Biztosítási kárszámok becslése

Diplomamunka

Írta: Martinek László

alkalmazott matematikus szak

Témavezető:

Dr. Arató Miklós, egyetemi docens
Valószínűségelméleti és Statisztika Tanszék
Eötvös Loránd Tudományegyetem, Természettudományi Kar

Eötvös Loránd Tudományegyetem
Természettudományi Kar
2010

Tartalomjegyzék

Ábrák jegyzéke

1. fejezet

Bevezető

A kötelező gépjármű-felelősségbiztosításokban széles körben elterjedt módszer, hogy a szerződők kockázati besorolását ún. bónusz-, ill. máluszosztályokkal végzik. Nevezetesen minden újonnan szerződő, azaz kezdő sofőr első évében egy kiinduló osztályban kezd, majd a következő évben átkerül egy másik osztályba attól függően, hogy balesetmentes éve volt-e vagy sem. Amennyiben nem okozott kárt az adott évben, jobb besorolásba kerül. Ha pedig az ő hibájából következően baleset következett be, akkor attól függően, hogy mennyi, rosszabb kategóriába kerül. A rosszabb osztályokban - melyeket hagyományosan máluszosztályoknak nevezünk - tartózkodó sofőrök díja nagyobb, mint a jobb osztályokban - bónuszosztályok - lévőknek. Ez természetes gondolat, hiszen a veszélyesebb sofőröket célszerű magasabb díjszabással büntetni, hiszen a biztosító várhatóan többet fog kifizetni kárrendezésre az ő közlekedésük folyományaként, mint a megbízható, jó vezetők miatt.

Emellett persze a díj kiszámításakor figyelembe szoktak venni több paramétert, mint például az autó motorjának lökettérfogatát, használatának célját (magán, ill. üzleti), a biztosított tartózkodási helyét (városi, ill. vidéki), életkorát stb. Ezek a jellemzők mind közrejátszanak a kockázati helyzet felmérésében, így a díjkalkulációban is. Mi most ezektől eltekintve fogjuk vizsgálni azt a központi kérdést, hogy vajon egy illető biztosított várhatóan mennyi kárt okoz egy év alatt. Ezt a λ értéket az ő kárgyakoriságának, más néven kárszámparaméterének fogjuk hívni. Fontos feltételezés lesz, hogy ezt az egyénre jellemző számot állandónak tekintjük, azaz életkorának változásával ez a paraméter nem változik.

Először ismertetjük röviden a magyar rendszer szabályait, és ismertetjük szükséges feltételezéseinket, többek között az egy év alatt okozott károk eloszlására, valamint az osztályok közötti mozgás folyamatának Markov-tulajdonságára. A második fejezet a paraméter maximum likelihood, ill. Bayes-i becslésére koncentrál, és a jelentős hangsúlyt ez utóbbi kapja. Ehhez feltételezünk egy a priori eloszlást

λ-ra, melynek paramétereit az állományunkból tudjuk becsülni, majd ezzel, és az egyénről szerzett információval az a posteriori várható értékre vagyunk kíváncsiak, amely λ becslése lesz. Ez az információ esetünkben kétféle lesz, egyrészt ismerhetjük a kártörténetet néhány évre visszamenőleg, azaz az elmúlt években okozott kárszámokat. A másik lehetséges információ az lesz, hogy a szerződő ismert számú év alatt egy ismert osztályba kerül. Ennek abból a szempontból van nagy jelentősége, hogy biztosítóváltáskor az új biztosító nem feltétlenül jut hozzá a kártörténethez, csak az eltöltött évek számát és az aktuális osztályt ismeri. Látni fogjuk, hogy ez a minimális információ többet mond, mint amennyire először gondolnánk. Ez abban nyilvánul meg, hogy meglepően jó becslést adunk ennek ismeretében az egyén kárgyakoriságára, azaz az a posteriori szórásnégyzet kicsi lesz.

A legérdekesebb kérdés ezen utóbbi információ alapján vett feltételes várható érték kiszámítása, hiszen ekkor a feltételes eloszlást nem tudjuk nevezetes paraméteres osztályba besorolni, sőt még a közvetlen mintavételezés sem lehetséges. Ezért lemondunk az analitikus megoldásról, és egy lehetséges alternatívát a későbbiekben ismertetett Monte-Carlo-típusú módszer formájában ismertetünk, mely nem túl gyorsan ugyan, de konvergál a keresett értékhez. Gyakorlati szempontból tekintve ez egy átlagos számítógépen 5 és 10 perc közötti számítási időt igényel. A probléma jellegét tekintve valójában ez még elfogadható, hiszen a már meglévő összes szerződőből kalkulált a priori paramétereket nem kell óráról órára változtatnunk.

Utolsó fejezetünkben keressük a lehetséges módokat a Monte-Carlo-típusú módszerünk gyorsításához. Egyrészt megvizsgáljuk, hogy az ún. importance sampling (fontossági mintavételezés) módszerét hogyan alkalmazhatnánk erre az esetre a mintavételi eloszlás és a rá alkalmazott függvény megváltoztatásával, melyből később átlagot számítunk, közelítve a várható értéket. Ez a gondolat természetes módon adódik abból a megállapításból, hogy bizonyos alacsonyabb osztályokhoz tartozó esetekben a generált mintának csak elenyésző hányadát használjuk fel. Másrészt az utolsó alfejezetben egy Markov-lánc-Monte-Carlo típusú módszerrel próbálunk mintát venni az a posteriori eloszlásból.

A szokásos, valószínűségszámításban és statisztikában elterjedt jelöléseket fogjuk alkalmazni, ill. ahol szükséges, ott röviden elmagyarázzuk azokat.

2. fejezet

Bónusz-málusz rendszerek

Mára a világ nagy részén a kötelező gépjármű-felelősségbiztosításban az ún. bónusz-málusz rendszereket vezették be, mint díjszámítási elvet. Lényege abban áll, hogy a biztosítottakat oly módon próbálják óvatosságra bírni, hogy a károkat okozók rosszabb besorolásba kerülnek, a kármentesen vezetőket pedig bónusszal jutalmazzák, azaz kedvezményt kapnak az alapdíjból.

Általánosan konstruálhatnánk olyan (végtelen terjedelmű) táblázatot, melyben a függőleges tengely az elmúlt évek számát, a vízszintes pedig az okozott károk számát jelöli, és ebben a t-edik sor k-adik eleme a t év alatt k kárt okozó sofőr $t+1$-edik évi díját jelöli. Ennek kezelése kissé - talán feleslegesen - bonyolult lenne, ezért mindenhol véges sok osztállyal számolnak. Általánosan jelöljük a legrosszabb osztályt C_1-gyel, a második legrosszabbat C_2-vel stb., a legjobbat pedig C_n-nel, ha n osztály van. Mi a magyar rendszerrel fogunk foglalkozni, melyben 15 osztályunk van: 4 málusz- (M1,...,M4), egy kiinduló (A0) és 10 bónuszosztály (B1,...,B10). Megjegyezzük emellett, hogy a később ismertetett módszerek kisebb módosítással teljesen hasonlóan alkalmazhatóak más országok rendszereire. Az itthoni szabályok értelmében minden egyes kármentes évvel eggyel jobb osztályba kerül a szerződő, de legfeljebb B10-be. Minden kár okozásával 2 osztállyal sorolják vissza, de legalább 4 kár esetén a legrosszabb M4-be kerül.

2.1. Markov-láncok

Fontos feltételezésünk egyrészt, hogy a t-edik évben az egyén B_t osztályba kerülése csak az előző évi osztályától függ, az előtörténettől nem. Más szóval feltesszük, hogy a különböző díjosztályokban való mozgás Markov-láncként modellezhető, azaz $P(B_t = C_i | B_{t-1}, \ldots, B_1) = P(B_t = C_i | B_{t-1})$. Másrészt feltesszük, hogy a biztosított egy biztosítási évben okozott kárszáma Poisson eloszlású valószínűségi változó

λ paraméterrel. Ez az egyénre vonatkozó ún. kárgyakoriság, az egy év alatt okozott kárszám várható értéke centrális szerepet fog játszani a következőkben. Ugyanis a biztosítónak fontos, hogy minél pontosabb becslést tudjon adni a valódi paraméterre, hiszen ez alapján tudja becsülni a várható károkat, ezzel együtt a kifizetéseket. Persze ehhez nem elég a károk számát tudni, hanem a károk, a kifizetendő összegek mértékét is közelíteni kell, amivel most nem foglalkozunk. További lényeges - a valóságnak nem megfelelő - feltételezésünk, hogy a λ kárgyakoriság az időben állandó, azaz a korral nem lesz sem ügyesebb, sem veszélyesebb a sofőr. A változó λ esetét duplán sztochasztikus, ún. Cox-folyamatokkal írhatnánk le, azaz mikor $\lambda(t)$ maga is sztochasztikus folyamat.

A fenti feltételezésekkel kapcsolatban mintegy pontosításként szükségünk van a következő definícióra.

2.1.1 Definíció (Homogén Markov-lánc) Legyen $\{B_n\}_{n\geq 0}$ diszkrét idejű sztochasztikus folyamat megszámlálható állapottérrel. Ha tetszőleges nemnegatív n egészre, és $i_0, i_1, \ldots, i_{n-1}, i, j$ állapotokra a

$$P(B_{n+1} = j \mid B_n = i, B_{n-1} = i_{n-1}, \ldots, B_0 = i_0) = P(B_{n+1} = j | B_n = i)$$

egyenlőség teljesül, akkor ezt a folyamatot Markov-láncnak nevezzük. Homogén a Markov-lánc, ha emellett még $P(B_{n+1} = j|B_n = i)$ független n-től. Ez esetben értelmes bevezetnünk erre egy n-től független jelölést, például $p_{ij} = P(B_{n+1} = j|B_n = i)$.

