.4	59.8
1938	88.6	83.5	55.9
1939	91.7	82.4	60.3
1940	93.3	83.0	64.1
1941	95.1	86.2	73.7
(World War II years excluded)			
1948	96.7	105.3	82.1
1949	96.7	102.0	83.1
1950	98.0	102.4	88.6
1951	96.1	105.4	88.3
1952	98.1	105.0	89.1
1953	99.1	102.6	92.1
1954	99.1	101.9	91.7
1955	99.8	100.8	96.5
1956	101.5	100.0	99.8
1957	99.9	99.8	99.9
1958	99.1	101.2	98.4
1959	101.0	98.8	101.8
1960	100.7	98.4	101.8
1961	100.8	98.8	103.1
1962	101.0	98.4	105.5

^a1957–1959 = 100.^bRetail prices of Bureau of Labor Statistics, deflated by dividing by consumer price index.^cPer capita disposable income, deflated by dividing by consumer price index.

these data:

Equation 1: $\log q = \alpha + \beta_1 \log p + \beta_2 \log y$

Equation 2: $\log q = \alpha + \beta_1 \log p + \beta_2 \log y + \beta_3 \log p \log y$

The last term in equation 2 is an interaction term that allows the price and income elasticities to vary. Equation 1 implies constant price and income elasticities.

The estimates of α , β_1 , β_2 , and β_3 for 1927–1942 and 1948–1962 separately and for both periods combined are presented in Table 4.4. Note that we have presented the t -ratios in parentheses because this is more convenient to interpret. If we are interested in significance tests, the t -ratios are more convenient and if we are interested in obtaining confidence intervals, it is convenient to have the standard errors.

The interaction term has a very low t -value in the equations for 1927–1942 and 1948–1962 separately. Hence we will consider only the equation for the combined data. For this equation the income coefficient has a low t -value and the sign also appears to be wrong at first sight. However, the equation does give positive income elasticities. The income elasticities are given by

$$\begin{aligned}\eta_{qy} &= \frac{d \log q}{d \log y} = \beta_2 + \beta_3 \log p \\ &= -0.718 + 0.211 \log p\end{aligned}$$

$\log p$ ranges from 4.33 to 4.66 in the data. Thus the income elasticity ranges from 0.195 to 0.265.

The price elasticity is given by

$$\eta_{qp} = \frac{d \log q}{d \log p} = \beta_1 + \beta_3 \log y = -0.996 + 0.211 \log y$$

Thus, as income increases, demand for food becomes more price inelastic. For the data in Table 4.3, $\log y$ ranges from 3.78 to 4.66. Thus the price elasticity ranges from -0.198 to -0.013 .

Table 4.4 Estimates of demand for food equations^a

	Equation 1			Equation 2 ^b		
	1927–1941	1948–1962	All Obs.	1927–1941	1948–1962	All Obs.
α	4.555 (22.67)	5.052 (5.61)	4.050 (29.65)	4.058 (0.54)	16.632 (0.61)	8.029 (4.47)
β_1 (price)	-0.235 (-4.41)	-0.237 (-1.54)	-0.120 (-2.95)	-0.123 (-0.07)	-2.745 (-0.47)	-0.996 (-2.51)
β_2 (income)	0.243 (10.63)	0.141 (3.03)	0.242 (17.85)	0.368 (0.20)	-2.416 (-0.40)	-0.718 (-1.66)
β_3 (interaction)				-0.028 (-0.07)	0.554 (0.43)	0.211 (2.22)
n	15	15	30	15	15	30
R^2	0.9066	0.8741	0.9731	0.9066	0.8762	0.9775
F^c	58.2	41.7	489	35.6	26.0	374.5
$10^2 \times \text{RSS}$	0.1151	0.0544	0.2866	0.1151	0.0535	0.2412
d.f.	12	12	27	11	11	26
$10^4 \times \hat{\sigma}^{2d}$	0.9594	0.4534	1.0613	1.0462	0.4866	0.9278

^aFigures in parentheses are t -ratios, not standard errors.

^bFrom equation 2, all the partial r^2 's are very low for 1927–1941 and 1948–1962, although the R^2 's are high.

^c F is an F -statistic for testing overall fit, that is, for R^2 . This is discussed in Section 4.8.

^d $\hat{\sigma}^2 = \text{RSS}/\text{d.f.}$ The importance of this is discussed in Section 4.9.