UNIVERSITAT DE BARCELONA
FACULTAT DE CIÈNCIES ECONÒMIQUES I EMPRESARIALS
DEPARTAMENT DE MATEMÀTICA ECONÒMICA, FINANCERA I ACTUARIAL

LA TEORIA DE LA CREDIBILITAT : EXTENSIONS I APLICACIONS

Lluís Bermúdez i Morata
Barcelona, maig 1997.

Taula 9.17 *Resultats de França pel mètode "two way"*

X_{js}	$s=1$	$s=2$	$s=3$	$s=4$	$s=5$	$s=6$	$s=7$	$s=8$	$s=9$	$s=10$
$j=1$	127.9	731.8	1953	724.7	344.3	64.3	86.9	50.7	0	0
$j=2$	339.1	1083	1097	523.6	191.2	112.1	130	114.9	0	0.1
$j=3$	740.3	2569	1959	614.5	521	14.3	7.1	89.2	0.1	0.2
$j=4$	686.8	2286	1075	319	11.8	24.8	104.9	91.9	0.1	0.1
$j=5$	405.4	1077	754.1	52.1	249.8	0	53.2	57.4	0.1	0.1
$j=6$	597.8	786.3	514.8	293.6	45	30.2	54	58.2	0	0.1
$j=7$	345.1	901.4	1040	0	176.9	29.9	53.3	57.5	0	0.1
$j=8$	136.5	675.9	577.3	188.6	122.4	20.7	36.9	39.8	0	0.1
$j=9$	372.6	285.8	593.5	180.7	117.3	19.8	35.4	38.2	0	0.1
$j=10$	706.6	1283	1178	358.8	232.9	39.3	70.2	75.7	0.1	0.1

Taula 9.18 *Resultats del Regne Unit pel mètode De Vylder*

X_{js}	$s=1$	$s=2$	$s=3$	$s=4$	$s=5$	$s=6$	$s=7$	$s=8$	$s=9$	$s=10$
$j=1$	102.6	227.2	175.5	117.4	0	71.1	0	26.4	55.9	0
$j=2$	1057	1348	360.4	76.8	31.2	47	68.7	5	0	0
$j=3$	723.7	1592	992.3	385.5	44.8	126.2	12.6	0	59.1	0
$j=4$	472	382.5	613.7	253.2	84.7	80.5	0	7.6	26.2	0
$j=5$	4.3	378.8	360.6	128.5	289.4	13.8	10.2	5.1	17.4	0
$j=6$	330.1	647.9	442.1	12.4	0	49.3	13.9	6.9	23.7	0
$j=7$	163.9	238.1	275.1	32.1	26.1	22.2	6.3	3.1	10.6	0
$j=8$	858.7	631.3	459.6	165.8	78.1	66.3	18.8	9.2	31.8	0
$j=9$	256.3	134	167.7	52.6	24.8	21	5.9	2.9	10.1	0
$j=10$	31	396.7	274.8	86.1	40.6	34.4	9.7	4.8	16.5	0

Taula 9.19 *Resultats del Regne Unit pel mètode De Vylder-Mack*

X_{js}	$s=1$	$s=2$	$s=3$	$s=4$	$s=5$	$s=6$	$s=7$	$s=8$	$s=9$	$s=10$
$j=1$	102.6	227.2	175.5	117.4	0	71.1	0	26.4	55.9	0
$j=2$	1057	1348	360.4	76.8	31.2	47	68.7	5	0	0
$j=3$	723.7	1592	992.3	385.5	44.8	126.2	12.6	0	52.2	0
$j=4$	472	382.5	613.7	253.2	84.7	80.5	0	8.7	29.8	0
$j=5$	4.3	378.8	360.6	128.5	289.4	13.8	14.5	7.1	24.6	0
$j=6$	330.1	647.9	442.1	12.4	0	43.5	12.3	6.1	20.9	0
$j=7$	163.9	238.1	275.1	32.1	23.8	20.2	5.7	2.8	9.7	0
$j=8$	858.7	631.3	459.6	147.6	69.5	59	16.7	8.2	28.3	0
$j=9$	256.3	134	185.8	58.2	27.4	23.3	6.6	3.2	11.2	0
$j=10$	31	590.2	408.8	128.1	60.4	51.3	14.5	7.1	24.6	0

Taula 9.20 *Resultats del Regne Unit pel mètode "two way"*

X_{js}	$s=1$	$s=2$	$s=3$	$s=4$	$s=5$	$s=6$	$s=7$	$s=8$	$s=9$	$s=10$
$j=1$	102.6	227.2	175.5	117.4	0	71.1	0	26.4	55.9	0
$j=2$	1057	1348	360.4	76.8	31.2	47	68.7	5	0	0
$j=3$	723.7	1592	992.3	385.5	44.8	126.2	12.6	0	58.3	0
$j=4$	472	382.5	613.7	253.2	84.7	80.5	0	8.9	30.5	0
$j=5$	4.3	378.8	360.6	128.5	289.4	13.8	12.8	6.3	21.8	0
$j=6$	330.1	647.9	442.1	12.4	0	43.2	12.2	6	20.8	0
$j=7$	163.9	238.1	275.1	32.1	24.2	20.5	5.8	2.9	9.9	0
$j=8$	858.7	631.3	459.6	161.2	75.9	64.5	18.2	9	31	0
$j=9$	256.3	134	186.9	58.6	27.6	23.4	6.6	3.3	11.2	0
$j=10$	31	526.6	364.7	114.3	53.9	45.7	12.9	6.4	21.9	0

9.3 Comparació entre models

Un cop vistos els resultats d'aquests cinc països pels tres mètodes estudiats: mètode de De Vylder (DV), de De Vylder-Mack (VM) i de "two way" (2W), ens preguntem dues coses: per una banda, quin d'aquests tres mètodes és el més apropiat?; i per altra banda, milloren els mètodes IBNR credibilístics els models d'IBNR clàssics? Per contestar a aquestes preguntes calia trobar un criteri per comparar els resultats. De fet, vam provar més d'un criteri de comparació. Finalment vam escollir-ne dos de diferents. El primer, utilitzat per Dannenburg, D. (1996) en la seva tesi doctoral, és la mitjana dels errors quadràtics:

$$MEQ = E\left[\left(X_{js} - X^*_{js}\right)^2\right] \tag{9.1}$$

Obviament, només trobarem els errors de les quanties de la part superior del triangle "run-off". A tal efecte, estimarem, amb el respectiu mètode, la part de dalt del triangle i calcularem la diferència amb les dades reals del triangle "run-off". En la següent taula, s'exposen els resultats d'aplicar aquest criteri als nostres resultats:

Taula 9.21 *Mitges dels errors al quadrat*

	B	NL	D	F	GB
DV	209.9	87.2	103.8	227.6	134
VM	211.1	90.5	110.9	238.9	157.6
2W	202.6	101.5	108.5	239	147.1
MQ	195.7	86.1	87.2	213.1	124.1
SA	189.9	109.3	121.3	183.7	152.4
SG	1104	176.1	890.9	236.5	618.6

A la primera columna, veiem les abreviatures dels mètodes utilitzats: els tres primers corresponen als models credibilístics, mentre que els altres són els següents models clàssics: el mètode de mínims quadrats de De Vylder (MQ), el de separació aritmètica (SA) i el de separació geomètrica (SG). A dalt de tot, veiem les abreviatures de les matrícules dels cinc països europeus analitzats.

A la vista dels anteriors resultats, podem veure que el millor mètode és el de mínims quadrats de De Vylder per als cinc països; resultat que no ens hauria de sorpendre, ja que el mètode mencionat consisteix precisament en trobar els resultats que minimitzin els errors quadràtics esperats. Sense comptar amb aquest mètode, podem veure que, en general, els mètodes de credibilitat tenen uns errors al quadrat menors que els clàssics. Amb això, ja podem destacar la millor aproximació dels nous mètodes envers els clàssics.

El que no sembla tan fàcil de respondre, és la primera pregunta que ens feiem al començament del capítol. Si només ens centrem en la part superior de la taula, que correspon als mètodes de credibilitat, observem que el mètode de De Vylder és el millor excepte per al primer país Bèlgica. Per tant no podem assegurar quin dels tres és el millor ja que segons quines siguin les dades, l'ajustament serà millor amb un mètode que amb un altre. De totes maneres, les diferències entre aquests tres mètodes són molt petites.

Amb només aquestes conclusions no ens podíem quedar, per tant, vam recòrrer a un altre criteri per escollir quin mètode s'ajusta més a les dades en cada cas. El mètode escollit, utilitzat per Goovaerts, M. et al. (1990), queda resumit en les taules que a continuació exposem per cada país. Partint, com abans, del triangle "run-off" original, eliminen la darrera diagonal, és a dir, les dades que corresponen al darrer període $j+s$. Amb les dades restants, estimem el triangle inferior, que inclou la diagonal eliminada, amb el mètode respectiu. D'aquesta manera, podem comparar les dades reals del darrer any $j+s$ amb l'estimació que hem fet d'aquest mateix any. En les següents taules, els resultats hi són plasmats de la següent manera. En primer lloc, tenim la primera fila que correspon a les dades reals de l'últim període conegut, són les dades que eliminem per recalcular la part inferior del triangle. Després d'aquests valors, les dues últimes columnes de la taula tenen el següent significat: la primera és la suma dels vuit valors estimats (o reals en el cas de la primera fila), i la segona, és la diferència entre l'anterior valor de les files posteriors i la suma de les dades reals, dividit per aquest últim valor. Aquest coeficient ens donarà una mesura de l'error comés en cada cas. Veiem-ho per als dos primers països (CL és l'abreviatura del metode Chain-ladder):

Taula 9.22 *Taula comparativa per Bèlgica*

real	1834	180.1	127.9	223.4	78.7	0	0	0	2445	0
DV	2365	696	206.9	149.5	42.6	36.4	0.4	0	3497	0.43
VM	2217	708.5	205.7	146.3	43	37	0.4	0	3358	0.37
2W	2324	705.6	203.1	144.7	42.4	34.5	0.4	0	3456	0.41
CL	2854	703.8	205	146.5	40.2	31.5	0.4	0	3981	0.63
MQ	2941	712	200.6	156.2	45.2	31	0.5	0	4086	0.67
SA	1995	720	202.4	137.7	35.5	34.7	0.3	0	3125	0.28
SG	3314	1173	218.8	198.1	7.6	1.5	0	0	4913	1.01

Taula 9.23 *Taula comparativa per als Països Baixos*

real	815.7	577.3	171.8	9.5	27.8	0.2	5.5	2.4	1610	0
DV	746.7	914	200.7	91.3	25.1	20.2	26.4	0.1	2025	0.26
VM	909	899	208	86	23.5	22.2	26.6	0.1	2175	0.35
2W	1053	966	220	75.5	18	22.2	28.3	0.1	2383	0.48
CL	1887	980.8	200.8	79.1	19.1	17.6	19.1	0	3204	0.99
MQ	1719	969.3	196.2	82.8	22	14.9	13.9	0	3018	0.87
SA	365.5	434.1	156.6	71.5	22.5	14.9	18.3	0	1083	−0.3
SG	352.5	384.3	102.6	54.1	17.1	12.7	10.6	0	934	−0.4

Per Bèlgica, el mètode que més s'ajusta és el de separació aritmètica, però sense tenir en compte aquest, els models de credibilitat donen una variació menor que la dels altres mètodes, a més, entre ells, el millor és el de De Vylder-Mack. De totes maneres, veurem que això no és igual per a tots els països estudiats.

Pels Països Baixos, en canvi, els resultats mostren que els mètodes de credibilitat són els millors, i entre ells, el de De Vylder presenta la variació més petita. Igual que abans, la diferència entre els models de credibilitat no és gran. Per tant, i com veurem per la resta de països, amb aquest criteri de comparació els models de credibilitat mostren la seva superioritat. A més, cal tenir en compte que en aquesta comparació ens fixem més en el darrer any, que és el que la companyia té més interés en conèixer.

Taula 9.24 *Taula comparativa per Alemanya*

real	238.8	225.5	0	0	1.9	0	0	0	466.3	0
DV	377	193.1	592.5	20.6	21.3	14.3	0	0	1218	1.61
VM	385.5	220.6	554.5	18.4	23.1	15.9	0	0	1198	1.57
2W	463.1	210.6	560	16.9	13.6	13.6	0	0	1278	1.74
CL	792.1	199	848.8	24.3	12	9.6	0	0	1886	3.04
MQ	747.3	141.3	761.1	19.1	13.2	8.9	0	0	1691	2.62
SA	217.4	164.7	326.6	23.8	58.1	7.5	0	0	798.2	0.71
SG	94.6	35.14	22.73	0	5.5	0	0	0	158	−0.7

Els coeficients de variació per aquest país són molt alts, la raó d'aquest problema són els valors reals per l'últim any, més de la meitat dels quals són iguals a zero. Si bé pot ser normal per als últims valors, el fet de que apareguin entre mig dispara molt l'error. De totes maneres, els models de credibilitat segueixen ajustant-se més que els clàssics. El mateix problema ens trobem en els següents països.

Taula 9.25 *Taula comparativa per França*

real	285.8	537.3	0	45	0	105	89.2	0	1062	0
DV	696	638.1	303.5	210	36.6	87.8	144.1	0.1	2116	0.99
VM	696	638	303.5	210	36.6	87.7	144.1	0.1	2116	0.99
2W	708.8	629.2	315.3	206.5	35.7	82.7	140	0.1	2118	0.99
CL	1115	537.9	307.6	162.8	30.4	73.6	141.7	0	2369	1.23
MQ	1188	553	291.3	140.3	23.6	57.3	151.1	0	2402	1.27
SA	713.1	660.1	330.2	242.7	54.5	136.7	181.8	0	2319	1.18
SG	679.5	602.8	248.8	211.3	34	70	220.1	90	2156	1.03

Taula 9.26 *Taula comparativa per al Regne Unit*

real	134	459.6	32.1	0	13.8	0	0	0	639.6	0
DV	283.5	511.1	53.6	61.7	46.1	20.2	31.5	190	1198	0.87
VM	275.1	472.4	51.5	57.8	57.9	22.4	29.3	177	1144	0.78
2W	279.7	492.5	52.3	58.2	53	22.5	31.3	177	1166	0.82
CL	376	625.6	64.6	66.1	41.7	20.5	33.1	232	1460	1.28
MQ	381.2	492.1	54.7	36.8	25.8	16.6	11.5	299	1318	1.06
SA	323.3	635.8	65.3	84.7	79	49.5	54.1	177	1438	1.25
SG	491.7	993	78.9	14.2	72.6	1.2	17.8	177	1846	1.89

En resum, si ens centrem amb les dades que eren més estables, les dels Països Baixos, veiem que els models de credibilitat s'ajusten més a les dades reals per l'últim any que els models clàssics. De totes maneres, s'ha pogut comprovar que és difícil classificar els mètodes IBNR.

Capítol 10

Conclusions

En aquest darrer capítol tenim l'objectiu d'exposar les conclusions a les que hem arribat durant la realització de la tesi doctoral. En primer lloc, les conclusions de la primera part d'aquest treball, Extensions de la Teoria de la Credibilitat, es refereixen a les primes de credibilitat amb recàrrec i als nous models de credibilitat de classificació creuada. I dins d'aquesta part, també s'inclou una aplicació sobre el ram de robatori de comerços a Espanya. En segon lloc, i per la segona part de la tesi doctoral, enumerem les conclusions d'aplicar els models de credibilitat al càlcul de reserves IBNR. Entre aquests models, hi figura el model IBNR "two way", que representa la conexió entre la primera i la segona part de la tesi. Com a la primera part, acabem aquesta segona part, amb una aplicació amb dades reals sobre el càlcul de reserves per una companyia que opera en diversos països europeus.

10.1 Primes de credibilitat amb recàrrec

En el primer capítol de la primera part, hem cregut convenient analitzar les primes de credibilitat incorporant-hi el recàrrec de seguretat que tota prima comercial ha de tenir. Es ben sabut que els models de credibilitat clàssics només es preocupen de donar una prima lineal ponderada entre la mitjana individual i la col·lectiva. Aquests models es basen en el criteri de l'esperança, o sigui, fixen la prima mitjançant el principi d'equivalència, és a dir, el pagament fixat que li correspon fer a l'assegurat ha de ser igual al valor esperat del pagament aleatori per sinistre de l'assegurador. Aquesta forma de definir la prima es coneix com a principi de prima pura i es justifica per les lleis dels grans nombres. Paradoxalment, aquest mètode entra en conflicte amb els resultats de la teoria de la ruïna, que afirmen que la prima pura, en general, no és suficient, més exactament: una prima que no excedeixi la prima pura resulta tècnicament ruïnosa amb probabilitat 1. Per tant, podem tenir en compte les consideracions de la teoria de la ruïna hem afegit o carregat una quantitat a la prima pura. Particularment, el percentatge sobre certa quantitat que sumarem a la prima pura es denomina recàrrec de seguretat i serà simbolitzat per α . Una de les quanties que podem escollir per afegir-hi un percentatge d'ella mateixa, és la variància del risc, que ens donarà una mesura de la variabilitat del risc. Per tant, el criteri de la variància consisteix en afegir a l'esperança un percentatge sobre la variància:

$$H(X) = E[X] + \alpha \cdot V[X] \tag{10.1}$$

A continuació, ens plantegem afegir la variància col·lectiva per a tots els riscos o bé la variància individual de cadascun, o com ens proposa la Teoria de la Credibilitat, una mitjana ponderada entre totes dues variàncies.

