

# Range Estimating

Dr. M. S. A. Vrijland

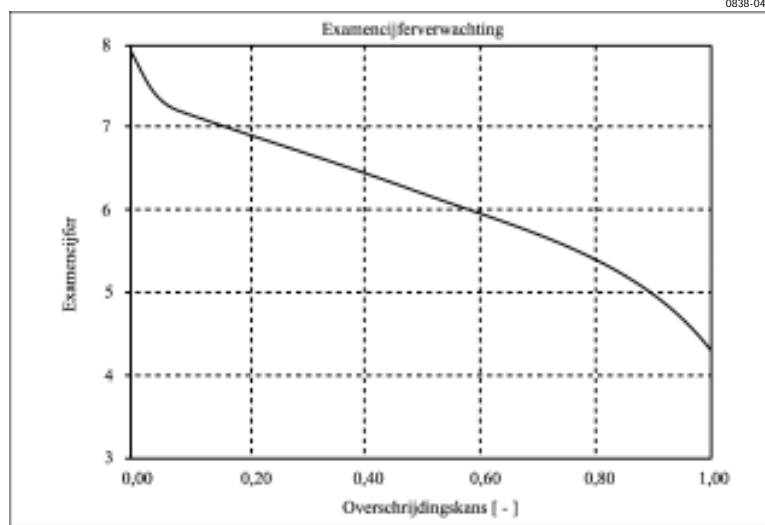
1.	Proloog	G3600- 3
2.	Range Estimating	G3600- 4
3.	Aselecte getallen	G3600- 5
4.	De Monte Carlo-methode	G3600- 7
4.1.	Numeriek integreren	G3600- 8
4.2.	Simuleren van data	G3600-10
4.2.1.	Normaal verdeeld toevalsgetal NRN	G3600-11
4.2.2.	Toevalsgetal met driehoeksverdeling TRN	G3600-12
4.2.3.	Andersverdeelde toevalsgetallen HRN	G3600-13
5.	De wet van Pareto	G3600-14
6.	Een voorbeeld	G3600-15
7.	Discussie	G3600-18
8.	Epiloog	G3600-19
9.	Literatuur	G3600-20



## 1. Proloog

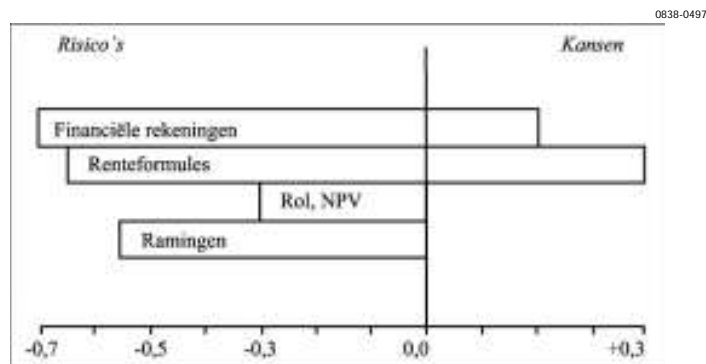
Een derdejaars studente in de gezelligste studentenstad van Nederland heeft de colleges *Proceseconomie* gevolgd, en wil aan het examen in dat vak meedoen.

Eén van de behandelde onderwerpen is *Range Estimating*. Zij besluit ter voorbereiding op het examen deze techniek nu eens toe te passen om haar slaagkans te ramen en om haar verdere inspanningen te kunnen optimaliseren. Het resultaat van haar Range Estimating activiteiten is vastgelegd in de figuren 1 en 2.



Figuur 1. 57% kans meer dan 6, 16% meer dan 7!

Figuur 1 laat zien dat zij ruim 40% kans heeft minder dan een 6 te scoren, en geen enkele kans heeft op een 8. Dat is haar niet genoeg. Zij wil haar kansen verbeteren door nog 30 uur studie in het vak te steken. Figuur 2 laat zien waar haar kansen en haar risico's liggen. Als de eerste vraag tegen zit kan haar cijfer met 0,7 punt dalen, als hij meezit 0,2 punt hoger uitvallen. Ze besluit op grond van de verkregen informatie alsnog 20 uur te besteden aan het onderwerp „Financiële Jaarrekeningen” en nog 10 uur aan „Renteformules”.



Figuur 2. Kans/risico analyse examencijfer.

## 2. Range Estimating

Voor de toepassing van de Range Estimating techniek zijn invoergegevens nodig. In dit geval doorliep de studente een procedure die volkomen analoog is met de wijze waarop de techniek gebruikt kan worden als hulpmiddel bij investeringsbeslissingen in de procesindustrie of elders.

Op grond van bestudering van vroegere examens verwachtte zij voor elk van de vier hoofdonderwerpen een vraagstuk. Per examen kunnen de vraagstukken in relatieve zwaarte verschillen. Gebruik makend van wat zij geleerd heeft stelde zij een tabel op met daarin per onderwerp het cijfer (op een schaal van 1 tot 10 waarop zijzelf haar eigen kennis schatte (MO). Daarbij legde zij ook de vermoedelijke weegfactor (punten uit 100) vast. Ook gaf zij zichzelf voor elk van de onderwerpen een cijferverwachting voor als het meezit (HI) en voor als het tegenzit (LO). De tabel zag er zo uit:

Onderwerp van het vraagstuk	MO	Punten	HI	LO
Financiële jaarrekeningen van ondernemingen	4	25 ± 5	6	1
Tijdwaarde van geld, renteformules	7	30 ± 5	9	3
Winstverwachtingscriteria (RoI, NPV, etc.)	8	20 ± 5	9	6
Investerings- en kostenramingen	7	Rest	8	4

Vervolgens ging zij een uurtje aan de gang met Lotus 123® en met een tekenprogramma. Dit leverde haar de figuren 1 en 2 op.

