

Het schatten van marktpenetratie en marktaandeel

Citation for published version (APA):

Wijnen, J. T. M. (1994). *Het schatten van marktpenetratie en marktaandeel*. (IWDE report; Vol. 9403). Technische Universiteit Eindhoven.

Document status and date:

Gepubliceerd: 01/01/1994

Document Version:

Uitgevers PDF, ook bekend als Version of Record

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.tue.nl/taverne

Take down policy

If you believe that this document breaches copyright please contact us at:

openaccess@tue.nl

providing details and we will investigate your claim.

Instituut
Wiskundige Dienstverlening
Eindhoven

RAPPORT IWDE 94-03

HET SCHATTEN VAN
MARKTPENETRATIE
EN
MARKTAANDEEL

J.Th.M. Wijnen

maart 1994



**HET SCHATTEN VAN
MARKTPENETRATIE
EN
MARKTAANDEEL**

J.Th.M.Wijnen

maart 1994

Inhoud

- 0. Inleiding
- 1. Penetratie
 - 1.0 Inleiding
 - 1.1 Intervalschatting voor penetratie
 - 1.2 Intervalschatting voor verschil in penetratie bij onafhankelijkheid
 - 1.3 Intervalschatting voor verschil in penetratie bij afhankelijkheid
 - 1.4 Het bepalen van de steekproefgrootte
- 2. Marktaandeel
 - 2.0 Inleiding
 - 2.1 Intervalschatting voor marktaandeel
 - 2.2 Intervalschatting voor verschil in marktaandeel bij onafhankelijkheid
 - 2.3 Intervalschatting voor verschil in marktaandeel bij afhankelijkheid

Isolatiesoort	Aantal 1993	Aantal 1992	Aantal 1991	Pen(%) 1993	Pen(%) 1992	Pen(%) 1991
Steenwol	43	96	60	25.9	24.6	22.1
Polystyreen	91	221	151	54.8	56.7	55.5
Polyurethaan	23	58	53	13.9	14.9	19.5
Overige	9	15	8	5.4	3.8	2.9
Totaal	166	390	272	100.0	100.0	100.0

Tabel 1. Marktpenetratie per isolatiesoort

1.1 Intervalschatting voor penetratie

Onder de veronderstelling dat de projecten aselekt gekozen zijn uit de populatie van projecten in Nederland geldt met een betrouwbaarheid van ongeveer 95% dat de penetratie (fractie projecten in de populatie) voor een bepaalde isolatiesoort ligt in het volgende interval:

$$p \pm 2\sqrt{(Cpq/n)}$$

Hierin is: p=fractie projecten met isolatiesoort in de steekproef
 q=1-p
 n=aantal projecten in de steekproef met n>30
 C=(N-n)/(N-1) met
 N=aantal projecten in de populatie

Het zal duidelijk zijn dat C≈1 als de steekproef klein is t.o.v. de populatie.

Een tweede opmerking is dat de factor 2 in de formule bepaald is door de gekozen betrouwbaarheid. Bij een betrouwbaarheid van 0.90 wordt deze factor 1.645, bij 0.99 wordt deze factor 2.58.

Voorbeeld

Voor isolatiesoort steenwol in 1993 geldt p=0.259 en n=166 (zie tabel 1).

Het interval wordt dan voor verschillende waarden van C:

C=1 :	0.259 ± 0.068	(in %:	25.9 ± 6.8)
C=0.9:	0.259 ± 0.065	(in %:	25.9 ± 6.5)
C=0.5:	0.259 ± 0.048	(in %:	25.9 ± 4.8)

1.2 Intervalschatting voor het verschil in penetratie bij onafhankelijkheid

Het is redelijk te veronderstellen dat schattingen voor de penetratie van een bepaalde isolatiesoort in verschillende jaren onafhankelijk zijn.

De intervalschatting voor het verschil van twee penetraties wordt dan gegeven door:

$$p_1 - p_2 \pm 2\sqrt{(C_1 p_1 q_1 / n_1 + C_2 p_2 q_2 / n_2)}$$

Hierin is voor $i=1$ of $i=2$:

- $q_i = 1 - p_i$
- n_i = aantal projecten in steekproef i met $n_i > 30$
- $C_i = (N_i - n_i) / (N_i - 1)$ met
- N_i = aantal projecten in populatie i

Deze intervallschatting kan gebruikt worden om te *toetsen* of de twee penetraties verschillend zijn: als het interval de waarde 0 *niet* bevat dan concluderen we dat de penetraties verschillend zijn. Dit betekent dat het onwaarschijnlijk is dat het gevonden verschil louter aan het toeval toe te schrijven is.

