Tijdschrift voor Economie en Management,
Vol. XXIV, nr. 2, 1979

P. VAN GOETHEM*
M.J. GOOVAERTS*

Actuarieel gebruik van de James-Stein Estimatietheorie.

INLEIDING

In een artikel van Fl. De Vylder [1] wordt de recente ontwikkeling van de credibiliteitstheorie gegeven. De idee waarop deze theorie stoelt kan als volgt verklaard worden. Onderstel dat we een portfolio van automobielverzekeringen bekijken, die op een bepaalde manier kan onderverdeeld worden in een aantal klassen.

Zij X de stochastische variable die bijvoorbeeld de gemiddelde schade per periode in de portfolio voorstelt en Y de stochastische variable die de gemiddelde schade per periode in één bepaalde klasse voorstelt, dan is het probleem hoe men door middel van waarnemingen van $X(x_1, s_2 \dots s_t)$ en $Y(y_1, y_2, \dots y_t)$ een zuivere premie kan bepalen voor de beschouwde klasse rekening houdend met de vastgestelde schadeervaringen in die klasse en in de ganse portfolio.

Men kan het gemiddelde van de klassen beschouwen

$$m_k = (y_1 + y_2 + \dots + y_t)/t$$

en dat van de ganse portfolio :

$$m_p = (x_1 + x_2 + \dots + x_t)/t$$

Een strategie die kan gevolgd worden is per kontrakt een premie m_k aanrekenen voor de risico's binnen de beschouwde klasse. Als alternatieve strategie zou men ook per kontrakt een premie m_p kunnen aanrekenen.

De respectievelijke auteurs zijn verbonden aan :
Departement Wiskunde UIA
Universiteitsplein 1, B-2620 wilrijk, Belgium.
Momenteel verbonden als actuariaatsinspecteur op de Controledienst voor de Verzekeringen.
Instituut voor Actuariële Wetenschappen, K.U.LKEUVEN, Dekenstraat 2, B-3000 Leuven, Belgium.
Wij danken ook Prof. Dr. W. Gochet voor zijn suggesties en opmerkingen.

Als de beschouwde klasse enkel relatief goede risico's bevat in vergelijking met de risico's van de portfolio dan zullen bij het aanrekenen van een premie m_p deze kontrakten relatief teveel betalen en systematisch bijbetalen voor de slechte risico's. Daartegenover staat het feit dat als het verleden uitmaakt dat de beschouwde klasse een «goede» klasse is dit geenszins perfecte aaduidingen geeft voor het goed blijven van de beschouwde klasse in de toekomst. Aanrekenen van een premie m_k is dus evenmin aanvaardbaar. Dit leidt ons tot de in actuariële middens welgekende credibiliteitsformule.

$m = zm_k + (1\text{-}z)m_p$ met $0 < z < 1$

Dit geeft het ontstaan aan de zogenaamde credibiliteitstheorie, die een aanzienlijk pakket uitmaakt van het onderzoeksdomein van de moderne actuaris. Voor recente ontwikkelingen wordt verwezen naar referentie [2]. Voor verdere literatuurverwijzing zie referentie [3]. Het is de bedoeling van de huidige bijdrage om aan de hand van twee praktische voorbeelden aan te tonen dat een James-Stein estimator perfect kan aangewend worden om het doel van de credibiliteitstheorie te verwezenlijken en dit voor een *reële* portfolio, waarin verzekeringskontrakten betreffende diverse risico's worden opgenomen.

Het reële risico van al de verzekeringstakken samen is immers een belangrijker maatstaf dan het risico per tak. James-Stein estimatoren houden juist rekening met het globale risico zoals zal blijken uit hetgeen volgt. De voorbeelden en behandeling van deze theorie zullen meer duidelijkheid brengen in deze beweringen.[1]

I. MEERVOUDIGE RISICO's.

A. Een verzekeringsonderneming ziet zich steeds geplaatst voor het probleem : bepaal de premie die gevraagd dient te worden om een gegeven risico te kunnen dekken. Als abstraktie gemaakt wordt van de opslagen dan wordt de risicopremie bepaald.