In de hierboven beschreven situatie heeft de studente Range Estimating toegepast om een onzekerheid te kwantificeren, om haar risico's en kansen vast te stellen.

Op grond hiervan kon zij een prioriteitsvolgorde voor aandachtpunten bepalen.

Hieronder zal worden beschreven op welke principes Range Estimating berust, en wat de sterke en zwakke punten van de techniek zijn.

### 3. Aselecte getallen

Het eenvoudigste hulpmiddel voor een verloting onder zes deelnemers is een klassieke dobbelsteen. Elk van de getallen 1 tot en met 6 heeft een precies even grote kans als resultaat van de worp naar boven te komen.

Voor een verloting onder 36 deelnemers zouden we een roulette kunnen gebruiken. Dat is zelfs duidelijk beter dan een enkele worp met zes dobbelstenen!

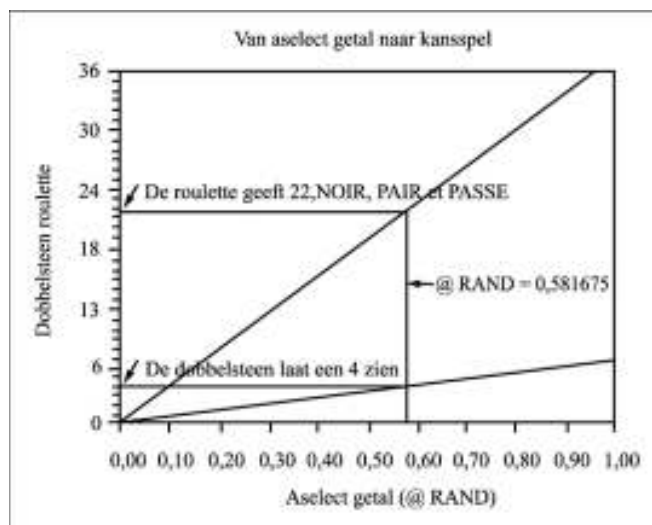
De uitkomst van een worp met een dobbelsteen of van een draai van de roulette kan gesimuleerd worden met behulp van aselecte getallen (ook wel toevalsgetallen of *random numbers* genoemd).

We beginnen met een aselect getal  $A$  (@RAND), met een willekeurige waarde tussen 0 en 1, en vermenigvuldigen dat met 6. Als het produkt  $6 \times A$  tussen 3 en 4 ligt komt dat overeen met een dobbelsteenworp van 4.

De roulette kan evenzo gesimuleerd worden door een aselect getal  $A$  te vermenigvuldigen met 37 (!).

Bij  $A \geq 0,972972972$ . . . is de inzet voor de bank. ( $37 \Rightarrow 0$ )

In Figuur 3 is te zien hoe een aselect getal  $A$ , bijvoorbeeld 0,581675, vertaald wordt naar een dobbelsteenworp 4 of een roulette-draai 22. Het aselecte getal is uitgezet op de X-as, het resultaat van de vertaling op de Y-as. De „vertaallijnen” in deze twee eenvoudige voorbeelden zijn rechten, met helling 6 (dobbelsteen) respectievelijk 37 (roulette). Elke simulatie waarbij kans een rol speelt gaat uit van aselecte getallen. Per gesimuleerd object moet worden nagegaan op welke wijze een aselect getal tot een uitkomst van de simulatie omgerekend moet worden. De wijze van omrekenen is grafisch vast te leggen met een „vertaallijn” in een XY-diagram.



Figuur 3. Van aseleectgetal naar uitkomst kansspel.

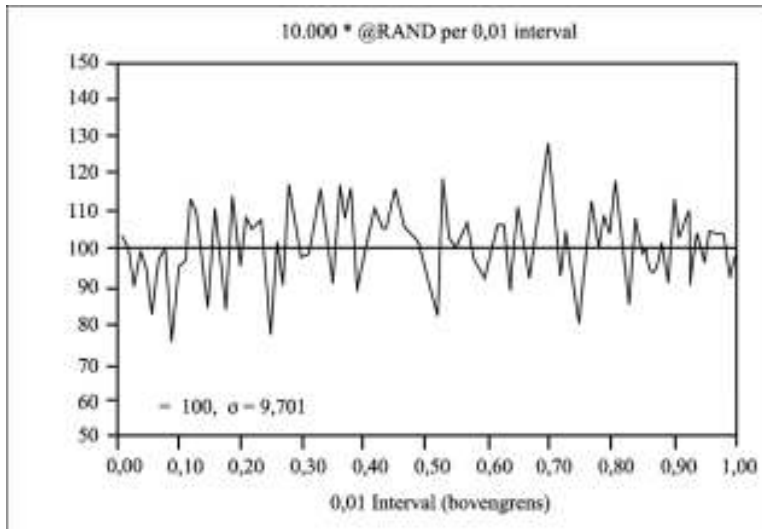
Niemand kan zelf een reeks aseleecte getallen bedenken. Daarom zijn in elk boek over kansrekening en statistiek tabellen met aseleecte getallen te vinden.

Een heel klein deel van een tabel met aseleecte getallen staat in figuur 4. In de praktijk wordt echter nauwelijks meer gebruik gemaakt van dergelijke verzamelingen: aseleecte getallen zijn dankzij een (pseudo) random number generator op de computer (mainframe of PC) beschikbaar.

