

STICHTING
MATHEMATISCH CENTRUM
2e BOERHAAVESTRAAT 49
AMSTERDAM
REKENAFDELING

R 239

Analyse van Nederlandse huishoudrekeningen.

door

[Staf van de Rekenafdeling]



[1952]

Analyse van Nederlandse Huishoudrekeningen.

Toelichting bij de af te leveren resultaten.

1. Code.

Het onderzoek strekt zich uit over 32 goederencatagorieën, elk voor 5 groepen van gezinnen.

Het verbruik per eenheidsconsument - y - wordt geschat door 6 verschillende functies.

Tijdens de berekeningen en bij de afgedrukte resultaten wordt onderstaande code toegepast.

Goederencatagorieën.

GC = 1	Peulvruchten	12	Margarine	23	Kippeneieren
2	Appelen	13	Boter	24	Zachte zeep
3	Peren	14	Rundvlees	25	Huishoudzeep
4	Aardbeien	15	Kalfsvlees	26	Zeeppoeder
5	Kersen	16	Varkensvlees	27	Toiletzeep
6	Sinaasappelen	17	Spek	28	Scheerzeep
7	Bananen	18	Vleeswaren	29	Sigaren
8	Kristalsuiker	19	Verse zeevis	30	Sigaretten
9	Thee	20	Koemelk	31	Tabak
10	Koffie	21	Karnemelk	32	Schoenen
11	Cacao	22	Kaas		

Groepen.

Gr = 1	Hoofdarbeiders in grotere gemeenten
2	Hoofdarbeiders in kleine gemeenten
3	Handarbeiders in grotere gemeenten
4	Handarbeiders in kleine gemeenten
5	De groepen 1,2,3 en 4 gezamenlijk

Functies.

F = 1	$y = a + bx + cx^2 + kn$
2	$y = a + bx + kn$
3	$y = a + bx + cx^2$
4	$y = a + bx$
5	$y = \alpha + \beta \log x + \delta \log n$
6	$y = \alpha + \beta \log x$

2. Schaalfactoren

Teneinde bij de berekening optimale nauwkeurigheid te bereiken worden tweeërlei schaalfactoren ingevoerd, nl. in de te verklaren variabelen en in de verklarende variabelen.

Te verklaren variabelen.

De verbruikte hoeveelheden per gezin - yn - worden per goederencatagorie met een factor f vermenigvuldigd, zodanig, dat deze voor alle goederencatagorieën gemiddeld van ongeveer dezelfde grootte worden.

Deze factoren zijn als volgt:

GC	f	GC	f	GC	f	GC	f	GC	f
1	10	8	1	15	10	21	1	27	10
2	10	9	10	16	1	22	1	28	1
3	10	10	10	17	1	23	1	29	1
4	10	11	10	18	1	24	10	30	1
5	10^2	12	1	19	10	25	10	31	10
6	10	13	1	20	10^{-1}	26	10	32	10^2
7	10	14	1						

Verklarende variabelen.

De variabelen in het rechterlid (1) en (2) worden met zodanige factoren m vermenigvuldigd dat deze onderling en hun quadraten en producten zoveel mogelijk van dezelfde grootte- orde worden.

Deze factoren worden aldus gekozen:

Variabele	m
1	10^{-1}
x	10^{-4}
x^2	10^{-7}
n	10^{-2}
log x	10^{-1}
log n	10^{-1}

In de optredende matrices en vectoren verschijnen deze schaalfactoren als de elementen m_k ($k=1(1)4,3$ of 2) van diagonaal matrices M, waarmede deze matrices en vectoren per functie links - of rechts - vermenigvuldigd worden.