Onderstel dat een onderneming n risico's van verschillende aard (brand, diefstal, hagel, enz.) die onafhankelijk zijn, wil dekken, dan dient zij hiervoor een risicopremie vast te stellen voor elk afzonderlijk. Deze risicopremie kan zij echter maar schatten en zij wenst deze schatting zo «goed mogelijk» uit te voeren. Daarvoor moet zij dus beschikken over een steekproef van resultaten uit het verleden.

Om de risicopremie voor het i-de risico te bepalen zal zij dus gebruik maken van een steekproef betreffende dit risico en dit zal zij voor elk afzonderlijk doen zonder rekening te houden met de resultaten van de andere risico's. Zij kan ook echter bij de bepaling van de i-de risicopremie rekening houden met een steekproef van resultaten van de andere risico's. Bepaalt zij nu op deze twee manieren de gezamenlijke premie dan zal, onder zekere voorwaarden, de tweede werkwijze een kleinere afwijking van de geschatte waarden (voor de premies) t.o.v. de juiste waarden opleveren dan de eerste methode. Voor andere resultaten in dit verband wordt verwezen naar [5] dat een bevattelijk overzicht geeft van de resultaten in dit domein.

B. Om de begrippen die gesuggereerd werden beter te omschrijven zullen we een wiskundig model opstellen van dit probleem en aan de hand daarvan de gevolgen nagaan. Het model dat we behandelen zal iets algemener zijn dan strikt noodzakelijk. De bedoeling is echter de verwantschap aan te tonen van de James-Stein estimator met de Bayes estimator. Daarom zal het principe van de Bayes estimator beschreven worden. De beschrijving van het model luidt als volgt.

De onderneming wenst risico's te dekken die van verschillende aard zijn en in elke referentie periode zullen de schadebedragen $X_1, \ldots, X_n$ uitgekeerd worden. De grootheden zijn toevallige veranderlijken die een zekere verdeling volgen. De premie wordt bepaald door een operator te laten werken op deze verdelingsfuncties. Om de risicopremies te bepalen dient men dus voor elke verdeling haar gemiddelde te zoeken. Deze gemiddelden $P_1, \ldots, P_n$ stellen dan de premie incasso's voor die nodig zijn om een evenwicht te bekomen tussen de inkomsten van de maatschappij en haar verwachte uitgaven. Deze getallen $P_1, \ldots,$ P_n zijn niet gekend, wel kunnen deze geschat worden aan de hand van resultaten uit voorbije referentieperioden.

De toevallige veranderlijke X_i is een stochastische som van onafhankelijke en gelijkverdeelde schadebedragen $Y_{i,k}$:

$$X_i = Y_{i,1} + \ldots + Y_{i,N_i}$$

Onderstellen we verder dat N_i Poisson verdeeld is, dan volgt X_i benaderend een normale verdeling met gemiddelde $E(N_i)\,E(Y_i)$ en variantie

$$E(N_i)\,\sigma^2(Y_i) + \sigma^2(N_i)\,E(Y_i)^2.$$

Deze laatste onderstelling is niet noodzakelijk om een James-Stein estimator te bepalen, het is alleen een verrechtvaardiging om deze estimator in de actuariële risicotheorie aan te wenden aangezien het hiervoor beschreven model het klassieke ingrediënt van de risicotheorie is.

Om het model wiskundig hanteerbaar te maken zullen we nog volgende onderstelling maken :

(1) De stochastische variabelen X_i zijn normaal verdeeld met gemiddeld M_i en variantie D

(2) De grootheden M_i zijn onbekend en kunnen opgevat worden als resultaten uit een normale verdeling met gemiddelde μ en variantie A, die bekende grootheden zijn.

Stellen we nog $m_i = M_i - \mu$ dan zijn de m_i genomen uit een $N(O,A)$ verdeling.

Alvorens het probleem verder te onderzoeken past het om even stil te staan bij deze onderstellingen. Wat voorwaarde (2) betreft, deze is niet noodzakelijk om een James-Stein estimator te bepalen. Inderdaad kunnen we volstaan met gemiddelden M_i die vast zijn. Nu is het verder ook zo dat de voorwaarde (2) moeilijk kan geverifieerd worden aan de hand van meetresultaten alleen. Het besluit dat deze geldig is in de praktische voorbeelden zal dan ook genomen worden op basis van andere overwegingen dan alleen kwantitatieve resultaten. Zij wordt hier ondersteld omdat het ook de bedoeling is de Bayes estimator in de vergelijking te betrekken.

