AKTUARSKA MATEMATIKA II

SADRŽAJ

Faculty and Institute of Actuaries zahvaljuje sljedećim osobama koje su pomogle u sastavljanju ovog materijala:

Claudia Chappell
Iain Currie
David Dickson
Roger Gray
Michael Hosking
Georgina Ivers
Athol Korabinski
Jim Tindale
Richard Verall
Howard Waters

POGLAVLJE 1 - TEORIJA ODLUČIVANJA

Nastavni ciljevi:

(i)1. Odrediti optimalne strategije po teoriji igara.

(i)2. Objasniti što se podrazumijeva pod funkcijom odlučivanja i funkcijom rizika.

(i)3. Primjeniti kriterije odlučivanja za odrediti koje funkcije odlučivanja su najbolje u odnosu na specificiran kriterij. Posebno razmatrati minimax kriterij i Bayesov kriterij.

0 Uvod

U donošenju odluka ili izbora čini se racionalnim izabrati opciju koja ima matematičko očekivanje “koje najviše obećava”, t.j., izabrati ili opciju koja minimizira očekivani gubitak ili maksimizira očekivani povrat.
Nažalost, ovaj jednostavni pristup donošenju odluka ne radi uvijek, budući da je često teško pridružiti bilo numeričku vrijednost rezultatu nekog ishoda, bilo vjerojatnosti tog ishoda.
Ovo poglavlje promatra teoriju igara i statističke igre kao uvod u teoriju odlučivanja.

1 Teorija igara

1.1 Igre zbroja nula s dva igrača

Za igru kažemo da je zbroja nula kada postoji konflikt među igračima, te štogod jedan igrač izgubi u igri, drugi igrač dobiva. Moguće je promatrati igre za više od dva igrača koje nisu zbroja nula, ali one izlaze iz okvira ovog poglavlja. U teoriji igara pretpostavlja se da u odabiru svoje strategije niti jedan igrač ne zna što će drugi igrač načiniti, i jednom kada je strategija odabrana ne može se mijenjati.
Igre se klasificiraju po broju strategija dostupnih svakom igraču. Na primjer, ako svaki igrač mora izabrati između dvije strategije, tada će to biti 2×2 igra. Ako jedan igrač ima 4 opcije, a drugi 5, tada će to biti 4×5 ili 5×4 igra.

Ovdje ćemo promatrati samo igre u kojima svaki igrač ima samo konačno mnogo opcija.

1.1.1 Oznake

Dva igrača označavat ćemo kao Igrač A i Igrač B.
Opcije Igrača A označavat ćemo s I,II,III ....
Opcije Igrača B označavat ćemo s 1,2,3 ....
Igra se može predstaviti matricom, npr., 2×2 igra može se predstaviti kao:

		Igrač A	
		I	II
Igrač B	1	L(I,I)	L(I,I)
	2	L(I,I)	L(I,I)

Elementi matrice zovu se isplate (payoffs) (t.j., gubitak za Igrača A i dobitak za Igrača B).
Iznosi $L(I,1), L(II,1), L(I,2)$ i $L(II,2)$ su vrijednosti funkcije gubitka određene igre. $L(I,1)$ predstavlja gubitak Igrača A (i dobitak Igrača B) kada Igrač A izabere opciju I, a Igrač B izabere opciju 1, itd.
Teorija igra pokušava naći optimalne strategije, t.j. one koje su najprofitabilnije za svakog igrača, i odgovarajuću isplatu ili vrijednost igre.

1.1.2 Sedlasta točka

Par strategija bit će u ravnoteži ako i samo ako je element $L(a,b)$ koji odgovara ravnoteži istovremeno najveći u stupcu i najmanji u retku. Takva strategija naziva se sedlastom točkom.
Na primjer, 3×3 matrica

		Igrač A		
		I	II	III
	1	6	2	2
Igrač B	2	4	3	5
	3	-3	1	0

ima sedlastu točku u drugom retku i drugom stupcu, ali 2×2 matrica

$$\begin{pmatrix} 2 & -2 \\ -2 & 2 \end{pmatrix}$$

nema sedlastih točaka.

1.1.3 Dominacija

U matrici isplate jedna strategija (predstavljena retkom ili stupcem) dominira drugu strategiju ako je izbor prve strategije barem toliko dobar kao izbor druge strategije, a u nekim slučajevima i bolji. Igrač A tražit će manji gubitak, a Igrač B veći dobitak. Dominirana strategija može se u igri uvijek odbaciti.
Na primjer u (4×3) igri

		Igrač A			
		I	II	III	IV
	1	2	0	1	4
Igrač B	2	1	2	1	4
	3	4	1	3	2

Kako Igrač A traži manji gubitak, strategija II dominira strategiju IV, jer će gubitak Igrača A uvijek biti manji ako je izabrana strategija II, a ne strategija IV. Stoga Igrač A nikada neće izabrati strategiju IV, te je možemo odbaciti. Igra se reducira na 3×3 igru

		Igrač A		
		I	II	III
	1	2	0	1
Igrač B	2	1	2	1
	3	4	1	3

Međutim, kako Igrač B traži veći dobitak, strategija 3 sada dominira strategiju 1, jer je dobitak Igrača B uvijek veći ako izabere strategiju 3, a ne strategiju 1. Stoga se igra reducira na 3×2 igru

		Igrač A		
		I	II	III
Igrač B	2	1	2	1
	3	4	1	3

2 Statističke igre

U statističkom zaključivanju se odluke o populaciji, kao što su očekivanje ili varijanca neke karakteristike, temelje na uzoračkim podacima. Statističko zaključivanje može se stoga promatrati kao igra između Prirode koja kontrolira relevantne osobine populacije i statističara koji pokušava donijeti odluke o populaciji.

Jedan način na koji se statističke igre razlikuju od teorije igara je da u teoriji igara svaki igrač izabire strategiju bez znanja o tome što će učiniti protivnik. U statističkoj igri statističar ima uzoračke podatke koji daju neku informaciju o izboru Prirode.

Primjer:

Statističaru je rečeno da novčić ima ili pismo na jednoj i grb na drugoj strani, ili dva grba. Statističar ne može pregledati novčić, ali može opažati jedno bacanje novčića i vidjeti da li je pao grb ili pismo. Statističar tada mora odlučiti da li novčić ima dva grba ili ne. Ako statističar donese krivu odluku plaća kaznu u iznosu £1. U slučaju ispravne odluke nema kazne (niti nagrade).

Ignorirajući činjenicu da statističar opaža jedno bacanje novčića, problem se može promatrati kao sljedeća igra

		Statističar (Igrač A)	
		a_1	a_2
Priroda	θ_1	0	1
(Igrač B)	θ_2	1	0

gdje je

θ_1 = "stanje prirode" je da novčić ima dva grba

θ_2 = "stanje prirode" je da je novčić balansiran

a_1 = odluka statističara je da novčić ima dva grba

a_2 = odluka statističara je da je novčić balansiran

Međutim, statističar zna što se dogodilo prilikom bacanj novčića, t.j., statističar zna da li je slučajna varijabla (recimo) x primila vrijednost $x = 0$ (grb) ili $x = 1$ (pismo). Statističar želi iskoristiti tu informaciju u izboru između a_1 i a_2, te treba funkciju odlučivanja koja određuje koju akciju poduzeti u svakom od slučaja.

$$d_1(x) = \begin{cases} a_1 \text{ kada je } x = 0 \\ a_2 \text{ kada je } x = 1 \end{cases}$$

ili $d_1(0) = a_1$ i $d_1(1) = a_2$.
Svrha indeksa je razlikovanje različitih funkcija odlučivanja. Druge moguće funkcije odlučivanja mogu biti

$$d_2(0) = a_1 \quad \text{i} \quad d_2(1) = a_1$$

t.j., uvijek izabrati a_1 bez obzira na ishod eksperimenta,

$$\text{ili} \quad d_3(0) = a_2 \quad \text{i} \quad d_3(1) = a_2$$
$$\text{ili} \quad d_4(0) = a_2 \quad \text{i} \quad d_4(1) = a_1$$

Ulazi u tablici daju odgovarajuće vrijednosti funkcije gubitka

		Statističar	
		a_1	a_2
Priroda	θ_1	$L(a_1, \theta_1)$	$L(a_2, \theta_1)$
	θ_2	$L(a_1, \theta_2)$	$L(a_2, \theta_2)$

Jedna opcija je izabrati a_1 kada je $x = 0$ i a_2 kada je $x = 1$, što se može iskazati kao

$$R(d_1, \theta_j) = \mathbb{E}[L(d_1(x), \theta_j)]$$

gdje je $L(d_1(x), \theta_j)$ funkcija gubitka u slučaju odluke $d_1(x)$ i strategije Prirode θ_j.
Očekivanje se računa u odnosu na slučajnu varijablu x i

$$\begin{array}{ll} \text{uz } \theta_1 & \mathbb{P}(x=0) = 1 \\ & \mathbb{P}(x=1) = 0 \\ & \\ \text{uz } \theta_2 & \mathbb{P}(x=0) = 1/2 \\ & \mathbb{P}(x=1) = 1/2 \end{array}$$

To daje

$$\begin{array}{lclclcl}
R(d_1,\theta_1) & = & 1\ L(a_1,\theta_1) + 0\ L(a_2,\theta_1) & = & 1\times 0 + 0 \times 1 & = & 0 \\
R(d_1,\theta_2) & = & 1/2\ L(a_1,\theta_2) + 1/2\ L(a_2,\theta_2) & = & 1/2\times 0 + 1/2 \times 0 & = & 1/2 \\
R(d_2,\theta_1) & = & 1\ L(a_1,\theta_1) + 0\ L(a_1,\theta_1) & = & 1\times 0 + 0 \times 0 & = & 0 \\
R(d_2,\theta_2) & = & 1/2\ L(a_1,\theta_2) + 1/2\ L(a_1,\theta_2) & = & 1/2\times 1 + 1/2 \times 1 & = & 1 \\
R(d_3,\theta_1) & = & 1\ L(a_2,\theta_1) + 0\ L(a_2,\theta_1) & = & 1\times 1 + 0 \times 1 & = & 1 \\
R(d_3,\theta_2) & = & 1/2\ L(a_2,\theta_2) + 1/2\ L(a_2,\theta_2) & = & 1/2\times 0 + 1/2 \times 0 & = & 0 \\
R(d_4,\theta_1) & = & 1\ L(a_2,\theta_1) + 0\ L(a_1,\theta_1) & = & 1\times 1 + 0 \times 0 & = & 1 \\
R(d_4,\theta_2) & = & 1/2\ L(a_2,\theta_2) + 1/2\ L(a_1,\theta_2) & = & 1/2\times 0 + 1/2 \times 1 & = & 1/2
\end{array}$$

To daje sljedeću 4×2 igru zbroja nula za dva igrača gdje su isplate jednake odgovarajućim vrijednostima funkcije rizika

		Statističar			
		d_1	d_2	d_3	d_4
Priroda	θ_1	0	0	1	1
	θ_2	1/2	1	0	1/2

Vidi se da d_1 dominira d_2, te da d_3 dominira d_4, tako da se d_2 i d_4 mogu odbaciti (kaže se da su nedopustive). To nije iznenađujuće, budući da one

impliciraju da se prihvati a_1 (odnosno da novčić ima dva grba) iako je opaženo pismo.
Igra se reducira na 2×2 igru zbroja nula za dva igrača.

		Statističar	
		d_1	d_3
Priroda	θ_1	0	1
	θ_2	1/2	0

Lako se može vidjeti da je Priroda zlonamjeran protivnik.
Optimalna strategija za statističara je randomizirana strategija s vjerojatnostima 2/3, odnosno 1/3, za d_1, odnosno d_3.
Vrijednost igre je 33p.
Ako se Prirodu ne promatra kao zlonamjernog protivnika, potreban je neki drugi kriterij za izbor između d_1 i d_3. To se diskutira u sljedećoj točki.

3 Kriteriji odlučivanja

Općenito je moguće pronaći najbolju funkciju odlučivanja samo u odnosu na neke kriterije. Razmatrat ćemo dva kriterija.

3.1 Minimax kriterij

Po minimax kriteriju se bira funkcija odlučaivanja za koju je $R(d, \theta)$, maksimiziran po θ, minimalan.
Primjenimo li minimax kriterij na primjer iz Odjeljka 2 sa d_2 i d_4, maksimalni rizik za d_1 je 1/2, a za d_3 je 1. Stoga d_1 minimizira maksimalni rizik.

3.2 Bayesov kriterij

Ako se Θ promatra kao slučajna varijabla, po Bayesovom kriteriju izabire se funkcija odlučivanja za koju je $\mathbb{E}[R(d, \theta)]$ minimalno, gdje se očekivanje računa u odnosu na Θ. Za kriterij je potrebno da se Θ promatra kao slučajna varijabla s danom distribucijom.

Primjena Bayesovog kriterija na primjer iz Odjeljka 2 zahtijeva pridruživanje vjerojatnosti strategijama Prirode θ_1 i θ_2. Ako je $\mathbb{P}(\theta_1) = p$ i $\mathbb{P}(\theta_2) = 1 - p$, tada je Bayesov rizik za

$$\begin{array}{lll} d_1 & \text{jednak} & 0.p + 1/2(1-p) = 1/2(1-p) \\ d_3 & \text{jednak} & 1.p + 0(1-p) = p \end{array}$$

Stoga kada je $p > 1/3$ Bayesov rizik od d_1 je manji od Bayesovog rizika za d_3, te se d_1 preferira pred d_3.
Kada je $p < 1/3$ Bayesov rizik od d_3 je manji od Bayesovog rizika za d_1, te se d_3 preferira pred d_1.
Kada je $p = 1/3$ oba Bayesova rizika su jednaka i bilo koji od d_1 i d_3 se može izabrati.

POGLAVLJE 2 - BAYESOVSKA STATISTIKA

Nastavni ciljevi:

(ii)1. Upotrijebiti Bayesov teorem za izračun jednostavnih uvjetnih vjerojatnosti.

(ii)2. Objasniti što se podrazumijeva pod apriori distribucijom, aposteriori distribucijom i konjugiranom apriori distribucijom.

(ii)3. Izvesti aposteriori distribuciju parametra u jednostavnim slučajevima.

(ii)4. Objasniti što se podrazmijeva pod funkcijom gubitka.

(ii)5. Upotrijebiti jednostavnu funkciju gubitka za izvod Bayesovskih procjena parametara.

0 Uvod

Bayesovska filozofija podrazumijeva potpuno drugačiji pristup statistici. Ovdje razmatramo Bayesovsku verziju procjene za osnovne situacije koje se tiču procjene parametara kada je dan slučajni uzorak iz određene distribucije. Klasična metoda uključjue metodu maksimalne vjerodostojnosti.
Fundamentalna razlika između Bayesovske i klasične metode je taj da se parametar θ promatra kao slučajna varijabla.
U klasičnoj statistici je θ fiksna, ali nepoznata veličina. To vodi do poteškoća kao što su pažljiva interpretacija potrebna za klasične intervale povjerenja, gdje je interval slučajan. Čim su opaženi podaci i izračunat je numerički interval, više nema vjerojatnosti. Ne može se izreći tvrdnja kao $\mathbb{P}(10.45 < \theta < 13.26) = 0.95$, jer θ nije slučajna varijabla.
U Bayesovskoj statistici ne pojavljuju se takve poteškoće i mogu se izricati vjerojatnosne tvrdnje o vrijednosti parametra θ.

1 Bayesov teorem

Ako $B_1, B_2, \ldots, B_k$ tvore potpun sustav događaja u prostoru elementarnih događaja S i $\mathbb{P}(B_i) \neq 0$ za $i = 1, 2, \ldots, k$, tada za svaki događaj A u S takav da $\mathbb{P}(A) \neq 0$

$$\mathbb{P}(B_r \mid A) = \frac{\mathbb{P}(A \mid B_r)\mathbb{P}(B_r)}{\mathbb{P}(A)} \quad \text{gdje je} \quad \mathbb{P}(A) = \sum_{i=1}^{k} \mathbb{P}(A \mid B_i)\mathbb{P}(B_i)$$

za $r = 1, 2, \ldots, k$.

1.1 Primjer

Proizvođači dostavljaju odjeću trgovcu na malo. 60% robe dolazi od proizvođača 1, 30% od proizvođača 2 i 10% od proizvođača 3. 10% odjeće proizvođača 1 ima grešku, 5% od proizvođača 2 ima grešku i 15% od proizvođača 3 ima grešku. Kolika je vjerojatnost da komad odjeće s greškom dolazi od proizvođača 3?
Rješenje:
Neka je A događaj da komad odjeće ima grešku. Neka je B_i događaj da komad odjeće dolazi od proizvođača i.
Uvrstimo li brojke u Bayesovu formulu,

$$\mathbb{P}(B_3 \mid A) = \frac{(0.15)(0.1)}{(0.1)(0.6) + (0.05)(0.3) + (0.15)(0.1)} = \frac{0.015}{0.09} = 0.167$$

Iako proizvođač 3 dostavlja samo 10% odjeće trgovcu na malo, gotovo 17% robe s greškom dolazi od tog proizvođača.

2 Apriori i aposteriori distribucije

Pretpostavimo da je $\underline{X} = (X_1, X_2, \ldots, X_n)$ slučajni uzorak iz populacije specificirane vjerojatnosnom funkcijom gustoće $f(x; \theta)$ i potrebno je procijeniti θ.
Budući da je θ slučajna varijabla, imat će distribuciju. To dozvoljava upotrebu bilo kakvog dostupnog znanja o vrijednosti od θ prije skupljanja bilo kakvih podataka. To znanje kvantificiramo izražavajući ga kao apriornu distribuciju od θ.
Nakon prikupljanja odgovarajućih podataka, određuje se aposteriorna distribucija od θ i to tvori temelj zaključivanja o θ.

2.1 Notacija

Budući daje θ slučajna varijabla, u stvari bismo je trebali označavati velikim slovom Θ, a njenu apriornu distribuciju kao $f_\Theta(\theta)$. Međutim, zbog jednostavnosti nećemo praviti razliku između Θ i θ, a gustoću ćemo jednostavno označavati s $f(\theta)$. Uočite da govoreći o gustoći impliciramo da je θ neprekidna. U većini primjena to će biti slučaj, jer čak kada je X diskretna (kao binomna ili Poissonova), parametar (p ili λ) mijenjat će se u neprekidnom prostoru ($(0, 1)$, odnosno $(0, \infty)$).
Također će gustoća populacije ili vjerojatnosna funkcija biti označena s $f(x \,|\, \theta)$ umjesto ranijeg $f(x; \theta)$, budući da predstavlja uvjetnu distribuciju od X uz dano θ.

2.2 Određivanje aposteriorne gustoće

Pretpostavimo da je $\underline{X}$ slučajni uzorak iz populacije specificirane s $f(x \,|\, \theta)$, te da θ ima apriornu gustoću $f(\theta)$.
Aposteriorna gustoća od $\theta \,|\, \underline{X}$ određuje se primjenom osnovne definicije uvjetne gustoće:

$$f(\theta \,|\, \underline{X}) = \frac{f(\theta, \underline{X})}{f(\underline{X})} = \frac{f(\underline{X} \,|\, \theta) f(\theta)}{f(\underline{X})}\,.$$

Uočite da je $f(\underline{X}) = \int f(\underline{X} \,|\, \theta) f(\theta)\, d\theta$. Taj rezultat je neprekidna verzija Bayesovog teorema iz osnovne vjerojatnosti.
Često je prikladnije izraziti taj rezultat pomoću vrijednosti statistike kao što je $\overline{X}$ umjesto vrijednosti uzorka $\underline{X}$. Tako, na primjer,

$$f(\theta \,|\, \overline{X}) = \frac{f(\theta, \overline{X})}{f(\overline{X})} = \frac{f(\overline{X} \,|\, \theta) f(\theta)}{f(\overline{X})}\,.$$

U praksi je to ekvivalentno.
Koristan način da se izrazi aposteriorna gustoća je upotreba proporcionalnosti. $f(\underline{X})$ ne uključuje θ, te je samo konstanta potrebna za pravu gustoću čiji je integral 1, tako da je

$$f(\theta \,|\, \underline{X}) \propto f(\underline{X} \,|\, \theta) f(\theta)\,.$$

Također uočite da je $f(\underline{X} \,|\, \theta)$ zajednička gustoća uzorka, te tako ništa drugo nego vjerodostojnost. Prema tome je aposteriorna gustoća proporcionalna produktu vjerodostojnosti i apriorne gustoće.

Za danu vjerodostojnost, ako apriorna distribucija vodi do aposteriorne distribucije koja pripada istoj familiji kao i apriorna distribucija, tada se apriorna distribucija naziva konjugiranom apriornom vjerodostojnosti.

3 Funkcija gubitka

Da bismo dobili procjenitelj za θ, najprije moramo specificirati funkciju gubitka. To je mjera "gubitka" koji nastaje upotrebom $g(\underline{X})$ kao procjenitelja za θ. Traži se funkcija gubitka koja je nula u slučaju točne procjene, t.j., $g(\underline{X}) = \theta$, i koja je pozitivna i neopadajuća kada se $g(\underline{X})$ udaljava od θ. Postoji jedna vrlo često upotrebljavana funkcija gubitka koja se naziva gubitkom kvadratne greške. Dvije druge se također upotrebljavaju u praksi.
Tada je Bayesov procjenitelj onaj $g(\underline{X})$ koji minimizira očekivani gubitak s obzirom na aposteriornu distribuciju.
Glavna funkcija gubitka je kvadratni gubitak definirana s

$$L(g(\underline{x}), \theta) = [g(\underline{x}) - \theta]^2 \,,$$

i povezana je sa srednjom kvadratnom greškom iz klasične statistike.
Druga funkcija gubitka je gubitak apsolutne greške definiran s

$$L(g(\underline{x}), \theta) = |g(\underline{x}) - \theta| \,.$$

Treća funkcija gubitka je "0/1" ili "sve ili ništa" gubitak definiran s

$$L(g(\underline{x}), \theta) = \begin{cases} 0 & \text{ako je } g(\underline{x}) = \theta \\ 1 & \text{ako je } g(\underline{x}) \neq \theta \end{cases}$$

Bayesovi procjenitelji koji se dobivaju minimiziranjem očekivanog gubitka za svaku of tih funkcija redom su očekivanje, medijan i mod aposteriorne distribucije, svaki od kojih je mjera lokacije aposteriorne distribucije.
Očekivani aposteriorni gubitak (*E*xpected *P*osterior *L*oss) je

$$EPL = \mathbb{E}\{L(g(\underline{x}), \theta)\} = \int L(g(\underline{x}), \theta) f(\theta \,|\, x) \, d\theta \,.$$

3.1 Kvadratni gubitak

Zbog jednoatavnosti pisat ćemo g umjesto $g(\underline{x})$.

$$\text{Dakle } EPL = \int (g-\theta)^2 f(\theta\,|\,x)\,d\theta$$
$$\frac{d}{dg}EPL = 2\int (g-\theta) f(\theta\,|\,x)\,d\theta$$
$$\text{izjednačimo s nulom } \Rightarrow g\int f(\theta\,|\,x)\,d\theta = \int \theta f(\theta\,|\,x)\,d\theta$$
$$\text{ali } \int f(\theta\,|\,x)\,d\theta = 1 \quad \therefore g = \int \theta f(\theta\,|\,x)\,d\theta = \mathbb{E}(\theta\,|\,x)\,.$$

To očito minimizira EPL.
$\therefore$ Bayesovski procjenitelj uz kvadratni gubitak je očekivanje aposteriorne distribucije.

3.2 Gubitak apsolutne greške

Opet ćemo umjesto $g(\underline{x})$ pisati g.

$$\text{Dakle } EPL = \int |g-\theta| f(\theta\,|\,\underline{x})\,d\theta$$

Pretpostavimo li da je θ u $(-\infty,\infty)$, tada je

$$EPL = \int_{-\infty}^{g} (g-\theta) f(\theta\,|\,\underline{x})\,d\theta + \int_{g}^{\infty} (\theta-g) f(\theta\,|\,\underline{x})\,d\theta$$
$$\therefore \frac{d}{dg}EPL = \int_{-\infty}^{g} f(\theta\,|\,\underline{x})\,d\theta - \int_{g}^{\infty} f(\theta\,|\,\underline{x})\,d\theta$$
$$\text{[Prisjetimo se } \frac{d}{dy}\int_{a(y)}^{b(y} f(x,y)\,dx = \int_{a(y)}^{b(y)} \frac{\partial}{\partial y} f(x,y)\,dx + b'(y)f(b(y),y) - a'(y)f(a(y),y)]$$
$$\text{izjednačimo s nulom } \Rightarrow \int_{-\infty}^{g} f(\theta\,|\,\underline{x})\,d\theta = \int_{g}^{\infty} f(\theta\,|\,\underline{x})\,d\theta$$
$$\text{to jest, } \mathbb{P}(\theta \le g) = \mathbb{P}(\theta \ge g)\,,$$

što specificira medijan aposteriorne distribucije.

3.3 Gubitak sve ili ništa

Ovdje ne možemo upotrijebiti pristup preko deriviranja. Umjesto toga koristit ćemo direktan pristup s graničnim argumentom.

Promotimo

$$L(g(\underline{x}, \theta) = \begin{cases} 0 & \text{ako je } g - \epsilon < \theta < g + \epsilon \\ 1 & \text{inače,} \end{cases}$$

tako da u limesu kada $\epsilon \to 0$ to teži traženoj funkciji gubitka.

$$EPL = 1 - \int_{g-\epsilon}^{g+\epsilon} f(\theta \,|\, \underline{x})\, d\theta = 1 - 2\epsilon f(g \,|\, \underline{x}) \quad \text{za male } \epsilon\,.$$

To se minimizira ako uzmemo da je g jednak modu od $f(\theta \,|\, \underline{x})$.

3.4 Primjer

Za procjenu binomne vjerojatnosti θ iz jednog opažanja X s apriornom distribucijom od θ koja je beta distribucija s parametrima α i β, ispitajte oblik aposteriorne distribucije od θ i odredite Bayesovski procjenitelj od θ za kvadratni gubitak.
Rješenje:
Koristit ćemo pristup preko proporcionalnosti i ispustiti sve konstante kada je to prikladno.

$$\text{Apriorna: } f(\theta) \propto \theta^{\alpha-1}(1-\theta)^{\beta-1}, \text{ispuštena konstanta } \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\,.$$

$$\text{Vjerodostonost: } f(X \,|\, \theta) \propto \theta^{x}(1-\theta)^{n-x}, \text{ispuštena konstanta } \binom{n}{x}\,.$$

$$\therefore f(\theta \,|\, X) \propto \theta^{x}(1-\theta)^{n-x} \cdot \theta^{\alpha-1}(1-\theta)^{\beta-1} = \theta^{x+\alpha-1}(1-\theta)^{n-x+\beta-1}\,.$$

Vidi se da je, do na odgovarajuću konstantu proporcionalnosti, ovo gustoća beta slučajne varijable. Slijedi neposredan zaključak da je aposteriorna distribucija od θ uz dano X beta distribucija s parametrima $(x+\alpha)$ i $(n-x+\beta)$. Vidimo da aposteriorna distribucija i apriorna distribucija pripadaju istoj familiji distribucija. Stoga je beta distribucija konjugirana apriorna distribucija za binomnu distribuciju.
Bayesovski procjenitelj za kvadratni gubitak je očekivanje te distribucije, to jest,

$$\frac{x+\alpha}{(x+\alpha)+(n-x+\beta)} = \frac{x+\alpha}{n+\alpha+\beta}\,.$$

POGLAVLJE 3 - DISTRIBUCIJE ŠTETA

Nastavni ciljevi:

(iii)1. Opisati svojstva statističkih distribucija prikladnih za modeliranje individualnih i skupnih (agregatnih) šteta.

(iii)2. Izvesti momente i funkcije izvodnice momenata (kada postoje) distribucija šteta uključujući gama, eksponencijalnu, Paretovu, generaliziranu Paretovu, normalnu, lognormalnu, Weibullovu i Burrovu distribuciju.

(iii)3. Primjeniti principe statističkog zaključivanja za izbor prikladne distribucije štete za skup šteta.

(iii)4. Objasniti pojmove franšize i granice samopridržaja.

(iii)5. Opisati operacije jednostavnih oblika kvotnog reosiguranja i reosiguranja viška šteta.

(iii)6. Izvesti distribuciju i odgovarajuće momente iznosa šteta koje isplaćuju osiguratelj i reosiguratelj kod franšiza i reosiguranja.

0 Uvod

0.1 Osigurateljni okvir

Ukupni iznos šteta u određenom vremenskom periodu je veličina od fundamentalnog značaja za pravilno vođenje osigurateljnog društva. Ključna pretpostavka svih ovdje proučavanih modela je da se pojavljivanje šteta i iznosi šteta mogu studirati odvojeno. Prema tome, šteta nastaje po nekom jednostavnom modelu vremenskog pojavljivanja događaja, a tada se iznos štete bira iz distribucije koja opisuje iznos štete.

0.2 Statistička pozadina

Za opis distribucija slučajnih varijabli upotrebljava se niz statističkih tehnika. Cilj je opisati varijaciju u iznosima šteta nalaženjem distribucije šteta koja adekvatno opisuje stvarne štete koje se pojavljuju. Uobičajeno se to može napraviti na dvije razine.

Na prvoj razini može se pretpostaviti da se štete pojavljuju kao realizacije iz poznate distribucije. Na primjer, moguće je pretpostaviti da logaritam iznosa šteta slijedi, do razumne aproksimacije, normalnu distribuciju s poznatim očekivanjem i poznatom standardnom devijacijom. Poznavanje procesa iznosa šteta bilo bi tada potpuno, i zanimanje bi se tada usredotočilo na posljedice u osiguranju. Na primjer, štete iznad izvjesne razine mogle bi pokrenuti neke reosigurateljne sporazume ili se štete ispod izvjesne razine ne bi prijavile ukoliko je na polici ugovorena franšiza.
Točne distribucije šteta skoro nikad neće biti poznate u praksi. Na toj drugoj razini standardna je metoda pretpostaviti da je distribucija šteta član neke familije. Sada se moraju procijeniti parametri familije korištenjem podataka o iznosima šteta pomoću odgovarajuće metode kao što je metoda maksimalne vjerodostojnosti. Komplikacije će se pojaviti ukoliko su velike štete ograničene (reosiguranje) ili se neke male štete nisu prijavile (franšiza).
Provedena su mnoga proučavanja o vrsti distribucije koja se može koristiti za opis varijacija u iznosima šteta. Opći zaključak je da su distribucije šteta sklone imati pozitivnu asimetričnost i dugi rep.

1 Procjena i prilagodba

Metode maksimalne vjerodostojnosti, momenata i percentila mogu se koristiti za prilagodbu distribucija podacima. Prilagodba distribucije može se formalno testirati upotrebom χ^2 testa. Metoda percentila skicirana je u Sekciji 2.3.; druge metode i χ^2 test pokrivene su u Predmetu C1. Formule za gustoće, momente i funkcije izvodnice momenata (kada postoje) za distribucije iz ovog poglavlja dane su u *Formulae and Tables for Actuarial Examinations.*

1.1 Eksponencijalna distribucija

Slučajna varijabla X ima eksponencijalnu distribuciju s parametrom λ ako je

$$F(x) = 1 - e^{-\lambda x}$$

i pišemo $X \sim \exp(\lambda)$.
Moguća je upotreba metode maksimalne vjerodostojnosti (ML) ili metode momenata za procjenu parametra eksponencijalne distribucije.

1.2 Paretova distribucija

Slučajna varijabla X ima Paretovu distribuciju s parametrima α i λ ako je

$$F(x) = 1 - \left(\frac{\lambda}{\lambda + x}\right)^{\alpha}$$

i pišemo $X \sim Pa(\alpha, \lambda)$.
Lagano se provjeri da Paretova distribucija ima vjerojatnosnu funkciju gustoće

$$f(x) = \frac{\alpha\lambda^{\alpha}}{(\lambda + x)^{\alpha+1}}, \ x > 0, \alpha > 0, \lambda > 0\,.$$

U slučaju Paretove distribucije lagano je primjeniti metodu momenata, ali će na taj način dobivene procjene imati prilično velike standardne greške, uglavnom stoga što S^2, varijanca uzorka, ima vrlo veliku varijancu. Međutim, metoda daje početne procjene za efikasnije metode procjene koje možda nije tako lako primjeniti, kao metodu maksimalne vjerodostojnosti, gdje je potrebno koristiti numeričke metode.

1.3 Weibullova distribucija

Paretova distribucija je distribucija s gornjim repom koji teži k nuli kao potencija od x. To daje distribuciju s puno težim repom od eksponencijalne. Izrazi za gornje repove eksponencijalne i Paretove distribucije su

$$\begin{array}{ll} \text{eksponencijalna} & \mathbb{P}(X > x) = \exp(-\lambda x) \\ \\ \text{Paretova} & \mathbb{P}(X > x) = (\lambda/(\lambda + x))^{\alpha}\,. \end{array}$$

Postoje daljnje mogućnosti. Stavimo

$$\mathbb{P}(X > x) = \exp(-\lambda x^{\gamma}), \ \gamma > 0\,.$$

Sada imamo dva slučaja. Ako je $\gamma < 1$, dobit ćemo distribuciju s težinom repa između eksponencijalne i Paretove distribucije, dok ćemo za $\gamma > 1$ dobiti distribuciju s lakšim repom od eksponencijalne ($\gamma = 1$ je eksponencijalna distribucija). Repna distribucija definira Weibullovu distribuciju, vrlo fleksibilnu distribuciju koja se može upotrijebiti kao model za štete u osiguranju, obično sa $\gamma < 1$. Slučajna varijabla X ima Weibullovu distribuciju s parametrima c i γ ako je

$$F(x) = 1 - \exp(-cx^{\gamma})$$

i pišemo $X \sim W(c, \gamma)$. (Uočite promjenu s λ na c; to je notacija korištena u *for Actuarial Examination*). Vjerojatnosna funkcija gustoće od $W(c, \gamma)$ distribucije je

$$f(x) = c\gamma x^{\gamma-1} \exp(-cx^{\gamma}),\ x > 0, c > 0, \gamma > 0\,.$$

Nije elementarno primjeniti niti metodu momenata niti metodu maksimalne vjerodostojnosti ako su i c i γ nepoznati (iako, ako je dostupno računalo, kao što je slučaj u praksi, jednadžbe su dovoljno jednostavne). U slučaju kada γ ima poznatu vrijednost γ^*, metoda maksimalne vjerodostojnosti je dovoljno lagana.

Funkcija distribucije $W(c, \gamma)$ distribucije je elementarna funkcija, te se na tome može zasnivati jednostavna metoda procjene za c i γ. Metoda se sastoji od izjednačavanja izabranih uzoračkih percentila i funkcije distribucije; na primjer, izjednačite kvartile, 25-ti i 75-ti percentil, s percentilima populacije. To odgovara načinu kako se izjednačavaju momenti uzorka i momenti populacije kod metode momenata. O toj metodi govorit ćemo kao o metodi percentila.

Kod metode momenata koriste se prva dva momenta ako postoje dva nepoznata parametra, i to se čini intuitivno razumnim (iako teorijska osnova za to nije tako čista). Na sličan način, ukoliko treba procijeniti jedan parametar koristit će se medijan. S dva parametra, procedura je nejasnija, ali se donji i gornji kvartil čine razboritim izborom.

Primjer

Procijenite c i γ u Weibullovoj distribuciji upotrebom metode percentila, gdje je prvi kvartil uzorka 401, a treći kvartil uzorka 2836.75.

Rješenje

Dvije jednadžbe za c i γ su

$$\begin{aligned} F(401) &= 1 - \exp(-c\,401^{\gamma}) &&= 0.25 \\ F(2836.75) &= 1 - \exp(-c\,2836.75^{\gamma}) &&= 0.75 \end{aligned}$$

koje se mogu prepisati kao

$$\begin{aligned} -c\,401^{\gamma} &= \log(3/4) \\ -c\,2836.75^{\gamma} &= \log(1/4)\,. \end{aligned}$$

Dijeljenjem nalazimo $\tilde{\gamma} = 0.8038$, te stoga $\tilde{c} = 0.002326$, gdje $\sim$ označava percentilni procjenitelj. Uočimo da je $\tilde{\gamma}$ manje od 1, što indicira deblji rep nego kod eksponencijalne distribucije.

1.4 Gama distribucija

Slučajna varijabla X ima gamu distribuciju s parametrima α i λ ako je

$$f(x) = \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} \exp(-\lambda x)\,, x > 0,$$

što pišemo $X \sim G(\alpha, \lambda)$. Očekivanje i varijanca od X su

$$\mathbb{E}(X) = \frac{\alpha}{\lambda}; \quad \mathrm{Var}(X) = \frac{\alpha}{\lambda^2}\,.$$

Budući da momenti imaju jednostavan oblik, vrlo je lako primjeniti metodu momenata. MLE (procjenitelj maksimalne vjerodostojnosti) za gama distribuciju ne može se dobiti u zatvorenom obliku (tj., pomoću elementarnih funkcija), ali se procjenitelji momenata mogu upotrijebiti kao početne procjene u potrazi za procjenama maksimalne vjerodostojnosti.
Pogodnije je dobiti procjene maksimalne vjerodostojnosti za gamu distribuciju upotrebom drugačije parametrizacije. Stavimo $\mu = \alpha/\lambda$ i procijenimo parametre α i μ. Tada dobijemo MLE za λ kao $\hat{\lambda} = \hat{\alpha}/\hat{\mu}$. Tu se koristi svojstvo invarijantnosti procjenitelja maksimalne vjerodostojnosti.

1.5 Lognormalna distribucija

Definicija lognormalne distribucije vrlo je jednostavna: X ima lognormalnu distribuciju ako $\log(X)$ ima normalnu distribuciju. Kada je $\log(X) \sim N(\mu, \sigma^2)$, $X \sim LN(\mu, \sigma^2)$.
Procjena za lognormalnu distribuciju je jednostavna, budući da se μ i σ^2 mogu procijeniti upotrebom log transformiranih podataka (tj. logaritama podataka). Neka su $x_1, x_2, \ldots, x_n$ opažene vrijednosti, i stavimo $y_i = \log x_i$. Procjenitelji maksimalne vjerodostojnosti za μ i σ^2 su $\bar{y}$ i s_y^2, gdje indeks y označava da je varijanca uzorka izračunata za y vrijednosti.

1.6 Miješane distribucije

Eksponencijalna distribucija je jedan od najjednostavnijih modela za osigurateljne štete. Pretpostavimo da za svaku osobu u velikom osigurateljnom portfelju štete nastaju po nekoj eksponencijalnoj distribuciji. Praktično znanje skoro svakog osigurateljnog portfelja otkriva da će se očekivanja tih raznih distribucija razlikovati među osiguranicima. Prema tome je opis šteta

u portfelju taj da svaka šteta slijedi svoju eksponencijalnu distribuciju, tj., da eksponencijalne distribucije imaju očekivanja koja se razlikuju od osobe do osobe.
Sada je potrebno pronaći opis varijacije među pojedinačnim očekivanjima. Jedan način da se to učini je pretpostavka da očekivanja eksponencijalne distribucije sama slijede neku distribuciju. U eksponencijalnom slučaju prikladna je sljedeća pretpostavka. Neka je $\lambda_i = 1/\theta_i$ recipročna vrijednost srednje štete i-tog osiguranika. Pretpostavimo da se varijacija među λ_i može opisati poznatom gama distribucijom $G(\alpha, \delta)$. tj., pretpostavimo $\lambda \sim G(\alpha, \delta)$ gdje je

$$f(\lambda) = \frac{\delta^\alpha}{\Gamma(\alpha)} \lambda^{\alpha-1} \exp(-\delta\lambda) \, , \lambda > 0 \, .$$

Posebno uočite da je to vjerojatnosna funkcija gustoće po λ s poznatim vrijednostima α i δ.
Ova formulacija ima puno zajedničkog s onom korištenom kod Bayesovske procjene. Zaista, fundamentalna ideja u Bayesovskoj procjeni je da se parametar koji nas zanima (ovdje λ) može tretirati kao slučajna varijabla s poznatom distribucijom. Uočite, međutim, da ovdje nije svrha procjeniti individualni λ_i, već opisati ukupne štete po cijelom portfelju. Kada se o $G(\alpha, \delta)$ govori kao o apriornoj distribuciji, Bayesovska procjena može se koristiti za procjenu individualnih λ_i. U ovom problemu opisa šteta po cijelom portfelju, $G(\alpha, \delta)$ distribucija koristi se za usrednjenje eksponencijalnih distribucija; o $G(\alpha, \delta)$ govori se kako o distribuciji miješanja, a o rezultirajućoj distribuciji štete kao o miješanoj distribuciji.
Marginalna distribucija od X je

$$\begin{aligned}
f_X(x) &= \int f_{X,\lambda}(x, \lambda) \, d\lambda \\
&= \int f_\lambda(\lambda) f_{X|\lambda}(x|\lambda) \, d\lambda \\
&= \int_0^\infty \frac{\delta^\alpha}{\Gamma(\alpha)} \lambda^{\alpha-1} \exp(-\delta\lambda) \times \lambda \exp(-\lambda x) \, d\lambda \\
&= \frac{\delta^\alpha}{\Gamma(\alpha)} \int_0^\infty \lambda^\alpha \exp\{-(x+\delta)\lambda\} \, d\lambda \\
&= \frac{\delta^\alpha}{\Gamma(\alpha)} \times \frac{\Gamma(\alpha+1)}{(x+\delta)^{\alpha+1}} \quad (G(\alpha+1, x+\delta) \text{ integral}) \\
&= \frac{\alpha\delta^\alpha}{(x+\delta)^{\alpha+1}} \, , \; x > 0,
\end{aligned}$$

što se može prepoznati kao Paretova distribucija $Pa(\alpha, \delta)$. Ovaj rezultat daje lijepu interpretaciju Paretove distribucije: $Pa(\alpha, \delta)$ pojavljuje se kada se eksponencijalne distribucije šteta usrednjuju pomoću $G(\alpha, \delta)$ distribucije miješanja.

1.7 Poopćenja Paretove distribucije

Funkcija distribucije Paretove distribucije $Pa(\alpha, \delta)$ je

$$F(x) = 1 - \frac{\lambda^\alpha}{(\lambda + x)^\alpha}\,.$$

Može se uvesti daljnji parametar γ stavivši

$$F(x) = 1 - \frac{\lambda^\alpha}{(\lambda + x^\gamma)^\alpha}\,.$$

To je funkcija distribucije transformirane Paretove ili Burrove distribucije. Dodatni parametar daje posebnu fleksibilnost kada je potrebna prilagodba podacima. Budući da funkcija distribucije postoji u zatvorenom obliku, moguće je prilagoditi Burrovu distribuciju podacima upotrebom metode percentila. Metoda maksimalne vjerodostojnost će sigurno zahtijevati upotrebu kompjutorskog softwarea koji omogućuje nelinearnu optimizaciju.
Drugo poopćenje Paretove distribucije koristi prije razmatranu ideju miješane distribucije. Ako su štete eksponencijalne s očekivanjem $1/\lambda$ i $\lambda \sim G(\alpha, \delta)$, tada je marginalna distribucija šteta $Pa(\alpha, \delta)$. To se može poopćiti ako se pretpostavi da su štete $G(k, \lambda)$ i $\lambda \sim G(\alpha, \delta)$. Ako je $k = 1$, tada je jasno da se dobije Paretova distribucija. Za općenit k, marginalna distribucija štete, X, je

$$\begin{aligned}
f_X(x) &= \int f_\lambda(\lambda) f_{X|\lambda}(x|\lambda)\, d\lambda \\
&= \int \frac{\delta^\alpha}{\Gamma(\alpha)} \lambda^{\alpha-1} \exp(-\delta\lambda) \times \frac{\lambda^k}{\Gamma(k)} x^{k-1} \exp(-\lambda x)\, d\lambda \\
&= \frac{\delta^\alpha x^{k-1}}{\Gamma(\alpha)\Gamma(k)} \int \lambda^{\alpha+k-1} \exp(-(\delta + x)\lambda)\, d\lambda \\
&= \frac{\Gamma(\alpha + k)\delta^\alpha}{\Gamma(\alpha)\Gamma(k)} \frac{x^{k-1}}{(\delta + x)^{\alpha+k}}\,,\ x > 0,
\end{aligned}$$

gdje je zadnji integral izračunat prepoznavanjem $G(\alpha + k, \delta + x)$ integranda. To je vjerojatnosna funkcija gustoće generalizirane Paretove distribucije.
Momenti generalizirane Paretove distribucije mogu se dobiti ili direktnim računanjem $\int x f(x)\, dx$ ili upotrebom argumenta uvjetnog očekivanja. Što se procjene tiče, funkcija distribucije ne postoji u zatvorenom obliku, pa metoda percentila nije na raspolaganju. Može se koristiti metoda maksimalne vjerodostojnosti, ali je opet potreban kompjutorski software. Metoda momenata daje početne procjene za sve iterativne metode.

2 Reosiguranje

Osigurateljno društvo mora pokriti štete u cijelosti, ali da se pokrije od velikih šteta, i samo može uzeti policu. Takva polica zove se reosiguranje. Za potrebe ovog poglavlja pretpostavit ćemo da je reosigurateljni ugovor jednog od sljedeća dva jednostavna tipa: individualno reosiguranje viška štete ili kvotno reosiguranje.

2.1 Reosiguranje viška štete

Kod reosiguranja viška štete, društvo će u cijelosti platiti štetu do iznosa M koji se zove samopridržaj; svaki iznos iznad M pokrit će reosiguratelj.
Ugovor o reosiguranju viška štete može se zapisati na sljedeći način: ako je iznos štete X, tada će društvo platiti Y gdje je:

$$\begin{aligned} Y &= X \qquad \text{ako je } X \leq M \\ Y &= M \qquad \text{ako je } X > M. \end{aligned}$$

Reosiguratelj plaća iznos $Z = X - Y$.
To na osigurateljevu obvezu utječe na dva očita načina:

(i) smanjuje se očekivani isplaćeni iznos;

(ii) smanjuje se varijanca isplaćenog iznosa.

Oba zaključka su jednostavne posljedice činjenice da reosiguranje viška štete ograničava odozgo velike štete.
Sada se mogu dobiti i očekivanje i smanjenje očekivanja iznosa koji isplaćuje osiguratelj pri reosiguranju viška štete. Uočite da je očekivani iznos koji

isplaćuje osiguratelj bez reosiguranja

$$\mathbb{E}(X) = \int_0^\infty x f(x)\, dx \tag{1}$$

gdje je $f(x)$ vjerojatnosna funkcija gustoće iznosa štete X. Uz samopridržaj M očekivanje postaje

$$\mathbb{E}(Y) = \int_0^M x f(x)\, dx + M\mathbb{P}(X > M)\,. \tag{2}$$

Općenitije, funkcija izvodnica momenata od Y, iznosa koji isplaćuje osiguratelj, je

$$M_Y(t) = \mathbb{E}(e^{tY}) = \int_0^M e^{tx} f(x)\, dx + e^{tM}\mathbb{P}(X > M)\,. \tag{3}$$

Formule (1) i (2) pokazuju glavnu teškoću pri reosiguranju viška štete. U (2) integral je nepotpun (tj., granice su mu od 0 do M, a ne ∞). (Ista poteškoća pojavit će se pri razmatranju franšize.) Kod reosiguranja viška štete postoji metoda kojom se nepotpuni integral pretvori u potpuni. Korištenjem (2)

$$\begin{aligned} \mathbb{E}(Y) &= \int_0^\infty x f(x)\, dx - \int_M^\infty x f(x)\, dx + M \int_M^\infty f(x)\, dx \\ &= \mathbb{E}(X) - \int_M^\infty (x - M) f(x)\, dx. \end{aligned}$$

Potpun integral može se dobiti supstitucijom $z = x - M$. Prema tome

$$\mathbb{E}(Y) = \mathbb{E}(X) - \int_0^\infty z f(z + M)\, dz\,, \tag{4}$$

što je jednostavna, ali važna formula.
Smanjenje očekivanog iznosa štete je

$$\int_0^\infty z f(z + M)\, dz. \tag{5}$$

Postoji također i problem inflacije. Pretpostavimo da inflacija povećava iznose šteta za faktor k, ali da se samopridržaj M ne mijenja. Kakve će to posljedice imati na ugovor?