Voorbeeld

De intervallschatting voor het verschil van de penetraties voor steenwol in de jaren 1993 (p_1) en 1991 (p_2) is:

$$C_1 = C_2 = 1 : \quad 0.259 - 0.221 \pm 2\sqrt{(0.001156 + 0.000633)} = 0.038 \pm 0.085$$

De waarde 0 ligt in dit interval dus de penetraties voor steenwol in de jaren 1993 en 1991 zijn niet significant verschillend. Dit betekent dat het gevonden verschil (0.038 of 3.8%) door toeval veroorzaakt zou kunnen zijn.

$$C_1 = C_2 = 0.5 : \quad 0.259 - 0.221 \pm 2\sqrt{(0.000578 + 0.000316)} = 0.038 \pm 0.060$$

Ook als de steekproeven de halve populaties omvatten is het verschil niet significant.

1.3 Intervallschatting voor het verschil in penetratie bij afhankelijkheid

De penetraties van twee verschillende isolatiesoorten in hetzelfde jaar zijn afhankelijk, immers de som van alle penetraties in een jaar is 100 (in %). Bij intervallschattingen moet dat in rekening worden gebracht. Dat leidt tot het volgende interval:

$$p_1 - p_2 \pm 2\sqrt{(C(p_1 q_1 + p_2 q_2 + 2p_1 p_2) / n)}$$

- Hierin is:
- $q_i = 1 - p_i$
 - n = aantal projecten in de steekproef met $n > 30$
 - $C = (N - n) / (N - 1)$ met
 - N = aantal projecten in de populatie

Voorbeeld

De intervallschatting voor het verschil in penetratie van steenwol (p_1) en polystyreen (p_2) in 1993 wordt voor verschillende waarden van C :

$$C = 1 : \quad 0.259 - 0.548 \pm 2\sqrt{((0.1919 + 0.2477 + 0.2839) / 166)} = -0.289 \pm 0.132$$

$$C=0.5: \quad -0.289 \pm 0.093$$

Ook hier kan getoetst worden: 0 ligt niet in het interval (voor beide waarden van C) dus de penetraties van steenwol en polystyreen zijn verschillend.

1.4 Het bepalen van de steekproefgrootte

Op basis van bovenstaande formules kan de steekproefgrootte worden berekend die nodig is om een intervallschatting te krijgen met een van tevoren gekozen lengte (L).

Voor de lengte van het schattingsinterval voor de penetratie geldt:

$$L=4\sqrt{(pq/n)}$$

Hieruit volgt:

$$n=16pq/L^2$$

Voorbeeld

Om een intervallschatting voor de penetratie van steenwol te krijgen met een lengte van hoogstens 0.10 moet de steekproefgrootte minstens zijn:

$$n=16*0.1919/0.01=307$$

Hier is C=1 genomen. Voor C<1 moet ook de populatiegrootte in rekening worden gebracht. Voor die intervallengte geldt dan:

$$L=4\sqrt{(pq(N-n)/n(N-1))}$$

Voor n geldt dan:

$$n=16pqN/(16pq+L^2(N-1))$$

Voorbeeld

Voor N=2500 en L=0.10 moet de steekproefgrootte zijn:

$$n=16*0.1919*2500/(16*0.1919+0.01*2499)=274$$

Merk op dat $C=(2500-274)/2499=0.89$

Voor het verschil tussen penetraties kunnen eveneens steekproefgrootte(n) worden uitgerekend, waarbij de intervallengte van tevoren is vastgelegd. Betreft het penetraties van dezelfde isolatiesoort in verschillende jaren dan moet daarbij de verhouding van de twee steekproefgrootten worden vastgelegd; meestal neemt men $n_1=n_2=n$.

Er geldt dan:

$$n=16(p_1q_1+p_2q_2)/L^2$$

Voorbeeld

Voor steenwol in de jaren 1993 en 1991 geldt bij L=0.10:

$$n=16*(0.1919+0.1722)/0.01=583$$

Om een interval met lengte 0.10 te krijgen voor het verschil in penetratie van steenwol tussen 1993 en 1991 moeten dus in elk jaar 583 projecten worden verzameld.

