

„INFORMATION ECONOMICS”: INLEIDING EN KRITIEK

door Dr. J. P. C. Kleijnen

Samenvatting

Information Economics is gebaseerd op statistische beslissingstheorie, en beoogt de waarde van informatie te berekenen. Die berekening wordt geïllustreerd via het eenvoudige voorbeeld van een detailhandelaar die moet beslissen hoeveel eenheden hij zal inkopen. Wij berekenen het effect op zijn winst van het installeren van een perfect respectievelijk een onnauwkeurig informatiesysteem. Dit voorbeeld illustreert de diverse beginselen van information economics. Wij bespreken een aantal beperkingen die optreden indien men deze benadering wil toepassen op Management Informatiesystemen (MIS).

1 Inleiding

Information economics is een wetenschappelijke benadering die het economisch nut van informatie probeert te berekenen op basis van de *statistische beslissingstheorie* à la Raiffa en Schlaifer [8]. In dit artikel zullen wij eerst de benadering illustreren aan de hand van een eenvoudig voorbeeld. Vervolgens behandelen wij de algemene concepten die deze benadering oplevert. Tenslotte evalueren wij de toepasbaarheid van deze benadering bij de economische evaluatie van informatie-systemen.

De volgende stochastische wetten vormen de basis van de berekeningen in information economics. Stel er zijn twee gebeurtenissen A en B, dan voldoet de voorwaardelijke kans aan

$$P(A|B) = \frac{P(AB)}{P(B)} \quad \text{mits } P(B) > 0 \quad (1)$$

waarbij $A|B$ betekent „A mits B optreedt”, en AB betekent „A en B treden beide op”. Tevens geldt

$$P(B) = P(B|A_1) \cdot P(A_1) + P(B|A_2) \cdot P(A_2) + \dots \\ \dots + P(B|A_n) \cdot P(A_n) \quad (2)$$

waarbij de gebeurtenis A is opgesplitst in zijn „deelgebeurtenissen” A_i , bijv. indien A betekent „kind” dan kan A opgesplitst worden in „jongen” (A_1) of „meisje” (A_2). Uit vergelijking (1) en (2) volgt *Bayes* theorema:

$$P(A_i|B) = \frac{P(A_i) \cdot P(B|A_i)}{P(B|A_1) \cdot P(A_1) + \dots + P(B|A_n) \cdot P(A_n)} \quad (i = 1, \dots, n) \quad (3)$$

waarbij $P(A_i)$ de „a priori” kans is, en $P(A_i|B)$ de a posteriori kans, nadat namelijk bekend is dat B is opgetreden.

2 Voorbeeld

Het volgend voorbeeld is, op enkele details na, ontleend aan Verhelst [9]. Een detailhandelaar moet elke dag beslissen hoeveel produkteenheden (zeg a stuks) hij wenst in te kopen, $a \leq 6$,. Hij verkoopt e stuks $a \leq 10$,.

Hij kan geen voorraad aanhouden. Derhalve geldt

$$\begin{aligned} \text{Winst} &= 10e - 6a && \text{mits } e \leq a \\ &= 10a - 6a = 4a && \text{mits } e \leq a \end{aligned} \tag{4}$$

Dit is weergegeven in de opbrengstenmatrix van Tabel 1. In deze tabel staan ook de a priori kansen, $P(e)$, die de handelaar toekent aan de verschillende waarden van e .

Tabel 1: Opbrengstenmatrix.