Így az állapotokon való bolyongásnak definiálhatjuk az átmenetvalószínűség-mátrixát, melyben az i-edik sor j-edik eleme annak a valószínűségét mondja meg, hogy az i-edik osztályból milyen valószínűséggel kerülünk át következő évre a j-edik osztályba. Rögzített λ-ra az $M(\lambda)$-val jelölt sztochasztikus mátrixra[1] $M(\lambda) \in M_{15}(\mathbb{R})$, és a következő alakot ölti:

[1] minden sorban a sorösszeg 1-gyel egyenlő, hiszen bármelyik osztályban is legyünk, valamelyikbe átkerülünk a következő évben

$$M(\lambda)=\begin{pmatrix}
1-e^{-\lambda} & e^{-\lambda} & 0 & \dots & 0 & 0 & 0\\
1-e^{-\lambda} & 0 & e^{-\lambda} & \ddots & 0 & 0 & 0\\
1-e^{-\lambda} & 0 & 0 & \ddots & 0 & 0 & 0\\
1-e^{-\lambda} & \lambda e^{-\lambda} & 0 & \ddots & 0 & 0 & 0\\
1-(\lambda+1)e^{-\lambda} & 0 & \lambda e^{-\lambda} & \ddots & 0 & 0 & 0\\
1-(\lambda+1)e^{-\lambda} & \lambda^2\cdot e^{-\lambda}/2! & 0 & \ddots & 0 & 0 & 0\\
1-(\lambda^2/2!+\lambda+1)e^{-\lambda} & 0 & \lambda^2\cdot e^{-\lambda}/2! & \ddots & 0 & 0 & 0\\
1-(\lambda^2/2!+\lambda+1)e^{-\lambda} & \lambda^3\cdot e^{-\lambda}/3! & 0 & \ddots & 0 & 0 & 0\\
1-(\lambda^3/3!+\lambda^2/2!+\lambda+1)e^{-\lambda} & 0 & \lambda^3\cdot e^{-\lambda}/3! & \ddots & 0 & 0 & 0\\
\vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots\\
1-(\lambda^3/3!+\lambda^2/2!+\lambda+1)e^{-\lambda} & 0 & 0 & \ddots & 0 & 0 & e^{-\lambda}\\
1-(\lambda^3/3!+\lambda^2/2!+\lambda+1)e^{-\lambda} & 0 & 0 & \dots & \lambda e^{-\lambda} & 0 & e^{-\lambda}
\end{pmatrix}$$

A kezdeti eloszlás világos, hogy $\pi_0 = (0,\dots,0,1,0,\dots,0)^T$, ahol az ötödik elem 1, a többi 0. Hiszen mindenki az 5. osztályban (A0) kezd, más szóval 1 valószínűséggel C_5-ben lesz, 0 valószínűséggel máshol. Továbbá $\pi_0 M^t$ azt mondja meg, hogy t év elteltével milyen valószínűséggel lesz az egyes osztályokban, azaz t lépés után milyen diszkrét eloszlást kapunk az osztályokon.

Itt érdemes egy rövid kitérőt tennünk a sztochasztikus mátrixokról:

Egy $n\times n$-es mátrixot sztochasztikusnak mondunk, ha elemei $[0,1]$-beli valós számok, és minden sorban a sorösszeg 1-gyel egyenlő. Továbbá emlékeztetünk a spektrálelméletből a mátrixok spektrálfelbontására, mely szerint az M mátrix előállítható UDV^T alakban, ahol U és V a bal, ill. jobb oldali sajátvektorokból képezett (unitér) mátrixok, D pedig diagonális mátrix M sajátértékeivel a főátlóban. Ily módon M tetszőleges egész t-edik hatványa felírható $M^t = \sum_{i=1}^{n} \lambda_i^t u_i v_i^T$ alakban [2], ugyanis minden $i \neq j$ esetén $u_i v_j^T = 0$. Mivel pillanatnyi célunk, hogy az átmenetvalószínűség-mátrix hatványainak konvergenciájáról mondjunk valamit, ez jelentős mértékben motiválhat minket. Gondoljunk csak arra, hogy λ_i-ktől elvárjuk az 1-nél kisebb abszolút értékűséget, máskülönben a hatvány növekedtével λ_i^t tartana végtelenbe. Ennek tükrében megemlítünk egy fontos tételt, nevezetesen:

Tétel 2.1.2 (Perron-Frobenius) *Legyen A $r\times r$-es nemnegatív (elemenként nemnegatív), irreducibilis és aperiodikus mátrix. A-nak létezik λ_1 pozitív valós sajátértéke 1 algebrai, és szintén 1 geometriai multiplicitással, és minden $j \in \{2,\dots,r\}$-re $\lambda_1 > |\lambda_j|$, továbbá a λ_1-hez tartozó bal és jobb oldali sajátvektor felírható úgy, hogy*

[2] nagyon vigyázzunk, hiszen itt nem a kárgyakoriságot, hanem M sajátértékeit jelöljük λ_i-kel

$u_1^T v_1 = 1$ teljesüljön. Teljesüljön a sajátértékekre $\lambda_1 > |\lambda_2| \geq \ldots \geq |\lambda_r|$, és ha $|\lambda_2| = |\lambda_j|$ valamely $j \geq 3$-ra, akkor $m_2 \geq m_j$, ahol m_j a λ_j algebrai multiplicitása.

Ekkor $A^n = \lambda_1^n v_1 u_1^T + O(n^{m_1-1}|\lambda_2|^n)$. Továbbá ha A még sztochasztikus is, akkor $\lambda_1 = 1$.

2.1.3 Megjegyzés 1) Az irreducibilitás azt fejezi ki, hogy bármely állapotból el lehet jutni bármely állapotba. A periódus az adott állapotba visszatérés lépésszámainak legnagyobb közös osztója, így egy állapot aperiodikus, ha ez a legnagyobb közös osztó 1. Erre igaz, hogy osztálytulajdonság, tehát ha irreducibilis Markov-láncot tekintünk, melynek egy állapotáról könnyen látható az aperiodikusság, akkor az egész aperiodikus.

2) $O(f(n))$ olyan függvénye n-nek, melyhez létezik $0 < \alpha \leq \beta < \infty$ valós számok, hogy $\alpha f(n) \leq O(f(n)) \leq \beta f(n)$ teljesül elég nagy n esetén.

Könnyen meggondolható, hogy a bónusz-málusz rendszerünk irreducibilis Markov-láncot határoz meg. Ugyanis egy lehetséges okoskodás szerint bárhonnan eljuthatunk $M4$-be, onnan pedig egyesével a többi osztályba. Az aperiodikusság osztálytulajdonsága miatt elég belátni, hogy $M4$ aperiodikus állapot. Ez szintén egyszerű, mert például 2, ill. 3 lépésben is visszaérhetünk ide. Az átmenetvalószínűség-mátrix sztochasztikussáiga nyilvánvaló. Ezek szerint teljesülnek rá a Perron-Frobenius-tétel feltételei, így létezik a $\lim_{t\to\infty} M^t = v_1 u_1^T = \underline{1}\pi^T$ határérték, ahol $v_1 = \underline{1}$ csupa 1 vektor. Más szóval létezik stacionárius eloszlás, melyet π-vel jelölünk. A konvergencia sebességét pedig a második legnagyobb abszolút értékű sajátérték, és annak multiplicitása határozza meg. Be kell látnunk, hogy ez a konvergencia esetünkben nem elég gyors ahhoz, hogy figyelmen kívül hagyhassuk a kisebb hatványok közötti eltéréseket. Tehát $M(\lambda)$ hatványait szükségszerűen külön kell kezelnünk a valószerű hatványokra emelés mellett. Ez a valószerűség már szubjektív dolog, de elég arra gondolnunk, hogy a szerződők nagy része 30 évnél nem régebben lépett be a rendszerbe, és ezalatt a hatvány alatt nagyobb λ-ra (pl. $0,2$-re) még nem stacionarizálódik az eloszlás. Tehát meg kell különböztetnünk egymástól a későbbi becsléseknél azokat az eseteket, amikor valaki 15 vagy 17 éve van jelen a rendszerben, ezzel szemben már nem érdemes a 45 vagy 50 éve biztosítottak között különbséget tenni.

2.2. Várható díj

Természetes kérdés, hogy a szerződő mennyi díjat fog kifizetni a biztosítónak az első t évben. Mivel ez t valószínűségi változó összege, ezért csak várható értékkel számolhatunk, ráadásul hallgatólagosan feltesszük, hogy nem változtat biztosítót

2.1. táblázat. Díjak (egységben)

M4	$d_1 = 2$
M3	$d_2 = 1.6$
M2	$d_3 = 1.35$
M1	$d_4 = 1.15$
A0	$d_5 = 1$
B1	$d_6 = 0.95$
B2	$d_7 = 0.9$
B3	$d_8 = 0.85$
B4	$d_9 = 0.8$
B5	$d_{10} = 0.75$
B6	$d_{11} = 0.7$
B7	$d_{12} = 0.65$
B8	$d_{13} = 0.6$
B9	$d_{14} = 0.55$
B10	$d_{15} = 0.5$

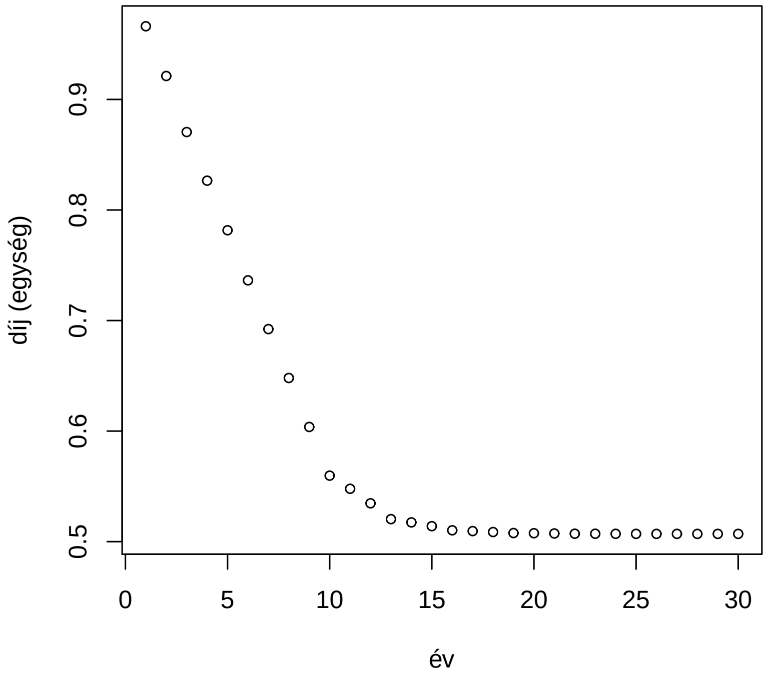
ez alatt az idő alatt. A különböző osztályokhoz tartozó díjakat az 2.1 táblázatban olvashatjuk, ahol A0-ban egységnyit fizetünk évente.

Jelölje P_i az i-edik évben fizetendő díjat, amely egy diszkrét valószínűségi változó, és a fenti díjértékeket veszi fel az évhez tartozó valószínűséggel. Kezdetben $P_1 = 1$ a π_0 eloszlásból adódóan. A következő évben az eloszlást a π_0^T sorvektor és az átmenetvalószínűség-mátrix szorzata adja, majd minden egyes év elteltével egy újabb jobbról szorzással kapjuk π_t-t. Ezzel az első t év várható összdíja

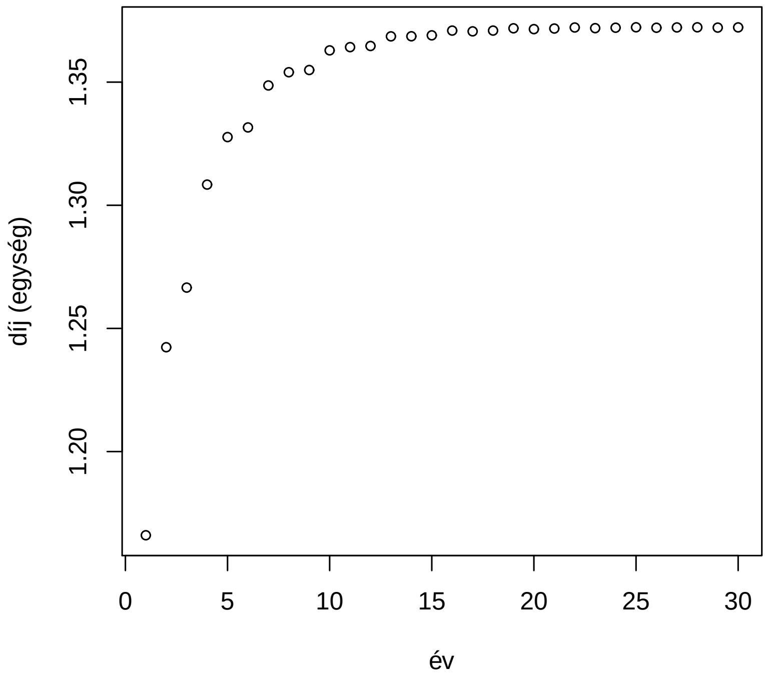
$$E(P_1 + \ldots + P_t) = \sum_{j=1}^{t} EP_j = \sum_{j=1}^{t} \pi_0^T \cdot M(\lambda)^{j-1} \cdot d = \pi_0^T \left(\sum_{j=0}^{t-1} M(\lambda)^j \right) d,$$

ahol d oszlopvektor elemei a fent ismertetett díjak.