Iznos štete je kX, a iznos Y koji plaća osiguratelj je

$$\begin{array}{ll} Y = kX & \text{ako je } kX \leq M \\ Y = M & \text{ako je } kX > M. \end{array}$$

Očekivani iznos koji isplaćuje osiguratelj je

$$\mathbb{E}(Y) = \int_0^{M/k} kxf(x)\,dx + M\mathbb{P}(X > M/k)\,. \tag{6}$$

Ista metoda upotrebljena u (4) može se iskoristiti za pretvaranje (6) u potpun integral. Jednadžba (6) može se zapisati kao

$$\begin{aligned} \mathbb{E}(Y) &= \int_0^{\infty} kxf(x)\,dx - \int_{M/k}^{\infty} kxf(x)\,dx + M\int_{M/k}^{\infty} f(x)\,dx \\ &= k\mathbb{E}(X) - k\int_{M/k}^{\infty} (x - M/k)f(x)\,dx\,. \end{aligned}$$

Novi očekivani iznos koji isplaćuje osiguratelj je

$$\mathbb{E}(Y) = k(\mathbb{E}(X) - \int_0^{\infty} yf(y + M/k)\,dy)\,. \tag{7}$$

Uočite da integral u (7) ima isti oblik kao onaj u (4).
Važan opći zaključak je da novi očekivani iznos šteta (7) nije k puta očekivani iznos šteta bez inflacije (4).
Sličan pristup može se primjeniti u situacijama gdje je granica samopridržaja povezana s nekim indeksom cijena.
Promotrimo problem procjene pri reosiguranju viška štete. Pretpostavimo da podacima o štetama pokazuju samo štete koje je osiguratelj stvarno isplatio. Tipični podaci o štetama mogli bi biti

$$x_1, x_2, M, x_3, M, x_4, x_5, \ldots \tag{8}$$

i zahtjeva se procjena temeljne distribucije bruto šteta. Metoda momenata nije dostupna, budući da se ne može izračunati čak niti očekivani iznos šteta. S druge strane, ponekad će se bez promjene moći koristiti metoda percentila. To će biti moguće kada je samopridržaj M visok te će (mali broj) reosigurateljnih šteta utjecati samo na više percentile uzorka.
Statistički termin za uzorak oblika (8) je cenzurirani uzorak. Općenito, cenzurirani uzorak pojavljuje se kada su neke vrijednosti točno zabilježene, dok

je za druge vrijednosti poznato samo da prelaze neku određenu vrijednost, ovdje samopridržaj M.
Metoda maksimalne vjerodostojnosti može se primijeniti na cenzurirani uzorak. Funkcija vjerodostojnosti sastoji se od dva dijela. Točno zabilježene vrijednosti doprinose faktor

$$L_1(\underline{\theta}) = \prod_1^n f(x_i; \underline{\theta})$$

gdje je n šteta, x_i, poznato točno. Cenzurirane vrijednosti doprinose faktor

$$L_2(\underline{\theta}) = \prod_1^m \mathbb{P}(X > M), \quad \text{tj. } [\mathbb{P}(X > M)]^m$$

gdje ima m reosigurateljnih šteta. Potpuna funkcija vjerodostojnosti je

$$L(\underline{\theta}) = \prod_1^n f(x_i; \underline{\theta}) \times [1 - F(M; \underline{\theta})]^m \tag{9}$$

gdje je $F(\cdot; \underline{\theta})$ funkcija distribucije šteta.
Promatrajmo sada reosiguranje s točke gledišta reosiguratelja. Reosiguratelj može imati zabilježene samo štete koje su veće od M. Ako je šteta manja od M, reosiguratelj ne mora čak niti znati da je šteta nastala. Problem reosiguratelja je dakle procjena temeljne distribucije šteta kada se opažaju samo štete veće od M. Statistički termin je da reosiguratelj opaža štete iz odrezane distribucije.
Pretpostavimo da temeljna distribucija iznosa šteta ima vjerojatnosnu funkciju gustoće $f(x)$ i funkciju distribucije $F(x)$. Pretpostavimo da reosiguratelj ima informacije samo o štetama većim od samopridržaja M, te da ima zabilježeno $z = x - M$. Što je vjerojatnosna funkcija gustoće $g(z)$ iznosa Z koji plaća reosiguratelj?
Argument ide ovako:

$$\begin{aligned}
\mathbb{P}(Z < z) &= \mathbb{P}(X < z + M | X > M) \\
&= \int_M^{z+M} \frac{f(x)}{1 - F(M)}\, dx \\
&= \frac{F(z + M) - F(M)}{1 - F(M)}\,.
\end{aligned}$$

Deriviranjem po z dobivamo vjerojatnosnu funkciju gustoće reosigurateljevih šteta

$$g(z) = \frac{f(z+M)}{1-F(M)}\,,\ z > 0. \tag{10}$$

2.2 Kvotno reosiguranje

Kod kvotnog reosiguranja osiguratelj isplaćuje fiksan omjer štete, kolika god bila šteta. Upotrebom istih oznaka kao gore, ugovor o kvotnom reosiguranju može se zapisati kako slijedi: ako je iznos štete X, tada će društvo isplatiti Y gdje je

$$Y = \alpha X \quad 0 < \alpha < 1\,.$$

Parametar α se zove samopridržaj. Uočite da se termin samopridržaj koristi i kod reosiguranja viška štete i kod kvotnog reosiguranja.
Budući da je iznos po šteti X koji isplaćuje osiguratelj jednak αX, a iznos koji isplaćuje reosiguratelj jednak $(1-\alpha)X$, distribucije oba ta iznosa jednostavno se nalaze zamjenom varijabli.

3 Franšiza

Police osiguranja s ugovorenom franšizom uobičajene su kod osiguranja motornih vozila, te kod mnogih drugih osiguranja imovine i osiguranja od nesreće. Kod ove vrste police, osiguranik pristaje snositi puni iznos štete do iznosa L, koji se zove franšiza. Ako je iznos štete X veći od L, osiguranik će potraživati samo $X - L$. Ako je Y stvarni iznos koji isplaćuje osiguratelj, tada je

$$\begin{array}{ll} Y = 0 & \text{ako je } X \le L \\ Y = X - L & \text{ako je } X > L. \end{array}$$

Jasno je da će premija za policu s franšizom biti manja nego za policu bez franšize.
Pozicija osiguratelja za policu s ugovorenom franšizom potpuno je ista poziciji reosiguratelja kod reosiguranja viška štete. Pozicija osiguranika, barem što se tiče šteta, potpuno je ista poziciji osiguratelja s ugovorenim reosiguranjem viška štete.

POGLAVLJE 4 - MODELI RIZIKA

Nastavni ciljevi:

(iv)1. Konstruirati odgovarajuće modele kratkoročnih osigurateljnih ugovora pomoću broja odštetnih zahtjeva i iznosa individualnih šteta.

(iv)2. Opisati glavne pojednostavljujuće pretpostavke u temelju modela iz (iv)1.

(iv)3. Izvesti funkciju izvodnicu momenata za sumu od N nezavisnih slučajnih varijabli; posebno kada N ima binomnu, Poissonovu, geometrijsku ili negativnu binomnu distribuciju.

(iv)4. Definirati složenu Poissonovu distribuciju i pokazati da zbroj nezavisnih slučajnih varijabli sa složenom Poissonovom distribucijom opet ima složenu Poissonovu distribuciju.

(iv)5. Izvesti očekivanje, varijancu i koeficijent asimetrije za složenu binomnu, složenu Poissonovu i složenu negativnu binomnu slučajnu varijablu.

(iv)6. Izvesti formule za funkcije izvodnice momenata i momente skupnih šteta kroz dano vremensko razdoblje za modele iz (iv)1. pomoću odgovarajućih funkcija za distribuciju broja odštetnih zahtjeva i iznosa šteta, te navesti matematičke pretpostavke u temelju tih formula.

(iv)7. Ponoviti (iv)5. za osiguratelja i reosiguratelja za jednostavne oblike kvotnog reosiguranja i reosiguranja viška štete.

(v)1. Izvesti rekurzivne formule za izračun distribucije skupnih šteta u slučajevima gdje su iznosi šteta distribuirani na pozitivnim cijelim brojevima, a brojevi odštetnih zahtjeva imaju binomnu, Poissonovu ili negativnu binomnu distribuciju.

(v)2. Aproksimirati distribuciju skupnih šteta normalnom distribucijom pomoću prilagodbe momenata.

(v)3. Aproksimirati distribuciju skupnih šteta translatiranom gama distribucijom pomoću prilagodbe momenata.

1 Modeli kratkoročnih osigurateljnih ugovora

1.1 Osnovni model

Mnogi oblici neživotnog osiguranja, kao na primjer osiguranje motornih vozila, mogu se smatrati kratkoročnim ugovorima, isto kao i neki oblici životnog osiguranja, kao na primjer ugovor o grupnom osiguranju života i jednogodišnje osiguranje života.
Kratkoročni osigurateljni ugovor može se definirati sljedećim svojstvima:

- Polica traje fiksno i relativno kratko vremensko razdoblje, tipično jednu godinu.
- Osigurateljno društvo prima od osiguranika premiju.
- Za uzvrat, osiguratelj isplaćuje štete nastale po polici za vrijeme trajanja police.

Po isteku police osiguranik može, ali ne mora, obnoviti policu; ako je polica obnovljena premija koju plaća osiguranik može, ali ne mora, biti ista kao za proteklo razdoblje.
Osiguratelj može prenijeti dio premije reosiguratelju; za uzvrat, reosiguratelj će pokriti osiguratelju dio troškova štete za vrijeme trajanja police po nekoj dogovorenoj formuli.
Važno obilježje kratkoročnog osigurateljnog ugovora je to da je premija određena na razini pokrića šteta nastalih samo za vrijeme (kratkog) trajanja police. To je u suprotnosti s policama životnog osiguranja gdje porast inteziteta smrtnosti s dobi znači da će godišnje premije u ranim godinama biti više nego dovoljne za pokriće očekivanih šteta u ranim godinama. Višak iznosa će se tada akumulirati kao pričuva za upotrebu u kasnijim godinama kada će same premije biti nedovoljne za pokriće očekivanih troškova šteta.
Određenije, promatrat ćemo kratkoročni osigurateljni ugovor za pokriće rizika. Rizik uključuje bilo individualnu policu, bilo određenu grupu polica. Zbog lakše terminologije, pretpostavit ćemo da je trajanje ugovora jedna godina, ali bi isto tako moglo biti bilo koje drugo kratko razdoblje, na primjer šest mjeseci. Slučajna varijabla S označava skupne štete koje isplaćuje osiguratelj u godini po tom riziku. Konstruirat ćemo modele za tu slučajnu varijablu S.
U sljedeće dvije sekcije proučavat ćemo modele kolektivnog rizika. Kasnije, u sekciji 4, ideja modela kolektivnog rizika proširena je na model individualnog rizika. Prvi korak u konstrukciji modela kolektivnog rizika je zapisati S

pomoću broja odštetnih zahtjeva u godini, označenim slučajnom varijablom N, i iznosa svake individualne štete. Neka slučajna varijabla X_i označava iznos i-te štete. Tada je

$$S = \sum_{i=1}^{N} X_i \tag{1.1}$$

gdje je suma jednaka nuli ako je N jednak nuli. Ta dekompozicija od S dozvoljava odvojeno promatranje broja odštetnih zahtjeva i isnosa šteta. Praktična prednost toga je da faktori koji utječu na broj odštetnih zahtjeva i iznos šteta mogu biti različiti. Uzmimo na primjer osiguranje motornih vozila. Dulje razdoblje lošeg vremena može imati značajan utjecaj na broj odštetnih zahtjeva, a mali ili nikakav utjecaj na distribuciju individualnih iznosa šteta. S druge strane, inflacija može imati značajni utjecaj na trošak popravka automobila, te stoga na distribuciju individualnih iznosa šteta, ali mali ili nikakav utjecaj na broj odštetnih zahtjeva.
Problem koji će se proučavati je izvod momenata i distribucije od S pomoću momenata i distribucija od N i X_i-ova. Oba će se proučavati sa i bez jednostavnih oblika reosiguranja. Također će se proučavati odgovarajući problem za reosiguratelja, tj., izvod momenata i distribucije skupnih šteta koje isplaćuje reosiguratelj u godini po tom riziku.

1.2 Diskusija pojednostavljenja u osnovnom modelu

Model za kratkoročno osiguranje opisan u prošloj podsekciji sadrži cijeli niz pojednostavljanja u odnosu na stvarno osiguranje. Prvo od njih je da je uobičajeno pretpostaviti da su momenti, i ponekad distribucije, od N i X_i-ova poznati sa sigurnošću. U praksi bi se to vjerojatno procijenilo iz nekih relevantnih podataka metodama već proučenim u Poglavlju 3.
Drugo pojednostavljenje je pretpostavka, barem implicitna, da se štete rješavaju manje više odmah nakon što nastanu, tako da je, na primjer, osigurateljev profit poznat na kraju godine. U praksi postoji barem kratka odgoda u rješavanju šteta, a u nekim slučajevima odgoda može iznositi i više godina. To će specijalno biti tako kada je teško odrediti razmjer štete, na primjer ako se šteta treba odrediti na sudu.
Model općenito ne spominje troškove rješavanja šteta. Pretpostavlja se da premija pokriva štete, te da sadrži dodatak za profit. U praksi će premija koju plaća osiguranik također uključivati dodatak za troškove. Na jednostavan je način moguće uključiti u model dodatak za troškove.

Važan element u modelima za dugoročno osiguranje je stopa inflacije, budući da će, kao što je gore objašnjeno, višak premija biti investiran u razvoj pričuve. Kamata je relativno manje važno, iako još uvijek važno, obilježje kratkoročnog osiguranja. Moguće je u model za kratkoročno osiguranje uključiti kamatu, ali je uobičajeno zanemariti je, barem u elementarnim modelima.

1.3 Oznake i pretpostavke

Kroz cijelo poglavlje vrijedit će sljedeće dvije važne pretpostavke:

- slučajne varijable $\{X_i\}_{i=1}^{\infty}$ su nezavisne i jednako distribuirane,
- slučajna varijabla N nezavisna je s $\{X_i\}_{i=1}^{\infty}$.

Riječima, te pretpostavke znače da:

1. iznos individualnih šteta ne utječe na broj odštetnih zahtjeva
2. iznos jedne individualne štete ne utječe na iznos druge individualne štete
3. distribucija iznosa individualnih šteta se ne mijenja kroz (kratko) trajanje police.

U ovom poglavlju pretpostavlja se da sve štete imaju nenegativan iznos, tako da je $\mathbb{P}(X_i \leq x) = 0$ za $x < 0$. Mnoge formule u ovom poglavlju biti će izvedene upotrebom funkcija izvodnica momenata (od sada skraćeno na f.i.m.) od S, N i X_i. Te f.i.m. označavat će se redom s $M_S(t)$, $M_N(t)$ i $M_X(t)$, i pretpostavit će se da postoje za neke pozitivne vrijednosti varijable t. Općenito se ne može pretpostaviti egzistencija f.i.m. nenegativne slučajne varijable. Na primjer, f.i.m. Paretove i lognormalne distribucije ne postoje niti za jednu pozitivnu vrijednost od t. Međutim, sve formule izvedene u ovom poglavlju pomoću f.i.m. mogu se izvesi, iako ne tako lako, bez pretpostavke da f.i.m. postoji za pozitivne vrijednosti od t.
Sa $G(x)$, odnosno $F(x)$, označavat će se funkcije distribucije od S, odnosno X_i, tako da je

$$G(x) = \mathbb{P}(S \leq x) \quad \text{i} \quad F(x) = \mathbb{P}(X_i \leq x)\,.$$

Zbog jednostavnosti će se često pretpostaviti da postoji gustoća od $F(x)$ i označavat će se s $f(x)$. U slučajevima kada ta gustoća ne postoji, tako da X_i ima diskretnu ili miješanu neprekidno/diskretnu distribuciju, izraze kao

$$\int_0^\infty x f(x)\, dx$$

treba interpretirati na odgovarajući način. Značenje bi uvijek trebalo biti jasno iz konteksta.
Sa m_k, $k = 1, 2, 3, \ldots$, označavat će se k-ti moment oko nule od X_i.

2 Modeli kolektivnog rizika

2.1 Modeli kolektivnog rizika

Prisjetimo se iz Sekcije 1.1 da S predstavlja zbroj od N slučajnih varijabli X_i, gdje X_i označava iznos i te štete. Dakle,

$$S = X_1 + X_2 + \cdots + X_N$$

i $S = 0$ ako je $N = 0$.
Uočite da se promatra broj odštetnih zahtjeva N iz kolektivnog rizika (a ne broj odštetnih zahtjeva po individualnoj polici). To daje ime "model kolektivnog rizika". Unutar tog okvira mogu se razviti opći izrazi za funkciju distribucije, očekivanje, varijancu i f.i.m. od S.
Izraz za $G(x)$, funkciju distribucije od S, može se izvesti promatrajući događaj $\{S \leq x\}$. Uočite da ako se taj događaj dogodi, tada se jedan i samo jedan od sljedećih događaja mora dogoditi

$$\begin{array}{lll}
 & \{S \leq x \text{ i } N = 0\} & \text{(tj. nema šteta)} \\
\text{ili} & \{S \leq x \text{ i } N = 1\} & \text{(tj. jedna šteta iznosa } \leq x) \\
\text{ili} & \{S \leq x \text{ i } N = 2\} & \text{(tj. dvije štete s ukupnim iznosom } \leq x) \\
 & \vdots & \\
\text{ili} & \{S \leq x \text{ i } N = r\} & \text{(tj. } r \text{ šteta s ukupnim iznosom } \leq x) \\
 & \vdots &
\end{array}$$

itd. Ti događaji su međusobno disjunktni i unija im je $\{S \leq x\}$. Prema tome

$$\{S \leq x\} = \bigcup_{n=0}^{\infty} \{S \leq x \text{ i } \ N = n\}$$

te stoga

$$\begin{aligned}\mathbb{P}(S \leq x) &= \sum_{n=0}^{\infty} \mathbb{P}(S \leq x \text{ i } N = n) \\ &= \sum_{n=0}^{\infty} \mathbb{P}(N = n)\mathbb{P}(S \leq x | N = n)\,.\end{aligned}$$

Uočite sada da ako je $N = n$, tada je S zbroj fiksnog broja n slučajnih varijabli $\{X_i\}_{i=1}^{N}$, te je stoga

$$\mathbb{P}(S \leq x | N = n) = F^{n^*}(x)$$

gdje je $F^{n^*}(x)$ n-struka konvolucija distribucije $F(x)$. (Uočite da je $F^{1^*}(x)$ upravo $F(x)$, a zbog pogodnosti se definira da je $F^{0^*}(x)$ jednako 1 za sve nenegativne vrijednosti od x. Inače je $F^{0^*}(x) = 0$.) Prema tome

$$G(x) = \sum_{n=0}^{\infty} \mathbb{P}(N = n) F^{n^*}(x). \tag{2.1}$$

Formula (2.1) je opći izraz za funkciju distribucije od S. Nisu specificirane niti distribucija od N niti distribucija od X_i.
Kada je X_i distribuirana na pozitivnim cijelim brojevima, uočite da je tada lako izračunati $\mathbb{P}(S = x)$ za $x = 1, 2, 3, \ldots$, budući da je

$$\begin{aligned}\mathbb{P}(S = x) &= G(x) - G(x-1) \\ &= \sum_{n=1}^{\infty} \mathbb{P}(N = n)\{F^{n^*}(x) - F^{n^*}(x-1)\}\,,\end{aligned}$$

tj.,

$$\mathbb{P}(S = x) = \sum_{n=1}^{\infty} \mathbb{P}(N = n) f_x^{n^*} \tag{2.2}$$

gdje je $f_x^{n^*} = F^{n^*}(x) - F^{n^*}(x-1)$ vjerojatnosna funkcija od $\sum_{i=1}^{n} X_i$. Kao i u slučaju neprekidnih slučajnih varijabli X_i, $\mathbb{P}(S = 0) = \mathbb{P}(N = 0)$.
U Sekciji 3 govorit će se o točnim i aproksimativnim metodama za računanje $G(x)$. Za aproksimativne metode potrebni su momenti od S. O njima se raspravlja dolje.

Za izračunavanje momenata od S, koriste se rezultati o uvjetnom očekivanju, i uvjetuje se na broj odštetnih zahtjeva N. Za nalaženje $\mathbb{E}[S]$ primjeni se identitet

$$\mathbb{E}[S] = \mathbb{E}[\mathbb{E}[S|N]]\,.$$

Sada $\mathbb{E}[S|N = n] = \sum_{i=1}^{n} \mathbb{E}[X_i] = nm_1$. Stoga je $\mathbb{E}[S|N] = Nm_1$, i

$$\mathbb{E}[S] = \mathbb{E}[Nm_1] = \mathbb{E}[N]m_1\,. \tag{2.3}$$

Formula (2.3) ima vrlo prirodnu interpretaciju. Ona kaže da je očekivani iznos skupnih šteta jednak produktu očekivanog broja odštetnih zahtjeva i očekivanog iznosa individualne štete.
Da bismo pronašli izraz za $\mathrm{Var}[S]$, primjenjujemo identitet

$$\mathrm{Var}[S] = \mathbb{E}[\mathrm{Var}[S|N]] + \mathrm{Var}[\mathbb{E}[S|N]]\,.$$

$\mathrm{Var}[S|N]$ može se pronaći upotrebom činjenice da su iznosi individualnih šteta nezavisni. Sada

$$\mathrm{Var}[S|N = n] = \mathrm{Var}\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} \mathrm{Var}[X_i] = n(m_2 - m_1^2)$$

te stoga $\mathrm{Var}[S|N] = N(m_2 - m_1^2)$. Zato

$$\mathrm{Var}[S] = \mathbb{E}[N(m_2 - m_1^2)] + \mathrm{Var}[Nm_1]\,,$$

tj.

$$\mathrm{Var}[S] = \mathbb{E}[N](m_2 - m_1^2) + \mathrm{Var}[N]m_1^2\,. \tag{2.4}$$

Za razliku od izraza za $\mathbb{E}(S)$, formula (2.4) nema prirodnu interpretaciju. Varijanca od S izražena je pomoću očekivanja i varijance i od N i od X_i.
Funkcija izvodnica momenata os S također se nalazi upotrebom uvjetnog očekivanja. Po definiciji, $M_S(t) = \mathbb{E}(\exp\{tS\})$, te je

$$M_S(t) = \mathbb{E}[\mathbb{E}[\exp\{tS\}|N]]\,. \tag{2.5}$$

Sada je $\mathbb{E}[\exp\{tS\}|N = n] = \mathbb{E}[\exp\{tX_1 + tX_2 + \cdots + tX_n\}]$, i budući da su $\{X_i\}_{i=1}^{n}$ nezavisne slučajne varijable,

$$\mathbb{E}[\exp\{tX_1 + tX_2 + \cdots + tX_n\}] = \prod_{i=1}^{n} \mathbb{E}[\exp\{tX_i\}]\,.$$

Također, budući da su $\{X_i\}_{i=1}^n$ jednako distribuirane, one imaju zajedničku f.i.m., $M_X(t)$, tako da je

$$\prod_{i=1}^{n} \mathbb{E}[\exp\{tX_i\}] = \prod_{i=1}^{n} M_X(t) = M_X(t)^n \, .$$

Zato je

$$\mathbb{E}[\exp\{tS\}|N] = [M_X(t)]^N \, . \tag{2.6}$$

Dakle, supstitucijom (2.6) u (2.5),

$$M_S(t) = \mathbb{E}(M_X(t)^N) = \mathbb{E}[\exp\{N \log M_X(t)\}] = M_N(\log M_X(t)) \, . \tag{2.7}$$

Prema tome, f.i.m. od S predstavljena je pomoću f.i.m. od N i od X_i. Kao i kod prethodnih rezultata, distribucije od N ili od X_i nisu specificirane.
Postoji jedan specijalan slučaj koji je od nekog interesa. To je kada sve štete imaju isti fiksan iznos.
Na primjer, možemo imati portfelj sa jednogodišnjim osiguranjima života sa istim osiguranim iznosom. Pretpostavimo da je iznos štete B s vjerojatnošću jedan, tj., $\mathbb{P}(X_i = B) = 1$, te stoga $m_1 = B$ i $m_2 = B^2$. Tada je S distribuirano na $0, B, 2B, \ldots$. U stvari, $S = NB$, te je

$$\mathbb{P}(S \leq Bx) = \mathbb{P}(N \leq x)$$

i distribucija od S slijedi distribuciju od N. Formule (2.3) i (2.4) daju očekivanje i varijancu od S, ali kako je $S = NB$, lakše je primjetiti da je $\mathbb{E}[S] = \mathbb{E}[N]B$ i $\mathrm{Var}[S] = \mathrm{Var}[N]B^2$.
Sljedeće tri sekcije razmatraju složene distribucije korištenjem različitih modela za broj odštetnih zahtjeva N.

2.2 Složena Poissonova distribucija

Promatrajmo prvo skupne štete u slučaju da N ima Poissonovu distribuciju s očekivanjem λ, u oznaci $N \sim P(\lambda)$. Tada S ima složenu Poissonovu distribuciju s parametrima λ i $F(x)$. Rezultati o N koji su nam potrebni za tu distribuciju su

$$\mathbb{E}[N] = \mathrm{Var}[N] = \lambda$$
$$M_N(t) = \exp\{\lambda(e^t - 1)\} \, .$$

Uočite da su ti rezultati dani u Tablicama.

Ti se rezultati mogu kombinirati s rezultatima iz Sekcije 2.1 kako slijedi:
iz (2.3)

$$\mathbb{E}[S] = \lambda m_1 \tag{2.8}$$

iz (2.4)

$$\mathrm{Var}[S] = \lambda m_2 \tag{2.9}$$

i iz (2.7),

$$M_S(t) = \exp\{\lambda(M_X(t) - 1)\}\,. \tag{2.10}$$

Rezultati za očekivanje i varijancu imaju vrlo jednostavan oblik. Uočite da je varijanca od S izražena pomoću drugog momenta oko nule od X_i (a ne pomoću varijance od X_i).
Da bismo pokazali da je treći centralni moment od S jednak λm_3, upotrijebit ćemo funkciju izvodnicu kumulanti (formula (0.2) s $n = 3$), tj.,

$$\mathbb{E}[(S - \lambda m_1)^3] = \frac{d^3}{dt^3} \log M_S(t)_{|t=0}\,.$$

Nadalje, $M_S(t) = \lambda(M_X(t) - 1)$, te je

$$\frac{d^3}{dt^3} \log M_S(t)_{|t=0} = \lambda \frac{d^3}{dt^3} \log M_X(t)_{|t=0} = \lambda m_3\,,$$

tj., $\mathbb{E}[(S - \lambda m_1)^3] = \lambda m_3$, pa je koeficijent asimetrije $= \lambda m_3/(\lambda m_2)^{3/2}$.
Taj rezultat pokazuje da je distribucija od S pozitivno asimetrična, budući da je m_3 treći moment oko nule od X_i, te je stoga veći od nule jer je X_i nenegativna slučajna varijabla. Uočite da je distribucija od S pozitivno asimetrična čak i ako je distribucija od X_i negativno asimetrična. Koeficijent asimetrije od S je $= \lambda m_3/(\lambda m_2)^{3/2}$, te stoga teži prema 0 kada $\lambda \to \infty$. Stoga je za velike vrijednosti od λ, distribucija od S skoro simetrična.
Vrlo važno svojstvo složene Poissonove distribucije je da je zbroj nezavisnih složenih Poissonovih slučajnih varijabli i sama složena Poissonova slučajna varijabla. Formalna tvrdnja tog svojstva je kako slijedi.
Neka su $S_1, S_2, \ldots, S_n$ nezavisne slučajne varijable, te neka S_i ima složenu Poissonovu distribuciju s parametrima λ_i i $F_i(x)$.
Definiramo $A = S_1 + S_2 + \cdots + S_n$. Tada A ima složenu Poissonovu distribuciju s parametrima Λ i $F(x)$, gdje je

$$\Lambda = \sum_{i=1}^{n} \lambda_i \quad \text{i} \quad F(x) = \frac{1}{\Lambda} \sum_{i=1}^{n} \lambda_i F_i(x)\,.$$

To je vrlo važan rezultat i na njega ćemo se kasnije pozivati kao na Rezultat 1.

Da bismo dokazali taj rezultat, uočite prvo da je $F(x)$ težinsko usrednjenje funkcija distribucije, i da su te težine pozitivne i imaju zbroj jedan. To znači da je $F(x)$ funkcija distribucije i ta distribucija ima f.i.m.

$$M(t) = \int_0^\infty \exp(tx) \frac{1}{\Lambda} \sum_{i=1}^{n} \lambda_i f_i(x)\, dx$$

gdje je $f_i(x)$ gustoća od $F_i(x)$. Zato je

$$M(t) = \frac{1}{\Lambda} \sum_{i=1}^{n} \lambda_i \int_0^\infty \exp\{tx\} f_i(x)\, dx = \frac{1}{\Lambda} \sum_{i=1}^{n} \lambda_i M_i(t) \tag{2.11}$$

gdje je $M_i(t)$ f.i.m. distribucije $F_i(x)$.

Neka $M_A(t)$ označava f.i.m. od A. Tada je $M_A(t) = \mathbb{E}[\exp\{tA\}]$ $= \mathbb{E}[\exp\{tS_1 + tS_2 + \cdots + tS_n\}]$. Zbog nezavisnosti od $\{S_i\}_{i=1}^n$,

$$M_A(t) = \prod_{i=1}^{n} \mathbb{E}(\exp\{tS_i\})\,.$$

Budući da su S_i složene Poissonove slučajne varijable, njihove f.i.m. imaju oblik dan formulom (2.10), tako da je

$$\mathbb{E}[\exp\{tS_i\}] = \exp\{\lambda_i(M_i(t) - 1)\}\,.$$

Zato je

$$\begin{aligned} M_A(t) &= \prod_{i=1}^{n} \exp\{\lambda_i(M_i(t) - 1)\} \\ &= \exp\left\{\sum_{i=1}^{n} \lambda_i(M_i(t) - 1)\right\} \end{aligned}$$

tj.,

$$M_A(t) = \exp\{\Lambda(M(t) - 1)\} \tag{2.12}$$

gdje je $\Lambda = \sum_{i=1}^n \lambda_i$ i $M(t) = \frac{1}{\Lambda}\sum_{i=1}^n \lambda_i M_i(t)$. Zbog 1-1 veze među distribucijama i funkcijama izvodnicama momenata, formula (2.12) pokazuje da A ima složenu Poissonovu distribuciju s Poissonovim parametrom Λ. Po (2.11), distribucija iznosa individualne štete je $F(x)$.

2.3 Složena binomna distribucija

U nekim okolnostima je binomna distribucija prirodan izbor za N. Na primjer, kod police o grupnom osiguranju života koja pokriva n života, distribucija broja smrti u godini je binomna, ako se pretpostavi da svaki osiguran život ima isti intenzitet smrtnosti, te da su životi nezavisni u odnosu na smrtnost. Oznaka $N \sim b(n,q)$ koristi se za binomnu distribuciju N. Ključni rezultati za tu distribuciju su

$$\mathbb{E}[N] = nq$$
$$\mathrm{Var}[N] = nq(1-q)$$
$$M_N(t) = (qe^t + 1 - q)^n \,.$$

Uočite da su ti rezultati dani u Tablicama.
Kada N ima binomnu distribuciju, S ima složenu binomnu distribuciju. Važna činjenica kod izbora binomne distribucije za N je da postoji gornja ograda, n, za broj odštetnih zahtjeva.
Kada je $N \sim b(n,q)$, izrazi za očekivanje, varijancu i f.i.m. od S nalaze se pomoću n, q, m_1, m_2 i $M_X(t)$.
Formule (2.3) i (2.4) daju očekivanje i varijancu

$$\mathbb{E}[S] = nqm_1 \tag{2.13}$$

$$\mathrm{Var}[S] = nq(m_2 - m_1^2) + nq(1-q)m_1^2 = nqm_2 - nq^2m_1^2\,. \tag{2.14}$$

Konačno, formula (2.7) daje f.i.m.

$$M_S(t) = (qM_X(t) + 1 - q)^n \,.$$

Treći centralni moment nalazi se iz funkcije izvodnice kumulanti (iskoristite formulu (0.2)):

$$\begin{aligned}
&\frac{d^3}{dt^3}\log M_S(t) = \frac{d^3}{dt^3} n\log(qM_X(t)+p) \qquad \text{gdje je } p = 1-q \\
&= \frac{d^2}{dt^2}\left\{ nq\left(\frac{d}{dt}M_X(t)\right)(qM_X(t)+p)^{-1}\right\} \\
&= \frac{d}{dt}\left\{ nq\left(\frac{d^2}{dt^2}M_X(t)\right)(qM_X(t)+p)^{-1} - n\left(q\frac{d}{dt}M_X(t)\right)^2 (qM_X(t)+p)^{-2}\right\} \\
&= nq\left(\frac{d^3}{dt^3}M_X(t)\right)(qM_X(t)+p)^{-1} - 3nq^2\left(\frac{d^2}{dt^2}M_X(t)\right)(qM_X(t)+p)^{-2}\left(\frac{d}{dt}M_X(t)\right) \\
&+ 2n\left(q\frac{d}{dt}M_X(t)\right)^3 (qM_X(t)+p)^{-3}\,.
\end{aligned}$$

Supstitucija $t = 0$ daje

$$\mathbb{E}[(S - nqm_1)^3] = nqm_3 - 3nq^2m_2m_1 + 2nq^3m_1^3 \,. \tag{2.15}$$

Koeficijent asimetrije dan je formulom

$$\frac{nqm_3 - 3nq^2m_2m_1 + 2nq^3m_1^3}{(nqm_2 - nq^2m_1^2)^{3/2}} \,.$$

Iz formule (2.15) može se izvesti da je moguće da složena binomna distribucija bude negativno asimetrična. Najjednostavnija ilustracija te činjenice je slučaj kada su svi iznosi šteta jednaki B. Tada je $S = BN$ i

$$\mathbb{E}[(S - \mathbb{E}[S])^3] = B^3\mathbb{E}[(N - \mathbb{E}[N])^3]$$

tako da je koeficijent asimetrije od S faktor koeficijenta asimetrije od N. Ako je $q > 0.5$, binomna distribucija je negativno asimetrična.

2.4 Složena negativna binomna distribucija

Završni izbor distribucije za N je negativna binomna distribucija koja ima vjerojatnosnu funkciju

$$\mathbb{P}(N = n) = \binom{k+n-1}{n} p^k q^n \quad \text{za } n = 0, 1, 2, \ldots .$$

Parametri distribucije su $k(> 0)$ i p, gdje je $p + q = 1$ i $0 < p < 1$. Ta distribucija se označava s $NB(k, p)$. Kada je $N \sim NB(k, p)$, tada

$$\begin{aligned} \mathbb{E}[N] &= kq/p \\ \mathrm{Var}[N] &= kq/p^2 \\ M_N(t) &= p^k(1 - qe^t)^{-k} \end{aligned}$$

Specijalan slučaj $k = 1$ vodi na geometrijsku distribuciju. Još jednom, uočite da su ti rezultati dani u Tablicama.

Negativna binomna distribucija je alternativa Poissonovoj za N. Jedna prednost koju negativna binomna distribucija ima pred Poissonovom distribucijom je ta da joj je varijanca veća od očekivanja. Te dvije veličine jednake su za Poissonovu distribuciju. Prema tome, negativna binomna distribucija može se bolje prilagoditi podacima kod kojih je varijanca uzorka veća od očekivanja uzorka. To je čest slučaj u praksi. U Sekciji 5.2 razmatra se situacija koja dovodi do negativne binomne distribucije za N. Kada N ima negativnu binomnu distribuciju, S ima složenu negativnu binomnu distribuciju.

Izrazi za očekivanje, varijancu i f.i.m. od S za slučaj $N \sim NB(k,p)$ izlaze odmah iz formula (2.3), (2.4) i (2.7):

$$\begin{aligned}
\mathbb{E}[S] &= \frac{kq}{p}m_1 \\
\mathrm{Var}[S] &= \frac{kq}{p}(m_2 - m_1^2) + \frac{kq}{p^2}m_1^2 = \frac{kq}{p}m_2 + \frac{kq^2}{p^2}m_1^2 \\
M_S(t) &= \frac{p^k}{(1-qM_X(t))^k}\,.
\end{aligned}$$

Kao i prije, treći centralni moment od S može se naći iz funkcije izvodnice kumulanti od S kako slijedi:

$$\begin{aligned}
\frac{d}{dt}\log M_S(t) &= \frac{d}{dt}(k\log p - k\log(1-qM_X(t))) \\
&= \frac{kq}{1-qM_X(t)}\left(\frac{d}{dt}M_X(t)\right).
\end{aligned}$$

Tada je

$$\frac{d^2}{dt^2}\log M_S(t) = kq^2\left(\frac{d}{dt}M_X(t)\right)^2\frac{1}{(1-qM_X(t))^2} + \frac{kq}{1-qM_X(t)}\left(\frac{d^2}{dt^2}M_X(t)\right)$$

i

$$\begin{aligned}
\frac{d^3}{dt^3}\log M_S(t) &= 3kq^2\left(\frac{d}{dt}M_X(t)\right)\left(\frac{d^2}{dt^2}M_X(t)\right)\frac{1}{(1-qM_X(t))^2} \\
&+ \frac{2kq^3}{(1-qM_X(t))^3}\left(\frac{d}{dt}M_X(t)\right)^3 + \frac{kq}{1-qM_X(t)}\left(\frac{d^3}{dt^3}M_X(t)\right).
\end{aligned}$$

Supstitucija $t = 0$ u treću derivaciju daje

$$\mathbb{E}[(S-\mathbb{E}[S])^3] = \frac{3kq^2m_1m_2}{p^2} + \frac{2kq^3m_1^3}{p^3} + \frac{kqm_3}{p}\,. \tag{2.16}$$

Parametri k i p su pozitivni, kao i momenti od $F(x)$. Stoga iz formule (2.16) slijedi da je složena negativna binomna distribucija pozitivno asimetrična. Koeficijent asimetričnosti može se naći iz $\mathbb{E}((S-\mathbb{E}[S])^3)/(\mathrm{Var}[S])^{3/2}$.

2.5 Distribucije skupne štete uz kvotno i reosiguranje viška štete

2.5.1 Kvotno reosiguranje

Distribucija broja odštetnih zahtjeva za reosiguratelja ista je distribuciji broja odštetnih zahtjeva za osiguratelja, budući da svaki isplaćuje definirani omjer svake štete. Za samopridržaj α $(0 \leq \alpha \leq 1)$, iznosi individualnih šteta za osiguratelja distribuirani su kao αX_i, a za reosiguratelja kao $(1-\alpha)X_i$. Iznosi skupnih šteta distribuirani su kao αS, odnosno $(1-\alpha)S$.

2.5.2 Reosiguranje viška štete

Iznos koji isplaćuje osiguratelj po i-toj šteti uz reosiguranje viška štete sa samopridržajem M je $Y_i = \min(X_i, M)$.
Iznos koji plaća reosiguratelj je $Z_i = \max(0, X_i - M)$.
Prema tome, osigurateljeve skupne štete, mogu se predstaviti kao

$$S_O = Y_1 + Y_2 + \cdots + Y_N$$

a reosigurateljeve skupne štete kao

$$S_R = Z_1 + Z_2 + \cdots + Z_N \,. \tag{2.17}$$

Ako je, na primjer, $N \sim P(\lambda)$, S_O ima složenu Poissonovu distribuciju s Poissonovim parametrom λ, a iznosi individualnih šteta distribuirani su kao Y_i. Slično, S_R ima složenu Poissonovu distribuciju s Poissonovim parametrom λ, a iznosi individualnih šteta distribuirani su kao Z_i. Uočite, međutim, da ako je $F(M) > 0$, kao što je obično slučaj, tada Z_i prima vrijednost 0 s vjerojatnošću različitom od nule (naime $F(M)$). Drugim riječima, 0 se računa kao mogući iznos štete za reosiguratelja. S praktičnog stanovišta, ta definicija od S_R je umjetna. Osiguratelj će znati opaženu vrijednost od N, ali će reosiguratelj vjerojatno znati samo broj šteta iznad samopridržaja M, budući da će osiguratelj prijaviti reosiguratelju samo štete iznad samopridržaja.

Primjer

Iznos godišnjih skupnih šteta iz rizika ima složenu Poissonovu distribuciju s Poissonovim parametrom 10. Iznosi individualnih šteta uniformno su distribuirani na $(0, 2000)$. Osiguratelj je sklopio ugovor o reosiguranju viška štete sa samopridržajem 1 600. Izračunajte očekivanje, varijancu i koeficijent asimetričnosti skupnih šteta osiguratelja i reosiguratelja pri tom ugovoru o reosiguranju.

Rješenje

Neka su S_O i S_R kao gore. Da bismo pronašli $\mathbb{E}[S_O]$ računamao $\mathbb{E}[Y_i]$. Vrijedi

$$\mathbb{E}[Y_i] = \int_0^M x f(x)\,dx + M\mathbb{P}(X_i > M)$$

gdje je $f(x) = 0.0005$ funkcija gustoće od $U(0, 2000)$, a $M = 1\,600$. To daje

$$\mathbb{E}[Y_i] = \frac{0.0005x^2}{2}\Big|_0^M + 0.2M = 960$$

tako da je

$$\mathbb{E}[S_O] = 10\mathbb{E}[Y_i] = 9\,600\,.$$

Da bismo pronašli $\mathrm{Var}[S_O]$ izračunajmo $\mathbb{E}[Y_i^2]$ iz

$$\begin{aligned}\mathbb{E}[Y_i^2] &= \int_0^M x^2 f(x)\,dx + M^2\mathbb{P}(X_i > M)\\ &= \frac{0.0005x^3}{3}\Big|_0^M + 0.2M^2 = 1\,194\,666.7\end{aligned}$$

tako da je

$$\mathrm{Var}[S_O] = 10\mathbb{E}[Y_i^2] = 11\,946\,667.$$

Da bismo našli koeficijent asimetrije osigurateljevih šteta, računamo $\mathbb{E}[Y_i^3]$ iz

$$\begin{aligned}\mathbb{E}[Y_i^3] &= \int_0^M x^3 f(x)\,dx + M^3\mathbb{P}(X_i > M)\\ &= \frac{0.0005x^4}{4}\Big|_0^M + 0.2M^3\\ &= 1\,638\,400\,000\end{aligned}$$

tako da je

$$\mathbb{E}[(S_O - \mathbb{E}(S_0))^3] = 10\mathbb{E}[Y_i^3] = 16\,384\,000\,000$$

a koeficijent asimetrije je

$$\begin{aligned}&= 16\,384\,000\,000/(11\,946\,667)^{3/2}\\ &= 0.397\,.\end{aligned}$$

Da bismo pronašli $\mathbb{E}[S_R]$, uočite da je očekivani iznos godišnje skupne štete iz rizika jednak 10 000, tako da je

$$\mathbb{E}[S_R] = 10\,000 - \mathbb{E}[S_O] = 400\,.$$

Da bismo pronašli $\mathrm{Var}[S_R]$ izračunajmo $\mathbb{E}[Z_i^2]$ iz

$$\begin{aligned}
\mathbb{E}[Z_i^2] &= \int_M^{2000} (x-M)^2 f(x)\,dx \\
&= \int_0^{2000-M} 0.0005y^2\,dy \qquad \text{gdje je } y = x - M \\
&= \frac{0.0005y^3}{3}\Big|_0^{2000-M} \\
&= 10\,666.7
\end{aligned}$$

tako da je

$$\mathrm{Var}[S_R] = 10\mathbb{E}[Z_i^2] = 106\,667\,.$$

Da bismo pronašli koeficijent asimetrije reosigurateljevih šteta, izračunajmo $\mathbb{E}[Z_i^3]$ iz

$$\begin{aligned}
\mathbb{E}[Z_i^3] &= \int_M^{2000} (x-M)^3 f(x)\,dx \\
&= \int_0^{2000-M} 0.0005y^3\,dy \qquad \text{gdje je } y = x - M \\
&= 3\,200\,000
\end{aligned}$$

tako da je

$$\mathbb{E}[(S_R - \mathbb{E}[S_R])^3] = 10\mathbb{E}[Z_i^3] = 32\,000\,000$$

a koeficijent asimetrije je

$$\begin{aligned}
&= 32\,000\,000/(106\,667)^{3/2} \\
&= 0.92\,.
\end{aligned}$$

Reosigurateljeve skupne štete također se mogu predstaviti kao

$$S_R = W_1 + W_2 + \cdots + W_{NR} \tag{2.18}$$

gdje slučajna varijabla NR označava stvarni broj reosiguravateljevih isplata (različitih od nule).

Na primjer, pretpostavimo da je gornji rizik proizveo sljedećih osam iznosa šteta u određenoj godini:

$$403 \;\; 1490 \;\; 1948 \;\; 443 \;\; 1866 \;\; 1704 \;\; 1221 \;\; 823\,.$$

Tada je u formuli (2.17) opažena vrijednost od N jednaka 8, te treća, peta i šesta šteta zahtijevaju reosiguravateljeve isplate od 348, 266 i 104. Reosigurratelj "isplaćuje" 0 za ostalih pet šteta.
U formuli (2.18), opažena vrijednost od NR je 3, i opažene vrijednosti od W_1, W_2 i W_3 su 348, 266 i 104. Uočite da je opažena vrijednost od S_R jednaka (tj. 718) uz obje definicije.
Poglavlje 3, Sekcija 3, pokazuje da W_i ima funkciju gustoće

$$p(w) = \frac{f(w+M)}{1-F(M)}\,, \quad w > 0\,.$$

Da bi se odredila distribucija od S_R definirane formulom (2.18) potrebna je distribucija od NR. Ta se može pronaći kao slijedi. Definirajmo

$$NR = I_1 + I_2 + \cdots + I_N$$

gdje N (kao i obično) označava broj odštetnih zahtjeva iz rizika. I_j je indikatorska slučajna varijabla koja prima vrijednost 1 ako reosiguratelj isplaćuje iznos različit od nule po j-toj šteti, a inače prima vrijednost 0. Prema tome NR daje broj reosigurateljevih isplata. Budući da I_j prima vrijednost 1 samo ako je $X_j > M$,

$$\begin{aligned} \mathbb{P}(I_j = 1) &= \mathbb{P}(X_j > M) = \pi, \text{ recimo, i} \\ \mathbb{P}(I_j = 0) &= 1 - \pi\,. \end{aligned}$$

Nadalje, f.i.m. od I_j je

$$M_I(t) = \pi \exp\{t\} + 1 - \pi$$

te po formuli (2.7), f.i.m. od NR je

$$M_{NR}(t) = M_N(\log M_I(t))\,.$$

Primjer
Nastavljajući gornji primjer i korištenjem formule (2.18) kao model za S_R, može se vidjeti da S_R ima složenu Poissonovu distribuciju s Poissonovim parametrom $0.2 \times 10 = 2$.

Individualne štete W_j imaju funkciju gustoće

$$p(w) = \frac{f(w+M)}{1-F(M)} = 0.0005/0.2 = 0.0025, \quad \text{za } 0 < w < 400$$

tj., W_i su uniformno distribuirane na $(0, 400)$. Nadalje, $\mathbb{E}[W_i] = 200$, $\mathbb{E}[W_i^2] = 53\,333.33$ i $\mathbb{E}[W_i^3] = 16\,000\,000$, što daje isti rezultat kao prije.
Prema tome, postoje dva načina za određivanje i računanje distribucije od S_R.

3 Točno i približno računanje $G(x)$ za model kolektivnog rizika

3.1 Uvod

U ovoj sekciji proučavaju se računanja i aproksimacije od $G(x)$, funkcije distribucije skupnih šteta za model kolektivnog rizika. U nekim slučajevima moguće je vrlo jednostavno pronaći funkciju distribucije $G(x)$, na primjer ako sve štete imaju isti iznos, ali će pretpostavke u takvim slučajevima obično biti prerestriktivne da bi bile od praktičnog interesa. U Sekciji 3.2 izvedena je rekurzivna formula za računanje $G(x)$. U Sekciji 3.3 $G(x)$ je aproksimirana normalnom distribucijom prilagodbom momenata. Konačno, u Sekciji 3.4 $G(x)$ je aproksimirana (translatiranom) gama distribucijom prilagodbom momenata. U Sekciji 3.2 pretpostavljeno je da su poznate distribucije broja odštetnih zahtjeva i iznosa štete; u Sekcijama 3.3. i 3.4 pretpostavljeno je da su poznata samo prva dva ili tri momenta distribucije.

3.2 Rekurzivna formula za $G(x)$

U ovoj podsekciji pretpostavlja se da je distribucija iznosa individualnih šteta, $F(x)$, diskretna distribucija na pozitivnim cijelim brojevima. To znači da su moguće vrijednosti za individualne štete $1, 2, 3, \ldots$, te su stoga moguće vrijednosti za skupne štete $0, 1, 2, 3, \ldots$. Uz tu pretpostavku funkcija distribucije iznosa individualnih šteta nema funkciju gustoće (budući da je svaka X_i diskretna slučajna varijabla).
Sljedeće oznake koristit će se za vjerojatnosne funkcije individualnih iznosa

šteta, odnosno skupnih šteta.

$$\begin{aligned} f_k &= \mathbb{P}(X_i = k) \qquad && k = 1, 2, 3, \ldots \\ g_k &= \mathbb{P}(S = k) \qquad && k = 0, 1, 2, \ldots \end{aligned}$$

Ako distribucija individualnih iznosa šteta *nije* diskretna na pozitivnim cijelim brojevima, što je čest slučaj, ta se distribucija uvijek može aproksimirati diskretnom distribucijom.
Uz gornje pretpostavke se sada problem računanja $G(x)$ svodi na računanje g_k za $k \leq x$. Pretpostavlja se da je

- distribucija broja odštetnih zahtjeva poznata

- distribucija iznosa individualnih šteta (tj. f_k) poznata.