Afzet (e)	Inkoop (a)			$P(e)$
	0	1	2	
0	0	-6	-12	0.10
1	0	+4	- 2	0.50
2	0	+4	+ 8	0.40

Uit Tabel 1 berekenen wij de verwachte winst voor elke „inkoopstrategie”. Bijvoorbeeld, indien de handelaar altijd $a = 2$ zou inkopen, dan geldt

$$\begin{aligned} \&(\text{Winst} \mid a = 2) &= -12 . P (e = 0) - 2 . P (e = 1) + 8 . P (e = 2) \\ &= -12 (0.10) - 2 (0.50) + 8 (0.40) \\ &= -1.2 - 1 + 3.2 = 1 \end{aligned} \tag{5}$$

waar $\&$ betekent „verwachte waarde” (gemiddelde). Zo kunnen wij ook berekenen dat $\& (\text{winst} \mid a = 0) = 0$ en $\& (\text{winst} \mid a = 1) = 3$. Derhalve is het optimaal om altijd één eenheid in te kopen.

Vervolgens beschouwen wij het nut van het volgende informatiesysteem. De handelaar belt elke dag zijn klanten op teneinde te informeren hoeveel stuks zij willen kopen. Eerst zullen wij aannemen dat die informatie perfect is, vervolgens dat die informatie onnauwkeurigheden vertoont.

Een perfect informatiesysteem betekent dat de boodschappen (y) van dat systeem precies voorspellen welke waarde voor e zal gelden:

$$\begin{aligned} P (e = i \mid y = k) &= 1 && \text{indien } i = k \\ &= 0 && \text{indien } i \neq k \\ &&& (i, k = 0, 1, 2) \end{aligned} \tag{6}$$

Derhalve kiest de handelaar die waarde voor de inkopen die de winst maximaliseert, d.w.z. in de relevante rij van Tabel 1 kiest hij, zeg a^* , welke het rijelement maximaliseert; zie Tabel 2. In die tabel hebben wij nog eens herhaald welke

Tabel 2. Perfecte informatie

$y = e$	a^*	Winst	$P(e)$
0	0	0	0.10
1	1	4	0.50
2	2	8	0.40

waarden van $e (= y)$ wij verwachten volgens Tabel 1. Derhalve is de verwachte winst bij perfecte informatie

$$\begin{aligned} \mathbb{E}(\text{winst} | \text{Perfecte info}) &= 0 (0.10) + 4 (0.50) + \\ &\quad + 8 (0.40) = 5.2 \end{aligned} \quad (7)$$

Zonder een informatiesysteem was de verwachte winst 3 (aannemende dat wij een optimale strategie kozen, $a^* = 1$). Derhalve is de bruto-opbrengst van een perfect informatiesysteem, $5.2 - 3 = 2.2$.

Vervolgens gaan wij uit van de meer realistische aanname dat het informatiesysteem niet perfect is, zodat vergelijking (6) niet meer geldt. Gewoonlijk moeten wij $P(e|y)$ berekenen uit $P(y|e)$, via Bayes theorema in vergelijking (3). Een voorbeeld is in de appendix uitgewerkt. Stel dat Tabel 3 resulteert. Merk op dat de a priori kans was $P(e = 0) = 0.10$ terwijl nu de a posteriori kans (nadat de boodschap $y = 0$ is ontvangen) is toegenomen tot $P(e = 0 | y = 0) = 0.44$.

Tabel 3. A posteriori kansen $P(e|y)$

y	e		
	0	1	2
0	0.44	0.31	0.25
1	0.04	0.78	0.18
2	0.03	0.26	0.71

Indien de boodschap is $y = 0$, en de handelaar bestelt a , dan volgt uit Tabel 1 en 3:

$$\text{Winst } (a = 0 | y = 0) = 0$$

$$\text{Winst } (a = 1 | y = 0) = -6(0.44) + 4(0.31) + 4(0.25) = -0.38$$

$$\text{Winst } (a = 2 | y = 0) = -12(0.44) - 2(0.31) + 8(0.25) = -3.88 \quad (8)$$