2.2.1 Példa Lássunk két példát 30 éves időhorizonton a várható díj értékére. Az előzőek miatt a stacionárius eloszlás létezésével kapcsolatban tetszőleges λ paraméterre ezen várható díj értéke konvergens. Az első (2.1) példán egy jó (vagy szerencsés) sofőr $\lambda = 0,04$, a másikon (2.2) egy rossz (vagy peches) sofőr $\lambda = 0,5$ kárgyakorisággal látható.



2.1. ábra. éves várható díj $\lambda = 0,04$ mellett



2.2. ábra. éves várható díj $\lambda = 0,5$ mellett

3. fejezet

Becslések

A problémát az okozza, hogy nem ismerjük az egyén kárgyakoriságát, ráadásul arra néha rendkívül csekély mértékben rendelkezésre álló információkból kell becslést adnunk. Erre egy, a gyakorlatban is jelentkező példa, hogy a szerződő teljes kártörténetéhez biztosítóváltáskor az új biztosító nem jut hozzá, tehát csak annyit tudunk, hogy t évet töltött a rendszerben, és ezalatt egy bizonyos C_i osztályba jutott el, és persze kihasználjuk, hogy mindenki A0-ban kezdett. Egy másik példa, hogy a teljes kártörténet helyett csak az előző néhány évi áll rendelkezésre, ráadásul a rendszerben eltöltött évek száma is ismeretlen lehet. Ezeket a lehetőségeket fogjuk most részletesen vizsgálni. Először az első információ alapján fogjuk λ-t vizsgálni több szemszögből. Ehhez egy fontos jelölésünk lesz $B_t = c$ annak az eseménynek a kifejezésére, hogy t év elteltével [1] az illető a c osztályba kerül.

3.0.2 Megjegyzés Néha a $B_t = C_i$ jelölést fogjuk használni, ha szeretnénk hangsúlyozni, hogy melyik a szóban forgó osztály.

3.1. Maximum likelihood becslés

Keressük a likelihood-egyenletet λ-ra, ami azzal ekvivalens, hogy a $B_t = C_i$ osztály valószínűségét szeretnénk maximalizálni λ-ban, azaz a feladat

$$\max_{\lambda} \pi_0^T \cdot M(\lambda)^t \cdot e_i = \max_{\lambda} M(\lambda)^t_{(5,i)}.$$

Itt $e_i = (0, \ldots, 0, 1, 0, \ldots, 0)^T$ olyan vektor, melynek i-edik eleme 1, a többi 0, tehát az 5. sor i-edik elemét kell maximalizálni. Tegyük hozzá, hogy λ-nak megengedünk tetszőleges nemnegatív értéket. Rögtön felmerül a kérdés, hogy miként lehet jól kezelni ilyen mátrix t-edik hatványát, hiszen λ függvényében formálisan kifejezni

[1] a Markov-lánc állapotterén tett t lépéssel, más szóval a $t + 1$-edik biztosítási évében

nagyobb kitevőkre egészen reménytelennek tűnik. Ha jobban meggondoljuk, egy $f(\lambda) : \mathbb{R}_0^+ \to [0,1]^{15\times 15}$ függvény esetén $f^t(\lambda)$-t kell maximalizálnunk az $(5,i)$-edik elemében. Ehhez vesszük a szóban forgó elem λ szerinti deriváltját, melynek értéke a keresett $\hat{\lambda}$-ra egyenlő 0-val. Emlékeztetünk, hogy mátrix deriváltja az eredetivel megegyező méretű mátrix, melyben az elemeket úgy kapjuk, hogy a megfelelő elemeket egyenként deriváljuk.

Állítás 3.1.1 *Az n-edik hatvány deriváltjára teljesül az* $(M^n)' = \sum_{k=0}^{n-1} M^k M' M^{n-k-1}$ *összefüggés.*

Bizonyítás. Ismert összefüggés, hogy $(AB)' = A'B + AB'$, így $(M^n)' = M'M^{n-1} + M(M^{n-1})'$. Ugyanezt M^{n-1}-re alkalmazva indukcióval kapjuk a fenti állítást. □

Ezzel a likelihood egyenlet a

$$\left(\sum_{k=0}^{n-1} M^k \cdot M' \cdot M^{n-k-1}\right)(\lambda)_{(5,i)} = 0$$

alakot ölti. Egy nem túl előnyös tulajdonsága a módszernek, hogy a legrosszabb osztályra ennek nem létezik maximuma, ami számolás nélkül könnyen meggondolható. Hiszen minél nagyobb a várható kárszám, annál nagyobb valószínűséggel lesz valaki a legrosszabb osztályban. Hasonló a helyzet a B10 legjobb díjosztállyal, ahol a valószínűség maximumát $\lambda = 0$ mellett veszi fel. Gondoljuk meg, hogy nem szerencsés 0 kárgyakorisággal venni azokat a szerződőket, akik akár a biztosítási állomány $70 - 80\%$-át is kitehetik.

A 3.1 táblázat a kezdettől számítva t év elteltével a c osztályban található biztosított kárgyakoriságára vonatkozó maximum likelihood becslés értékeit tartalmazza 3 tizedesjegyre kerekítve. Ahol NA található, abba az osztályba annyi idő alatt lehetetlen eljutni.

3.2. Bayes-i megközelítés

Ebben a részben azzal a Bayes-i feltételezéssel élünk, hogy maga a kárgyakoriság is valószínűségi változó, ezt jelöljük Λ-val. A gyakorlatban előforduló esetek nagy többségében jogosan választjuk ezen a priori eloszlást gamma eloszlásúnak - azaz $\Lambda \sim \Gamma(\alpha, \beta)$ -, ezért mi most erre szorítkozunk. Ezt a gamma eloszlást más néven keverő eloszlásnak nevezzük. Rögzített $\Lambda = \lambda$-ra a szerződő $X \sim Poisson(\lambda)$ számú kárt okoz, így $P(X = k|\Lambda = \lambda) = \frac{\lambda^k}{k!}e^{-\lambda}$. Vajon mi lesz a feltétel nélküli eloszlás?

Állítás 3.2.1 *A* $\Gamma(\alpha, \beta)$ *keverőjű keverék Poisson-eloszlású valószínűségi változó eloszlása Negatív Binomiális* $(\alpha, \frac{\beta}{1+\beta})$ *paraméterrel.*

3.1. táblázat. Maximum likelihood becslések 1-20 éve szerződőkre

t\c	M4	M3	M2	M1	A0	B1	B2	B3	B4	B5	B6	B7	B8	B9	B10
1	∞	NA	1.000	NA	NA	0.000	NA	NA	NA	NA	NA	NA	NA	NA	NA
2	∞	1.137	NA	0.500	NA	NA	0.000	NA	NA	NA	NA	NA	NA	NA	NA
3	∞	0.857	0.701	NA	0.334	NA	NA	0.000	NA	NA	NA	NA	NA	NA	NA
4	∞	1.124	0.542	0.513	NA	0.250	NA	NA	0.000	NA	NA	NA	NA	NA	NA
5	∞	0.897	0.776	0.415	0.407	NA	0.200	NA	NA	0.000	NA	NA	NA	NA	NA
6	∞	0.817	0.606	0.602	0.340	0.337	NA	0.167	NA	NA	0.000	NA	NA	NA	NA
7	∞	0.947	0.524	0.477	0.495	0.289	0.288	NA	0.143	NA	NA	0.000	NA	NA	NA
8	∞	0.835	0.687	0.414	0.401	0.421	0.252	0.252	NA	0.125	NA	NA	0.000	NA	NA
9	∞	0.824	0.577	0.556	0.352	0.348	0.367	0.223	0.223	NA	0.111	NA	NA	0.000	NA
10	∞	0.873	0.543	0.466	0.473	0.310	0.309	0.325	0.201	0.201	NA	0.100	NA	NA	0
11	∞	0.814	0.642	0.429	0.401	0.415	0.279	0.279	0.292	0.182	0.183	NA	0.091	NA	0
12	∞	0.822	0.569	0.531	0.368	0.356	0.370	0.254	0.254	0.266	0.167	0.167	0.084	0.084	0
13	∞	0.837	0.565	0.464	0.460	0.328	0.322	0.336	0.233	0.234	0.188	0.154	0.106	0.077	0
14	∞	0.808	0.618	0.449	0.404	0.410	0.298	0.296	0.302	0.216	0.211	0.162	0.138	0.097	0
15	∞	0.816	0.569	0.516	0.385	0.363	0.371	0.275	0.273	0.272	0.200	0.194	0.110	0.128	0
16	∞	0.818	0.579	0.467	0.452	0.344	0.332	0.340	0.255	0.254	0.212	0.187	0.136	0.102	0
17	∞	0.805	0.603	0.467	0.408	0.407	0.315	0.307	0.307	0.239	0.230	0.187	0.164	0.126	0
18	∞	0.810	0.572	0.507	0.402	0.369	0.372	0.291	0.286	0.279	0.222	0.212	0.131	0.154	0
19	∞	0.809	0.585	0.471	0.447	0.360	0.340	0.342	0.272	0.267	0.229	0.208	0.158	0.121	0
20	∞	0.803	0.595	0.480	0.414	0.405	0.329	0.316	0.312	0.255	0.242	0.205	0.179	0.146	0

Bizonyítás. A Gamma-eloszlás sűrűségfüggvényét $f_\Lambda(\lambda)$-val jelölve

$$P(X=k)=\int_0^\infty P(X=k|\Lambda=\lambda)\cdot f_\Lambda(\lambda)d\lambda=\int_0^\infty \frac{\lambda^k}{k!}e^{-\lambda}\cdot\frac{\lambda^{\alpha-1}\beta^\alpha e^{-\beta\lambda}}{\Gamma(\alpha)}d\lambda=$$

$$=\int_0^\infty \frac{\lambda^{\alpha+k-1}\beta^\alpha e^{-(1+\beta)\lambda}}{k!\Gamma(\alpha)}d\lambda=\frac{\beta^\alpha\Gamma(\alpha+k)}{k!\Gamma(\alpha)(1+\beta)^{\alpha+k}}\cdot\int_0^\infty\frac{\lambda^{\alpha+k-1}(1+\beta)^{\alpha+k}e^{-(1+\beta)\lambda}}{\Gamma(\alpha+k)}d\lambda=,$$

használjuk ki, hogy az integrálban a $\Gamma(\alpha+k,\beta+1)$ eloszlás sűrűségfüggvénye található, így értéke 1, továbbá a gamma függvény $\Gamma(z+1)=z\Gamma(z)$ összefüggését, ezzel tovább

$$=\frac{\beta^\alpha\cdot(\alpha+k-1)\cdot\ldots\cdot\alpha\cdot\Gamma(\alpha)}{k!\Gamma(\alpha)(1+\beta)^{\alpha+k}}=\binom{\alpha+k-1}{k}\cdot\left(\frac{\beta}{1+\beta}\right)^\alpha\cdot\left(\frac{1}{1+\beta}\right)^k.$$

□

3.2.2 Megjegyzés A $\Gamma(\alpha,\beta)$ eloszlású valószínűségi változó sűrűségfüggvénye $f(x)=\frac{x^{\alpha-1}\cdot\beta^\alpha\cdot e^{-\beta x}}{\Gamma(\alpha)}$ alakú, ahol $\Gamma(\alpha)=\int_0^\infty t^{\alpha-1}e^{-t}dt$, továbbá a várható értéke $\frac{\alpha}{\beta}$, szórásnégyzete $\frac{\alpha}{\beta^2}$.