Prije dokaza, ili čak tvrdnje formule, potrebna je pretpostavka o distribuciji broja odštetnih zahtjeva N. Označimo $\mathbb{P}(N = r)$ sa p_r, i pretpostavimo da postoje brojevi a i b takvi da je

$$p_r = (a + b/r)p_{r-1} \quad \text{za } r = 1, 2, 3, \ldots. \tag{3.1}$$

Sve tri distribucije broja odštetnih zahtjeva proučavane u Sekciji 2 zadovoljavaju Pretpostavku (3.1). Formule za g_r su:

$$g_0 = p_0 \tag{3.2}$$

$$g_r = \sum_{j=1}^{r} (a + bj/r) f_j g_{r-1} \quad \text{za } r = 1, 2, \ldots. \tag{3.3}$$

Formula (3.2) slijedi iz činjenice da je minimalni iznos štete jednak jedan. Skupne štete biti će nula ako i samo ako ne nastane niti jedna šteta.
Za dokaz formule (3.3), koristit će se sljedeće tri formule: za $n = 2, 3, \ldots$

$$\mathbb{E}\left[X_1 | \sum_{i=1}^{n} X_i = r\right] = r/n \tag{3.4}$$

$$\mathbb{E}\left[X_1 | \sum_{i=1}^{n} X_i = r\right] = \sum_{j=1}^{r} j f_j f_{r-j}^{(n-1)^*} / f_r^{n^*} \tag{3.5}$$

$$p_n f_r^{n^*} = \sum_{j=1}^{r-1} (a + bj/r) f_j p_{n-1} f_{r-j}^{(n-1)^*} \tag{3.6}$$

Obje formule (3.4) i (3.5) vrijede za svaku vrijednost r za koju $f_r^{n^*}$ nije nula; (3.6) vrijedi za $r = 1, 2, \ldots$, bez obzira da li je $f_r^{n^*}$ nula ili ne.
Formula (3.4) slijedi iz simetrije: $X_1, X_2, \ldots, X_n$ su jednako distribuirane, tako da ako je njihov zbroj r, očekivanje svake od njih tada mora biti r/n.
Da biste vidjeli zašto vrijedi (3.5), uočite da je $f_j f_{r-j}^{(n-1)^*}/f_r^{n^*}$ (uvjetna) vjerojatnost da je X_1 jednako j uz dano da je $\sum_{j=1}^n X_j$ jednako r. (U (3.5) je pretpostavljeno da vjerojatnost da je $\sum_{j=1}^n X_j$ jednaka r, tj. $f_r^{n^*}$, nije nula.) Uz dano da je $\sum_{j=1}^n X_j$ jednako r, vrijednost od X_1 ne može biti veća od r. Stoga je desna strana od (3.5) zbroj, po svim vrijednostima koje X_1 može primiti, tih vrijednosti pomnoženih s vjerojatnošću da X_1 poprimi tu vrijednost, uvjetno na $\sum_{j=1}^n X_j$ jednako r. To je tada jednako lijevoj strani od (3.5).
Sada izvodimo (3.6). Prvo primjetite da (3.6) vrijedi ako je $f_r^{n^*}$ jednako nula, budući da u tom slučaju za svaku vrijednost od $j = 1, 2, \ldots, r$, ili f_j ili $f_{r-j}^{(n-1)^*}$, ili oba, moraju biti nula. Dakle, ako je $f_r^{n^*}$ nula, obje strane u (3.6) su nula. Pretpostavimo sada da $f_r^{n^*}$ nije nula. Tada je

$$\begin{aligned}
p_n f_r^{n^*} &= p_{n-1}(a + b/n) f_r^{n^*} && \text{po (3.1)} \\
&= p_{n-1}\mathbb{E}[a + bX_1/r \mid \textstyle\sum_{i=1}^n X_i = r] f_r^{n^*} && \text{po (3.4)} \\
&= p_{n-1} \textstyle\sum_{j=1}^r (a + bj/r) f_j f_{r-j}^{(n-1)^*} && \text{po (3.5)} \\
&= p_{n-1} \textstyle\sum_{j=1}^{r-1} (a + bj/r) f_j f_{r-j}^{(n-1)^*} && (\text{zbog } f_0^{(n-1)^*} = 0)
\end{aligned}$$

Konačno, sada možemo izvesti (3.3). Za $r = 1, 2, \ldots$

$$\begin{aligned}
g_r &= \textstyle\sum_{n=1}^\infty p_n f_r^{n^*} && \text{po (2.2)} \\
&= p_1 f_r + \textstyle\sum_{n=1}^\infty p_{n+1} f_r^{(n+1)^*} \\
&= (a + b) p_0 f_r + \textstyle\sum_{n=1}^\infty \sum_{j=1}^{r-1} (a + bj/r) f_j p_n f_{r-j}^{n^*} && \text{po (3.6)} \\
&= (a + b) g_0 f_r + \textstyle\sum_{j=1}^{r-1} (a + bj/r) f_j \sum_{n=1}^\infty p_n f_{r-j}^{n^*} \\
&= (a + b) g_0 f_r + \textstyle\sum_{j=1}^{r-1} (a + bj/r) f_j g_{r-j} && \text{po (2.2)} \\
&= \textstyle\sum_{j=1}^r (a + bj/r) f_j g_{r-j}
\end{aligned}$$

što dokazuje formulu (3.3).

U važnom specijalnom slučaju kada N ima Poissonovu distribuciju, tako da je $a = 0$ i $b = \lambda$, te se formule pojednostavljuju na

$$\begin{aligned} g_0 &= e^{-\lambda} \\ g_r &= \tfrac{\lambda}{r} \textstyle\sum_{j=1}^{r} j f_j g_{r-j} \,. \end{aligned}$$

3.3 Normalna aproksimacija za $G(x)$

Rekurzivna formula za računanje $G(x)$ o kojoj smo govorili u Sekciji 3.2 je koristan alat, ali ima neke nedostatke. Kao prvo, u nekim primjenama ipak može biti potrebna značajna količina kompjutorskog vremena za izračun vrijednosti od $G(x)$. Drugo, formula se ne može koristiti ako nisu poznate, ili se ne mogu dovoljno precizno procijeniti, distribucije od N i X_i.

U ovoj podsekciji sve što se pretpostavlja o S je da su poznati, ili da se pouzdano mogu procijeniti, njezino očekivanje i varijanca. Budući da razne distribucije imaju isto očekivanje i varijancu, $G(x)$ se ne može izračunati upotrebom samo te informacije. Jedan način da se u takvoj situaciji aproksimira $G(x)$ je da se pretpostavi da je S približno normalno distribuirana.

Formalnije, neka je $\Phi(z)$ funkcija distribucije normalno distribuirane slučajne varijable s očekivanjem 0 i varijancom 1, tako da je

$$\Phi(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} \exp\{-x^2/2\}\, dx \,.$$

Neka sada μ i σ^2 označavaju očekivanje i varijancu od S. Pretpostavka u ovoj sekciji je da je S približno normalno distribuirana s očekivanjem μ i varijancom σ^2 tako da je za svaki x

$$G(x) = \mathbb{P}(S \leq x) = \mathbb{P}((S-\mu)/\sigma \leq (x-\mu)/\sigma) \approx \Phi((x-\mu)/\sigma)\,.$$

Lagano je dobiti vrijednosti vjerojatnosti za normalnu distribuciju.

S je zbroj slučajnog broja nezavisnih, jednako distribuiranih slučajnih varijabli. Centralni granični teorem sugerira da bi normalna aproksimacija mogla biti razumna. Čim je veća (očekivana) vrijednost od N (broja slučajnih varijabli koje se sumiraju), tim se očekuje bolja aproksimacija.

3.4 Translatirana gama aproksimacija za $G(x)$

Pretpostavimo sada da su poznata prva tri momenta od S, a ne samo prva dva, ili da se mogu procijeniti s razumnom pouzdanoččću. Drugi način aproksimacije distribucije od S je pomoću translatirane gama distribucije. Neka redom μ, σ^2 i β označavaju očekivanje, varijancu i koeficijent asimetrije od S. Kod translatirane gama aproksimacije pretpostavlja se da S ima približno istu distribuciju kao slučajna varijabla $k+Y$, gdje je k konstanta, a Y ima gama distribuciju s parametrima α i δ. Parametri k, α i δ izabrani su tako da $k+Y$ ima ista prva tri momenta kao i S. Uočite da je $k+Y$ naprosto gama slučajna varijabla, Y, čije su vrijednosti translatirane za pozitivni ili negativni iznos k. Jedan razlog zašto translatirana gama distribucija općenito daje bolju prilagodbu nego normalna distribucija je da je gama distribucija pozitivno asimetrična, isto kao i S u mnogim praktičnim situacijama.
Koeficijent asimetričnosti, β, gama distribucije s parametrima α i δ je $2/\sqrt{\alpha}$. Izjednačavanje koeficijenta asimetrije, varijance i očekivanja od S i $k+Y$ daje sljedeće tri formule

$$\begin{aligned} \beta &= 2/\sqrt{\alpha} \\ \sigma^2 &= \alpha/\delta^2 \\ \mu &= k+\alpha/\delta \end{aligned}$$

iz kojih se mogu izračunati α, δ i k iz poznatih vrijednoti β, σ^2 i μ.
Razlog za aproksimaciju distribucije od S translatiranom gama, ili normalnom, distribucijom, je taj da može biti lakše dobiti vrijednosti za veličine kao što je $\mathbb{P}(a < k+Y < b)$ nego za $\mathbb{P}(a < S < b)$. Vjerojatnosti za gama distribuciju brzo se mogu dobiti pomoću većine statističkih paketa.

4 Model individualnog rizika

U tom modelu promatra se portfelj koji se sastoji od fiksnog broja rizika. Pretpostavlja se

- da su ti rizici nezavisni
- iznosi šteta iz tih rizika nisu jednako distribuirane slučajne varijable
- broj rizika se ne mijenja za vrijeme trajanja osigurateljnog pokrića.

Kao i prije, skupne štete iz tog portfelja označavat će se s S. Dakle,

$$S = Y_1 + Y_2 + \cdots + Y_n$$

gdje Y_j označava iznos šteta nastalih po j-tom riziku, a n označava broj rizika. Moguće je da po nekim rizicima ne nastane šteta. Prema tome, neke od opaženih vrijednosti od $\{Y_j\}_{j=1}^n$ mogu biti 0.
Za svaki od rizika imamo sljedeće pretpostavke:

– broj odštetnih zahtjeva po j-tom riziku, N_j, je ili 0 ili 1 (4.1)

– vjerojatnost štete iz j-tog rizika je q_j (4.2)

Ako šteta nastane po j-tom riziku, iznos štete označava se slučajnom varijablom X_j. Neka $F_j(x)$, μ_j i σ_j^2 označavaju redom funkciju distribucije, očekivanje i varijancu od X_j.
Pretpostavka (4.1) je vrlo restriktivna. Ona znači da model dozvoljava jednu štetu po riziku. To uključuje rizike kao što je jednogodišnje osiguranje za slučaj smrti, ali isključuje mnogo tipova općih polica osiguranja. Na primjer, ne postoje restrikcije na broj odštetnih zahtjeva u godini po polici osiguranja kućanstva.
Postoje tri glavne razlike između tog modela i modela kolektivnog rizika:

(1) Broj rizika u portfelju je specificiran. U modelu kolektivnog rizika nije bilo potrebno specificirati taj broj, niti pretpostaviti da je fiksan za vrijeme trajanja osigurateljnog pokrića (čak niti kada je bilo pretpostavljeno $N \sim b(n, q)$.)

(2) Broj odštetnih zahtjeva po svakom individualnom riziku je ograničen. Takvog ograničenja nije bilo u modelu kolektivnog rizika.

(3) Pretpostavljeno je da su individualni rizici nezavisni. U modelu kolektivog rizika bilo je pretpostavljeno da su iznosi individualnih šteta nezavisni.

Pretpostavke (4.1) i (4.2) kažu da je $N_j \sim b(1, q_j)$. Prema tome, distribucija od Y_j je složena binomna, uz individualne štete distribuirane kao X_j. Iz formula (2.13) i (2.14) odmah slijedi da je

$$\mathbb{E}[Y_j] = q_j \mu_j \tag{4.3}$$

$$\mathrm{Var}[Y_j] = q_j \sigma_j^2 + q_j(1 - q_j)\mu_j^2 \tag{4.4}$$

S je zbroj od n nezavisnih složenih binomnih slučajnih varijabli. U Sekciji 2.4 vidjeli smo da ne postoji opći rezultat o distribuciji takvog zbroja. Ta se distribucija može zapisati samo kada su složene binomne slučajne varijable jednako distribuirane, uz to što su nezavisne. Pod nekim uvjetima je moguće, ali komplicirano, izračunati funkciju distribucije od S. Međutim, lako je pronaći očekivanje i varijancu od S.

$$\mathbb{E}[S] = \mathbb{E}\left[\sum_{j=1}^{n} Y_j\right] = \sum_{j=1}^{n} \mathbb{E}[Y_j] = \sum_{j=1}^{n} q_j \mu_j \tag{4.5}$$

Pretpostavka da su individualni rizici nezavisni potrebna je da bi se zapisalo

$$\mathrm{Var}[S] = \mathrm{Var}\left[\sum_{j=1}^{n} Y_j\right] = \sum_{j=1}^{n} \mathrm{Var}[Y_j] = \sum_{j=1}^{n}(q_j\sigma_j^2 + q_j(1-q_j)\mu_j^2) \tag{4.6}$$

U specijalnom slučaju kada je $\{Y_j\}_{j=1}^{n}$ niz jednako distribuiranih, te i dalje nezavisnih, slučajnih varijabli, za svaku policu su vrijednosti q_j, μ_j i σ_j^2 identične, recimo q, μ i σ^2. Također, $F_j(x)$ ne ovisi o j, recimo $F(x)$. Dakle, S je složena binomna, s binomnim parametrima n i q, te individualnim štetama s funkcijom distribucije $F(x)$. U ovom specijalnom slučaju, to se svodi na model kolektivnog rizika, te se iz (4.5) i (4.6) može vidjeti da je

$$\begin{aligned}
\mathbb{E}[S] &= nq\mu \\
\mathrm{Var}[S] &= nq\sigma^2 + nq(1-q)\mu^2
\end{aligned}$$

što odgovara redom (2.13) i (2.14).

5 Varijabilnost/neizvjesnost parametara

5.1 Uvod

Do sada smo proučavali modele rizika uz pretpostavku da su sa sigurnošću poznati parametri, to jest momenti a u nekim slučajevima čak i distribucije, broja odštetnih zahtjeva i iznosa individualnih šteta. Općenito ti parametri neće biti poznati, nego će ih trebati procijeniti iz odgovarajućeg skupa podataka. U ovoj sekciji vidjet će se kako se ranije uvedeni modeli mogu proširiti tako da dozvoljavaju varijabilnost/neizvjesnost parametara. To će biti učinjeno promatranjem niza primjera. Većina, ali ne svi, od tih primjera razmatrat će neizvjesnoat u distribuciji broja odštetnih zahtjeva, budući da je to, a ne distribucija iznosa individualnih šteta, dobilo više pažnje u aktuarskoj literaturi. Svi primjeri biti će zasnovani na Poissonovoj distribuciji broja odštetnih zahtjeva.

5.2 Varijabilnost u heterogenom portfelju

Promotrimo portfelj koji se sastoji od n nezavisnih polica. Skupne štete nastale po i-toj polici označene su slučajnom varijablom S_i, gdje S_i ima složenu Poissonovu distribuciju s parametrima λ_i i $F(x)$. Uočite da je, zbog jednostavnosti, pretpostavljeno da je distribucija iznosa individualnih šteta, $F(x)$, jednaka za sve police. U ovom primjeru pretpostavlja se da je distribucija iznosa individualnih šteta, tj. $F(x)$, poznata, ali su vrijednosti Poissonovih parametara, tj. λ_i, nepoznate. U ovoj podsekciji pretpostavlja se da su λ_i (uzoračke vrijednosti) nezavisnih slučajnih varijabli s istom (poznatom) distribucijom. Drugim riječima, $\{\lambda_i\}_{i=1}^n$ se tretira kao skup nezavisnih i jednako distribuiranih slučajnih varijabli s poznatom distribucijom. To znači da ako je iz portfelja izabrana polica na slučajan način, tada se pretpostavlja da njezin Poissonov parametar nije poznat, ali se o njemu mogu stvarati vjerojatnosne tvrdnje. Na primjer, "postoji 50% šanse da je njezin Poissonov parametar između 3 i 5". Važno je razumjeti da je Poissonov parametar police izabrane iz portfelja fiksan broj; problem je da je taj broj nepoznat.

5.2.1 Primjer

Pretpostavimo da su Poissonovi parametri polica iz portfelja nepoznati, ali jednako vjerojatni 0.1 i 0.3.

(i) Pronađite očekivanje i varijancu (pomoću m_1 i m_2) skupnih šteta slučajno odabrane police iz portfelja.

(ii) Pronađite očekivanje i varijancu (pomoću m_1, m_2 i n) skupnih šteta cijelog portfelja.

Moglo bi pomoći da se o ovome razmišlja kao o modelu dijela portfelja osiguranja motornih vozila. Police u cijelom portfelju podijeljene su po svojim vrijednostima faktora za određivanje premije kao što su "dob vozača", "tip vozila", te čak i "proteklo iskustvo o štetama". Police u razmatranom dijelu portfelja imaju jednake vrijednosti tih faktora za određivanje premije. Međutim, postoje neki faktori, kao na primjer "vještina vožnje", koji se ne mogu lako mjeriti, te ih se ne može eksplicitno uračunati. Pretpostavlja se da su neki od osiguranika u tom portfelju "dobri" vozači, a ostali su "loši" vozači. Distribucija iznosa individualnih šteta jednaka je za sve vozače, ali je broj nastalih šteta za "dobre" vozače manji (prosječno 0.1 godišnje) nego za "loše" vozače (prosječno 0.3 godišnje). Pretpostavlja se da je poznato, moguće iz nacionalnih podataka, da je osiguranik u tom dijelu portfelja s jednakom vjerojatnošću "dobar" vozač ili "loš" vozač, ali da se ne može znati da li je određeni osiguranik "dobar" vozač ili "loš" vozač.

Rješenje

Neka je λ_i, $i = 1, 2, \ldots, n$, Poissonov parametar i-te police u portfelju. $\{\lambda_i\}_{i=1}^n$ se promatra kao skup nezavisnih i jednako distribuiranih slučajnh varijabli, svaka sa sljedećom distribucijom

$$\begin{aligned} \mathbb{P}(\lambda_i = 0.1) &= 0.5 \\ \mathbb{P}(\lambda_i = 0.3) &= 0.5\,. \end{aligned}$$

Iz toga

$$\begin{aligned} \mathbb{E}[\lambda_i] &= 0.2 \\ \mathrm{Var}[\lambda_i] &= 0.01\,. \end{aligned}$$

(i) Momenti od λ_i mogu se izračunati uvjetovanjem na vrijednost od λ_i. Budući da $S_i|\lambda_i$ ima složenu Poissonovu distribuciju, formule (2.8) i (2.9) mogu se upotrijebiti za

$$\mathbb{E}[S_i] = \mathbb{E}[\mathbb{E}[S_i|\lambda_i]] = \mathbb{E}[\lambda_i m_1] = 0.2 m_1$$

$$\begin{aligned} \mathrm{Var}[S_i] &= \mathbb{E}[\mathrm{Var}[S_i|\lambda_i]] + \mathrm{Var}[\mathbb{E}[S_i|\lambda_i]] \\ &= \mathbb{E}[\lambda_i m_2] + \mathrm{Var}[\lambda_i m_1] \\ &= 0.2 m_2 + 0.01 m_1^2\,. \end{aligned}$$

(ii) Slučajne varijable $\{S_i\}_{i=1}^n$ su nezavisne i jednako distribuirane, svaka sa distribucijom od S_i danom u dijelu (i). Dakle, možemo upotrijebiti rezultat iz (i) za

$$\begin{aligned}
\mathbb{E}\left[\sum_{i=1}^{n} S_i\right] &= n\mathbb{E}[S_i] = 0.2nm_1 \\
\mathrm{Var}\left[\sum_{i=1}^{n} S_i\right] &= n\mathrm{Var}[S_i] = 0.2nm_2 + 0.01nm_1^2 \,.
\end{aligned}$$

5.2.2 Primjer

Pretpostavimo sa su Poissonovi parametri individualnih polica izvučeni iz gama distribucije s parametrima α i δ. Nađite distribuciju broja odštetnih zahtjeva nastalih po slučajno odabranoj polici iz portfelja.

Rješenje

Neka N_i označava broj odštetnih zahtjeva po i- toj polici u portfelju, i neka je λ_i njen Poissonov parametar. Tada N_i ima Poissonovu distribuciju s parametrom λ_i, ali je problem (po pretpostavci) da je vrijednost od λ_i nepoznata. Što *jeste* poznato je distribucija iz koje je λ_i odabrano. Problem se može sažeti kako slijedi:

Uz dano

$$N_i|\lambda_i \sim P(\lambda_i) \text{ i } \ \lambda_i \sim G(\alpha, \delta)$$

nađite marginalnu distribuciju od N_i.

Ovaj se problem može riješiti uklanjanjem uvjetovanja na uobičajeni način:

Za $x = 0, 1, 2, \ldots$

$$\begin{aligned}
\mathbb{P}(N_i = x) &= \int_0^\infty \exp\{-\lambda\} \frac{\lambda^x}{x!} \frac{\delta^\alpha}{\Gamma(\alpha)} \lambda^{\alpha-1} \exp\{-\delta\lambda\} \, d\lambda \\
&= \frac{\delta^\alpha}{\Gamma(\alpha) x!} \int_0^\infty \exp\{-\lambda(\delta+1)\} \lambda^{x+\alpha-1} \, d\lambda \,.
\end{aligned}$$

Integral se računa uspoređujući integrand s gama gustoćom, tako da je

$$\mathbb{P}(N_i = x) = \frac{\delta^\alpha}{\Gamma(\alpha) x!} \frac{\Gamma(x+\alpha)}{(\delta+1)^{x+\alpha}}$$

što pokazuje da je marginalna distribucija od N_i negativna binomna s parametrima α i $\delta/(\delta+1)$.

5.3 Varijabilnost u homogenom portfelju

Sada razmatramo drugačiji primjer. Pretpostavimo, kao i do sada, da imamo portfelj sa n polica. Skupne štete nastale po pojedinačnoj polici imaju složenu Poissonovu distribuciju s parametrima λ i $F(x)$. Ti parametri su isti za sve police u portfelju. Kada bi vrijednost od λ bila poznata, skupne štete nastale po različitim policama bile bi međusobno nezavisne. Pretpostavljeno je da vrijednost od λ nije poznata, možda zato jer se mijenja iz godine u godinu, ali da *postoje* neki pokazatelji vjerojatnosti da će λ biti u danom skupu vrijednosti. Kao i u prethodnom primjeru, pretpostavlja se, zbog jednostavnosti, da nema neizvjesnosti oko momenata distribucije iznosa individualnih šteta, tj. oko $F(x)$. Neizvjesnost oko vrijednosti od λ može se modelirati shvaćajući λ kao slučajnu varijablu (s poznatom distribucijom).

5.3.1 Primjer

Pretpostavimo da će Poissonov parametar, λ, biti jednak 0.1 ili 0.3 s otprilike jednakom vjerojatnošću.

(i) Izračunajte očekivanje i varijancu (pomoću m_1 i m_2) skupnih šteta po slučajno odabranoj polici iz portfelja.

(ii) Izračunajte očekivanje i varijancu (pomoću m_1, m_2 i n) skupnih šteta iz cijelog portfelja.

Rješenje
Koristeći iste oznake kao i prije, neka S_i označava skupne štete po i-toj polici u portfelju. Situacija se može sažeti kako slijedi:
Slučajne varijable $\{S_i|\lambda\}_{i=1}^n$ su nezavisne i jednako distribuirane, svaka sa složenom Poissonovom distribucijom s parametrima λ i $F(x)$. Slučajna varijabla λ ima sljedeću distribuciju

$$
\begin{aligned}
\mathbb{P}(\lambda = 0.1) &= 0.5 \\
\mathbb{P}(\lambda = 0.3) &= 0.5\,.
\end{aligned}
$$

(i) Uvjetovanjem na vrijednost od λ,

$$
\begin{aligned}
\mathbb{E}[S_i] &= \mathbb{E}[\mathbb{E}(S_i|\lambda)] = \mathbb{E}[\lambda m_1] = 0.2m_1 \\
\mathrm{Var}[S_i] &= \mathbb{E}[\mathrm{Var}(S_i|\lambda)] + \mathrm{Var}[\mathbb{E}(S_i|\lambda)] = \mathbb{E}[\lambda m_2] + \mathrm{Var}[\lambda m_1] \\
&= 0.2m_2 + 0.01m_1^2\,.
\end{aligned}
$$

(ii)

$$\begin{aligned}
\mathbb{E}\left[\sum_{i=1}^{n} S_i\right] &= n\mathbb{E}[S_1] = 0.2nm_1 \quad (\text{zbog } \{S_i\}_{i=1}^{n} \text{ jednako distribuirane}) \\
\mathrm{Var}\left[\sum_{i=1}^{n} S_i\right] &= \mathbb{E}\left[\mathrm{Var}\left(\sum_{i=1}^{n} S_i|\lambda\right)\right] + \mathrm{Var}\left[\mathbb{E}\left(\sum_{i=1}^{n} S_i|\lambda\right)\right] \\
&= \mathbb{E}[n\lambda m_2] + \mathrm{Var}[n\lambda m_1] \\
&= 0.2nm_2 + 0.01n^2m_1^2\,.
\end{aligned}$$

Korisno je usporediti odgovore za gornji primjer s onima iz Primjera 5.2.1. Očekivane vrijednosti su u svim slučajevima iste, kao i varijance kada se razmatra pojedinačna polica (dio (i)). Razlika nastaje kada se razmatraju varijance za više od jedne police (dio(ii)). U tom slučaju drugi primjer daje veću varijancu. Važno je razumjeti razlike (i sličnosti) između ta dva primjera. Praktična situacija u kojoj je drugi odgovarajući bio bi portfelj polica koje osiguravaju zgrade u nekom području. Broj odštetnih zahtjeva može ovisiti, uz druge faktore, i o vremenu kroz godinu; neuobičajeno velik broj oluja rezultira visokim očekivanim brojem odštetnih zahtjeva (tj. visokom vrijednošću od λ) i obratno za sve police zajedno.

5.4 Varijabilnost u broju odštetnih zahtjeva i iznosima šteta i neizvjesnost parametara

Ova sekcija sadrži još dva primjera. Prvi primjer je prilično kompliciran i sadrži neizvjesnost o iznosima šteta, kao i neizvjesnost o broju odštetnih zahtjeva.

5.4.1 Primjer

Osigurateljno društvo modelira štete od oluja nastale po osigurateljnoj polici kućanstva koristeći sljedeće pretpostavke.
Pretpostavlja se da broj oluja u svakoj godini, K, ima Poissonovu distribuciju s parametrom λ.
Pretpostavlja se da broj odštetnih zahtjeva prijavljenih po i-toj oluji, N_i, $i = 1, 2, \ldots, K$, ima Poissonovu distribuciju s parametrom Θ_i.
Pretpostavlja se da su parametri Θ_i, $i = 1, 2, \ldots, K$, nezavisne i jednako distribuirane slučajne varijable s $\mathbb{E}(\Theta_i) = n$ i $\mathrm{Var}(\Theta_i) = s_1^2$.

Iznos j-te štete nastale u i-toj oluji, X_{ij}, $j = 1, 2, \ldots, N_i$, ima lognormalnu distribuciju s parametrima μ_i i σ, gdje se pretpostavlja da je σ poznat. Pretpostavlja se da su očekivani iznosi šteta, $\Lambda_i = \exp(\mu_i + \sigma^2/2)$ nezavisne i jednako distribuirane slučajne varijable s očekivanjem p i varijancom s_2^2. Također se pretpostavlja da su Θ_i i Λ_i nezavisne.

(i) Pokažite da je $\mathbb{E}[X_{ij}] = p$ i $\mathrm{Var}[X_{ij}] = \exp\{\sigma^2\}(p^2 + s_2^2) - p^2$.

(ii) Neka S_i označava skupne štete nastale iz i-te oluje, tako da je $S_i|\{\Theta_i, \Lambda_i\}$ složena Poissonova slučajna varijabla.

Pokažite da je $\mathbb{E}[S_i] = np$ i

$$\mathrm{Var}[S_i] = (p^2 + s_2^2)(n^2 + s_1^2)(n^2 + s_1^2 + n\exp\{\sigma^2\}) - n^2p^2 \,.$$

(iii) Nađite izraz za očekivanje i varijancu godišnjih skupnih štete nastalih iz svih oluja.

Rješenje

(i)

$$\begin{aligned}
\mathbb{E}[X_{ij}] &= \mathbb{E}[\mathbb{E}(X_{ij}|\Lambda_i)] = \mathbb{E}[\Lambda_i] = p \\
\mathrm{Var}[X_{ij}] &= \mathbb{E}[\mathrm{Var}(X_{ij}|\Lambda_i)] + \mathrm{Var}[\mathbb{E}(X_{ij}|\Lambda_i)] \\
&= \mathbb{E}[\Lambda_i^2(\exp\{\sigma^2\} - 1)] + \mathrm{Var}(\Lambda_i) \\
&= (p^2 + s_2^2)(\exp\{\sigma^2\} - 1) + s_2^2 \\
&= (p^2 + s_2^2)\exp\{\sigma^2\} - p^2
\end{aligned}$$

(ii)

$$\mathbb{E}[S_i] = \mathbb{E}[\mathbb{E}(S_i|\Theta_i, \Lambda_i)] = \mathbb{E}[\Theta_i\Lambda_i] = np$$

budući da su Θ_i i Λ_i nezavisne.

Nadalje, budući da $S_i|\{\Theta_i, \Lambda_i\}$ ima složenu Poissonovu distribuciju,

$$\mathrm{Var}[S_i|\Theta_i, \Lambda_i] = \Theta_i\mathbb{E}[X_{ij}^2|\Lambda_i] = \Theta_i(\Lambda_i^2\exp\{\sigma^2\}) \,,$$

te stoga

$$\mathbb{E}[\mathrm{Var}(S_i|\Theta_i, \Lambda_i)] = n(p^2 + s_2^2)\exp\{\sigma^2\} \,.$$

Također,

$$\begin{aligned}\mathrm{Var}[\mathbb{E}(S_i|\Theta_i,\Lambda_i)] &= \mathrm{Var}[\Theta_i,\Lambda_i] = \mathbb{E}[\Theta_i^2\Lambda_i^2] - n^2p^2 \\ &= (n^2+s_1^2)(p^2+s_2^2) - n^2p^2\,.\end{aligned}$$

Zadnja dva rezultata zajedno daju

$$\mathrm{Var}[S_i] = (n^2+s_1^2)(p^2+s_2^2) - n^2p^2 + n(p^2+s_2^2)\exp\{\sigma^2\}\,.$$

(iii) Neka slučajna varijabla R označava godišnje skupne štete nastale iz svih oluja. Tada se R može zapisati

$$R = \sum_{i=1}^{K} S_i$$

gdje K ima Poissonovu distribuciju, a slučajne varijable $\{S_i\}$ su nezavisne jednako distribuirane.

Dakle, R ima složenu Poissonovu distribuciju, te je stoga

$$\begin{aligned}\mathbb{E}[R] &= \lambda\mathbb{E}(S_i) = \lambda np \\ \mathrm{Var}[R] &= \lambda\mathbb{E}(S_i^2) = \lambda(\mathrm{Var}[S_i] + \mathbb{E}[S_i]^2) \\ &= \lambda(p^2+s_2^2)(n^2+s_1^2+n\exp\{\sigma^2\})\,.\end{aligned}$$

5.4.2 Primjer

Svake godine osigurateljno društvo izda izvjestan broj osigurateljnih polica za kućanstvo, za svaku od kojih je godišnja premija £80. Skupne godišnje štete nastale po pojedinačnoj polici imaju složenu Poissonovu distribuciju; Poissonov parametar je 0.4, a individualni iznosi šteta imaju gama distribuciju s parametrima α i λ. Trošak rješavanja štete je slučajna varijabla uniformno distribuirana između £50 i £$b\,(>$ £50). Iznos troškova nezavisan je od iznosa pridružene štete. Slučajna varijabla S predstavlja ukupne skupne štete i troškove u godini iz tog portfelja. Može se pretpostaviti da S ima približno normalnu distribuciju.

(i) Pretpostavimo da je

$$\alpha = 1;\ \lambda = 0.01;\ b = 100.$$

Pokažite da društvo mora godišnje prodati barem 884 police da bi s 99%-tnom sigurnošću prihod od premija nadmašio štete i troškove.

(ii) Pretpostavimo sada da vrijednosti od α, λ i b nisu sa sigurnošću poznate, nego da mogu biti bilo gdje u sljedećim intervalima

$$0.95 \leq \alpha \leq 1.05;\ 0.009 \leq \lambda \leq 0.011;\ 90 \leq b \leq 110.$$

Promatrajući što bi za osigurateljno društvo bila najgora moguća kombinacija vrijednosti od α, λ i b, izračunajte broj polica koje društvo mora prodati da bi s barem 99%-tnom sigurnošću prihod od premija nadmašio štete i troškove.

Rješenje

Neka je X_i iznos i-te štete, a Y_i iznos pridruženih troškova. Neka je N ukupan broj odštetnih zahtjeva za portfelj, i neka je n broj polica u portfelju. Tada N ima Poissonovu distribuciju s parametrom $0.4n$ i S se može zapisati kao

$$S + \sum_{i=1}^{N}(X_i + Y_i)$$

gdje je $\{X_i + Y_i\}_{i=1}^{\infty}$ niz nezavisnih i jednako distribuiranih slučajnih varijabli, nezavisnih od N. Odavde se može vidjeti da S ima složenu Poissonovu distribuciju gdje $(X_i + Y_i)$ predstavljaju "iznos i-te individualne štete". Sada se mogu iskoristiti standardni rezultati za zapis sljedećih formula za momente od S:

$$\begin{aligned}
\mathbb{E}[S] &= 0.4n\mathbb{E}[X_i + Y_i] \\
\mathrm{Var}[S] &= 0.4n\mathbb{E}[(X_i + Y_i)^2] \\
&= 0.4n(\mathbb{E}[X_i^2] + 2\mathbb{E}[X_iY_i] + \mathbb{E}[Y_i^2])\,.
\end{aligned}$$

Momenti od X_i i Y_i su pomoću α, λ i b kako slijedi:

$$\begin{aligned}
\mathbb{E}[X_i] &= \alpha/\lambda & \mathbb{E}[Y_i] &= (b+50)/2 \\
\mathbb{E}[X_i^2] &= \alpha(\alpha+1)/\lambda^2 & \mathbb{E}[Y_i^2] &= (b^2 + 50b + 2500)/3
\end{aligned}$$

$$\mathbb{E}[X_iY_i] = \mathbb{E}[X_i]\mathbb{E}[Y_i]$$

gdje zadnja relacija slijedi iz nezavisnosti od X_i i Y_i.

(i) Stavimo sada

$$\alpha = 1;\ \lambda = 0.01;\ b = 100$$

u te formule da bismo pokazali da je

$$\mathbb{E}[S] = 70n \text{ i } \operatorname{Var}[S] = 127.80^2 n\,.$$

Stoga S ima približno normalnu distribuciju s očekivanjem $70n$ i standardnom devijacijom $127.80\sqrt{n}$. Prihod od premija je $80n$ i traži se najmanja vrijednost od n takva da je

$$\mathbb{P}(S < 80n) \geq 0.99\,.$$

Standardizacijom od S kao što je uobičajeno kod normalnih distribucija, to postaje

$$\mathbb{P}\left[\frac{(S-70n)}{127.80\sqrt{n}} < \frac{(80n-70n)}{127.80\sqrt{n}}\right] \geq 0.99\,.$$

Gornja 99%-tna točka standardne normalne distribucije je 2.326, te je stoga uvjet za n

$$\frac{(80n-70n)}{127.80\sqrt{n}} \geq 2.326$$

što daje

$$n \geq 883.7 \quad (\text{ili } n \geq 884 \text{ do sljedećeg većeg cijelog broja}).$$

(ii) Za reosigurateljno društvo je najgora moguća kombinacija vrijednosti od α, λ i b ona koja daje najviše moguće vrijednosti za $\mathbb{E}[S]$ i $\operatorname{Var}[S]$. Da bismo to vidjeli, neka μ i σ označavaju očekivanje i standardnu devijaciju skupnih šteta i troškova po pojedinačnoj polici. I μ i σ će biti funkcije od α, λ i b, te

$$\mathbb{E}[S] = n\mu \text{ i } \operatorname{Var}[S] = n\sigma^2\,.$$

Slijedivši iste korake kao u dijelu (i), uvjet na n je

$$\frac{(80-\mu)\sqrt{n}}{\sigma} \geq 2.326$$

što postaje

$$n \geq [2.326\sigma/(80-\mu)]^2\,.$$

Dakle, najviša vrijednost za n rezultira iz najviših vrijednosti za μ i σ (uz uvjet da je najviša vrijednost za μ manja od 80). Sada uočite da je

$$\mu = 0.4\mathbb{E}[X_i + Y_i] \text{ i } \sigma^2 = 0.4\mathbb{E}[(X_i + Y_i)^2]\,.$$

Iz formula za momente od X_i i Y_i danih gore, μ i σ su maksimalni kada su α i b najveći mogući, a λ najmanji moguć, tj. kada je

$$\alpha = 1.05; \quad \lambda = 0.009; \quad b = 110.$$

Te kombinacije vrijednosti daju

$$\mu = 78.67 \text{ i } \sigma = 144.14$$

tako da n mora biti barem 63 546 da bi osigurateljno društvo barem 99% bilo sigurno da će prihod od premija nadmašiti štete i troškove.

POGLAVLJE 5 - TEORIJA NESOLVENTNOSTI

Nastavni ciljevi:

(vi)1. Objasniti što se podrazumijeva pod procesom ukupnih šteta i procesom toka novca za rizik.

(vi)2. Definirati vjerojatnost propasti u beskonačnom/konačnom vremenu i neprekidnom/diskretnom vremenu i navesti i objasniti odnose između različitih vjerojatnosti propasti.

(vi)3. Definirati Poissonov proces, izvesti distribuciju broja događaja u danom vremenskom intervalu, izvesti distribuciju vremena između događaja, te primjeniti te rezultate.

(vi)4. Definirati složen Poissonov proces i izvesti momente i funkciju izvodnicu momenata tog procesa.

(vi)5. Definirati koeficijent prilagodbe za složen Poissonov proces i za diskretno vremenske procese koji nisu složeni Poissonovi, izračunati ga u jednostavnim slučajevima i izvesti aproksimaciju.

(vi)6. Navesti Lundbergovu nejednakost i objasniti značaj koeficijenta prilagodbe.

(vi)7. Opisati utjecaj promjene vrijednosti parametara na vjerojatnost propasti, u konačnom i beskonačnom vremenu.

(vi)8. Analizirati utjecaj reosiguranja na koeficijent prilagodbe i vjerojatnost propasti.

0 Uvod

Tehnički detalj koji će biti potreban u ovom poglavlju:
Funkcija $f(x)$ opisana je kao $o(x)$ kada x ide prema nuli, ako

$$\lim_{x\to 0} \frac{f(x)}{x} = 0\,.$$

1 Osnovni pojmovi

1.1 Oznake

U četvrtom poglavlju proučavali smo ukupne štete nastale u portfelju polica kroz jedan vremenski period. U aktuarskoj literaturi često se koristi riječ "rizik" umjesto fraze "portfelj polica". U ovom poglavlju koristit ćemo oba termina tako da će riječ "rizik" značiti bilo jednu policu bilo kolekciju polica. U ovom poglavlju pomaknut ćemo razmatranja korak dalje promatrajući štete nastale u portfelju kroz sukcesivne vremenske periode. Potrebne su nam neke oznake.

$N(t)$	broj šteta nastalih u portfelju u vremenskom intervalu $[0, t]$ za sve $t \geq 0$
X_i	iznos i-te štete, $i = 1, 2, 3, \ldots$
$S(t)$	ukupna šteta u vremenskom intervalu $[0, t]$ za sve $t \geq 0$.

Ovdje je $\{X_i\}_{i=1}^{\infty}$ niz slučajnih varijabli. $\{N(t)\}_{t\geq 0}$ i $\{S(t)\}_{t\geq 0}$ su familije slučajnih varijabli, za svaki $t \geq 0$ po jedna; drugim riječima, $\{N(t)\}_{t\geq 0}$ i $\{S(t)\}_{t\geq 0}$ su slučajni procesi.

Vidi se da je $S(t) = \sum_{i=1}^{N(t)} X_i$ uz konvenciju daje suma jednaka nuli ako je $N(t)$ nula. Slučajni proces $\{S(t)\}_{t\geq 0}$ definiran gore zove se proces ukupnih šteta za rizik. Slučajne varijable $N(1)$, odn. $S(1)$, predstavljaju broj šteta, odn. ukupne štete, u portfelju u prvoj vremenskoj jedinici. Te dvije slučajne varijable odgovaraju slučajnim varijablama N i S uvedenim u poglavlju 4.

Osiguratelj ovog portfelja dobiva od osiguranika premije. U ovom trenutku prikladno je pretpostaviti, i ta pretpostavka će vrijediti kroz cijelo ppglavlje, da se premije primaju neprekidno i po konstantnoj stopi. Evo još malo oznaka:

c stopa uplata premija po jedinici vremena,

tako da je ukupna premija primljena u vremenskom intervalu $[0, t]$ jednaka ct. Također ćemo pretpostavit da je c strogo pozitivno.

1.2 Proces viška prihoda

Pretpostavimo da je u trenutku 0 osiguratelj na stranu stavio iznos novca za taj portfelj. Taj iznos novca naziva se početni višak i označavat ćemo ga s

U. Uvijek ćemo pretpostavljati da je $U \geq 0$. Osiguratelj treba početni višak prihoda, jer buduće premije možda neće biti dovoljne za pokriće budućih šteta. Osigurateljev višak u bilo koje buduće vrijeme $t > 0$ je slučajna varijabla, budući da njena vrijednost ovisi o iskustvu o štetama do trenutka t. Osigurateljev višak u trenutku t označavat ćemo s $U(t)$. Možemo napisati sljedeću formulu za $U(t)$:

$$U(t) = u + ct - S(t)\,.$$

Riječima ta formula kaže da osigurateljev višak u trenutku t jednak početnom višku plus prihod od premija do trenutka t minus ukupni iznos šteta do trenutka t. Uočite da početni višak i prihod od premija nisu slučajne varijable, jer su određene prije nego što počne proces rizika. Gornja formula vrijedi za sve $t \geq 0$ uz konvenciju da je $U(0)$ jednako U. Za dano t, $U(t)$ je slučajna varijabla, jer je $S(t)$ slučajna varijabla. Stoga je $\{U(t)\}_{t\geq 0}$ slučajni proces koji je poznat kao proces toka novca ili proces viška prihoda.
Slika 1. pokazuje jedan mogući ishod procesa viška. Štete se događaju u vremenima T_1, T_2, T_3, T_4 i T_5 i u tim trenucima višak trenutno pada za iznos štete. Između šteta višak raste po konstantnoj stopi c po jedinici vremena. Opisan model za osigurateljev višak uklučuje mnoga pojednostavljenja, isto kao i svaki model kompleksne situacije u stvarnom životu. Usprkos svojoj jednostavnosti, ovaj model daje zanimljiv uvid u matematiku osigurateljnih operacija.

1.3 Vjerojatnost propasti u neprekidnom vremenu

Sa slike 1. može se vidjeti da osigurateljev višak pada ispod nule kao rezultat štete u vremenu T_3. Govoreći na trenutak slobodnije, kada višak padne ispod nule osiguratelj je ostao bez novca, te se kaže da se dogodila *propast*. U ovom pojednostavljenom modelu, osiguratelj će željeti držati vjerojatnost tog događaja, t.j. vjerojatnost propasti, čim manjom mogućom, ili barem ispod predodređene granice. Još uvijek govoreći slobodnije, o propasti se može misliti da znači nesolventnost, iako je u praksi određivanje da li je osiguravajuće društvo nesolventno vrlo složen problem. Drugi način kako se može gledati na vjerojatnost propasti je da je to vjerojatnost da će u nekom budućem vremenu osiguravajuće društvo trebati dodatni kapital za financiranje tog određenog portfelja.

A sada preciznije. Definiramo sljedeće dvije vjerojatnosti:

$$\begin{aligned}\psi(U) &= \mathbb{P}[U(t) < 0, \text{za neki } t, 0 < t < \infty] \\ \psi(U,t) &= \mathbb{P}[U(\tau) < 0, \text{za neki } \tau, 0 < \tau \le t]\,.\end{aligned}$$

$\psi(U)$ je vjerojatnost propasti (uz dan početni višak U), a $\psi(U,t)$ je vjerojatnost propasti do trenutka t (uz dan početni višak U). Te vjerojatnosti se ponekad nazivaju vjerojatnost propasti u beskonačnom vremenu i vjerojatnost propasti u konačnom vremenu. Slijede neke važne logične relacije između tih vjerojatnosti za $0 < t_1 \le t_2$ i za $0 \le U_1 \le U_2$:

$$\psi(U_2,t) \le \psi(U_1,t) \tag{1.1}$$

$$\psi(U_2) \le \psi(U_1) \tag{1.2}$$

$$\psi(U,t_1) \le \psi(U,t_2) \le \psi(U) \tag{1.3}$$

$$\lim_{t\to\infty} \psi(U,t) = \psi(U) \tag{1.4}$$

Intuitivno objašnjenje tih realcija je sljedeće:
Čim je veći početni višak, manje je vjerojatno da će doći do propasti, bilo u konačnom vremenu, otkud (1.1), bilo u beskonačnom vremenu, otkud (1.2).
Za dani početnoi višak U, čim dulji vremenski period čekamo da li će doći do propasti, tim je vjerojatnost propasti veća, otkud (1.3).
Konačno, vjerojatnost propasti može se aproksimirati s vjerojatnosti propasti u konačnom vremenu t, ukoliko je t dovoljno velik, otkud (1.4).

1.4 Vjerojatnost propasti u diskretnom vremenu

Dvije vjerojatnosti propasti do sada promatane su neprekidno-vremenske vjerojatnosti propasti, jer provjeravamo propast u neprekidnom vremenu. U praksi, moguće je (i ponekad poželjno) provjeravati propast samo u diskretnim vremenskim intervalima.
Za dani vremenski interval, označimo ga s h, definiramo sljedeće dvije diskretno-vremenske vjerojatnosti propasti:

$$\begin{aligned}\psi_h(U) &= \mathbb{P}[U(t) < 0, \text{za neki } t, t = h, 2h, 3h, \ldots] \\ \psi_h(U,t) &= \mathbb{P}[U(\tau) < 0, \text{za neki } , t = h, 2h, \ldots, t-h, t]\,.\end{aligned}$$

Uočite da je u definiciji od $\psi_h(U,t)$ pretpostavljeno da je t cjelobrojni faktor od h. Slika 2. pokazuje istu realizaciju procesa viška kao i na slici 1. uz pretpostavku da proces provjeravamo samo u diskretnim vremenskim intervalima. Crne oznake pokazuju vrijednosti procesa viška u cjelobrojnim vremenskim intervalima (t.j., $h = 1$); crne oznake zajedno s bijelim oznakama pokazuju vrijednosti procesa viška u vremenskim intervalima duljine $1/2$.

Sa slike 2. može se vidjeti da u diskretnim vremenima sa $h = 1$ ne dolazi do propasti za tu realizaciju, ali se propast dogodi u diskretnom vremenu sa $h = 1/2$.