Deze matrices zijn de volgende:

F	M
1	$\begin{bmatrix} 10^{-1} & & & \\ & 10^{-4} & & \\ & & 10^{-7} & \\ & & & 10^{-2} \end{bmatrix}$
2	$\begin{bmatrix} 10^{-1} & & & \\ & 10^{-4} & & \\ & & 10^{-2} & \\ & & & \end{bmatrix}$
3	$\begin{bmatrix} 10^{-1} & & & \\ & 10^{-4} & & \\ & & 10^{-7} & \\ & & & \end{bmatrix}$
4	$\begin{bmatrix} 10^{-1} & & & \\ & 10^{-4} & & \\ & & & \\ & & & \end{bmatrix}$
5	$\begin{bmatrix} 10^{-1} & & & \\ & 10^{-1} & & \\ & & 10^{-1} & \\ & & & \end{bmatrix}$
6	$\begin{bmatrix} 10^{-1} & & & \\ & 10^{-1} & & \\ & & & \\ & & & \end{bmatrix}$

Als gevolg van de invoering van deze beide soorten schaalfactoren worden de formules (3) t/m (6) in onderstaande gewijzigde vorm geschreven om voor de berekeningen te kunnen dienen.

Uitgaande van XM i.p.v. X en fy i.p.v. y heeft men:

$$(3) \quad \begin{aligned} fy &= (XM)(M^{-1}f\beta) + fu \\ M^{-1}fb &= (MX'XM)^{-1}(MX'fy) \end{aligned}$$

$$(4) \quad \begin{aligned} f^2Ts^2 &= fy' \cdot fy - (fb'M^{-1})(MX'fy) \\ f^2\sigma^2 &= \frac{f^2Ts^2}{T-\hat{n}} \end{aligned}$$

$$(5) f^2M^{-1}C_bM^{-1} = (f^2\sigma^2)(MX'XM^{-1})$$

standaard-

$$f_{out} = f m_k^{-1} (C_b)_{kk}^{\frac{1}{2}}$$

$$(6) \quad 1-R_{rel}^2 = \frac{T}{T-\hat{n}} \cdot \frac{f^2Ts^2}{fy' \cdot fy} = \frac{T \cdot f^2\sigma^2}{fy' \cdot fy}$$

$$f\bar{y} = \frac{\sum fy_i}{T}$$

$$f^2s_y^2 = \frac{1}{T} fy' \cdot fy - (f\bar{y})^2$$

$$1-R_{rel}^2 = \frac{f^2\sigma^2}{f^2s_y^2}$$

3. Nauwkeurigheid.

Er wordt gestreefd naar een nauwkeurigheid in de eindresultaten van gemiddeld 3 à 4 significante cijfers. In het begin zijn wat meer cijfers meegenomen omdat bij matrixprocessen als deze vaak enkele cijfers vooraan wegvallen.

Mede omdat de plaats van het decimaalteken bij het afdrukken niet door een komma kan worden aangegeven is bij de berekening als conventie aangenomen, dat steeds met gehele getallen gewerkt wordt.

Hiertoe worden de optredende grootheden van de volgende factoren voorzien:

grootheid	factor
$MX'XM$	10^5
$(MX'XM)^{-1}$	10^5
$MX'fy$	10
$fy' \cdot fy$	10^{-2}
$M^{-1}fb$	10^2
$f^2\sigma^2$	1
$f^2M^{-1}C_bM^{-1}$	10^{-1}
$f m_k^{-1} (C_b)_{kk}^{\frac{1}{2}}$	10
R_{abs}^2	10^4
R_{abs}	10^4
$f\bar{y}$	10^2
$f^2s_y^2$	1
$f s_y$	10^2
R_{rel}^2	10^5
R_{rel}	10^4

4. Checks.

Op de berekening worden de volgende checks toegepast:

(a) ponsen van x, n en yn.

De geponste getallen worden per gezin afgedrukt en proefgelezen met de oorspronkelijke lijsten.

(b) log x en log n.

De opgezochte en geponste waarden worden per gezin afgedrukt en vergeleken met een lijst van onafhankelijk opgezochte waarden.

(c) matrices X'X.