Het probleem is nu M_i te schatten uitgaande van een steekproef van X_i :

$X_{i,1}, \ldots, X_{i,k}$.

Het gemiddelde $\overline{X}_i = \frac{1}{k} \sum_{j=1}^{k} X_{ij}$ is dan een voldoende statistiek en volgt een normale verdeling met gemiddelde M_i en variantie D/n. We zullen dus onderstellen dat de toevalige variabelen X_i reeds het gemiddelde zijn van de onafhankelijke observaties en dat M_i en D de parameters zijn van de verdeling van X_i. Er zal dus gehandeld worden alsof we over één observatie beschikken voor elke veranderlijke X_i, in plaats van k. Als m_i geschat wordt door δ_i dan gaan we uit van een kwadratische verliesfunctie (die in de credibiliteitstheorie overwegend wordt toegepast [9] :

$$L(m_i, \delta_i) = (\delta_i - m_i)^2$$

De risicofunctie is dan het gemiddeld verlies als de werkelijke waarde m_i is :

$$R(m_i, \delta_i) = E_{m_i} [L(m_i, \delta_i)]$$

waarbij E_{m_i} de verwachtingswaarde is met betrekking tot de voorwaardelijke verdeling van X_i gegeven $m(m_1, \ldots m_n)$, waarbij we toelaten dat X_i afhangt van deze m. Analoge regels gelden voor de schatting van de variantie D. Om technische redenen zal niet D geschat worden maar

$$B = \frac{D}{A+D}$$

De verdeling van m_i gegeven X_i, is normaal verdeeld met gemiddelde $(1\text{-}B)\ (X_i\text{-}\mu)$ en variantie $D(1\text{-}B)$.
Als de grootheid B onbekend is kan getracht worden deze te schatten, nl.

$$S = \sum_{i=1}^{n} (X_i - \mu)^2$$

is een voldoende statistiek voor B met verdeling $1/B \cdot \chi^2_n$.

De Bayesschatter voor m_i is dan $(1\text{-}B)\ (X_i\text{-}\mu)$ waaruit volgt dat de Bayesschatter voor

$M_i = \mu + (1\text{-}B)\ (X_i\text{-}\mu)$

en het Bayesrisico, dat een maat is voor de afwijking van de echte waarde, wordt gegeven door $1\text{-}B$.
De meest aanzienlijke schatter voor m_i wordt gegeven door $X_i\text{-}\mu$ die we zullen voorstellen door δ_i en deze heeft een risiko gelijk aan D.

Wanneer de verliesfunktie nu de som is van afzonderlijke kwadratische verliezen, d.w.z. $L = \sum_{i:1}^{n} (\delta_i - m_i)^e$, dan werd aangetoond door James en Stein [7, 8], dat voor $n \geqslant 3$ de estimator

$$\delta_i = [1 - \frac{(n\text{-}2)D}{S}]\ (X_i - \mu)$$

uniform beter is dan de meest aanzienlijke schatter (zie ook referenties 4-7).
Een schatter van deze gedaante waarbij de resultaten van de andere klassen optreden in het schatten van een parameter uit één klasse (langs S om) heet een James-Stein estimator.

James en Stein toonden aan dat het risiko van deze schatter strikt kleiner is dan nD, wat het gezamenlijke verlies is van de meest aanzienlijke schatters. Als een funktie van m_i wordt het risico $R(m, \delta)$ gegeven door

$1 - \frac{n-2}{n} \quad E_m\left(\frac{n-2}{S}\right)$ met $\delta = (\delta_1, \delta_2, \ldots, \delta_n)$.

Hetgeen het meest opvalt bij deze laatste methode is dat om de waarde m_i te schatten, men een schatter bekomt die niet alleen van X_i afhangt, maar ook van de observaties X_j $(j \neq i)$.

C. De verzekeringsonderneming in kwestie zou dus bij het bepalen van het i-de premie-incasso rekening houden met de observaties van de andere risico's. Werkt zij met een kwadratische verliesfunktie dan gaat haar «verlies in estimeren» kleiner zijn dan wanneer zij dit voor elk risico afzonderlijk zou doen.