Aquesta darrera línia, va ser seguida per Bühlmann, H. (1970) i Centeno, L.

(1989). El primer, partint del criteri de la variància:

$$E[\mu(\Theta) \,/\, \vec{X}_j] + \alpha \left(E[\sigma^2(\Theta) \,/\, \vec{X}_j] + Var[\mu(\Theta) \,/\, \vec{X}_j]\right) \tag{10.2}$$

estableix una fórmula de credibilitat amb tres components:

$$P_{t+1}\left(\vec{X}_j\right) = \underbrace{E\left[\mu(\Theta_j) \,/\, \vec{X}_j\right]}_{(a)} + \underbrace{\alpha \cdot E\left[\sigma^2(\Theta_j) \,/\, \vec{X}_j\right]}_{(b)} + \underbrace{\alpha \cdot Var\left[\mu(\Theta_j) \,/\, \vec{X}_j\right]}_{(c)} \tag{10.3}$$

on (a) és la part del valor esperat, (b) la part de variància i (c) la part de fluctuació. Bühlmann, va proposar estimadors de credibilitat per a cada part. Aquestes expressions són denominades, en el treball de Bühlmann, fórmules de credibilitat i substitueixen les primes de credibilitat dels models de credibilitat clàssics. Aquestes fórmules de credibilitat prenen el lloc a les primes pures de credibilitat. Pel que fa a les dues primeres parts, Bühlmann les construeix a partir de linealitzacions de les quantitats teòriques, procurant que la desviació quadràtica esperada romangui com més petita possible, és a dir, de la mateixa manera que en els models clàssics de la Teoria de la Credibilitat.

En resum, segons l'article de Bühlmann, tenim la següent aproximació a la prima de credibilitat:

$$\begin{aligned} P_{t+1}\left(\vec{X}_j\right) &\sim b \cdot \bar{X}_j + (1-b) \cdot E\left[\mu(\Theta_j)\right] \\ &+ \alpha\left(c \cdot \Sigma_j^2 + (1-c) \cdot E\left[\sigma^2(\Theta_j)\right]\right) \\ &+ \alpha\left((1-b) \cdot Var[\mu(\Theta_j)]\right) \end{aligned} \tag{10.4}$$

on $\bar{X}_j$ és la mitjana individual, $E\left[\mu(\Theta_j)\right]$ la col·lectiva i b el coeficient de credibilitat que les pondera. A més, Σ_j^2 és la variància individual de cada risc, $E\left[\sigma^2(\Theta_j)\right]$ la col·lectiva i c el corresponent factor de credibilitat.

Aquest model, basat en el model clàssic de Bühlmann, presenta un greu inconvenient, que és paral·lel al que ja presentava el model sense recàrrec. Suposa que totes les pòlisses tenen la mateixa importància dins de la cartera, independentment del seu volum. Per aquest motiu, Centeno, L. (1989) construeix un model de credibilitat amb recàrrec basant-se en el model de Bühlmann-Straub, que va aparèixer en el seu moment per superar la limitació que suposava el model de Bühlmann. Així doncs, Centeno incorpora els pesos w_{js} per a cada pòlissa. Seguint el mateix plantejament que Bühlmann, basant-se en les hipòtesis del model de Bühlmann-Straub, i després de determinar les tres components de la prima de credibilitat amb recàrrec, ens dóna la següent expressió per la prima del període $t_j + 1$:

$$\begin{aligned} P_{t_j+1}\left(\vec{X}_j\right) &= z_j \cdot \bar{X}_j + (1 - z_j) \cdot E[\mu(\Theta_j)] + \\ &+ \frac{\alpha}{w_{j,t_j+1}} c_j \cdot \Sigma_1^2 + (1 - c_j) \cdot E[\sigma^2(\Theta_j)] + \\ &+ \alpha \cdot (1 - z_j) \cdot Var\left[\mu(\Theta_j)\right] \end{aligned} \tag{10.5}$$

D'aquesta manera, s'elimina l'inconvenient anteriorment esmentat.

Tot i així, aquests dos models, resumits en el segon, tenen una sèrie d'inconvenients o problemes que calen ser tinguts en compte, i que a continuació enumerem:

1. A l'estimació de la part de fluctuació, aproximem per la mitjana de les pòlisses, perdent la informació individual de la variància.

2. A l'estimació de la part de la variància, apliquem l'assumpció de normalitat seguint la llei dels grans nombres.

3. L'estimador del coeficient de credibilitat (c_j) per a la part de la variància, s'obté amb unes dades no prou justificades. A més, tot i que és definit, en principi, per a cada j a la pràctica és un únic valor.

Per altra banda, els càlculs per trobar els estimadors de credibilitat són molt llargs i complexos.

Per millorar els anteriors models, Bauwelinckx, T et al. (1991) presenten un model totalment diferent als dos anteriors. Amb el mateix objectiu, trobar una prima de credibilitat amb recàrrec que contingui una fracció de la variància del risc com a recàrrec sobre la prima pura, segueixen un procediment molt més senzill i amb menys hipòtesis restrictives que l'anterior model de Centeno. A partir de la prima d'Esscher:

$$H(Y\,/\,X) \;=\; \frac{E\left[Y \cdot e^{hY}\,/\,X\right]}{E\left[e^{hY}\,/\,X\right]} \tag{10.6}$$

es construeix el nou model.

En primer lloc, es demostra que la prima d'Esscher, per a un recàrrec de seguretat (aquí denominat h) petit, s'aproxima al criteri de la variància:

$$H(Y\,/\,X) \;=\; E[Y\,/\,X] \,+\, h \cdot Var[Y\,/\,X] \,+\, O\left(h^2\right) \tag{10.7}$$

Normalment, el recàrrec de seguretat és inferior a la unitat, per tant, el seu valor al quadrat és un número més petit i, per tant, en podem prescindir. Per això, es desprècia $O(h^2)$. Així doncs, a partir de la prima d'Esscher, obtenim:

$$H(X_{t+1}\,/\,\vec{X}) \;=\; \frac{E\left[X_{t+1} \cdot e^{hX_{t+1}}\,/\,\vec{X}\right]}{E\left[e^{hX_{t+1}}\,/\,\vec{X}\right]} \tag{10.8}$$

Aquesta prima amb recàrrec la podem escriure, d'acord amb la nomenclatura de la credibilitat, com:

$$H(X_{t+1}\,/\,\vec{X}) \;=\; E[\mu(\Theta)\,/\,\vec{X}] \,+\, h \cdot \left(E[\sigma^2(\Theta)\,/\,\vec{X}] \,+\, Var[\mu(\Theta)\,/\,\vec{X}]\right) \,+\, O(h^2) \tag{10.9}$$

A partir d'aquí, evitant els problemes del model de Centeno, trobem els resultats de credibilitat a partir del desenvolupament en termes de h del numerador i el denominador de (10.8). A més d'evitar les aproximacions dels altres treballs, el nou model pot ser fàcilment estés al model jeràrquic de credibilitat de Jewell. Trobant l'estimador de credibilitat per a cada part de (10.8), tindrem una estimació que inclourà les tres parts dels models anteriors com podem veure a (10.9).

Amb totes aquestes noves eines, i fent servir les propietats dels estimadors de credibilitat semilineals, sota les hipòtesis del model de Bühlmann-Straub i sota l'assumpció de l'abundància de pesos naturals (w_{js}), la següent expressió és l'aproximació a la prima d'Esscher $H\left(X_{t+1}\,/\,\vec{X}\right)$ d'ordre h per aconseguir la prima de credibilitat amb recàrrec,

$$
\begin{aligned}
&(1 - z_j)\hat{m} + z_j X_{jw} + h\Bigg[(1 - z_j) X_{zw}^2 + z_j X_{jw}^2 + \\
+ &\left\{ \frac{2w_{\Sigma\Sigma}^2}{w_{*\Sigma}^2 - \sum_{j=1}^{k} w_{j\Sigma}^2} \left(Cov_B\left(X_{js}, X_{js}^2\right) - \frac{k-1}{k(t-1)} Cov_W\left(X_{js}, X_{js}^2\right)\right) \left(\frac{z_j}{\hat{a}}\right) - \right. \\
- &\frac{2z_j^2}{\hat{a} w_{j\Sigma}} \cdot \frac{w_{\Sigma\Sigma}}{k(t-1)} Cov_W\left(X_{js}, X_{js}^2\right) - 2\left(\frac{z_j^2}{\hat{a}}\right) \frac{w_{\Sigma\Sigma}^2}{w_{\Sigma\Sigma}^2 - \sum_{j=1}^{k} w_{j\Sigma}^2} \times \\
\times &\left.\left(Cov_B\left(X_{js}, X_{js}^2\right) - \frac{k-1}{k(t-1)} Cov_W\left(X_{js}, X_{js}^2\right)\right) \right\} \cdot (X_{jw} - \hat{m}) - \\
- &\left((1 - z_j)\hat{m} + z_j X_{jw}\right)^2 \Bigg]
\end{aligned}
\tag{10.10}
$$

Tot i lextensió de l'expressió, els càlculs per trobar la prima de credibilitat amb recàrrec són senzills.

Si bé des d'un punt de vista teòric sembla que aquest últim mètode millora els dos anteriors, fins ara no s'havia comprovat empíricament. Per aquest motiu hem programat amb APL tots tres models i n'hem comparat els resultats en una aplicació pràctica a l'apartat 2.5. A continuació, exposem alguns dels resultats que ens ajudaran a treure les principals conclusions. Els tres primers resultats són les mitjanes $(\bar{X}_j)$ i les variàncies individuals (Σ_j^2) per a cada pòlissa j . Els resultats de la part inferior de la taula, mostren els tants per cent d'augment sobre la prima pura que s'experimenta per la incorporació del recàrrec per a cada pòlissa i per als tres models analitzats: Bühlmann (1), Centeno (2) i Bauwenlickx et al. (3).

$\bar{X}_j$ (sense pesos)	19.21	10.94	19.81	11.14	15.18
$\bar{X}_j$ (amb pesos)	19.21	10.94	19.99	11.03	15.63
Σ_j^2	1.19	0.74	17.43	24.81	12.41
Model 1	15.24	24.99	18.02	32.52	21.68
Model 2	17.38	17.02	22.60	41.94	23.36
Model 3	9.89	5.52	24.43	44.90	14.94

Tot seguit, enumerem les conclusions d'aquesta aplicació:

1. Les cinc pòlisses tenen variàncies individuals molt diferents, aquesta característica ens permetrà veure quin dels tres models proporciona una prima de credibilitat amb recàrrec més acurada. A nosaltres ens interessa que, un cop trobada la prima pura de credibilitat per als tres models, el recàrrec aplicat vagi amb consonància amb la variància individual de cada pòlissa. És a dir, diferenciar les primes no només per la mitjana individual, sinó també per la variància de cada pòlissa.

2. El model de Bühlmann s'ajusta a les dades empíriques meys que el model de Centeno, en la mesura que el segon utilitza la informació del nombre de sinistres com a ponderació a l'hora de calcular la prima pura de credibilitat.

3. A més, el model de Bühlmann té un coeficient de credibilitat per la part de la variància $c = 0.1522$ quasi insignificant. Aquest fet fa que la variància col·lectiva sigui més important a l'hora de determinar la prima de credibilitat per la part de la variància. Per tant, la diferenciació que busquem per part de la variància quasi no existirà com podem veure en el quadre anterior, on es mostra que el tant per cent d'augment sobre la prima pura que s'experimenta per la incorporació del recàrrec és molt similar per a totes les pòlisses, quan d'acord amb les variàncies individuals hauria de ser més diferent. En canvi, en el model

de Centeno el mateix coeficient de credibilitat és més gran $c = 0.5516$. Per tant, la variància individual tindrà més del cinquanta per cent d'importància en la prima de credibilitat per la part de la variància. Això farà que el percentatge del recàrrec sigui més diferenciat per a cada pòlissa, tal i com nosaltres volem.

4. Per tant, sembla clar que el model de Centeno millora els resultats del model de Bühlmann. Ens queda per veure si el nou model, basat en la prima d'Esscher, ens dóna uns resultats més acurats que el de Centeno. Pel que fa a la part de la prima pura, tots dos tenen els mateixos resultats, ja que tots dos utilitzen la prima pura de credibilitat del model de Bühlmann-Straub.

5. Ens resta comprovar quin dels dos models proporciona un recàrrec de seguretat més addient segons les variàncies de cada pòlissa. En aquest cas, no podem comparar els coeficients de credibilitat per la part de la variància ja que el nou model no ens dóna aquesta informació. Per tant, ens haurem de fixar en el percentatge d'augment sobre la prima pura que s'experimenta per la incorporació del recàrrec de seguretat basat en la variància. Fàcilment, veiem com en aquest nou model la diferenciació és molt més acusada que en el model de Centeno.

6. En resum, i tal com podíem preveure per les millores teòriques introduïdes, el model de Bauwenlickx et al. ens proporciona une primes de credibilitat amb recàrrec més acurades des del punt de vista de la diferenciació per la variància individual.

10.2 Models de classificació creuada

No podíem deixar la primera part sense resoldre un dels problemes més importants amb que es troba la Teoria de la Credibilitat en aquests moments: la seva aplicabilitat al món actuarial. Per una banda, els models clàssics són difícilment interpretables en la majoria dels casos, la interpretació d'un model i de les seves respectives primes són, a més, un important criteri perquè un actuari decideixi quin model ha de fer servir. Per altra banda, en moltes ocasions les dades empíriques disponibles no s'adiuen als models clàssics, per exemple, en els models jeràrquics suposem que els diferents factors que hi intervenen estan ordenats jeràrquicament, quan a la realitat no sempre és així.

Aquests dos inconvenients poden ser superats mitjançant l'aplicació dels models de classificació creuada apareguts recentment. En efecte, aquests models suposen, per una part, la creació de models de credibilitat amb noves relacions entre els factors de risc, no necessàriament jeràrquiques, i a més, la generalització de tots els models de credibilitat, els clàssics i els nous, en una única nomenclatura i forma d'interpretació dels resultats; aquesta última molt més clara que en els models clàssics. En el capítol 3 recollim, en primer lloc, el model general de classificació creuada. Una versió del model general per P factors la podem presentar a partir de les variables aleatòries independents $\Xi^{(i_1 \dots i_q)}_{j_{i_1} \dots j_{i_q} s}$, cadascuna amb mitjana zero i variància $b^{(i_1 \dots i_q)}$. D'aquesta manera, podem escriure el risc $X_{j_1 \dots j_P s}$ com:

$$X_{j_1 \dots j_P s} = m + \sum_{q=1}^{P} \left(\sum_{i_1=1}^{P-q+1} \cdots \sum_{i_q=i_{q-1}+1}^{P} \Xi^{(i_1 \dots i_q)}_{j_{i_1} \dots j_{i_q} s} \right) + \Xi^{(1 \dots P, P+1)}_{j_1 \dots j_P s} \qquad (10.11)$$

amb la variància de l'últim terme igual a $S^2 / w_{j_1 \dots j_P s}$.