15	88	09	22	61
71	92	60	08	19
64	42	52	81	08
79	78	22	39	24
35	33	77	45	38

Figuur 4. Aseleecte getallen.

Figuur 5 laat zien hoe 10.000 aseleecte getallen tussen 0 en 1, gegenereerd met de @RAND-functie in de spreadsheet Lotus 123®, in een bepaalde cyclus verdeeld waren over 100 intervals tussen 0 en 1.



Figuur 5. Frequentieverdeling van 10.000 aselechte getallen in Lotus 123®.

Het gemiddelde over alle intervallen is exact 100, maar hoe kleiner het interval gekozen wordt, hoe groter de afwijkingen van het gemiddelde zullen zijn.

Het is duidelijk dat bovenstaand plaatje op zich ook weer een toevalstreffer is: worden er opnieuw 10.000 aselechte getallen gegenereerd dan zullen vergelijkbare pieken en dalen voorkomen, maar het aantal per interval zal afwijken.

De conclusie van deze paragraaf is dat kansen gesimuleerd kunnen worden met behulp van aselechte getallen, en dat het niet moeilijk is over aselechte getallen te beschikken.

#### 4. De Monte Carlo-methode

De Monte Carlo-methode is een op het gebruik van aselechte getallen berustende techniek, die toepassing vindt bij het numeriek integreren van complexe functies en voor het simuleren van data. Dankzij de steeds ruimere beschikbaarheid van steeds krachtigere computers biedt de methode vaak een aantrekkelijk alternatief. Soms zelfs de enige praktisch uitvoerbare oplossingsroute.

**G3600-8** Range Estimating

*4.1. Numeriek integreren*

Een aardig voorbeeld laat zien hoe het oppervlak van een cirkel bepaald kan worden door numerieke integratie met de Monte Carlo-methode.

Als we de oorsprong van het assenkruis in het middelpunt van de cirkel leggen voldoen alle punten op de omtrek van een cirkel met een straal  $r = 1$  meter aan de vergelijking  $x^2 + y^2 = 1$ .

Expliciteren van  $y$  geeft:

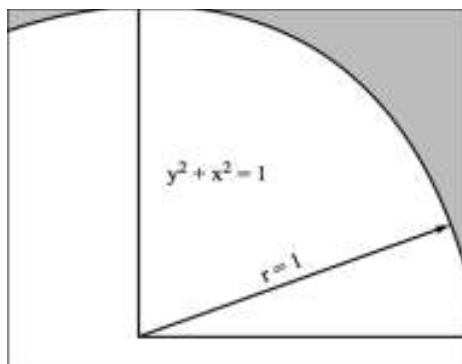
$$y = \pm \sqrt{1-x^2} \quad (1)$$

Om nu het oppervlak van de cirkel te bepalen kunnen we deze functie in het eerste kwadrant integreren van  $x = 0$  tot  $x = 1$ , en dan met 4 vermenigvuldigen. Het totale oppervlak van de cirkel is gelijk aan:

$$A_{\text{cirkel}} = 4 \times \int_{x=0}^{x=1} \sqrt{1-x^2} dx \quad (2)$$

Niet iedereen zal deze integraal zonder hulp kunnen oplossen (het gaat in feite om het klassieke probleem van de kwadratuur van de cirkel!). De Monte Carlo-methode biedt echter uitkomst:

0838-0500



*Figuur 6. Oppervlak van een cirkel.*

- Plaats de cirkel – met het middelpunt in de oorsprong – in een door vier raaklijnen – parallel met de assen – gevormd vierkant, en beschouw het eerste kwadrant. Figuur 6.



- Raadpleeg een bron van lineair tussen 0 en 1 verdeelde toevalsgetallen.  
Dit kan een tabel zijn, maar eenvoudiger is het de *random number generator* van bijvoorbeeld Lotus 123® (@RAND) te gebruiken.
- Genereer een punt (x y) met een randomgetal als x (x = @RAND) en een volgend randomgetal als y (y = @RAND).
- Bereken  $y' = \sqrt{1 - x^2}$ .
- Als  $y' \leq y$  ligt het toevalspunt binnen of op de cirkel, anders daarbuiten.
- Genereer een nieuw punt (x y) en herhaal de twee laatste stappen N-maal; houd daarbij bij hoe vaak het punt binnen of op de cirkel ligt, bijvoorbeeld P-maal.
- De verhouding P/N is gelijk aan de verhouding tussen het oppervlak van de cirkel en het omgeschreven vierkant. Voor  $r = 1$  m is het vierkant in het eerste kwadrant  $1 \text{ m}^2$  groot, en omvat het een kwart van het oppervlak van de cirkel. Het totale oppervlak van een cirkel met een straal van 1 meter, bepaald door numerieke integratie met de Monte Carlo-methode, is dus  $4 \times P/N \text{ m}^2$ .

Met een *spreadsheet* werd bovenstaande procedure driemaal doorlopen.

Met  $N = 10.000$  werd voor P/N respectievelijk gevonden: 0,7882 ; 0,7899 ; 0,7851. Vermenigvuldigen met 4 levert voor het oppervlak van een cirkel met een straal van 1 meter respectievelijk:  $A = 3,1528 \text{ m}^2$ ,  $A = 3,1596 \text{ m}^2$  en  $A = 3,1404 \text{ m}^2$ .

Deze waarden liggen ruim binnen 0,5% van de exacte waarde:  $A = \pi r^2 = 3,1416 \text{ m}^2$ .