Indien dus de boodschap is $y = 0$, dan is het optimaal om niets in te kopen, $a^* = 0$. Zo kunnen wij ook berekenen dat voor $y = 1$ geldt $a^* = 1$, met een winst van 3.56 en voor $y = 2$ geldt $a^* = 2$ met een winst van 4.92. De kans op een bepaalde y kunnen we ook uitrekenen (zie appendix) en bedraagt $P(y = 0) = 0.16$, $P(y = 1) = 0.45$ en $P(y = 2) = 0.39$. Derhalve geldt:

$$\begin{aligned} \mathbb{E}(\text{Winst} | \text{onnauwkeurige info}) &= 0 (0.16) + (3.56) (0.45) + \\ &\quad + (4.92) (0.39) = 3.52 \end{aligned} \quad (9)$$

De verwachte winst is slechts 3.52 en niet 5.2 zoals bij perfecte informatie; zie vergelijking (7). De bruto-opbrengst van het onnauwkeurige informatiesysteem is $3.52 - 3 = 0.52$, waarbij 3 de verwachte winst is zonder een informatiesysteem.

3 Information economics: algemene beginselen

Zelfs het eenvoudige voorbeeld van de vorige paragraaf toont dat information economics een zinvol raamwerk oplevert. De volgende beginselen blijken namelijk belangrijk te zijn.

- a Het *verrassingsgehalte* van informatie: Informatie die reeds verwacht wordt, genereert niet veel waarde. Statistisch wordt dit gemeten door het verschil tussen de a priori kansen en de a posteriori kansen. (Bij een a priori kans van 1 is additionele informatie waardeloos.) Eenvoudige voorbeelden kan men vinden in [2].
- b Het effect van de informatie op de *beslissing*: Indien wij dezelfde beslissing zouden nemen, onafhankelijk van de informatie, dan heeft additionele informatie geen nut. Bijvoorbeeld, stel dat de nauwkeurigheid van de informatie gespecificeerd door de kansen $P(y|e)$ in de appendix, worden vervangen door die van Tabel 4. Dan is de optimale strategie altijd $a^* = 1$, onafhankelijk van de boodschap y . Zie [2] voor een ander voorbeeld van het effect van de nauwkeurigheid van informatie.

Tabel 4: $P(y|e)$

y	$e = 0 \quad 1 \quad 2$		
0	0.4	0.2	0.2
1	0.4	0.4	0.4
2	0.2	0.4	0.4

- c Het effect van de beslissing op de *opbrengst*: De opbrengst kan ongevoelig zijn voor de beslissing. (In tabel 1 zouden de kolommen identiek kunnen zijn.) Een bekend voorbeeld vinden wij in het voorraadbeheer, waar de totale voorraadkosten weinig gevoelig zijn voor de juiste waarde van de kostenparameters die optreden in de wortel-formule voor de optimale bestelhoeveelheid.
- d De berekeningen in paragraaf 3 toonden duidelijk aan dat de beslissingen moeten worden aangepast aan de nauwkeurigheid van de *informatie*.

4 Beperkingen

Wij zien de volgende beperkingen bij het toepassen van information economics voor de evaluatie van (gecomputeriseerde) informatiesystemen t.b.v. managementbeslissingen (Management Information Systems MIS).

- 1 Een grote hoeveelheid *gegevens* is nodig als invoer voor het rekenproces. Tabel 1 toont dat wij dienen te beschikken over alle mogelijke beslissingsalternatieven (a), alle mogelijke toestanden (e), en de uitkomsten voor alle mogelijke combinaties van beslissingen en toestanden. Het is dus niet mogelijk een beslissingsstrategie te ontwikkelen naargelang de gebeurtenissen om actie vragen. Tabel 1 toont verder dat wij de a priori kansen, $P(e)$, moeten specificeren. Deze kansen mogen echter subjectieve schattingen zijn. Tabel 3 laat zien dat wij ook de kwaliteit van het informatiesysteem moeten kwantificeren. Het is echter alleszins redelijk dat, indien wij de waarde van het informatiesysteem willen kwantificeren, wij de kwaliteit van het informatiesysteem moeten kwantificeren.