En aquest treball només ens hem centrat en els models que poden desenvolupar-se a partir de dos factors. Del model general mem deduït dos models en concret, el model "two way" i el model additiu amb dos factors. A més, també podrem deduir-ne els clàssics de Bühlmann-Straub i el jeràrquic de dos nivells de Jewell. En primer lloc, i com a model amb dos factors més ampli, analitzem el model "two way". A partir de l'expressió (10.11), el podem enunciar de la manera següent:

$$X_{ijs} = m + \Xi^{(1)}_i + \Xi^{(2)}_j + \Xi^{(12)}_{ij} + \Xi^{(123)}_{ijs} \qquad (10.12)$$

Les variables aleatòries $\Xi_i^{(1)}$ i $\Xi_j^{(2)}$ denoten les principals característiques corresponents al primer i segon factor respectivament. $\Xi_{ij}^{(12)}$ representa la interacció entre els dos factors, i per últim, $\Xi_{ijs}^{(123)}$ denota la desviació sobre la mitjana, donats $\Xi_i^{(1)}$, $\Xi_j^{(2)}$ i $\Xi_{ij}^{(12)}$ en un període d'observació s . Suposem que el primer factor té I categories, mentre que el segon en té J. D'acord amb les hipòtesis dels models de components de variància, es dedueix la següent estructura de covariàncies:

$$Cov\,[X_{iju}\,,\,X_{kls}] \;=\; \delta_{i,k}b^{(1)} + \delta_{j,l}b^{(2)} + \delta_{ij,kl}b^{(12)} + \delta_{iju,kls}S^2\,/\,w_{kls}\;; \tag{10.13}$$

on $Var\left[\Xi_i^{(1)}\right] = b^{(1)}$; $Var\left[\Xi_j^{(2)}\right] = b^{(2)}$; $Var\left[\Xi_{ij}^{(12)}\right] = b^{(12)}$ i $Var\left[\Xi_{ijs}^{(123)}\right] = S^2\,/\,w_{j_1 j_2 s}$. En la secció 3.3.1 definim l'estimador de credibilitat per aquest model, així com els estimadors dels paràmetres estructurals necessaris per fer operacional aquest model.

A partir d'aquest, eliminant alguna de les seves components, podem definir la resta de models possibles amb dos factors. En primer lloc, eliminant la component mixta $\Xi_{ij}^{(12)}$, definim el model additiu:

$$X_{ijs} \;=\; m + \Xi_i^{(1)} + \Xi_j^{(2)} + \Xi_{ijs}^{(123)} \tag{10.14}$$

Les variables aleatòries $\Xi_i^{(1)}$ i $\Xi_j^{(2)}$ denoten, com abans, les principals característiques corresponents al primer i segon factor respectivament. De la mateixa manera que en l'anterior model es defineixen l'estimador de credibilitat i els estimadors dels paràmetres. Aquests dos models són la nova aportació dels models de classificació creuada a la Teoria de la Credibilitat. Però com hem avançat abans, els models clàssics també poden ser derivats a partir del model "two way".

Per exemple, el model de Bühlmann-Straub pot ser expressat de la manera següent:

$$X_{ijs} \;=\; m + \Xi_{ij}^{(12)} + \Xi_{ijs}^{(123)} \tag{10.15}$$

Les caracterítiques del risc, en aquest cas, venen explicades únicament per la variable aleatòria $\Xi_{ij}^{(12)}$ que representa la interacció entre els dos factors. Comparat amb el model "two way", hi manquen les dues variables aleatòries $\Xi_i^{(1)}$ i $\Xi_j^{(2)}$ que denotaven les característiques de cada factor independentment de l'altre. Un cop trobat l'estimador de credibilitat, es pot veure que coincideix amb l'obtingut en el model clàssic.

El model jeràrquic de dos nivells de Jewell pot ser definit de la següent manera a partir dels models de components de variància:

$$X_{ijs} = m + \Xi_i^{(1)} + \Xi_{ij}^{(12)} + \Xi_{ijs}^{(123)} \tag{10.16}$$

La variable aleatòria $\Xi_i^{(1)}$ denota les principals característiques corresponents al primer factor de risc. $\Xi_{ij}^{(12)}$ representa la interacció entre aquest primer factor i el segon factor, de manera que el segon factor té una dependència jeràrquica amb el primer. Fixem-nos que, a diferència del model "two way", la variable aleatòria $\Xi_j^{(2)}$ que denotava les característiques del segon factor exclusivament, no hi figura. De la mateixa manera podem fer que el nivell principal sigui per al segon factor i que el primer depengui jeràrquicament d'aquest:

$$X_{ijs} = m + \Xi_j^{(2)} + \Xi_{ij}^{(12)} + \Xi_{ijs}^{(123)} \tag{10.17}$$

En resum, amb aquesta nova manera d'enfocar la Teoria de la Credibilitat tenim dues possibilitats més de calcular les primes de credibilitat, i aquests dos nous models són comparables amb els models clàssics que, alhora, són més fàcilment interpretats.

Per posar punt final a aquest capítol, hem cregut convenient afegir-hi una aplicació pràctica per comprovar l'aplicabilitat del nou enfoc i la millora en el tractament de dades empíriques. Com a exemple per mostrar l'aplicabilitat dels models de credibilitat de classificació creuada, hem escollit unes dades oferides per una companyia belga d'assegurances sobre crèdits privats. La variable X_{ijs} analitzada fa referència a cada un dels pagaments, que la empresa asseguradora va haver de fer a diferents bancs, per cobrir les pèrdues causades pels clients d'aquests bancs que no van poder fer front o cancel·lar els seus prèstecs privats. Les observacions han estat dividides o catalogades d'acord amb la situació civil dels clients, i el temps que porten treballant en la seva feina o lloc de treball habitual. Així doncs, el primer factor de risc té a veure amb l'estat civil de la persona, i el segon factor amb l'antiguitat a la seva feina. Cada un d'aquests factors s'ha dividit amb tres categories diferents, $I = J = 3$ segons el model "two way" vist a la secció 3.3. L'estat civil del client pot ser solter, divorciat o separat, i la resta dels casos, és a dir, casats, ajuntats o vidus. Pel que fa referència al temps que porta a la seva feina, també hi han tres classes diferents: menys de dos anys, entre dos i deu anys, i més de deu anys.

A partir de les dades que tenim, les hem aplicat a cadascun dels quatre models que acabem de definir i n'hem tret les següents conclusions:

1. Els resultats obtinguts mitjançant el model "two way" són acceptables, $b^{(1)}$ i $b^{(2)}$ són prou elevats com per considerar que les seves respectives variables són significatives a l'hora d'explicar la composició dels estimadors de credibilitat. Paral·lelament, els coeficients de credibilitat $z_i^{(1)}$ i $z_j^{(2)}$ són elevats, i per tant més credibilitat és donada a l'experiència de les variables corresponent al primer i segon factor separadament. En canvi, $b^{(12)}$ és un valor molt petit, això fa que els seus respectius $z_{ij}^{(12)}$ siguin gairabé insignificants i, per tant, la informació que la variable mixta $\Xi_{ij}^{(12)}$ ens aporta és molt petita. Aquest fet, indica que el model utilitzat potser hauria de ser reduït al model additiu.

2. Els resultats obtinguts en el model additiu són molt semblants a l'anterior degut a la poca importància que es donava a $\Xi_{ij}^{(12)}$. En efecte, eliminant aquesta component mixta aconseguim uns valors de $b^{(1)}$ i $b^{(2)}$ una mica més elevats que en l'anterior model, és a dir, la petita informació que ens proporcionava la variable mixta queda repartida en cadascuna de les variables de cada factor. Per tant, els coeficients de credibilitat $z_i^{(1)}$ i $z_j^{(2)}$ són una mica més grans que abans. Així doncs, per calcular la prima de credibilitat per una pòlissa, únicament hem de sumar o restar la component referida a l'antiguitat

i la component referida a l'estat civil. Des del nostre punt de vista, creiem que aquest és el model més apropiat per aquestes dades, ja que els estimadors de credibilitat de les components $\Xi_i^{(1)}$ i $\Xi_j^{(2)}$ tenen un sustancial efecte a la prima de credibilitat a pagar pels nous clients.

3. Tot i l'anterior afirmació, també volíem aplicar les dades a la resta de models. Així doncs, si apliquem el model de Bühlmann-Straub, és a dir, eliminem les components $\Xi_i^{(1)}$ i $\Xi_j^{(2)}$, la prima de credibilitat només ens vindrà explicada per la component mixta. Evidentment això significa un pas enrere, escollint aquest model prescindim de les components que segons els models anteriors eren més significatives, i n'escollim la menys significativa per explicar tot el comportament de la variable X_{ijs} . La informació que recollirem serà veure cóm afecten conjuntament els dos factors de risc, estat civil i antiguitat. Cal senyalar que els resultats obtinguts són essencialment els mateixos que aplicant el model clàssic. Fins i tot, $b^{(12)}$ coincideix amb el paràmetre a del model clàssic.

4. Quan hem aplicat el model "two way" hem vist com les components $\Xi_i^{(1)}$ i $\Xi_j^{(2)}$ tenen un sustancial efecte a la prima de credibilitat a pagar per als nous clients, cap d'aquestes components pot ser eliminada com s'hauria de fer si apliquèssim el model de Jewell. Si apliquem el model de Jewell amb la següent estructura:

$$X_{ijs} = m + \Xi_i^{(1)} + \Xi_{ij}^{(12)} + \Xi_{ijs}^{(123)} \tag{10.18}$$

hem obtingut que la variància $b^{(1)}$ és nul·la, tal resultat significa que tot s'explicarà per la component mixta, igual que en el model de Bühlmann-Straub. Per altra banda, el valor nul de $b^{(1)}$ indica que el model no pot aportar res del factor estat civil sense anar acompanyat del factor antiguitat. Per aquesta raó ho fa conjuntament amb $\Xi_{ij}^{(12)}$. Igualment, si provem amb el segon factor com a principal nivell, això és:

$$X_{ijs} = m + \Xi_j^{(2)} + \Xi_{ij}^{(12)} + \Xi_{ijs}^{(123)} \tag{10.19}$$

també veurem que la variància $b^{(2)}$ és nul·la. Per tant, les mateixes conclusions es poden treure. A la vista dels resultats obtinguts, podem veure que el model jeràrquic no és un bon model per aquestes dades, dit d'una altra manera, els dos factors que intervenen no poden ser ordenats jeràrquicament. A més, amb el model Bühlmann-Straub la informació recollida és molt general. Per tant, aquestes dades ens permeten afirmar que els nous models de credibilitat s'ajusten molt més que els clàssics.

5. En resum, d'aquesta aplicació podem treure unes conclusions més generals que només les particulars de l'aplicació de cada model. Aquestes són les següents:

   (a) Tradicionalment s'ha vingut aplicant la Teoria de la Credibilitat com un mètode de tarificació *a posteriori* . En aquesta aplicació pràctica, on les dades de cada cel·la no representen dades d'un període determinat de temps, sinó un conjunt de pòlisses que tenen les mateixes característiques $(i\,,\,j)$, ens adonem que també pot ser aplicada a la tarificació *a priori* .

   (b) Un dels problemes més difícils amb que es troba la Teoria de la Credibilitat és l'estimació dels paràmetres estructurals. En els nous models de classificació creuada s'escullen arbitràriament uns estimadors que no tenen perquè ser els òptims, la qüestió no ha estat encara derivada, així utilitzem un dels molts estimadors no esbiaixats. De totes maneres, en el model clàssic de Bühlmann-Straub sí que s'havia demostrat quin estimador era l'òptim i si tenim en compte que l'estimador de $b^{(12)}$ és igual a l'obtingut en el clàssic (a) podem entreveure que el mètode aplicat per calcular els estimadors de les variàncies $b^{(1)}$, $b^{(2)}$ i $b^{(12)}$ no és dolent.

(c) Per últim, destacar que cada conjunt de dades pot ajustar-se millor a un model que a un altre. Amb això, ens venim a referir que no hi ha cap model millor que un altre, sinó que el model escollit té molt a veure amb l'adeqüació de cada conjunt de dades.

En el capítol 4 hem volgut fer una altra aplicació, aquesta sobre unes dades espanyoles, amb els models de credibilitat de que disposem, intentant buscar quin dels models proposats és el que més s'ajusta a les dades proposades. En aquest capítol hem aplicat els models credibilístics de classificació creuada al ram de robatori a comerços. Aquest és un ram on la diferenciació de pòlisses segons el risc que assumeixen és molt important, per aquest motiu, la Teoria de la Credibilitat pot representar una bona ajuda a l'hora de calcular les primes de les diferents pòlisses segons la seva pròpia experiència de sinistralitat. En aquest cas, hem escollit dos factors que creiem són els més importants per diferenciar unes pòlisses d'altres. El primer factor que hem considerat és el capital assegurat per a cada pòlissa, i el segon el tipus de comerç. El primer factor ha estat dividit en tres categories diferents, per tant $I = 3$, on la primera ($i = 1$) està formada per aquelles pòlisses que tenen un capital assegurat de menys de cinc milions de pessetes; la segona categoria ($i = 2$) és per aquells comerços que tenen un capital assegurat entre cinc i deu milions de pessetes; i, per últim, aquelles pòlisses amb el capital assegurat superior als deu milions de pessetes formen la tercera categoria ($i = 3$) del primer factor. Pel que fa al segon factor, hem escollit dotze tipus de comerços diferents ($J = 12$).

A continuació enumerem aquests dotze tipus,

$j = 1$	Restaurants, bars i cafeteries
$j = 2$	Bijuteries
$j = 3$	Moda i confecció
$j = 4$	Queviures
$j = 5$	Art
$j = 6$	Esports
$j = 7$	Electrodomèstics
$j = 8$	Ordinadors i material d'oficines
$j = 9$	So, vídeo i electrònica
$j = 10$	Supermercats
$j = 11$	Transports (agències)
$j = 12$	Viatges (agències)

Les dades que hem fet servir han estat recollides de les publicacions que l'associació "Investigación cooperativa entre entidades aseguradoras" (I.C.E.A.) va fer durant els anys d'observació de 1990 a 1994. Aquesta associació va fer aquests estudis a partir de les dades oferides per una vintena de companyies asseguradores, que treballen en aquest ram a Espanya. Aquests informes, sota el títol genèric de "La siniestralidad de robo en comercios", són una recopilació estadística del que ha estat aquest ram d'assegurances en aquests cincs anys a Espanya. Hem cregut interessant definir X_{ijs} com el cost mig de sinistralitat per a cada tipus de comerç segons el tram de capital assegurat. A més, en aquest ram, tant important és el cost mig per sinistre com el nombre de sinistres que hi han hagut per cada any d'observació. Per això, hem cregut convenient incorporar aquesta informació via pesos o ponderacions naturals per a la nostra aplicació. A continuació, enumerem els passos seguits per mirar de trobar quin model és el més apropiat per aquest conjunt de dades:

1. En primer lloc, hem aplicat el model més complet, amb totes les components

possibles, el model "two way". Un cop calculats els resultats, el primer que veiem és que les variàncies $b^{(1)}$ i $b^{(2)}$ són nul·les. Per tant, el model "two way" no sembla ser el que més s'ajusta a les dades, a més, opta per explicar-nos tota la variància de les dades per la component mixta dels dos factors, igual que faríem aplicant el model de Bühlmann-Straub. En resum, els resultats d'aplicar les dades del comerç espanyol al model "two way" no semblen gaire bons. Això ens porta a provar altres models com a solució.

2. Si el nostre interés es centra a conèixer l'efecte sobre la prima dels dos factors per separat, és a dir, eliminant la component mixta $\Xi_{ij}^{(12)}$, podem calcular les primes pel proper any utilitzant el model additiu. En aquest cas, veiem que les variàncies $b^{(1)}$ i $b^{(2)}$ ja no són nul·les, tot el contrari, totes dues són prou grans com per creure en la bondat d'aquest model aplicat sobre les nostres dades. Un cop eliminada la component mixta, hem forçat al nou model (additiu) a explicar la composició de la prima mitjançant les components dels dos factors per separat. Com hem dit anteriorment, aquest pot ser un bon model si el que volem és diferenciar les primes per aquests dos factors per separat i amb la mateixa importància. Recordem però, que en el model "two way" la component mixta era molt important, aquest resultat pot indicar-nos que existeix alguna relació entre els dos factors i que, per tant, la component mixta potser no ha de ser eliminada.

3. A continuació, apliquem les dades amb el model del Bühlmann-Straub per calcular les primes de credibilitat. El resultat, com podíem pensar, no difereix del trobat aplicant el model "two way" ja que tota la composició de les primes ve explicada per la component mixta. Arribats aquest punt, podem plantejar-nos recurrir als models jeràrquics.