Met een kleine aanpassing van het programma kon evensnel het volume van een bol met  $r = 1$  m bepaald worden:  $V = 4,147 \text{ m}^3$ ,  $V = 4,218 \text{ m}^3$  en  $V = 4,139 \text{ m}^3$ . Dit komt ook goed overeen met de exacte waarde:  $V = 4/3\pi r^3 = 4,189 \text{ m}^3$ .

Het aardige van de Monte Carlo-methode is dat multidimensioneel integreren met hetzelfde gemak uitvoerbaar is: even simpel als het oppervlak van een cirkel en het volume van een bol gevonden werden, kan worden vastgesteld dat een „vier-dimensionale bol” 31,67 respectievelijk 31,31 of 30,22 ( $10\pi$ )% inneemt van de omhullende vierdimensionale ruimte.

#### 4.2. Simuleren van data

De „uitkomst” van een investeringsraming of een kostprijsschatting is een getal dat is gevonden door een groot aantal invoergegevens bij elkaar op te tellen, met elkaar te vermenigvuldigen en/of te delen. Elk invoergegeven heeft zijn eigen onzekerheid. Het mathematisch bepalen van de onzekerheid in het eindresultaat die voortkomt uit de onzekerheden in de invoergegevens is vaak bewerkelijk.

Ook hier biedt de Monte-Carlomethode een aantrekkelijke oplossing.

Voorwaarde voor toepassing is dat van elk invoergegeven een gemiddelde of meest waarschijnlijke waarde bekend is en de kans op een afwijking van dat gemiddelde. Bovendien moeten alle variabelen onafhankelijk zijn.

Ten behoeve van een investeringsraming worden de kosten van een component X vaak opgegeven op één van twee wijzen:

- Het beschouwde apparaat kost  $f Y \pm 20\%$ .
- De meest waarschijnlijke prijs is  $f Y$  (= MOdaal), maar tenminste  $f (Y - a)$  (= LOw), en „zeker” niet meer dan  $f (Y + b)$  (= HIgh).

In het eerste geval zouden we kunnen aannemen dat de kosten van het apparaat *normaal verdeeld* zijn met Y als gemiddelde en  $0,10 \times Y$  als standaarddeviatie.

Bij de normale verdeling ligt 95% van de waarnemingen binnen twee standaarddeviaties van het gemiddelde. 20% komt dus (als Y met 95% zekerheid is opgegeven) overeen met twee standaarddeviaties. In het tweede geval is er sprake van een *driehoeksverdeling*: de praktijk leert dat de mens van nature optimist is, en dat afwijkingen naar boven (de kosten vallen tegen) vaker voorkomen dan afwijkingen naar beneden ( $b > a$ ).

De essentie van de Monte Carlo-datasimulatie is nu dat de uitkomst van het model (in het voorbeeld het eindbedrag van de raming) een groot aantal malen berekend wordt, telkens met door het toeval bepaalde invoerwaarden voor elke afzonderlijke component. Voor elke component zal daarbij de kansdichtheid rond de meest waarschijnlijke waarde het hoogst moeten zijn. Dit kan gerealiseerd worden met behulp van een *normaal verdeeld toevalsgetal (NRN)* respectievelijk een *toevalsgetal met driehoeksverdeling (TRN)*, beide te berekenen uitgaande van een aselekt getal tussen 0 en 1.

In elke simulatierun wordt aan elke factor die de einduitkomst mede bepaalt een voor die specifieke run eenmalig geldende numerieke waarde toegekend:

$$\text{Waarde} = \text{Gemiddelde} + \text{NRN} \times \text{Standaarddeviatie}$$

(bij normale kansverdeling)

of:

$$\text{Waarde} = \text{Lage waarde} + \text{TRN} \times (\text{Hoge waarde} - \text{Lage waarde})$$

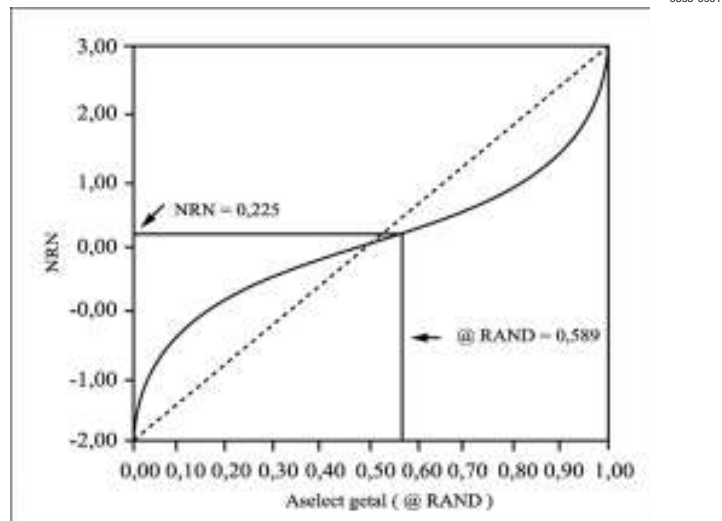
(bij driehoekige kansverdeling)

*4.2.1. Normaal verdeeld toevalsgetal NRN*

NRN is een normaal verdeeld toevalsgetal met een gemiddelde ( $\mu$ ) = 0 en een standaarddeviatie ( $\sigma$ ) = 1.