Dolje je navedeno pet relacija između diskretno vremenskih vjerojatnosti propasti za $0 \leq U_1 \leq U_2$ i za $0 \leq t_1 \leq t_2 < \infty$. Formule (1.5), (1.6), (1.7) i (1.8) su diskretno vremenske verzije formula (1.1), (1.2), (1.3) i (1.4), i njihova intuitivna objašnjenja su slična. Intuitivno objašnjenje od (1.9) proizlazi iz Slike 2.

$$\psi_h(U_2, t) \leq \psi_h(U_1, t) \tag{1.5}$$

$$\psi_h(U_2) \leq \psi_h(U_1) \tag{1.6}$$

$$\psi_h(U, t_1) \leq \psi_h(U, t_2) \leq \psi_h(U) \tag{1.7}$$

$$\lim_{t \to \infty} \psi_h(U, t) = \psi_h(U) \tag{1.8}$$

$$\psi_h(U, t) \leq \psi(U, t) \tag{1.9}$$

Intuitivno, za očekivati je da su sljedeće dvije relacije istinite, jer se vjerojatnost propasti u neprekidnom vremenu može aproksimirati s vjerojatnosti propasti u diskretnom vremenu, uz isti početni višak, i isti horizont t, ukoliko se dovoljno često provjerava propast, t.j., ukoliko je h dovoljno mali.

$$\lim_{h \to 0+} \psi_h(U, t) = \psi(U, t) \tag{1.10}$$

$$\lim_{h \to 0+} \psi_h(U) = \psi(U) \tag{1.11}$$

Formule (1.10) i (1.11) su istinite, ali su dokazi prilično zapetljani, te ih nećemo ovdje dati.

Za sada nismo načinili nikakve pretpostavke o distribuciji od $S(t)$. Uz neke pretpostavke mogu se dobiti detaljniji rezultati od vrlo općenith rezultata ovog odjeljka. Te pretpostavke uvodimo u sljedećem odjeljku.

2 Poissonovi i složeni Poissonovi procesi

2.1 Uvod

U ovom odjeljku uvest ćemo neke pretpostavke na proces broja šteta $\{N(t)\}_{t\geq 0}$, i na iznose šteta $\{X_i\}_{i=1}^{\infty}$. Pretpostavit ćemo da je proces broja šteta Poissonov proces, što vodi do složenog Poissonovog procesa $\{S(t)\}_{t\geq 0}$ za ukupne štete. Pretpostavke uvedene u ovom odjeljku vrijedit će i u ostatku ovog poglavlja.

2.2 Poissonov proces

Poissonov proces je primjer procesa brojenja. Ovdje nas zanima broj šteta proizašlih iz rizika. Budući da se štete broje kroz vrijeme, proces broja šteta mora zadovoljavati sljedeće uvjete:

(i) $N(0) = 0$, t.j., nema šteta u vremenu 0

(ii) za svaki $t > 0$, $N(t)$ mora biti cjelobrojan

(iii) za $s < t$, $N(s) < N(t)$, t.j., broj šteta kroz vrijeme je neopadajući

(iv) za $s < t$, $N(t) - N(s)$ predstavlja broj šteta koje se pojavljuju u intervalu (s, t).

Proces broja šteta $\{N(t)\}_{t\geq 0}$ je definiran kao Poissonov proces s parametrom λ ako su ispunjeni sljedeći uvjeti:

(i) $N(0) = 0$, i $N(s) \leq N(t)$ za $s < t$

(ii)

$$\begin{aligned}
&\mathbb{P}(N(t+h) = r \mid N(t) = r) = 1 - \lambda h + o(h) \\
&\mathbb{P}(N(t+h) = r+1 \mid N(t) = r) = \lambda h + o(h) \\
&\mathbb{P}(N(t+h) > r+1 \mid N(t) = r) = o(h)
\end{aligned} \tag{2.1}$$

(iii)

za $s < t$, broj šteta u intervalu $(s, t]$ nezavisan je od broja šteta do trenutka s (2.2)

Uvjet (ii) kaže da je u vrlo kratkom vremenskom intervalu duljine h jedini mogući broj šteta jednak nula ili jedan. Uočite da uvjet (ii) također povlači da broj šteta u vremenskom intervalu duljine h ne ovisi o tome kada taj interval počinje.
Razlog zašto se proces koji zadovoljava uvjete (i) do (iii) zove Poissonov proces je taj da za fiksnu vrijednost od t, slučajna varijabla $N(t)$ ima Poissonovu distribuciju s parametrom λt. To se dokazuje na sljedeći način:
Neka je $p_n(t) = \mathbb{P}(N(t) = n)$. Dokazat ćemo da je

$$p_n(t) = \exp(-\lambda t)\frac{(\lambda t)^n}{n!} \tag{2.3}$$

izvodeći i rješavajući "diferencijalno-diferencijsku" jednadžbu.
Za fiksnu vrijednost $t > 0$ i malu pozitivnu vrijednost h, uvjetujemo na broj šteta u trenutku t i pišemo

$$\begin{aligned} p_n(t+h) &= p_{n-1}(t)[\lambda h + o(h)] + p_n(t)[1 - \lambda h + o(h)] + o(h) \\ &= \lambda h p_{n-1}(t) + [1 - \lambda h]p_n(t) + o(h) \end{aligned}$$

Stoga je

$$p_n(t+h) - p_n(t) = \lambda h[p_{n-1}(t) - p_n(t)] + o(h) \tag{2.4}$$

i taj identitet vrijedi za $n = 1, 2, 3, \ldots$.
Podijelimo li sada (2.4) s h i pustimo li h prema nuli zdesna, dobivamo diferencijalno-diferencijsku jednadžbu

$$\frac{d}{dt}p_n(t) = \lambda[p_{n-1}(t) - p_n(t)]\,. \tag{2.5}$$

Za $n = 0$, identična analiza daje

$$\frac{d}{dt}p_0(t) = -\lambda p_0(t)\,. \tag{2.6}$$

Rješavamo za $p_n(t)$ uvodeći funkciju izvodnicu vjerojatnosti $G(s,t)$ definiranu sa

$$G(s,t) = \sum_{n=0}^{\infty} s^n p_n(t)$$

tako da je

$$\frac{d}{dt}G(s,t) = \sum_{n=0}^{\infty} s^n \frac{d}{dt}p_n(t)\,.$$

Pomnožimo sada (2.5) sa s^n i sumiramo preko svih vrijednosti n da bismo dobili

$$\sum_{n=1}^{\infty} s^n \frac{d}{dt} p_n(t) = \lambda \sum_{n=1}^{\infty} s^n p_{n-1}(t) - \lambda \sum_{n=1}^{\infty} s^n p_n(t) \, .$$

Dodamo li (2.6) gornjem identitetu dobivamo

$$\sum_{n=0}^{\infty} s^n \frac{d}{dt} p_n(t) = \lambda \sum_{n=1}^{\infty} s^n p_{n-1}(t) - \lambda \sum_{n=0}^{\infty} s^n p_n(t) \, .$$

što se može napisati kao

$$\frac{d}{dt} G(s,t) = \lambda s G(s,t) - \lambda G(s,t) \, ,$$

ili ekvivalentno

$$\frac{1}{G(s,t)} \frac{d}{dt} G(s,t) = \lambda(s-1) \, . \tag{2.7}$$

Budući daje lijeva strana od (2.7) jednaka derivaciji po t od $\log G(s,t)$, (2.7) možemo integrirati, te dobivamo

$$\log G(s,t) = \lambda t(s-1) + c(s)$$

gdje je $c(s)$ neka funkcija od s. $c(s)$ se može identificirati primjetimo li da je za $t = 0$, $p_0(t) = 1$ i $p_n(t) = 0$, $n = 1, 2, 3, \ldots$. Stoga je $G(s, 0) = 1$ i $\log G(s, 0) = 0 = c(s)$. Zato je

$$G(s,t) = \exp\{\lambda t(s-1)\}$$

što je funkcija izvodnica vjerojatnosti Poissonove distribucije s parametrom λt. Budući da postoji jedan-jedan veza između funkcija izvodnica vjerojatnosti i funkcija distribucije, slijedi da je distribucija od $N(t)$ Poissonova s parametrom λt.

Ovo proučavanje Poissonovog procesa završit ćemo razmatranjem distribucije vremena do prve štete i vremena između šteta.

Neka slučajna varijabla T_1 označava vrijeme prve štete. Za fiksnu vrijednost t, ako se niti jedna šteta nije pojavila do trenutka t imamo $T_1 > t$. Slijedi

$$\mathbb{P}(T_1 > t) = \mathbb{P}(N(t) = 0) = \exp\{-\lambda t\}$$

i

$$\mathbb{P}(T_1 \leq t) = 1 - \exp\{-\lambda t\}\,,$$

tako da T_1 ima eksponencijalnu distribuciju s parametrom λ.
Za $i = 2, 3, \ldots$, neka slučajna varijabla T_i označava vrijeme između $(i-1)$-ve i i-te štete. Tada je

$$\begin{aligned}
\mathbb{P}(T_{n*1} > t \,|\, \sum_{i=1}^{n} T_i = r) &= \mathbb{P}(\sum_{i=1}^{n+1} T_i > t + r \,|\, \sum_{i=1}^{n} T_i = r) \\
&= \mathbb{P}(N(t+r) = n \,|\, N(r) = n) \\
&= \mathbb{P}(N(t+r) - N(r) = 0 \,|\, N(r) = n)\,.
\end{aligned}$$

Po uvjetu (2.2),

$$\mathbb{P}(N(t+r) - N(r) = 0 \,|\, N(r) = n) = \mathbb{P}(N(t+r) - N(r) = 0)\,.$$

Konačno,

$$\mathbb{P}(N(t+r) - N(r) = 0) = \mathbb{P}(N(t) = 0) = \exp\{-\lambda t\}\,,$$

budući da broj šteta u vremenskom intervalu duljine r ne ovisi o tome kada taj interval počinje (uvjet (2.1)). Prema tome i vremena između događaja imaju eksponencijalnu distribuciju s parametrom λ.

2.3 Složeni Poissonov proces

U ovom odjeljku kombinirat ćemo Poissonov proces broja šteta s distribucijom iznosa šteta, te ćemo dobiti složeni Poissonov proces za proces ukupnih šteta definiran u odjeljku 1.1.
Načinit ćemo sljedeće tri važne pretpostavke:

- slučajne varijable $\{X_i\}_{i=1}^{\infty}$ su nezavisne i jednako distribuirane
- slučajne varijable $\{X_i\}_{i=1}^{\infty}$ su nezavisne od $N(t)$ za sve $t \geq 0$
- slučajni proces $\{N(t)\}_{t\geq 0}$ je Poissonov proces čiji parametar označavamo s λ.

U odjeljku (2.2) pokazano je da zadnja pretpostavka znači da za svaki $t \geq 0$, slučajna varijabla $N(t)$ ima Poissonovu distribuciju s parametrom λt, tako da

$$\mathbb{P}[N(t) = k] = \exp\{-\lambda t\} \frac{(\lambda t)^k}{k!} \quad \text{za } k = 0, 1, 2, \ldots .$$

Uz te pretpostavke, proces ukupnih šteta $\{S(t)\}_{t\geq 0}$ naziva se složen Poissonov proces s Poissonovim parametrom λ. Uspoređujući gornje pretpostavke s pretpostavkama u Poglavlju 4, odjeljak 1.3, i u Poglavlju 4, odjeljak 2.2, može se vidjeti da ako je $\{S(t)\}_{t\geq 0}$ složen Poissonov proces s Poissonovim parametrom λ, tada, za fiksnu vrijednost od $t \geq 0$, $S(t)$ ima složenu Poissonovu distribuciju s Poissonovim parametrom λt. (Uočite malu promjenu terminologije ovdje: "Poissonov parametar λ"postaje "Poissonov parametar λt" kada napravimo promjenu s procesa na distribuciju.)
Zajednička funkcija distribucije od X_i označavat će se s $F(x)$, i pretpostavit ćemo do kraja ovog poglavlja da je $F(0) = 0$, tako da su sve štete pozitivnog iznosa.
Vjerojatnosna funkcija gustoće od X_i, ako postoji, označavat će se s $f(x)$, a k-ti moment oko nule od X_i, ako postoji, označavat će se s m_k, tako da je

$$m_k = \mathbb{E}[X_i^k] \quad \text{za } k = 1, 2, 3, \ldots .$$

Kadgod postoji zajednička funkcija izvodnica momenata od X_i, njezina vrijednost u točki r označavat će se s $M_X(r)$.
Budući da za fiksnu vrijednost od t, $S(t)$ ima složenu Poissonovu distribuciju, iz Poglavlja 4, odjeljak 2.2, slijedi da proces $\{S(t)\}_{t\geq 0}$ ima očekivanje $\lambda t m_1$, varijancu $\lambda t m_2$ i funkciju izvodnicu momenata $M_S(r)$ gdje je

$$M_S(r) = \exp\{\lambda t(M_X(r) - 1)\} .$$

U nastavku ovog poglavlja napravit ćemo sljedeću (intuitivno razumljivu) pretpostavku o stopi premije:

$$c > \lambda m_1 \tag{2.8}$$

tako da je osigurateljeva premija (po jedinici vremena) veća od očekivane štete (po jedinici vremena). Ponekad ćemo c napisati kao

$$c = (1 + \theta)\lambda m_1$$

gdje je $\theta > 0$ dodatak na premiju.

2.4 Tehnikalija

U sljedećem odjeljku trebat će nam tehnički rezultat o $M_X(r)$ (funkcija izvodnica momenata distribucije pojedinačnih iznosa šteta), koji zbog pogodnosti pokazujemo ovdje.
U nastavku ovog odjeljka pretpostavit ćemo da postoji broj γ $(0 < \gamma \leq \infty)$ takav da je $M_X(r)$ konačno za sve $r < \gamma$ i

$$\lim_{r \to \gamma-} M_X(r) = \infty\,. \tag{2.9}$$

(Na primjer, ako X_i ima sliku ograničenu nekim konačnim brojem, tada će γ biti ∞; ako X_i ima eksponencijalnu distribuciju s parametrom α, tada će γ biti jednak α.)
U sljedećem odjeljku biti će potreban ovaj rezultat:

$$\lim_{r \to \gamma-} (\lambda M_X(r) - cr) = \infty\,. \tag{2.10}$$

Za konačan γ to slijedi odmah iz (2.9). Sada ćemo pokazati da (2.10) vrijedi kada je γ beskonačan. To zahtijeva malo više pažnje. Prvo uočimo da postoji pozitivan broj, recimo ϵ, takav da je

$$\mathbb{P}[X_i > \epsilon] > 0\,.$$

To je zato što su svi iznosi šteta pozitivni. Označimo tu vjerojatnost s π. Tada je

$$M_X(r) \geq e^{r\epsilon}\pi\,.$$

Stoga

$$\lim_{r \to \infty} (\lambda M_x(r) - cr) \geq \lim_{r \to \infty} (\lambda e^{r\epsilon}\pi - cr) = \infty\,.$$

3 Koeficijent prilagodbe i Lundbergova nejednakost

3.1 Lundbergova nejednakost

Lundbergova nejednakost tvrdi da je

$$\psi(U) \leq \exp\{-RU\}$$

gdje je U osigurateljev početni višak. R je parametar povezan s procesom viška i poznat je kao koeficijent prilagodbe. Njegova vrijednost ovisi o distribuciji ukupnih šteta i o stopi premije. Prije nego što definiramo R ilustrirat ćemo važnost rezultata i neka svojstva koeficijenta prilagodbe.
Slika 3. pokazuje grafove od $\exp\{-RU\}$ i $\psi(U)$ po U kada su iznosi šteta eksponencijalno distribuirani s očekivanjem 1, te kada je dodatak na premiju 10%. (Rješenje za R naći ćemo u odjeljku 3.2. Formula za $\psi(U)$ dana je u odjeljku 4.) Može se vidjeti da je za velike vrijednosti od U, $\psi(U)$ vrlo blizu gornjoj granici, tako da je $\psi(U) \simeq \exp\{-RU\}$.
U aktuarskoj literaturi, $\exp\{-RU\}$ se često koristi kao aproksimacija za $\psi(U)$.
R se može interpretirati kao mjera rizika. Čim je veća vrijednost od R, manja će biti gornja granica za $\psi(U)$. Stoga se očekuje da $\psi(U)$ opada kako R raste. R je funkcija parametara koji utječu na vjerojatnost propasti, i može se promatrati ponašanje od R kao funkcija tih parametara.
Slika 4. pokazuje graf od R kao funkciju dodatka na premiju θ, kada

(i) je distribucija iznosa šteta eksponencijalna s očekivanjem 10, i

(ii) su sve štete iznosa 10.

Uočite da je u oba slučaja R rastuća funkcija od θ. To nije iznenađujuće, jer očekujemo da je $\psi(U)$ opadajuća funkcija od θ, te budući da je $\psi(U) \simeq \exp\{-RU\}$, svaki faktor koji prouzrokuje pad u $\psi(U)$ će uzrokovati rast u R.
Uočite također da je vrijednost od R u slučaju eksponencijalno distribuiranih šteta manja od vrijednosti od R kada sve štete iznose 10. I opet, taj rezultat nije iznenađujući. Obje distribucije iznosa šteta imaju očekivanje 10, ali eksponencijalna distribucija ima veću varijabilnost. Veća varijabilnost se povezuje s većim rizikom, te se stoga očekuje veća vrijednost od $\psi(U)$ za eksponencijalnu distribuciju, te manja vrijednost od R. Ovaj primjer pokazuje da na R utječu dodatak na premiju i karakteristike distribucije pojedinačnih iznosa šteta. Sada ćemo definirati R i pokazati da općenito uključuje sve faktore koji utječu na proces viška.

3.2 Koeficijent prilagodbe - složeni Poissonovi procesi

Proces viška ovisi o početnom višku, o procesu ukupnih šteta i o stopi premija. Koeficijent prilagodbe je parametar pridružen procesu viška koji uzima u obzir dva od tih faktora: ukupne štete i prihod od premija. Koeficijent prilagodbe je mjera rizika za proces viška. Kada su ukupne štete složen Poissonov proces, koeficijent prilagodbe definira se pomoć Poissonovog parametra, funkcije izvodnice momenata iznosa pojedinačnih šteta i prihoda od premija po jedinici vremena.

Koeficijent prilagodbe, označen s R, definiramo kao jedinstveno pozitivno rješenje od

$$\lambda M_X(r) - \lambda - cr = 0\,. \tag{3.1}$$

Dakle, R je dan s

$$\lambda M_X(R) = \lambda + cR\,. \tag{3.2}$$

Uočite da jednažba (3.1) povkači da vrijednost koeficijenta prilagodbe ovisi o Poissonovom parametru, distribuciji pojedinačnih iznosa šteta i o stopi premije. Međutim, napišemo li $c = (1+\theta)\lambda m_1$, dobivamo

$$M_X(r) = 1 + (1+\theta)m_1 r$$

tako da je R nezavisan od Poissonovog parametra i jednoastavno ovisi o dodatku na premiju θ i o distribuciji pojedinačnih iznosa šteta.

Na sljedeći način se pokazaje da zaista postoji jedno pozitivno rješenje od (3.1).

Definiramo $g(r) = \lambda M_X(r) - \lambda - cr$ i promatramo graf od $g(r)$ na intervalu $[0, \gamma]$. Uočimo prvo da je $g(0) = 0$. Nadalje, $g(r)$ je opadajuća funkcija u $r = 0$, budući da je

$$\frac{d}{dr}\, g(r) = \lambda \frac{d}{dr}\, M_X(r) - c$$

tako da je derivacija od $g(r)$ u $r = 0$ jednaka $\lambda m_1 - c$ što je strogo manje od nule po pretpostavci (2.8).

Također se može pokazati da ako funkcija $g(r)$ ima lokalni ekstrem, to mora biti minimum funkcije. Druga derivacija je

$$\frac{d^2}{dr^2}\, g(r) 0 \lambda \frac{d^2}{dr^2}\, M_X(r)$$

što je uvijek strogo pozitivno. Dakle, može postoajti samo jedan lokalni ekstrem, budući da je svaki lokalni ekstrem minimum. Da bismo pokazali da postoji lokalni ekstrem uočimo iz (2.10) da je $\lim_{r \to \gamma-} g(r) = \infty$. Budući da je g opadajuća funkcija u $r = 0$, mora imati minimum, te je graf od g kao na Slici 5.

Dakle postoji jedinstven pozitivan broj R koji zadovoljava jednadžbu (3.1). Jednadžba (3.1) je implicitna jednadžba za R. Za neke oblike od $F(x)$ moguće ju je riješiti eksplicitno po R; inače jednadžbu moramo riješiti numerički.
Promotrimo eksponencijalnu distribuciju $F(x) = 1 - e^{-\alpha x}$.
Za tu distribuicju, $M_X(r) = \alpha/(\alpha - r)$, tako da

$$\begin{aligned} & \lambda + cR = \frac{\lambda}{\alpha - R} \qquad (3.3) \\ \Rightarrow\ & \lambda\alpha - \lambda R + cR\alpha - cR^2 = \lambda\alpha \\ \Rightarrow\ & R^2 - (\alpha - \frac{\lambda}{c})R = 0 \\ \Rightarrow\ & R = \alpha - \frac{\lambda}{c} \end{aligned}$$

budući da je R pozitivno rješenje od (3.1).
Ako je $c = (1+\theta)\lambda/\alpha$, tada je $R = \alpha\theta/(1+\theta)$.
Ako se jednadžba za R treba riješiti numerički, korisno je imati grubu ideju o vrijednosti od R. Jednadžba (3.2) može se upotrijebiti za nalaženje jednostavne gornje granice za R kako slijedi:

$$\begin{aligned} \lambda + cR &= \lambda M_X(R) \\ &= \lambda \int_0^\infty e^{Rx} f(x)\,dx \\ &> \lambda \int_0^\infty (1 + Rx + \frac{1}{2}R^2x^2) f(x)\,dx \\ &= \lambda(1 + Rm_1 + \frac{1}{2}R^2 m_2) \end{aligned}$$

tako da je $(c - \lambda m_1)R > R^2 m_2/2$, što daje

$$R < \frac{2(c - \lambda m_1)}{\lambda m_1} \qquad (3.4)$$

tako da je $R < 2\theta m_1/m_2$ uz $c = (1+\theta)\lambda m_1$. Uočite da ako je vrijednost od R mala, tada bi trebala biti vrlo blizu ovoj gornjoj međi, budući da bi aproksimacija za e^{Rx} trebala biti dobra.
Donja granica za R može se izvesti kada postoji gornja granica, recimo M, na iznos pojedinačnih šteta. Na primjer, ako su pojedinačne štete uniformno

distribuirane na $(0, 100)$, tada je $M = 100$. Taj rezultat se dokazuje na sličan način kao i rezultat (3.4). Donja granica se nalazi primjenom nejednakosti

$$\exp\{Rx\}\} \leq \frac{x}{M}\exp\{RM\} + 1 - \frac{x}{M} \quad \text{za } 0 \leq x \leq M\,. \tag{3.5}$$

Tu nejednakost dokazujemo razvijajući $\exp\{RM\}$ u red potencija.

$$\begin{aligned}
\frac{x}{M}\exp\{RM\} + 1 - \frac{x}{M} &= \frac{x}{M}\sum_{j=0}^{\infty}\frac{(RM)^j}{j!} + 1 - \frac{x}{M} \\
&= 1 + \sum_{j=1}^{\infty}\frac{R^j M^{j-1} x}{j!} \\
&\geq 1 + \sum_{j=1}^{\infty}\frac{(Rx)^j}{j!} \qquad \text{za } 0 \leq x \leq M \\
&= \exp\{Rx\}
\end{aligned}$$

Nejednakost (3.5) može se iskoristiti za izvod

$$R > \frac{1}{M}\log(\frac{c}{\lambda m_1})$$

kada pojedinačne štete imaju neprekidnu distribuciju na $(0, M)$.
Počinjemo od jednadžbe koja definira R.

$$\begin{aligned}
\lambda + cR &= \lambda\int_0^M \exp\{Rx\} f(x)\,dx \\
&\leq \lambda\int_0^M \left(\frac{x}{M}\exp\{RM\} + 1 - \frac{x}{M}\right) f(x)\,dx \\
&= \frac{\lambda}{M}\exp\{RM\}m_1 + \lambda - \frac{\lambda}{M}m_1
\end{aligned}$$

Preuredivši,

$$\begin{aligned}
\frac{c}{\lambda m_1} &\leq \frac{1}{RM}(\exp\{RM\} - 1) = 1 + \frac{RM}{2} + \frac{(RM)^2}{31} + \cdots \\
&< \exp\{RM\}
\end{aligned}$$

što daje $R > \dfrac{1}{M}\log(\dfrac{c}{\lambda m_1})$, kao što se i tražilo.
Režući razvoj u red od $\exp\{Rx\}$, mogu se naći i druge aproksimacije za R, specijalno kada je R mali.

3.3 Koeficijent prilagodbe - opći procesi ukupnih šteta

U odjeljku 3.2 dokazali smo postojanje koeficijenta prilagodbe za složeni Poissonov proces ukupnih šteta. U ovom odjeljku dokazujemo egzistenciju koeficijenta prilagodbe za opći proces ukupnih šteta.
Neka je $\{S_i\}_{i=1}^{\infty}$ niz nezavisnih jednako distribuiranih slučajnih varijabli.

$S_i \equiv$ ukupne štete iz rizika u periodu i

c konstantna premija koja se naplaćuje za osiguranje tog rizika

Uvodimo sljedeće pretpostavke:

$$c > \mathbb{E}[S_i] \tag{3.6}$$

Postoji broj $\gamma > 0$ tako da je

$$\lim_{r \to \gamma-} \mathbb{E}[e^{r(S_i - c)}] = \infty \tag{3.7}$$

$$S_i \text{ ima gustoću } h(x),\ -\infty < x < \infty\,. \tag{3.8}$$

Dokaz da postoji jedan i samo jedan pozitivan broj R takav da je

$$\mathbb{E}[e^{r(S_i - c)}] = 1$$

je kako slijedi.
Neka je $f(r) = \mathbb{E}[e^{r(S_i - c)}]$ za $-\infty < r < \gamma$. Tada

$$\begin{aligned} f(0) &= 1 \\ f'(0) &= \mathbb{E}[S_i - c] < 0 \\ f''(x) &> 0 \\ \lim_{r \to \gamma-} f(r) &= +\infty \end{aligned}$$

Napomene: Pretpostavimo da S_i ima Poissonovu distribuciju s Poissonovim parametrom λ i pojedinačnim iznosom štete danim slučajnom varijablom X. Tada

$$\begin{aligned} & \mathbb{E}[e^{r(S_i - c)}] = 1 \\ \Rightarrow\ & \mathbb{E}[e^{RS_i}] = e^{Rc} \\ \Rightarrow\ & e^{\lambda(M_X(R) - 1)} = e^{Rc} \\ \Rightarrow\ & \lambda M_X(R) = Rc + \lambda \end{aligned}$$

što je isto kao i 3.2.

4 Posljedica promjene vrijednosti parametara na vjerojatnsot propasti u konačnon i beskonačnom vremenu

4.1 Uvod

U ovom odjeljku diskutirat ćemo posljedice promjene parametara na vrijednosti od $\psi(U,t)$ i $\psi(U)$. Nećemo uvesti nikakvu novu teoriju i metode za dobivanje numeričkih vrijednosti za $\psi(U,t)$ neće se diskutirati. Svojstva od $\psi(U,t)$, i u nekim slučajevima od $\psi(U)$, ilustrirat će se nizom numeričih primjera. U tim primjerima načinit ćemo osnovne pretpostavke kao u prethodnom odjeljku. Specijalno, pretpostavit ćemo da je proces ukupnih šteta složen Poissonov proces. Dodatno ćemo pretpostavljati kroz odjeljke 4.3, 4.4 i 4.5 da vrijedi

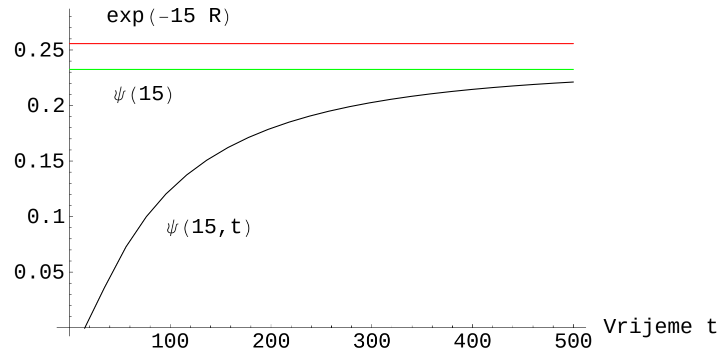
– Poissonov parametar za broj šteta jednak je 1 (4.1)
– očekivana vrijednost pojedinačne štete jednaka je 1 (4.2)
– pojedinačne štete imaju eksponencijalnu distribuciju (4.3)

U odjeljku 4.6 pretpostavit ćemo (4.2) i (4.3), ali ćemo dopustiti da se mijenja Poissonov parametar.

Posljedica prepostavke (4.1) je da je jedinica vremena izabrana tako da je očekivani broj šteta u jedinici vremena jednak 1. Zato je $\psi(U, 500)$ vjerojatnost propasti (uz početni višak U) kroz vremenski period u kojem očekujemo 500 šteta. Stvarni broj šteta u tom periodu ima Poissonovu distribuciju (s parametrom 500) i može poprimiti bilo koju nenegativnu cjelobrojnu vrijednost.

Posljedica preptostavke (4.2) je da je monetarna jedinica izabran tako da bude jednaka očekivanom iznosu jedne štete. Zato je $\psi(20, 500)$ vjerojatnost propasti (kroz vremenski period u kojem očekujemo 500 šteta) uz dan početni višak jednak 20 puta očekivani iznos jedne štete.

Prednost upotrebe eksponencijalne distribucije za pojedinačne štete (pretpostavka (4.3)) je da se i $\exp\{-RU\}$ i $\psi(U)$ mogu izračunati za te primjere. Vidi odjeljak 3 i odjeljak 4.2.



Slika 6:

4.2 Formula za $\psi(U)$ kada je $F(x)$ eksponencijalna distribucija

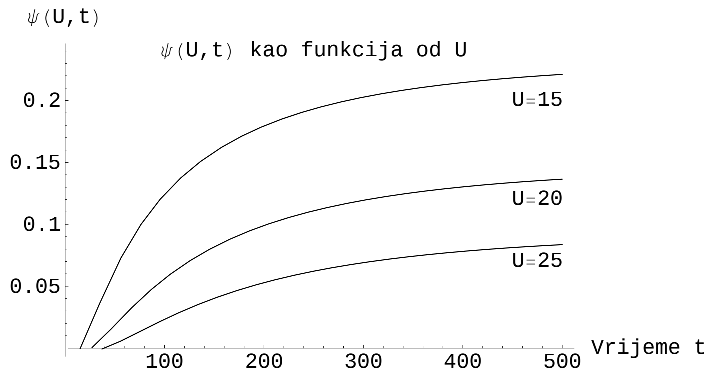
Formula zs $\psi(U)$ kada su pojedinačni iznosi šteta eksponencijalno distribuirani s očekivanjem 1, i kada je dodatak na premiju jednak θ, dana je u sljedećem rezultatu.
Kada je $F(x) = 1 - \exp\{-x\}$,

$$\psi(U) = \frac{1}{1+\theta} \exp\left\{-\frac{\theta U}{1+\theta}\right\} . \tag{4.4}$$

Program ne zahtijeva da se taj rezultat izvede ili zapamti.
Taj rezultat smo naveli za ilustraciju kakve posljedice, za tu konkretnu distribuciju, ima promjena parametara na vjerojatnost propasti.

4.3 $\psi(U, t)$ kao funkcija od t

Slika 6. pokazuje graf od $\psi(15, t)$ za $0 \leq t \leq 500$. Dodatak na premiju θ jednak je 0.1, tako da je prihod od premija po jedinici vremena 1.1. Također su na Slici 6. pokazani $\psi(15)$ (zelena linija) i $\exp\{-15R\}$ (crvena linija) za taj portfelj. Te zadnje dvije vrijednosti prikazane su kao paralelni pravci vremenskoj osi, budući da su njihove vrijednosti nezavisne od vremena.
Na Slici 6. trebamo uočiti sljedeće osobine:



Slika 7:

(i) $\psi(15, t)$ je rastuća funkcija od t (vidi formulu (1.3)),

(ii) za male vrijednosti od t, $\psi(15, t)$ raste vrlo brzo (njena vrijednost se udvostručuje kada t raste od 25 do 50, te se opet udvostručuje kada t raste od 50 do 100),

(iii) za velike vrijednosti od t, $\psi(15, t)$ raste polaganije i približava se asimptotski vrijednosti od $\psi(15)$ (vidi formulu (1.4)).

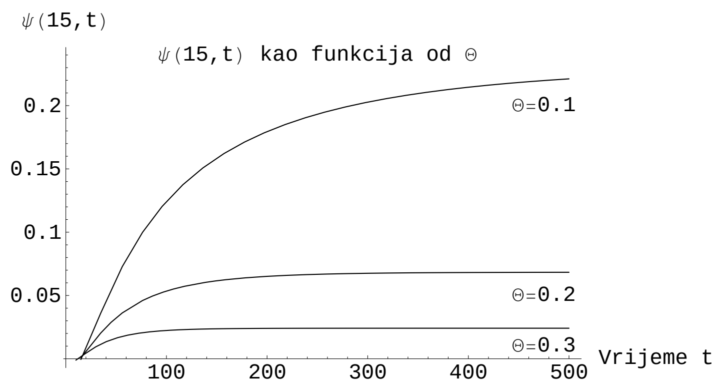
4.4 Vjerojatnost propasti kao funkcija početnog viška

Slika 7. pokazuje vrijednosti od $\psi(U, t)$ za $0 \le t \le 500$ za tri vrijednosti početnog viška $U = 15, 20$ i 25. Dodatak na premiju je 0.1 kao i na Slici 6. Za $U = 15$ graf od $\psi(U, t)$ je kao na Slici 6.
Na Slici 7. trebamo uočiti sljedeće osobine:

(i) svi grafovi imaju isti opći oblik,

(ii) povećanje vrijednosti od U smanjuje vrijednost od $\psi(U, t)$ za svaku vrijednost od t (vidi formulu (1.1)),

(iii) svaki od tri grafa približava se asimptotskoj granici kada t raste (kao što smo već primjetili za U jednako 15 u diskusiji Slike 6). Uočite da je $\psi(20) = 0.1476$ i $\psi(25) = 0.0937$.



Slika 8:

Formula (1.2) pokazuje da je $\psi(U)$ neopadajuća funkcija od U.
U slučaju eksponencijalno distribuiranih iznosa pojedinačnih šteta, derivacija po U od $\psi(U)$ jednaka je

$$\frac{d}{dU}\,\psi(U) = -\,\frac{\theta}{1+\theta}\,\psi(U)$$

što je negativno zbog $\theta > 0$. Zato je $\psi(U)$ opadajuća funkcija od U. To je preciznija tvrdnja od opće trvrdnje u formuli (1.2).
Intuitivno je jasno da bi $\psi(U,t)$ (te kao specijalna slučaj $\psi(U)$) trebala biti opadajuća funkcija od U. Povećanje od U predstavlja povećanje osigurateljevog viška bez odgovarajućeg povećanja u iznosu šteta. Dakle, povećanje od U predstavlja povećanje osigurateljeve sigurnosti, te će stoga smanjiti vjerojatnost propasti.

4.5 Vjerojatnost propasti kao funkcija dodatka na premiju

Slika 8. pokazuje vrijednosti od $\psi(15,t)$ za $0 \leq t \leq 500$ za tri vrijednosti dodatka na premiju, $\theta = 0.1, 0.2$ i 0.3. Graf od $\psi(15,t)$ za $\theta = 0.1$ jednak je grafu na Slici 6. i jednak jednom od grafova na Slici 7. Slika 8. je u mnogim aspektima slična Slici 7.
Na Slici 8. trebamo uočiti sljedeće osobine:

(i) grafovi od $\psi(15, t)$ svi imaju isti opći oblik,

(ii) povećanje vrijednosti od θ smanjuje vrijednost od $\psi(15, t)$ za svaku danu vrijednost od t; to je u stvari istinito za sve vrijednosti od U, i očigledan je rezultat, budući daje povećanje od θ ekvivalentno povećanju stope premije bez promjene procesa ukupnih šteta,

(iii) vidi se da je za $\theta = 0.2$ i 0.3, $\psi(15, t)$ manje-više konstantno za t veće od otprilike 150. Za $t_1 \leq t_2$, razlika $\psi(15, t_2) - \psi(15, t_1)$ predstavlja vjerojatnost da će se propast dogoditi između vremena t_1 i t_2. Zato za vrijednosti od θ jednake 0.2 i 0.3 (te za vrijednost početnog viška 15 i za dani proces ukupnih šteta), ako se propast uopće dogodi, puno je veća šansa da će se dogoditi prije vremena 150. To ćemo nadalje diskutirati u odjeljku 4.6.

Pomoću općenitog razmatranja jasno je da $\psi(U)$ mora biti neopadajuća funkcija od θ. U slučaju eksponencijalno distribuiranih šteta, $\psi(U)$ je opadajuća funkcija od θ.

$$\frac{d}{dU}\,\psi(U) = -(1+\theta)^{-1}\psi(U) - U(1+\theta)^{-2}\psi(U)$$

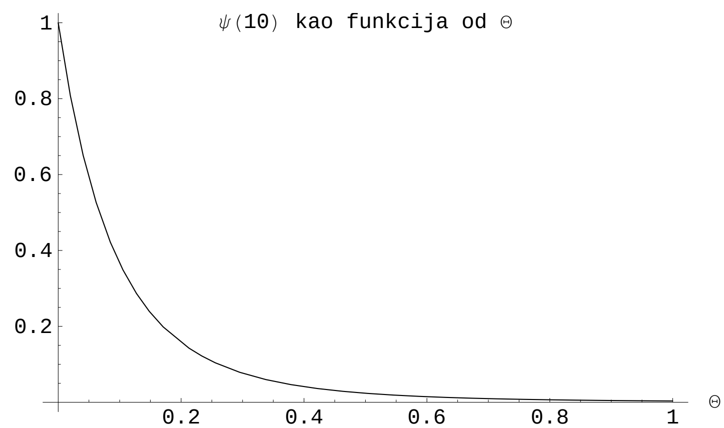
što je očito negativno, budući da su θ, U i $\psi(U)$ sve pozitivne veličine. Kako je derivacija manja od nule za sve vrijednosti od θ, $\psi(U)$ je opadajuća funkcija od θ.
Slika 9. pokazuje $\psi(10)$ kao funkciju od θ.

4.6 Vjerojatnost propasti kao funkcija Poissonovog parametra

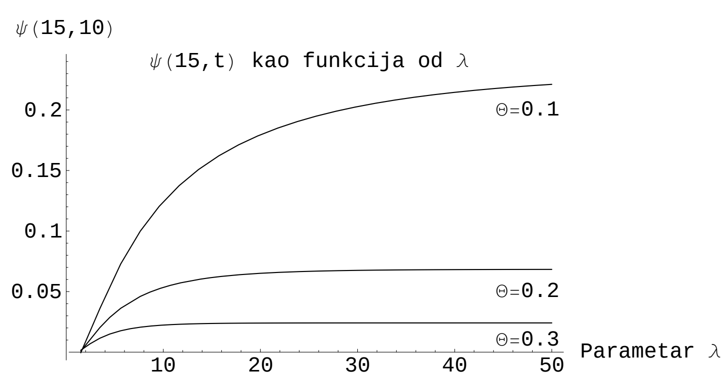
Slika 10. pokazuje $\psi(15, 10)$ kao funkciju od λ za tri vrijednosti dodatka na premiju $\theta = 0.1, 0.2$ i 0.3. Taj graf jednak je Slici 6. do na oznake na x-osi. To možemo objasniti promatrajući sljedeća dva rizika.

Rizik 1: ukupne štete su složen Poissonov proces s Poissonovim parametrom 1 i $F(x) = 1 - e^{-x}$. Dohodak od premije po jedinici vremena za pokriće tog rizika je $(1+\theta)$.

Rizik 2: ukupne štete su složen Poissonov proces s Poissonovim parametrom 0.5 i $F(x) = 1 - e^{-x}$. Dohodak od premije po jedinici vremena za pokriće tog rizika je $0.5(1+\theta)$.



Slika 9:



Slika 10:

Za jedinicu vremena uzimamo jednu godinu. Jedina razlika između tih rizika je da kod prvog rizika očekujemo dva put više šteta svake godine. To se reflektira u dvije premije.
Promotrimo Rizik 2 u novoj vremenskoj jedinici ekvivalentnoj dvije godine. Tada su distribucija ukupnih šteta i prihod od premija po jedinici vremena jednaki odgovarajućim veličinama za Rizik 1. Zato je vjerojatnost propasti u beskonačnom vremenu jednaka za oba rizika.
Puna linija na Slici 11. pokazuje jedan ishod procesa viška za Rizik 1 uz $\theta = 0.1$. Iscrtkana linija pokazuje isti proces viška kada je jedinica vremena jednaka dvije godine. To pokazuje da će svaki ishod procesa viška koji uzrokuje propast za Rizik 1, također uzrokovati propast za Rizik 2. Jedino je vrijeme (u godinama) do propasti različito. Mjereći vrijeme u godinama, vjerojatnost propasti do vremena 1 za Rizik 1 jednaka je vjerojatnosti propasti do vremena 2 za Rizik 2. To objašnjava zašto Slike 6. i 10. pokazuju iste funkcije. Na primjer, vrijednost od $\psi(15, 10)$ za $\lambda = 50$ (Slika 10.) jednaka je vrijednosti od $\psi(15, 500)$ za $\lambda = 1$ (Slika 6.).
Sada ćemo ispitati točku (iii) iz odjeljka 4.5, gdje smo uočili da su vrijednosti od $\psi(15, t)$ manje-više konstantne za vrijednosti od t veće od 150 kada je $\theta = 0.2$ i 0.3. Specijalno, promatrat ćemo situaciju kada je dodatak na premiju 0.2.
Promotrimo drugi proces ukupnih šteta koji je jednak procesu kojeg smo do sada razmatrali osim što je Poissonov parametar 150, a ne 1. (Taj drugi proces je u stvari identičan originalnom; sve što se dogodilo je da je promijenjena jedinica vremena.) Upotrijebimo ψ^* kao oznaku vjerojatnosti propasti tog drugog procesa, te neka ψ označava, kao i do sada, vjerojatnosti propasti originalnog procesa. Promjena vremenske jedinice znači da je za svaki $t \geq 0$

$$\psi^*(U, t) = \psi(U, 150t)\,,$$

ali nema nikakvog utjecaja na vjerojatnost propasti u beskonačnom vremenu (stavi $t = \infty$ u gornju relaciju), tako da je

$$\psi^*(U) = \psi(U)\,.$$

U (iii) gore smo uočili da je

$$\psi(15, 150) \simeq \psi(15)\,.$$

Odavde i iz dvije prethodne relacije vidi se da je

$$\psi^*(15, 1) \simeq \psi^*(15)\,.$$

Riječima, ta relacija kaže da je za drugi proces, uz početni višak 15, vjerojatnost propasti unutar jednog vremenskog perioda gotovo jednaka (u stari malo manja od) vjeroajatnosti propasti u beskonačnom vremenu. Taj zaključak bitno ovisi o činjenjici da je $\psi^*(15,t)$ neprekidno vremenska vjerojatnost propasti. Da bismo to vidjeli, promotrimo $\psi(15,1)$, što je upravo vjerojatnost da je za drugi proces višak na kraju jednog vremenskog perioda negativan. $\psi(15,1)$ može se približno izračunati ako pretpostavimo da ukupne štete u jednom vremenskom periodu, što ćemo označiti s $S^*(1)$, imaju normalnu distribuciju. Prisjetimo se da pojedinačne štete imaju eksponencijalnu distribuciju s očekivanjem 1, i da broj šteta u jednom vremenskom periodu ima Poissonovu distribuciju s očekivanjem 150. Odavde

$$\mathbb{E}[S^*(1)] = 150 \quad \text{i} \quad \mathrm{Var}[S^*(1)] = 300\,.$$

Koristeći tablice normalne distribucije

$$\begin{aligned} \psi(15,1) &= \mathbb{P}[S^*(1) > 15 + 1.2 \times 150] \\ &= \mathbb{P}[(S^*(1) - 150)/17.32 > 45/17.32] \\ &\simeq 0.005 \end{aligned}$$

Iz SLike 8. može se vidjeti je vrijednost od $\psi(15,150)$; te dakle od $\psi^*(15,1)$, otprilike 0.07, što je jako različito od (približne) vrijednosti diskretno vremenske vjerojatnosti propasti $\psi(15,1)$ izračunate gore.

4.7 Završne napomene

Kada su pojedinačni iznosi šteta eksponencijalni distribuirani, uočimo da ako je $\theta = 0$, tada je $\psi(U) = 1$ bez obzira na vrijednost od U. Taj rezultat je u stvari istinit bez obzira na oblik od $F(x)$. (Trivijalno slijedi da ako je $\theta < 0$, tada je $\psi(U) = 1$.) Drugim riječima, da vjerojatnost propasti u beskonačnom vremenu ne bude sigurna, bitno je da dodatak na premiju bude pozitivan.
Također primjetite da je kroz odjeljak bilo pretpostwvljeno da su pojedinačni iznosi šteta eksponencijalno distribuirani s očekivanjem 1. To očekivanje može se mjeriti u jedinicama od £100, £1,000 ili možda £1,000,000. Bez smanjenja općenitosti možemo staviti da parametar eksponencijalne distribucije bude 1 ukoliko korektno specificiramo monetarnu jedinicu. Jednostavno rečeno, vjerojatnost propasti kada je U 1£ jednaka je vjerojatnosti propast kada je U 100 penija. Možemo reći da je

$$\psi(U) \quad \text{kada je} \quad F(x) = 1 - e^{-\alpha x}$$

jednako kao i

$$\psi(\alpha U) \;\text{ kada je }\; F(x) = 1 - e^{-x}\,.$$

Drugim riječima, ako se očekivane štete po jedinici vremena povećaju za (faktor) α, to se i početni višak mora tako povećati želimo li da vjerojatnost propasti u beskonačnom vremenu ostane nepromijenjena.

5 Reosiguranje i propast

5.1 Uvod

Reosiguranje je jedna od opcija otvorenih osiguratelju ako želi smanjiti varijabilnost ukupnih šteta iz rizika. Očekuje se da smanjenje varijabilnosti poveća osigurateljevu sigurnost, te tako smanji i vjerojatnsot propasti. Budući da je teško naći eksplicitna rješenja za vjerojatnost propasti, promatrat ćemo utjecaj reosiguranja na koeficijent prilagodbe. Ako možemo naći reosiguranje koji maksimizira vrijednost koeficijenta prilagodbe, minimizirat će se gornja granica za vjerojatnost propasti u beskonačnom vremenu. Budući da je koeficijent prilagodbe mjera rizika, čini se razumnim ciljem maksimizirati njegovu vrijednost. Nadalje ćemo razmatrati utjecaj koeficijenta prilagodbe na kvotno reosiguranje i reosiguranje viška štete. Ti ugovori o reosiguranju definirani su u Poglavkju 4, odjeljak 2.

5.2 Maksimiziranje koeficijenta prilagodbe uz kvotno reosiguranje

Promotrimo prvo utjecaj kvotnog reosiguranja sa samopridržajem α na osigurateljev koeficijent prilagodbe. Kroz odjeljak 5.2, označavat ćemo osigurateljev prihod od premija po jedinici vremena, prije plaćanja reosigurateljne premije, sa $(1+\theta)\lambda m_1$, što predstavlja očekivane ukupne štete po jedinici vremena za složen Poissonov proces s dodatkom na premiju θ. Također ćemo pretpostaviti da se reosigurateljna premija računa kao $(1+\xi)(1-\alpha)\lambda m_1$. Budući da reosiguratelj plaća omjer $1-\alpha$ svake štete, $(1-\alpha)\lambda m_1$ predstavlja reosigurateljeve očekivane štete po jedinici vremena.
Zato je ξ dodatak na premiju koji koristi reosiguratelj. Osigurateljev prihod od premija nakon odbitka za reosiguranje je

$$[(1+\theta) - (1+\xi)(1-\alpha)]\lambda m_1\,. \tag{5.1}$$

Pretpostavit ćemo također da je $\xi \geq \theta$. Kada to ne bi bilo ispunjeno, bilo bi moguće da osiguratelj prebaci cijeli rizik na reosiguratelja i ostvari siguran profit. Da bi osigurateljev prihod od premija, nakon odbitka za reosiguranje, bio pozitivan mora biti

$$\begin{aligned} & 1+\theta > (1+\xi)(1-\alpha) \\ \text{t.j.} \quad & \alpha > (\xi-\theta)(1+\xi)\,. \end{aligned}$$

Postoji, međutim, važnije ograničenje za osiguratelja. Osigurateljev prihod od premija, nakon odbitka za reosiguranje, po jedinici vremena mora premašiti očekivane ukupne štete po jedinici vremena. U suprotnom je propast u beskonačnom vremenu sigurna (kao što smo primjetili u odjeljku 4.7). Nakon odbitka za reosiguranje, osigurateljeve očekivane ukupne štete po jedinici vremena su $\alpha\lambda m_1$. Stoga,

$$(1+\theta) - (1+\xi)(1-\alpha) > \alpha$$

ili

$$\alpha > 1 - \frac{\theta}{\xi}\,. \tag{5.2}$$

Nejednakost (5.2) specificira osigurateljev minimalni samopridržaj, budući da je

$$1 - \frac{\theta}{\xi} \geq \frac{\xi-\theta}{1+\xi} \;\text{ kada je }\; \theta \leq \xi$$

jedini zanimljiv slučaj. Ako su dodaci na premiju jednaki, tada nejednakost (5.2) postaje $\alpha > 0$. U tom slučaju postoji ugovor o dijeljenju rizika i moguć je svaki samopridržaj. Ako je, međutim, $\xi > \theta$, tada osiguratelj mora zadržati dio rizika. Promatrajmo prvo slučaj u kojem i osiguratelj i reosiguratelj koriste θ kao dodatak na premiju. Naći ćemo koeficijent prilagodbe kao funkciju samopridržaja α kada je $F(x) = 1 - e^{-0.1x}$.