- 2 De beslisser wordt geacht te streven naar een *optimum*. In ons voorbeeld streeft de handelaar naar een maximale winst. Information economics is ruimer in die zin dat het criterium niet in geld hoeft te worden uitgedrukt, maar ook in nutseenheden („utility”) kan worden geformuleerd. Dit leidt echter tot praktische problemen bij het specificeren van zo'n nutsfunctie. Realistischer is het om te veronderstellen dat de beslisser streeft naar het voldoen aan meerdere criteria, zgn. „satisficing” van de behavioral organisatieleer à la Cyert en March [1].
- 3 De *berekeningen* kunnen ingewikkeld worden en zelfs ernstige mathematische problemen oproepen voor bepaalde kansfuncties. Het is interessant dat de computer gebruikt kan worden om de subjectieve a priori kansen op te vragen, en te manipuleren via Bayes' theorema [7].
- 4 De *organisatie* die bestuurd moet worden, wordt verondersteld zeer simpel te zijn. In werkelijkheid vertonen organisaties meerdere, samenhangende processen (bijv. fabricageprocessen, financiële processen). Bovendien bestaat een organisatie uit meerdere beslissers, die met elkaar communiceren. Deze complicatie is onderzocht door Marschak en Radner [6] in hun „theory of teams”. Een andere dimensie van dit probleem is de dynamiek van systemen. Beslissingen beïnvloeden toekomstige gebeurtenissen. Bijvoorbeeld, indien voorraden zijn toegestaan, dan wordt het resultaat in een periode afhankelijk van beslissingen in vorige perioden. Dynamische aspecten zijn onderzocht door Feltham [3], maar net als Marschak en Radner's werk blijft zijn studie erg abstract (en gecompliceerd).
- 5 De analyse beschouwt slechts één aspect van informatie, namelijk de *nauwkeurigheid*. Elders [5] hebben wij laten zien dat in werkelijkheid een reeks van aspecten is te onderscheiden, zoals ouderdom van informatie, aggregatie, rapportering, enz.

Bovengenoemde beperkingen, met name (4) en (5), zijn niet zo ernstig in eenmalige beslissingen, zoals de beslissing om al of niet een nieuw produkt te lanceren. Eenmalige beslissingen vinden wij vooral op strategisch en tactisch gebied. (Er bestaan toepassingen van statistische beslissingstheorie op operationeel gebied, bijv. in de kwaliteitscontrole, maar dit betreft zeer beperkte „informatiesystemen”.) Informatiesystemen t.b.v. strategische en tactische beslissingen zijn typisch niet geautomatiseerd, zodat wij bij onze konklusie blijven dat de toepassing van information economics op MIS nog vele beperkingen kent.

5 Konklusies

Onze kritische beschouwing van information economics in de vorige paragraaf, betekent niet dat wij deze benadering als waardeloos terzijde willen schuiven. In paragraaf 3 hebben wij reeds benadrukt dat deze School ons althans een raamwerk verstrekt waarbinnen wij onze gedachten over economische evaluatie van MIS verder kunnen ontwikkelen. Tot nu toe is deze methode hoofdzakelijk ontwikkeld door statistici en wiskundige economen. Het lijkt gewenst dat thans informatici gaan onderzoeken of de methode kan worden uitgewerkt t.b.v. de evaluatie van MIS zodanig dat de bezwaren van paragraaf 4 vervallen. Dit hoeft niet perse te leiden tot een techniek die gedetailleerde antwoorden geeft. Wij verwachten eerder dat deze benadering een theoretische achtergrond oplevert waar-

binnen dat exacte berekeningen kunnen plaatsvinden via andere technieken, met name simulatie. Simulatie is immers een techniek die niet aan de beperkingen van de vorige paragraaf is onderworpen; zie de voorbeelden in [4] en [5]. Simulatie is echter niet meer dan een techniek, en heeft theorie nodig die vertelt *wat* we moeten onderzoeken!

Appendix: Van $P(y|e)$ naar $P(e|y)$.