Eenzelfde redenering zou opgaan als de maatschappij alleen haar verlies

$$\Sigma_i \, [(X_i - M_i)^2 t]_+ \qquad \text{waarbij} \quad Y_+ \quad \begin{aligned} &= Y \text{ als } Y > o \\ &O \text{ als } y < O \end{aligned}$$

zo klein mogelijk zou maken. De grootheid M_i die het gemiddelde is van de toevallige veranderlijke is het totaal premie incasso nodig om een actuarieel evenwicht te bekomen, zodat $P_i = M_i$.

Afgezien van juridische of boekhoudkundige bezwaren zou een premieberekeningsprincipe dat voor elk risico rekening houdt met observaties van de andere risico's een kleiner verwacht totaal verlies opleveren. De afzonderlijke verliezen per tak zouden wel groter kunnen zijn. In dit verband wordt verwezen naar de waarschuwing die staat in [4].

ii. credibiliteit.

A. Onderstel dat een maatschappij een portefeuille beheert die kan opgevat worden als een samenstelling van homogene risicoklassen. Die risicoklassen zijn niets anders dan deelverzamelingen die gekarakteriseerd worden door eenzelfde risicoparameter. In de meeste modellen die men beschouwt wordt de gemiddelde schade $\mu(\Theta)$ dan bekend ondersteld zodra men de risicoparameters kent ; deze is echter een toevallige veranderlijke Θ, die men alleen kan estimeren. Daarnaast kan ook nog de gemiddelde schade in de ganse portefeuille bekeken worden : ofwel onderstelt men dat deze bekend is m, ofwel dat deze kan geschat worden uit t referentieperiode : $\hat{m}$.

Om nu de gemiddelde schade in de klasse met risicoparameters Θ te schatten wordt door Europese actuarissen sedert lange tijd gebruik gemaakt van volgende uitdrukking [9] :

hetzij

$\mu(\Theta) = z\hat{m}_\Theta + (1\text{-}z)m$

hetzij

$\mu(\Theta) = z\hat{m}_\Theta + (1\text{-}z)\hat{m}$

naargelang men het portefeuillegemiddelde kent of moet schatten. In deze uitdrukkingen is $\hat{m}_\Theta$: een schatting voor de gemiddelde schade in de risicoklasse Θ gebaseerd op schadegevallen uit de voorbije referentie periodes. De koëfficiënt z bij $\hat{m}_\Theta$ heet dan de credibiliteitsindeks of de credibiliteit bij de geobserveerde klasse.

Deze indices worden bepaald door de afwijking van $\mu(\Theta)$ zo klein mogelijk te maken voor een gegeven verliesfunktie. Een voorbeeld van verliesfunktie kan de kwadratische funktie :

$L(\mu(\Theta), \hat{\mu}(\Theta)) = (\mu(\Theta) - \hat{\mu}(\Theta))^2$

zijn, zodat de variantie op $\hat{\mu}(\Theta)$ minimaal moet zijn ; uit deze voorwaarde bepaalt men dan z.

B. Keren we terug tot de James-Stein estimator. Een maatschappij bezit een portefeuille bestaande uit homogene deelklassen en voor elk van deze deelklassen dient zij een risicopremie vast te stellen. Onderstel dat $X_1, \ldots, X_n$ de schadebedragen zijn gedurende een referentieperiode in de klassen resp. 1, ..., n. Onderstel dat de maatschappij beschikt over een steekproef die zich uitstrekt over t periodes en dat de deelklassen ruim genoeg zijn om de normale verdeling voor de toevallige veranderlijken X_i te mogen onderstellen.

Als $\mu_1, \ldots, \mu_n$ de echte gemiddelden zijn in de klassen 1, ..., n dan wordt gevraagd estimatoren $\hat{\mu}_1, \ldots, \hat{\mu}_n$ te bepalen zodat de verwachte waarde van

$$\sum_{i=1}^{n} (\mu_i\text{-}\hat{\mu})^2$$

minimaal wordt.