Funkcija distribucije osigurateljevih pojedinačnih šteta, nakon odbitka za reosiguranje, je eksponencijalna s parametrom $0.1/\alpha$. To možemo vidjeti tako da uočimo da ako je $Y = \alpha X$, tada je

$$\mathbb{P}[Y \leq y] = \mathbb{P}[X \leq y/\alpha] = 1 - \exp\{-0.1y/\alpha\}\,.$$

Zato je jednadžba koja definira R (vidi formulu (3.2))

$$\begin{aligned}
& \lambda + (1+\theta)\lambda 10\alpha R = \lambda \int_0^\infty e^{Rx}(0.1/\alpha)e^{-0.1x/\alpha}\,dx \\
\Rightarrow \quad & 1 + (1+\theta)10\alpha R = \frac{1}{1-10\alpha R} && (5.3) \\
\Rightarrow \quad & R = \frac{\theta}{(1+\theta)10\alpha} \quad \text{za} \;\; 0 < \alpha \leq 1\,. && (5.4)
\end{aligned}$$

Vidi se da je R opadajuća funkcija od α. To ima smisla, jer čim je manji samopridržaj, to je manji rizik za reosiguratelja, te stoga oěkujemo da $\psi(U)$ raste, a R opada, po α.
Promotrimo sada što se događa kada je $\xi > 0$. U ostatku odjeljka 5.2 pretpostavimo da je

- distribucija pojedinačnih iznosa šteta jednaka $F(x) = 1 - e^{-0.1x}$
- osigurateljev dodatak na premiju jednak $\theta = 0.1$.

Pretpostavimo prvo da je reosigurateljev dodatak na premiju jednak $\xi = 0.2$, tako da je osigurateljev (neto) prihod od premije po jedinici vremena jednak $(12\alpha - 1)\lambda$. Nejednakost (5.2) pokazuje da osiguratelj mora zadržati barem 50% svake štete. Zbog toga ćemo vrijednost od α koja maksimizira vrijednost od R tražiti u intervalu $[0.5, 1]$. Jednadžba koja definira R je

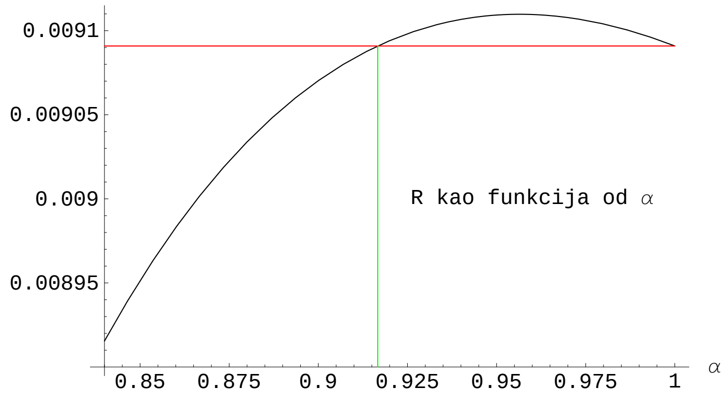
$$\lambda + (12\alpha - 1)\lambda R = \frac{\lambda}{1 - 10\alpha R}$$

- to slijedi iz (5.3), jer je samo premija različita - što vodi do

$$R = \frac{2\alpha - 1}{10(12\alpha^2 - \alpha)} \quad \text{za} \;\; 0.5 < \alpha \leq 1\,. \tag{5.5}$$

Kao i kada su dodaci na premiju jednaki, koeficijent prilagodbe ovisi o samopridržaju.
Tražimo vrijednost od α koja maksimizira R u (5.5).
Deriviramo R po α i dobijemo

$$\frac{dR}{d\alpha} = \frac{20(12\alpha^2 - \alpha) - (2\alpha - 1)10(24\alpha - 1)}{100(12\alpha^2 - \alpha)^2}\,.$$



Slika 12:

Nazivnik je uvijek pozitivan za vrijednosti od α u $[0.5, 1]$, tako da funkcija ima kritičnu točku kada je

$$20(12\alpha^2 - \alpha) = (2\alpha - 1)10(24\alpha - 1)$$

t.j., kada je $24\alpha^2 - 24\alpha + 1 = 0$.
Rješenja te kvadratne jednadžbe su 0.9564 i 0.0436, te je kritična točka koja nas zanima 0.9564. Promotrimo sljedeće vrijednosti:

α	R
0.5	0
0.9564	0.00911
1.0	0.00909

To pokazuje da R ima maksimum na $[0.5, 1]$ u 0.9564.
Slika 12. pokazuje R kao funkciju od α (danu s (5.5)) za vrijednosti od α veće od 0.85. Taj raspom vrijednosti od α odabran je da bi se naglasila važna svojstva grafa. Iscrtkana linija pokazuje vrijednost od R kada je $\alpha = 1$ (t.j., nema reosiguranja). Iz slike 12. može se vidjeti da postoji interval vrijednosti za α, $\beta < \alpha < 1$, takav da ako je samopridržaj u tom intervalu, vrijednost koeficijenta prilagodbe premašuje vrijednost kada je $\alpha = 1$. Vrijednost od β može se izračunati iz (5.5) izjednačivši vrijednost od R u β s vrijednosti od R u 1, što daje $\beta = 0.9167$. Strelica na slici 12. pokazuje tu vrijednost β.

U terminima maksimizacije koeficijenta prilagodbe, optimalni samopridržaj je $\alpha = 0.954$. Treba, međutim, primjetiti, da optimalnost u jednom smislu ne povlači optimalnost u nekom drugom. Na primjer, ako osiguratelj ne sklopi ugovor o reosiguranju, tada je očekivani profit po jedinici vremena $\theta\lambda m_1$ (t.j., λ, jer je $\theta = 0.1$ i $m_1 = 10$). Ako osiguratelj sklopi ugovor o reosiguranju sa samopridržajem 0.9564, tada je očekivani profit po jedinici vremena jednak 0.9128λ (t.j., prihod od premije, iz (5.1), manje očekivane štete).
Sada ćemo naći očekivani profit po jedinici vremena pomoću α i λ.
Već smo izračunali u (5.1), da kada je $\theta = 0.1$, $\xi = 0.2$ i $m_1 = 10$, osigurateljev neto prihod od premije jednak je $(12\alpha - 1)\lambda$. Osigurateljeve očekivane štete po jedinici vremena su $10\alpha\lambda$. Zato je osigurateljev očekivani profit po jedinici vremena jednak $(2\alpha - 1)\lambda$.
To pokazuje da je osigurateljev očekivani profit po jedinici vremena rastuća funkcija po α, te ako bi osiguratelj trebao izabrati α tako da maksimizira očekivani profit po jedinici vremena, izbor bi bio $\alpha = 1$. Ovaj primjer ilustrita opće stanovište - razina reosiguranja je balans između sigurnosti i profita.
Sada nalazimo vrijednost od α koja maksimizira R kada je reosigurateljev dodatak na premiju 0.3.
Iz (5.1), osigurateljev neto prihod od premije je $(13\alpha - 2)\lambda$, tako da je jednadžba koja definira R

$$\lambda + (13\alpha - 2)\lambda R = \frac{\lambda}{1 - 10\alpha R}$$

što vodi na

$$R = \frac{3\alpha - 2}{10(13\alpha^2 - 2\alpha)} \quad \text{za} \;\; 0.67 < \alpha \leq 1\,.$$

Po (5.2), osiguratelj mora zadržati barem 2/3 svake štete, tako da vrijednost od α koja maksimizira R tražimo u $[2/3, 1]$. Deriviranje daje

$$\frac{dR}{d\alpha} = \frac{30(13\alpha^2 - 2\alpha) - (3\alpha - 2)10(26\alpha - 2)}{100(13\alpha^2 - 2\alpha)^2}$$

i funkcija ima kritičnu točku kada je

$$30(13\alpha^2 - 2\alpha) = (3\alpha - 2)10(26\alpha - 2)$$

t.j. kada je $39\alpha^2 - 52\alpha + 4 = 0$.
Rješenja te kvadratne jednadžbe su 0.0820 i 1.2514, te nema kritičnih točaka u intervalu $[2/3, 1]$, i R je tom intervalu raste po α od 0 za $\alpha = 2/3$ do 0.00909

za $\alpha = 1$. Stoga je vrijednost od α koja maksimizira koeficijent prilagodbe jednaka 1.
Nije uvijek moguće povećati vrijednost koeficijenta prilagodbe sklapanjem ugovora o reosiguranju. Uočite da kada osiguratelj sklopi ugovor o reosiguranju, to smanjuje varijabilnost osigurateljevih ukupnih šteta. Međutim, kada je $\xi > 0$, osigurateljev dodatak na premiju, nakon odbitka za reosiguranje, opada, i očekujemo da vrijednost koeficijenta prilagodbe opada s dodatkom na premiju. Kada je reosigurateljev dodatak na premiju bio 0.3, smanjenje osigurateljeve sigurnosti uzrokovano smanjenjem dodatka na premiju ima veći utjecaj na koeficijent prilagodbe nego povećanje koje slijedi iz reosiguranje za sve vrijednosti od α.
Sada upotrebom (5.1) nalazimo osigurateljev dodatak na premiju, nakon odbitka za reosiguranje, i pokazujemo da je rastuća funkcija po α.
Dodatak na premiju nalazimo dijeljenjem očekivanog profita po jedinici vremena s očekivanim štetama po jedinici vremena. Očekivani profit po jedinici vremena je

$$[(1+\theta) - (1+\xi)(1-\alpha)]\lambda m_1 - \alpha\lambda m_1$$

tako da je dodatak na premiju

$$\begin{aligned} \theta' &= [(1+\theta) - (1+\xi)(1-\alpha) - \alpha]/\alpha \\ &= \xi - (\xi - \theta)/\alpha\,. \end{aligned}$$

Sada je $\dfrac{d\theta'}{d\alpha} = (\xi - \theta)/\alpha^2$ što je pozitivno zbog $\xi > 0$, tako da je θ' rastuća funkcija od α. Zato dodatak na premiju nakon odbitka za reosiguranje raste kada raste samopridržaj.

5.3 Maksimiziranje koeficijenta prilagodbe uz reosiguranje viška štete

U ovom odjeljku razmatrat ćemo utjecaj reosiguranja viška štete na koeficijent prilagodbe. Sljedeće pretpostavke vrijedit će u odjeljku 5.3:

- osigurateljev prihod od premije (prije reosiguranja) po jedinici vremena je $(1+\theta)\lambda m_1$
- premija za reosiguranje po jedinici vremena je $(1+\xi)\lambda\mathbb{E}(Z)$, gdje je $\xi \geq \theta$ reosigurateljev dodatak na premiju, a $Z = \max(0, X - M)$

Osigurateljeve pojedinačne štete imaju distribuciju $Y = \min(X, M)$, a osigurateljev prihod od premije, nakon odbitka za reosiguranje, je $c^* = (1 + \theta)\lambda m_1 - (1 + \xi)\lambda\mathbb{E}(Z)$, što daje jednadžbu koja definira R

$$\lambda + c^* R = \lambda \left(\int_0^M e^{Rx} f(x)\, dx + e^{RM}(1 - F(M)) \right) .$$

To je formula (3.2) s odrezanim iznosima šteta kao rezultat reosiguranja viška štete. Da bismo ilustrirali ideje, pogledajmo situaciju kada je $X \sim U(0, 20)$, tako da je $f(x) = 0.05$ za $0 < x < 20$. Tada za $0 < M < 20$

$$\mathbb{E}(Z) = \int_M^{20} (x - M)\, 0.05\, dx = 10 - M + 0.025 M^2 ,$$

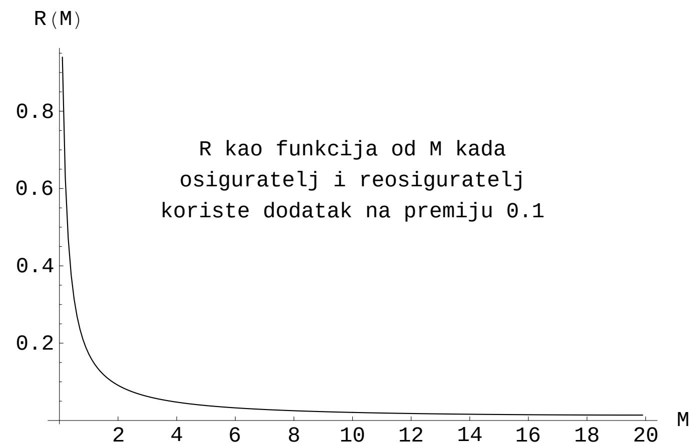
i

$$\begin{aligned} M_Y(R) &= \int_0^M e^{Rx} 0.05\, dx + e^{RM}(1 - 0.05M) \\ &= \frac{0.05}{R}(e^{RM} - 1) + e^{RM}(1 - 0.05M) . \end{aligned}$$

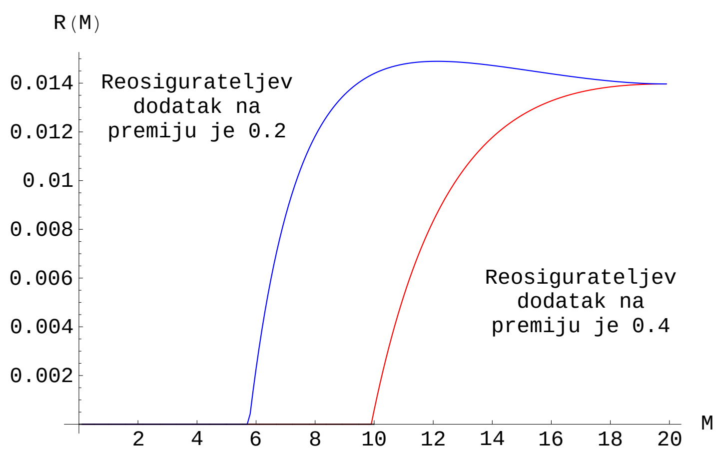
Jednadžba za R treba se numerički riješiti po R za dane vrijednosti od θ i ξ. Slika 13. pokazuje R kao funkciju od M za $\theta = \xi = 0.1$. Kao i u odjeljku 5.2, svaki samopridržaj je moguć kada su dodaci na premije jednaki. R je opadajuća funkcija od M.
Na Slici 13., R ide prema ∞ kada M ide prema 0.
Kada je $\theta < \xi$, postoji najmanji samopridržaj zbog istog razloga kao i u prethodnom odjeljku. Na primjer, kada je $\theta = 0.1$ i $\xi = 0.2$ osigurateljev neto prihod od premije, c^*, je $11\lambda - 1.2\lambda(10 - M + 0.025M^2)$ i to mora premašiti osigurateljeve očekivane štete, nakon odbitka za reosiguranje. Osigurateljeve neto štete jednake su $\lambda\mathbb{E}(X) - \lambda\mathbb{E}(Z)$, što daje $\lambda(M - 0.025M^2)$. Zato,

$$\begin{aligned} & -1 + 1.2M - 0.03M^2 > M - 0.025M^2 \\ \Rightarrow\ & M^2 - 40M + 200 < 0 \\ \Rightarrow\ & 5.8579 < M < 34.1421 \end{aligned}$$

Slijedi da je minimalni samopridržaj jednak 5.8579. Slično, kada je $\xi = 0.4$, minimalni samopridržaj je 10.
Slika 14. pokazuje R kao funkciju od M za sljedeće kombinacije od θ i ξ:



Slika 13:



Slika 14:

(a) $\theta = 0.1$ i $\xi = 0.2$ (plava linija)

(b) $\theta = 0.1$ i $\xi = 0.4$ (crvena linija)

Bez reosiguranja, t.j., za $M = 20$, osigurateljev koeficijent prilagodbe je 0.014 (bez obzira na reosigurateljev dodatak na premiju).
Iz Slike 14. može se vidjeti da za $\xi = 0.2$

$$\begin{aligned} R(M) &> R(20) \text{ za } 9.6 < M < 20 \\ R(M) &< R(20) \text{ za } M < 9.6 \end{aligned}$$

a za $\xi = 0.4$

$$R(M) < R(20) \text{ za } R < 20\,.$$

Zato je za $\xi = 0.2$ osiguratelju moguće povećati vrijednost koeficijenta prilagodbe sklapanjem ugovora o resoiguranju, ukoliko je samopridržaj iznad 9.6. Međutim, za $\xi = 0.4$, osiguratelj bi trebao zadržati cijeli rizik ako želi maksimizirati vrijednost koeficijenta prilagodbe. Kao i u slučaju kvotnog reosiguranja, osigurateljev očekivani profit po jedinici vremena smanjuje se uz reosiguranje.

POGLAVLJE 6 - TEORIJA POVJERENJA

Nastavni ciljevi:

(vii)1. Objasniti što se misli pod formulom za premiju povjerenja i opisati ulogu faktora povjerenja.

(vii)2. Objasniti Bayesovski pristup teoriji povjerenja i upotrijebiti ga za izvod premija povjerenja u jednostavnim slučajevima.

(vii)3. Objasniti empirijski Bayesov pristup teoriji povjerenja, i posebno njegove sličnosti sa i razlike od Bayesovskog pristupa.

(vii)4. Upotrijebiti empirijski Bayesov pristup za izvod formula za premiju povjerenja za dva standardna elementarna modela empirijske Bayesove teorije povjerenja, jedan koji uključuje i jedan koji ne uključuje volumene rizika (Modeli 1 i 2 u Formulae and Tables foe Actuarial Examinations).

(vii)5. Izložiti pretpostavke na kojima se temelje dva modela iz (vii)4.

(vii)6. Izračunati premije povjerenja za dva modela iz (vii)4.

0 Uvod

Rukovanje uvjetnim očekivanjima važna je tehnika u teoriji povjerenja, isto kao i u mnogim drugim područjima aktuarske znanosti. Neki od potrebnih rezultata u ovom poglavlju su:

> Za sve slučajne varijable X i Y (za koje postoje relevantni momenti) i za sve funkcije f (osim nekih specijalnih slučajeva koji nisu od praktičnog interesa)

(0.1) $$\mathbb{E}[X] = \mathbb{E}[\mathbb{E}[X|Y]]$$

(0.2) $$\mathbb{E}[f(Y)|Y] = f(Y)$$

$$\begin{aligned}\mathbb{E}[Xf(Y)] &= \mathbb{E}[\mathbb{E}[Xf(Y)|Y]] \\ &= \mathbb{E}[f(Y)\mathbb{E}[X|Y]]\,.\end{aligned} \tag{0.3}$$

Formula (0.1) je standardni rezultat. Formula (0.2) je intuitivno očigledna; ako je dano da Y poprima vrijednost, recimo, y, tada je poznato da $f(Y)$ mora primiti vrijednost $f(y)$. Prva jednakost u (0.3) je primjena formule (0.1); druga slijedi na isti način kao i (0.2). Ako je dano da Y poprima vrijednost, recimo, y, tada je poznato da $Xf(Y)$ mora primiti vrijednost $Xf(y)$. Dakle, štogod bila vrijednost od Y, $\mathbb{E}[Xf(Y)|Y]$ će biti jednako $f(Y)\mathbb{E}[X|Y]$.
Uvjetna nezavisnost je drugi važan pojam. Ako su dvije slučajne varijable X_1 i X_2 uvjetno nezavisne s obzirom na danu treću slučajnu varijablu Y, tada je:

$$\mathbb{E}[X_1X_2|Y] = \mathbb{E}[X_1|Y]\mathbb{E}[X_2|Y]\,. \tag{0.4}$$

Intuitivno to kaže da su i X_1 i X_2 ovisne o Y, ali ako je poznata vrijednost koju poprimi Y, tada su X_1 i X_2 nezavisne. To ne povlači da su X_1 i X_2 bezuvjetno nezavisne, tj. nezavisne ako nije poznata vrijednost koju primi Y. Prema tome, može se dogoditi da je:

$$\mathbb{E}[X_1X_2] \neq \mathbb{E}[X_1]\mathbb{E}[X_2]$$

iako vrijedi (0.4).

1 Povjerenje

1.1 Formula za premiju povjerenja

Osnovna ideja na kojoj se temelji formula za premiju povjerenja je intuitivno vrlo jednostavna i vrlo privlačna. Razmotrimo jedan krajnje jednostavan primjer.
Pretpostavimo da je lokalna vlast u malom gradu izvjestan broj godina upravljala voznim parkom od deset autobusa. Lokalna vlast želi osigurati taj vozni park u sljedećoj godini protiv šteta nastalih u nesrećama u kojima su uključeni ti autobusi. Potrebno je izračunati čistu (riziko) premiju za to osiguranje, tj., očekivane troškove šteta u nadolazećoj godini. Podaci za proteklih par godina za taj određeni vozni park autobusa pokazuju da su prosječni

godišnji troškovi šteta (za deset autobusa) bili £1 600. Pretpostavimo da pored te informacije postoje podaci o velikom broju voznih parkova autobusa kojima upravljaju lokalne vlasti unutar cijelog Ujedinjenog Kraljevstva koji pokazuju da su prosječni godišnji troškovi šteta po autobusu £250, tako da su prosječni godišnji troškovi šteta za vozni park od deset autobusa jednaki £2 500. Međutim, dok se iznos od £2 500 zasniva na puno više voznih parkova nego iznos od £1 600, neki od voznih parkova autobusa uključenih u taj velik skup podataka djeluju pod sasvim različitim uvjetima (za koje se smatra da utječu na broj i/ili veličinu šteta, na primjer u velikim gradovima ili ruralnim područjima) od određenog voznog parka koji je ovdje važan.
Postoje dva krajnja izbora za čistu premiju za sljedeću godinu:

(i) Može se odabrati £1 600 na temelju činjenice da se ta procjena zasniva na najodgovarajućim podacima, dok se procjena od £2 500 zasniva na manje relevantnim podacima, ili

(ii) Može se odabrati £2 500 na temelju činjenice da se to zasniva na više podataka, te je prema tome, u nekom smislu, pouzdaniji iznos.

Pristup povjerenja tom problemu je da se uzme težinska sredina ta dva krajnja odgovora, tj. da se čista premija izračuna kao:

$$Z \times 1\,600 + (1 - Z) \times 2\,500$$

gdje je Z neki broj između nula i jedan. Z se zove faktor povjerenja. Čisto kao ilustracija, pretpostavimo da je Z jednak 0.6, tako da je čista premija jednaka £1 960.
Na taj primjer ćemo se vratiti u sljedećoj sekciji da bismo ilustrirali neka pitanja, ali ćemo sada gornje ideje izraziti malo formalnije. Problem je procijeniti očekivane skupne štete, ili moguće, samo očekivani broj odštetnih zahtjeva u sljedećoj godini nastalih iz rizika. Pod rizikom se smatra pojedinačna polica ili grupa polica. Te police su tipično kratkoročne police, te će se zbog pogodnosti uzeti da je trajanje police jedna godina, iako bi moglo biti i kraće razdoblje. Sljedeće informacije su dostupne:

$\bar{X}$ je procjena očekivanih skupnih šteta/broja šteta za nadolazeću godinu zasnovana samo na podacima iz samog rizika

μ je procjena očekivanih skupnih šteta/broja šteta za nadolazeću godinu zasnovana na kolateralnim podacima, tj. na podacima o rizicima koji su slični, ali ne nužno identični, određenom riziku koji se promatra.

Formula za premiju povjerenja (ili procjenu povjerenjem skupnih šteta/broja šteta) za taj rizik je:

$$Z\bar{X} + (1 - Z)\mu \tag{1.1}$$

gdje je Z broj između nula i jedan, i zove se faktor povjerenja. Atraktivna obilježja formule za premiju povjerenja su njezina jednostavnost, i, u slučaju da su $\bar{X}$ i μ očito razumne alternative, lakoća kojom se može objasniti laiku.

1.2 Faktor povjerenja

Faktor povjerenja Z je samo težinski faktor. Njegova vrijednost odražava koliko se povjerenja daje podacima iz samog rizika ($\bar{X}$) u usporedbi s podacima iz veće grupe (μ), kao procjeni očekivanih skupnih šteta/broja šteta u sljedećoj godini - čim je viša vrijednost od Z, tim se više povjerenja daje $\bar{X}$ u usporedbi s μ i obratno. Tu ideju ćemo pojasniti na jednostavnom primjeru iz Sekcije 1.1.

Pretpostavimo da su podaci za određeni vozni park dostupni za više od pet godina. Na primjer, pretpostavimo da je procjena skupnih šteta u sljedećoj godini zasnovan na podacima iz tog voznog parka £1 600 kao i prije, ali da se sada ti podaci zasnivaju na desetgodišnjim podacima, a ne petogodišnjim. U tom slučaju se iznos od £1 600 smatra vjerodostojnijim od iznosa £2 500, što znači davanje faktoru povjerenja veće vrijednosti, recimo 0.75, a ne 0.6. Rezultirajuća procjena povjerenjem skupnih šteta bila bi £1 825.

Pretpostavimo sada da se iznos od £1 600 zasniva na petogodišnjim podacima, kao u Sekciji 1.1, ali da je iznos od £2 500 zasnovan na podacima o voznim parkovima koji djeluju u gradovima otprilike iste veličine kao onaj kojeg razmatramo, tj. da više ne sadrži podatke iz velikih gradova ili ruralnih područja. (Još uvijek se pretpostavlja da se iznos od £2 500 zasniva na značajno više podataka nego iznos £1 600.) U tom slučaju bi kolateralne podatke smatrali relevantnijim nego u Sekciji 1.1, te bi stoga faktor povjerenja bio odgovarajuće smanjen, na primjer sa 0.6 na 0.4, dajući premiju povjerenja £2 140.

Konačno, pretpostavimo da je situacija točno kao u Sekciji 1.1 osim da se iznos od £2 500 zasniva na podacima o voznim parkovima koji djeluju u Londonu i Glasgowu. U ovom slučaju kolateralni podaci se mogu smatrati manje relevantnim nego u Sekciji 1.1, te bi faktor povjerenja bio odgovarajuće povećan, na primjer sa 0.6 na 0.8 dajući premiju povjerenja £1 780.

Iz ovih jednostavnih primjera može se općenito vidjeti da je za očekivati da se faktor povjerenja ponaša kako slijedi:

(1.2) Čim je više podataka iz samog rizika, tim bi viša trebala biti vrijednost faktora povjerenja.

(1.3) Čim su relevantniji kolateralni podaci, tim bi niža trebala biti vrijednost faktora povjerenja.

Završna primjedba o faktoru povjerenja je ta da dok bi njegova vrijednost trebala odražavati količinu dostupnih podataka iz samog rizika, njegova vrijednost ne bi trebala ovisiti o stvarnim podacima iz samog rizika, tj. o vrijednosti $\bar{X}$. Kada bilo dopušteno da Z ovisi o $\bar{X}$, tada bi se svaka procjena skupnih šteta/broja šteta, recimo ϕ, koja prima vrijednosti između $\bar{X}$ i μ mogla napisati u obliku (1.1) izborom Z jednako $(\phi - \mu)/(\bar{X} - \mu)$.
Preostaje problem kako mjeriti relevantnost kolateralnih podataka i kako izračunati faktor povjerenja Z. Postoje dva pristupa tim problemima: Bayesovski pristup povjerenja i empirijska Bayesova teorija povjerenja.

2 Bayesovsko povjerenje

2.1 Uvod

Ovo proučavanje Bayesovskog pristupa teoriji povjerenja ilustrirano je promatranjem dva modela: Poisson/gama modela i normalna/normalna modela.

2.2 Poisson/gama model

Pretpostavimo da je potrebno procijeniti frekvenciju rizika, tj. očekivani broj odštetnih zahtjeva u nadolazećoj godini. Problem se može sažeti kako slijedi:
Pretpostavlja se da broj odštetnih zahtjeva u svakoj godini ima Poissonovu distribuciju s parametrom λ.
Vrijednost od λ je nepoznata, ali su moguće procjene njene vrijednosti kao na primjer "postoji 50% šansi da je vrijednost od λ između 50 i 150".
Preciznije, prije nego što su dostupni bilo kakvi podaci iz rizika, mišljenje o vrijednosti od λ je da ima gama(α, β) distribuciju.
Podaci o riziku su sada dostupni i pokazuju broj odštetnih zahtjeva u svakoj od proteklih n godina.

Ovaj problem ulazi točno u okvir Bayesovske statistike, i formalno se može sažeti kako slijedi:

(2.1) Slučajna varijabla X predstavlja broj šteta iz rizika u sljedećoj godini.

(2.2) Distribucija od X ovisi o fiksnoj, ali nepoznatoj, vrijednosti parametra λ.

(2.3) Uvjetna distribucija od X uz dano λ je Poisson(λ).

(2.4) Apriorna distribucija od λ je gama(α, β).

(2.5) $x_1, x_2, \ldots, x_n$ su protekle opažene vrijednosti od X. (Ti podaci označavat će se s $\underline{x}$.)

Problem je procijeniti λ uz dane podatke $\underline{x}$, i željena procjena je Bayesova procjena s obzirom na kvadratni gubitak, tj. $\mathbb{E}[\lambda|\underline{x}]$.
Aposteriorna distribucija od λ uz dano $\underline{x}$ je gama($\alpha + \sum_{i=1}^{n} x_i, \beta + n$) i

$$\mathbb{E}[\lambda|\underline{x}] = (\alpha + \sum_{i=1}^{n} x_i)/(\beta + n)\,. \tag{2.6}$$

Pogledamo li pažljivije formulu (2.6), vidimo da je očekivani broj šteta $\sum_{i=1}^{n} x_i/n$, a očekivani broj šteta zasnovan na apriornim vjerovanjima je očekivanje apriorne game distribucije, α/β.
Izraz za aposteriorno očekivanje (2.6) može se zapisati u obliku povjerenja:

$$\mathbb{E}[\lambda|\underline{x}] = Z\left[\sum_{i=1}^{n} x_i/n\right] + (1 - Z)\alpha/\beta \tag{2.7}$$

gdje je

$$Z = n/(\beta + n)\,. \tag{2.8}$$

Pretpostavimo sada da nikakvi podaci nisu dostupni iz samog rizika (tj. pretpostavimo $n = 0$) tako da je $Z = 0$. Jedina dostupna informacija za procjenu λ bila bi apriorna distribucija. Najbolja procjena od λ bilo bi tada očekivanje apriorne distribucije koje je α/β.
S druge strane, da su samo podaci iz rizika dostupni za procjenu λ, očigledna procjena bila bi $\sum_{i=1}^{n} x_i/n$. (To je procjena maksimalne vjerodostojnosti od

λ. Uočite da je ta procjena, koja ima ulogu od $\bar{X}$ u formuli premije povjerenja (1.1), linearna funkcija opaženih vrijednosti $x_1, x_2, \ldots, x_n$. To će biti važno u razmatranju empirijskog Bayesovog povjerenja u Sekcijama 3 i 4.)

Vrijednost od Z ovisi o količini dostupnih podataka za rizik n, i kolateralnoj informaciji, kroz β. To je ohrabrujuće s obzirom na primjedbu na kraju Sekcije 1.2. Kako n raste, uzoračka greška od $\sum_{i=1}^{n} x_i/n$ kao procjene za λ pada.

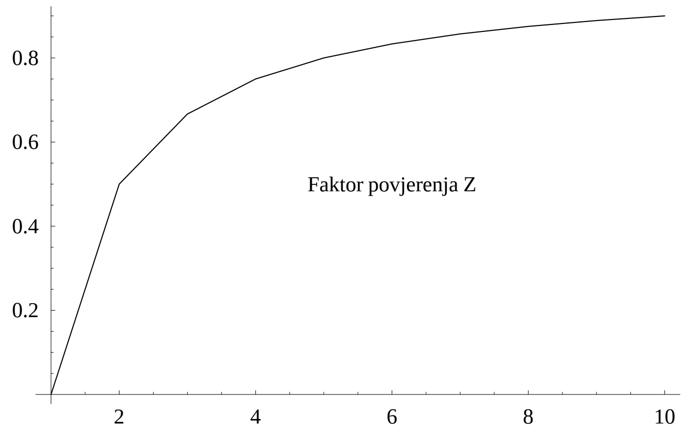
Slično, β odražava varijancu apriorne distribucije za λ. Tako Z odražava relativnu pouzdanost dviju alternativnih procjena od λ.

Imajući te tri točke na umu, Bayesovska procjena od λ dana s (2.7) ima oblik težinske sredine procjene od λ zasnovane samo na podacima iz samog rizika i procjene od λ zasnovane na nekoj drugoj informaciji. To je točno oblik procjene povjerenjem dane formulom (1.1) (gdje je "kolateralnim podacima" dana precizna interpretacija apriorne distribucije). Uočite da za ovaj model faktor povjerenja, Z, nije više nejasno definirana veličina kao u Sekciji 1.2; Z je dan točno formulom (2.8).

2.3 Numeričke ilustracije Poisson/gama modela

U ovoj sekciji razmatrat će se neki jednostavni primjeri u vezi s Poisson/gama modelom pomoću kojih ćemo donijeti neke daljnje zaključke o tom modelu.

Okružje i problem su isti kako u Sekciji 2.2. Vrijednost od λ je 150, tako da broj šteta nastalih iz rizika svake godine ima Poisson(150) distribuciju. U praksi neće biti poznata istinska vrijednost od λ. U radu kroz ovaj primjer pretpostavit će se samo da je poznata apriorna distribucija od λ, koju ćemo privremeno uzeti da bude gama(100,1). (Uočite da ta distribucija ima očekivanje 100, i standardnu devijaciju 10.) Točan broj šteta koje nastaju iz

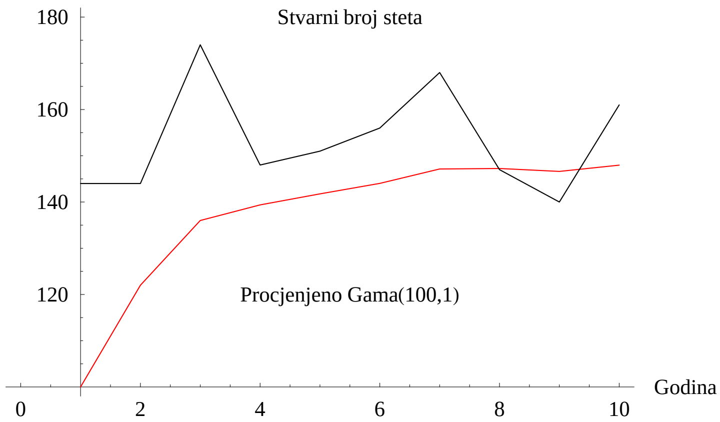


Slika 1:

tog rizika svake godine je kako slijedi

Godina	Broj šteta
1	144
2	144
3	174
4	148
5	151
6	156
7	168
8	147
9	140
10	161

Slika 1 pokazuje faktor povjerenja u uzastopnim godinama, a slika 2 pokazuje procjenu povjerenjem broja odštetnih zahtjeva u uzastopnim godinama za gama(100,1) apriornu distribuciju za λ.
Sljedeće jednostavne primjedbe vrijede za slike 1 i 2. Kao prvo, iz slike 1 može se vidjeti da faktor povjerenja raste s vremenom. To odgovara komentaru na kraju Sekcije 1.2; kako vrijeme prolazi više se podataka prikuplja iz samog rizika i, čim je više podataka iz rizika, tim bi viši trebao biti faktor povjerenja zbog rastuće pouzdanosti podataka u procjeni pravog ali nepoznatog očekivanog broja šteta za rizik. (Matematički, činjenica da Z

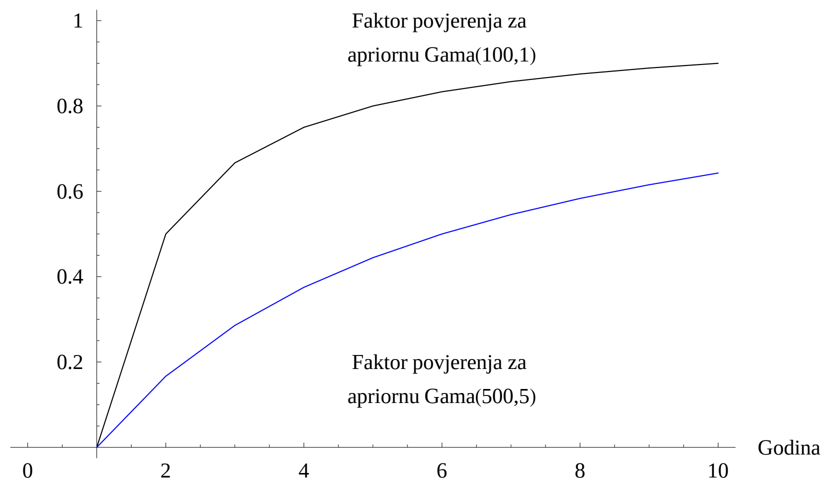


Slika 2:

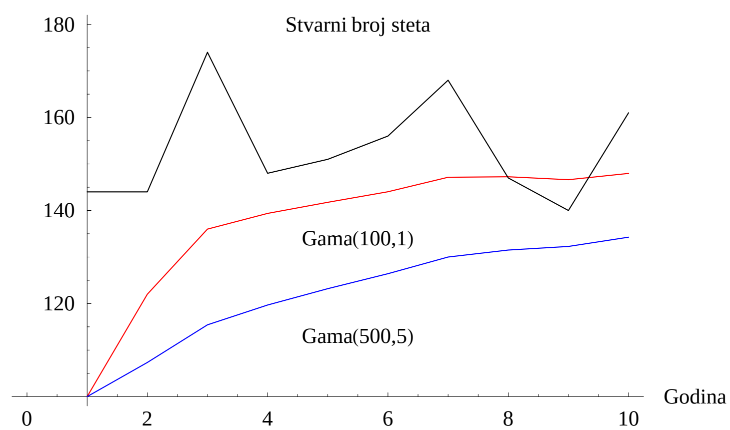
raste s vremenom za taj određen model proizlazi jednostavno iz toga da je $n/(\beta + n)$ rastuća funkcija od n za svaku pozitivnu vrijednost β.) Druga primjedba odnosi se na broj šteta prikazan na slici 2. Početna procjena je 100, očekivanje apriorne distribucije od λ. Međutim, pokazuje se da je to slaba procjena budući da su svi stvarni brojevi šteta oko 150, i nijedan u prvih deset godina nije niži od 140. Slika 2 pokazuje kako procijenjeni broj šteta raste s vremenom sve dok ne dostigne razinu stvarnog broja šteta nakon osam godina. Porast je posljedica progresivno veće težine, tj. povjerenja, koje se daje podacima iz samog rizika i odgovarajuće manje težine koja se daje kolateralnim podacima, tj. apriornoj distribuciji od λ.

Pretpostavimo sada da je apriorna distribucija od λ gama(500,5), a ne gama(100,1). Slike 3 i 4 pokazuju redom faktor povjerenja i procijenjeni broj šteta u uzastopnim godinama za apriornu gama(500, 5) distribuciju, u oba slučaja prikazanih za usporedbu zajedno s vrijednostima za apriornu gama(100, 1) distribuciju.

Slike 3 i 4 pokazuju da faktor povjerenja i procijenjeni broj šteta za apriornu gama(500,5) distribuciju imaju ista opća obilježja kao i kod apriorne gama(100,1) distribucije, tj. rastući faktor povjerenja i rastući procijenjeni broj šteta. Najočiglednija razlika između ta dva slučaja je da faktor povjerenja raste sporije za apriornu gama(500,5) distribuciju nego za apriornu gama(100,1) distribuciju. (Matematički je to jednostavna posljedica više vrijednosti za β kod gama(500,5) distribucije. Vidi (2.8).) To obilježje može



Slika 3:



Slika 4:

se objasniti pomoću teorije povjerenja kako slijedi. Procijenjen broj šteta zasnovan na kolateralnim podacima, tj. očekivanje apriorne distribucije za λ, isto je za obje apriorne distribucije, naime 100. Međutim, standardna devijacija apriorne distribucije niža je za apriornu gama(500, 5) distribuciju, tj. $\sqrt{20} = 4.472$, nego za apriornu gama(100,1) distribuciju, tj. 10. Veličina standardne devijacije apriorne distribucije može se interpretirati kao pokazatelj koliko povjerenja se daje početnoj procjeni broja šteta; čim je manja standardna devijacija apriorne distribucije, tim se više vjeruje u pouzdanost te početne procjene. Budući da u Bayesovskom povjerenju apriorna distribucija predstavlja "kolateralne podatke", gornja tvrdnja može se izraziti kao "čim je manja standardna devijacija apriorne distribucije, tim se relevantnijim smatraju kolateralni podaci". Uz tu interpretaciju, očekuje se da će manja standardna devijacija apriorne distribucije rezultirati manjim faktorom povjerenja (vidi komentar (1.3) na kraju Sekcije 1.2), a to se vidi na slici 3.

Evo i posljednje, ali ne nevažne, opće napomene o tome što je načinjeno u Sekciji 2. Problem je bio procijeniti očekivani broj šteta u nadolazećoj godini. Budući da X predstavlja broj šteta u nadolazećoj godini, to bi se naivno moglo interpretirati kao zahtjev za procjenom od $\mathbb{E}[X]$. To se može jednostavno izračunati kako slijedi:

$$\begin{aligned} \mathbb{E}[X] &= \mathbb{E}[\mathbb{E}[X|\lambda]] \\ &= \mathbb{E}[\lambda] \\ &= \alpha/\beta\,. \end{aligned}$$

U stvari se zahtijeva $\mathbb{E}[X|\underline{x}]$. Da bi se vidjelo da α/β nije razuman odgovor za taj problem, pogledajte sliku 2. Prihvatiti α/β kao odgovor za problem značilo bi procijeniti broj šteta svake godine sa 100 u slici 2. To bi očigledno rezultiralo slabijim procjenama nego procjenama povjerenjem pokazanim na slici 2.

2.4 Normalna/normalna model

U ovoj sekciji razmatramo drugi model.

Problem u ovoj sekciji je procijeniti čistu premiju, tj. očekivane skupne štete za rizik. Neka slučajna varijabla X predstavlja skupne štete za rizik u nadolazećoj godini. Pretpostavlja se sljedeće.

(2.9) Distribucija od X ovisi o fiksnoj, ali nepoznatoj, vrijednosti parametra θ.

(2.10) Uvjetna distribucija od X uz dano θ je $N(\theta, \sigma_1^2)$.

(2.11) Neizvjesnost oko vrijednosti od θ modelira se na uobičajeni Bayesovski način smatrajući θ slučajnom varijablom.

(2.12) Apriorna distribucija od θ je $M(\mu, \sigma_2^2)$.

(2.13) Vrijednosti od μ, σ_1 i σ_2 su poznate.

(2.14) n proteklih opaženih vrijednosti od X označene su s $x_1, x_2, \ldots, x_n$, ili kraće, $\underline{x}$.

Kada bi vrijednost od θ bila poznata, ispravna čista premija za taj rizik bila bi $\mathbb{E}[X|\theta]$, što je, iz (2.10), jednako θ. Problem je procijeniti $\mathbb{E}[X|\theta]$ uz dano $\underline{x}$ i upotrijebit će se, kao i kod Poisson/gamma modela, Bayesovska procjena u odnosu na kvadratni gubitak. To znači da će procjena biti $\mathbb{E}[\mathbb{E}[X|\theta]|\underline{x}]$, što je isto što i $\mathbb{E}[\theta|\underline{x}]$.
Aposteriorna distribucija od θ uz dano $\underline{x}$ je:

$$N\left(\frac{\mu\sigma_1^2 + n\sigma_2^2\bar{x}}{\sigma_1^2 + n\sigma_2^2}, \frac{\sigma_1^2\sigma_2^2}{\sigma_1^2 + n\sigma_2^2}\right)$$

gdje je

$$\bar{x} = \sum_{i=1}^{n} x_i/n$$

i

$$\begin{aligned}\mathbb{E}[\theta|\underline{x}] &= \frac{\mu\sigma_1^2 + n\sigma_2^2\bar{x}}{\sigma_1^2 + n\sigma_2^2} \\ &= \frac{\sigma_1^2}{\sigma_1^2 + n\sigma_2^2}\mu + \frac{n\sigma_2^2}{\sigma_1^2 + n\sigma_2^2}\bar{x} \\ &= Z\bar{x} + (1-Z)\mu \qquad (2.15)\end{aligned}$$

gdje je

$$Z = \frac{n}{n + \sigma_1^2/\sigma_2^2}\,. \qquad (2.16)$$

Formula (2.15) je procjena povjerenjem za $\mathbb{E}[\theta|\underline{x}]$, budući da je težinska sredina dviju procjena: prva, $\bar{x}$, je procjena maksimalne vjerodostojnosti zasnovana isključivo na podacima iz samog rizika, a druga, μ, je najbolja moguća procjena ako nisu dostupni nikakvi podaci iz rizika.
Uočite da je, isto kao i kod Poisson/gama modela, procjena zasnovana isključivo na podacima iz samog rizika linearna funkcija opaženih vrijednosti.
Evo nekoliko daljnjih primjedbi o faktoru povjerenja Z danog formulom (2.16). Kao prvo, Z je uvijek između nula i jedan. Drugo, on je rastuća funkcija od n, broja dostupnih podataka. Konačno, on je rastuća funkcija od σ_2, standardne devijacije apriorne distribucije. To su točno obilježja koja se očekuju od faktora povjerenja.

2.5 Daljnje napomene o normalna/normalna modelu

U Sekciji 2.4 razmatran je normalna/normalna model unutar okvira Bayesovske statistike. U ovoj sekciji razmatrat ćemo isti model, uz iste pretpostavke, ali na malo drugačiji način.
Razlog zašto se to čini je taj da će neke primjedbe pomoći pri razmatranju empirijske Bayesovske teorije povjerenja u Sekciji 3.
U ovoj sekciji, isto kao i u Sekciji 2.4, problem je procijeniti očekivane skupne štete nastale svake godine iz rizika. Neka su:

$$X_1, X_2, \ldots, X_n, X_{n+1}, \ldots$$

slučajne varijable koje predstavljaju skupne štete u uzastopnim godinama. Pretpostavlja se sljedeće:

(2.17) Distribucija svake od X_j ovisi o vrijednosti fiksnog, ali nepoznatog, parametra θ.

(2.18) Uvjetna distribucija od X_j uz dano θ je $N(\theta, \sigma_1^2)$.

(2.19) Uz dano θ, slučajne varijable $\{X_j\}$ su nezavisne.

(2.20) Apriorna distribucija od θ je $N(\mu, \sigma_2^2)$.

(2.21) Već su opažene vrijednosti od $X_1, \ldots, X_n$ i potrebno je procijeniti očekivane skupne štete u nadolazećoj, tj. $(n+1)$-voj, godini.

Važno je uvidjeti da su pretpostavke i problem naznačeni gore potpuno isti kao i pretpostavke i problem naznačeni u Sekciji 2.4. Malo drugačije oznake koriste se u ovoj sekciji: u Sekciji 2.4, $X_1, \ldots, X_n$ bili su označeni s $x_1, \ldots, x_n$, jer se pretpostavilo da su njihove vrijednosti poznate, a X_{n+1} bio je označen samo X. Pretpostavke (2.17), (2.18) i (2.20) bile su eksplicitne u Sekciji 2.4. Pretpostavka (2.21) također je bila eksplicitna u Sekciji 2.4, uz upravo objašnjenu promjenu notacije. Jedina od gornjih pretpostavki koja nije bila eksplicitna u Sekciji 2.4 je (2.19).

Nakon što smo naglasili da je sve isto kao i u Sekciji 2.4, razmatrat ćemo neke posljedice gornjih pretpostavki. Neke od važnih posljedica su:

(2.22) Uz dano θ, slučajne varijable $\{X_j\}$ su jednako distribuirane, te također nezavisne.

(2.23) Slučajne varijable $\{X_j\}$ su (bezuvjetno) jednako distribuirane.

(2.24) Slučajne varijable $\{X_j\}$ nisu (bezuvjetno) nezavisne.