Aan de hand van een voorbeeld zullen wij laten zien, hoe wij $P(e|y)$ kunnen berekenen vanuit $P(y|e)$, gebruikmakend van Bayes' theorema in vergelijking (3). Laten wij aannemen dat de handelaar informatie verkrijgt door elke dag 50 klanten te bellen, terwijl zijn hele klantenbestand 1000 klanten omvat. Stel dat elke klant met een kans p één produkteenheid koopt.

(Die kans p is onbekend en verandert van dag tot dag.) Indien x aangeeft hoeveel een klant koopt, dan geldt

$$\begin{aligned} P(x = 1) &= p \\ P(x = 0) &= 1-p \end{aligned} \tag{A.1}$$

Neem aan dat op een bepaalde dag 20 van de 50 klanten toezeggen te kopen, d.w.z. de boodschap y wordt

$$y = \sum_{j=1}^{50} x_j = 20 \tag{A.2}$$

Stel dat in het verleden m waarden van p werden waargenomen, p_i ($i = 1, \dots, m$), met bekende a priori kansen $P(p_i)$. De binomiale verdeling toont ons dan

$$P(y = 20 \mid p = p_i) = \binom{50}{20} (p_i)^{20} (1-p_i)^{30} \tag{A.3}$$

Vergelijking (A.3) kunnen wij uitrekenen voor elk van de m waarden p_i . Dit levert $P(y = y_k \mid p = p_i)$. Vervolgens passen wij vergelijking (3) toe met $A_i = p_i$ en $B = y_k$, hetgeen ons $P(p_i \mid y_k)$ oplevert. In het voorbeeld van paragraaf 2 werken wij niet met de kans p_i maar met de absolute aantallen e_i .

Stel dat Tabel 5 geldt. Uit vergelijking (2) volgt dan:

$$P(y = 0) = 0.7 (0.1) + 0.1(0.5) + 0.1(0.4) = 0.16$$

Tabel 5: $P(y|e)$

y	e		
	0	1	2
0	0.7	0.1	0.1
1	0.2	0.7	0.2
2	0.1	0.2	0.7
$P(e)$	0.10	0.50	0.40

Op dezelfde manier berekenen wij $P(y = 1) = 0.45$ en $P(y = 2) = 0.39$. Vergelijking (3) levert dan, bijv.

$$P(e = 0 \mid y = 0) = \frac{(0.1)(0.7)}{0.16} = 0.44$$

Aldus is tabel 3 berekend.

Referenties

1. CYERT, R. M. and J. G. MARCH,
A behavioral theory of the firm. Prentice Hall, Inc., Englewood Cliffs, New Jersey, 1963.
2. EMERY, J. C.,
Cost-benefit analysis of information systems. SMIS Workshop Report No. 1, The Society for Management Information Systems, 1971 (second printing March 1973).
3. FELTHAM, G. A.,
The value of information. The Accounting Review, 43, no. 4, Oct. 1968, pp. 684-696.
4. KLEIJNEN, J. P. C.,
Het economisch nut van nauwkeurige informatie: simulatie van ondernemingsbepalingen en informatie. Maandblad voor Accountancy en Bedrijfshuishoudkunde, 50, no. 8, Sept. 1966, pp. 453-470.
5. KLEIJNEN, J. P. C.,
Computers and profit: Quantification of Financial Benefits of Information (provisional title). To be published 1977.
6. MARSCHAK, J. and R. RADNER,
Economic theory of teams. Yale University Press, New Haven, 1972.
7. NOVICK, M. R. and P. H. JACKSON,
Statistical methods for educational and psychological research. McGraw-Hill Book Company, New York, 1974.
8. RAIFFA, H.,
Decision analysis: introductory lectures on choices under uncertainty. Addison-Wesley Publishing Company, Reading, 1968.
9. VERHELST, M.,
Contribution to the analysis of organizational information systems and their benefits. Faculteit der Economische Wetenschappen, Universiteit van Leuven, Leuven, 1974.