Als het gaat om een intervallschatting voor het verschil in penetratie van twee verschillende isolatiesoorten in hetzelfde jaar moet de steekproefgrootte worden berekend uit:

$$n=16(p_1q_1+p_2q_2+2p_1p_2)/L^2$$

Voorbeeld

Voor een intervallschatting met lengte 0.10 voor het verschil in penetratie van steenwol en polystyreen in 1993 is een steekproef nodig ter grootte:

$$n=16*(0.1919+0.2477+0.2839)/0.01=1158$$

Rekening houdend met de grootte van de populatie worden de formules wat ingewikkelder. Gezien bovenstaande uitkomsten en het feit dat de populatiegrootte van de orde 2500 à 3000 zal zijn is het nuttig om ook deze formules te vermelden en te zien wat het effect is op de vereiste steekproefgrootte. Eerst weer het verschil tussen jaren. Voor de steekproefgrootte geldt dan (we nemen $n_1=n_2=n$):

$$n=16(p_1q_1N_1/(N_1-1)+p_2q_2N_2/(N_2-1))/(L^2+p_1q_1/(N_1-1)+p_2q_2/(N_2-1))$$

Voorbeeld

Voor steenwol in 1993 en 1991 geeft dat voor $N_1=N_2=2500$:

$$n=16*(0.1919*2500/2499+0.1722*2500/2499)/(0.01+0.1919/2499+0.1722/2499)=575$$

Hierbij hoort $C=(2500-575)/2499=0.77$. Als de populatie niet groot is (t.o.v. de steekproef) zijn minder waarnemingen nodig om een intervallschatting te krijgen met dezelfde lengte.

Bij afhankelijke penetraties geldt:

$$n=16(p_1q_1+p_2q_2+2p_1p_2)/(L^2+(p_1q_1+p_2q_2+2p_1p_2)/(N-1))$$

Voorbeeld

Voor steenwol en polystyreen in 1993 ($N=2500$) wordt:

$$n=16*(0.1919+0.2477+0.2839)/(0.01+(0.1919+0.2477+0.2839)/2499)=1125$$

Hierbij is $C=(2500-1125)/2499=0.55$. Er zijn weliswaar minder projecten nodig dan bij een grote populatie, maar de steekproef omvat bijna de helft van de populatie.

2. Marktaandeel

2.0 Inleiding

Het marktaandeel (MA) van een bepaalde isolatiesoort is het quotiënt van de bijbehorende (dak)oppervlakte en de totale (dak)oppervlakte. Berekent men dit quotiënt uit de in de steekproef opgenomen projecten dan wordt een schatting verkregen die systematisch afwijkt van het werkelijke marktaandeel (in de populatie). Die systematische afwijking is te verwaarlozen als de verhouding van de standaardafwijking van het gemiddelde en het gemiddelde zelf, zowel voor teller als voor noemer, klein genoeg is (ongeveer 0.1). Omdat standaardafwijking en gemiddelde van de dakoppervlakten vaak van dezelfde orde van grootte zijn moet het aantal projecten met die isolatiesoort minstens ongeveer honderd zijn. Dat is niet altijd het geval, zodat de hierna gegeven intervallschattingen dan minder goed bruikbaar zijn. Ook de standaardafwijking van het marktaandeel wordt dan niet goed geschat, zodat ook de opgegeven betrouwbaarheid niet correct is.

De wijze waarop de intervallschattingen worden berekend berust op de wet van de voortplanting van fouten. Onder de eerder genoemde voorwaarden zijn deze intervallschattingen wel goed bruikbaar. Bij de bespreking van het marktaandeel gebruiken we de gegevens in onderstaande tabel. Deze gegevens komen eveneens van de door Rockwool verstrekte spreadsheetfiles.

	1993		1992		1991	
	steenwol	overige	steenwol	overige	steenwol	overige
aantal	43	123	90	293	59	211
gemiddelde	1842	1370	1819	1523	2651	2141
totaal	79200	168519	163725	446271	156400	451795
stand.afw.	2006	1632	2397	1680	3245	2713
MA (in%)	32.0	68.0	26.8	73.2	25.7	74.3

Tabel 2. Marktaandeel van steenwol

2.1 Intervallschatting voor marktaandeel

Het marktaandeel van een isolatiesoort wordt geschat door

$$MA = x_i / (x_i + y_i)$$

Hierin is: x_i = totale oppervlakte isolatiesoort
 y_i = totale oppervlakte overige isolatiesoorten