4. En efecte, el fet que en el model "two way" les components per separat no hagin

estat utilitzades en la composició de la prima no vol dir que no es pugui provar combinar els mateixos factors d'una altra manera, per exemple, jerarquitzar els factors, és a dir, donar més importància a un factor que a l'altre, contràriament del que feiem a l'aplicar el model "two way". Si apliquem el model de Jewell amb la següent estructura:

$$X_{ijs} = m + \Xi_i^{(1)} + \Xi_{ij}^{(12)} + \Xi_{ijs}^{(123)} \tag{10.20}$$

és a dir, posant el primer factor (tram de capital assegurat) com a nivell principal i fent dependre el tipus de comerç del capital assegurat, veiem que la variància $b^{(12)}$ és nul·la. Aquest resultat no ens aporta res d'interessant al nostre estudi, ja que eliminem per complet el segon factor. En canvi, si provem amb el segon factor com a principal nivell, això és, escollir el segon factor (tipus de comerç) com el pricipal factor a tenir en compte, deixant reduïda a la component mixta la participació del primer factor:

$$X_{ijs} = m + \Xi_j^{(2)} + \Xi_{ij}^{(12)} + \Xi_{ijs}^{(123)} \tag{10.21}$$

L'elevat valor de les variàncies $b^{(2)}$ i $b^{(12)}$ donen validesa al model jeràrquic de dos nivells que hem escollit. Per tant, és bona la idea de diferenciar, en primer lloc, per a cada tipus de comerç i dins d'un tipus de comerç en concret subdividir pels tres trams de capital assegurat. Així doncs, en aquest model igual que en el model additiu, tenim en compte els dos factors com nosaltres volíem. La diferència es troba en el tractament que es fa del primer factor, en el model additiu aquest factor té el mateix nivell o importància que el segon. En canvi, en aquest model jeràrquic, el primer factor depén, o està en un nivell inferior, del segon factor.

5. En resum, podem concluir que des del nostre punt de vista aquest últim model pot ser el més apropiat amb aquestes dades. De totes maneres, cal senyalar que

només disposem de cinc anys, quatre en algunes cel·les, d'experiència. Aquest fet, segons la variabilitat de les dades pot tindre efectes negatius a l'hora d'estimar les primes de credibilitat pel proper any. Creiem que amb uns anys més d'experiència els resultats que obtindríem podrien ajustar-se molt millor als models plantejats. En aquesta aplicació, tornem a tindre observacions per a cada any en cada cel·la, per tant, el resultat obtingut té més a veure amb un criteri de tarificació *a posteriori* .

10.3 Models IBNR de credibilitat

La segona part de la tesi comença amb el capítol 5, amb una introducció al món del càlcul de reserves IBNR. Com ja hem anunciat, a la fi de l'any comptable, en una companyia d'assegurances, normalment, resten alguns sinistres per pagar, i és per aquests que es creen les reserves per sinistres. A més, per regla general, no totes les quanties dels sinistres són conegudes al final de l'any comptable, i per això, també ha de ser calculada una reserva. Ens referim a les reserves per sinistres pendents de liquidació i/o pagament i a les reserves per sinistres pendents de declaració. En una companyia d'assegurances, l'habilitat per estimar les seves reserves correctament és d'una gran importància. Un dels propòsits d'aquesta estimació és el d'assegurar una correcta imatge del passiu de la companyia en el balanç comptable. Per aquesta raó, en els papers per a la presentació de l'estat comptable anual, i en les fitxes tècniques de les primes, es dóna tanta atenció a l'estimació de les reserves per sinistres.

Un fet comú, com a punt inicial, en tots els models per calcular les reserves per sinistres és el denominat triangle "run-off". Aquest triangle, és la manera més natural de mostrar les dades dels pagaments futurs a partir de cada any d'origen, i s'anomena "triangle" perquè aquesta és la forma en què apareixen les dades, a causa de que com més llunyà estigui l'any d'origen, més gran serà el seu historial. En aquest primer capítol d'introducció, expliquem com es forma aquest triangle de dades.

Definim X_{js} com el risc a tenir en consideració corresponent al període d'origen j ($j = 1, 2, ..., k$) i en el seu s -èsim ($s = 1, 2, ..., t$) període de desenvolupament. Suposarem, com és habitual, que tenim el mateix nombre de períodes d'origen i de desenvolupament $k = t = J$. Després de J períodes, coneixem les observacions de la variable aleatòria X_{js} per $j + s \leq J + 1$. Basant-nos amb aquestes dades conegudes, volem prediure els futurs valors per X_{js} , o sigui, per $j + s \geq J + 2$. Per exemple, considerem una cartera de pòlisses per incendi, i X_{js} les quanties pagades per aquest motiu. Si els períodes són anuals i comencem al 1991, per $j = 1$, cada risc X_{js} és igual a la suma de quanties pagades per sinistres de pòlisses contractades a l'any d'origen j , però saldades a l'any de desenvolupament s . Per exemple, X_{24} és la quantia total dels pagaments fets al 1995 corresponents a les pòlisses que varen ser venudes l'any 1992. Gràficament aquesta composició es pot veure en la següent taula:

	1	2	3	$\dots$	$J-1$	J
1	X_{11}	X_{12}	X_{13}	$\dots$	$X_{1,J-1}$	X_{1J}
2	X_{21}	X_{22}	X_{23}	$\dots$	$X_{2,J-1}$	
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$		
$J-2$	$X_{J-2,1}$	$X_{J-2,2}$	$X_{J-2,3}$			
$J-1$	$X_{J-1,1}$	$X_{J-1,2}$				
J	$X_{J,1}$					

Les realitzacions conegudes van apareixent horitzontalment per a cada any d'origen que avancen, al mateix temps, verticalment. Ademés, els riscos $X_{j,J+1-j}$ paral·lels a la diagonal $X_{J1}, ..., X_{1J}$ ens donen les observacions pagades al mateix any d'observació $j + s$. Com ja hem avançat, el problema de les reserves IBNR, consisteix en l'extrapolació de les realitzacions conegudes del triangle anterior per calcular les observacions futures desconegudes, i convertir el triangle amb quadrat. Per altra banda, també expliquem breument l'evolució del tractament d'aquestes reserves fins el dia d'avui.

En els dos següents capítols (6 i 7) analitzem els models de credibilitat per al càlcul de les reserves IBNR. En primer lloc, en el capítol 6, s'analitza el model de De Vylder (1982) que va ser la primera aplicació de la Teoria de la Credibilitat al càlcul de reserves IBNR. El problema que es planteja és el d'estimar $E[X_{js} / \Theta_j]$,

$$E[X_{js} / \Theta_j] = y_s \cdot \beta(\Theta_j) \tag{10.22}$$

Per tant, només ens cal estimar y_s i trobar un estimador de credibilitat per $\beta(\Theta_j)$. Per això, disposem de les variables observades X_{js} . Es pot reduir el problema a un de Teoria de la Credibilitat amb les següents hipòtesis:

(DV.1) Les pòlisses $j = 1, 2, \ldots, k$, o sigui, els parells $(\Theta_1, \vec{X}_1)$, $(\Theta_2, \vec{X}_2), \ldots$, $(\Theta_k, \vec{X}_k)$ són independents.

(DV.2) $E\left[\vec{X}_j / \Theta_j\right] = y_s \cdot \beta(\Theta_j)$

(DV.3) $Cov\left[\vec{X}_j / \Theta_j\right] = \dfrac{r^2}{w_j}\beta^2(\Theta_j) \cdot I = \sigma_j^2(\Theta_j) \cdot I$

(DV.4) Equidistribució i independència de les variables estructurals $\Theta_1, \Theta_2, \ldots, \Theta_k$.

Sota aquestes condicions, es pot aplicar el model de regressió de Hachemeister (vegeu apèndix A) per estimar $\beta(\Theta_j)$ i, amb algunes generalitzacions, els resultats d'aquest model són vàlids.

L'estimador de credibilitat B_j per $\beta(\Theta_j)$ és igual a

$$B_j = (1 - z_j)b + z_j\hat{b}_j \tag{10.23}$$

on

$$\hat{b}_j = \left(y_j'y_j\right)^{-1} y_j'X_j = \left(\sum_{s\in T_j} y_s X_{js}\right) / \left(\sum_{s\in T_j} y_s^2\right)$$

$$\begin{aligned} z_j &= \frac{a}{a + S_j^2 v_j} \\ S_j^2 &= E[\sigma_j^2(\Theta_j)] = S^2 / w_j \\ v_j &= \left(y_j' y_j\right)^{-1} = \frac{1}{\sum\limits_{s \in T_j} y_s^2} \end{aligned}$$

També donem els estimadors dels paràmetres estructurals y_s , S^2 i a .

En resum, de

$$E[X_{js} / \Theta_j] = y_s \cdot \beta(\Theta_j) \tag{10.24}$$

y_s ens aporta la informació dels anys de desenvolupament i $\beta(\Theta_j)$ la dels anys d'origen. Cap variable ens aporta quelcom sobre els anys de pagament (diagonals del triangle). De totes maneres, per possibles canvis en aquests anys de les diagonals, com per exemple la inflació, el mateix De Vylder ens proporciona una possible solució: deflactar les dades i tornar a calcular-ho de la mateixa manera. Hadidi, N. (1985) preveu que en certes ocasions l'actuari voldria combinar les dades del triangle "run-off" amb la seva pròpia intuició. Aquestes dues extensions són explicades al final del capítol 6.

El model de De Vylder va ser millorat i generalitzat per Mack, T. (1990). En el model de De Vylder, (DV.3) no sembla real en la majoria dels casos, com es pot veure en el capítol 7. En el nou model, s'intenta superar aquest inconvenient a través de (VM.4). A més es presenta el model com un cas especial del model de credibilitat de Bühlmann-Straub més simple que el de Hachemeister, en el qual De Vylder es va basar per muntar el seu model. Mack, T. (1990) fa les següents hipòtesis:

(VM.1) Les pòlisses $j = 1, 2, ..., k$, o sigui, els parells $(\Theta_1 , \vec{X}_1), (\Theta_2 , \vec{X}_2), ...,$ $(\Theta_k , \vec{X}_k)$ són independents entre sí, i hi ha equidistribució i independència de les variables estructurals $\Theta_1 , \Theta_2 , ..., \Theta_k$.

(VM.2) $Cov[X_{js} , X_{ju} / \Theta_j] = 0$ per $s \neq u$

(VM.3) $E\,[X_{js} / \Theta_j] = y_s \cdot \beta(\Theta_j)$

(VM.4) $Var\,[X_{js}\,/\,\Theta_j] \;=\; y_s \cdot \sigma^2(\Theta_j)\,/\,w_j$

Les modificacions de Mack no canvien l'estructura interna, així de

$$E[X_{js}\,/\,\Theta_j] \;=\; y_s \cdot \beta(\Theta_j) \tag{10.25}$$

y_s ens segueix aportant la informació dels anys de desenvolupament i $\beta(\Theta_j)$ la dels anys d'origen. Cap variable, com abans, ens aporta quelcom sobre els anys de pagament (diagonals del triangle).

Finalment, podem ajuntar els dos models en un de sol, el model De Vylder-Mack, que assumeix les següents hipòtesis:

(VM.1) Les pòlisses $j = 1, 2, \ldots, k$, o sigui, els parells $(\Theta_1\,,\,\vec{X}_1)\,,\,(\Theta_2\,,\,\vec{X}_2)\,,\ldots,$ $(\Theta_k\,,\,\vec{X}_k)$ són independents entre sí i hi ha equidistribució i independència de les variables estructurals $\Theta_1\,,\,\Theta_2\,,\,\ldots,\,\Theta_k$.

(VM.2) $Cov[X_{js}\,,\,X_{ju}\,/\,\Theta_j] \;=\; 0$ per $s \neq u$

(VM.3) $E\,[X_{js}\,/\,\Theta_j] \;=\; y_s \cdot \beta(\Theta_j)$

(VM.4)' $Var\,[X_{js}\,/\,\Theta_j] \;=\; y_s^{2-\alpha} \cdot \sigma^2(\Theta_j)\,/\,w_j \qquad \alpha \;=\; 0\,,\,1\,,\,2.$

On α pot valer 0, 1, 2. Si $\alpha = 2$ es correspon amb el model de De Vylder, on la variància no depén de s . En canvi, si $\alpha = 1$ (model de Mack) la hipòtesi resultant és la nostra favorita, on la distribució de pagaments o els canvis d'avaluació són els mateixos per tots els anys de desenvolupament. Si $\alpha = 0$ obtenim la hipòtesi de l'alternativa general que s'analitza a 7.3.3.

10.4 Tancant el cercle

Arribats a aquest punt, després de veure l'aplicabilitat dels models de credibilitat en el càlcul de reserves IBNR, era fins a cert punt lògic, utilitzar els nous models de credibilitat de classificació creuada per al càlcul d'IBNR. Fruit d'aquesta unió neix el model IBNR "two-way", desenvolupat al capítol 8. La independència entre els pagaments corresponents a diferents períodes d'origen, assumida en la majoria dels models IBNR de credibilitat, no és sempre una assumpció real, com ja fou indicat a Sundt, B. (1979). Per exemple, en un particular any d'observació (o de pagament) el nombre de sinistres que corresponen a diferents anys d'origen estan també relacionats d'alguna manera com a conseqüència, per exemple, del temps o la inflació. De Vylder i Mack bàsicament, assumeixen que tals influències no són aleatòries i les modelen a partir dels paràmetres y_{js} . En aquest nou model, tenim en compte aquestes dependències entre pagaments de diferents anys d'origen, però pagades al mateix any d'observació, a través de la introducció d'uns nous paràmetres. Cada període d'observació correspon a la diagonal del triangle "run off" (de l'inferior esquerre al superior dret) que conté les realitzacions de X_{js} amb el mateix valor $j + s$. La $(j+s)$-èsima diagonal és llavors caracteritzada per la variable $\Theta^{(d)}_{j+s}$, mentrestant les característiques derivades de cada any d'origen són ara representades per la variable aleatòria $\Theta^{(o)}_{j}$. D'aquesta manera, el model de De Vylder-Mack pot ser extés a un model de classificació creuada "two way":

(TW.1) Les variables $\Theta^{(o)}_{j}$ són independents i estan idènticament distribuïdes. Les variables $\Theta^{(d)}_{j+s}$ també ho són, ademés són independents de $\Theta^{(o)}_{j}$. Les variàncies de X_{js} són finites.

(TW.2) $E\left[X_{js} \,/\, \Theta^{(o)}_{j} \,,\, \Theta^{(d)}_{j+s}\right] \;=\; y_{js} \cdot \mu(\Theta^{(o)}_{j} \,,\, \Theta^{(d)}_{j+s})$

(TW.3) $E[Cov[X_{js} \,,\, X_{ju} \,/\, \Theta^{(o)}_{j} \,,\, \Theta^{(d)}_{j+s}]] \;=\; \delta_{su} \cdot y_{js}^{2-\alpha} \cdot S^2 \,/\, w_j$

$$\text{on } \delta_{su} = \begin{cases} 1 & s = u \\ 0 & s \neq u \end{cases} \text{, i } \alpha = 0\,,1\,,2.$$

Les variables $\Theta^{(o)}_{j}$ i $\Theta^{(d)}_{j+s}$ recullen la informació sobre les dependències que s'estableixen entre els sinistres de cada any d'origen $(\Theta^{(o)}_{j})$ i entre els pagaments que es fan al mateix any d'observació ($\Theta_{j+s}^{(d)}$). En resum,

$$\text{(TW.2)'} \quad E\left[X_{js} \,/\, \Theta_j^{(o)}\,,\, \Theta_{j+s}^{(d)}\right] \;=\; y_s \cdot \mu(\Theta_j^{(o)}\,,\, \Theta_{j+s}^{(d)})$$

on y_s aporta la informació dels anys de desenvolupament, com feiem en els dos models anteriors. En canvi, $\mu(\Theta_j^{(o)}\,,\, \Theta_{j+s}^{(d)})$ ens aporta la informació dels anys d'origen i, com a novetat, la informació dels anys de pagament o observació. Aquesta segona part té l'estimador de credibilitat U_{js} que pot ser escrit com:

$$U_{js}^{*} \;=\; 1 \,+\, \Xi_j^{(o)*} \,+\, \Xi_{j+s}^{(d)*} \,+\, \Delta_{js}^{*} \tag{10.26}$$

D'aquesta manera, ens trobem amb un model additiu "two way". Finalment, l'estimador de credibilitat per X_{js} és

$$E\left[X_{js} \,/\, \Theta_j^{(o)}\,,\, \Theta_{j+s}^{(d)}\right] \;=\; y_s \cdot \left(1 \,+\, \Xi_j^{(o)*}\right) \tag{10.27}$$

Cal tindre en compte que $\Xi_j^{(o)*}$ conté la informació dels anys d'origen ($\Theta_j^{(o)}$) i també la dels anys de pagament ($\Theta_{j+s}^{(d)}$). Per tant, suposa una millora respecte el model De Vylder-Mack.