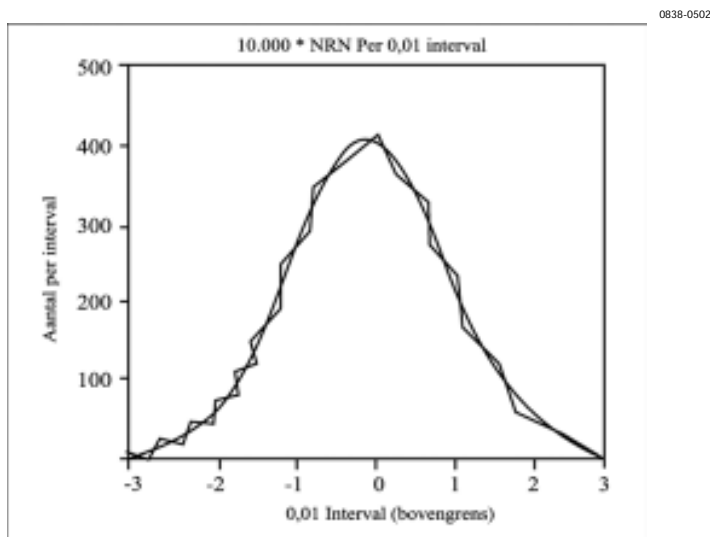
Een aselekt getal tussen 0 en 1, bijvoorbeeld @RAND in Lotus 123®, kan worden vertaald in een NRN met een algoritme van Berryman en Himmelblau (1971):

$$\text{NRN} = \sqrt{-2 \ln(@\text{RAND})} \times \cos(2\pi @\text{RAND}) \tag{3}$$



*Figuur 7. NRN uit een aselekt getal.*

Een vertaallijn van aselect getal naar NRN is gegeven in figuur 7. Duidelijk is in figuur 7 te zien hoe de helling van de vertaallijn het kleinst is rond  $@RAND = 0,5$ . Er zal dus veel vaker een NRN berekend worden met een waarde rond 0 dan één met een extreme waarde rond + of  $-2,5$ .



Figuur 8. Toets van NRN.

Dat de gegenereerde NRN's inderdaad normaal verdeeld zijn rond  $\mu = 0$ , met  $\sigma = 1$ , is mooi te zien in figuur 8. Die grafiek geeft de frequentieverdeling van 10.000 met bovenstaand algoritme berekende NRN's, in intervals van 0,01. Een tweede run zou weer geringe – maar niet identieke – schommelingen laten zien rond de ideale frequentieverdeling zoals die ook ingetekend is.

(In Lotus 123® versie 5 is het aantal statistische @functies overigens sterk uitgebreid. Nu kan een NRN simpelweg worden opgeroepen met de @NORMAL functie.)

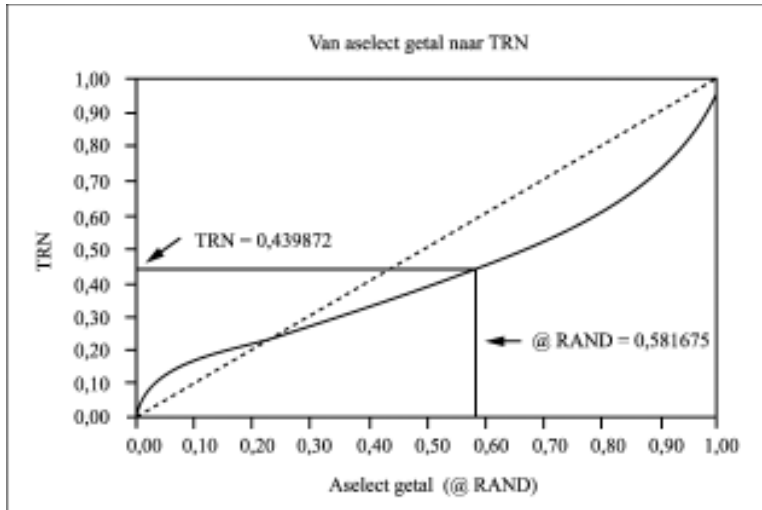
#### 4.2.2. Toevalsgetal met driehoeksverdeling TRN

TRN is een toevalsgetal met driehoeksverdeling; de modus (MO) ligt links (of rechts) van het gemiddelde; TRN ligt tussen 0 en 1. Een aselect getal tussen 0 en 1, bijvoorbeeld  $@RAND$  in Lotus 123®, kan worden vertaald in een TRN met een algoritme van Sprow (1967):

$$\text{TRN} = \sqrt{\text{MO} \times @\text{RAND}} \quad (@\text{RAND} \leq \text{MO}) \quad (4)$$

$$\text{TRN} = 1 - \sqrt{(1-\text{MO}) \times (1-@\text{RAND})} \quad (@\text{RAND} \geq \text{MO})$$

0838-0503



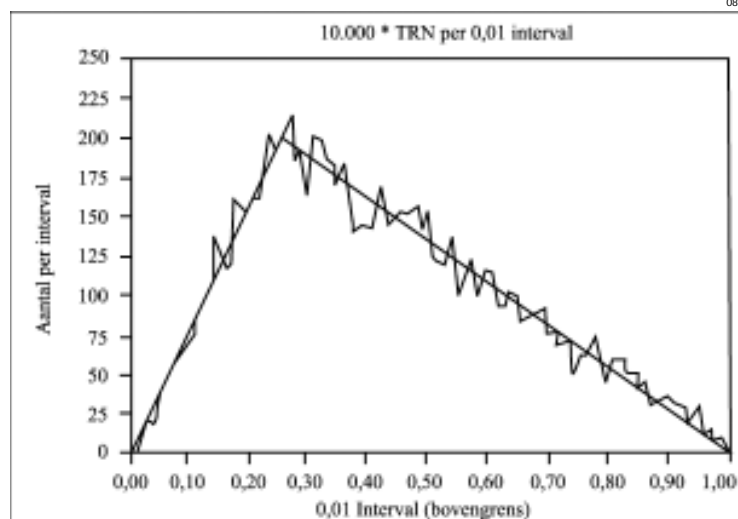
Figuur 9. TRN uit een aselekt getal.