We zullen nog onderstellen dat $\mu_1, \ldots, \mu_n$ genomen zijn uit een normale verdeling $(N(\mu, A)$. Deze onderstelling is niet noodzakelijk en dient om het verband te leggen met de Bayes estimator, zoals behandeld in de vorige paragraaf. De reden dat we opnieuw deze onderstelling maken is dat we beide problemen naast elkaar willen plaatsen zodat duidelijk blijkt dat de wiskundige behandeling in beide gevallen dezelfde is.

Verder zullen we door $X_i = \frac{1}{t} \sum_{i=1}^{t} X_i$, en het gemiddelde voorstellen van de observaties uit de i-de klasse voor t periodes en onderstellen dat de variantie van X_i gelijk is aan D_i, wat een bekende grootheid is.

Stellen we nog $\bar{X} = \frac{1}{n} \sum_{i=1}^{n} X_i$ het gemiddelde van de n klassen (m.a.w. $\bar{X} = \hat{m}$)

Naar analogie met het eerste probleem volgt nu dat $X_i - \mu \ N(O, D_i)$ verdeling volgt en stellen we

$$S_1 = \sum_{i=1}^{n} (X_i - \mu)^e \text{ en } S = \sum_{i=1}^{n} (X_i - \bar{X})^2$$

In dit geval wordt dus het portefeuillegemiddelde gekend verondersteld en wordt de James-Stein estimator voor μ_i :

$$\hat{\mu}_i = \mu + \left(1 - \frac{(n-2)D_i}{S_1}\right)(X_i - \mu)$$

Deze estimator heeft zoals werd aangetoond door James en Stein [5] een uniform beter verlies dan wanneer X_i zou genomen worden.

Herschrijven van deze estimator levert nog :

$$\hat{\mu} = \frac{(n-2)D_i}{S_1}\mu + \left(1 - \frac{(n-2)D_i}{S_1}\right) X_i$$

waaruit dus blijkt dat de «credibiliteits»-gedaante terug gevonden wordt voor de James-Stein estimator.

In het geval dat μ niet gekend is en dient geschat te worden door X wordt als James-Stein schatter gevonden [4] :

$$\hat{\mu}_i = \bar{X} + \left(1 - \frac{(n-3)D_i}{S}\right)(X_i - \bar{X})$$

of nog

$$\hat{\mu}_i = \frac{(n-3)D_i}{S}\bar{X} + \left(1 - \frac{(n-3)D_i}{S}\right) X_i$$

Dezelfde opmerking als in vorig geval is hier geldig.

Vermelden we ook nog dat in die gevallen waar

$$(1 - \frac{(n\text{-}2)D^i}{S_1}) \text{ resp}$$

$$(1 - \frac{(n\text{-}3)D^i}{S})$$

negatief is een «betere» estimator gevonden wordt als we deze faktor vervangen door O.

Tenslotte zij nog vermeld dat in het artikel van Efron en Moris [5] bewezen wordt dat, zodra $n \geqslant 3$, de James-Stein estimator «beter» is dan de Bayes estimator voor μ_i.

Zonder verder in te gaan in wiskundige technieken, zij nog vermeld dat de Bayes estimator voor μ_i in het vorige model gegeven wordt door

$$\hat{\mu}_i' = (1 - \frac{D_i}{A+D_i}) \qquad (X_i - \mu) + \mu$$

$$= \frac{D_i}{A+D_i} \quad \mu + \quad (1 - \frac{D_i}{A+D_i}) X_i.$$

M.a.w. hier wordt de credibiliteit teruggevonden zoals die voldoende bekend is aan actuarissen. [vergelijk hiervoor de uitdrukking voor b op blz. 101 in ref. [9] met deze voor δ_i^* in ref.[5] op blz. 120.]

Ten slotte weze nog opgemerkt dat men zich niet hoeft te beperken tot het geval van normale verdelingen. Mits eenvoudige aanpassingen kunnen resultaten bekomen worden voor diverse combinaties van verdelingen.

Praktisch voorbeeld.

Onderstel dat een verzekeraar zich gaat toeleggen om de volgende risico's te dekken voor kleine en middelgrote bedrijven : brand (A), ontploffing (B), stormschade (C), natuurevenementen met uitsluiting van hagel en vorst (D), hagel en vorst schade (E), alle andere schaden (F).