3.3. Az a priori paraméterek becslése

Ezzel elsődleges feladatunk a rendelkezésekre álló $\eta_1,\eta_2,\ldots,\eta_K$ éves kárszámadatok alapján becslést adni az α,β paraméterekre. Erről a mintáról feltesszük, hogy független, és azonos negatív binomiális eloszlásból származik. Fontos az állomány homogenitása, ugyanis ha példaként említve rossz sofőrök káradatait a valós aránytól nagyobb százalékban használnánk, akkor az rontana a paramétereken. Két lehetőséget mutatunk be, egyrészt a momentum módszert, másrészt a maximum likelihood módszert.

Momentum Módszer: Legyen $\eta_1\sim NegBinom(r,p)$, melyre

$$P(\eta_1=k)=\binom{k+r-1}{r-1}(1-p)^rp^k=\frac{\Gamma(r+k)}{k!\Gamma(r)}(1-p)^rp^k$$

(ahol $k=0,1,\ldots$). Ezzel $E\eta_1=\frac{rp}{1-p}$ és $D^2\eta_1=\frac{rp}{(1-p)^2}$. Definiáljuk minden pozitív egész i-re az $M_i=\frac{1}{K}\sum_{j=1}^\infty j^i\cdot|\{\eta_k=j\}|$ tapasztalati momentumokat, ahol $|\{\eta_k=j\}|$ a j db kárt okozó szerződések számát jelöli, továbbá legyen $S^2=M_2-M_1^2$. Innen $M_1=\frac{rp}{1-p}$ és $S^2=\frac{rp}{(1-p)^2}$, ami a $\hat p=1-\frac{M_1}{S^2}$ és $\hat r=M_1\cdot\frac{1-\hat p}{\hat p}=\frac{M_1^2}{S^2-M_1}$ becslésekhez vezet.

Maximum likelihood-módszer: A likelihood-függvény

$$L(r,p)=P_{r,p}(\eta_1=n_1,\ldots,\eta_K=n_K)=\prod_{i=1}^K\frac{\Gamma(r+n_i)}{n_i!\cdot\Gamma(r)}(1-p)^rp^{n_i},$$

azaz a loglikelihood

$$l(r,p) = rK\ln(1-p) + \ln p \cdot \sum_{i=1}^{K} n_i + \sum_{i=1}^{K}\sum_{m=0}^{n_i-1} \ln(r+m) - \sum_{i=1}^{K} \ln n_i!,$$

majd a két változó szerint deriválva az

$$\frac{\partial l(r,p)}{\partial p} = -rK\cdot\frac{1}{1-p}+\frac{1}{p}\sum_{i=1}^{K} n_i = 0 \text{ } \textit{és} \text{ } \frac{\partial l(r,p)}{\partial r} = K\cdot\ln(1-p)+\sum_{i=1}^{K}\sum_{m=0}^{n_i-1}\frac{1}{r+m} = 0$$

egyenletrendszert kell megoldanunk. A likelihood-egyenletnek egyértelműen létezik megoldása, ha $S^2 > M_1$.

3.3.1 Megjegyzés Fontos, hogy végezzünk mindezek után hipotézisvizsgálatot, ezzel ellenőrizzük, hogy X valóban negatív binomiális eloszlású. Ugyanis előfordulhat, hogy rossz feltételezésekkel éltünk, és nem jó a modell. (Ehhez megfelelő eszközt a χ^2-próbák között találunk.)

3.4. Monte-Carlo típusú módszer

Ezzel konkrét α, β feltételezés birtokában becsüljük λ-t. A rendelkezésre álló információk az ismert $B_0 = A0$, és a t évvel későbbi $B_t = c$ állapot. A Bayes-tétel alapján a feltételes sűrűségfüggvény

$$f_{\Lambda|B_t=c}(\lambda|c) = \frac{P(B_t = c|\Lambda = \lambda)\cdot f_\Lambda(\lambda)}{\int\limits_0^\infty P(B_t = c|\Lambda = \lambda)\cdot f_\Lambda(\lambda)d\lambda}.$$

Ezzel λ-ra a becslés Λ feltételes eloszlás szerinti, azaz a posteriori várható értéke lesz, ami

$$\lambda_t \stackrel{jel}{=} \int_0^\infty \lambda \cdot f_{\Lambda|B_t=c}(\lambda|c)d\lambda.$$

Írhattuk volna az általánosabb, eloszlással felírt alakot, de a gamma eloszlás abszolút folytonossága miatt nem baj, ha sűrűségfüggvényt használunk. A nehézséget az okozza, hogy a $P(B_t = c|\Lambda = \lambda)$ feltételes valószínűség rögzített λ-ra ugyan könnyen számolható mátrixhatványozással, de λ függvényében reménytelennek tűnik meghatározni. Hiszen ez egyenlő $M(\lambda)^t_{(5,|c|)}$-vel, ahol $|c|$ a c osztály oszlopához tartozó indexet jelöli, ezért a közvetlen integrálásról le kell mondanunk. Tehát nemhogy a feltételes sűrűségfüggvényt nem tudjuk kifejezni, még az általa meghatározott eloszlásból sem tudunk véletlen mintát generálni, amelynek átlagát véve közelíthetnénk a kívánt várható értéket a nagy számok törvényének értelmében.

Ehelyett Monte-Carlo-típusú módszert konstruálunk a következőképpen. Készítsünk egy (ξ_n, η_n) párokból álló sorozatot úgy, hogy ξ_n értékét Λ eloszlásából, nevezetesen $\Gamma(\alpha, \beta)$-ból generáljuk. Továbbá generáljunk egy Markov-láncot $\Lambda = \xi_n$ szerint, amely egy t lépésből álló út az állapottéren. Ha ezen ξ_n mesterkélt kárgyakoriság mellett $B_t = c$-be jutunk, akkor η_n értékét 1-nek választjuk, különben 0-nak. A Markov-tulajdonság miatt elég, ha vesszük az $M(\xi_n)$ átmenetvalószínűség-mátrix t-edik hatványának $(5, |c|)$ elemét, és ekkora valószínűséggel 1-et, különben 0-t vesz fel η_n. Ezek szerint η_n egy ξ_n-től függő $Binom(1, p)$ változó, ahol $p = M(\xi_n)^t_{(5,|c|)}$. A gondolatmenetet az motiválja, hogy a véletlen generált ξ_n értékeket, melyeket a valódi λ értékének meghatározásához választunk, megritkítsuk. Méghozzá oly módon, hogy a megmaradók átlaga már jól közelítse az a posteriori várható értéket. Más szóval, a valószínűtlenebb ξ_n-eket jó eséllyel kidobjuk az átlagolásból. Ne feledkezzünk meg a (ξ_n, η_n) és (ξ_m, η_m) párok egymástól való függetlenségétől, hiszen a ξ_n-eket egymástól függetlenül választottuk a gamma eloszlásból. Ugyanez a függetlenség a párokon belül persze már egyáltalán nem mondható el. Szükségünk lesz a következő két állításra.

Állítás 3.4.1 $\lambda_t = E(\Lambda | B_t = c) = E(\xi_1 | \eta_1 = 1)$. *(Mostantól ha nem szükséges, az indexet elhagyjuk ξ-ről és η-ról, hiszen mint valószínűségi változók, eloszlásuk megegyezik.)*

Bizonyítás. A teljes várható érték tételével

$$E(\xi\eta) = E(\xi\eta | \eta = 1) \cdot P(\eta = 1) + E(\xi\eta | \eta = 0) \cdot P(\eta = 0) =$$

$$= E(\xi | \eta = 1) \cdot P(\eta = 1) = E(\xi | \eta = 1) \cdot E(\eta),$$

azaz $E(\xi | \eta = 1) = \frac{E(\xi\eta)}{E(\eta)}$. Továbbá

$$E(\xi\eta) = E(\xi \cdot \chi_{\{B_t = c\}}) = E(E(\xi \cdot \chi_{\{B_t = c\}} | \xi)) = E(\xi \cdot P(B_t = c | \xi)) =$$

$$= \int_0^\infty x \cdot P(B_t = c | \xi = x) \cdot f_\xi(x) dx,$$

ahol ξ eloszlása megegyezik Λ eloszlásával, ezért $f_\xi \equiv f_\Lambda$. Másrészt - ξ és Λ eloszlásának egyezőségét és a teljes valószínűség tételét felhasználva -

$$E(\eta) = P(\eta = 1) = P(B_t = c) = \int_0^\infty P(B_t | \Lambda = x) \cdot f_\Lambda(x) dx.$$

Ezzel megkaptuk a kívánt egyenlőséget, hiszen a fentieket összevetve $E(\xi | \eta = 1) = \frac{E(\xi\eta)}{E(\eta)} = \lambda_t$. $\square$

Állítás 3.4.2 *$E\left(\frac{\sum_{n=1}^{N}\xi_n\eta_n}{\sum_{n=1}^{N}\eta_n}\right) = E(\xi|\eta=1)\cdot(1-(1-P(\eta=1))^N)$, azaz a $\Phi_N = \frac{\sum_{n=1}^{N}\xi_n\eta_n}{\sum_{n=1}^{N}\eta_n}$ becslés a feltételes várható értékre aszimptotikusan torzítatlan.*