Per últim, en el capítol 9 hem aplicat unes dades reals als diferents mètodes d'IBNR per treure'n comparacions i conclusions. Per portar a terme aquesta aplicació hem escollit les dades d'una companyia belga que opera per tota Europa. Les dades pertanyen a un contracte d'assegurança de crèdit utilitzat per diversos estats: Bèlgica, Països Baixos, Alemanya, França i el Regne Unit. Les dades del triangle "run-off" són trimestrals. Un cop exposades les dades, mostrem els resultats que hem obtingut pels tres models analitzats en aquesta tesi doctoral. Per això, hem programat aquests models amb llenguatge APL com es pot veure a l'apèndix C. Un cop vistos els resultats d'aquests cinc països pels tres mètodes estudiats: mètode de De Vylder (DV), de De Vylder-Mack (VM) i de "two way" (2W), ens preguntem dues coses: per una banda, quin d'aquests tres mètodes és el més apropiat?; i per altra banda, milloren els mètodes IBNR credibilístics els models d'IBNR clàssics?

Per contestar a aquestes preguntes calia trobar un criteri per comparar els resultats. Finalment n'hem escollit dos de diferents:

1. Comparar la mitja dels errors quadràtics de la part superior dels triangles "run-off".

2. Partint, com abans, del triangle "run-off" original, eliminen la darrera diagonal, és a dir, les dades que corresponen al darrer període $j + s$. Amb les dades restants, estimem el triangle inferior, que inclou la diagonal eliminada, amb el mètode respectiu. D'aquesta manera, podem comparar les dades reals del darrer any $j + s$ amb l'estimació que hem fet d'aquest mateix any.

Un cop fetes les comparacions pertinents, podem arribar a les següents conclusions:

1. Segons el primer mètode de comparació, podem veure que el millor mètode és el de mínims quadrats de De Vylder per als cinc països; resultat que no ens hauria de sorpendre, ja que el mètode mencionat consisteix precisament en trobar els resultats que minimitzin els errors quadràtics esperats. Sense comptar amb aquest mètode, podem veure que, en general, els mètodes de credibilitat tenen uns errors al quadrat menors que els clàssics. Amb això, ja podem destacar la millor aproximació dels nous mètodes envers els clàssics.

2. Amb el segon mètode, també podem observar que els models de credibilitat, en general, funcionen millor que els mètodes clàssics.

3. El que no sembla tan fàcil de respondre, és la primera pregunta que ens feiem sobre quin dels tres models de credibilitat és el millor. Per als dos mètodes de comparació no es poden treure unes conclusions prou clares, per tant, no podem assegurar quin dels tres és el millor ja que segons quines siguin les dades, l'ajustament serà millor amb un mètode que amb un altre. De totes maneres, les diferències entre aquests tres mètodes són molt petites.

Per finalitzar aquestes conclusions, creiem que els resultats d'aquesta tesi demostren la millora, que per algunes aplicacions, representen els models de credibilitat. Per fer un progrés en aquesta direcció, la comunicació entre els actuaris "teòrics" i els "pràctics" ha de ser més important que fins el dia d'avui. La Teoria de la Credibilitat no ha de quedar al marge del món assegurador, tot treball que vagi amb la línia d'aplicar aquesta teoria a la vida quotidiana de les companyies d'assegurances, és de vital importància perquè la Teoria de la Credibilitat creixi en nombre d'aplicacions i no només en el pla teòric.

Per altra banda, les possibles conexions amb altres disciplines estadístiques han de ser promugudes per enriquir els dos camps de recerca. Alguns exemples d'aquestes col·laboracions, han estat enunciats en la introducció d'aquest treball, on les principals línies de recerca en el camp credibilístic es defineixen.

APÈNDIXS

Apèndix A

Models de credibilitat

A.1 Model de Bühlmann

Els fonaments de la Teoria de la Credibilitat van ser establerts per Bühlmann, H. al 1965 quan va presentar la seva fòrmula de credibilitat de distribució lliure, publicada el 1967, basada en el criteri dels mínims quadrats. Encara que les hipòtesis que va introduir han estat objecte de crítica degut a la seva rigidesa, la formulació del seu model fou un important pas pel desenvolupament de la Teoria de la Credibilitat , i ha donat lloc a l'aparició de models cada vegada més perfeccionats.

A.1.1 Variables rellevants

El model de Bühlmann utilitza les següents variables:

- Θ_j : És el paràmetre de risc per la pòlissa j-èsima. És una variable estructural, que descriu les característiques del risc del contracte j-èsim, amb $j = 1, 2, \ldots, k$; éssent k el nombre total de pòlisses de la cartera. En la majoria dels casos aquest paràmetre és desconegut o inobservable, d'aquí, que sigui considerat com a variable aleatòria desconeguda.

- X_{js} : Variable que indica l'experiència de reclamacions per la pòlissa j-èsima en el període s-èsim, on $j = 1, 2, \ldots, k$ i $s = 1, 2, \ldots, t$, on t és el nombre de períodes observats per a cada pòlissa. També és una variable aleatòria, però amb realitzacions observables. De vegades s'interpreta com l'import mitjà de les indemnitzacions per sinistre.

Així doncs, el model de Bühlmann s'estructura segons el següent esquema:

	$j = 1$	$j = 2$	$\ldots$	$j = k$
	Θ_1	Θ_2	$\ldots$	Θ_k
1	X_{11}	X_{21}	$\ldots$	X_{k1}
2	X_{12}	X_{22}	$\ldots$	X_{k2}
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
t	X_{1t}	X_{2t}	$\ldots$	X_{kt}

on la pòlissa j-èsima ve explicada pel vector:

$$(\Theta_j, X_{j1}, X_{j2}, \ldots, X_{jt}) = (\Theta_j, \vec{X}_j)$$

Per tant:

- cada columna representa, per exemple, la quantia mitjana de reclamacions d'un tipus Θ_j de pòlissa per cada període t.

- i cada fila la quantia mitjana de reclamacions en un període per cadascuna de les diferents pòlisses.

A.1.2 Hipòtesis del model

Les hipòtesis assumides per Bühlmann, H. (1967) són dues:

(B.1) Les pòlisses $j = 1, 2, \ldots, k$, o sigui, els parells $(\Theta_1, \vec{X}_1), (\Theta_2, \vec{X}_2), \ldots,$ $(\Theta_k, \vec{X}_k)$ són independents i estan idènticament distribuïts.

(B.2) Per cada contracte $j = 1, 2, \ldots, k$, i per un Θ_j donat, les variables condicionades: $X_{j1} / \Theta_j, X_{j2} / \Theta_j, \ldots, X_{jt} / \Theta_j$ són independents i estan idènticament distribuïdes.

La primera hipòtesi indica l'existència d'independència entre les pòlisses (riscos) i equivalència exterior de les mateixes o equidistribució de les variables d'estructura, o dit d'un altra manera, s'està assumint que les pòlisses són independents entre elles, i *a priori* , tant és escollir una pòlissa com una altra.

La segona hipòtesi expressa l'existència d'independència dins de cada risc i homogeneïtat en el temps, per tant, s'assumeix implícitament que no hi ha aprenentatge, que un individu no millora ni empitjora amb el temps.

Així doncs, podem expresar les hipòtesis de forma alternativa, com:

(B.1) Les pòlisses $j = 1, 2, \ldots, k$, o sigui, els parells $(\Theta_1, \vec{X}_1), (\Theta_2, \vec{X}_2), \ldots,$ $(\Theta_k, \vec{X}_k)$ són independents i estan idènticament distribuïts.

(B.2) (a) $E[X_{js} / \Theta_j] = \mu(\Theta_j) \qquad s \in (1, 2, \ldots, t)$

les observacions esperades no depenen del període en que ens situem, degut a l'existència d'homogeneïtat en el temps. Bühlmann defineix $\mu(\Theta_j)$ com la prima de risc individual.

(b) $Cov[\vec{X}_j / \Theta_j] = \sigma^2(\Theta_j) \cdot I \qquad s \in (1, 2, \ldots, t)$

on I és la matriu identitat, de dimensió $(t \times t)$, $\sigma^2(\Theta_j) = Var[X_j / \Theta_j]$ i $\vec{X}_j$ és un vector de dimensió $(t \times 1)$, éssent $\vec{X}_j = (X_{j1}, X_{j2}, \ldots, X_{jt})$.

A.1.3 Notació prèvia

En primer lloc:

$$\mu(\Theta_j) \;=\; E[X_{js} \,/\, \Theta_j]$$

És la *prima de risc individual* per una pòlissa concreta, la j -èsima, o sigui, és la quantitat esperada de reclamacions individuals; que no és res més que l'esperança de les reclamacions observades condicionada a un paràmetre de risc fix.

$$m \;=\; E[X_{js}] \;=\; E\Big[\mu(\Theta_j)\Big]$$

Denota la quantitat mitjana esperada de reclamacions pel conjunt de la cartera. És la *prima de risc col·lectiva*; o sigui, el valor esperat de totes les primes de risc individuals.

$$a \;=\; Var\Big[E[X_{js} \,/\, \Theta_j]\Big] \;=\; Var\Big[\mu(\Theta_j)\Big]$$

Medeix la variància o dispersió que hi ha entre les primes de risc individuals. És un indicador de la heterogeneïtat de la cartera, mesurada a través de la quantitat de reclamacions esperades.

$$S^2 \;=\; E\Big[Var[X_{js} \,/\, \Theta_j]\Big] \;=\; E\Big[\sigma^2(\Theta_j)\Big]$$

És el valor esperat de la dispersió total de les dades de reclamacions de la cartera, ja que $Var[X_{js} \,/\, \Theta_j]$ reflexa la variància de la variable experiència de reclamacions per una pòlissa concreta en el temps.

D'altra banda:

1. $Cov[X_{jr}\,,\, X_{js}] \;=\; a + \delta_{rs} \cdot S^2$

2. $Cov[\mu(\Theta_j)\,,\, X_{js}] \;=\; a$

on δ_{rs} és el símbol Kronecker:

$$\delta_{rs} = \begin{cases} 0 & \text{si} \quad r \neq s \\ 1 & \text{si} \quad r = s \end{cases}$$

A.1.4 Estimació de la prima de risc individual

L'objectiu de Bühlmann, H. (1967) era estimar de la millor manera possible la prima de risc individual $\mu(\Theta_j)$. Per fer-ho, va decidir que les primes fossin lineals i, per tant, l'objectiu seria seleccionar la millor prima dins de la classe restringida de les primes lineals de la forma:

$$c_{j0} + \sum_{s=1}^{t} c_{js} \cdot X_{js} \tag{A.1}$$

A més va demostrar que si l'anterior expressió es la millor aproximació lineal per $\mu(\Theta_j)$, també ho serà per l'estimador posterior de Bayes:

$$E\left[\left(\mu(\Theta_j) \,/\, X_{j1}\,,\, X_{j2}\,,\, \dots\,,\, X_{jt}\right)\right]$$

Per tant, la millor aproximació lineal per $\mu(\Theta_j)$ es calcula trobant els valors dels coeficients $c_{j0}\,,\, c_{j1}\,,\, \dots\,,\, c_{jt}$ tals que minimitzin

$$L(c_{j0}\,,\, c_{j1}\,,\, \dots\,,\, c_{jt}) = E\left[\left(\mu(\Theta_j) - c_{j0} - \sum_{s=1}^{t} c_{js} \cdot X_{js}\right)^2\right] \tag{A.2}$$

d'acord amb el criteri de la màxima precisió mínimquadràtica.

Per tant, primer derivem l'expressió anterior respecte els $t+1$ coeficients que tenim i igualem a zero, obtenint un sistema d'equacions homogeni de $t+1$ equacions amb $t+1$ incògnites :

$$\begin{cases} E\left[\mu(\Theta_j) - c_{j0} - \sum_{s=1}^{t} c_{js} \cdot X_{js}\right] = 0 & \\ E\left[X_{jr} \cdot \left(\mu(\Theta_j) - c_{j0} - \sum_{s=1}^{t} c_{js} \cdot X_{js}\right)\right] = 0 & r = 1\,,2\,,\dots\,,t \end{cases} \tag{A.3}$$

De manera que els valors de c_{j0}, c_{j1}, ..., c_{jt} que minimitzen (A.2) són:

$$c_{j0} = m \cdot (1 - z) \tag{A.4}$$

$$c_{j1} = c_{j2} = \dots = c_{jt} = c = \frac{z}{t} \tag{A.5}$$

on $z = \dfrac{a \cdot t}{S^2 + a \cdot t}$

D'aquí que, segons (A.5), l'estimador ajustat de credibilitat per $\mu(\Theta_j)$ és:

$$\hat{\mu}(\Theta_j) = c_{j0} + \sum_{s=1}^{t} c_{js} \cdot X_{js} = c_{j0} + c \cdot \sum_{s=1}^{t} X_{js}$$

I substituint c_{j0} i c pels respectius valors trobats, (A.4) i (A.5), resulta que:

$$\hat{\mu}(\Theta_j) = (1 - z) \cdot m + z \cdot \bar{X} \tag{A.6}$$

on:

$$\bar{X} = \sum_{s=1}^{t} \frac{X_{js}}{t}$$

Resumint, la prima de credibilitat calculada no és res més que una ponderació entre la que seria la prima, si només ens fixèssim en l'experiència individual ($\bar{X}$), i la prima col·lectiva (m). En cada cas, z , anomenada factor de credibilitat, indicarà en quin grau intervindrà cadascuna d'elles en la prima de credibilitat.

A.1.5 Propietats del factor de credibilitat

El factor de credibilitat és un número comprés entre 0 i 1 , ambdós inclosos. Les propietats que verifica aquest factor són les següents:

1.- Si $\quad t \to \infty \quad \Longrightarrow \quad z \to 1$

Quan t creix, aleshores z tendeix a la unitat. Això significa que si tenim més informació sobre el risc, més gran serà la ponderació relativa donada a l'experiència individual. En el cas de tenir experiència completa, és obvi, que tindrem una total confiança en la prima de risc individual.

2.- Si $a = 0 \implies z = 0$

Com a és un indicador de la heterogeneïtat dins la cartera, si és nul, significa que no hi ha heterogeneïtat i m és, per tant, el millor estimador lineal per $\mu(\Theta_j)$.

3.- Si $a \to \infty \implies z \to 1$

Que a tendeix a infinit ens indica que el col·lectiu és extremadament heterogeni, per tant, les diferències observades entre les reclamacions registrades d'un risc a un altre són importants. En aquest cas, $\bar{X}$ és el millor estimador lineal.

4.- Si $S^2 \to \infty \implies z \to 0$

I per últim, quan S^2 tendeix a infinit , això és, si existeix molta variància en la sinistralitat individual, el millor estimador per la prima de risc individual és la col·lectiva.

A.1.6 Estimació dels paràmetres estructurals

Norberg, R. (1979), anomena m, a i S^2 paràmetres estructurals. Utilitza aquest terme per indicar que, a la pràctica, aquests paràmetres han de ser estimats a partir de les observacions empíriques, i així poder calcular z i $\hat{\mu}(\Theta_j)$. Per l'obtenció dels paràmetres seguirem l'article de Norberg, R. (1979), on els paràmetres que volem estimar són:

$$m = E[X_{js}] = E\Big[\mu(\Theta_j)\Big]$$

$$a = Var\Big[E[X_{js} \,/\, \Theta_j]\Big] = Var\Big[\mu(\Theta_j)\Big]$$

$$S^2 = E\Big[Var[X_{js} \,/\, \Theta_j]\Big] = E\Big[\sigma^2(\Theta_j)\Big]$$

i els estimadors proposats per Norberg :

- Per la mitjana poblacional:

$$\hat{m} = \frac{1}{k} \cdot \sum_{j=1}^{k} \bar{X}_j = \frac{1}{k} \cdot \sum_{j=1}^{k} \sum_{s=1}^{t} \frac{X_{js}}{t}$$

- Pel valor esperat de la dispersió total:

$$\hat{S}^2 = \frac{1}{k} \cdot \sum_{j=1}^{k} \frac{1}{t-1} \cdot \sum_{s=1}^{t} (X_{js} - \bar{X}_j)^2$$

- Per la variació de la prima pura verdadera individual entre els riscos de la població:

$$\hat{a} = \frac{1}{k-1} \cdot \sum_{j=1}^{k} (\bar{X}_j - \hat{m})^2 - \frac{1}{t} \cdot \hat{S}^2$$

Els tres estimadors són no esbiaxats i consistents.

Resumint, direm que el Model de Bühlmann tan sols és adequat en el cas d'utilitzar quantitats de reclamacions deflactades o sense tendència, degut a la homogeneïtat en el temps que es considera, i a més es considera implícitament que les pòlisses tenen el mateix impacte en el negoci.

A.2 Model de Bühlmann-Straub

Aquest model (1970), no es res més que una senzilla ampliació del model de Bühlmann. És un model que considera el temps homogeni, però amb observacions ponderades, i és probablement el model més utilitzat a la pràctica.