De premie die zal gevraagd worden is evenredig met het verzekerd kapitaal en de uitkeringen zullen gebeuren volgens de evenredigheidsregel.

Aan de hand van statistisch materiaal, dat men herleid heeft op een basis index, heeft men volgende resultaten gevonden (in miljoenen)

Jaren takken	1965	1966	1967	1968	1969
A	7.97	9.64	7.83	9.82	10.77
B	14.42	10.66	5.11	11.22	14.12
C	6.85	9.61	6.64	9.91	11.46
D	5.39	10.41	10.96	13.42	9.19
E	10.27	5.23	10.77	13.40	13.05
F	8.12	9.56	8.01	9.71	10.53
	1970	1971	1972	1973	1974
A	10.64	6.60	9.24	7.95	9.99
B	13.74	1.31	9.41	16.47	11.71
C	11.26	4.61	8.94	6.83	10.17
D	5.34	11.45	5.53	7.04	14.09
E	9.13	5.56	11.21	5.72	7.13
F	10.42	6.95	9.21	8.11	9.85

Deze gegevens werden bekomen door die polissen te nemen die gedurende deze tienjarige periode konden geobserveerd worden. De totale verzekerde kapitalen bedroegen voor elk van de takken 5000 miljoen aan de basis indeks.
De gemiddelde jaarlijkse schade die op deze manier wordt bekomen voor elk der takken bedraagt voor

A : 9.045
B : 10.817
C : 8.628
D : 9.282
E : 9.147
F : 9.047

Dit geeft dan de koëfficiënt aan waaruit de risikopremie zou bestaan voor 5000 F verzekerd kapitaal per tak indien we het gemiddelde als schatter nemen.
Om de James-Stein estimator te bepalen dient nog een schatting gekend te zijn van de varianties van de onderscheiden branches. Hiervoor vinden we respektievelijk voor :

$\hat{\sigma}_A^2 = 1.72/6 = 0.29$
$\hat{\sigma}_B^2 = 19.03/6 = 3.17$
$\hat{\sigma}_C^2 = 4.67/6 = 0.28$
$\hat{\sigma}_D^2 = 9.86/6 = 1.64$
$\hat{\sigma}_E^2 = 8.53/6 = 1.42$
$\hat{\sigma}_F^2 = 1.28/6 = 0.21$

en voor $\hat{S} = \sum_{i:l}^{6} (X_i - \hat{\mu})^2$ vinden we 2.9 daar $\hat{\mu} = 9.33$

Als James-Stein estimator zullen we hier de estimator nemen van de gedaante :

$$\hat{\mu}_i = \overline{X} \; + (1 - \; \frac{(n\text{-}3)D_i}{S} \; + (X_i - \overline{X})$$

Aangezien het aantal variabelen zeer klein is, is dit hier ook aangewezen.

Als estimator voor tak A vinden we nu :

$$\hat{\mu}_A = \frac{(6-3) \text{ x } 0.29}{2.90} \quad \text{x } 9.32 + \quad (1 - \frac{(6\text{-}3) \text{ x } 0.29}{2.90}) \text{ x } 9.045 = 9.13$$

Door analoge bewerkingen vinden we :

$\hat{\mu}_B = 9.33$
$\hat{\mu}_C = 8.83$
$\hat{\mu}_D = 9.33$
$\hat{\mu}_E = 9.33$
$\hat{\mu}_F = 9.11$

Vergelijken we beide estimators dan vinden we respectievelijk de gemiddelden en de James-Stein schatters

9.045	resp.	9.13
10.817		9.33
8.628		8.83
9.282		9.33
9.147		9.33
9.047		9.11

Als we nu een tipje van de sluier oplichten en verklaren hoe we deze getallen uit de steekproef bekomen werden, dan kunnen we beide estimators naar waarde schatten.

a) De gemiddelden voor de takken A t/m F werden genomen uit een N (10, 0.25) verdeling en hiervoor namen we voor

A	als gemiddelde	9.49
B		10.19
C		9.36
D		9.96
E		9.84
F		9.43

Deze getallen werden gegenereerd op een zakrekenmachine.