Bizonyítás. Az állításban szereplő egyenlőség bal oldala

$$= \sum_{\substack{\eta_{i_1}=\ldots=\eta_{i_j}=1\\ \eta_{i_{j+1}}=\ldots=\eta_{i_N}=0}} E\left(\frac{\sum_{n=1}^{N}\xi_n\eta_n}{\sum_{n=1}^{N}\eta_n}\middle|\eta_{i_1}=1,\ldots,\eta_{i_j}=1,\eta_{i_{j+1}}=0,\ldots,\eta_{i_N}=0\right)\cdot$$

$$\cdot P(\eta_{i_1}=1,\ldots,\eta_{i_N}=0).$$

Itt az összegzés η_i-k összes lehetséges $0,1$ értékére történik, azaz 2^N értéket adunk össze. Megjegyezzük továbbá, hogy a pozitív valószínűséggel előforduló $\eta_1=\ldots=\eta_N=0$ eseményre a feltételes várható értékben szereplő $\frac{0}{0}$ hányados értékét 0-nak tekintjük, hiszen ekkor 0 darab ξ átlagát vesszük, így az átlag is 0. Ezt a 2^N-1 tagú összegzési feltételt - elhagyva az összes $\eta_i=0$ esetet - a továbbiakban az egyszerűség kedvéért R_η-val fogjuk jelölni. Ugyanis ki fogjuk használni, hogy létezik legalább egy η_i, ami nem 0. Az egyenlőséget folytatva

$$= \sum_{R_\eta} E\left(\frac{\xi_{i_1}+\ldots+\xi_{i_j}}{j}\middle|\eta_{i_1}=1,\ldots,\eta_{i_N}=0\right)\cdot P(\eta_{i_1}=1)\cdot\ldots\cdot P(\eta_{i_N}=0).$$

Felhasználtuk, hogy η_i-k egymástól független valószínűségi változók, így a $P(R_\eta)$ valószínűség szorzatra bomlik. Tovább

$$= \sum_{R_\eta}\left[E\left(\frac{\xi_{i_1}}{j}\middle|\eta_{i_1}=1\right)+\ldots+E\left(\frac{\xi_{i_j}}{j}\middle|\eta_{i_j}=1\right)\right]\cdot P(\eta_{i_1}=1)\cdot\ldots\cdot P(\eta_{i_N}=0),$$

ahol a szögletes zárójelben lévő feltételes várható értékek nem különböznek egymástól, ezért az egész felcserélhető egy tag j-szeresére, ezzel

$$= \sum_{R_\eta}\left[E\left(\xi_{i_1}|\eta_{i_1}=1\right)\right]\cdot P(\eta_{i_1}=1)\cdot\ldots\cdot P(\eta_{i_N}=0) =$$

$$= E(\xi_1|\eta_1=1)\cdot\sum_{R_\eta}\underbrace{P(\eta_{i_1}=1)}_{=:p}\cdot\ldots\cdot\underbrace{P(\eta_{i_N}=0)}_{=:1-p} =$$

$$= E(\xi_1|\eta_1=1)\cdot\sum_{j=1}^{N}\binom{N}{j}\cdot p^j\cdot(1-p)^{N-j} = E(\xi|\eta=1)\cdot(1-(1-p)^N).$$

A továbbiakban az $1-P(\eta=1)=P(\eta=0)$ valószínűséget q-val fogjuk jelölni. $\square$

Következmény 3.4.3 *Az előző két állításból azt kaptuk, hogy Φ_N várható értéke aszimptotikusan megegyezik Λ a posteriori várható értékével, λ_t-vel, amit szeretnénk meghatározni.*

Tehát a $\frac{\Phi_N}{1-q^N}$ sorozat várható értéke minden N-re megegyezik λ_t-vel. Alkalmazzuk a sorozatra a Csebisev-egyenlőtlenséget:

$$P\left(\left|\frac{\Phi_N}{1-q^N} - E\left(\frac{\Phi_N}{1-q^N}\right)\right| > \varepsilon\right) = P\left(\left|\frac{\Phi_n}{1-q^N} - \lambda_t\right| > \varepsilon\right) < \frac{D^2\left(\frac{\Phi_N}{1-q^N}\right)}{\varepsilon^2}$$

tetszőleges pozitív ε esetén. A felső korlátról szeretnénk mondani valamit, ezért számítsuk ki a számlálóban található szórásnégyzetet. Rögtön megjegyezzük, hogy elég Φ_N szórásnégyzetét kiszámítani, hiszen az $\frac{1}{1-q^N}$ konstans kiemelhető, így a számláló az $\frac{1}{(1-q^N)^2} \cdot D^2(\Phi_N)$ alakot ölti.

Lemma 3.4.4 $D^2(\Phi_N) = D^2(\xi|\eta = 1) \cdot J_{N,p} + E^2(\xi|\eta = 1) \cdot (1 - q^N)q^N$.

BIZONYÍTÁS. $E^2(\Phi_N) = (1 - q^N)^2 \cdot E^2(\xi|\eta = 1)$-et már meghatároztuk. Most vizsgáljuk a 2. momentumot:

$$E(\Phi_N^2) = E\left(\frac{[(\xi_1\eta_1)^2 + \ldots + (\xi_N\eta_N)^2] + 2 \cdot \sum_{k\neq l} [\xi_k\eta_k \cdot \xi_l\eta_l]}{\eta_1^2 + \ldots + \eta_N^2 + 2 \cdot \sum_{k\neq l} \eta_k\eta_l}\right) =$$

$$= \sum_{R_\eta} E\left(\frac{(\xi_1\eta_1)^2 + \ldots + (\xi_N\eta_N)^2 + 2 \cdot \sum_{k\neq l} \xi_k\eta_k \cdot \xi_l\eta_l}{j + 2 \cdot \binom{j}{2}}\right|$$

$$\left|\eta_{i_1} = \ldots = \eta_{i_j} = 1, \eta_{i_{j+1}} = \ldots = \eta_{i_N} = 0\right) \cdot$$

$$\cdot P(\eta_{i_1} = \ldots = \eta_{i_j} = 1, \eta_{i_{j+1}} = \ldots = \eta_{i_N} = 0)$$

A szorzat második tagjában az $\eta_{i_1} = \ldots = \eta_{i_j} = 1, \eta_{i_{j+1}} = \ldots = \eta_{i_N} = 0$ eseményt a rövidség kedvéért csak $\eta\ felt.$-ként fogjuk feltüntetni. Ezzel

$$= \sum_{R_\eta} \left[E\left(\frac{j \cdot \xi_1^2}{j^2}\middle|\eta_1 = 1\right) + \right.$$

$$\left. +2 \cdot E\left(\frac{\sum_{\substack{i_k \neq i_l \\ i_k, i_l \in \{i_1,\ldots,i_j\}}} \xi_{i_k}\xi_{i_l}}{j^2}\middle|\eta_{i_1} = \ldots = \eta_{i_j} = 1, \eta_{i_{j+1}} = \ldots = \eta_{i_N} = 0\right)\right] \cdot P(\eta\ felt.)$$

Ez utóbbi lépésben kihasználtuk, hogy ξ_i-k független, azonos eloszlásúak, így négyzeteik is, ezzel j darabot összeadva feltételesen ξ_1^2 j-szeresére cserélhetjük. Hasonlóan a második feltételes várható érték megegyezik $E\left(\frac{\binom{j}{2}}{j^2} \cdot \xi_1\xi_2 \middle| \eta_1 = \eta_2 = 1\right) =$

$\frac{j-1}{2j}E\left(\xi_1\xi_2\middle|\eta_1=\eta_2=1\right)$-gyel. Vigyázzunk, ugyanis a második feltételes várható értékben implicite feltettük, hogy létezik legalább két η_i, amelynek értéke nem 0, így erre gondolnunk kell az összegzéskor. Vegyük észre, hogy $j=1$ esetben nincs gond, hiszen $\frac{j-1}{2j}$-vel, azaz 0-val szorzódik $E(\xi_1\xi_2|\eta_1=\eta_2=1)$. Tehát

$$=\sum_{R_\eta}\left[\frac{1}{j}E(\xi_1^2|\eta_1=1)+\frac{j-1}{j}\cdot E(\xi_1\xi_2|\eta_1=\eta_2=1)\right]\cdot P(\eta\ felt.)=$$

$$=\sum_{R_\eta}\left[\underbrace{\frac{1}{j}E(\xi_1^2|\eta_1=1)-\frac{1}{j}E^2(\xi_1|\eta_1=1)}_{\frac{1}{j}\cdot D^2(\xi_1|\eta_1=1)}+E^2(\xi_1|\eta_1=1)\right]\cdot P(\eta\ felt.)=$$

$$=D^2(\xi|\eta=1)\cdot\sum_{j=1}^{N}\binom{N}{j}\frac{1}{j}\cdot p^j(1-p)^{N-j}+E^2(\xi|\eta=1)\cdot\sum_{j=1}^{N}\binom{N}{j}p^j(1-p)^{N-j}$$

Az előbb ismét kihasználtuk ξ_i-k függetlenségét, egyúttal korrelálatlanságát, ugyanis így $cov(\xi_1,\xi_2)=0$-ból adódóan

$$E(\xi_1\xi_2|\eta_1=\eta_2=1)=E(\xi_1|\eta_1=1)\cdot E(\xi_2|\eta_2=1)=E^2(\xi_1|\eta_1=1).$$

Bevezetve a $J_{N,p}:=\sum_{j=1}^{N}\binom{N}{j}\frac{1}{j}\cdot p^j(1-p)^{N-j}$ jelölést Φ_N szórásnégyzetére a következőt kaptuk:

$$D^2(\Phi_N)=D^2(\xi|\eta=1)\cdot J_{N,p}+E^2(\xi|\eta=1)\cdot(1-q^N-(1-q^N)^2)=$$
$$=D^2(\xi|\eta=1)\cdot J_{N,p}+E^2(\xi|\eta=1)\cdot(1-q^N)q^N.$$

□

Következmény 3.4.5

$$P\left(\left|\frac{\Phi_N}{1-q^N}-\lambda_t\right|>\varepsilon\right)<\frac{D^2(\xi|\eta=1)\cdot J_{N,p}+E^2(\xi|\eta=1)\cdot(1-q^N)q^N}{\varepsilon^2}.$$

Most vizsgáljuk meg a fenti $J_{N,p}$ összeget. Ez majdnem egy $X\sim Binom(N,p)$ eloszlású valószínűségi változó reciprokának várható értéke azzal a módosítással, hogy a 0 értéket nem engedjük meg. Megjegyezzük, hogy ekkor $(1-p)>0$ miatt a várható érték nem is létezne. Először anélkül, hogy explicite meghatároznánk $J_{N,p}$-t, ezen binomiális észrevétel alapján azon sejtésünket próbáljuk igazolni, mely szerint nagy N-re közel lesz $\frac{1}{Np}$-hez. A motiváció abban áll, hogy egy (N,p) paraméterű binomiális eloszlású valószínűségi változó várható értéke $N\cdot p$. Ezt a heurisztikus elképzelést a következőkben pontosítjuk.