De vertaallijn van aselekt tussen 0 en 1 naar TRN tussen 0 en 1 voor  $\text{MO} = 0,25$  is gegeven in figuur 9. De helling is het vlakst rond  $@\text{RAND} = \text{MO} = 0,25$ , zodat de frequentie van TRN's rond de 0,25 ook het hoogst zal zijn. In figuur 10 is dat mooi te zien: ook nu weer schommelingen rond de ideale lijnen. De grootte van de schommelingen is inherent aan het aantal runs en aan de grootte van de intervals en ook reproduceerbaar; de exacte ligging van de schommelingen uiteraard niet.

#### 4.2.3. Andersverdeelde toevalsgetallen HRN

Bij het uitvoeren van Monte Carlo-simulaties kan men zich meestal goed redden met de symmetrische en de asymmetrische verdelingen zoals hierboven besproken. Andere frequentieverdelingen zijn natuurlijk denkbaar, en, mits een algoritme voor het berekenen uit een aselekt getal beschikbaar is, even bruikbaar.

De computersoftware REP/PC van Decision Sciences (Curran, 1989) voert een Monte Carlo-simulatie uit op basis van een eigen (*confidential and undisclosed*) heuristiek. Een vertaallijn van aselekt



Figuur 10. Toets van TRN.

getal naar een HRN (heuristisch toevalsgetal) zal dan in meerdere of mindere mate afwijken van die voor de bovenbeschreven NRN en TRN. De onderliggende principes zijn uiteraard dezelfde.

Dat een HRN betrouwbaarder resultaten zou geven dan een NRN of een TRN valt moeilijk te bewijzen, bij gebrek aan cijfers uit de praktijk. Wel stelt Curran (1990) dat dankzij zijn heuristiek het zogenaamde *iatrogene risico* vermeden wordt. Dit is het risico van een te lage raming, dat kan voortkomen uit ondeskundig gebruik van de Monte Carlo-methode in een *spreadsheet* (al dan niet opgetuigd met een *add in* zoals @RISK of Crystal Ball) door iemand met onvoldoende kennis van de statistiek.

## 5. De wet van Pareto

Bij een ingewikkeld model, zoals een uitgebreide investerings- of kostprijsraming, zal de einduitkomst niet door alle factoren in dezelfde mate worden beïnvloed. Onder verwijzing naar Pareto (de uit de economie welbekende 20/80 regel) werkt Curran (1989) in zijn REP/PC met een drempelwaarde voor de minimale, door één enkel element te veroorzaken afwijking in de einduitkomst.

Op basis van empirische gegevens stelt hij de grens op 0,5% van de einduitkomst voor *conceptual estimates* en op 0,2% voor *detailed estimates*. Kan een element de einduitkomst niet voor ten minste de drempelwaarde beïnvloeden, dan wordt het bij de risico/kans analyse niet meegenomen.

## 6. Een voorbeeld

In de conceptuele fase van een procesontwerp wordt de *Onsite* investering in een 50 ktpa azijnzuurplant geraamd met behulp van de Monte Carlo-methode.

Daarbij wordt uitgegaan van de kosten van de *Major Equipment*, te bepalen op basis van het DACE-prijzenboekje en/of prijsindicaties van leveranciers.

Op basis van ervaring en/of *expert interviews* worden aan elk apparaat naast de meest waarschijnlijke waarde een maximaal en een minimaal waarschijnlijke waarde (met bijbehorende waarschijnlijkheid) toegekend. Voor de reactor wordt dit bijvoorbeeld met 95% zekerheid 350 k\$  $\pm$  30%, normaal verdeeld. In een enkele Monte Carlo-run wordt de kostprijs van die reactor dan berekend als 350 + NRN  $\times$  52,5 u\$. Dit laatste bedrag is de standaarddeviatie, overeenkomend met de helft van 30% van 350 k\$. De uitkomst in de enkele run is (bijvoorbeeld) 343 u\$ (zie ook Tabel I).

Wil de reactor kunnen functioneren, dan moet hij opgesteld en aangesloten worden. Dit vraagt om een fundering, pijpen, instrumentatie, enzovoort. De reactor vormt dan als het ware de kern van een module, waarbij een complete *plant* geacht wordt te bestaan uit naadloos (en kosteloos) aaneengeschakelde modules. Om de kosten van een module te bepalen worden de kosten van de kale apparatuur elk vermenigvuldigd met een individuele *Hand-* of opstelfactor. In te zetten waarden zijn spaarzaam in de literatuur te vinden, bijvoorbeeld bij Cran (1981).

Laat de *Hand*-factor voor de reactor in kwestie de waarde 2,60  $\pm$  30% hebben, in dit voorbeeld eveneens normaal verdeeld en gegeven met 95% zekerheid.

De waarde van de *Hand*-factor in dezelfde Monte Carlo-run is dan bijvoorbeeld 3,04.

De kosten van de reactormodule in die enkele Monte Carlo-run worden dan geschat op  $3,04 \times 343 = 1044$  k\$. Evenzo worden in dezelfde run de kosten bepaald van de overige modules van de azijn-

zuurfabriek. Optellen van de modulekosten geeft dan een uitkomst voor de *onsite*-investering in de complete *plant*.