(2.22) je neposredna posljedica pretpostavke (2.18), koja kaže da uz dano θ, svaka X_j ima istu $N(\theta, \sigma_1^2)$ distribuciju. (2.23) je druga posljedica pretpostavke (2.18); formalnije, vrijedi sljedeća formula za bezuvjetnu funkciju distribucije od X_j:

$$\mathbb{P}(X_j \leq y) = \int_{-\infty}^{\infty} \frac{1}{\sigma^2\sqrt{2\pi}} \exp\{-(\theta-\mu)^2\} \Phi\left(\frac{y-\theta}{\sigma_1}\right) d\theta$$

gdje je $\Phi(\cdot)$ standardizirana normalna funkcija distribucije. Taj izraz je isti za svaku vrijednost od j, te je tvrdnja (2.23) istinita. Tvrdnja (2.24) ne mora biti odmah očigledna, ali se može pokazati kako slijedi.

Upotrebom formule (0.1) i činjenice da su, uz dano θ, X_1 i X_2 uvjetno nezavisne:

$$\begin{aligned}
\mathbb{E}[X_1X_2] &= \mathbb{E}[\mathbb{E}[X_1X_2|\theta]] \\
&= \mathbb{E}[\mathbb{E}[X_1|\theta]\mathbb{E}[X_2|\theta]] \\
&= \mathbb{E}[\theta^2] \qquad \text{(zbog } \mathbb{E}[X_1|\theta] = \mathbb{E}[X_2|\theta] = \theta) \\
&= \mu^2 + \sigma_2^2
\end{aligned}$$

Ideja iz ovog argumenta ponovno će se koristiti kasnije.

Sada, kada bi X_1 i X_2 bile bezuvjetno nezavisne

$$\mathbb{E}[X_1X_2] = \mathbb{E}[X_1]\mathbb{E}[X_2]\,.$$

Međutim, ponovnom upotrebom (0.1):

$$\mathbb{E}[X_1] = \mathbb{E}[\mathbb{E}[X_1|\theta]] = \mathbb{E}[\theta] = \mu\,.$$

Slično, $\mathbb{E}[X_2] = \mu$. Stoga:

$$\mathbb{E}[X_1 X_2] = \mu^2 + \sigma_2^2 \neq \mathbb{E}[X_1]\mathbb{E}[X_2]$$

To pokazuje da X_1 i X_2 nisu bezuvjetno nezavisne. Veza između X_1 i X_2 je ta da su njihova očekivanja izabrana iz zajedničke distribucije. Ako je to očekivanje, θ, poznato, tada se veza kida, i postoji uvjetna nezavisnost.

2.6 Diskusija Bayesovskog pristupa povjerenju

Pristup korišten u Poisson/gama i normalna/normalna modelima u suštini je isti - jedina važna razlika je u pretpostavkama o distribuciji. Pristup se može sažeti kako slijedi:

(i) Postavljen je problem, na primjer odrediti distribuciju veličine štete ili broja odštetenih zahtjeva. U oba slučaja cilj je procijeniti neku veličinu koja karakterizira tu distribuciju, na primjer očekivani broj šteta ili očekivani iznos šteta.

(ii) Načine se neke pretpostavke o modelu unutar Bayesovskog okvira, na primjer pretpostavi se da je distribucija broja odštetnih zahtjeva Poissonova, te da nepoznati parametar Poissonove distribucije ima gama distribuciju sa specificiranim parametrima

(iii) Izvede se (Bayesovska) procjena određene veličine.

(iv) Tada se pokaže da je ta procjena u obliku procjene povjerenjem, tj. u obliku formule (1.1).

Taj pristup je bio vrlo uspješan u ova dva slučaja. Pojam kolateralnih podataka postavljen je vrlo precizno (interpretirajući ih pomoću apriorne distribucije), i dobivene su formule za računanje faktora povjerenja. Koji su, dakle, nedostaci tog pristupa?
Prva poteškoća je u tome da li je Bayesovski pristup problemu prihvatljiv, i, ako jeste, koje vrijednosti pridružiti parametrima apriorne distribucije. Na primjer, iako Poisson/gama model daje formulu (2.8) za računanje faktora povjerenja, ta formula uključuje parametar β. Nije se razmatralo kako bi se mogla odabrati vrijednost za β. Bayesovski pristup izboru vrijednosti parametara apriorne distribucije je tvrdnja da oni sumiraju subjektivni stupanj vjerovanja u moguće vrijednosti veličine koja se procjenjuje, na primjer, očekivanje broja šteta, λ, za Poisson/gama model.

Druga poteškoća je ta da čak ako problem ulazi u Bayesovski okvir, Bayesovski pristup ne mora funkcionirati, u smislu da se procjena koju daje ne mora moći preurediti u oblik procjene povjerenjem. To se može ilustrirati korištenjem uniformne apriorne distribucije s Poissonovom distribucijom za broj šteta.

3 Empirijska Bayesova teorija povjerenja: model 1

3.1 Uvod

Empirijska Bayesova teorija povjerenja je ime dano određenom pristupu problemima iz Sekcije 1. Taj pristup je doveo do golemog broja različitih modela s raznim stupnjevima kompleksnosti. U ovom poglavlju proučavat će se dva takva modela . U ovoj sekciji proučavat će se najjednostavniji mogući model. Iako taj model, na kojeg ćemo se pozivati kao na Model 1, nije vrlo koristan u praksi, on tvori dobar uvod u principe na kojoj se temelji empirijska Bayesova teorija povjerenja (EBTP) (engl. empirical Bayes credibility theory - EBCT). Model naročito pokazuje sličnosti i razlike između empirijskog Bayesovog i čisto Bayesovskog pristupa teoriji povjerenja. U Sekciji 4 proučavat će se proširenje Modela 1 koje je puno korisnije u praksi.

3.2 Model 1: specifikacija

U ovoj sekciji postavit će se pretpostavke za Model 1 EBTP-a. Model 1 može se smatrati poopćenjem normalna/normalna modela. To će se detaljnije razmatrati kasnije u sekciji.
Problem koji nas zanima je procjena čiste premije, ili možda frekvencije šteta, za rizik. Neka $X_1, X_2, \ldots$ označavaju skupne štete tog rizika u uzastopnim razdobljima. Preciznija tvrdnja problema je da je, nakon što su opaženi $X_1, X_2, \ldots, X_n$, potrebno procijeniti vrijednost od X_{n+1}. Od sada nadalje, $X_1, X_2, \ldots, X_n$ označavat će se s $\underline{X}$.
Sljedeće pretpostavke vrijedit će za Model 1 EBTP-a.

(3.1) Distribucija svake X_j ovisi o parametru, u oznaci θ, čija je vrijednost fiksna (i ista za sve X_j-ove), ali nepoznata.

(3.2) Uz dano θ, X_j su nezavisni i jednako distribuirani.

Parametar θ se zove parametar rizika. On može biti realan broj, kao u Sekciji 2, ili općenitija veličina kao što je skup realnih brojeva.
Posljedica ove dvije pretpostavke je:

(3.3) slučajne varijable $\{X_j\}$ su nezavisne i jednako distribuirane.

Važno je uočiti da:

(3.4) X_j nisu (nužno) bezuvjetno nezavisne.

Gornje pretpostavke i posljedice su sve, ili učinjene, ili uočene za normalna/normalna model Bayesovskog povjerenja u Sekciji 2.5. Vidi (2.17), (2.22), (2.23), (2.24).
Uvedimo neke oznake. Definiramo $m(\theta)$ i $s^2(\theta)$ kako slijedi:

$$\begin{aligned} m(\theta) &= \mathbb{E}[X_j|\theta] \\ s^2(\theta) &= \mathrm{Var}[X_j|\theta]\,. \end{aligned}$$

Dvije stvari treba uočiti o $m(\theta)$ i $s^2(\theta)$. Prva je da, budući da su, uz dano θ, X_j jednako distribuirani, niti $m(\theta)$ niti $s^2(\theta)$ ne ovise o j, kao što i sugerira njihova notacija. Druga je da su, budući da se θ tretira kao slučajna varijabla, $m(\theta)$ i $s^2(\theta)$ slučajne varijable.
Kada bi bile poznate vrijednost od θ i distribucija od X_j uz dano θ, očigledna procjena za očekivane skupne štete, ili očekivani broj šteta, u svakoj budućoj godini bila bi $m(\theta)$. Budući da je pretpostavljeno da je θ nepoznat, problem je:

(3.5) procijeniti $m(\theta)$ uz dano $\underline{X}$.

Sličnosti između Modela 1 EBTP-a i normalna/normalna modela mogu se sažeti kako slijedi:

(i) Uloga od θ jednaka je u oba modela: ona karakterizira temeljne distribucije procesa koji se modeliraju, na primjer distribucije skupnih šteta za svaku poslovnu godinu. Vidi (2.17) i (3.1).

(ii) Pretpostavke koje se odnose na bezuvjetnu distribuciju od X_j su iste: X_j su jednako distribuirani u oba slučaja. Vidi (2.23), (2.24), (3.3) i (3.4).

(iii) Pretpostavke koje se odnose na (uvjetnu) distribuciju od X_j uz dano θ su iste: X_j su uvjetno nezavisne u oba slučaja. Vidi (2.19), (2.22) i (3.2).

Model 1 EBTP-a može se smatrati poopćenjem normalna/normalna modela. Određene točke u kojima se razlikuje, tj. u kojima EBTP poopćava normalna/normalna model su

(i) $\mathbb{E}[X_j|\theta]$ je neka funkcija od θ, $m(\theta)$, za Model 1 EBTP-a, a jednostavno je θ za normalna/normalna model. Stoga:

$$\mathrm{Var}[m(\theta)]\ \{\text{EBTP}\}\ \text{odgovara}\ \sigma_2^2 (= \mathrm{Var}[\theta])\ \{\text{normalna/normalna}\}\quad .$$

(ii) $\mathrm{Var}[X_j|\theta]$ je funkcija od θ, $s^2(\theta)$, za Model 1 EBTP-a, a konstanta je, σ_1^2 za normalna/normalna model. Stoga.

$$\mathbb{E}[s^2(\theta)]\ \text{odgovara}\ \sigma_1^2 (= \mathrm{Var}[X_j|\theta])\ .$$

(iii) Normalna/normalna model ima vrlo precizne pretpostavke o distribuciji od X_j uz dani θ, koja je $N(\theta, \sigma_1^2)$, i o distribuciji od θ, koja je $N(\mu, \sigma_2^2)$. Model 1 EBTP-a nema takvih distribucijskih pretpostavki.

(iv) Parametar rizika, θ, je realan broj za normalna/normalna model, ali može biti općenitija veličina za Model 1 EBTP-a.

3.3 Model 1: izvod premije povjerenja

U prošloj sekciji proučavane su pretpostavke Modela 1 EBTP-a i naglašene su sličnosti tog modela i normalna/normalna modela. U ovoj sekciji dobit će se rješenje problema sažetog u (3.5), procjene od $m(\theta)$ uz dane podatke $\underline{X}$. Ako se prihvati Bayesovski pristup tom problemu, očigledna procjena od $m(\theta)$ uz dano $\underline{X}$ bila bi aposteriorno očekivanje:

$$\mathbb{E}[m(\theta)|\underline{X}]\,. \tag{3.6}$$

To je očita procjena, budući da je to Bayesovska procjena od $m(\theta)$ s obzirom na funkciju kvadratnog gubitka. Problem s (3.6) kao s procjenom je da ne vodi uvijek do odgovora koji se može izraziti u obliku premije povjerenja. Poželjan je odgovor na problem postavljen u (3.5) onaj koji ima privlačan oblik formule povjerenja (1.1). Formula (1.1) uključuje procjenu zasnovanu na podacima iz samog rizika, u oznaci $\bar{X}$, te nekim drugim veličinama, Z i μ. U Poisson/gama i normalna/normalna modelu procjena $\bar{X}$ je bila posebno jednostavna, budući da je bila linearna u opaženim podacima. Ako je željena procjena linearna u opaženim podacima, odabrat će se najbolji linearni procjenitelj. Drugim riječima, procjena od $m(\theta)$ biti će ograničena na oblik:

$$a_0 + a_1X_1 + a_2X_2 + \cdots + a_nX_n \tag{3.7}$$

gdje su $a_0, a_1, \ldots, a_n$ konstante izabrane tako da je (3.7) "najbolja" linearna procjena od $m(\theta)$ među svim procjenama oblika (3.7). Potrebno je odrediti što se smatra "najboljim" u ovom slučaju. Razumni kriterij je izabrati konstante tako da se minimizira očekivana kvadratna udaljenost između $m(\theta)$ i procjenitelja od $m(\theta)$, tj. (3.7). (To je razumno, jer odgovara kriteriju za Bayesovsku procjenu s obzirom na funkciju kvadratnog gubitka.) Put kojim će se riješiti problem postavljen u (3.5) može se sažeti kako slijedi:
$m(\theta)$ procijenit će se funkcijom oblika:

$$a_0 + a_1X_1 + a_2X_2 + \cdots a_nX_n$$

gdje su konstante $a_0, a_1, \ldots, a_n$ odabrane tako da minimiziraju:

$$\mathbb{E}[(m(\theta) - a_0 - a_1X_1 - a_2X_2 - \cdots - a_nX_n)^2]. \tag{3.8}$$

Problem ćemo riješiti deriviranjem (3.8) s obzirom na $a_0, a_1, \ldots, a_n$ i izjednačavanjem svake derivacije s nulom. Prije toga potrebni su neki tehnički rezultati.

(i) Upotrebom (0.1) i (0.3) i definicije od $m(\theta)$:

$$\begin{aligned}
\mathbb{E}[X_j m(\theta)] &= \mathbb{E}[\mathbb{E}[X_j m(\theta)|\theta]] \\
&= \mathbb{E}[m(\theta)\mathbb{E}[X_j|\theta]] \\
&= \mathbb{E}[m(\theta)^2] \\
&= \mathrm{Var}[m(\theta)] + (\mathbb{E}[m(\theta)])^2
\end{aligned}$$

(ii) Upotrebom (0.1) i činjenice da su, uz dano θ, X_j i X_k nezavisne za $j \neq k$:

$$\begin{aligned}
\mathbb{E}[X_j X_k] &= \mathbb{E}[\mathbb{E}[X_j X_k|\theta]] \\
&= \mathbb{E}[\mathbb{E}[X_j|\theta]\mathbb{E}[X_k|\theta]] \\
&= \mathbb{E}[m(\theta)m(\theta)] \\
&= \mathrm{Var}[m(\theta)] + (\mathbb{E}[m(\theta)])^2
\end{aligned}$$

(iii) Upotrebom (0.1) i definicije od $s^2(\theta)$:

$$\begin{aligned}\mathbb{E}[X_j^2] &= \mathbb{E}[\mathbb{E}[X_j^2|\theta]] \\ &= \mathbb{E}[\mathrm{Var}[X_j|\theta] + (\mathbb{E}[X_j|\theta])^2] \\ &= \mathbb{E}[s^2(\theta)] + \mathbb{E}[m(\theta)^2] \\ &= \mathbb{E}[s^2(\theta)] + \mathrm{Var}[m(\theta)] + (\mathbb{E}[m(\theta)])^2\end{aligned}$$

Sada deriviramo (3.8) s obzirom na a_0 i izjednačimo s nulom. To daje sljedeću jednadžbu:

$$\mathbb{E}[m(\theta) - a_0 - \sum_{j=1}^{n} a_j X_j] = 0$$

što, po definiciji od $m(\theta)$, daje:

$$a_0 = \mathbb{E}[m(\theta)](1 - \sum_{j=1}^{n} a_j) \tag{3.9}$$

Sada deriviramo (3.8) s obzirom na a_k, $k \neq 0$, i izjednačimo derivaciju s nulom. To daje:

$$\mathbb{E}[X_k(m(\theta) - a_0 - \sum_{j=1}^{n} a_j X_j)] = 0$$

otkud:

$$\mathrm{Var}[m(\theta)]+(\mathbb{E}[m(\theta)])^2 - a_0\mathbb{E}[m(\theta)] - \sum_{j=1}^{n} a_j(\mathrm{Var}[m(\theta)]+(\mathbb{E}[m(\theta)])^2) - a_k\mathbb{E}[s^2(\theta)] = 0$$

Sređivanje zadnje jednadžbe daje:

$$a_k\mathbb{E}[s^2(\theta)] = (1 - \sum_{j=1}^{n} a_j)(\mathrm{Var}[m(\theta)] + (\mathbb{E}[m(\theta)])^2) - a_0\mathbb{E}[m(\theta)] \tag{3.10}$$

Jednadžba (3.10) vrijedi za $k = 1, 2, 3, \ldots, n$. Važno za uočiti u (3.10) je da vrijednost od a_k ne ovisi o k. Drugim riječima, (3.10) pokazuje da je:

$$a_1 = a_2 = \cdots = a_n \, .$$

Razlog tome je da je problem simetričan u opaženim vrijednostima podataka $X_1, X_2, \ldots, X_n$. Označimo zajedničku vrijednost od $a_1, a_2, \ldots, a_n$ sa Z/n tako da je:

$$Z = \sum_{j=1}^{n} a_j$$

a procjenitelj se može zapisati kao:

$$a_0 + \sum_{j=1}^{n} a_j X_j = a_0 + Z\bar{X} \tag{3.11}$$

gdje je:

$$\bar{X} = \sum_{j=1}^{n} X_j/n\,.$$

Jednadžbe (3.9) i (3.10) su sada dvije linearne jednadžbe s dvije nepoznanice, a_0 i Z. Rješenje tih dviju jednadžbi je:

$$a_0 = (1 - Z)\mathbb{E}[m(\theta)] \tag{3.12}$$

$$Z = \frac{n}{n + \mathbb{E}[s^2(\theta)]/\mathrm{Var}[m(\theta)]}\,. \tag{3.13}$$

Prema tome, rješenje problema procjene od $m(\theta)$, uz dano $\underline{X}$, dano je desnom stranom u (3.11) sa a_0 i Z danim redom formulama (3.12) i (3.13).
Ukratko, procjena od $m(\theta)$ uz dano $\underline{X}$ dana Modelom 1 EBTP-a je:

$$(1 - Z)\mathbb{E}[m(\theta)] + Z\bar{X}$$

gdje je

$$\bar{X} = \sum_{j=1}^{n} X_j/n$$

i

$$Z = \frac{n}{n + \frac{\mathbb{E}[s^2(\theta)]}{\mathrm{Var}[m(\theta)]}} \tag{3.14}$$

Prva i najvažnija stvar koju treba primjetiti o rješenju je da je u obliku procjene povjerenjem. Drugim riječima, desna strana od (3.11) ima oblik formule (1.1) sa:

$$\begin{array}{ll} \mathbb{E}[m(\theta)] & \text{u ulozi od } \mu, \text{ i} \\ \sum_{j=1}^{n} X_j/n & \text{u ulozi od } \bar{X}\,. \end{array}$$

Druga stvar koju treba uočiti je sličnost gornjeg rješenja i rješenja za normalna/normalna model, te, posebno, sličnost među formulama za faktor povjerenja, tj. (2.16) i (3.13). Formula (3.13) je direktna generalizacija formule (2.16), budući da se $\mathbb{E}[s^2(\theta)]$ i $\mathrm{Var}[m(\theta)]$ mogu smatrati generalizacijama redom od σ_1^2 i σ_2^2. Sličnosti između Modela 1 EBTP-a i normalna/normalna modela vodile su do sličnosti u rezultirajućim odgovorima.
Konačno, treba uočiti da formula za procjenu povjerenjem (3.14) uključuje tri parametra, $\mathbb{E}[m(\theta)]$, $\mathbb{E}[s^2(\theta)]$ i $\mathrm{Var}[m(\theta)]$, koji su tretirani kao da su poznati. EBTP ne pravi nikakve pretpostavke o distribuciji parametra rizika θ, različito od Bayesovskog pristupa povjerenju, ali je potrebno poznavati vrijednosti ta tri parametra. Činjenica da se ta tri parametra odnose na momente prvog i drugog reda, a ne na više momente ili neke druge veličine, je posljedica izbora vrijednosti $a_0, a_1, \ldots, a_n$ u (3.7) tako da minimiziraju očekivanu vrijednost kvadrata razlike. Način na koji se ti parametri mogu procijeniti razmatra se u sljedećoj sekciji.
Sada ćemo promatrati malo drugačiji scenarij.
Neka su $\alpha_0, \alpha_1, \ldots, \alpha_n$ realni brojevi i promatrajmo:

$$\mathbb{E}[(\mathbb{E}[m(\theta)|\underline{X}] - \alpha_0 - \sum_{j=1}^{n} \alpha_j X_j)^2] \tag{3.15}$$

(i) Pokazat će se da se (3.15) može zapisati kao

$$\mathbb{E}[(m(\theta) - \alpha_0 - \sum_{j=1}^{n} \alpha_j X_j)^2] - 2\mathbb{E}[AB] - \mathbb{E}[A^2]$$

gdje je

$$A = m(\theta) - \mathbb{E}[m(\theta)|\underline{X}]$$

i

$$B = \mathbb{E}[m(\theta)|\underline{X}] - \alpha_0 - \sum_{j=1}^{n} \alpha_j X_j\,.$$

(ii) Pokazat će se da je:

$$\mathbb{E}[AB] = 0$$

(iii) Nadalje će se pokazati da su vrijednosti od $\alpha_0, \alpha_1, \ldots, \alpha_n$ koje minimiziraju (3.15) jednake vrijednostima od $a_0, a_1, \ldots, a_n$ koje minimiziraju (3.8).

(i) $\mathbb{E}[(m(\theta)-\alpha_0-\sum_{j=1}^{n}\alpha_j X_j)^2]$ može se zapisati kao $\mathbb{E}[(A+B)^2]$ gdje su A i B kao u tvrdnji dijela (i). Kvadriranje binoma daje traženi odgovor.

(ii) Upotrebom formule (0.1):

$$\mathbb{E}[AB] = \mathbb{E}[\mathbb{E}[AB|\underline{X}]]$$

Uočimo sada da je B funkcija od $\underline{X}$, tako da je upotrebom formule (0.3):

$$\mathbb{E}[\mathbb{E}[AB|\underline{X}]] = \mathbb{E}[B\mathbb{E}[A|\underline{X}]]$$

Ali,

$$\begin{aligned}\mathbb{E}[A|\underline{X}] &= \mathbb{E}[m(\theta) - \mathbb{E}[m(\theta)|\underline{X}]|\underline{X}] \\ &= \mathbb{E}[m(\theta)|\underline{X}] - \mathbb{E}[m(\theta)|\underline{X}] \\ &= 0\,,\end{aligned}$$

budući da je, opet upotrebom (0.3),

$$\mathbb{E}[\mathbb{E}[m(\theta)|\underline{X}]|\underline{X}] = \mathbb{E}[m(\theta)|\underline{X}]$$

Stoga je

$$\mathbb{E}[AB] = \mathbb{E}[B \times 0] = 0$$

što se i tražilo.

(iii) Kao rezultat dijelova (i) i (ii), (3.15) može se zapisati kao:

$$\mathbb{E}[(m(\theta) - \alpha_0 - \sum_{j=1}^{n} \alpha_j X_j)^2] - \mathbb{E}[A^2]$$

Uočite sada da A, te stoga i $\mathbb{E}[A^2]$, nije funkcija od $\alpha_0, \alpha_1, \ldots, \alpha_n$.

To znači da su vrijednosti od $\alpha_0, \alpha_1, \ldots, \alpha_n$ koje minimiziraju (3.15) ujedno i vrijednosti koje minimiziraju:

$$\mathbb{E}[(m(\theta) - \alpha_0 - \sum_{j=1}^{n} \alpha_j X_j)^2]$$

Ali taj izraz je upravo (3.8) (s α umjesto a). To se upravo tražilo.

Zanimljivost gornjeg rezultata je da pokazuje da bi se isti odgovor dobio traženjem najbolje linearne procjene od $\mathbb{E}[m(\theta)|\underline{X}]$ umjesto najbolje linearne procjene od $m(\theta)$. To nije iznenađujuće, jer

(i) kroz ovu sekciju "najbolji" je značilo minimiziranje očekivanog kvadrata razlike, i

(ii) $\mathbb{E}[m(\theta)|\underline{X}]$ je najbolja procjena od $m(\theta)$ među *svim* funkcijama od $\underline{X}$, a (3.14) daje najbolju procjenu od $m(\theta)$ među svim linearnim funkcijama od $\underline{X}$.

3.4 Model 1: procjena parametara

U ovoj sekciji će se upotpuniti problem procjene od $m(\theta)$ uz dano $\underline{X}$ tako da se pokaže kako procijeniti $\mathbb{E}[m(\theta)]$, $\text{Var}[m(\theta)]$ i $\mathbb{E}[s^2(\theta)]$. Da bi se to napravilo, potrebne su daljnje pretpostavke: dostupni su podaci iz nekog drugog rizika koji je sličan, ali ne identičan, originalnom riziku. To zahtijeva postavljanje problema u malo većoj općenitosti, dodatne pretpostavke, i manju promjenu notacije. Važna razlika između čistog Bayesovskog pristupa povjerenju i EBTP-a je da za prvi nisu potrebni podaci za procjenu parametara, dok za drugi jesu.

Kao i u Sekciji 3.2 i Sekciji 3.3, i nadalje nas zanima procjena čiste premije, ili očekivanog broja odštetnih zahtjeva, za određeni rizik, te da je taj rizik jedan od N rizika u skupu (engl. collective). Pod skupom se misli na familiju različitih rizika koji su povezani na način koji će biti objašnjen kasnije u ovoj sekciji. Zbog jednostavnosti, pretpostavimo da je određen rizik koji nas zanima rizik broj 1 u tom skupu. Pretpostavlja se da su za svaki od tih N rizika opažane skupne štete, ili frekvencije šteta, za svaku od proteklih n godina. Neka X_{ij} označava skupne štete, ili broj šteta, za rizik broj i, $i = 1, 2, \ldots, N$, u godini j, $j = 1, 2, \ldots, n$.

Te vrijednosti sažete su u sljedećoj tablici:

Svaki red Tablice 1 predstavlja opažanja za različiti rizik; prvi red, koji se odnosi na rizik broj 1, je skup opaženih vrijednosti označen s $\underline{X}$ u Sekciji 3.2 i Sekciji 3.3. (Uočite da su $X_1, X_2, \ldots, X_n$ iz Sekcije 3.2 i Sekcije 3.3 sada postali $X_{11}, X_{12}, \ldots, X_{1n}$.)

U Sekciji 3.2 načinjene su dvije pretpostavke, (3.1) i (3.2), o vezi između opaženih vrijednosti za pojedinačan rizik koji se tada razmatrao. U ovoj ćemo sekciji načiniti iste pretpostavke za svaki od N rizika u skupu. Te su pretpostavke dane u (3.16) i (3.17) dolje.

Za svaki rizik i, $i = 1, 2, \ldots, N$:

(3.16) Distribucija svake X_{ij}, $j = 1, 2, \ldots, n$, ovisi o vrijednosti parametra, u oznaci θ_i, čija je vrijednost fiksna (i ista za svaki j), ali nepoznata.

(3.17) Uz dano θ_i, X_{ij}, $j = 1, 2, \ldots, n$, su nezavisne i jednako distribuirane.

Uočite da je parametar rizika, označen s θ u Sekciji 3.2 i Sekciji 3.3, sada označen s θ_i, te da je implikacija ove dvije pretpostavke da parametri rizika imaju različite vrijednosti za različite rizike. (Međutim, kao i u Sekciji 3.2 i Sekciji 3.3, parametar rizika za dani rizik ne mijenja vrijednost iz godine u godinu.) Te dvije pretpostavke pokazuju nešto o vezi unutar svakog reda Tablice 1, ali ne pokazuju ništa o vezi među različitim recima, tj. među različitim rizicima u skupu. Sljedeća pretpostavka pokazuje nešto o vezi među različitim rizicima u skupu.

(3.18) Za sve $i \neq k$, parovi (θ_i, X_{ij}) i (θ_k, X_{km}) su nezavisni i jednako distribuirani.

Ta pretpostavka pokazuje da su redovi Tablice 1 nezavisni jedan od drugoga. Dvije neposredne posljedice ove pretpostavke su:

(3.19) Za $i \neq k$, X_{ij} i X_{km} su nezavisne i jednako distribuirane.

(3.20) Parametri rizika $\theta_1, \theta_2, \ldots, \theta_N$ su nezavisni i jednako distribuirani.

Veza među različitim rizicima, tj. recima u tablici, je rezultat pretpostavke da su parametri rizika, $\theta_1, \theta_2, \ldots, \theta_N$, jednako distribuirani. Intuitivno, to znači da kada bi na neki način bile poznate vrijednosti od $\theta_2, \theta_3, \ldots, \theta_N$, tada bi bilo poznato nešto o distribuciji od θ_i, te bi stoga bilo poznato nešto o θ_1, ili, barem, o distribuciji iz koje dolazi.

Funkcije $m(\cdot)$ i $s^2(\cdot)$ uvedene su u Sekciji 3.2. Zadržat ćemo iste definicije i primijeniti ih na sve rizike u skupu:

$$\begin{aligned} m(\theta_i) &= \mathbb{E}[X_{ij}|\theta_i] \\ s^2(\theta_i) &= \mathrm{Var}[X_{ij}|\theta_i] \end{aligned}$$

Uočite da, isto kao u Sekciji 3.2, niti $\mathbb{E}[X_{ij}|\theta_i]$ niti $\mathrm{Var}[X_{ij}|\theta_i]$ ne ovise o j, budući da su, uz dano θ_j, slučajne varijable $X_{i1}, X_{i2}, \ldots, X_{in}$ jednako distribuirane. Uočite također da, budući da su $\theta_1, \theta_2, \ldots, \theta_N$ jednako distribuirane, $\mathbb{E}[m(\theta_i)]$, $\mathbb{E}[s^2(\theta_i)]$ i $\mathrm{Var}[m(\theta_i)]$ ne ovise o i. To su točno parametri

$\mathbb{E}[m(\theta)]$, $\mathbb{E}[s^2(\theta)]$ i $\mathrm{Var}[m(\theta)]$ iz Sekcije 3.2 i Sekcije 3.3, za procjenu kojih će se koristiti skup.
Potrebno je još oznaka. Označimo:

$$\sum_{j=1}^{n} X_{ij}/n \text{ sa } \; \bar{X}_i$$

i

$$\sum_{i=1}^{N} \bar{X}_i/N (= \sum_{i=1}^{N}\sum_{j=1}^{n} X_{ij}/(Nn)\,) \text{ sa } \; \bar{X}\,.$$

Uočite da što je sada označeno s $\bar{X}_i$, bilo u Sekciji 3.2 i Sekciji 3.3 označeno s $\bar{X}$. Važno je prepoznati razliku između značenja simbola $\bar{X}$ u ovoj sekciji i ranijim sekcijama.
Ova nova notacija koristit će se za reformulaciju procjene povjerenjem čiste premije, ili broja šteta, nastalih iz rizika 1 iz skupa, za nadolazeću godinu:

$$(1-Z)\mathbb{E}[m(\theta)] + Z\bar{X}_1$$

gdje je

$$\bar{X}_1 = \sum_{j=1}^{n} X_{1j}/n$$

i

$$Z = \frac{n}{n + \mathbb{E}[s^2(\theta)]/\mathrm{Var}[m(\theta)]}\,. \tag{3.21}$$

Važno je shvatiti da je formula (3.21) ista kao formula (3.14), ali da koristi notaciju $\bar{X}_1$ i X_{1j} umjesto $\bar{X}$ i X_j.
Sada se mogu dobiti procijenitelji za $\mathbb{E}[m(\theta)]$, $\mathbb{E}[s^2(\theta)]$ i $\mathrm{Var}[m(\theta)]$. Ti procijenitelji biti će funkcije od $\{\{X_{ij}\}_{j=1}^{n}\}_{i=1}^{N}$, čije će vrijednosti biti poznate kada se izračuna procjena povjerenjem od $m(\theta_1)$.
Svaki redak Tablice 1 odgovara fiksnoj vrijednosti od θ. Imajući to i definicije od $m(\theta_i)$ i $s^2(\theta_i)$ na umu, očigledni procjenitelji za $m(\theta_i)$ i $s^2(\theta_i)$ su redom:

$$\bar{X}_i \quad \text{i} \quad (n-1)^{-1}\sum_{j=1}^{n}(X_{ij}-\bar{X}_i)^2\,.$$

Nadalje, $\mathbb{E}[m(\theta)]$ je "prosjek" (po distribuciji od θ) vrijednosti od $m(\theta)$ za različite vrijednosti od θ. Očigledan procjenitelj za $\mathbb{E}[m(\theta)]$ je prosjek procjena od $m(\theta_i)$ za $i = 1, 2, \ldots, N$. Drugim riječima, procjenitelj za $\mathbb{E}[m(\theta)]$ je $\bar{X}$. Slično, $\mathbb{E}[s^2(\theta)]$ je "prosjek" vrijednosti od $s^2(\theta)$, te je stoga prosjek procjena od $s^2(\theta_i)$ očigledan procjenitelj, a to je:

$$N^{-1} \sum_{i=1}^{N} (n-1)^{-1} \sum_{j=1}^{n} (X_{ij} - \bar{X}_i)^2 \,.$$

Za svaki redak Tablice 1, $\bar{X}_i$ je procjena od $m(\theta_i)$, $i = 1, 2, \ldots, N$. Stoga se može pomisliti da bi opažena varijanca tih vrijednosti, tj.

$$(N-1)^{-1} \sum_{i=1}^{N} (\bar{X}_i - \bar{X})^2$$

bila očigledan procjenitelj za $\mathrm{Var}[m(\theta)]$. Na nesreću, može se pokazati da je to pristran procjenitelj od $\mathrm{Var}[m(\theta)]$. Može se pokazati da se nepristran procjenitelj od $\mathrm{Var}[m(\theta)]$ može dobiti oduzimanjem popravnog člana u gornjoj formuli. Nepristran procjenitelj za $\mathrm{Var}[m(\theta)]$ je:

$$(N-1)^{-1} \sum_{i=1}^{N} (\bar{X}_i - \bar{X})^2 - (Nn)^{-1} \sum_{i=1}^{N} (n-1)^{-1} \sum_{j=1}^{n} (X_{ij} - \bar{X}_i)^2 \,.$$

Popravni član je n^{-1} puta procjenitelj za $\mathbb{E}[s^2(\theta)]$.
Ti procjenitelji mogu se sažeti kako slijedi:

	Parametar	**Procjenitelj**
(3.22)	$\mathbb{E}[m(\theta)]$	$\bar{X}$
(3.23)	$\mathbb{E}[s^2(\theta)]$	$N^{-1} \sum_{i=1}^{N} (n-1)^{-1} \sum_{j=1}^{n} (X_{ij} - \bar{X}_i)^2$
(3.24)	$\mathrm{Var}[m(\theta)]$	$(N-1)^{-1} \sum_{i=1}^{N} (\bar{X}_i - \bar{X})^2$ $-(Nn)^{-1} \sum_{i=1}^{N} (n-1)^{-1} \sum_{j=1}^{n} (X_{ij} - \bar{X}_i)^2$

Te su formule dane u Tablicama.

Važno je uočiti da iako je $\mathrm{Var}[m(\theta)]$ nenegativan parametar, budući da je varijanca, procjenitelj dan s (3.24) može biti negativan. Formula (3.24) je razlika dva člana. Svaki od članova je nenegativan, ali njihova razlika ne mora biti. Ako u praksi (3.24) daje negativnu vrijednost, prihvaćena procedura je procijeniti $\mathrm{Var}[m(\theta)]$ s 0. Strogo govoreći, to znači da je procjenitelj za $\mathrm{Var}[m(\theta)]$ maksimum od 0 i vrijednosti dane s (3.24). Iako (3.24) daje nepristranu procjenu od $\mathrm{Var}[m(\theta)]$, uzimanje maksimuma od 0 i (3.24) ne daje nepristranu procjenu. Međutim, taj pragmatičan pristup izbjegava besmislenu procjenu za $\mathrm{Var}[m(\theta)]$.
Parametar $\mathbb{E}[s^2(\theta)]$ također mora biti nenegativan, ali će njegov procjenitelj, dan formulom (3.23), uvijek biti nenegativan, te stoga (3.23) ne zahtijeva prilagođivanje.
Može se pokazati da su procjenitelji za $\mathbb{E}[m(\theta)]$, $\mathbb{E}[s^2(\theta)]$ i $\mathrm{Var}[m(\theta)]$ nepristrani. Dokaz izlazi iz okvira ovog predmeta.
Promotrimo sada formulu (3.13) za faktor povjerenja Modela 1 EBTP-a. Različiti "sastojci" te formule imaju sljedeće definicije ili interpretacije:

n	je broj podataka u odnosu na rizik
$\mathbb{E}[s^2[\theta)]$	je prosječna varijabilnost podataka iz godine u godinu za pojedinačan rizik, tj. prosječna varijabilnost unutar redaka Tablice 1
$\mathrm{Var}[m(\theta)]$	je varijabilnost prosječnih vrijednosti podataka za različite rizike, tj. varijabilnost očekivanja redaka u Tablici 1.

Gledajući formulu (3.13), može se primjetiti sljedeće:

(i) Z je uvijek između nula i jedan.

(ii) Z je rastuća funkcija od n. To je za očekivati - čim je više podataka iz samog rizika, tim ćemo se više pouzdati u procjenu povjerenjem čiste premije ili broja šteta.

(iii) Z je padajuća funkcija od $\mathbb{E}[s^2(\theta)]$. To je za očekivati - čim je veća vrijednost od $\mathbb{E}[s^2(\theta)]$, tim su varijabilniji, te stoga manje pouzdani, podaci iz samog rizika u odnosu na podatke iz drugih rizika u skupu.

(iv) Z je rastuća funkcija od $\mathrm{Var}[m(\theta)]$. To je za očekivati - čim je veća vrijednost od $\mathrm{Var}[m(\theta)]$, u odnosu na $\mathbb{E}[s^2(\theta)]$, tim je veća varijabilnost među različitim rizicima u skupu, te je stoga manje vjerojatno da će drugi rizici u skupu sličiti riziku koji nas zanima, i manje se treba osloniti na podatke iz drugih rizika.

Čini se da postoji kontradikcija u ovoj sekciji. U Sekciji 1.2 tvrdilo se da faktor povjerenja ne smije ovisiti o podacima iz rizika za koji se određuje premija. Međutim, ti podaci su se koristili za procjenu od $\mathrm{Var}[m(\theta)]$ i $\mathbb{E}[s^2(\theta)]$, čije vrijednosti se tada koriste za računanje vrijednosti od Z. Objašnjenje je da u principu faktor povjerenja Z, dan formulom (3.21), ne ovisi o stvarnim podacima iz rizika za koji se određuje premija. Na nesreću, formula za Z uključuje dva parametra, $\mathrm{Var}[m(\theta)]$ i $\mathbb{E}[s^2(\theta)]$, čije su vrijednosti nepoznate, ali se, u praksi, mogu procijeniti iz podataka iz samog rizika i drugih rizika u skupu.

Komentirajmo, nadalje, još malo pretpostavke modela. Dana je pretpostavka koja se odnosi na jednaku distribuiranost X_{ij} i iz različitih rizika i iz istog rizika. Ako X_{ij} dolaze iz različitih rizika, pretpostavljeno je da su oni (bezuvjetno) jednako distribuirani (vidi (3.19)). Ako X_{ij} dolaze iz istog rizika, i, pretpostavljeno je da su oni bezuvjetno jednako distribuirani (vidi (3.3)) i uvjetno, uz dano θ_i, jednako distribuirani (vidi (3.2) i (3.17)). Posljedica tih pretpostavki je da niti $\mathbb{E}[X_{ij}|\theta_i]$ niti $\mathrm{Var}[X_{ij}|\theta_i]$ ne ovisi o j. Ta posljedica bila je jedini korak u izvodu procjene povjerenjem (3.21) koji je zahtijevao pretpostavke o jednakoj distribuciji od X_{ij}. U stvari, pretpostavke (3.17) i (3.18) (i odgovarajuća pretpostavka iz Sekcije 3.2, (3.2)) mogle su biti zamijenjene slijedećim pretpostavkama i još uvijek bi bilo moguće izvesti istu procjenu povjerenjem.

(3.25) Uz dano θ_i, X_{ij}, $j = 1, 2, \ldots, n$, su nezavisni za svaki rizik i, $i = 1, 2, \ldots, N$.

(3.26) Za $i \neq k$, parovi (θ_i, X_{ij}) i (θ_k, X_{kl}) su nezavisni i parametri rizika $\theta_1, \theta_2, \ldots, \theta_N$, su također jednako distribuirani

(3.27) Za svaki rizik, i, $i = 1, 2, \ldots, N$, niti $\mathbb{E}[X_{ij}|\theta_i]$ niti $\mathrm{Var}[X_{ij}|\theta_i]$ ne ovise o j.

Pretpostavke (3.25) i (3.26) oslabljene su verzije od (3.17) i (3.18). Pretpostavka (3.27) se sada mora uključiti kao posebna pretpostavka, budući da nije posljedica od (3.25) i (3.26).

Nijedan od rezultata ili formula u Sekciji 3 ne bi bio promijenjen uz te slabije pretpostavke. Razlog zbog kojeg su načinjene jače pretpostavke, kao što je učinjeno u Sekciji 3.2 i Sekciji 3.3, je taj da se prikaz načini malo lakšim.

Slabije pretpostavke su pokazane zbog razloga da to pomaže uspostavljanju veze između Modela 1 EBTP-a i Modela 2 EBTP-a u Sekciji 4.

4 Empirijska Bayesova teorija povjerenja: Model 2

4.1 Uvod

U ovoj sekciji će se tehnike empirijskog Bayesovog povjerenja primijeniti na drugi, ponešto kompliciraniji model. Format će biti potpuno isti kao u Sekciji 3. Prvo, u Sekciji 4, postavit će se problem i izložiti pretpostavke. Problem će biti isti kao u Sekciji 3, tj. procijeniti čistu premiju, ili očekivani broj šteta, nastalih po riziku u nadolazećoj godini. Pretpostavke će biti malo drugačije od onih u Sekciji 3. Nadalje, u Sekciji 4.3, izvest će se procjena povjerenjem čiste premije ili očekivanog broja šteta. Konačno, u Sekciji 4.4, diskutirat će se metoda procjene vrijednosti parametara koji su dio procjene povjerenjem. Držeći se terminologije iz Sekcije 3, o modelu iz ove sekcije govorit će se kao o Modelu 2 EBTP-a.

4.2 Model 2: specifikacija

Problem je procijeniti očekivane skupne štete, ili očekivani broj šteta, nastalih po riziku u nadolazećoj godini. Neka slučajne varijable $Y_1, Y_2, \ldots$ predstavljaju skupne štete ili broj šteta nastalih po tom riziku u uzastopnim godinama. Pretpostavlja se da su vrijednosti od $Y_1, Y_2, \ldots, Y_n$ već opažene i potrebno je procijeniti očekivanu vrijednost od Y_{n+1}. Za sada problem izgleda potpuno isto kao i problem u Sekciji 3. Važna razlika između Modela 1 EBTP-a i Modela 2 EBTP-a je ta da Model 2 uključuje jedan dodatni parametar poznat kao volumen rizika, P_j. Intuitivno, vrijednost od P_j mjeri "količinu posla" u godini j. Na primjer, P_j može predstavljati dohodak od premije za rizik u godini j, ili broj polica koje tvore rizik u godini j. Važno je uočiti da se pretpostavlja da je vrijednost od P_{n+1} poznata na početku godine $n+1$. Nadalje, novi niz slučajnih varijabli, $X_1, X_2, \ldots$, definira se kako slijedi:

$$X_j = Y_j / P_j \qquad j = 1, 2, \ldots .$$

Slučajne vrijable X_j predstavljaju skupne štete, ili broj šteta, u godini j, standardiziranih tako da se ukloni efekt različitih razina poslova u različitim godinama. Pretpostavke koje specificiraju Model 2 su sljedeće:

(4.1) Distribucija svakog X_j ovisi o vrijednosti parametra, θ, čija je vrijednost ista za svaki j, ali nepoznata.

(4.2) Uz dano θ, X_j su nezavisne (ali ne nužno jednako distribuirane).

(4.3) $\mathbb{E}[X_j|\theta]$ ne ovisi o j.

(4.4) $P_j \mathrm{Var}[X_j|\theta]$ ne ovisi o j.

Kao i u prethodnoj sekciji, θ se zove parametar rizika za rizik, i, isto kao za Model 1, može biti realan broj ili općenitija veličina kao što je vektor realnih brojeva. Pretpostavka (4.1) je standardna pretpostavka za sve modele povjerenja koji se ovdje razmatraju. Vidi pretpostavke (2.2), (2.9) i (3.1). Petpostavka (4.2) odgovara pretpostavci (3.2) Modela 1 EBTP-a, ali uočite da je (4.2) *malo slabija* od (3.2). Pretpostavka (4.2) ne zahtijeva da X_j budu uvjetno (uz dano θ) jednako distribuirani, već samo uvjetno nezavisni. U Modelu 2 ne postoji pretpostavka da su X_j bezuvjetno, ili uvjetno uz dano θ, jednako distribuirani.

Uspoređujući gornje pretpostavke s pretpostavkama (3.25), (3.26) i (3.27) Modela 1 EBTP-a, vidimo da ako su svi P_j jednaki 1, tada je Model 2 EBTP-a točno isti kao Model 1 EBTP-a.

Pretpostaviši (4.3) i (4.4), možemo definirati $m(\theta)$ i $s^2(\theta)$ kako slijedi:

$$\begin{aligned} m(\theta) &= \mathbb{E}[X_j|\theta] \\ s^2(\theta) &= P_j \mathrm{Var}[X_j|\theta]\,. \end{aligned}$$

Definicija od $m(\theta)$ točno odgovara definiciji za Model 1 EBTP-a u Sekciji 3, ali je definicija od $s^2(\theta)$ malo drugačija.

Da bismo dobili bolji uvid u pretpostavke (4.3) i (4.4), promatrajmo sljedeći primjer. Pretpostavimo da se rizik koji razmatramo svake godine sastoji od različitog broja nezavisnih polica, te da je broj polica u godini j jednak P_j. Pretpostavimo da skupne štete u jednoj godini nastale po pojedinačnoj polici imaju očekivanje $m(\theta)$ i varijancu $s^2(\theta)$, gdje su $m(\cdot)$ i $s^2(\cdot)$ funkcije od θ, i θ je fiksan, ali nepoznat, parametar rizika za sve te police. Neka Y_j označava skupne štete nastale po svim policama na snazi u godini j. Tada je:

$$\begin{aligned} \mathbb{E}[Y_j] &= P_j m(\theta) \\ \mathrm{Var}[Y_j] &= P_j s^2(\theta) \\ \mathbb{E}[X_j] &= m(\theta) \\ P_j \mathrm{Var}[X_j] &= s^2(\theta) \end{aligned}$$

Taj primjer zadovoljava pretpostavke (4.3) i (4.4).

4.3 Model 2: izvod premije povjerenjem

U protekloj sekciji je problem postavljen prilično neodređeno kao procjena očekivane vrijednosti od Y_{n+1}, uz dane vrijednosti od $Y_1, Y_2, \ldots, Y_n$. Budimo sada malo precizniji. Veličina koju treba procijeniti je očekivanje (uz dano θ) od Y_{n+1}. To je dano s $P_{n+1}m(\theta)$. Budući da je u modelu P_{n+1} poznato na početku godine $n+1$, problem je procijeniti $m(\theta)$. Dostupni podaci su vrijednosti svakog Y_j i odgovarajućeg P_j za $j = 1, 2, \ldots, n$. Zbog pogodnosti, ti će podaci biti označeni s $\underline{X}$. Kao i u Sekciji 3, traži se procjenitelj koji je linearan u opaženim podacima, i, opet kao u Sekciji 3, kao procjenitelj od $m(\theta)$ biti će odabrana linearna funkcija od $X_1, X_2, \ldots, X_n$ koja je najbolja procjena od $m(\theta)$ u smislu najmanje očekivane razlike kvadrata. (Uočite ovdje da je linearna funkcija od $X_1, X_2, \ldots, X_n$ također linearna funkcija od $Y_1, Y_2, \ldots, Y_n$, budući da je svaki Y_j konstantan faktor odgovarajućeg X_j.) Stavivši sve zajedno, problem je procijeniti $m(\theta)$ funkcijom oblika

$$a_0 + a_1X_1 + a_2X_2 + \cdots + a_nX_n$$

gdje su konstante $a_0, a_1, \ldots, a_n$ odabrane tako da minimiziraju:

$$\mathbb{E}[(m(\theta) - a_0 - a_1X_1 - a_2X_2 - \cdots a_nX_n)^2] \tag{4.5}$$

Taj problem ima točno isti oblik kao i problem u Sekciji 3, iako su definicija od X_j i neke od pretpostavki različite. Metoda rješavanja problema biti će ista kao u Sekciji 3: derivira se (4.5) po $a_0, a_1, \ldots, a_n$ i svaka od derivacija izjednači s nulom.

Evo nekih tehničkih rezultata koji će biti od pomoći u rješavanju tog problema.