Lemma 3.4.6 *Az* $\frac{1}{p(N+1)}-2q^N<J_{N,p}<\frac{2}{p(N+1)}$ *alsó és felső becslés érvényes* $J_{N,p}$*-re. Megj.: Az alsó becslés csak bizonyos, később látott határtól véve igaz.*

Bizonyítás. Az alsó becslés

$$J_{N,p} = \frac{1}{N+1}\cdot\sum_{j=1}^{N}\binom{N}{j}\frac{N+1}{j}p^j(1-p)^{N-j} \geq \frac{1}{N+1}\cdot\sum_{j=1}^{N}\binom{N}{j}\frac{N+1}{j+1}p^j(1-p)^{N-j} =$$

$$= \frac{1}{N+1}\cdot\sum_{j=1}^{N}\binom{N+1}{j+1}p^j(1-p)^{N-j} = \frac{1}{p(N+1)}\cdot\sum_{j=2}^{N+1}\binom{N+1}{j}p^j(1-p)^{N+1-j} =$$

$$= \frac{1}{p(N+1)}\cdot\left(1-(1-p)^{N+1}-(N+1)p(1-p)^N\right) = \frac{1}{p(N+1)} - \frac{q^{N+1}}{p(N+1)} - q^N,$$

ahol - a szokásos $1-p=q$ jelöléssel - $\frac{q^{N+1}}{p(N+1)} + q^N = q^N\cdot\left(\frac{q+p+Np}{p(N+1)}\right) = q^N\left(\frac{1+Np}{p+Np}\right) < 2q^N$, ha $\frac{1+Np}{p+Np} < 2$, azaz ha $N > \frac{1-2p}{p}$. Ezzel $J_{N,p} > \frac{1}{p(N+1)} - 2q^N$.

A felső becslésnél felhasználjuk az $\frac{1}{j} = \frac{1}{j(j+1)} + \frac{1}{j+1}$ azonosságot:

$$J_{N,p} = \sum_{j=1}^{N}\binom{N}{j}\frac{1}{j(j+1)}p^j(1-p)^{N-j} + \sum_{j=1}^{N}\binom{N}{j}\frac{1}{j+1}p^j(1-p)^{N-j} =$$

$$= \frac{1}{N+1}\cdot\sum_{j=1}^{N}\binom{N+1}{j+1}\cdot\frac{1}{j}p^j(1-p)^{N-j} + \frac{1}{N+1}\cdot\sum_{j=1}^{N}\binom{N+1}{j+1}p^j(1-p)^{N-j} =$$

$$= \frac{1}{p(N+1)}\cdot\sum_{j=2}^{N+1}\binom{N+1}{j}\cdot\underbrace{\left(1+\frac{1}{j-1}\right)}_{\leq 2}p^j(1-p)^{N+1-j} \leq$$

$$\leq \frac{2}{p(N+1)}\cdot\sum_{j=2}^{N+1}\binom{N+1}{j}p^j(1-p)^{N+1-j} < \frac{2}{p(N+1)}.$$

□

Ennek felhasználásával egyszerűen javíthatunk felső becslésen, mégpedig felhasználva az $\frac{1}{j-1} \leq \frac{3}{j}$ összefüggést $j \geq 2$ esetén.

$$J_{N,p} = \sum_{j=1}^{N}\binom{N}{j}\cdot\frac{1}{j+1}\cdot p^j(1-p)^{N-j} + \sum_{j=1}^{N}\binom{N}{j}\frac{1}{j(j+1)}\cdot p^j(1-p)^{N-j} \leq$$

$$\leq \frac{1}{p(N+1)} + \sum_{j=1}^{N}\frac{1}{p(N+1)}\binom{N+1}{j+1}\frac{1}{j}p^{j+1}(1-p)^{(N+1)-(j+1)} =$$

$$= \frac{1}{p(N+1)} + \frac{1}{p(N+1)}\cdot\sum_{j=2}^{N+1}\binom{N+1}{j}\frac{1}{j-1}p^j(1-p)^{N+1-j} \leq$$

$$\leq \frac{1}{p(N+1)}\cdot\left(1+3\cdot\sum_{j=2}^{N+1}\binom{N+1}{j}\frac{1}{j}p^j(1-p)^{N+1-j}\right) \leq \frac{1}{p(N+1)}\left(1+3\cdot\frac{2}{Np}\right) =$$

$$= \frac{1}{p(N+1)} + \frac{6}{N(N+1)p^2}.$$

Következmény 3.4.7 *Összefoglalva azt kaptuk, hogy tetszőleges* $0 < p < 1$ *esetén* $\frac{1}{(N+1)p} - 2q^N < J_{N,p} < \frac{1}{(N+1)p} + \frac{6}{N(N+1)p^2}$ *(az alsó határra vonatkozóan* $N > \frac{1}{p} - 2$ *esetén). Más szóval* $(N+1)p \cdot J_{N,p} \overset{N\to\infty}{\longrightarrow} 1$, *azaz a* $J_{N,p}$ *összeg aszimptotikusan* $\frac{1}{Np}$ *nagyságrendű, amit igazolni akartunk.*

A fentiek ismeretében szeretnénk adott ε-ra olyan lehető legkisebb N_0-t választani, hogy az eltérésre vonatkozó valószínűség kisebb legyen egy szintén általunk megadott $\delta > 0$ értéknél. Tehát hány (ξ_n, η_n) párt kell generálnunk annak érdekében, hogy a becslésnek az a posteriori eloszlástól vett eltérése abszolútértékben ne haladja meg ezen ε értéket, csak legfeljebb egy kis δ valószínűséggel.

Mielőtt nekikezdenénk, említsünk meg egy lemmát, aminek a későbbiekben hasznát vesszük:

Lemma 3.4.8 $J_{N,p} = q^N \cdot \sum\limits_{n=1}^{N} \frac{1-q^n}{nq^n}$.

BIZONYÍTÁS.

$$\sum_{j=1}^{N} \binom{N}{j} \frac{1}{j} p^j (1-p)^{N-j} = \sum_{j=1}^{N} \left[\sum_{n=j}^{N} \binom{n-1}{j-1} \right] \frac{1}{j} p^j (1-p)^{N-j} =$$

$$= \sum_{n=1}^{N} \sum_{j=1}^{n} \binom{n-1}{j-1} \frac{1}{j} p^j (1-p)^{N-j} = \sum_{n=1}^{N} \sum_{j=1}^{N} \binom{n}{j} \frac{1}{n} p^j (1-p)^{N-j} =$$

$$= \sum_{n=1}^{N} \frac{1}{n} \sum_{j=1}^{n} \binom{n}{j} p^j (1-p)^{N-j} = \sum_{n=1}^{N} \frac{1}{n} (1-p)^{N-n} \sum_{j=1}^{n} \binom{n}{j} p^j (1-p)^{n-j} =$$

$$= \sum_{n=1}^{N} \frac{1}{n} (1-p)^{N-n} (1-(1-p)^n) = (1-p)^N \cdot \sum_{n=1}^{N} \frac{1}{n} \cdot \frac{1-(1-p)^n}{(1-p)^n} =$$

$$\overset{1-p=q}{=} q^N \cdot \sum_{n=1}^{N} \frac{1}{n} \cdot \frac{1-q^n}{q^n},$$

ami az állítást igazolja. □

A könnyebb átláthatóság kedvéért vezessük be a $D^2 := D^2(\xi|\eta = 1)$ és $E^2 := E^2(\xi|\eta = 1)$ jelöléseket. A fenti felső becslésre a $\frac{D^2 \cdot J_{N,p} + E^2 \cdot (1-q^N) q^N}{\varepsilon^2 (1-q^N)^2} =$

$$(*) = \frac{D^2 \cdot J_{N,p}}{\varepsilon^2 (1-q^N)^2} + \frac{E^2 q^N}{\varepsilon^2 (1-q^N)} = \leq \delta$$

egyenlőtlenséget kell megoldanunk N-re.

$$(*) \overset{(1)}{\leq} \frac{4D^2}{\varepsilon^2} \cdot J_{N,p} + \frac{2E^2}{\varepsilon^2} \cdot q^N \overset{(2)}{\leq} \frac{4D^2 + 2E^2}{\varepsilon^2} \cdot J_{N,p} \overset{(3)}{\leq}$$

$$\leq \frac{4D^2 + 2E^2}{\varepsilon^2} \cdot \left(\frac{1}{Np} + \frac{6}{N^2 p^2} \right) \leq \delta.$$

Ahol kihasználtuk a következőket:

(1) Igaz, ha $q^N \leq \frac{1}{2}$, ami pontosan akkor teljesül, ha $2 \leq (\frac{1}{q})^n$, azaz $N \geq \log_{\frac{1}{q}} 2 = \frac{\log 2}{\log \frac{1}{q}}$.

(2) Ehhez arra van szükségünk, hogy $J_{N,p} \geq q^N$ legyen, ami valamely N-re biztosan bekövetkezik, hiszen ez előbbi $\frac{1}{N}$, míg utóbbi exponenciális nagyságrendben tart 0-hoz. A lemma alapján tudjuk, hogy elegendő $q^N \sum_{n=1}^{N} \frac{1-q^n}{nq^n} > q^N$-et belátnunk. Leosztva mindkét oldalt q^N-nel és kicsit átrendezve kapjuk, hogy $\sum_{n=1}^{N} \frac{1}{nq^n} > 1 + \sum_{n=1}^{N} \frac{1}{n}$ kell. A bal oldal nyilván nem nő, ha az összegzésben az $\frac{1}{n}$ tagokat végig $\frac{1}{N}$-re változtatjuk, ezzel a könnyen összegezhető $\left(\frac{1}{q}\right)^n$ tagok maradnak. Ezzel a bal oldal nem kisebb, mint $\frac{1}{N} \cdot \sum_{n=1}^{N} \left(\frac{1}{q}\right)^n = \frac{1}{N} \cdot \frac{1-q^N}{q^N(1-q)}$. Ha ez nagyobb $1 + \log(N+1)$-nél, akkor a kívánt egyenlőtlenség teljesül, ugyanis $\log(N+1)$ felülről becsüli a pozitív egészek reciprokösszegét N-ig. Tehát elegendő olyan N-et választani, amelyre már teljesül az $\frac{1}{N} \cdot \frac{1-q^N}{q^N(1-q)} > 1 + \log(N+1)$ egyenlőtlenség. Ez az egyenlőtlenség már jól kezelhető, q ismeretében N alsó korlátját megadhatjuk ezen feltétel mellett. Megjegyezzük azonban, hogy a gyakorlati kérdések többségében a $q = 0,9999$ értéknél nagyobb valószínűség nem jellemző, hogy előforduljon, és ekkor az $N = 40\ 000$ érték kielégítőnek bizonyul.

(3) $J_{N,p}$ becslésében csak növelünk, ha $N+1$ helyett mindenhol N-et írunk.