In Tabel I is het resultaat van een enkele Monte Carlo-raming van de *onsite*-investering in een azijnzuurfabriek weergegeven. Daarbij zijn in dit voorbeeld eenvoudigheidshalve alle kosten en alle *Hand*-factoren als normaal verdeeld rond een gemiddelde gekozen. Wordt een driehoekige (of nog andersoortige) verdeling juist geacht, dan is het model natuurlijk eenvoudig aan te passen met eerder beschreven procedures.

Azijnzuurplant	Kosten fob		MC-Sim	Handfactor		MC-Sim	Module MC-Sim
	[k\$]	+/-	[k\$]	[-]	+/-	[-]	[k\$]
Onsites Estimate	350	105,0	343	2,60	0,78	3,04	1044
Reactor	700	70,0	745	1,70	0,17	1,62	1210
Compressor	800	120,0	747	2,70	0,41	2,78	2073
Kolommen	100	10,0	97	2,10	0,21	2,13	207
Vaten							
Warmtewisselaars	275	27,5	265	2,20	0,22	2,21	586
Pompen	75	7,5	75	2,60	0,26	2,49	186
	Uitkomst enkele Monte Carlo-run [k\$]:						5307

Tabel I. De uitkomst van een enkelvoudige Monte Carlo-simulatie.

In een Monte Carlo-raming wordt bovenbeschreven procedure een groot aantal malen doorlopen. Van de uitkomsten van N (bijvoorbeeld 1000) runs kan een gemiddelde met bijbehorende standaarddeviatie worden bepaald. Tabel II geeft het resultaat van vijf Monte Carlo-ramingen van 1000 runs elk.

Vijf Monte Carlo Ramingen Onsites [k\$]; N = 1000	#1	#2	#3	#4	#5	Theoretische Waarde
Onsite Investering	5279	5273	5264	5255	5288	5270
Standaarddeviatie	308	318	315	325	326	315

Tabel II. Vijfmaal een Monte Carlo-raming van de Onsite investering in een 50 ktpa azijnzuurfabriek.

In het gekozen voorbeeld is voor alle factoren een normale verdeling aangenomen. Dat betekent dat de theoretische waarde van de uitkomst eenvoudig te berekenen is:  $350 \times 2,60 + 700 \times 1,70 + 800 \times 2,70 + 100 \times 2,10 + 275 \times 2,20 + 75 \times 2,60 = 5270$  k\$.

Te constateren valt dat de uitkomsten van de Monte Carlo-ramin-



gen ook bij het relatief geringe aantal van 1000 runs minder dan 0,5% afwijken van de verwachte waarde.

Ook voor de standaarddeviatie is in dit eenvoudige voorbeeld een theoretische waarde te berekenen. Daarbij wordt gebruik gemaakt van twee formules uit de foutenleer voor het vermenigvuldigen en het optellen van  $A \pm a$  en  $B \pm b$ . (a en b zijn daarbij absolute getallen, geen percentages.)

Voor het produkt P geldt:

$$P \pm p = A \times B \pm \sqrt{(A \times b)^2 + (a \times B)^2} \quad (5)$$

en voor de som:

$$S \pm s = A + B \pm \sqrt{a^2 + b^2} \quad (6)$$

of meer algemeen:

$$S \pm s = \sum_{i=1}^{i=n} (A_i \pm a_i) = \sum_{i=1}^n A_i \pm \sqrt{\sum_{i=1}^n a_i^2} \quad (7)$$

Voor de kosten van de reactormodule wordt met formule 5 berekend:

$$350 \times 2,60 \pm ((350 \times 0,78)^2 + (105 \times 2,60)^2)^{1/2} = 910 \pm 386 \text{ k\$}$$

Voor de overige modules worden de  $\pm$  marges op de modules op dezelfde manier berekend. Ze zijn respectievelijk: 168, 458, 30, 86 en 28 k\$.

Met formule 7 wordt de theoretische fout in de totale investering bepaald op:  $(386^2 + 168^2 + 458^2 + 30^2 + 86^2 + 28^2)^{1/2} = 630 \text{ k\$}$ . Deze 630 k\$ volgt uit de met een 95% zekerheidsgrens opgegeven waarden voor de fob-prijzen en de *Hand*-factoren, en heeft zelf ook een zekerheid van 95%. De gehele on-site investering ligt dus met 95% zekerheid tussen  $5270 - 630$  en  $5270 + 630 \text{ k\$}$ .

630 k\$ is dan tweemaal de standaarddeviatie, die dus 315 k\$ draagt.

De uit de Monte Carlo-simulaties gevonden standaarddeviatie wijkt eveneens minder dan 0,5% af van de theoretische waarde.

Voor een niet-symmetrische driehoeksverdeling is de mediaan-gemiddelde waarde volgens Vrijland (1986) te vinden uit  $(LO + MO + HI)/3$ .

Ook als driehoeksverdelingen gebruikt worden in het simulatiemodel zijn 1000 runs voldoende om de theoretische waarde tot op 0,5% te benaderen.

Het resultaat van een uitgebreide Monte Carlo-simulatie wordt inzichtelijk weergegeven in een plaatje als figuur 11. Deze grafiek laat zien hoe met een Monte Carlo-raming de investering in de azijnzuurfabriek werd bepaald op 5245 k\$, met een overschrijdingskans van 50%. Wil de investeerder weten op welk bedrag hij moet rekenen om de overschrijdingskans terug te brengen tot 20%, dan moet hij een investeringsraming aanhouden van 5530 k\$. Hij kan ook uitgaan van 5245 k\$ mits hij  $5530 - 5245 = 285$  k\$ als „onvoorzien” in zijn raming opneemt. Er zij nadrukkelijk op gewezen dat de post „onvoorzien” dient ter compensatie van de kosten opgeroepen door marges in de financiële uitgangspunten en door onvolledigheid van de overeengekomen omvang van een opdracht, voorzover niet veroorzaakt door orderwijzigingen of door overmacht (Anonymus, 1989).