Va aparèixer per superar la limitació que suposava el model de Bühlmann quan considerava que totes les pòlisses tenien igual importància dins de la cartera, independentment del seu volum, limitació que en aquest model deixa de ser-ho a l'introduir una ponderació o pes natural per a cada pòlissa, ampliant d'aquesta manera el seu camp d'aplicació. Les observacions esperades segueixen sent homogènies en el temps, però la variància deixa de ser constant, passant a dependre del període considerat via els pesos o ponderacions naturals. Un altre element important a tenir en compte, és que el factor de credibilitat ja no és constant per tota la cartera, com ho era al model anterior, sinó que cada pòlissa te ara el seu propi factor de credibilitat. Tot i això, aquest model, com l'anterior, només és apte en el cas de treballar amb dades deflactades o sense tendència.

A.2.1 Variables rellevants

A més de les dues que ja apareixien en el model de Bühlmann, el present model incorpora una nova variable:

- Θ_j : És el paràmetre de risc per la pòlissa j-èsima. És una variable estructural, que descriu les característiques del risc del contracte j-èsim, amb $j = 1, 2, ..., k$; éssent k el nombre total de pòlisses de la cartera. En la majoria dels casos aquest paràmetre és desconegut o inobservable, d'aquí, que sigui considerat com a variable aleatòria desconeguda.

- X_{js} : Variable que indica l'experiència de reclamacions per la pòlissa j-èsima en el període s-èsim, on $j = 1, 2, ..., k$ i $s = 1, 2, ..., t$, on t és el nombre de períodes observats per a cada pòlissa. També és una variable aleatòria, però amb realitzacions observables.

- w_{js} : Variable que indica els pesos o ponderacions naturals, que són nombres positius coneguts, amb $j = 1, 2, ..., k$ i $s = 1, 2, ..., t$.

Normalment, X_{js} representa la quantia mitjana de reclamacions o sinistres i w_{js} sol representar el nombre de reclamacions o sinistres. Encara que també X_{js} pot representar el ràtio de pèrdues i en aquest cas w_{js} serien els corresponents volums de primes. La incorporació d'aquestes ponderacions amplia considerablement el camp d'aplicació del model.

Amb el següent esquema es pot resumir la informació que disposem :

	$j = 1$	$j = 2$	$\ldots$	$j = k$
	Θ_1	Θ_2	$\ldots$	Θ_k
1	$X_{11}(w_{11})$	$X_{21}(w_{21})$	$\ldots$	$X_{k1}(w_{k1})$
2	$X_{12}(w_{12})$	$X_{22}(w_{22})$	$\ldots$	$X_{k2}(w_{k2})$
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
t	$X_{1t}(w_{1t})$	$X_{2t}(w_{2t})$	$\ldots$	$X_{kt}(w_{kt})$

on X_{js} serà el cost mitjà per sinistre i w_{js} el nombre de sinistres de cada pòlissa que ha tingut reclamacions. En aquest cas podem escriure:

$$X_{js} = \frac{\sum_{k=1}^{w_{js}} X_{jk}}{w_{js}}$$

O també:

- cada columna representa, per exemple, la quantia mitjana de reclamacions d'un tipus Θ_j de pòlissa per a cada període t .
- cada fila la quantia mitjana de reclamacions en un període per cadascuna de les diferents pòlisses.
- i entre parèntesi el nombre de sinistres per a cada element X_{js} .

A.2.2 Hipòtesis del model

Les hipòtesis assumides per aquest model són:

(BS.1) Les pòlisses $j = 1, 2, \ldots, k$, o sigui, els parells $(\Theta_1 , \vec{X}_1)$, $(\Theta_2 , \vec{X}_2)$, ..., $(\Theta_k , \vec{X}_k)$ són independents i estan idènticament distribuïts.

(BS.2) Per cada contracte $j = 1, 2, ..., k$, i per un Θ_j donat, les variables condicionades: X_{j1} / Θ_j , X_{j2} / Θ_j , ..., X_{jt} / Θ_j són independents i estan idènticament distribuïdes. Aquesta hipòtesi, com en el model anterior, es pot expressar com:

(a) $E[X_{js} / \Theta_j] = \mu(\Theta_j) \qquad s \in (1, 2, ..., t)$

(b) $Cov[X_{jr}, X_{js} / \Theta_j] = \delta_{rs} \cdot \frac{1}{w_{js}} \cdot \sigma^2(\Theta_j) \qquad s \in (1, 2, ..., t)$

El model manté la independència dins i entre les pòlisses, com en el model de Bühlmann. En canvi, encara que les observacions esperades segueixen sent homogènies en el temps:

$$E[X_{js} / \Theta_j] = \mu(\Theta_j) \qquad s \in (1, 2, ..., t)$$

la variància sí que depén del període en què ens situem, a través dels pesos o ponderacions naturals, ja que:

$$Var[X_{js} / \Theta_j] = \frac{1}{w_{js}} \cdot \sigma^2(\Theta_j)$$

A.2.3 Estimació de la prima de risc individual

L'objectiu del model de Bühlmann-Straub és el mateix que el del model anterior, estimar de la millor manera possible la prima de risc individual $\mu(\Theta_j)$. A més, seguirem el mateix procediment, seleccionar la millor prima dins de la classe restringida de les primes lineals, utilitzant per això el procediment dels mínims quadrats. També seguirem la mateixa notació de l'apartat A.1.3.

Recordem que estem interessats en trobar la millor aproximació que sigui lineal per a $E[\mu(\Theta_j) / X_{j1}, X_{j2}, ..., X_{jt}]$ que, com hem vist, es troba calculant els valors $c_{j0}, c_{j1}, ..., c_{jt}$ que minimitzin:

$$L(c_{j0}, c_{j1}, ..., c_{jt}) = E\left[\left(\mu(\Theta_j) - c_{j0} - \sum_{s=1}^{t} c_{js} \cdot X_{js}\right)^2\right]$$

Després de seguir uns passos molts similars als fets en el model de Bühlmann, arribem a trobar que els valors buscats són:

$$c_{j0} = m \cdot (1 - z_j) \tag{A.7}$$

$$c_{js} = \frac{z_j}{w_{j\Sigma}} \cdot w_{js} \tag{A.8}$$

on

$$z_j = \frac{a \cdot \sum_{s=1}^{t} w_{js}}{S^2 + a \cdot \sum_{s=1}^{t} w_{js}} = \frac{a \cdot w_{j\Sigma}}{S^2 + a \cdot w_{j\Sigma}}$$

i

$$w_{j\Sigma} = \sum_{s=1}^{t} w_{js}$$

Per tant, l'estimador ajustat de credibilitat per $\mu(\Theta_j)$ és:

$$\begin{aligned} \hat{\mu}(\Theta_j) &= c_{j0} + \sum_{s=1}^{t} c_{js} \cdot X_{js} \\ &= m \cdot (1 - z_j) + \sum_{s=1}^{t} \frac{z_j}{w_{j\Sigma}} \cdot w_{js} \cdot X_{js} \\ &= m \cdot (1 - z_j) + z_j \cdot \sum_{s=1}^{t} \frac{w_{js} \cdot X_{js}}{w_{j\Sigma}} \end{aligned}$$

I si simbolitzem per:

$$X_{jw} = \sum_{s=1}^{t} \frac{w_{js} \cdot X_{js}}{w_{j\Sigma}}$$

l'estimador per la prima de risc individual resulta ser:

$$\hat{\mu}(\Theta_j) = (1 - z_j) \cdot m + z_j \cdot X_{jw} \tag{A.9}$$

A.2.3.1 Propietats del factor de Credibilitat

En aquest cas, el factor de credibilitat z_j ja no és una constant com en el model anterior, sinó que depén de j , a través de $w_{j\Sigma}$, amb $j = 1, 2, \ldots, k$. Així doncs, un element important a tenir en compte és que cada pòlissa té el seu propi factor de credibilitat que és definit de la següent manera:

$$z_j = \frac{a \cdot w_{j\Sigma}}{S^2 + a \cdot w_{j\Sigma}}$$

De l'expressió de z_j es desprén que quan el valor de w_j és elevat, z_j està pròxim a 1, ja que quan t creix disposem de més informació sobre el risc, i més gran serà la ponderació relativa donada a la experiència individual. No obstant, cal assenyalar que l'elevat valor de w_j , també pot ser causat perquè els pesos credibilístics w_{js} siguin elevats durant els t períodes observats sense necessitat d'allargar el període d'observació. El factor de credibilitat z_j segueix verificant la resta de les propietats que verificava en el model de Bühlmann:

1.- Si $a = 0 \implies z_j \to 0$

2.- Si $a \to \infty \implies z_j \to 1$

3.- Si $S^2 \to \infty \implies z_j \to 0$

A.2.4 Estimació dels paràmetres estructurals

Com ja sabem, m, a i S^2 són els paràmetres estructurals a estimar per poder realitzar els càlculs de z_j i $\hat{\mu}(\Theta_j)$. Primer, introduirem algunes notacions convencionals, a efectes de presentar les fórmules d'aquests estimadors d'una manera més entenedora. Suposant que X_{js} representa la quantia mitjana de reclamacions (cost mitjà per prima) i w_{js} representa el nombre de reclamacions (volum de primes), les següents expressions es defineixen com:

1.- $\mathbf{w_{j\Sigma}}$: suma, per a cada contracte j , del nombre de primes en els t períodes observats:

$$w_{j\Sigma} = \sum_{s=1}^{t} w_{js}$$

2.- $\mathbf{w_{\Sigma\Sigma}}$: volum total de primes:

$$w_{\Sigma\Sigma} = \sum_{j=1}^{k} w_{j\Sigma}$$

3.- $\mathbf{X_{jw}}$: mitjana ponderada, segons el volum de primes (w_{js}), per a cada tipus de contracte j :

$$X_{jw} = \sum_{s=1}^{t} \frac{w_{js}}{w_{j\Sigma}} \cdot X_{js}$$

4.- $\mathbf{X_{ww}}$: mitjana poblacional, comportament mitjà de la sinistralitat a la població, per a tot tipus de contracte j :

$$X_{ww} = \sum_{j=1}^{k} \frac{w_{j\Sigma}}{w_{\Sigma\Sigma}} \cdot X_{jw} = \sum_{j=1}^{k} \sum_{s=1}^{t} \frac{w_{js}}{w_{\Sigma\Sigma}} \cdot X_{js}$$

5.- $\mathbf{X_{zw}}$: mitjana poblacional ponderada, segons el factor de credibilitat z_j :

$$X_{zw} = \sum_{j=1}^{k} \frac{z_j}{z_\Sigma} \cdot X_{jw} = \sum_{j=1}^{k} \sum_{s=1}^{t} \frac{z_j}{z_\Sigma} \cdot \frac{w_{js}}{w_{\Sigma\Sigma}} \cdot X_{js}$$

Amb aquesta notació podem passar a definir els estimadors de:

$$\begin{aligned} m &= E[X_{js}] = E\big[\mu(\Theta_j)\big] \\ a &= Var\big[E[X_{js} \,/\, \Theta_j]\big] = Var\big[\mu(\Theta_j)\big] \\ S^2 &= E\big[Var[X_{js} \,/\, \Theta_j]\big] = E\big[\sigma^2(\Theta_j)\big] \end{aligned}$$

que són:

- Per la mitjana poblacional:

$$\hat{m} = X_{zw}$$

- Pel valor esperat de la dispersió total:

$$\hat{S}^2 = \frac{1}{k \cdot (t-1} \cdot \sum_{j=1}^{k} \sum_{s=1}^{t} w_{js} \cdot (X_{js} - X_{jw})^2$$

- Per la variació de la prima pura verdadera individual entre els riscos de la població:

$$\hat{a} = \frac{w_{\Sigma\Sigma}}{w_{\Sigma\Sigma}^2 - \sum_{j=1}^{k} w_{j\Sigma}^2} \cdot \left(\sum_{j=1}^{k} w_{j\Sigma} \cdot (X_{jw} - X_{ww})^2 - (k - 1) \cdot \hat{S}^2 \right)$$

Els tres estimadors que acabem de veure no són esbiaixats.

A.3 Model jeràrquic de Jewell

Aquest model es pot considerar com una generalització del Model de Bühlmann-Straub. Jewell, W. (1975c) considera que la cartera existent pot ser dividida en un cert nombre p de subcarteres, on cadascuna d'elles està caracteritzada per un paràmetre de risc, que descriu les diferències existents entre les diferents subcarteres, ja que cada subcartera està formada per un grup de pòlisses amb unes característiques comunes. No obstant, cada pòlissa té també unes característiques específiques que la diferencien de les demés pòlisses de la subcartera i que venen recollides per un altre paràmetre de risc. Es tracta doncs d'un model jeràrquic a dos nivells, encara que es pot generalitzar a més nivells. Igual que en el model de Bühlmann-Straub, les observacions són homogènies en el temps, i la variància depén del temps a través dels pesos naturals.

A.3.1 Variables rellevants

Les variables d'aquest model són:

- Θ_p : És el paràmetre de risc que caracteritza la subcartera p amb $p = 1, 2, \ldots, P$; on P és el nombre total de subcarteres en què s'ha dividit la cartera. Es tracta d'una variable aleatòria inobservable.

- Θ_{pj} : És el paràmetre de risc que descriu les característiques de la pòlissa j-èsima i que pertany a la subcartera p amb $j = 1, 2, \ldots, k_p$; on k_p és el nombre total de pòlisses que constitueixen la subcartera p . També es tracta d'una variable aleatòria inobservable.

- X_{pjs} : Variable que indica l'experiència de reclamacions per la pòlissa j-èsima de la subcartera p , en el període s-èsim, amb $p = 1, 2, \ldots, P$; $j = 1, 2, \ldots, k_p$ i $s = 1, 2, \ldots, t_{pj}$, on t_{pj} és el nombre de períodes observats per a cada pòlissa de la subcartera p .

- w_{pjs} : Variable que indica els pesos o ponderacions naturals de les variables observables $Xpjs$; són nombres positius coneguts

Així doncs, una subcartera p ve explicada pel següent conjunt de variables: Θ_p, Θ_{pj}, X_{pjs} i un contracte pj per : Θpj, X_{pjs}

El següent esquema pot resultar clarificador:

	Subcarteres						
	$p = 1$			$\ldots$	$p = P$		
	Θ_1			$\ldots$	Θ_P		
	Pòlisses						
	$pj = 11$	$\ldots$	$pj = 1k_p$	$\ldots$	$pj = P1$	$\ldots$	$pj = Pk_p$
	Θ_{11}	$\ldots$	Θ_{1k_p}	$\ldots$	Θ_{P1}	$\ldots$	Θ_{Pk_p}
$s = 1$	X_{111}	$\ldots$	X_{1k_p1}		X_{P11}	$\ldots$	X_{Pk_p1}
$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\ldots$	$\vdots$	$\ddots$	$\vdots$
$s = t$	X_{11t}	$\ldots$	X_{1k_pt}		X_{P1t}	$\ldots$	X_{Pk_pt}

A.3.2 Hipòtesis del model

Les hipòtesis assumides pel model són:

(J.1) Les subcarteres $p = 1, 2, \ldots, P$ són independents, és a dir, les ternes següents $(\Theta_p\,, \Theta_{pj}\,, X_{pjs})$ són independents.

(J.2) També les pòlisses, o sigui, els parells $(\Theta_{pj}\,, X_{pjs})$ donada una subcartera p són independents i estan idènticament distribuïts.

(J.3) Per cada contracte $p = 1, 2, \ldots, P$ i $j = 1, 2, \ldots, k_p$, i per cada $(\Theta_p, \Theta_{pj}$ donat, les observacions condicionades: $X_{pj1} \,/\, \Theta_p, \Theta_{pj}\,, \ldots, X_{pjt_{pj}} \,/\, \Theta_p, \Theta_{pj}$ són independents.

(J.4) Per tot p , j i s :

(a) $E[X_{pjs} \,/\, \Theta_p, \Theta_{pj}] \;=\; \mu(\Theta_p, \Theta_{pj})$

(b) $Var[X_{pjs} \,/\, \Theta_p, \Theta_{pj}] \;=\; \delta_{rs} \cdot \frac{1}{w_{pjs}} \cdot \sigma^2(\Theta_p, \Theta_{pj})$

on $\mu(\Theta_p, \Theta_{pj}$ i $\sigma^2(\Theta_p, \Theta_{pj})$ no depenen dels subíndexs p , j i s

Com hem dit anteriorment, les observacions esperades són homogènies en el temps, i la variància depén del període considerat a través de les ponderacions naturals.

Paral·lelament, definim $\mu(\Theta_p)$ com:

$$\mu(\Theta_p) \;=\; E[\mu(\Theta_p, \Theta_{pj}) \,/\, \Theta_p] \;=\; E[X_{pjs} \,/\, \Theta_p]$$

que no és res més que el valor esperat per a tots el elements de la subcartera p en el temps, o dit d'una altra manera, és la prima de risc per a cada subcartera p .