Ezt N szerint másodfokú egyenletté rendezve

$$\delta \cdot N^2 - \frac{4D^2 + 2E^2}{\varepsilon^2 p} \cdot N - \frac{6}{p^2} = 0 \ -t$$

kell megoldanunk, annak is a nagyobb gyökét keressük. Tehát legyen

$$N' := \frac{\frac{4D^2+2E^2}{\varepsilon^2 p} + \sqrt{\left(\frac{4D^2+2E^2}{\varepsilon^2 p}\right)^2 + \frac{4\cdot\delta\cdot 6}{p^2}}}{2\delta} = \frac{4D^2 + 2E^2 + \sqrt{(4D^2+2E^2)^2 + 24\delta\varepsilon^2}}{2\varepsilon^2 p\delta} <$$

$$< \frac{(4D^2+2E^2) + (4D^2+2E^2) + \varepsilon\sqrt{24\delta}}{2\varepsilon^2 p\delta} = \frac{4D^2+2E^2}{\varepsilon^2 p\delta} + \frac{\sqrt{6\delta}}{\varepsilon p\delta} <$$

Most használjuk a $E(\xi^2|\eta = 1) = D^2(\xi|\eta = 1) + E^2(\xi|\eta = 1)$ összefüggést, amellyel $2E^2$ helyett a kétszeresét véve felülről becsülhetjük:

$$< \frac{4 \cdot E(\xi^2|\eta = 1)}{\varepsilon^2 p\delta} + \frac{\sqrt{6}}{\varepsilon p\sqrt{\delta}}$$

Még nem vagyunk készen, hiszen a becslésben szereplő feltételes várható értékről nem mondtunk semmit. Erre a következő felső becslést adhatjuk:

$$E(\xi^2|\eta = 1) = \frac{E(\xi^2\eta^2)}{P(\eta = 1)} \leq \frac{E(\xi^2)}{P(\eta = 1)} = \frac{\frac{\alpha}{\beta^2} + \frac{\alpha^2}{\beta^2}}{P(\eta = 1)} = \frac{\alpha(1+\alpha)}{\beta^2 \cdot p},$$

amit a $\xi \sim \Gamma(\alpha, \beta)$ alapján ismertünk. Az előbbieket összefoglalva N' értékére a

$$\left\lceil \frac{4\alpha(1+\alpha)}{\varepsilon^2\delta\beta^2 p^2} + \frac{\sqrt{6}}{\varepsilon p\sqrt{\delta}} \right\rceil$$

biztosan megfelelő határ lesz. Egy dologgal persze adósak maradtunk, mivel a p értékről valójában nem tudjuk, hogy pontosan mennyi, azt viszont tudjuk, hogy η várható értékével egyenlő. A nagy számok erős törvényének értelmében $\frac{\sum_{n=1}^{N} \eta_n}{N}$ tart a várható értékhez 1 valószínűséggel és L^1-ben. Ez a konvergencia ráadásul exponenciális sebességű:

A Hoeffding-egyenlőtlenséget használjuk:

Tétel 3.4.9 (Hoeffding-egyenlőtlenség) *Legyenek $X_1, X_2, \ldots, X_n$ független, azonos eloszlású valószínűségi változók, melyek mindegyikére $X_i \in [a_i, b_i]$ 1 valószínűséggel ($a_i < b_i$-k valós korlátok). Az $S_n = \sum_{i=1}^{n} X_i$ jelölés mellett tetszőleges $\varepsilon > 0$-ra*

$$P\left(|S_n - ES_n| \geq \varepsilon\right) \leq 2 \cdot e^{-\frac{2\varepsilon^2}{\sum_{i=1}^{n}(b_i - a_i)^2}}.$$

Ezt felírva X_i helyett η_n szereposztással minden $\varepsilon' > 0$-ra

$$P\left(\left|\frac{\sum_{n=1}^{N'} \eta_n}{N'} - E(\eta_1)\right| \geq \varepsilon'\right) = P\left(\left|\sum_{n=1}^{N'} \eta_n - N' \cdot E(\eta_1)\right| \geq \varepsilon' \cdot N'\right) =$$

$$\leq 2 \cdot e^{-\frac{2 \cdot N'^2 \varepsilon'^2}{N'}} = 2 \cdot e^{-N' \cdot 2\varepsilon'^2}.$$

Így alkalmasan választott N-re p-nek jó közelítését kapjuk, és ezt használjuk az előbbi becsléshez. Ilyen feltételek mellett megválasztott N_0-ra, N'-re és p-re azt mondhatjuk, hogy

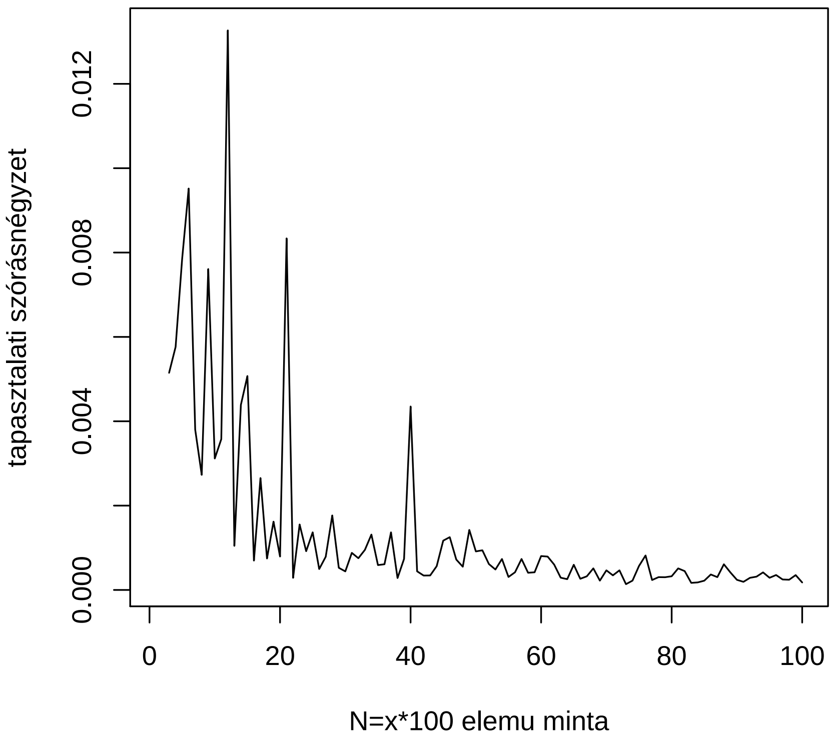
$$P\left(P\left(\left|\frac{\sum_{n=1}^{N} \xi_n \eta_n}{\sum_{n=1}^{N} \eta_n} - \lambda_t\right| > \varepsilon\right) < \delta\right) \geq 1 - \varepsilon^*.$$

Példaként készítsünk egy táblázatot, amelyben az i-edik sor j-edik eleme azt mondja meg, hogy amennyiben egy - nálunk - új szerződő $9 + i$ évet töltött el, és ezalatt C_j állapotba került, mi lesz a rá vonatkozó kárgyakoriság becslése. Ezzel 10-től 28 évig vesszük t értékét. Ehhez be kell állítanunk az állományból származtatott α, β paramétereket, melyekre példánkban az $\alpha = 1,7$ és $\beta = 18$ feltételezéssel élünk. Így a táblázat a következő lesz: 3.2 (NA ismét a lehetetlen események helyén áll).

3.2. táblázat. Bayes-féle becslések a kárgyakoriságra

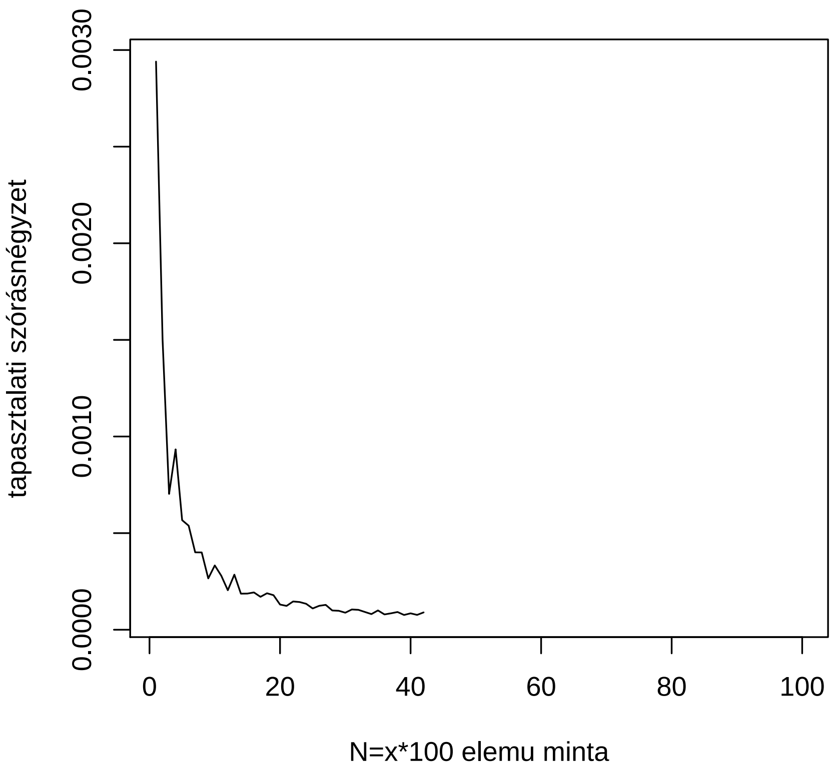
t\c	M4	M3	M2	M1	A0	B1	B2	B3	B4	B5	B6	B7	B8	B9	B10
10	0.274	0.259	0.216	0.205	0.214	0.169	0.170	0.167	0.131	0.129	NaN	0.096	NaN	NaN	0.061
11	0.273	0.251	0.252	0.206	0.211	0.207	0.162	0.161	0.174	0.128	0.126	NaN	0.093	NaN	0.059
12	0.288	0.256	0.234	0.235	0.196	0.195	0.192	0.159	0.156	0.165	0.122	0.122	0.094	0.089	0.057
13	0.286	0.265	0.238	0.223	0.228	0.187	0.186	0.203	0.150	0.153	0.132	0.121	0.100	0.084	0.067
14	0.299	0.275	0.267	0.225	0.216	0.223	0.181	0.181	0.187	0.149	0.143	0.119	0.111	0.095	0.066
15	0.316	0.269	0.279	0.248	0.208	0.207	0.203	0.173	0.173	0.162	0.141	0.138	0.100	0.108	0.064
16	0.297	0.289	0.258	0.245	0.241	0.205	0.203	0.206	0.169	0.166	0.144	0.138	0.106	0.094	0.071
17	0.305	0.280	0.272	0.243	0.227	0.236	0.197	0.198	0.184	0.162	0.152	0.131	0.123	0.106	0.069
18	0.327	0.275	0.266	0.261	0.231	0.223	0.226	0.191	0.189	0.172	0.157	0.145	0.105	0.118	0.068
19	0.316	0.298	0.271	0.255	0.243	0.210	0.217	0.210	0.184	0.180	0.151	0.151	0.114	0.104	0.073
20	0.311	0.288	0.276	0.254	0.239	0.241	0.210	0.204	0.194	0.179	0.161	0.143	0.126	0.111	0.072
21	0.327	0.296	0.271	0.255	0.243	0.226	0.224	0.204	0.199	0.183	0.164	0.154	0.111	0.124	0.072
22	0.325	0.302	0.281	0.265	0.260	0.228	0.225	0.223	0.194	0.187	0.155	0.159	0.120	0.106	0.074
23	0.327	0.302	0.296	0.266	0.246	0.248	0.226	0.215	0.203	0.185	0.170	0.148	0.125	0.113	0.074
24	0.322	0.310	0.287	0.272	0.246	0.235	0.237	0.216	0.203	0.193	0.169	0.157	0.115	0.122	0.073
25	0.331	0.322	0.296	0.280	0.263	0.240	0.227	0.223	0.203	0.193	0.158	0.164	0.121	0.110	0.075
26	0.333	0.315	0.297	0.284	0.256	0.254	0.232	0.222	0.210	0.193	0.171	0.147	0.124	0.117	0.075
27	0.344	0.321	0.293	0.279	0.256	0.252	0.239	0.228	0.202	0.195	0.173	0.161	0.116	0.122	0.075
28	0.345	0.324	0.302	0.284	0.271	0.261	0.241	0.231	0.212	0.194	0.163	0.160	0.123	0.115	0.075