(i) Upotrebom (0.1) i (0.3) i definicije od $m(\theta)$:

$$\begin{aligned}
\mathbb{E}[X_jm(\theta)] &= \mathbb{E}[\mathbb{E}[X_jm(\theta)|\theta]] \\
&= \mathbb{E}[m(\theta)\mathbb{E}[X_j|\theta]] \\
&= \mathbb{E}[m(\theta)^2] \\
&= \mathrm{Var}[m(\theta)] + (\mathbb{E}[m(\theta)])^2
\end{aligned}$$

(ii) Upotrebom (0.1) i činjenice da su, uz dano θ, X_j i X_k nezavisne za $j \neq k$:

$$\begin{aligned}
\mathbb{E}[X_j X_k] &= \mathbb{E}[\mathbb{E}[X_j X_k|\theta]] \\
&= \mathbb{E}[\mathbb{E}[X_j|\theta]\mathbb{E}[X_k|\theta]] \\
&= \mathbb{E}[m(\theta)^2] \\
&= \mathrm{Var}[m(\theta)] + (\mathbb{E}[m(\theta)])^2
\end{aligned}$$

(iii) Upotrebom (0.1) i definicije od $s^2(\theta)$:

$$\begin{aligned}
\mathbb{E}[X_j^2] &= \mathbb{E}[\mathbb{E}[X_j^2|\theta]] \\
&= \mathbb{E}[\mathrm{Var}[X_j|\theta] + (\mathbb{E}[X_j|\theta])^2] \\
&= \mathbb{E}[s^2(\theta)]/P_j + \mathbb{E}[m(\theta)^2] \\
&= \mathbb{E}[s^2(\theta)]/P_j + \mathrm{Var}[m(\theta)] + (\mathbb{E}[m(\theta)])^2
\end{aligned}$$

Sada:

$$\frac{\partial}{\partial a_0}(\mathbb{E}[(m(\theta) - a_0 - a_1X_1 - a_2X_2 - \cdots - a_nX_n)^2]) = 0$$

tj.

$$\mathbb{E}[m(\theta) - a_0 - a_1X_1 - a_2X_2 - \cdots - a_nX_n] = 0$$

što, po definiciji od $m(\theta)$, daje:

$$a_0 = \mathbb{E}[m(\theta)](1 - \sum_{j=1}^{n} a_j)\,. \tag{4.6}$$

Deriviranjem po a_k, $k = 1, 2, \ldots, n$:

$$\begin{aligned}
&\frac{\partial}{\partial a_k}(\mathbb{E}[(m(\theta) - a_0 - a_1X_1 - a_2X_2 - \cdots - a_nX_n)^2]) = 0 \quad \rightarrow \\
&\mathbb{E}[X_k(m(\theta) - a_0 - a_1X_1 - a_2X_2 - \cdots - a_nX_n)] = 0\,.
\end{aligned}$$

Nakon sređivanja to daje:

$$\mathrm{Var}[m(\theta)]+(\mathbb{E}[m(\theta)])^2-a_0\mathbb{E}[m(\theta)]-\sum_{j=1}^{n} a_j(\mathrm{Var}[m(\theta)]+(\mathbb{E}[m(\theta)])^2)-a_k\mathbb{E}[s^2(\theta)]/P_k = 0$$

što, nakon daljnjeg sređivanja i upotrebe (4.6) daje:

$$a_k = P_k \mathrm{Var}[m(\theta)](1 - \sum_{j=1}^{n} a_j)/\mathbb{E}[s^2(\theta)]\,. \tag{4.7}$$

Zbrajanjem obiju strana u (4.7) od $k = 1$ do $k = n$ daje:

$$\sum_{k=1}^{n} a_k = (\sum_{j=1}^{n} P_j)(1 - \sum_{j=1}^{n} a_j)\mathrm{Var}[m(\theta)]/\mathbb{E}[s^2(\theta)]\,.$$

Rješavanje te jednadžbe po $\sum_{j=1}^{n} a_j$ i uvrštavanjem rješenja u (4.6) i (4.7) daje:

$$a_0 = \frac{\mathbb{E}[m(\theta)]\mathbb{E}[s^2(\theta)]/\mathrm{Var}[m(\theta)]}{\sum_{j=1}^{n} P_j + \mathbb{E}[s^2(\theta)]/\mathrm{Var}[m(\theta)]} \tag{4.8}$$

$$a_k = \frac{P_k}{\sum_{j=1}^{n} P_j + \mathbb{E}[s^2(\theta)]/\mathrm{Var}[m(\theta)]}\,. \tag{4.9}$$

Stavljanje vrijednosti za a_0 i a_k, $k = 1, 2, \ldots, n$ u izraz:

$$a_0 + a_1 X_1 + a_2 X_2 + \cdots + a_n X_n$$

pokazuje da je rješenje problema, tj. "najbolja" linearna procjena od $m(\theta)$ uz dano $\underline{X}$, dana s

$$\frac{\mathbb{E}[m(\theta)]\mathbb{E}[s^2(\theta)]/\mathrm{Var}[m(\theta)] + \sum_{j=1}^{n} Y_j}{\sum_{j=1}^{n} P_j + \mathbb{E}[s^2(\theta)]/\mathrm{Var}[m(\theta)]}$$

To se može atraktivnije zapisati kao:

$$Z\bar{X} + (1 - Z)\mathbb{E}[m(\theta)] \tag{4.10}$$

gdje je

$$\bar{X} = \sum_{j=1}^{n} P_j X_j \,/\, \sum_{j=1}^{n} P_j$$

i

$$Z = \frac{\sum_{j=1}^{n} P_j}{\sum_{j=1}^{n} P_j + \mathbb{E}[s^2(\theta)]/\mathrm{Var}[m(\theta)]}\,.$$

To pokazuje sličnosti i razlike sa rezultatima Modela 1.
Evo nekoliko napomena o rješenju problema.

(i) Vrijednosti od a_k nisu nužno jednake za $k = 1, 2, \ldots, n$ (vidi formulu (4.9)). To znači da rješenje za Model 2 EBTP-a nije simetrično po X_k kao što je bilo za Model 1 EBTP-a (vidi napomene nakon formule (3.10)). To nije iznenađujuće, budući da podaci za Model 2 EBTP-a nisu simetrični - različite godine imaju različite volumene rizika. Kada bi svi P_k bili jednaki, tako da podaci budu simetrični, tada bi, kao što se može vidjeti iz (4.9), svi a_k bili jednaki za $k = 1, 2, \ldots, n$. Međutim, kada bi svi P_k bili jednaki, Model 2 EBTP-a bi, u stvari, bio jednak Modelu 1 EBTP-a.

(ii) Kada bi svi P_k bili jednaki jedan, rješenje dano s (4.10) bilo bi potpuno isto kao i rješenje dano s (3.14). To tako i treba biti, budući da ako su svi P_k jednaki 1, tada je Model 2 EBTP-a potupno jednak Modelu 1 EBTP-a.

(iii) Kao i kod Modela 1 EBTP-a, rješenje dano s (4.10) uključuje tri parametra, $\mathbb{E}[m(\theta)]$, $\mathbb{E}[s^2(\theta)]$ i $\mathrm{Var}[m(\theta)]$. Način na koji se ti parametri mogu procijeniti biti će objašnjen u sljedećoj sekciji.

4.4 Model 2: procjena parametara

Procedura za procjenu parametara $\mathbb{E}[m(\theta)]$, $\mathbb{E}[s^2(\theta)]$ i $\mathrm{Var}[m(\theta)]$ za Model 2 EBTP-a slijedi točno iste korake kao procedura za Model 1 EBTP-a koju smo proučavali u Sekciji 3.4.

Sada se pretpostavlja da je rizik koji nas zanima jedan iz skupa od N rizika, te da postoje podaci u obliku danom u Sekciji 4.2 za svaki od tih N rizika za svaku od proteklih n godina. Ti podaci sastoje se od vrijednosti za skupne štete, ili broj šteta, te odgovarajući volumen rizika. Neka je Y_{ij} slučajna varijabla koja označava skupne štete, ili broj šteta, za rizik broj i u godini j, $j = 1, 2, \ldots, n$, $i = 1, 2, \ldots, N$, te neka P_{ij} označava odgovarajući volumen rizika.

Za sve i i j definiramo:

$$X_{ij} = Y_{ij}/P_{ij}$$

Podaci su sažeti u sljedećoj tablici, koja odgovara Tablici 1 Modela 1 EBTP-a: Zbog jednostavnosti je pretpostavljeno, kao što je učinjeno u Sekciji 3., da je rizik koji nas zanima rizik broj 1 u tom skupu. To znači da što je bilo označeno s Y_j, P_j i X_j u Sekciji 4.2 i Sekciji 4.3, sada je označeno redom s Y_{1j}, P_{1j} i X_{1j} u ovoj sekciji. Problem je procijeniti očekivanu vrijednost od $X_{1,n+1}$ i rješenje problema već je dano formulom (4.10), uz prisjećanje da su X_j i P_j u (4.10) sada označeni s X_{1j} i P_{1j}. Svrha podataka iz drugih rizika u skupu je samo pomoć pri procjeni parametara $\mathbb{E}[m(\theta)]$, $\mathbb{E}[s^2(\theta)]$ i $\mathrm{Var}[m(\theta)]$ koji se pojavljuju u (4.10). Drugi rizici u skupu zadovoljavaju iste pretpostavke (4.1), (4.2), (4.3) i (4.4) kao i rizik broj 1. Te pretpostavke su kako slijedi:

(4.11) Distribucija svakog X_{ij} ovisi o vrijednosti parametra, θ_i, čija je vrijednost ista za svaki j, ali nepoznata.

(4.12) Uz dano θ_i, X_{ij} su nezavisne (ali ne nužno jednako distribuirane).

(4.13) Postoji funkcija $m(\cdot)$ takva da je $m(\theta_i) = \mathbb{E}[X_{ij}|\theta_i]$.

(4.14) Postoji funkcija $s^2(\cdot)$ takva da je $s^2(\theta_i) = P_{ij}\mathrm{Var}[X_{ij}|\theta_i]$.

Te četiri pretpostavke pokazuju da svaki rizik u skupu zadovoljava iste pretpostavke kao i određeni rizik koji nas zanima. Sljedeće dvije pretpostavke, koje odgovaraju (3.18), (3.19) i (3.20) u Sekciji 3.4, pokazuju vezu *između* različitih rizika u skupu.

(4.15) Parametri rizika, $\theta_1, \theta_2, \ldots, \theta_N$, promatrani kao slučajne varijable, su nezavisni i jednako distribuirani.

(4.16) Za $i \neq k$, su parovi (θ_i, X_{ij}) i (θ_k, X_{km}) nezavisni.

Uočite da, budući da su θ_i jednako distribuirani, vrijednosti od $\mathbb{E}[m(\theta_i)]$, $\mathbb{E}[s^2(\theta_i)]$ i $\mathrm{Var}[m(\theta_1)]$ ne ovise o i, pa stoga mogu biti označeni redom sa $\mathbb{E}[m(\theta)]$, $\mathbb{E}[s^2(\theta)]$ i $\mathrm{Var}[m(\theta)]$, kao i u Sekciji 4.2 i Sekciji 4.3.

Kao i u Sekciji 3.4, potrebno je još oznaka. Označimo:

$$\sum_{j=1}^{n} P_{ij} \quad \text{sa} \quad \bar{P}_i \tag{4.17}$$

$$\sum_{i=1}^{N} \bar{P}_i \quad \text{sa} \quad \bar{P} \tag{4.18}$$

$$(Nn-1)^{-1} \sum_{i=1}^{N} \bar{P}_i(1 - \bar{P}_i/\bar{P}) \quad \text{sa} \quad P^* \tag{4.19}$$

$$\sum_{j=1}^{n} P_{ij} X_{ij}/\bar{P}_i \quad \text{sa} \quad \bar{X}_i \tag{4.20}$$

$$\sum_{i=1}^{N} \bar{P}_i \bar{X}_i/\bar{P} = \sum_{i=1}^{N} \sum_{j=1}^{n} P_{ij} X_{ij}/\bar{P} \quad \text{sa} \quad \bar{X}\,. \tag{4.21}$$

Te oznake su sve dane u Tablicama. Uočite da što je bilo označeno s $\bar{X}$ u (4.10), sada je oznčeno s $\bar{X}_1$, a da $\bar{X}$ sada ima drugu definiciju. (To je točno što se dogodilo u Sekciji 3.4.) Uočite, također, da su $\bar{X}_i$ i $\bar{X}$ težinske sredine od X_{ij}, gdje su težine volumeni rizika P_{ij}.
Uz tu novu notaciju, procjena povjerenjem čiste premije, ili broja šteta, u nadolazećoj godini za rizik broj 1 u skupu originalno dana formulom (4.10) može se reformulirati kao

$$Z\bar{X}_1 + (1-Z)\mathbb{E}[m(\theta)] \tag{4.22}$$

gdje je

$$\bar{X}_1 = \sum_{j=1}^{n} P_{1j} X_{1j} \,/\, \sum_{j=1}^{n} P_{1j}$$

i

$$Z = \frac{\sum_{j=1}^{n} P_{1j}}{\sum_{j=1}^{n} P_{1j} + \mathbb{E}[s^2(\theta)]/\mathrm{Var}[m(\theta)]}\,.$$

Važno je uvidjeti da je to isto kao i formula (4.10), ali je zapisano korištenjem notacije iz ove sekcije umjesto notacije prethodne sekcije.
Mogu se predložiti nepristrani procjenitelji za $\mathbb{E}[m(\theta)]$, $\mathbb{E}[s^2(\theta)]$ i $\mathrm{Var}[m(\theta)]$, zasnovani na opaženim vrijednostima od $\{\{Y_{ij}, P_{ij}\}_{j=1}^{n}\}_{i=1}^{N}$.

Ti procjenitelji su:

	Parametar	**Procjenitelj**
(4.23)	$\mathbb{E}[m(\theta)]$	$\bar{X}$
(4.24)	$\mathbb{E}[s^2(\theta)]$	$N^{-1}\sum_{i=1}^{N}(n-1)^{-1}\sum_{j=1}^{n}P_{ij}(X_{ij}-\bar{X}_i)^2$
(4.25)	$\mathrm{Var}[m(\theta)]$	$P^{*^{-1}}\left\{(Nn-1)^{-1}\sum_{i=1}^{N}\sum_{j=1}^{n}P_{ij}(X_{ij}-\bar{X})^2 - N^{-1}\sum_{i=1}^{N}(n-1)^{-1}\sum_{j=1}^{n}P_{ij}(X_{ij}-\bar{X}_i)^2\right\}$

Uočite sljedeće o tim procjeniteljima:

(i) Oni imaju isti oblik kao i procjenitelji za Model 1 EBTP-a. Pogledajte ponovno formule (3.22), (3.23) i (3.24). Posebno, ako su svi P_{ij} jednaki 1, tada su dva skupa procjenitelja identična.

(ii) U praksi se može dogoditi da formula (4.25) daje negativnu vrijednost čak iako $\mathrm{Var}[m(\theta)]$ mora biti nenegativna. U takvoj se situaciji $\mathrm{Var}[m(\theta)]$ procjeni nulom. Slična stvar o procjeni varijance u Modelu 1 EBTP-a razmatrala se u Sekciji 3.4.

(iii) Formule (4.23), (4.24) i (4.25) su sve dane u Tablicama.

(iv) Dokazi da su (4.23), (4.24) i (4.25) nepristrani izlaze iz okvira ovog predmeta.

		Godina			
		1	2	$\cdots$	n
Broj	1	X_{11}	X_{12}	$\cdots$	X_{1n}
rizika	2	X_{21}	X_{22}	$\cdots$	X_{2n}
	.	.	.	$\cdots$	.
	.	.	.	$\cdots$	.
	N	X_{N1}	X_{N2}	$\cdots$	X_{Nn}

Tablica 1:

		Godina			
		1	2	$\cdots$	n
Broj	1	Y_{11}, P_{11}	Y_{12}, P_{12}	$\cdots$	Y_{1n}, P_{1n}
rizika	2	Y_{21}, P_{21}	Y_{22}, P_{22}	$\cdots$	Y_{2n}, P_{2n}
	.	.	.	$\cdots$	.
	.	.	.	$\cdots$	.
	N	Y_{N1}, P_{N1}	Y_{N2}, P_{N2}	$\cdots$	Y_{Nn}, P_{Nn}

Tablica 2:

POGLAVLJE 7 - JEDNOSTAVAN SUSTAV ISKUSTVENOG UTVRĐIVANJA PREMIJA

Nastavni ciljevi:

(viii)1. Objasniti rad jednostavnog sustava iskustvenog utvrđivanja premija zasnovanog na frekvenciji odštetnih zahtjeva i izračunati stacionarnu distribuciju broja osiguranika na svakom nivou popusta.

(viii)2. Izračunati vjerojatnost (koristeći jednostavne kriterije) da će osiguranik zahtijevati odštetu.

(viii)3. Ponoviti (viii)1, primjenivši rezultate iz (viii)2.

0 Uvod

Prilikom određivanja premije koju osiguranik treba platiti, mnoga osigurateljna društva koriste informaciju o broju odštetnih zahtjeva osiguranika u proteklim godinama. To se čini na osnovama da to daje bolju indikaciju vjerojatnosti da će osiguranik u budućnosti imati odštetni zahtjev.
To je naročito često u slučaju osiguranja motornih vozila, gdje većina, ako ne i sva osigurateljna društva koriste sustav bonusa (engl. No Claims Discount system - NCD system). (Neki osiguratelji također koriste taj sustava za druge tipove osiguranja koa što su kućanstvo, grupno osiguranje života i zdravstveno pokriće (engl. medical cover)).
Sustav bonusa funkcionira tako da daje osiguraniku popust na normalnu premiju koji je neposredno povezan s brojem “godina bez odštetnog zahtjeva” koje je osiguranik imao.
Kod odluke da li prijaviti štetu, osiguranik treba tada razmatrati posljedice na premiju sljedeće godine. Jedan od razloga za uvođenje sustava bonusa je, stoga, obeshrabrivanje malih odštetnih zahtjeva. Osiguranik neće prijaviti štetu, ako je manja od idućeg povećanja premije. Dakle, sustav bonusa bi trebao smanjiti broj malih odštetnih zahtjeva prijavljenih osigurateljnom društvu. To će smanjiti troškove odštetnih zatjeva i poravnati smanjenje prihoda od premija. Još važnije, to će smanjiti troškove obrada odštetnih zahtjeva. Čim je manje zahtjeva, tim se manji režijski troškovi mogu naplatiti

osiguraniku. Reduciranjem broja malih odštetnih zahtjeva, društvo smanjuje broj zahtjeva čija obrada stoji nerazmjerno veliki iznos, u odnosu na troškove štete. To će učiniti premije osigurateljnog društva kompetitivnijima.

1 Definicija sustava bonusa

1.1 Kategorije popusta

Postoje dva dijela sustava bonusa: kategorije popusta i pravila za prelaženje iz jedne kategorije u drugu. Dodatno, da bi se ispitivala svojstva sustava bonusa potrebno je poznavati vjerojatnost za svaku godinu da će osiguranik prijaviti štetu.
O kategorijama se često govori kao o broju "godina bez odštetnog zahtjeva". Međutim, pravila za kretanje među kategorijama su obično takva da se u stvari ne odnose na broj godina od odštetnog zahtjeva. Umjesto da zahtjev rezultira vraćanjem u kategoriju bez popusta, uobičajeno je da se osiguranik jednostavno pomakne u kategoriju na nižem nivou popusta.

1.1.1 Primjer

Promotrimo sustav bonusa s tri kategorije:

kategorija	popust %
0	0
1	25
2	40

U kategoriji 0 osiguranik plaća punu premiju koja će u praksi biti različita za pojedince ovisno o njihovim osobnim prilikama (kao npr. dob) koje se koriste kao faktor određivanja premija. Zbog jednostavnosti ćemo promatrati skup osiguranika koji se mogu smatrati jednakim na osnovu faktora za određivanje premije. U tom slučaju, puna premija biti će ista za sve police u portfelju.
U kategoriji 1, osiguranik plaća samo 75% pune premije, a u kategoriji 2 samo 60% pune premije.
Ako osiguranik ne prijavi štetu tokom godine, on ili ona prelazi u sljedeću kategoriju većeg popusta. Ako se prijavi jedna ili više šteta, on ili ona silazi jednu kategoriju niže (ili ostaje na nula popusta).

Uočite: u praksi može biti pet ili šest kategorija i prijava štete može rezultirati silaskom za više od jedne kategorije.

1.1.2 Prijelazna matrica

Ovaj sustav je lakše analizirati ako se zapiše u matematičkijem obliku. Očekivani omjer osiguranika u kategoriji i označen je s π_i. Uočite da je $\sum \pi_i = 1$. Također, omjeri u kategorijama popusta mogu se predstaviti vektorom $\tilde{\pi} = (\pi_0, \pi_1, \pi_2, \ldots, \pi_n)$.
Sada se može zapisati vjerojatnost prijelaza osiguranika iz kategorije i u kategoriju j iz jedne u drugu godinu.
Najjednostavniji način da se to predstavi je prijelazna matrica

$$P = \begin{bmatrix} p_{00} & p_{01} & p_{02} & \cdots \\ p_{10} & p_{11} & p_{12} & \cdots \\ p_{20} & p_{21} & p_{22} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}.$$

p_{ij} je vjerojatnost da osiguranik prijeđe iz kategorije i u kategoriju j.

1.1.3 Distribucija osiguranika

Prijelazna matrica može se koristiti za procjenu očekivanog broja osiguranika u svakoj kategoriji popusta za svaku godinu. Pretpostavimo da svi osiguranici počinju iz kategorije 0 u godini 1. U godini 1, $\pi_0 = 1$ i omjeri u svakoj kategoriji dani su vektorom koji bi za tri kategorije popusta bio

$$\tilde{\pi}^{(1)} = (1, 0, 0).$$

U godini 2, očekivani omjeri u svakoj kategoriji dani su vektorom koji se dobije množenjem $\tilde{\pi}^{(1)}$ s P:

$$\tilde{\pi}^{(2)} = \tilde{\pi}^{(1)} P$$

Slično $\tilde{\pi}^{(n+1)} = \tilde{\pi}^{(n)} P$.

2 Analiza stacionarnosti

2.1 Stacionarna distribucija

Moguće je nastaviti računanje $\pi^{(n)}$- ova za velike vrijednosti od n. Uz razumne pretpostavke će $\pi^{(n)}$ težiti prema graničnoj vrijednosti kada $n \to \infty$.

Kada se to dogodi, sustav je dostigao ekvilibrij ili stacionarnost. Ta granična vrijednost se označava s $\tilde{\pi}$.
Puštanje $n \to \infty$ daje $\tilde{\pi} = \tilde{\pi}P$.
Primjetivši da je $\sum \pi_i = 1$, taj skup jednadžbi može se riješiti i pronaći $\tilde{\pi}$.

2.1.1 Primjer

Da bismo ilustrirali kako se određuje stacionarna distribucija, promatrajmo primjer s tri kategorije popusta u Sekciji 1.1.1 gore.
Ako je vjerojatnost da osiguranik neće prijaviti štetu 0.9, $\tilde{\pi}$ je rješenje od

$$(\pi_0, \pi_1, \pi_2) \begin{bmatrix} 0.1 & 0.9 & 0 \\ 0.1 & 0 & 0.9 \\ 0 & 0.1 & 0.9 \end{bmatrix} = (\pi_0, \pi_1, \pi_2)\,,$$

to jest, $(0.1\pi_0 + 0.1\pi_1, 0.9\pi_0 + 0.1\pi_2, 0.9\pi_1 + 0.9\pi_2) = (\pi_0, \pi_1, \pi_2)$. To se može zapisati kao tri simultane jednadžbe.

$$0.1\pi_0 + 0.1\pi_1 = \pi_0 \quad (1)$$

$$0.9\pi_0 + 0.1\pi_2 = \pi_1 \quad (2)$$

$$0.9\pi_1 + 0.9\pi_2 = \pi_2 \quad (3)$$

Budući da imamo tri jednadžbe s tri nepoznanice, čini se da se jednadžbe mogu riješiti po π_0, π_1 i π_2. Problem je u tome da su samo dvije od tri jednadžbe linearno nezavisne. Međutim, također je poznato da je $\pi_0 + \pi_1 + \pi_2 = 1$, te se to može iskoristiti kao treća jednadžba. Rješavanje daje

$$\pi_0 = \frac{1}{91}, \quad \pi_1 = \frac{9}{91}, \quad \pi_2 = \frac{81}{91}\,.$$

2.2 Heterogenost portfelja

Jedan od razloga kojim se opravdava korištenje sustava bonusa je taj da rezultira automatskim određivanjem premija. Drugim riječima, osiguranik koji manje puta prijavljuje štetu plaća manje nego onaj koji više prijavljuje štetu. Iako je to očito istinito, čim se sustavi bonusa razmatraju podrobnije, uskoro se vidi da većina njih ne radi tako dobro kao što smo se nadali, te da premije koje na kraju osiguranici plaćaju nisu proporcionalne vjerojatnostima prijavljivanja štete.

Djelomično je razlog tome mali broj kategorija popusta i relativno niske razine ponuđenih popusta. Ali, razlog je također u niskim vjerojatnostima događanja šteta te stoga visokim vjerojatnostima da svi osiguranici dostignu u nekom trenutku maksimalnu razinu popusta.
Uz dane vjerojatnosti prijava šteta za sve osiguranike, matematički bi bilo moguće odrediti sustav bonusa koji bi dugoročno rezultirao time da svi osiguranici plaćaju čistu premiju direktno proporcionalnu njihovim vjerojatnostima štete. Međutim, time bi bilo izuzetno teško upravljati i bilo bi izuzetno teško za razumijeti.
Sljedeći primjer uzima situaciju sa suprotnog ekstrema pretpostavljajući da postoje samo dva moguća tipa osiguranika i samo tri kategorije popusta. Međutim, čak i u toj jednostavnoj situaciji nije lako proizvesti sustav koji pridružuje premije vjerojatnostima štete.

2.2.1 Primjer

Pretpostavimo da umjesto da imamo isto toliko vjerojatnosti šteta kao i osiguranika, postoje samo "dobri" vozači i "loši" vozači. Vjerojatnost da dobar vozač prijavi štetu je niža (recimo 0.1) od vjerojatnosti da loš vozač prijavi štetu (recimo 0.2). Mogu se naći očekivani omjeri dobrih i loših vozača u svakoj kategoriji nakon većeg broja godina.
Stacionarna distribucija za loše vozače je $\left(\frac{1}{21}, \frac{4}{21}, \frac{16}{21}\right)$.
Sada se mogu usporediti prosječne premije koje plaćaju dobri vozači i loši vozači. Budući da loši vozači prijavljuju dvostruko više šteta nego dobri vozači, za očekivati je da njihove čiste premije (tj. zanemarujući troškove i dodatak za profit) budu dvostruko veće od onih za dobre vozače (u prosjeku, pretpostavljajući da je distribucija veličina šteta jednaka za dobre i loše vozače). Pretpostavimo da je puna premija jednaka c. Tada je prosječna čista premija koju plaća dobar vozač jednaka

$$\frac{1}{91} \times c + \frac{9}{91} \times 0.75c + \frac{81}{91} \times 0.6c = 0.619c\,,$$

a prosječna čista premija koju plaća loš vozač jednaka

$$\frac{1}{21} \times c + \frac{4}{21} \times 0.75c + \frac{16}{21} \times 0.6c = 0.648c\,.$$

Tako je, usprkos činjenici da loši vozači s dvostrukom većom vjerojatnošću prijavljuju štetu nego dobri vozači, premija koju oni plaćaju samo marginalno veća (u prosjeku).

(U stvari u ovom jednostavnom primjeru je maksimalni dopustiv popust nedovoljan da se naplate "idealne" dvostruke premije. Međutim, uz malo algebre mogu se odrediti popusti koji bi rezultirali time da loši vozači plaćaju dvostruke premije od onih za dobre vozače.
Za vjerojatnosti šteta u ovom primjeru, zajedno sa samo tri kategorije popusta, potreban popust u kategoriji 2 bi bio barem 98.2%. Ta situacija bi se mogla popraviti ako bi postojalo više kategorija, ali bi kompleksnost počela rasti).

3 Posljedice sustava bonusa na sklonost prijavi štete

3.1 Ponovno određivanje prijelaznih vjerojatnosti

U tome što je do sada učinjeno pretpostavka je bila da je vjerojatnost kojom vozač prijavljuje štetu ista, neovisno u kojoj je on ili ona kategoriji popusta. Osiguranik može pri odlučivanju da li prijaviti štetu ili ne uzeti u obzir porast budućih premija. To se može razmatrati uspoređivanjem promjena u premijama u slučaju prijave štete.
Na primjer, promatrajmo vozača koji plaća punu premiju u sustavu bonusa sa tri kategorije kao u Sekciji 1.1.1, te pretpostavimo da je puna premija £500.
Ako se u prvoj godini (ili sljedećim godinama) ne prijavi šteta, buduće premije biti će $£375, £300, £300, \ldots$. Međutim, ako se šteta prijavi u prvoj godini, buduće premije biti će $£500, £375, £300 \ldots$. Ukupna dodatna premija nakon prijave štete je prema tome £200. Slični iznosi mogu se izračunati i za druge kategorije popusta.
Vjerojatnosti da u svakoj kategoriji osiguranik stvarno prijavi štetu biti će stoga različite.
Te razlike su bile izračunate gledajući unaprijed do godine u kojoj se dostigne maksimalni popust. Moguće je da osiguranik neće gledati tako daleko u budućnost zbog razloga da ionako očekuju drugu štetu u tom periodu. (Ekstremno, osiguranik može zanemariti sve takve račune zbog razloga da će još šteta biti učinjeno za vrijeme tekuće police). Broj budućih godina koje se razmatraju zove se osiguranikov horizont, i sklonost prijavi štete će također ovisiti o tom horizontu.

3.2 Računanje prijelaznih vjerojatnosti

Iz ovoga se može vidjeti da vjerojatnost da osiguranik pretrpi gubitak (tj. ima nesreću) nije ista kao vjerojatnost prijave štete. U slučaju da se dogodi nesreća, osiguranik može pretrpjeti gubitak u obliku oštećenja vozila i vlasništva, te kompenzacije ozlijeđenim osobama. Nakon nesreće osiguranik može (zanemarujući posljedice franšize) prijaviti štetu za gubitke koju će platiti osiguratelj. Ako je poznata distribucija šteta, može se izračunati vjerojatnost prijave štete nakon nesreće.
Na primjer, promotrimo osiguranika iz prethodnog primjera koji ima beskonačan horizont, trenutno je u kategoriji s 25% popusta i upravo je imao nesreću. Šteta će biti prijavljena samo ako su troškovi nesreće veći od £275. Ako slučajna varijabla X predstavlja trošak nesreće, tada

$$\mathbb{P}(\text{prijava štete} \,|\, \text{nesreća }) = \mathbb{P}(X > 275)\,.$$

Budući da je pretpostavljeno da je distribucija od X poznata, ta vjerojatnost se može izračunati.

3.2.1 Primjer

Osiguratelj motornih vozila radi po sustavu bonusa opisanog u Sekciji 1.1.1. s razinama popusta 0%, 25% i 40%. U slučaju jedne ili više prijava štete u godini, osiguranik sljedeće godine prelazi u sljedeći niži nivo popusta, ili ostaje na 0% popusta. U slučaju godine bez prijave štete, osiguranik sljedeće godine prelazi u sljedeći viši nivo popusta, ili ostaje na 40% popusta. Vjerojatnost nesreće tokom godine za dobre vozače je 0.1. Vjerojatnost nesreće tokom godine za loše vozače je 0.2. Vjerojatnost da bilo koji vozač ima dvije ili više nesreća tokom godine je toliko mala da se može pretpostaviti da je nula. Troškovi, u funtama, popravka nakon nesreće imaju lognormalnu distribuciju s parametrima $\mu = 5$ i $\sigma = 2$. Godišnja premija za osiguranika uz 0% popusta je £500. Osiguranik prijavljuje štetu nakon nesreće ako i samo ako je trošak popravka veći od razlike između:

(a) zbroja tri premije plative na sljedeća tri datuma godišnje obnove police, u slučaju da je prijavljena šteta za trošak popravka, i

(b) zbroja tri premije u slučaju da nije prijavljena šteta za trošak popravka.

U oba ova slučaja osiguranik pretpostavlja da on ili ona neće ima daljnjih nesreća prije datuma trećeg obnavljanja police.

(i) Za svaki nivo popusta izračunajte troškove popravka ispod kojih osiguranik neće prijaviti štetu.

(ii) Za svaki nivo popusta izračunajte vjerojatnost da će osiguranik prijaviti štetu nakon nesreće.

(iii) Izračunajte omjer dobrih i loših vozača za svaki nivo popusta pretpostavljajući da su ti omjeri dostigli stacionarnost.

Rješenje

(i) Tri godine je dovoljno dugo da vozač iz najniže kategorije popusta dospije u najvišu kategoriju popusta. Stoga su troškovi popravka ispod kojeg osiguranik neće prijaviti štetu isti kao u Sekciji 3.1, naime:

Nivo popusta 0%: £200

Nivo popusta 25%: £275

Nivo popusta 40%: £75.

(ii) $\mathbb{P}(\text{prijava štete} \,|\, \text{nesreća}) = \mathbb{P}(\text{trošak popravka} > x)$ gdje je x iznos nađen u (i).

Neka je X = trošak popravka. Tada je $X \sim$ lognormalna i $\log X \sim N(\mu, \sigma^2)$. Stoga zahtijevamo

$$\begin{aligned}\mathbb{P}(X > x) &= \mathbb{P}(\log X > \log x) \\ &= 1 - \Phi\left(\frac{\log x - \mu}{\sigma}\right)\end{aligned}$$

Za svaki nivo popusta je vjerojatnost prijave štete u slučaju nesreće:

0% popusta

$$1 - \Phi\left(\frac{\log 200 - 5}{2}\right) = 1 - \Phi(0.149) = 0.441$$

25% popusta

$$1 - \Phi\left(\frac{\log 275 - 5}{2}\right) = 1 - \Phi(0.308) = 0.379$$

40% popusta

$$1 - \Phi\left(\frac{\log 75 - 5}{2}\right) = 1 - \Phi(-0.341) = 0.633$$

(iii)

$$\mathbb{P}(\text{prijava štete}) = \mathbb{P}(\text{prijava štete} \,|\, \text{nesreća})\mathbb{P}(\text{ nesreća})$$

Prijelazna matrica za dobre vozače je

$$P = \begin{bmatrix} 0.0441 & 0.9559 & 0 \\ 0.0379 & 0 & 0.9621 \\ 0 & 0.0633 & 0.9367 \end{bmatrix}.$$

Stacionarna distribucija je rješenje od $\tilde{\pi}P = \tilde{\pi}$.

To daje sljedeće jednadžbe:

$$\begin{aligned} 0.0441\pi_0 + 0.0379\pi_1 &= \pi_0 &\quad (4)\\ 0.9559\pi_0 + 0.0633\pi_2 &= \pi_1 &\quad (5)\\ 0.9621\pi_1 + 0.9367\pi_2 &= \pi_2 &\quad (6) \end{aligned}$$

Iz jednadžbe (4),

$$\pi_1 = 25.222\pi_0$$

Iz jednadžbe (6)

$$\pi_2 = 15.199\pi_1 = 383.350\pi_0$$

Također,

$$\begin{aligned} &\pi_0 + \pi_1 + \pi_2 = 1 \\ \therefore\quad &\pi_0 + 25.222\pi_0 + 383.350\pi_0 = 1 \\ \therefore\quad &\pi_0 = 0.0024, \quad \pi_1 = 0.0616, \quad \pi_2 = 0.9360. \end{aligned}$$

Prijelazna matrica za loše vozače je

$$P = \begin{bmatrix} 0.0882 & 0.9118 & 0 \\ 0.0758 & 0 & 0.9242 \\ 0 & 0.1266 & 0.8734 \end{bmatrix}.$$

Rješavanje stacionarne jednadžbe $\tilde{\pi}P = \tilde{\pi}$ daje

$$\pi_0 = 0.0099, \quad \pi_1 = 0.1193, \quad \pi_2 = 0.8708.$$

POGLAVLJE 8 - ANALIZA RAZVOJNIH TROKUTOVA

Nastavni ciljevi:

(x)1. Definirati razvojni faktor i pokazati kako se skup pretpostavljenih razvojnih faktora može koristiti za projekciju budućeg razvoja trokuta odgode.

(x)2. Opisati i primjeniti osnovnu metodu ulančanih ljestvica za upotpunjenje trokuta odgode.

(x)3. Pokazati kako se osnovna metoda ulančanih ljestvica može prilagoditi tako da naknada za inflaciju postane eksplicitnom.

(x)4. Diskutirati alternativne načine za izvod razvojnih faktora koji mogu biti prikladni za upotpunjenje trokuta odgode.

(x)5. Opisati i primjeniti metodu prosječnog troška po šteti za upotpunjenje trokuta odgode.

(x)6. Opisati i primjeniti Bornhuetter-Fergusonovu metodu za procjenu iznosa neriješenih šteta.

(x)7. Diskutirati pretpostavke u temelju primjene metoda u (x)1 do (x)6 gore.

0 Uvod

0.1 Porijeklo razvojnih trokutova

Razvojni trokuti (trokuti odgode, engl. delay triangles) obično se pojavljuju u tipovima osiguranja (naročito neživotnog osiguranja) gdje može proteći neko vrijeme nakon štete dok se potpuno ne saznaju iznosi šteta koji se trebaju isplatiti. Važno je da se štete pripišu godini u kojoj je polica zaključena. Osigurateljno društvo treba znati koliki iznos šteta je obavezno platiti da bi moglo izračunati koliki ima višak. Međutim, može proći puno godina prije no što sazna točne ukupne štete. Postoji mnogo razloga za odgode do konačnog rješavanja ukupnih iznosa štete. Odgoda se može pojaviti prije prijave štete i/ili između prijave i konačnog podmirenja štete.

Jasno je da iako osigurateljno društvo ne zna točne brojke za ukupne štete svake godine, ono mora pokušati procijeniti tu brojku čim pouzdanije i točnije moguće.

0.2 Prikaz podataka o štetama

Postoji nekoliko načina predstavljanja podataka o štetama koji naglašavaju različite aspekte podataka. Ovdje će podaci biti prikazani kao trokut, što je najčešće korištena metoda. Godina u kojoj je zaključen posao i u kojoj je osiguratelj bio u riziku zove se godina nastanka štete. Broj godina do isplate zove se odgoda, ili razvojni period. Podaci o štetama podijeljeni su po godini nastanka štete i po razvojnoj godini. Sljedeća tablica je primjer podataka o štetama prikazanih pomoću godine nastanka štete i razvojne godine. U nekim tipovima osiguranja može biti relevantno promatrati razvoj šteta po mjesecima ili kvartalima, ali principi ostaju nepromijenjeni.

0.2.1 Primjer

Kumulativne isplate štete

Godina nastanka štete	**Razvojna godina**				
	0	**1**	**2**	**3**	**4**
1992	786	1410	2216	2440	2519
1993	904	1575	2515	2796	
1994	995	1814	2880		
1995	1220	2142			
1996	1182				

Slika 1:

Brojke su kumulativne i predstavljaju ukupan iznos isplaćen do kraja svake razvojne godine. Bile su sastavljene nakon završetka godine nastanka štete 1996. Za godinu nastanka štete 1996., prikazane su samo isplate s odgodom

0. Za godinu nastanka štete 1995., prikazane su isplate s odgodom 0 i 1, i tako dalje.
Zadatak je odrediti koliki iznosi se još trebaju isplatiti u odnosu na danu godinu nastanka štete. To se može učiniti za 1996. gledajući prethodne godine nastanka štete. Ako kumulativne isplate rastu na sličan način, vjerojatno je da bi mogle biti oko 3788 kroz četiri godine. Ta brojka dobivena je uz pretpostavku da je godina nastanka štete 1996. slična po uzorku isplata godini nastanka štete 1992., te uz procjenu kumulativnih isplata na kraju razvojne godine 4 pomoću

$$1182 \times \frac{2519}{786} = 3788\,.$$

To nije nužno "najbolja" procjena, ali je moguće popuniti donji trokut na Slici 1 usporedbom sadašnjih brojki s prošlim iskustvom. Taj proces je glavni objekt ovog poglavlja.

1 Projekcije korištenjem razvojnih faktora

1.1 Uzorci razvoja

Osnovna pretpostavka u procjeni neriješenih šteta odnosi se na uzorak razvoja. Najjednostavnija pretpostavka je da će isplate nastajati na sličan način svake godine nastanka štete. Proporcionalni porast u poznatim kumulativnim isplatama iz jedne godine u drugu, može se tada iskoristiti za računanje očekivanih kumulativnih isplata za buduće razvojne godine.
Međutim, kao što pokazuje sljedeći primjer, postoji više mogućnosti za omjer koji treba koristiti za projiciranje budućih šteta.
Napomena: Omjer koji se koristi za projiciranje budućih šteta zove se razvojni faktor.

1.1.1 Primjer

Svaka godina nastanka štete od 1992. do 1995. ima drugačiji omjer za porast kumulativnih isplata od razvojne godine 0 do razvojne godine 1. Nije jasno koji od njih je "ispravan" za projiciranje unaprijed za godinu nastanka štete 1993. Za konzervativnu procjenu kumulativnih isplata, možda bi bilo najbolje uzeti najveći omjer - t.j. 1.823.

Proporcionalni porasti kumulativnih isplata

Godina nastanka štete	**Razvojna godina** **0**		**1**		**2**		**3**		**4**
1992	786	*1.794*	1410	*1.572*	2216	*1.101*	2440	*1.032*	2519
1993	904	*1.742*	1575	*1.597*	2515	*1.112*	2796		
1994	995	*1.823*	1814	*1.558*	2880				
1995	1220	*1.756*	2142						
1996	1182								

Slika 2:

Međutim, neka vrsta usrednjenja omjera čini se prikladnijom. Moguće je koristiti jednostavnu aritmetičku sredinu:

$$\frac{1.794 + 1.742 + 1.823 + 1.756}{4} = 1.779\,.$$

Nedostatak toga je da ne uzima u obzir da godine u kojima ima više šteta daju više informacija. Stoga, čim je veći iznos šteta, tim je veće povjerenje koje možete imati u omjer. To sugerira korištenje težinske sredine i uobičajen izbor težina su kumulativne vrijednosti šteta.

Godina nastanka štete	**Omjer**	**Težina**
1992	1.794	786
1993	1.742	904
1994	1.823	995
1995	1.756	1220

$$\frac{1.794 \times 786 + 1.742 \times 904 + 1.823 \times 995 + 1.756 \times 1220}{786 + 904 + 995 + 1220} = 1.777$$

Ta metoda procjene omjera koji opisuju uzorak razvoja zove se metoda ulančanih ljestvica (engl. chain ladder method). Najefikasniji način računanja omjera dan je u sljedećoj sekciji.

1.2 Metoda ulančanih ljestvica

Ova metoda računanja razvojnih omjera pokazana je u sljedećem primjeru:

1.2.1 Primjer

Prisjetimo se da je omjer u godini nastanka štete 1992. izračunat kako slijedi:

$$1.794 = \frac{1410}{786}$$

Omjeri za druge godine nastanka štete izračunati su na sličan način. Brojnik jednadžbe na kraju sekcije 1.1 može se stoga zapisati kao

$$\begin{aligned} & \frac{1410}{786} \times 786 + \frac{1575}{904} \times 904 + \frac{1814}{995} \times 995 + \frac{2142}{1220} \times 1220 \\ = \; & 1410 + 1575 + 1814 + 2142 \, . \end{aligned}$$

Stoga se razvojni faktor može izračunati korištenjem kumulativnih šteta u razvojnim godinama 0 i 1:

$$\frac{1410 + 1575 + 1814 + 2142}{786 + 904 + 995 + 1220} \, .$$

Ime ove metode vjerojatno proizlazi iz "ljestvičastih" operacija koje su ulančane preko razvojnih godina. Razvojni faktori za tehniku ulančanih ljestvica mogu se pronaći za svaku razvojnu godinu dodavanjem odgovarajućeg broja članova. To je ilustrirano na slici 3.
Razvojni faktori izračunati su za svaku razvojnu godinu. Sada je moguće projicirati unaprijed svaku godinu nastanka štete.
Za godinu nastanka štete 1996. projekcije kumulativnih šteta su

$$\begin{array}{lllllllllll} 1182 & \times & 1.777 & & & & & & & = & 2100 \\ 1182 & \times & 1.777 & \times & 1.856 & & & & & = & 3331 \\ 1182 & \times & 1.777 & \times & 1.856 & \times & 1.107 & & & = & 3688 \\ 1182 & \times & 1.777 & \times & 1.856 & \times & 1.107 & \times & 1.032 & = & 3806 \end{array}$$

Za godinu nastanka štete 1995., počinje se od razvojne godine 1 i koriste samo zadnja 3 povezna omjera.
Uočite da se ne može napraviti projekcija za prvu godinu nastanka štete, budući da nije moguće projicirati poslije najviše razvojne godine.

GNŠ	**RG**				
	0	**1**	**2**	**3**	**4**
1992	786	1410	2216	2440	2519
1993	904	1575	2515	2796	
1994	995	1814	2880		
1995	1220	2142			
1996	1182				

$$\frac{6941}{3905} = 1.777 \qquad \frac{7611}{4799} = 1.586 \qquad \frac{5236}{4731} = 1.107 \qquad \frac{2519}{2440} = 1.032$$

Slika 3:

Projekcije kumulativnih isplata

Godina nastanka štete	**Razvojna godina**				
	0	**1**	**2**	**3**	**4**
1992					
1993					2885
1994				3188	3290
1995			3397	3761	3881
1996		2100	3331	3688	3806

Slika 4:

Pričuva koju treba držati na kraju 1996. je zbroj, kroz sve godine nastanka štete za koje je napravljena projekcija, razlika među kumulativne isplate na kraju razvojne godine 4 i zadnjeg poznatog upisa u razvojnom trokutu za tu godinu nastanka štete.
Dakle, iz slika 1 i 4, pričuva na kraju 1996. je:

$$(2885 - 2796) + (3290 - 2880) + (3881 - 2142) + (3806 - 1182) = 4862.$$

Uočite da diskontna stopa nije primjenjena na isplate u različitim godinama.

1.3 Provjera modela

Tehnika ulančanih ljestvica primarno se koristi za procjenu razvoja kumulativnih isplata šteta. Međutim, korisno je provjeriti da li je razumno prilagođena podacima o štetama koje su već zaprimljene. Da bismo to ilustrirali, pogledajmo podatke na slici 2.
Da bismo provjerili koliko je dobra tehnika ulančanih ljestvica, u sljedećem primjeru razmatrat ćemo štete za godine nastanka štete 1992.-1995.

1.3.1 Primjer

Ako su stvarne štete u razvojnoj godini 0 kako slijedi:

1992	786
1993	904
1994	995
1995	1220

Razvojni faktori izračunati u Sekciji 2.2 bili su 1.777, 1.586, 1.107 i 1.032. Koristimo li ih, možemo dobiti procjene kumulativnih isplata šteta u svakoj razvojnoj godini. Posebno je zanimljivo usporediti te procjene sa stvarnim vrijednostima danim na slici 2.
Stoga, tablica na slici 5 daje "prilagođene" vrijednosti korištenjem tehnike ulančanih ljestvica.
Sada se mogu usporediti slike 1 i 5. Međutim, kada se promatra prilagodba modela, preporučljivije je gledati poraste u kumulativnim isplatama. To daje osjetljiviji test.
Porasti kumulativnih isplata s razvojnim godinama (stvarnim i prilagođenim) dani su na slici 6.

Prilagođene kumulativne isplate štete

Godina nastanka štete	**Razvojna godina**				
	0	**1**	**2**	**3**	**4**
1992	786	1397	2215	2452	2531
1993	904	1606	2548	2820	
1994	995	1768	2804		
1995	1220	2168			

Slika 5:

Razvojna godina

		0	**1**	**2**	**3**	**4**
1992	stvarni	786	624	806	224	79
	prilagođeni	786	611	818	237	79
	greška	−−	*13*	*-12*	*-13*	*0*
1993	stvarni	904	671	940	281	
	prilagođeni	904	702	942	272	
	greška	−−	*-31*	*-2*	*9*	
1994	stvarni	995	819	1066		
	prilagođeni	995	773	1036		
	greška	−−	*46*	*30*		
1995	stvarni	1220	922			
	prilagođeni	1220	948			
	greška	−−	*-26*			

Slika 6:

Nijedna greška nije toliko velika da bi sugerirala netočnost modela.
Međutim, usprkos toj provjeri potpuno je moguće da su dobivene procjene slab vodič za budućnost.