A.3.3 Notació prèvia

En primer lloc:

$$m_p \;=\; \mu(\Theta_p) \;=\; E[\mu(\Theta_p, \Theta_{pj}) \,/\, \Theta_p] \;=\; E[X_{pjs} \,/\, \Theta_p]$$

És el valor esperat de les pòlisses de la subcartera p , amb $p = 1, 2, \ldots, P$.

$$m \;=\; E[\mu(\Theta_p)] \;=\; E\big[\mu(\Theta_p,\,\Theta_{pj})\big] \;=\; E[X_{pjs}]$$

És l'esperança conjunta per a la totalitat de la cartera.

$$a \;=\; E\big[Var(\mu(\Theta_p,\,\Theta_{pj})\,/\,\Theta_p)\big]$$

Medeix el grau de variabilitat (heterogeneïtat) esperat dintre de les subcarteres.

$$S^2 \;=\; E\big[\sigma^2(\Theta_p,\,\Theta_{pj}))\big]$$

Medeix l'heterogeneïtat esperada en el temps de l'experiència de reclamacions.

$$b \;=\; Var[\mu(\Theta_p)]$$

Medeix l'heterogeneïtat entre les diferents subcarteres, és una novetat del model.

A més, es defineixen:

$$w_{p\Sigma\Sigma} \;=\; \sum_{j=1}^{k_p} w_{pj\Sigma} \;=\; \sum_{j=1}^{k_p}\sum_{s=1}^{t_{pj}} w_{pjs}$$

$$z_{pj} \;=\; \frac{a\cdot w_{pj\Sigma}}{S^2 + a\cdot w_{pj\Sigma}}$$

És el factor de credibilitat en el nivell dels contractes, factor que coincideix amb l'obtingut en el model de Bühlmann-Straub.

$$z_p \;=\; \frac{b\cdot z_{p\Sigma}}{a + b\cdot z_{p\Sigma}} \qquad ; \qquad z_{p\Sigma} \;=\; \sum_{j=1}^{k_p} z_{pj}$$

És el factor de credibilitat en el nivell de les subcarteres, novetat en aquest model.

$$\begin{aligned} X_{pjw} &= \sum_{s=1}^{t_{pj}} \frac{w_{pjs}\cdot X_{pjs}}{w_{pj\Sigma}} \\ X_{pzw} &= \sum_{j=1}^{k_p} \frac{z_{pj}\cdot X_{pjw}}{z_{p\Sigma}} \\ X_{zzw} &= \sum_{p=1}^{P} \frac{z_p\cdot X_{pzw}}{z_\Sigma} \qquad ; \qquad z_\Sigma \;=\; \sum_{p=1}^{P} z_p \end{aligned}$$

D'altra banda, es troben els següents resultats per a les covariàncies:

- $Cov\Big[\mu(\Theta_p,\, \Theta_{pj}),\, X_{qis}\Big] \;=\; \delta_{pq}(\delta_{ij}\cdot a \,+\, b)$

- $Cov\Big[\mu(\Theta_p),\, X_{qjw}\Big] \;=\; \delta_{pq}\cdot b$

- $Cov\Big[X_{pjs},\, X_{pjs'}\Big] \;=\; \dfrac{\delta_{ss'}\cdot S^2}{w_{pjs}} \,+\, a \,+\, b$

- $Cov\Big[X_{pjs},\, X_{pj's'}\Big] \;=\; \delta_{jj'}\cdot\left(\dfrac{\delta_{ss'}\cdot S^2}{w_{pjs}} \,+\, a\right) \,+\, b$

- $Cov\Big[X_{pjs},\, X_{qj's'}\Big] \;=\; 0 \qquad\qquad p \neq q$

A.3.4 Estimadors de credibilitat lineals

Un cop introduïts tots els elements, passem al problema que ens ocupa. El model jeràrquic de Jewell ens proporcionarà, d'una banda, estimadors de credibilitat lineals per $\mu(\Theta_p)$, pel nivell de les subcarteres; i d'una altra, per $\mu(\Theta_p,\, \Theta_{pj})$ en el nivell de les pòlisses.

Estem interessats en trobar els estimadors de credibilitat, no homogenis, per les primes de risc individuals i per a la seva obtenció, ens centrarem en una pòlissa concreta, la pj . Seguirem el mateix procediment que en el model de Bühlmann-Straub. En aquest cas, però, haurem d'aplicar aquest procediment dos cops, com ara veurem. Al tractarse d'un model a dos nivells, la pòlissa pj forma part de la cartera considerada a través de la subcartera p a la qual pertany, de manera que primer hem d'obtindre l'estimador de credibilitat per $\mu(\Theta_p,\, \Theta_{pj})$. Per això, el problema que es planteja és el següent:

$$\text{Min} \quad E\left[\left(\mu(\Theta_p,\, \Theta_{pj}) \,-\, c_0 \,-\, \sum_{i=1}^{k_p}\sum_{s=1}^{t_{pi}} c_{pis}\cdot X_{pis}\right)^2 \Big/\, \Theta_p\right] \tag{A.10}$$

Aquest proces de minimització té la mateixa solució que el trobat al model de Bühlmann-Straub, però restringit a la subcartera p , així doncs:

$$\hat{\mu}(\Theta_p,\, \Theta_{pj}) \;=\; z_{pj}\cdot X_{pjw} \,+\, (1 \,-\, z_{pj})\cdot m_p \tag{A.11}$$

Però per obtindre aquest estimador per la pòlissa pj , ens cal cercar la prima de risc per la subcartera p : $\mu(\Theta_p) = m$. Per trobar-la, únicament cal aplicar un altre cop el model de Bühlmann-Straub per estimar la prima de risc per la subcartera p dins del total de la cartera, on el problema a considerar és:

$$\text{Min} \quad E\left[\left(\mu(\Theta_p) - c_0 - \sum_{q=1}^{P}\sum_{j=1}^{k_p}\sum_{s=1}^{t_{qj}} c_{qjs} \cdot X_{qjs}\right)^2\right] \tag{A.12}$$

Seguint un procés similar al dels altres models vistos, arribem a trobar que els valors que minimitzen l'expressió [A.12] són:

$$c_0 = m \cdot (1 - z_p) \tag{A.13}$$

$$c_{pjs} = \frac{c_{pj} \cdot w_{pjs}}{w_{pj\Sigma}} \qquad c_{pj} = z_p \cdot \frac{z_{pj}}{z_{p\Sigma}} \tag{A.14}$$

Substituïm els valors en l'expressió lineal de la prima de risc, i obtenim:

$$\hat{\mu}(\Theta_p) = (1 - z_p) \cdot m + z_p \cdot X_{pzw} \tag{A.15}$$

Un cop trobat aquest resultat ja el podem substituir a la prima de risc per a la pòlissa pj , per tant:

$$\hat{\mu}(\Theta_p\,, \Theta_{pj}) = z_{pj} \cdot X_{pjw} + (1 - z_{pj}) \cdot \left(\frac{a \cdot m + b \cdot \sum_{j=1}^{k_p} z_{pj} \cdot X_{pjw}}{a + b \cdot z_{p\Sigma}}\right) \tag{A.16}$$

on:

$$z_{pj} = \frac{b \cdot w_{pj\Sigma}}{S^2 + a \cdot w_{pj\Sigma}}$$

A.3.5 Estimació dels paràmetres estructurals

Per poder calcular els dos estimadors trobats és necessari que estimem prèviament els paràmetres estructurals m , S^2 , a i b . Els estimadors no esbiaixats proposats per aquests paràmetres són els següents:

- Per l'esperança conjunta de la cartera:

$$\hat{m} = X_{zzw} = \sum_{p=1}^{P} \frac{z_p \cdot X_{pzw}}{z_\Sigma}$$

- Pel valor S^2 , que medeix la heterogeneïtat esperada de l'experiència de reclamacions en el temps:

$$\hat{S}^2 = \frac{\sum_{p=1}^{P} \sum_{j=1}^{k_p} \sum_{s=1}^{t_{pj}} w_{pjs} \cdot (X_{pjs} - X_{pjw})^2}{\sum_{p=1}^{P} \sum_{j=1}^{k_p} (t_{pj} - 1)}$$

- Per a que medeix la heterogeneïtat esperada dintre de les subcarteres:

$$\hat{a} = \frac{\sum_{p=1}^{P} \sum_{j=1}^{k_p} z_{pj} \cdot (X_{pjw} - X_{pzw})^2}{\sum_{p=1}^{P} (k_p - 1)}$$

- Per b que medeix la heterogeneïtat esperada entre les diferents subcarteres:

$$\hat{b} = \frac{\sum_{p=1}^{P} z_p \cdot (X_{pzw} - X_{zzw})^2}{P - 1}$$

A.4 Model de regressió de Hachemeister

El model que ens ocupa, és una generalització del model de Bühlmann-Straub. Hachemeister, C. (1975) es va plantejar la necessitat d'establir estimacions per les tendències que presentaven certes dades per l'efecte de la inflació. Per això, va proposar un estimador que fos una ponderació entre l'estimador individual i un estimador pel conjunt de les dades, objectiu final de tota la Teoria de la Credibilitat, tot utilitzant un model de regressió lineal. Ell no va formular explícitament les hipòtesis del model, va ser Norberg, R. (1979) qui les va formular i es va ocupar de l'estimació dels paràmetres estructurals.

El model de regressió de Hachemeister està molt lligat al model de Bühlmann-Straub, de fet hi està inspirat. Si en aquest darrer model ja es permet la variació en el temps de la $Var[X_{js} / \Theta_j]$, en el model de Hachemeister també es permesa la variació de $E[X_{js} / \Theta_j]$. Per tant, la hipòtesi d'homogeneïtat en el temps es substituïda per una altra de tipus polinòmic i d'aquesta manera permet tenir en compte els efectes de la inflació o d'altres tendències.

A.4.1 Variables rellevants

Les variables d'aquest model són:

- Θ_j : És el paràmetre de risc que descriu les característiques de la pòlissa j-èsima. Es tracta d'una variable aleatòria inobservable.

- X_{js} : Variable que indica l'experiència de reclamacions per la pòlissa j-èsima, en el període s-èsim,amb $j = 1, 2, \dots, k$ i $s = 1, 2, \dots, t_j$, éssent t_j el nombre de períodes observats per cada pòlissa.

- w_{pjs} : Variable que indica els pesos o ponderacions naturals de les variables observables X_{js} ; són nombres positius coneguts.

- Y_j : Matriu definida prèviament de dimensió (t_j, n) on s'ha de verificar que $n < t_j$. La n més una unitat indica el grau de la major tendència polinòmica prevista per a les diferents pòlisses que integren la cartera.

- v_j : matriu donada, de dimensió (t_j, t_j) i semidefinida positiva.

Només les dues últimes són novetat d'aquest model.

A.4.2 Hipòtesis del model

Les hipòtesis assumides per aquest model són:

(H.1) Les pòlisses $j = 1, 2, ..., k$, o sigui, els parells $(\Theta_1 , \vec{X}_1) , (\Theta_2 , \vec{X}_2) , ... ,$ $(\Theta_k , \vec{X}_k)$ són independents i estan idènticament distribuïts.

(H.2) Per a cada contracte $j = 1, 2, ..., k$, i tot $s = 1, 2, ..., t_j$ tenim:

(a) $E[X_j \,/\, \Theta_j] \;=\; Y_j \cdot \beta(\Theta_j) \qquad s \in (1 , 2 , ... , t_j)$

On $\beta(\Theta_j)$ és un vector de regressió desconegut de dimensió $(n, 1)$, Y_j és una matriu donada de dimensió (t_j, n) i X_j és un vector columna $(t_j, 1)$. La introducció de les matrius Y_j reemplaça la rígida assumpció d'homogeneïtat en el temps que, fins ara havíem suposat, per una altra de tipus polinòmic. Per exemple, quan $n = 2$ és de tipus lineal:

$$X_j = \begin{pmatrix} X_{j1} \\ X_{j2} \\ \vdots \\ X_{jt} \end{pmatrix} ; \qquad Y_j = \begin{pmatrix} 1 & t_j \\ 1 & t_j - 1 \\ \vdots & \vdots \\ 1 & 1 \end{pmatrix}$$

(b) $Cov[X_{js} 0\, \Theta_j] \;=\; v_j \cdot \sigma^2(\Theta_j) \qquad s \in (1 , 2 , ... , t_j)$

El model manté la independència dins i entre les pòlisses, com en el model de Bühlmann-Straub. En canvi, les observacions esperades ja no segueixen sent homogènies en el temps. És a dir, per a cada període s podem tenir diferent valor esperat. Com en el model de Bühlmann-Straub la variància segueix depenent del període en què ens situem, a través dels pesos o ponderacions naturals, ja que, encara que admet que v_j sigui una matriu semidefinida positiva amb covariàncies diferents de zero, a la llarga ens quedem amb la matriu diagonal dels pesos del model inspirador d'aquest:

$$Var[X_{js} \,/\, \Theta_j] \;=\; \frac{1}{w_{js}} \cdot \sigma^2(\Theta_j)$$

A.4.3 Estimació de la prima de risc individual

L'objectiu del model segueix sent estimar, de la millor manera possible, les primes de risc individual. Com ja hem apuntat, Hachemeister va creure que la millor manera era a través d'un model de regressió. Escollint aquest mètode, ja tenia solucionat, en part, el problema. Només havia de calcular els coeficients de regressió i la matriu de covariàncies. El seu plantejament va ser el següent: Estem interessats en estimar la prima de risc individual per la pòlissa j en el període s :

$$E[X_{js} \,/\, \Theta_j] \;=\; Y'_{js} \cdot \beta(\Theta_j) \;=\; \mu_{js}(\Theta_j)$$

La millor aproximació lineal per la prima de risc individual, l'obtenim calculant els valors a_0 i de a_{ir} amb $i = 1, ..., k$ i $r = 1, ..., t_i$ que minimitzin la següent esperança total:

$$E\left[\left(\mu_{js}(\Theta_j) \;-\; a_0 \;-\; \sum_{i=1}^{k}\sum_{r=1}^{t_i} a_{ir} \cdot X_{ir}\right)^2\right]$$

Un cop hem fet tot el procés de minimització, l'estimador de credibilitat lineal òptim es pot escriure com:

$$\hat{\mu}_{js}(\Theta_j) \;=\; \hat{\beta}(\Theta_j) \cdot Y_{js} \tag{A.17}$$

on

$$\hat{\beta}(\Theta_j) \;=\; \beta' \cdot (I \;-\; z_j) \;+\; \hat{\beta'}_j \cdot z_j \tag{A.18}$$

és un vector de dimensió ($1, n$) i se'l coneix amb el nom d'estimador de credibilitat del vector de regressió $\beta(\Theta_j)$ on:

$$\begin{aligned}
\beta \;&=\; E[\beta(\Theta_j)] \\
\hat{\beta'}_j \;&=\; X'_j v_j^{-1} Y_j \left(Y'_j v_j^{-1} Y_j\right)^{-1} \\
z_j \;&=\; \Gamma \cdot \left(\Gamma \;+\; S^2 w_j\right)^{-1} \\
\Gamma \;&=\; Cov\,[\beta(\Theta_j)\,,\, \beta'(\Theta_j)]
\end{aligned}$$

A.4.4 Estimació dels paràmetres estructurals

Fem servir com estimadors dels paràmetres estructurals β, S^2 i Γ els proposats per De Vylder, F. Aquests són els passos per obtindre'ls:

1. Donats $X_j(t_j, 1)$, $v_j(t_j, t_j)$ i $Y_j(t_j, n)$, primer es calculen:

$$\hat{\beta}_j = X_j' v_j^{-1} Y_j \left(Y_j' v_j^{-1} Y_j \right)^{-1}$$

$$S^2 = \frac{1}{k \cdot (t-n)} \sum_{j=1}^{k} \left(X_j - Y_j \hat{\beta}_j \right)' \cdot v_j^{-1} \cdot \left(X_j - Y_j \hat{\beta}_j \right)$$

2. Per poder calcular la resta d'estimadors, és necessari fer un procés iteratiu. Es dóna un valor inicial elevat per al paràmetre Γ. Un cop fixat aquest, es calcula $\hat{\beta}_z$ i a continuació $\hat{\Gamma}$ i així iterativament fins que el valor de Γ s'estabilitza. Els pseudoestimadors per a les variables que intervenen en aquest procés tenen els següents estimadors:

$$\hat{\beta}_z = \left(\sum_{j=1}^{k} z_j \right)^{-1} \sum_{j=1}^{k} z_j \hat{\beta}_j$$

$$\hat{\Gamma} = \frac{1}{k-1} \sum_{j=1}^{k} z_j \left(\hat{\beta}_j - \hat{\beta}_z \right) \cdot \left(\hat{\beta}_j - \hat{\beta}_z \right)'$$

A.5 Propietats dels estimadors de credibilitat

Dues propietats tenen els estimadors de credibilitat: la propietat de linealitat i la d'iteractivitat. La primera significa que l'estimador de credibilitat d'una combinació lineal de variables aleatòries és igual a la mateixa combinació lineal dels estimadors de credibilitat de cada una d'aquelles variables aleatòries separadament. Això es verifica en el següent teorema:

Teorema A.1 *Considerant certes variables aleatòries* $M_1, \ldots, M_N$ *, i uns coeficients fixos* $c_0, \ldots, c_N$ *. Llavors l'estimador de credibilitat de* $L = c_0 + c_1 M_1 +$

... + $c_N M_N$ *és igual a* $L^* = c_0 + c_1 M_1^* + ... + c_N M_N^*$ *, on* M_i^* *és l'estimador de credibilitat de* M_i $(i = 1, ..., N)$ *.*

Per altra banda, la propietat d'iteractivitat vol dir, parlant rudement, que si un estimador de credibilitat s'utilitza per predir un risc condicionat donat una certa variable aleatòria, aleshores l'estimador de credibilitat aplicat a aquest estimador de credibilitat condicional és igual a l'estimador de credibilitat no condicionat del risc. Vegem el teorema següent:

Teorema A.2 *Considerant el vector de variables aleatòries* $\vec{M}$ *per al qual l'estimador de credibilitat és* $\vec{M}^*$ *. Definim a més* $\vec{M}_\omega$ *com l'estimador de credibilitat de* $\vec{M}$ *donades unes realitzacions* ω *d'una certa variable aleatòria* Ω *. I finalment, definim* $\vec{M}_\Omega$ *com la variable aleatòria obtinguda al reemplaçar* ω *amb* ω *. Llavors, l'estimador de credibilitat de* $\vec{M}_\Omega$ *és igual a* $\vec{M}^*$ *.*

Apèndix B

Models I.B.N.R.