De quadraten en kruisproducten der verklarende variabelen worden per gezin machinaal gechecked op

$$\frac{(nx^2)^2}{n^2 \cdot x^4} = 1 \text{ en } \frac{(\log x \log n)^2}{\log^2 x \cdot \log^2 n} = 1.$$

De sommaties worden tweemaal uitgevoerd.

(d) matrices (X'X)⁻¹.

De inversies worden met de hand uitgevoerd, waarbij de gebruikelijke checkmethoden worden toegepast.

De geponste elementen worden afgedrukt en proefgelezen.

(e) de vectoren X'y.

Aan de matrices X' wordt een extra kolom toegevoegd, bestaande uit een lineaire combinatie van de andere kolommen. Dit leidt tot een extra element in X'y, dat eenzelfde combinatie van de overige elementen is, hetgeen per groep en per goederencategorie gechecked is.

De sommaties worden bovendien tweemaal uitgevoerd en de daarbij geponste totaalkaarten afgedrukt en proefgelezen.

(f) de vectoren b.

Aan (X'X)⁻¹ wordt een somrij toegevoegd, hetgeen een som-element in b geeft, waarop elke b gechecked wordt.

Aan de goederencategorieen werd GC33 toegevoegd waarvan de X'y gevormd wordt uit de som van de andere 32. Per groep en per functie geeft dit eenzelfde somcheck op de b.

(g) substitutie van de parameters.

Per groep, functie en goederencategorie wordt gechecked op $X'y = X'X \cdot b$ waaruit blijkt dat door de toegepaste matrixinversie en -vermenigvuldiging gemiddeld een decimaal vectoren gaat.

(h) σ^2 en R.

Een globale check op σ^2 en R wordt verkregen door deze grootheden eveneens te berekenen voor een benadering met slechts een parameter (de gemiddelde y-waarden).

(i) R en $(C_b)_{kk}^{\frac{1}{2}}$.

De berekeningen van de vierkantswortels voor deze grootheden worden gechecked door de quadraten daarvan machinaal met de oorspronkelijke getallen te vergelijken

5. Afdrukken der resultaten.

De onderscheidene resultaten worden als volgt afgedrukt, alle in 12-voud:

(a) $10MX'fy$

Sortering: verticaal naar GC, Gr, F.
Alleen de vectoren voor F=1 en F=5
zijn afgedrukt; de overige ontstaan hieruit door
weglaten van 1 of 2 elementen.

(b) $10^5 MX'XM$

Sortering: verticaal naar Gr, F, rijnummer.
horizontaal naar kolomnummer.

(c) $10^5 (MX'XM)^{-1}$

Sortering: verticaal naar Gr, F, rijnummer.
horizontaal naar kolomnummer.

Negatieve getallen zijn aangeduid door een sterretje
erachter.

(d) $10^4 R_{abs}$

Sortering: verticaal naar GC, F.
horizontaal naar Gr.

(e) $10^4 R_{rel}$

Sortering: verticaal naar GC, F.
horizontaal naar Gr.

In gevallen met $R_{rel}^2 < 0$ is alleen een sterretje af-
gedrukt in plaats van R_{rel} .

(f) $10^2 M^{-1} f_b, 10 f_m k^{-1} (C_b)_{kk}^2, 10^4 R_{rel}, f^2 \sigma^2, 10^2 f \bar{y}, 10^2 f s_y$

Sortering: verticaal naar GC, Gr en dan als volgt:

F					
1	a	b	c	k	R
	s_a	s_b	s_c	s_k	σ^2
2	a	b		k	R
	s_a	s_b		s_k	σ^2
3	a	b	c		R
	s_a	s_b	s_c		σ^2
4	a	b			R
	s_a	s_b			σ^2
5	α	β		f	R
	s_α	s_β		s_f	σ^2
6	α	β			R
	s_α	s_β			σ^2
	-				
	y	s_y			

Negatieve getallen zijn aangeduid door een sterretje er-
 achter. Parameter en standaardfouten zijn zodanig afgedrukt,
 dat de plaats van het decimaalteken voor beide dezelfde is.