Tenslotte werden de getallen uit de steekproef op dezelfde manier gegenereerd uit de volgende verdelingen :

resp. $N(9.49\ ;\ 1.96)$, $N(10.19\ ;\ 18.49)$, $N(9.36\ ;\ 5.29)$
$N(9.96\ ;\ 17.64)$, $N(\ 9.84\ ;\ 15.21)$, $N(9.43\ ;\ 1.2)$

Als de som der kwadratische afwijkingen van de gemiddelden vinden we nu : 2.2136 en voor die der James-Stein schatters : 1.9095, hetgeen reeds het voordeel aantoont van de James-Stein schatter. Dit zijn de afwijkingen t.o.v. de exakte waarden.

Ter illustratie kunnen we ook even de Bayes estimaties bekijken. Daarvoor onderstellen we een a priori verdeling die $N(10\ ;\ 0.25)$ is, terwijl voor de D_i-waarden de schattingen genomen worden.
Op die manier bekomt men nu voor de verschillende takken de respektievelijke schattingen.

A :	9.55
B :	10.06
C :	9.35
D :	9.90
E :	9.87
F :	9.48

Tenslotte kunnen we deze resultaten nog vergelijken met diegene die men zou bekomen hebben met behulp van de schattingen voor μ en A, nl. :

$\mu =$ 9.33
$A =$ 0.48

en we vinden dan respektievelijk

A :	9.15
B :	9.52
C :	8.89
D :	9.32
E :	9.28
F :	9.13

Als som van de kwadratische afwijkingen vinden we nu respektievelijk : 0.0276 en 1.5986.

Hieruit kunnen we besluiten dat hoewel de James-Stein estimaties gemiddeld beter zijn dan de meeste andere schatters, dit niet steeds zo is in een bijzonder geval. Dit doet evenwel niets af aan de waarde van de

James-Stein schatters ; er dient namelijk rekening mee gehouden worden met het feit dat de Bayes estimatie gebruik maakt van informatie a priori.

Als besluit kunnen we formuleren dat in geval een kontrakt wordt afgesloten waarin de zes risiko's gedekt worden, de maatschappij een premie ontvangst zal vertonen die in haar totaliteit minder afwijkingen zal vertonen zo zij gebruik maakt van de James-Stein estimaties dan van de voor de hand liggende rekenkundige gemiddelden.

REFERENTIES

[1] Fl. De Vylder, «Le développement récent de la théorie de la crédibilité» (verschijnt in het *Tijdschrift van de «Koninklijke Vereniging der Belgische Actuarissen»).*

[2] Kahn, P.M., *Credibility : Theory and applications.* Proceedings of Actuarial Research Conference on Credibility, Berkeley, September 1974 Academic Press, New York 1975.

[3] N. De Pril, L. D'Hooge, M.J. Goovaerts, «A bibliography on credibility theory and its applications», *Journ. Comp. Applied Math.*, Vol. II, No. 1, pp. 55-62.

[4] Efron, B. en Moris, C., *Stein's paradox in statistics.* Scientific American, May 1977, 119-127.

[5] Efron, B. en Moris, C., Stein's estimation and its competitors — an empirical Bayes approach. *Journ. Amer. Statist. Assoc.* 68 (1973) 117-130.

[6] Efron, B. en Moris, C., Combining possibly related estimation problems *Journ. Royal Statist. Soc.* B 35 (1973) 379-421.

[7] James, W. en Stein, G., *Estimation with quadratic loss.* Proc. Fourth Berkeley Symp. on Math. Statist. and Prob., Univ. Calif. Press, 1961. 361-379.

[8] Stein, C., *Inadmissibility of the usual estimator for the mean of a multivariate normal distribution.* Proc. Third Berkeley Symp. on Mathem. Statist. and Prob., Univ. of California Press, 1956, vol. 1, 197-206.

[9] Bühlmann H., *Mathematical methods in Risk Theory*, Springer Verlag, 1970.

SUMMARY

Although it is not our intention to demystificate the James-Stein Estimation nor to give full proofs concerning the link that exists between this estimation and the Bayes estimation we will give some examples of the most representative applications to actuarial sciences. The James-Stein estimation can be applied as a justification for the well-known credibility formulae from which we derive the rule for a premium calculation.

An example is given where we are showing the power of the James-Stein estimation relative to the standard estimation of the mean and the Bayes estimation.