1.4 Druge metode izvoda razvojnih faktora

Moguće je izračunate razvojne faktore prilagoditi u svijetlu drugih informacija. Toj metodi koja koristi ranije znanje može se formalno pristupiti, ali je to češće ad hoc prilagodba. Mogu postojati dobri razlozi za promjenu razvojnih faktora. Na primjer, preinake u računovodstvenim metodama ili administraciji odštetnih zahtjeva mogu promjeniti brzinu rješavanja šteta. To može dovesti do promjena u razvojnim faktorima i bilo bi razborito da se to odrazi na procjene budućih isplata šteta. Bilo da su izračunati neposredno iz podataka, bilo da su određeni ekspertnim znanjem, razvojni faktori se uvijek koriste na isti način za procjenu isplata neriješenih šteta.
Metoda ulančanih ljestvica može se upotrijebiti i na trokut podataka omjera šteta umjesto na kumulativne isplate, gdje je omjer šteta (engl. loss ratio) za danu razvojnu godinu i godinu nastanka štete kumulativna isplata do uključivo te razvojne godine podijeljena s ukupnim premijama u odnosu na danu godinu nastanka štete.

1.5 Pretpostavke u temelju metode

Tehnika ulančanih ljestvica zasniva se na pretpostavci da će se isplate za svaku godinu nastanka štete razvijati na isti način. Drugim riječima, isti razvojni faktori primjenjuju se za projekciju neriješenih šteta za svaku godinu nastanka štete. Promjene u stopi nastanka šteta mogu se uključiti samo "ručnom prilagodbom" razvojnih faktora.

2 Prilagodba za inflaciju

2.1 Metoda ulančanih ljestvica prilagođena za inflaciju

2.1.1 Inflacija u proteklom razdoblju

Inflacija će utjecati na isplate u razvojnom trokutu po kalendarskoj godini isplate. U ovdje promatranom modelu, pretpostavit će se da je godišnja stopa inflacije ista za sve štete unutar određene kalendarske godine isplate. Svaka

kalendarska godina isplate odgovara dijagonali u trokutu. Za ilustraciju, pogledajte opet sliku 1.
Kada se prilagođuje za inflaciju, potrebno je promatrati isplate u svakoj kalendarskoj godini, a ne kumulativne isplate. Prvi korak je izračunati priraste isplata iz kumulativnih isplata oduzimanjem duž svakog retka. Ista operacija bila je provedena u Sekciji 2.2 i sljedeća slika može se usporediti sa slikom 6.
Primjer
Slika 7 daje priraste (ili nekumulativne) isplate šteta za podatke sa slike 1.

Prirasti isplata šteta u novčanim iznosima

Godina nastanka štete			**Razvojna godina**		
	0	**1**	**2**	**3**	**4**
1992	786	624	806	224	79
1993	904	671	940	281	
1994	995	819	1066		
1995	1220	922			
1996	1182				

Slika 7:

Pretpostavimo da su godišnje stope inflacije korištene za isplate šteta kroz 12 mjeseci sve do sredine dane godine kao što slijedi:

1993	5.1%
1994	6.4%
1995	7.3%
1996	5.4%

Zbog jednostavnosti je također pretpostavljeno da se isplate vrše sredinom svake kalendarske godine. Sada je moguće izračunati indeks za pretvorbu svih isplata na cijene sredinom 1996.
Isplate sa slike 7 sada se mogu prilagoditi korištenjem stopa inflacije. Slika 8 daje podatke o prirastima isplata prilagođenim za inflaciju.

Prirasti isplata šteta u cijenama sredinom 1996.

Godina nastanka štete	**Razvojna godina**				
	0	**1**	**2**	**3**	**4**
1992	994	751	912	236	79
1993	1088	759	991	281	
1994	1125	863	1066		
1995	1286	922			
1996	1182				

Slika 8:

Sada je jednostavno formirati tablicu kumulativnih isplata prilagođenih za inflaciju na koju se može primjeniti tehnika ulančanih ljestvica.

Predviđanja kumulativnih isplata po cijenama sredinom 1996. dana su na slici 9.

Predviđanja kumulativnih isplata po cijenama sredinom 1996.

Godina nastanka štete	**Razvojna godina**				
	0	**1**	**2**	**3**	**4**
1993					3203
1994				3341	3431
1995			3383	3701	3801
1996		2048	3138	3433	3526

Slika 9:

2.1.2 Inflacija u budućem razdoblju

Predviđanja kumulativnih isplata ne uzimaju, međutim, u obzir inflaciju u budućem razdoblju. Da bi se prognozirale stvarne isplate, potrebno je pretpostaviti buduću stopu inflacije. Ponovno je potrebna pretvorba sa kumulativnih isplata na nekumulativne podatke prije njihove prilagodbe za inflaciju u budućem razdoblju, slično kao kod uzimanja u obzir inflacije u proteklom razdoblju.

Primjer

Primjenimo na podatke sa slike 9 godišnju stopu inflacije od 10% (na dan 30. lipnja). Dobivamo revidirana predviđanja kumulativnih isplata šteta kao što slijedi:

Predviđanja kumulativnih isplata šteta u novčanim iznosima

Godina nastanka štete			**Razvojna godina**		
	0	**1**	**2**	**3**	**4**
1993					2888
1994				3196	3305
1995			3435	3820	3953
1996		2135	3454	3847	3983

Slika 10:

Pričuva koju treba držati na kraju 1996. je 5129.

2.1.3 Pretpostavke u temelju metode

Ključna pretpostavka u temelju metode je da je za svaku godinu nastanka štete, iznos isplaćenih šteta, u realnoj vrijednosti, u svakoj razvojnoj godini konstantan omjer ukupnih šteta, u realnoj vrijednosti, iz te godine nastanka štete.

Eksplicitno se pretpostavlja inflacija u proteklom i budućem razdoblju.

3 Metoda prosječnog troška po šteti

Ova metoda zasebno razmatra dva ključna elementa ukupnih odštetnih zahtjeva, naime broj odštetnih zahtjeva i prosječne iznose šteta.

3.1 Opis metode

Ova metoda zahtijeva razvoj tablica i za ukupne štete i za broj odštetnih zahtjeva.
Treća razvojna tablica, za prosječne iznose šteta, tada se formira dijeljenjem brojki u odgovarajućim elementima prve dvije tablice.
Sljedeći korak je projiciranje brojki u tablici prosječnih šteta i tablici broja odštetnih zahtjeva, korištenjem bilo bruto faktora (engl. grossing-up factors), bilo razvojnih faktora.
Konačno, projicirane krajnje štete mogu se izračunati množenjem, za svaku godinu nastanka štete, projiciranih brojki za prosječne iznose šteta i broja odštetnih zahtjeva.
Tada se može izračunati pričuva oduzimanjem svih dosadašnjih isplata s obzirom na štete koje se odnose na podatke u tablici.
Dolje je dan primjer koji ilustrira taj proces.

3.2 Primjena metode

Metoda prosječnog troška po šteti nije jedinstveno definirana. Stoga se može jednako primjeniti na kohortu godine nastanka štete, bilo za isplaćene, bilo za nastale štete, ili na kohortu godine prijave štete. Važno je, međutim, osigurati da oblik podataka za broj odštetnih zahtjeva odgovara onom za ukupan iznos šteta (tj., isplaćene štete odgovaraju broju riješenih šteta, a nastale štete odgovaraju prijavljenom broju).
Naravno, važno je primjenjivati metodu samo na podatke za koje se razvoj smatra stabilnim i stoga prikladnim za projiciranje u budućnost. Vjerojatnije je da će prijavljivanje štete biti stabilnije od rješavanja, stoga se donji primjer odnosi na nastale štete.
Nadalje, metoda nije jedinstveno definirana u odnosu na korištenje određenog oblika bruto faktora ili razvojnih faktora. Međutim, metoda bruto faktora se općenito smatra jednostavnijom i u primjeru koji slijedi koristi se u svojem jednostavnom prosječnom obliku.

Konačno, gore opisana metoda zanemaruje bilo kakvu prilagodbu za inflaciju. To se, međutim, može učiniti na točno isti način kao kod prilagodbe osnovne metode ulančanih ljestvica i tvorbe metode ulančanih ljestvica prilagođene na inflaciju (tj., ako su korišteni podaci prilagođeni za inflaciju, jednostavno bi se zahtijevalo da se indeks buduće inflacije primjeni na nekumulativne projicirane prosječne iznose šteta prije množenja s projiciranim brojem odštetnih zahtjeva). U praksi se normalno radi prilagodba za inflaciju.

3.3 Primjer

Kumulativni podaci nastalih šteta, po godinama nastanka štete i prijavljenom razvoju

		RG						
		0	1	2	3	4	5	krajnje
GNŠ	1	2777	3264	3452	3594	3719	3717	3717
	2	3252	3804	3973	4231	4319		
	3	3725	4404	4779	4946			
	4	4521	5422	5676				
	5	5369	6142					
	6	5818						

Broj prijavljenih odštetnih zahtjeva, po godinama nastanka štete i prijavljenom razvoju

		RG						
		0	1	2	3	4	5	krajnje
GNŠ	1	414	460	482	488	492	494	494
	2	453	506	526	536	539		
	3	494	548	572	582			
	4	530	588	615				
	5	545	605					
	6	557						

Dijeljenje svakog elementa u prvoj tablici s odgovarajućim elementom u drugoj tablici daje akumulirane prosječne nastale troškove po šteti.

Prosječni nastali troškovi po šteti, po godinama nastanka štete i prijavljenom razvoju

		RG 0	1	2	3	4	5	krajnje
GNŠ	1	6.708	7.096	7.162	7.365	7.559	7.524	7.524
	2	7.179	7.518	7.553	7.894	8.013		
	3	7.540	8.036	8.355	8.498			
	4	8.530	9.221	9.229				
	5	9.851	10.152					
	6	10.445						

Te tablice vode do bruto faktora i projiciranih krajnjih brojki danih u sljedećoj tablici (projekcije se zasnivaju na podvučenim jednostavnim prosjecima bruto faktora).

Prosječni iznosi šteta

		RG 0	1	2	3	4	5	krajnje
GNŠ	1	6.708	7.096	7.162	7.356	7.559	7.524	7.524
		89.2%	94.3%	95.2%	97.9%	100.5%	100.0%	
	2	7.179	7.518	7.553	7.894	8.013		7.973
		90.0%	94.3%	94.7%	99.0%	**100.5%**		
	3	7.540	8.036	8.355	8.498			8.632
		87.4%	93.1%	96.8%	**98.45%**			
	4	8.530	9.221	9.229				9.657
		88.4%	95.5%	**95.57%**				
	5	9.851	10.152					10.766
		91.5%	**94.3%**					
	6	10.445						11.697
		89.3%						

Broj odštetnih zahtjeva

		RG						
		0	1	2	3	4	5	krajnje
	1	414	460	482	488	492	494	494
		83.8	93.1	97.6	98.8	99.6	**100.0%**	
	2	453	506	526	536	539		541
		83.7	93.5	97.2	99.1	**99.6%**		
	3	494	548	572	582			588
GNŠ		84.0	93.2	97.3	**98.95%**			
	4	530	588	615				632
		84.0	93.2	**97.37%**				
	5	545	605					649
		84.1	**93.25%**					
	6	557						11.664
		83.92%						

Ukupna krajnja šteta je stoga zbroj sljedećih projiciranih iznosa za svaku godinu nastanka štete:

GNŠ	Prosječni trošak po šteti	×	Broj zahtjeva	=	Projicirana procjena štete
1	7.524		494		3717
2	7.973		541		4313
3	8.623		588		5076
4	9.657		632		6103
5	10.766		649		6987
6	11.697		664		7767

Ukupna projicirana procjena štete = 33963

Ukoliko do sada isplaćene štete iznose 20334, ukupna potrebna pričuva bila bi 13629.

3.4 Pretpostavke u temelju modela

Budući da ne postoji jedinstven način definiranja metode, ne postoji niti jedinstven skup pretpostavki. Posebno će pretpostavke koje se odnose na inflaciju ovisiti o korištenim podacima.
Općenito, međutim, pretpostavke su da su za svaku godinu nastanka štete i broj i prosječni iznos šteta u odnosu na svaku razvojnu godinu u konstantnom omjeru sa ukupnim brojem i iznosom šteta za tu godinu nastanka štete.
Konačno, korisno je spomenuti sljedeće: da bi pretpostavke vrijedile za ovu metodu, normalno je da one također vrijede i za jednostavniju metodu koja primjenjuje ukupan, a ne prosječan, iznos šteta, kao što je metoda ulančanih ljestvica.

4 Omjeri šteta

Omjer nastalih šteta i prihodovanih premija kroz definirani period zove se omjer šteta.
Ispitivanja omjera šteta za svaku od nekoliko različitih godina nastanka štete obično bi pokazala izvjesnu konzistentnost, uz uvjet da nisu postojale iskrivljenosti, i naročito, nikakve značajne promjene stopa premija.
Očekivani omjer šteta će također biti dio izvoda osnova za premije.
Stoga je logično da se omjer šteta, baziran na prošlim podacima, pogledima preuzimača rizika, ili tržišnim podacima, može koristiti kao osnova za procjenu mogućih gubitaka, te zato i neriješenih šteta. Međutim, samo za sebe to je gruba mjera zbog fluktuacija prirođenih svakom opažanju šteta.

5 Bornhuetter-Fergusonova (B-F) metoda

5.1 Koncept Bornhuetter-Fergusonove metode

Bornhuetter-Fergusonova metoda udružuje omjer šteta s metodom projekcije. Ona stoga poboljšava grubu upotrebu omjera šteta uzimajući u obzir informaciju dobivenu zadnjim razvojnim uzorkom šteta, dok dodatak omjera šteta metodi projekcije služi za stabilnost protiv iskrivljenja u razvojnom uzorku.
Koncepti u temelju metode su:

- Kakve god štete su se već razvile u odnosu na danu godinu nastanka štete, daljni uzorak razvoja slijedit će onaj već doživljen za druge godine nastanka štete.

- Protekli razvoj za danu godinu nastanka štete ne pruža nužno bolji ključ za buduće štete od općenitijeg omjera šteta.

5.2 Opis metode

U svojem najjednostavnijem obliku koncept vodi do sljedećeg pristupa izračunima:

1. Odrediti početnu procjenu ukupnih krajnjih šteta za svaku godinu nastanka štete korištenjem premija i omjera šteta.

2. Podijeliti te procjene faktorima projekcije (f) određenim, na uobičajen način, iz tablice razvoja šteta. To su efektivno procjene šteta koje su se trebale razviti do danas.

3. Oduzeti te iznose od odgovarajućih ukupnih krajnjih šteta da bi se dobila procjena iznosa šteta koji se još trebaju razviti.

Jasno, ta tri koraka mogu se kombinirati i izraziti kao:

$$\text{Budući razvoj šteta} = \text{Premije} \times \text{Procjenjeni omjer šteta} \times (1 - 1/f)$$

Kako je konačna procjena krajnje štete zasnovana na opaženim podacima i na početnoj procjeni koja zanemaruje opažanja, na tu metodu možemo gledati kao na metodu s Bayesovskim pristupom.

5.3 Primjena metode

U svom izvornom obliku, Bornhuetter-Fergusonova metoda bila je primjenjena na razvoj nastalih šteta. Međutim, isto tako se može primjeniti na razvoj isplaćenih šteta, korištenjem bilo kohorte godine nastanka štete, bilo kohorte osigurateljne godine.

Nadalje, izvorna projekcija načinjena je korištenjem pristupa ulančanih ljestvica, iako se lako mogu primjeniti i alternativni razvojni faktori ili bruto faktori (g) (tj. g bi zamijenio $1/f$ u gornjem izazu).

U izvornom obliku također nije bilo eksplicitne prilagodbe za inflaciju, iako se metoda može prilagoditi na sličan način kao i druge metode.

Doljnji primjer zasnovan je na izvornom obliku metode, ali će ispitivač očekivati da studenti mogu primjeniti metodu na isplaćene štete.

5.4 Primjer

Prvi korak je odrediti razvojne faktore korištenjem iste metode kao kod metode ulančanih ljestvica.

Kumulativni podaci nastalih šteta, po godinama nastanka štete i prijavljenom razvoju

		RG						
		0	1	2	3	4	5	*krajnje*
	1	2866	3334	3503	3624	3719	3717	3717
	2	3359	3889	4033	4231	4319		
	3	3848	4503	4779	4946			
GNŠ	4	4673	5422	5676				
	5	5369	6142					
	6	5818						
UKUPNO		25933	23290	17991	12801	8038	3717	
UKUPNO - zadnji no:		20115	17148	12315	7855	3719		
OMJER (r)		1.158	1.049	1.039	1.023	0.999	1.000	
RAZVOJNI FAKTOR (f)		1.290	1.114	1.062	1.022	0.999	1.000	

Nadalje, očekivani krajni omjer šteta, recimo 83%, primjenjuje se na prihodovane premije (PP) da bi se dobila početna procjena ukupne krajnje štete (KŠ) za svaku godinu nastanka štete.

Početna procjena ukupnih krajnjih šteta, po godini nastanka štete

GNŠ	1	2	3	4	5	6
PP	4486	5024	5680	6590	7482	8502
KŠ (0.83PP)	3723	4170	4714	5470	6210	7057

(NB: U ovom primjeru je očekivani omjer štete uzet kao onaj doživljen za potpuno razvijenu prvu godinu nastanka štete. To je učinjeno isključivo na osnovu nedostatka drugih informacija).
Sljedeći korak je primjena razvojnih faktora za procjenu krajnjih šteta i dodatak nastalih šteta koje su već prijavljene.

Revidirana procjena ukupnih krajnjih šteta, po godini nastanka štete

GNŠ	6	5	4	3	2	1
f	1.290	1.114	1.062	1.022	0.999	1.000
$1-1/f$	0.225	0.102	0.058	0.022	-0.001	0
Početna KŠ	7057	6210	5470	4714	4170	3723
Nastajuća obveza	1588	633	317	104	-4	0
Prijavljena obveza	5818	6142	5676	4946	4319	3717
Krajnja obveza	7406	6775	5993	5050	4315	3717

Ukupna krajnja obveza u odnosu na tih šest godina nastanka štete je stoga 33256.
Ako do danas isplaćene štete iznose 20344, ukupna potrebna pričuva bila bi 12922.

5.5 Pretpostavke u temelju metode

Pretpostavke ponovno ovise o tome da li se koristi izvorna ili popravljena verzija metode.
Za izvornu metodu pretpostavke su iste kao kod osnovne metode ulančanih ljestvica, zajedno s pretpostavkom da je procijenjen omjer štete odgovarajući.

POGLAVLJE 9 - GENERALIZIRANI LINEARNI MODELI

Nastavni ciljevi:

(x)1. Definirati eksponencijalnu familiju distribucija. Pokazati da se sljedeće distribucije mogu napisati u tom obliku: binomna, Poissonova, eksponencijalna, gama, normalna.

(x)2. Navesti očekivanje i varijancu eksponencijalne familije, i definirati funkciju varijance i parametar skaliranja.

(x)3. Objasniti što se podrazumijeva pod funkcijom veze i kanonskom funkcijom veze, pozivajući se na distribucije iz 1.

(x)4. Objasniti što se podrazumijeva pod varijablom, faktorom koji uzima kategorijske vrijednosti i interaktivnim članom. Definirati linearni predviđatelj, ilustrirati njegov oblik u jednostavnim modelima, uključujući polinomijalne modele i modele s faktorima.

(x)5. Definirati devijaciju i skaliranu devijaciju i navesti kako se mogu procijeniti parametri generaliziranog linearnog modela. Opisati kako se može izabrati odgovarajući model koristeći analizu devijacije i ispitujući značajnost parametara.

(x)6. Definirati Poissonove i devijacijske ostatke i opisati kako ih se može upotrijebiti.

1 Uvod

Generalizirane linearne modele (GLM) možemo promatrati kao proširenje linearnih modela koji su bili razmatrani u Predmetu 101. Specijalno, vidjet ćemo da je model regresije jednostavan GLM, i mnoge ideje iz modeliranja regresije upotrebljavat će se u raspravi o GLM. Bitna razlika je da sada dozvoljavamo da distribucija podataka ne bude normalna. To je naročito važno u aktuarskom radu gdje podaci vrlo često nemaju normalnu distribuciju. Na primjer, u smrtnosti, Poissonova distribucija se koristi za modeliranje intenziteta smrtnosti, μ_x, a binomna distribucija za početnu stopu smrtnosti, q_x.

U neživotnom osiguranju, Poissonova distribucija često se koristi za modeliranje frekvencije šteta, a gama ili lognormalna distribucija za veličinu šteta. Svrha analize podataka obično se sastoji u odlučivanju koje varijable ili faktori su važni predviđatelji promatranog rizika, te zatim u kvantificiranju relacije između predviđatelja i rizika da bi se odredile odgovarajuće razine premija. Nastavni cilj pokriva temeljnu teoriju generaliziranih linearnih modela potrebnu za gore spomenute primjene.
GLM povezuje varijablu koju želimo predviđati (nazvanu varijabla odziva), s varijablama faktora (nazvane predviđateljima, kovarijatama (engl. covariates) ili nezavisnim varijablama) o kojima imamo informaciju. Da bismo to učinili, potrebno je najprije definirati distribuciju odziva. Tada se kovarijate mogu povezati s odzivom dopuštajući slučajnu varijaciju podataka. Stoga je prvi korak proučiti opći oblik distribucija (poznat kao eksponencijalne familije) koji se upotrebljava u GLM.

2 Eksponencijalne familije

Distribucija slučajne varijable Y pripada eksponencijalnoj familiji ako njena gustoća ima sljedeći oblik:

$$f_Y(y;\theta,\phi) = \exp\left[\frac{y\theta - b(\theta)}{a(\phi)} + c(y,\phi)\right] \tag{2.1}$$

gdje su a, b i c funkcije.
Uočite da to nije jedinstveno, te da se drugdje mogu vidjeti malo drugačije definirane eksponencijalne familije.
Imamo dva parametra u gornjoj definiciji: θ, koji se naziva "prirodnim" parametrom, je relevantan za model zbog odnosa odziva (Y) i kontroliranih varijabli, i ϕ poznat kao parametar skaliranja. Da bismo motivirali te definicije i kasniji razvoj, promotrimo prvo normalnu distribuciju.

2.1 Normalna distribucija

$$\begin{aligned} f_Y(y;\theta,\phi) &= \frac{1}{\sqrt{2\pi\sigma^2}}\exp\left[\frac{-(y-\mu)^2}{2\sigma^2}\right] \\ &= \exp\left[\frac{y\mu - \mu^2/2}{\sigma^2} - \frac{1}{2}\left(\frac{y^2}{\sigma^2} + \log 2\pi^2\right)\right] \end{aligned}$$

što je oblika (2.1) sa

$$\begin{aligned}\theta &= \mu\\ \phi &= \sigma^2\\ a(\phi) &= \phi\\ b(\theta) &= \theta^2/2\\ c(y,\phi) &= -\frac{1}{2}\left(\frac{y^2}{\sigma^2}+\log 2\pi\sigma^2\right).\end{aligned}$$

Stoga je prirodni parametar normalne distribucije jednak μ, a parametar skaliranja je σ^2.
Promotrimo log-vjerodostojnu funkciju, $l(y;\theta,\phi)=\log(f_Y(y;\theta,\phi))$. Koristit ćemo ju kasnije pri procjeni GLM. Trenutno trebamo dva vrlo dobro znana rezultata iz statističke teorije:

$$\mathbb{E}\left[\frac{\partial l}{\partial\theta}\right]=0 \qquad \text{i} \qquad \mathbb{E}\left[\frac{\partial^2 l}{\partial\theta^2}\right]+\mathbb{E}\left[\left(\frac{\partial l}{\partial\theta}\right)^2\right]=0$$

Upotrebom tih rezultata u (2.1), može se pokazati da su očekivanje i varijanca od Y jednaki:

$$\mathbb{E}[Y]=b'(\theta) \qquad \text{i} \qquad \mathrm{Var}(Y)=a(\phi)b''(\theta)\,,$$

gdje crtica označava derivaciju s obzirom na θ. To su vrlo korisni rezultati koje ćemo sada detaljnije proučavati. Prvo, promatramo li normalnu distribuciju, možemo izvesti očekivanje i varijancu

$$\begin{aligned} b(\theta)=\theta^2/2 &\quad \text{te stoga} \quad \mathbb{E}[Y]=b'(\theta)=\theta=\mu\\ a(\phi)=\phi &\quad \text{te stoga} \quad \mathrm{Var}[Y]=a(\phi)b''(\theta)=\phi=\sigma^2\,.\end{aligned}$$

Općenito, uočite da očekivanje ne ovisi o ϕ, te kada predviđamo Y važan je θ. Također, varijanca podataka ima dvije komponente: jednu koja uključuje parametar skaliranja, i drugu koja određuje način na koji varijanca ovisi o očekivanju. Za normalnu distribuciju, varijanca ne ovisi o očekivanju (zbog $b''(\theta)=1$), ali ćemo kod drugih distribucija vidjeti da varijanca ovisi o očekivanju.
Da bi se naglasila ovisnost o očekivanju, varijanca se često piše kao $\mathrm{Var}(Y)=a(\phi)V(\mu)$, gdje je "funkcija varijance" definirana sa $V(\mu)=b''(\theta)$.
Sada ćemo promatrati druge distribucije eksponencijalne familije. Uočite da uz malu zloupotrebu notacije koristimo f i za neprekidne i zs diskretne distribucije.

2.2 Poissonova distribucija

$$f_Y(y;\theta,\phi) = \frac{\mu^y e^{-\mu}}{y!} = \exp[y\log\mu - \log y!]$$

što je oblika (2.1) sa

$$\begin{aligned} \theta &= \log\mu \\ \phi &= 1, \text{ te zato } a(\phi) = 1 \\ b(\theta) = e^{\theta} & \\ c(y,\phi) &= -\log y! \end{aligned}$$

Zato je prirodni parametar Poissonove distribucije $\log\mu$, očekivanje je $\mathbb{E}[Y] = b'(\theta) = e^{\theta} = \mu$, a funkcija varijance je $V(\mu) = b''(\theta) = e^{\theta} = \mu$. Funkcija varijance nam kaže da je varijanca proporcionalna očekivanju. Vidimo da je varijanca u stvari *jednaka* očekivanju, jer je $a(\phi) = 1$.

2.3 Binomna distribucija

To je malo nespretnije, budući da prvo moramo binomnu slučajnu varijablu podijeliti s n. Pretpostavimo, dakle, da je $Z \sim$ binomna (n,μ). Stavimo $Y = \frac{Z}{n}$, tako da je $Z = nY$. Distribucija od Z je $f_Z(z;\theta,\phi) = \binom{n}{z}\mu^z(1-\mu)^{n-z}$, te supstituirajući za z, distribucija od Y je

$$\begin{aligned} f_Y(y;\theta,\phi) &= \binom{n}{ny}\mu^{ny}(1-\mu)^{n-ny} \\ &= \exp\left[n(y\log\mu + (1-y)\log(1-\mu)) + \log\binom{n}{ny}\right] \\ &= \exp\left[n\left(y\log\left(\frac{\mu}{1-\mu}\right) + \log(1-\mu)\right) + \log\binom{n}{ny}\right] \end{aligned}$$

što je oblika (2.1) sa

$$\begin{aligned} \theta &= \log\left(\frac{\mu}{1-\mu}\right) \ (\text{uočite da je inverz od toga } \mu = \frac{e^{\theta}}{1+e^{\theta}}\ ), \\ \phi &= n \\ a(\phi) &= \frac{1}{\phi} \\ b(\theta) &= \log(1+e^{\theta}) \\ c(y,\phi) &= \log\binom{n}{ny}. \end{aligned}$$

Zato je prirodni parametar binomne distribucije $\log\left(\frac{\mu}{1-\mu}\right)$, očekivanje je

$$\mathbb{E}[Y] = b'(\theta) = \frac{e^\theta}{1+e^\theta}\,,$$

a funkcija varijance je

$$V(\mu) = b''(\theta) = \frac{e^\theta}{(1+e^\theta)^2} = \mu(1-\mu)\,.$$

2.4 Gama distribucija

Najbolji način za razmatranje gama distribucije je zamijeniti parametre iz α i λ u α i $\mu = \frac{\alpha}{\lambda}$, t.j., $\lambda = \frac{\alpha}{\mu}$.

$$\begin{aligned} f_Y(y;\theta,\phi) &= \frac{\lambda^\alpha}{\Gamma(\alpha)} y^{\alpha-1} e^{-\lambda y} = \frac{\alpha^\alpha}{\mu^\alpha \Gamma(\alpha)} y^{\alpha-1} e^{\frac{y\alpha}{\mu}} \\ &= \exp\left[\left(-\frac{y}{\mu} - \log\mu\right)\alpha + (\alpha-1)\log y + \alpha\log\alpha - \log\Gamma(\alpha)\right] \end{aligned}$$

što je oblika (2.1) sa

$$\begin{aligned} \theta &= -\frac{1}{\mu} \\ \phi &= \alpha \\ a(\phi) &= \frac{1}{\phi} \\ b(\theta) &= -\log(-\theta) \\ c(y,\phi) &= (\phi-1)\log y + \phi\log\phi - \log\Gamma(\phi)\,. \end{aligned}$$

Stoga je, ignorirajući negativan predznak, prirodni parametar gama distribucije jednak $1/\mu$. Očekivanje je $\mathbb{E}[Y] = b'(\theta) = -1/\theta = \mu$. Funkcija varijance je $V(\mu) = b''(\theta) = 1/\theta^2 = \mu^2$, te je varijanca jednaka μ^2/α.

Konačno, lognormalna distribucija često se upotrebljava u neživotnom osiguranju za modeliranje distribucija veličine šteta. To se može uključiti u okvir GLM, jer ako je $Y \sim$ lognormalna, tada je $\log Y \sim$ normalna. Ako trebamo koristiti lognormalnu distribuciju, podatke prvo trebamo logaritmirati, te se tada može primjeniti modeliranje normalnom distribucijom.

3 Funkcije veze i linearni predviđatelji

Odnos između odziva i kontroliranih varijabli definiran je kroz $\mathbb{E}[Y]$. Razmatramo li prvo linearni model regresije za normalno distribuirane podatke, to se može napisati kako slijedi:

$$Y \sim N(\mu, \sigma^2)$$

gdje je $\mu = \beta_0 + \beta_1 x$. Uočite da su sastavni dijelovi modela

1. *Distribucija podataka*

U ovom slučaju to je normalna distribucija, ali će biti proširena do bilo koje distribucije koja se može napisati kao eksponencijalne familije.

2. *"Linearni predviđatelj"*

Linearni predviđatelj je funkcija kontroliranih varijabli. U ovom slučaju to je $\beta_0 + \beta_1 x$.

3. *"Funkcija veze"*

Nužno je povezati srednji odziv s linearnim predviđateljem. U ovom slučaju veza je direktna jednakost: $\mathbb{E}[Y]$ = linearni predviđatelj. Općenito uzimamo neku funkciju srednjeg odziva i tu funkciju zovemo funkcijom veze.
Spojimo li 2. i 3., imamo općenito odnos

$$g(\mu) = \eta$$

gdje je g funkcija veze i η je linearni predviđatelj.
Da bismo definirali GLM trebamo specificirati sve tri gornje komponente. U praksi, distribucija podataka obično je specificirana na početku (često definirana podacima), linearni predviđatelj može se odabrati ovisno o tome što se smatra odgovarajućim ili pogodnim, i tada se nalazi najbolja struktura modela promatrajući raspon linearnih predviđatelja. Naravno, to nisu pravila kojih se moramo pridržavati: moguće je da odgovara više od jedne distribucije, i te se moraju ispitivati prije donošenja konačne odluke. Može biti nejasno koju funkciju veze treba upotrijebiti, te se opet treba ispitati raspon funkcija.
Sada ćemo detaljnije promatrati funkcije veze i linearne predviđatelje.

3.1 Funkcije veze

Da bi funkcija veze odgovarala modelu potrebno je, tehnički, da bude diferencijabilna i invertibilna. Osim tih osnovnih uvjeta, veći broj funkcija su odgovarajuće za gornje distribucije. Za svaku distribuciju se prirodna, ili kanonska, funkcija veze definira sa $g(\mu) = \theta(\mu)$. Stoga je kanonska funkcija veze za svaki od gornjih slučaja

Normalna	Identiteta	$g(\mu) = \mu$
Poissonova	Log	$g(\mu) = \log(\mu)$
Binomna	Logit	$g(\mu) = \log\left(\frac{\mu}{1-\mu}\right)$
Gama	Inverz	$g(\mu) = \frac{1}{\mu}$

Te funkcije veze dobro rade za svaku od gornjih distribucija, ali ih nije nužno koristiti u svakom slučaju. Na primjer, možete koristiti identitetu kao funkciju veze u vezi s Poissonovom distribucijom, možete koristiti log funkciju veze za podatke koji imaju gama distribuciju, i tako dalje. Međutim, morate razmotriti posljedice izbora funkcije veze na moguće vrijednosti od μ. Na primjer, ako podaci imaju Poissonovu distribuciju, tada μ mora biti pozitivan. Ako upotrijebite log funkciju veze, tada je $\eta = \log(\mu)$ i $\mu = e^{\eta}$. Stoga je sigurno da će μ biti pozitivan bez obzira koju vrijednost (pozitivnu ili negativnu) poprimi linearni predviđatelj. To neće biti tako upotrijebite li identitetu kao funkciju veze.
Postoje i druge funkcije veze i za određene namjere modeliranja mogu biti vrlo složene. Kao temelj aktuarskih primjena, gornje četiri funkcije su često dovoljne.

3.2 Linearni predviđatelj

Kovarijate (također zvane kontrolne varijable ili varijable predviđanja), ulaze u model pomoću linearnih predviđatelja. Ovdje se također pojavljuju parametri koje treba procijeniti. U slučaju pravca za jednu kontrolnu varijablu x, linearni predviđatelj je $\beta_0 + \beta_1 x$, te da bi se taj model prilagodio potrebno je procijeniti faktore β_0 i β_1. U tom slučaju važna je stvarna vrijednost od x. Primjer varijable tog tipa koja se redovito pojavljuje u aktuarskoj praksi je dob osiguranika. Drugi glavni tip kovarijate je faktor koji uzima kategorijsku vrijednost. Na primjer, spol osiguranika je ili muški ili ženski, što tvori faktor s dvije kategorije (ili nivoa). Taj tip kovarijate može se parametrizirati tako da linearni predviđatelj ima član α_1 za muško, a član α_2 za žensko. Tako model koji uključuje učinak dobi i učinak spola na osiguranika može imati linearni predviđatelj

$$\alpha_i + \beta x$$

gdje je $i = 1$ za muško i $i = 2$ za žensko. Uočite da je parametar β_0 redundantan i nije uključen (može se procijeniti zasebno od α_1 i α_2). Uočite također da je učinak dobi osiguranika jednak za muške i ženske. Može se dopustiti *interakcija* između dvije kovarijate, dobi i spola, i imati linearni predviđatelj

$$\alpha_i + \beta_i x\,.$$

U ovom slučaju je učinak dobi osiguranika različit za muške i ženske. Faktor za sebe se naziva *glavnim učinkom*. Također je moguće imati interakciju između dva faktora koja dopušta da učinak jednog faktora na varijablu odziva ovisi o vrijednosti drugog faktora. Kada se u modelu koristi interaktivni član, moraju se uključiti oba glavna učinka. Model s glavnim učincima za dva faktora i njihovu interakciju ima linearni predviđatelj

$$\alpha_i + \beta_i + \gamma_{ij}\,.$$

Postoji notacija za specificiranje tih modela, koja je za gornje modele kako slijedi:

model	linearni predviđatelj
dob	$\beta_0 + \beta_1 x$
spol	α_i
dob + spol	$\alpha_i + \beta x$
dob + spol + dob. spol	$\alpha_i + \beta_i x$
dob* spol	$\alpha_i + \beta_i x$

Zadnja dva modela su ekvivalentna, a pokazana su odvojeno za ilustraciju upotrebe notacije za modele. I drugi modeli se mogu prilagoditi, uključujući, na primjer, model za dob bez slobodnog člana. Modeli se mogu specificirati na sličan način, i to se obično radi direktno unutar statističkih kompjutorskih paketa kao što su GLIM ili S-Plus.

Općenito, aktualna vrijednost varijable ulazi u linearni predviđatelj, dok za faktore postoji parametar za svaki nivo koji taj faktor može imati. Raspon linearnih predviđatelja je vrlo širok, i dolje su dani neki daljnji primjeri. Zahtjev je da bude linearan *u parametrima*.

Najjednostavniji model za varijablu je pravac, a to se može proširiti do polinoma, do funkcija te varijable, i do linearnih predviđatelja koji uključuju više od jedne varijable. To ilustrira donja tablica. Dob i trajanje se tretiraju kao varijable, dok su dob i grupa za određivanje premije vozila faktori. Ako postoji više od jednog faktora u modelu, tada uključivanje člana interakcije implicira da učinak svakog faktora ovisi o nivou drugog faktora.

model	linearni predviđatelj
dob	$\beta_0 + \beta_1 x$
dob + dob^2	$\beta_0 + \beta_1 x + \beta_2 x^2$
dob + trajanje	$\beta_0 + \beta_1 x + \beta_2 x_2$
log(dob)	$\beta_0 + +\beta \log x$
spol	α_i
grupa za određivanje premije vozila	β_j
spol + grupa za određivanje premije vozila	$\alpha_i + \beta_j$
spol* grupa za određivanje premije vozila	$\alpha_i + \beta_j$

4 Devijacija u prilagodbi modela

Parametri u GLM se obično procjenjuju upotrebom procjene maksimalne vjerodostojnosti. Funkcija log-vjerodostojnosti, $l(y;\theta,\phi) = \log(f_Y(y;\theta,\phi))$, ovisi o parametrima u linearnom predviđatelju preko funkcije veze. Stoga se procjene parametara pomoću maksimalne vjerodostojnosti mogu dobiti maksimiziranjem l u odnosu na parametre u linearnom predviđatelju. Aproksimacije standardnih pogrešaka parametara također se mogu dobiti upotrebom teorije asimptotske maksimalne vjerodostojnosti. Proces izbora modela također koristi metode koje su aproksimacije temeljene na teoriji maksimalne vjerodostojnosti, i u ovom odjeljku skiciramo taj proces.
Zasićeni model definira se kao model u kojem je toliko parametara koliko i opažanja, tako da su prilagođene vrijednosti jednake opaženim vrijednostima. Skalirana devijacija definira se kao dvostruka razlika između log-vjerodostojnosti modela kojeg razmatramo (poznat kao sadašnji model) i zasićenog modela. Devijacija za sadašnji model, D_M, definira se tako da je

$$\text{skalirana devijacija} = \frac{D_M}{\phi}\,.$$

Odluka koji model koristiti obično započinje razmatranjem devijacija za niz modela. Što je manja devijacija, to je model bolji s točke gledišta prilagodbe

modela. To se može ilustrirati promatrajući slučaj kada su podaci normalno distribuirani.
U tom slučaju je log-vjerodostojnost uzorka veličine n jednaka

$$\begin{aligned} l(y;\theta,\phi) &= \sum_{i=1}^{n} \log(f_Y(y_i;\theta_i,\phi)) \\ &= -\frac{n}{2}\log 2\pi\sigma^2 - \sum_{i=1}^{n} \frac{-(y_i-\theta_i)^2}{2\sigma^2} \end{aligned}$$

Za zasićeni model se parametar θ_i procjenjuje s y_i, te tako drugi član iščezava. Stoga je skalirana devijacija (dvostruka razlika između vrijednosti log-vjerodostojnosti sadašnjeg i zasićenog modela) jednaka

$$\sum_{i=1}^{n} \frac{(y_i-\hat{\theta}_i)^2}{\sigma^2}$$

gdje je $\hat{\theta}_i$ prilagođena vrijednost za sadašnji model. Devijacija je (podsjetimo se da je parametar skaliranja $\phi = \sigma^2$) dobro poznat zbroj kvadrata ostataka:

$$\sum_{i=1}^{n} (y_i-\hat{\theta}_i)^2 .$$

Stoga se može vidjeti da je devijacija mjera prilagodbe modela. Za normalno distribuirane podatke, skalirana devijacija ima χ^2-distribuciju. Do sada smo ignorirali problem parametra skaliranja koji također treba procijeniti. Za normalne podatke, uobičajena procedura je uzeti kvocijent zbroja kvadrata i upotrijebiti F-test (kao u analizi varijance). U slučaju podataka koji nisu normalno distribuirani, parametar skaliranja može biti poznat (na primjer, za Poissonovu distribuciju $\phi = 1$), i devijacija je samo aproksimativno (u stvari asimptotski) χ^2 distribucija. Zbog tih razloga uobičajena procedura je usporedba dva modela promatranjem razlike u devijaciji i uspoređivanje s χ^2 distribucijom. Značajna vrijednost (na nivou 5%) χ^2 distribucije sa ν stupnjava slobode je približno 2ν. Stoga, ako želimo odlučiti da li je model 2 (koji ima p parametara i devijaciju S_2) značajno poboljšanje u odnosu na model 1 (koji ima q parametara i devijaciju S_1), možemo usporediti $S_1 - S_2$ s χ^2_{p-q} distribucijom. Kao aproksimacija, preferirat ćemo model 2 ako je $S_1 - S_2 > 2(p-q)$. Važna točka je da se ova metoda usporedbe može upotrijebiti samo za *ugniježdene* modele. Drugim riječima, model 1 mora biti podmodel modela 2. Zato možemo uspoređivati dva modela za koje je distribucija podataka i funkcija veze ista, ali linearni predviđatelj ima dodatni parametar u modelu 2. Na primjer $\beta_0+\beta_1 x$ i $\beta_0+\beta_1 x+\beta_2 x^2$. Ali na ovaj način ne možemo uspoređivati različite distribucije podataka ili funkcije veze, ili, na primjer, kada su linearni predviđatelji $\beta_0+\beta_1 x+\beta_2 x^2$ i $\beta_0+\beta_2 \log x$. Treba biti jasno da možemo mjeriti važnost faktora ispitujući skalirane devijacije, ali ne možemo upotrijebiti gore skiciranu proceduru testiranja.

Ukratko, tablica devijacija za ugnježdene modele može se koristiti za ispitivanje skupa mogućih modela da bi se odredilo koji su faktori, interakcije, varijable ili funkcije varijabli važni predviđatelji varijable odziva.

4.1 Primjer

Za ilustraciju upotrebe devijacije u izboru modela, promatramo skup podataka koji se odnose na štete motornih vozila. Postoje tri faktora: dob osiguranika (do), starost vozila (sv) i premijska grupa vozila (gv). Ovi podaci su korišteni u knjizi P. McCullagh i J. Nelder, "Generalized Linear Models", koja je vjerojatno najraširenija referentna knjiga u ovom području. Koristimo analizu iz te knjige. Devijacije za potpun raspon mogućih modela pokazane su u doljnjoj tablici.

Model	Devijacija	Stupnjevi slobode	Razlike	
1	638.32	122		
			81.13	7
do	557.19	115		
			226.5	3
do + gv	330.65	112		
			204.1	3
do + gv + sv	126.51	109		
			34.32	21
do + gv + sv + do.gv	92.191	88		
			18.77	21
do + gv + sv + do.gv + do.sv	73.416	67		
			3.89	9
do + gv + sv + do.gv + do.sv + gv.sv	69.524	58		
			69.524	58
do + gv + sv + do.gv + do.sv + gv.sv + do.gv.sv	0	0		

Uočite da se prvi model sastoji jednostavno od konstante (1 parametar), a zadnji ima toliko parametara koliko i podataka, te je stoga prilagodba savršena. Zadnji model također sadrži i član s trostrukom interakcijom. U praksi, kod prilagođivanja modela određivanja premija takve je modele najbolje izbjegavati.

Pri odlučivanju koji model(i) adekvatno objašnjava podatke, treba ispitivati razlike u devijacijama i stupnjevima slobode. Redoslijed kojim se članovi dodaju modelu utječe na rezultate, te se u praksi može gledati nekoliko redoslijeda tako da se ne propusti ništa važno. Na primjer, svaki glavni učinak može se prilagoditi sam za sebe, umjesto da se svaki dodaje modelu kao što je napravljeno gore.

Ako je razlika u devijacijama veća od dvostrukih stupnjeva slobode, tada je dodani član značajan u objašnjenju varijacije u odzivu. Stoga vidimo da se svaki od glavnih učinaka čini značajnim i treba biti upotrebljen u modelu. Međutim, niti jedan od interaktivnih članova ne čini se naročito značajnim.

5 Analiza ostataka i ocjena prilagodbe modela

Nakon što je upotrebom svih devijacija nađen mogući model, treba ga provjeriti gledajući ostatke i značajnost parametara. Ostaci se zasnivaju na razlikama između opaženih odziva, y, i prilagođenih odziva, $\hat{\mu}$. Prilagođeni odzivi dobiju se primjenom inverza funkcije veze linearnog predviđatelja na prilagođene vrijednosti parametara. Na primjer, upotrijebimo li log funkciju vezu i linearni perdviđatelj koji ima samo dva glavna učinka, prilagođene vrijednosti mogu se dobiti iz

$$\hat{\mu} = e^{\alpha_j + \beta_j} .$$

Pearsonovi ostaci definiranu su kao $\frac{y-\hat{\mu}}{\sqrt{\mathrm{Var}(\hat{\mu})}}$, dok su ostaci devijacija definirani kao produkt predznaka od $y - \hat{\mu}$ i drugog korijena doprinosa od y devijaciji. Stoga je ostatak devijacije jednak $\mathrm{sign}(y - \hat{\mu})d_i$, gdje je devijacija $\sum d_i$.
Pearsonov ostatak, koji se često upotrebljava za normalno distibuirane podatke, ima nedostatak da mu je distribucija često nesimetrična za podatke koji nisu normalni. To čini teškom interpretaciju dijagrama ostataka. Za ostatke devijacija vjerojatnije je da budu simetrično distribuirani i da imaju približno normalnu distribuciju, te im se stoga daje prednost u aktuarskim primjenama. Za normalno distribuirane podatke su Pearsonovi ostaci i ostaci devijacije identični.
Nakon što se upotrijebi tablica devijacija za identifikaciju modela (ili više njih) koji može biti prikladan za razmatrane podatke, model se treba dalje ispitivati promatranjem dijagrama ostataka i značajnosti parametara. Pretpostavke GLM zahtijevaju da ostaci ne pokazuju nikakav uzorak. Prisutnost uzorka povlači da nešto nedostaje u relaciji između predviđatelja i odziva. Ako je to slučaj, treba pokušati s drugom specifikacijom modela. Za nalaženje uzoraka i outliera potrebno je ispitivati dijagrame ostataka u odnosu na varijable i faktore u modelu.
Također je potrebno ispitivati histogram ostataka, ili neki drugi sličan dijagnostički dijagram, da bi se odredilo da li su opravdane pretpostavke o distribuciji. Također je korisno ispitati značajnost parametara. Drugim riječima, treba odrediti da li je svaki parametar značajno različit od nule. Ako nije, moguće je da se model može pojednostavniti. Približne standardne greške parametara mogu se dobiti upotrebom asimptotske teorije maksimalne vjerodostojnosti. Ugrubo, pokazatelj značajnosti parametara

dan je dvostrukom standardnom greškom. Stoga, ako je

$$\hat{\beta} > 2 \text{ standardna greška}(\hat{\beta})$$

parametar je značajan i treba ga zadržati u modelu. Inače, moguće je da ga se može odbaciti. Treba primjetiti da se u nekim slučajevima može učiniti da je parametar nepotreban po tom kriterijuj, ali da model bez njega ne daje dovoljno dobru prilagodbu podacima.
Ponekad je nužna prilična fleksibilnost u interpretaciji testova temeljenih na statističkom zaključivanju da bi se stiglo do prikladnog modela. Stoga se na gore navedene interpretacije devijacije, ostataka i značajnosti parametara treba gledati kao korisne vodiče u odabiru modelu, a ne kao na pravila kojih se treba pridržavati.

Literatura

1. N. Bowers et al., *Actuarial Mathematics*, 2nd edition, Society of Actuaries (1997)

2. H. Bühlmann, *Mathematical methods in risk theory*, Springer, Heidelberg (1970)

3. C. D. Daykin, T. Pentikäinen, M. Pesonen, *Practical risk theory for actuaries*, Chapman & Hall, London (1994)

4. E. De Vylder, *Advanced Risk Theory: A Self-Contained Introduction*, University of Brussels (1996)

5. *Foundations of Casualty Actuarial Science*, 4th edition, Casualty Actuarial Society (2001)

6. R. V. Hogg, S. A. Klugmann, *Loss Distributions*, Wiley, New York (1984)

7. I. B. Hossack, J. H. Pollard, B. Zehnwirth, *Introductory statistics with applications in general insurance*, 2nd edition, Cambridge University Press (1999)

8. S. A. Klugman, H. H. Panjer, G. E. Wilmott, *Loss Models; From Data to Decisions*, Wiley (1998)

9. P. McCullagh, J. A. Nelder, *Generalized Linear Models*, 2nd edition, Chapmann & Hall (1989)

10. T. Rolski, H. Schmidli, V. Schmidt, J. Teugels, *Stochastic Processes for Insurance and Finance*, Wiley (1998)

11. E. Straub, *Non-Life Insurance Mathematics*, Springer (1997)