## 7. Discussie

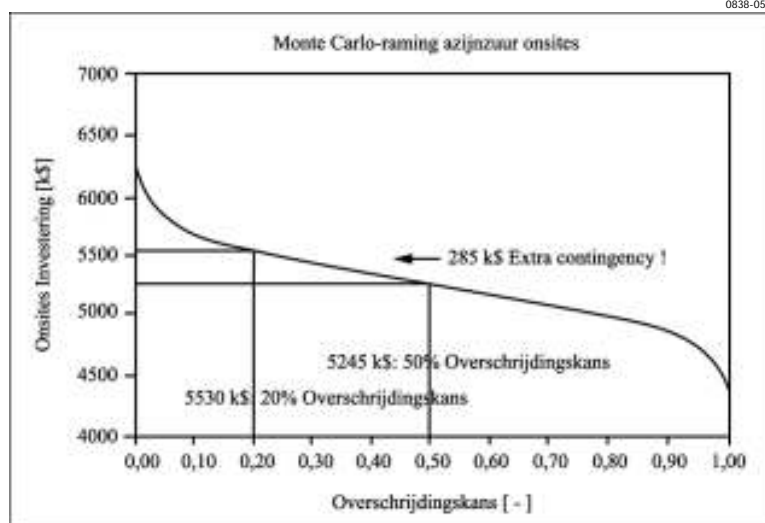
In het bovenstaande zijn de principes waarop *Range Estimating* berust behandeld. Met twee voorbeelden is gedemonstreerd welke output de techniek oplevert.

*Range Estimating* is geen beschermde naam, maar als over de mogelijkheden gesproken wordt heeft men daarbij vaak de door de Amerikaan M. W. Curran in 1972 voor McDonnell Douglas ontwikkelde software „REP®” (Range Estimating Program) of het recentere „REP/PC®” (voor de PC) in gedachten. Dit programma wordt sindsdien op ruime schaal in uiteenlopende takken van industrie, van *Aerospace* en *Aircraft* tot *Transport* en *Utilities*, met daartussenin onder meer *Banking (Financial Engineering!)*, *Chemicals*, *Insurance* en *Process* gebruikt.

Het programma combineert Monte Carlo met Pareto en eigen (vertrouwelijke) heuristiek. De gebruiker blijven daardoor beslissingen over te gebruiken kansdichtheidsverdelingen bespaard: hij hoeft alleen maar voor elke component de grenzen aan te geven waarbinnen de prijs ligt. In principe zijn, afgezien van die heuristiek (waarvan overigens nauwelijks een grote afwijkende invloed op de uitkomst te verwachten is), met Lotus 123® overeenkomstige resultaten te verkrijgen. Dat betekent niet dat REP geen nuttige functie zou vervul-

len: het moet gezien worden als een gebruiksvriendelijk programma waarmee mogelijke budgetoverschrijdingen en kansen en risico's opgespoord en gekwantificeerd kunnen worden zonder dat daarvoor diepgaande statistische kennis nodig is.

(Een tekst als deze schrijven is in beginsel ook heel goed mogelijk zonder tekstverwerkingssoftware als WordPerfect®, maar zo'n programma maakt het toch wel gemakkelijker.)



*Figuur 11. De azijnzuurplant vraagt een investering van 5530 k\$ met 20% op overschrijding, mits de scope ongewijzigd blijft.*

De invoergegevens, zoals de kostenrange van apparatuur, grondstoffen, arbeid, enzovoort blijven uiteraard de zorg en de verantwoordelijkheid van de gebruiker.

## 8. Epiloog

De studente in het vak Proceseconomie wist met aanvullende studie haar kennis van het vak te vergroten tot op het niveau van een 8. Het zat haar niet tegen, zodat zij haar examen inderdaad aflegde met een 8 als resultaat. Omdat zij door haar Monte Carlo-raming van het te verwachten cijfer haar kennis in praktijk bracht, en daarmee zo duidelijk begrip voor de doelstellingen van het vak aan de dag legde, kreeg zij een punt extra. Resultaat: een fraaie 9!

## 9. Literatuur

- Anonymus, *Onvoorzien in relatie tot de risicobepaling bij investeringsbegrotingen*, Handboek Cost Engineers, DACE/Samsom, 1989- (losbladig).
- Berryman, J. E. en D. M. Himmelblau, *Effect of Stochastic Inputs and Parameters on Process Analysis and Design*, Ind.Eng.Chem. Process Des.Develop., 1971 10(4) 441-449.
- Cran, J., *Improved Factored Method Gives Better Preliminary Cost Estimates*, Chemical Engineering 1981, April 6, 65-79.
- Curran, M. W., *Range Estimating, Measuring Uncertainty and Reasoning with Risk*, Cost Engineering 1989 31(3) 18-26.
- Curran, M. W., *Range Estimating Reduces Iatrogenic Risk*, 1990 AACE Transactions, K.3.1-3.
- Sprow, F. B., *Evaluation of Research Expenditures Using Triangular Distribution Functions and Monte Carlo Methods*, Industrial and Engineering Chemistry 1967, 59(7), 35-38.
- Vrijland, M. S. A. en H. J. Fonteijn, *The Monte Carlo Method in Cost Estimations*, 1986 AACE Transactions, A2, 1-7.