B.1 Mètode de Chain-Ladder

B.1.1 Dades

Assumim que el triangle "run-off" ens ve donat amb les entrades de C_{js}, la quantia acumulada pagada fins l'any s dels sinistres ocurreguts a l'any d'inici j.

Gràficament:

	1	2	3	$\dots$	$t-1$	t
1	C_{11}	C_{12}	C_{13}	$\dots$	$C_{1,t-1}$	C_{1t}
2	C_{21}	C_{22}	C_{23}	$\dots$	$C_{2,t-1}$	
$\vdots$	$\vdots$	$\vdots$	$\vdots$			
$t-2$	$C_{t-2,1}$	$C_{t-2,2}$	$C_{t-2,3}$			
$t-1$	$C_{t-1,1}$	$C_{t-1,2}$				
t	$C_{t,1}$					

B.1.2 Model i hipòtesis

El mètode de Chain-ladder és un procediment per completar l'anterior triangle i convertir-l'ho amb rectangle, per així, calcular la quantitat a reservar. La hipòtesi fonamental és que les columnes del triangle són proporcionals, a part de fluctuacions aleatòries. Per tant, si tenim que $C_{13} / C_{14} = 2/3$ també tindrem, aproximadament, que $C_{23} / C_{24} = C_{33} / C_{34} = ... = 2/3$.

Nosaltres podem dir que les hipòtesis impliquen que el triangle és estable en el transcurs del temps, és a dir, que el mètode no deixa possibilitat a canvis interns o externs que causin un canvi en l'estructura. Les causes que podrien alterar la hipòtesi d'homogeneïtat en el temps són, per exemple:

1. Un canvi administratiu en el tractament dels sinistres.

2. Un canvi de la jurisprudència, que faci elevar o disminuir les quanties dels sinistres.

3. Un increment de la inflació, que pot fer que la quantia dels sinistres retardats augmentin més que proporcionalment.

4. I d'altres

B.1.3 Càlculs

Tot el que necessitem, donat el triangle anterior, és un estimador del factor de proporcionalitat entre dues columnes. Definim $\hat{p}_{s,s+1}$ com el factor estimat per passar de la columna s a la columna $s+1$ i ve donat per:

$$\hat{p}_{s,s+1} = \frac{\sum_{j=1}^{t-s} C_{j,s+1}}{\sum_{j=1}^{t-s} C_{j,s}} \tag{B.1}$$

i per a completar el triangle, calculem els $\hat{C}_{js}\,(t-j < s \leq t)$ restants de la següent manera:

$$\hat{C}_{js} = C_{j,t-j+1} \cdot \prod_{s=t-j+1}^{j-1} \hat{p}_{s,s+1} \tag{B.2}$$

Si ens donen $C_{1\infty}$ els estimadors per $C_{j\infty}$ es calculen de la mateixa manera.

B.1.4 Exemple numèric

Donat el següent triangle:

	1	2	3	4	5
1	23.2	33.8	37.3	38.9	39.1
2	25.8	37.3	42.9	45.6	
3	22.1	30.3	30.7		
4	35.9	43.0			
5	34.9				

Per completar el triangle, en primer lloc s'han de calcular els factors de proporcionalitat, així, aplicant (B.1) trobem:

$$\begin{aligned}\hat{p}_{12} &= 1.350\\ \hat{p}_{23} &= 1.094\\ \hat{p}_{34} &= 1.054\\ \hat{p}_{45} &= 1.005\end{aligned}$$

I fent servir (B.2) es completa el triangle:

	1	2	3	4	5
1	23.2	33.8	37.3	38.9	39.1
2	25.8	37.3	42.9	45.6	45.8
3	22.1	30.3	30.7	32.2	32.5
4	35.9	43.0	47.0	49.6	49.8
5	34.9	47.1	51.5	54.3	54.6

Per exemple, per calcular $\hat{C}_{4,4}$:

$$\begin{aligned} \hat{C}_{4,4} &= C_{4,2} \cdot \hat{p}_{23} \cdot \hat{p}_{34} = \\ &= 43.0 \cdot 1.094 \cdot 1.054 = 49.6 \end{aligned}$$

Si es suposa que $C_{1\infty} = 39.3$, aleshores podem calcular aquesta columna multiplicant la darrera columna per 39.3/39.1.

	$C_{j\infty}$
1	39.3
2	46.1
3	32.7
4	50.1
5	54.8

B.2 Mètode dels mínims quadrats de De Vylder

B.2.1 Dades

Suposem donat el triangle "run-off" amb la variable c_{js} sense acumular. Per exemple, c_{js} és la quantia total pagada a l'any de desenvolupament s dels sinistres ocurreguts en l'any j . En aquest cas, no és necessari que la meitat superior del triangle sigui completa, això ens pot permetre eliminar dades dubtoses.

	1	2	3	$\ldots$	$t-1$	t
1	c_{11}	c_{12}	c_{13}	$\cdots$	$c_{1,t-1}$	c_{1t}
2	c_{21}	c_{22}	c_{23}	$\cdots$	$c_{2,t-1}$	
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$		
$t-2$	$c_{t-2,1}$	$c_{t-2,2}$	$c_{t-2,3}$			
$t-1$	$c_{t-1,1}$	$c_{t-1,2}$				
t	$c_{t,1}$					

B.2.2 Model i hipòtesis

El model està basat en l'aproximació que fem de c_{js} :

$$c_{js} = x_j \, p_s$$

on x_j és la quantia total de tots els sinistres ocurreguts a l'any j i p_s és la proporció fixa de la quantia anterior que correspon a l'any de desenvolupament s i per tant s'ha de cumplir la següent relació:

$$\sum_{s=1}^{t} p_s = 1$$

Del que es tracta és d'estimar aquestes dues variables a partir de les dades donades pel triangle.

Per tant, assumim que les columnes del triangle són proporcionals, tal com feiem en el mètode de chain-ladder, el ràtio de la columna h i la columna l és $p_h \, / \, p_l$.

De Vylder incorpora l'efecte d'una inflació constant en el seu model. Aquesta afecta als sinistres a través del corresponent any de pagament, per tant aproxima c_{js} per $x_j \, p_s \, \lambda^{j+s-2}$, on λ és el ràtio (constant) d'inflació. Així doncs,

$$x_j \, p_s \, \lambda^{j+s-2} = x'_j \, p'_s$$

amb

$$x'_j = x_j \lambda^{j-1} \qquad \text{i} \qquad y'_s = y_s \lambda^{s-1}$$

Ja veiem que no és possible estimar λ únicament amb les dades del triangle; però al menys el model no queda invalidat davant la presència d'un ràtio d'inflació constant.

B.2.3 Càlculs

Nosaltres hem de buscar x_i i p_j tals que minimitzin:

$$\sum (c_{js} - x_j p_s)^2$$

Les solucions d'aquest problema són les donades per les següents equacions:

$$x_j = \frac{\sum\limits_j c_{js} p_j}{\sum\limits_j p_j^2} \qquad \text{(B.3)}$$

$$p_s = \frac{\sum\limits_s c_{js} x_s}{\sum\limits_s x_s^2} \qquad \text{(B.4)}$$

on els sumatoris són definits només per les cel.les conegudes del triangle "run-off". Per completar el triangle podem procedir iterativament: començant per un arbitrari valor de p_s i aleshores utilitzant successivament (A4), (A3), (A4),... A la pràctica, sempre es garanteix una solució única.

B.2.4 Exemple numèric

Per l'exemple escollim el mateix triangle que hem utilitzat fins ara, però en la seva forma no acumulativa.

	1	2	3	4	5
1	23.2	10.6	3.5	1.6	0.2
2	25.8	11.5	5.6	2.7	
3	22.1	8.2	0.4		
4	35.9	7.1			
5	34.9				

Després de cinc iteracions i cumplint $\sum_s = 1$, obtenim:

	1	2	3	4	5
x_j	37.6	42.1	34.4	53.8	54.5
p_s	0.64	0.21	0.09	0.05	0.05

A partir d'aquests resultats calculem la resta del triangle:

	1	2	3	4	5
1	24.06	8.00	3.28	2.05	0.20
2	26.94	8.96	3.67	2.30	0.22
3	22.04	7.33	3.00	1.88	0.18
4	34.45	11.46	4.69	2.94	0.29
5	34.9	11.61	4.75	2.98	0.29

Per poder comparar amb el model anterior, mostrem la solució en la seva forma acumulada.

	1	2	3	4	5
1	24.1	32.1	35.3	37.4	37.6
2	26.9	35.9	39.6	41.9	42.1
3	22.0	29.4	32.4	34.2	34.4
4	34.5	45.9	50.6	53.5	53.8
5	34.9	46.5	51.3	54.2	54.5

B.3 Mètode de separació aritmètica de Taylor

B.3.1 Dades

Assumim que el triangle "run-off" ens ve donat amb les entrades de s_{js}, la mitja de sinistres. Si tenim en compte que c_{js} és la quantia pagada l'any s dels sinistres ocurreguts a l'any d'inici j, fixem-nos que no són quanties acumulades:

$$s_{js} = \frac{c_{js}}{n_j} \tag{B.5}$$

on n_j és el nombre total de sinistres ocurreguts a l'any j que estimem al final de l'any de desenvolupament $s = 1$. Gràficament:

	1	2	3	$\ldots$	$t-1$	t
1	s_{11}	s_{12}	s_{13}	$\ldots$	$s_{1,t-1}$	s_{1t}
2	s_{21}	s_{22}	s_{23}	$\ldots$	$s_{2,t-1}$	
$\vdots$	$\vdots$	$\vdots$	$\vdots$			
$t-2$	$s_{t-2,1}$	$s_{t-2,2}$	$s_{t-2,3}$			
$t-1$	$s_{t-1,1}$	$s_{t-1,2}$				
t	$s_{t,1}$					

B.3.2 Model i hipòtesis

Suposem que cada s_{js} , lliure de fluctuacions aleatòries, pot ser escrit com:

$$E[s_{js}] = r_s \cdot \lambda_{j+s} \tag{B.6}$$

on λ_{j+s} és un índex extern per a cada període de pagament $j+s$, que pot representar influències monetàries o d'inflació entre d'altres. En canvi, els r_s indiquen com es distribueixen els pagaments de sinistres al llarg dels anys de pagament $j+s$. Nosaltres suposem que aquesta distribució no depén del respectiu any d'origen j i per tant, la següent relació resulta:

$$\sum_{s=1}^{t} r_s = 1$$

Des de que construïm el triangle "run-off" a partir de l'estimació del nombre final de sinistres d'un any determinat d'origen (n_j), implícitament, assumim que tal estimació no és problemàtica. Per tant, tractem amb un assegurador que no es preocupa pels retards en la declaració dels sinistres, sinó que ho fa pel retard entre la declaració i el pagament dels sinistres. Per tant, aquesta situació és la d'un assegurador que es preocupa pels IBNER, no els IBNR.

Per altra banda, per completar el triangle i convertir-l'ho amb rectangle, també necessitem l'estimació de la inflació futura. El coneixement de la inflació passada ens pot ajudar en aquest punt.

Per últim, el mètode pot fracassar si l'any d'origen j té alguna influència en els pagaments del sinistres, ja que això no es contempla a (B.6). Exemples d'aquestes possibles influències són:

1. Un canvi administratiu en el tracrament dels sinistres.

2. Un canvi en la protecció oferta per les pòlisses, causant un canvi en el nivell de les quanties dels sinistres.

3. Un canvi en la freqüència dels sinistres.

4. I d'altres

Ademés, volem posar èmfasi sobre la fòrmula:

$$s_{js} = \frac{c_{js}}{n_j}$$

Com hem dit anteriorment, n_j és el nombre total de sinistres ocurreguts a l'any j estimat al final de l'any de desenvolupament $s = 1$. Per això, no haurem de reemplaçar-les per estimacions més recents o pel nombre exacte, si és possible, perquè pot causar un biaix si no es aplicat a tots els anys d'origen.

B.3.3 Càlculs

Per trobar els valors $\hat{\lambda}$ i $\hat{r}$ ho farem de manera recursiva d'una manera molt simple. Per començar, sigui v_s la suma de la columna s i d_s la suma respecte les diagonals del triangle "run-off":

$$v_s = \sum_{j=1}^{t-s+1} s_{js} \qquad \text{i} \qquad v_s = \sum_{j=1}^{s} s_{j,s-j+1} \tag{B.7}$$

I a continuació:

$$\begin{aligned}
\hat{\lambda}_t &= d_t \\
\hat{r}_t &= \frac{v_t}{\hat{\lambda}_t} \\
\hat{\lambda}_{t-1} &= \frac{d_{t-1}}{(1 - \hat{r}_t)} \\
\hat{r}_{t-1} &= \frac{v_{t-1}}{(\hat{\lambda}_t + \hat{\lambda}_{t-1})} \\
&\cdots \\
\hat{\lambda}_1 &= \frac{d_1}{(1 - \sum_{h=2}^{t} \hat{r}_h)} \\
\hat{r}_1 &= \frac{v_1}{(\sum_{h=1}^{t} \hat{\lambda}_h)}
\end{aligned}$$

Per completar el triangle:

$$\hat{s}_{js} = \hat{r}_s \cdot \hat{\lambda}_{j+s} \qquad (j = 1, ..., t;\ s = 1, ..., t - j + 1$$

B.3.4 Exemple numèric

Sigui el triangle original:

	1	2	3	4	5
1	88	43.6	51	54.2	15.6
2	93.2	45	64.2	54.8	
3	109	69.2	57.4		
4	122.4	63.4			
5	136.8				

Sumant columnes i diagonals tenim:

	1	2	3	4	5
v_j	549.4	221.2	172.6	108.9	15.6
d_j	88	136.8	205	309.9	328

Per tant, recursivament trobem:

	5	4	3	2	1
$\hat{\lambda}_j$	328	325.4	260.9	229.2	218.1
$\hat{r}_j$	0.048	0.167	0.189	0.193	0.403

que ens permet ajustar el triangle original:

	1	2	3	4	5
1	88	44.3	49.3	54.3	15.6
2	92.5	50.5	61.4	54.7	
3	105.3	63	61.9		
4	131.3	63.4			
5	132.3				

Per completar-l'ho, només ens cal extrapolar els valors passats de λ per mitjanes o per regressió log-lineal. Així doncs,

	6	7	8	9
$\hat{\hat{\lambda}}_j$	380.8	428	481	540.5

Aquests valors ens permeten convertir el triangle en rectangle:

	1	2	3	4	5
1	88	44.3	49.3	54.3	15.6
2	92.5	50.5	61.4	54.7	18.1
3	105.3	63	61.9	63.5	20.4
4	131.3	63.4	71.9	71.4	22.9
5	132.3	73.7	80.8	80.2	25.7