Voor de variantie van het geschatte marktaandeel (s^2) geldt onder de eerder genoemde voorwaarden bij benadering:

$$s^2 = (C_x n_x s_x^2 y_i^2 + C_y n_y s_y^2 x_i^2) / (x_i + y_i)^4$$

Hierin is: n_x = aantal projecten met isolatiesoort in de steekproef
 n_y = aantal projecten met overige isolatiesoorten in de steekproef
 s_x = standaardafwijking isolatiesoort
 s_y = standaardafwijking overige isolatiesoorten
 $C_x = 1 - n_x/N_x$ met
 N_x = aantal projecten met isolatiesoort in de populatie
 $C_y = 1 - n_y/N_y$ met
 N_y = aantal projecten met overige isolatiesoorten in de populatie

Een intervalschatting voor het marktaandeel met een betrouwbaarheid van ongeveer 0.95 wordt dan berekend met:

$$MA \pm 2s$$

Voorbeeld

Voor steenwol in 1993 wordt de schatting van het marktaandeel:

$$MA = 79200/(79200 + 168519) = 0.320$$

Met $C_x=C_y=1$ wordt de schatting van de variantie:

$$s^2 = (43*2006^2*168519^2 + 123*1632^2*79200^2)/(79200 + 168519)^4 = 0.001851$$

Het interval wordt dan:

$$0.320 \pm 2\sqrt{0.001851} = 0.320 \pm 0.086$$

Met waarden van C_x en C_y kleiner dan 1 worden uiteraard kleinere intervallen verkregen.

2.2 Intervalschatting voor verschil in marktaandeel bij onafhankelijkheid

Omdat schattingen voor marktaandelen in verschillende jaren onafhankelijk zijn wordt een intervalschatting voor het verschil in marktaandeel van een isolatiesoort in een jaar en een isolatiesoort in een ander jaar:

$$MA_1 - MA_2 \pm 2\sqrt{(s_1^2 + s_2^2)}$$

Voorbeeld

Voor steenwol in 1992 berekenen we ($C_x=C_y=1$):

$$MA = 163725/(163725 + 446271) = 0.268$$

$$s^2 = (90*2397^2*446271^2 + 293*1680^2*163725^2)/(163725 + 446271)^4 = 0.000904$$

Het interval voor het verschil in marktaandeel in 1993 en 1992 wordt dan:

$$0.320 - 0.268 \pm 2\sqrt{(0.001851 + 0.000904)} = 0.052 \pm 0.105$$

Hieruit kan de conclusie worden getrokken dat er geen significant verschil is tussen de marktaandelen van steenwol in de jaren 1993 en 1992 (de waarde 0 ligt in het interval).

2.3 Intervalschatting voor verschil in marktaandeel bij afhankelijkheid

Net zoals bij penetratie het geval is, zijn de schattingen voor marktaandelen van verschillende isolatiesoorten in hetzelfde jaar niet onafhankelijk: de noemers van beide quotiënten zijn hetzelfde. Voor de variantie van het verschil in marktaandeel (s_v^2) geldt onder eerder genoemde voorwaarden het volgende:

$$s_v^2 = (C_x n_x (2y_i + z_i)^2 s_x^2 + C_y n_y (2x_i + z_i)^2 s_y^2 + C_z n_z (x_i - y_i)^2 s_z^2) / (x_i + y_i + z_i)^4$$

Hierin is: x_i = totale oppervlakte van de ene isolatiesoort in de steekproef
 y_i = totale oppervlakte van de andere isolatiesoort in de steekproef
 z_i = totale oppervlakte van de overige isolatiesoorten in de steekproef

De overige symbolen zijn al eerder ingevoerd.

Het interval wordt dan:

$$MA_x - MA_y \pm 2s_v$$

Voorbeeld

Om het interval te berekenen voor het verschil in marktaandeel van steenwol en polystyreen in 1993 maken we gebruik van de gegevens in de volgende tabel.

	steenwol	polystyreen	overige	totaal 1993
aantal	43	91	32	166
gemiddelde	1841	1316	1523	1492
totaal	79200	119785	48734	247719
stand.afw.	2006	1405	2172	1742
MA(in%)	32.0	48.3	19.7	-

Tabel 3. Marktaandelen steenwol en polystyreen in 1993

Het verschil in marktaandeel heeft voor $C_x=C_y=C_z=1$ een geschatte variantie:

$$s_v^2 =$$

$$(43*(2*119785+48734)^2*2006^2+91*(2*79200+48734)^2*1405^2+32*(79200-119785)^2*2172^2)/247719^4$$

$$= 0.005932$$

Het interval wordt:

$$0.320 - 0.483 \pm 2\sqrt{0.0059302} = -0.163 \pm 0.154$$

De marktaandelen steenwol en polystyreen verschillen significant.