A módszer hátulütője, hogy alacsonyabb osztályok esetén az alacsony $\frac{\alpha}{\beta}$ várható érték miatt nagy lesz Φ_N szórásnégyzete, ugyanis kevés η_n veszi fel az 1 értéket. Ez a probléma magasabb osztályoknál már nem jelentkezik, hiszen pontosan arról van szó, hogy az a priori eloszlásból generált kárgyakoriságok túlnyomó többségben a jó sofőrökre jellemzőek, tehát ott sok η_n lesz 1. Ennek egy lehetséges kiküszöböléséről a későbbiekben ejtünk szót, most lássunk két példát a tapasztalati szórásnégyzet alakulására N mintaelemszám függvényében. A 3.1 ábrán az $M4$ osztályhoz és $t = 15$ évhez tartozó becslés tapasztalai szórásnégyzetét ábrázoltuk, ahol az x tengely beosztásán látható érték 100-szorosának megfelelő mintaelemszámmal dolgoztunk. A 3.2 ábrán ugyanezen év mellett, csak a $B9$ osztályhoz tartozó tapasztalati szórásnégyzetek láthatóak $N = 100 \cdot x$ függvényében. Ez utóbbi abban különbözik az előbbitől, hogy a függőleges tengely tartományát negyedére csökkentettük. Látható, hogy a $B9$ esetben $N = 2800$-tól már 10^{-4} alatti lesz minden érték, majd $N = 4500$-tól a gép számára kezelhetetlenül kicsi szórásnégyzeteket kapunk. ($\alpha = 1,7$ és $\beta = 18$)



3.1. ábra. Φ_N tapasztalati szórásnégyzete $t = 15$ és $c = M4$ esetén

Ugyanezt elvégezhetjük minden osztályra, és azt tapasztaljuk várakozásunknak megfelelően, hogy minél jobb bónuszosztályt veszünk, annál kevesebb minta generálása elég a becslés stabilizálásához. Ne felejtsük el, hogy ezek Φ_N-re vonatkozó



3.2. ábra. Φ_N tapasztalati szórásnégyzete $t = 15$ és $c = B9$ esetén

szórásnégyzetek, a torzítatlan becslésünk a paraméterekre pedig $\frac{\Phi_N}{1-q^N}$ volt, tehát ezt mindig figyelembe kell vennünk a gyakorlati számításoknál.

3.4.10 Megjegyzés A szórásnégyzeteket a $D^2(\Phi_N) = D^2(\xi|\eta = 1) \cdot J_{N,p} + + E^2(\xi|\eta = 1) \cdot (1 - q^N) q^N$ összefüggés felhasználásával becsültük a következőképpen:

$$\hat{D^2}(\Phi_N) =$$

$$\frac{\sum\limits_{n=1}^{N} \eta_n}{\sum\limits_{n=1}^{N} \eta_n - 1} \left[\frac{\sum\limits_{n=1}^{N} \xi_n^2 \eta_n^2}{\sum\limits_{n=1}^{N} \eta_n^2} - \left(\frac{\sum\limits_{n=1}^{N} \xi_n \eta_n}{\sum\limits_{n=1}^{N} \eta_n} \right)^2 \right] \cdot \hat{J}_{N,p} + \left(\frac{\sum\limits_{n=1}^{N} \xi_n^2 \eta_n^2}{\sum\limits_{n=1}^{N} \eta_n^2} \right)^2 \cdot \Big(1 - (1-\hat{p})^N\Big) \cdot (1-\hat{p})^N,$$

ahol a jobb oldal első törtje korrigáló tényező, másrészt $\hat{p} = \frac{\sum_{n=1}^{N} \eta_n}{N}$, és $\hat{J}_{N,p} = J_{N,\hat{p}}$.

3.4.11 Példa Lássunk egy példát az előzőekben hosszan ismertetett módszer szerinti becslésre. R-program segítségével generáltunk egy 5000 elemű homogén kárszámmintát, melynek segítségével az a priori paramétereket becsültük meg. Mindezt tettük oly módon, hogy legelőször előzetesen megadott α, β értékek által karakterizált Γ-eloszlásból származtattunk 5000 elemű mintát, azaz a személyre szabott kárgyakoriságok paraméterét. Majd ezeket a $\lambda_1, \ldots, \lambda_{5000}$ értékeket felhasználva készítettük el az 5000 elemű Poisson-eloszlású mintát, melyben minden egyes érték a hozzá tartozó $Poisson(\lambda_i)$ eloszlásból lett mintavételezve ($i = 1, \dots, 5000$). Ebből a maximum likelihood módszer mellett döntve megkaptuk az ismeretlen $\hat{\alpha}, \hat{\beta}$ paraméterbecsléseket. Ne felejtsük el, hogy a Gamma-eloszlás β skálaparamétere és a negatív binomiálisnál kapott valószínűség közti összefüggés $\frac{\beta}{1+\beta} = p$, azaz a kapott p-ből $\beta = \frac{p}{1-p}$. Ezzel esetünkben az $\alpha = 1,7$ és $\beta = 18$ értékekre jutottunk, melyeket egyébként mindvégig példaként használtunk.

Ezután az előző mintát magunk mögött hagyva vettünk újabb, most 100 elemű mintát a $\Gamma(1.7, 18)$-eloszlásból. Az a feltételezésünk, hogy 100 új szerződő vált a mi biztosítónkhoz, akik mind 12 éve biztosítottak a magyar gépjármű-felelősségbiztosítás keretein belül, és ezek az ő paramétereik személyre szabva, melyeket természetesen nem ismerünk. Majd szimuláltuk 12 év elteltét mindenkinél Markov-láncokkal, melynek következtében a 100 új szerződő a következő arányban került a különböző osztályokba:

M4	M3	M2	M1	A0	B1	B2	B3	B4	B5	B6	B7	B8	B9	B10
0	0	0	0	0	0	0	2	1	0	18	2	2	43	32

Ennek ismeretében minden egyes ügyfélhez elkészíthető a paraméterének $\{B_t = c\}$ feltétel szerinti becslése, melyet összesítve készítünk, hiszen kíváncsiak vagyunk, hogy erre a 100 fős ügyfélkörre milyen eltérés lesz a valóság és becslésünk között. A $t = 12$ év mellett minden oszálynak kiolvasható a táblázatunkból a hozzá tartozó a posteriori kárgyakoriság, így ezeket csak összeszorozzuk a fenti létszámokkal, majd összegzünk. Ezzel a példában a következő évre $8,7$ kárt jósoltunk, majd mindenki paraméterével ismét generáltunk Poisson-eloszlású mintát, amik pontosan az elkövetkezendő év kárszámait hivatottak reprezentálni. Azt kaptuk, hogy 92 szerződő nem okozott kárt, és 8-an egyet, így 1-nél többet nem regisztráltunk. Láthatjuk, hogy ebben az esetben ezen kevés információ is megfelelő kiindulópontnak bizonyult.

3.4.12 Példa Az előzőhöz hasonlóan nézzünk még egy esetet, most legyen 500 új szerződőnk 16 éves háttérrel. Az osztályok megoszlása a következő:

M4	M3	M2	M1	A0	B1	B2	B3	B4	B5	B6	B7	B8	B9	B10
0	1	2	2	6	4	11	4	20	12	17	71	33	24	293

A szimulált következő évben 457-en semmi, 38-an egy és 5-en két kárt okoztak, azaz összesen 48 kár keletkezett. Emellett a becslésünk erre a csoportra 51.

3.5. Becslések ismert előző évi kárszám esetén

Új irányban vizsgálódva tegyük fel, hogy az előzőekben látottakkal ellentétben más információ áll a rendelkezésünkre. Nevezetesen a biztosított előző évi kárszáma, melyet k-val jelölünk, és természetesen a jelenlegi osztálya. Ha kétséges lenne a kronológia, szemléljük úgy a helyzetet, hogy éppen ebben a pillanatban váltott biztosítót az illető személy, és a múlt évi produkciójának történetét felhasználva próbálunk következtetni az ideire. Nem ismerjük azonban, hogy hány évet töltött el a bónusz-málusz rendszerben, azaz t ismeretlen. Megjegyezzük, hogy nyilvánvaló következtetéseket természetesen mindig le lehet vonni. Ilyen például, ha B8-ba vettük át a szerződőt, természetesen nem lehet 2 éve biztosítása stb. Ezekről a kizárható eseményekről a maximum likelihood becslésnél látott táblázat tanúskodik.

A teljesség kedvéért megemlítjük a maximum likelihood módszerrel történő becslést. Az előző, t. évi kárszámot az η_t valószínűségi változó jelöli, melynek értékéről ismert, hogy k. Ezzel a $P(\eta_t = k) = \frac{\lambda^k}{k!}e^{-\lambda}$ valószínűséget kell maximalizálnunk λ-ban, azaz $\log \lambda^k - \log k! + \log e^{-\lambda}$ deriváltja $\frac{k}{\lambda} - 1 = 0$, azaz $\hat{\lambda} = k$.

Bayes-féle becsléssel ez a következőképpen néz ki. Legyen Λ $\Gamma(\alpha, \beta)$ eloszlású valószínűségi változó, azaz sűrűségfüggvénye $f_\Lambda(s) = \frac{\beta^\alpha s^{\alpha-1}e^{-\beta s}}{\Gamma(\alpha)}$, melynek tartója a pozitív félegyenes és $P(X_t = k|\Lambda = \lambda) = \frac{\lambda^k}{k!}e^{-\lambda}$. A becslést ezúttal $\hat{\lambda}$ fogja jelölni. Keressük a $\hat{\lambda} = E(\Lambda|\eta_t = k)$ a posteriori várható értéket. Ezúttal lényegesen könnyebb dolgunk van, hiszen λ függvényében explicit előállítást ismerünk a feltételes valószínűségekre, így a Bayes-tétel felhasználásával a feltételes sűrűségfüggvény

$$f_{\Lambda|\eta_t}(s|k) = \frac{P(\eta_t = k|\Lambda = s) \cdot f_\Lambda(s)}{P(\eta_t = k)} =$$

$$= \frac{\frac{s^k}{k!}e^{-s}\beta^\alpha s^{\alpha-1}e^{-\beta s}}{P(\eta_t = k)} = s^{k+\alpha-1} \cdot e^{-(1+\beta)s} \cdot \frac{\beta^\alpha}{k!P(\eta_t = k)}.$$

Ennek alakjából rögtön következtethetünk, hogy $\Gamma(k+\alpha, 1+\beta)$ eloszláshoz tartozó a sűrűségfüggvény, ugyanis egy (-1)-nél nagyobb hatványra emelt s, egy skálaparaméteres exponenciális tag, valamint egy konstans szorzatából tevődik össze. Tehát megvan a konkrét eloszlás a paramétereivel, melynek várható értéke $\frac{\alpha+k}{\beta+1}$.

3.5.1 Megjegyzés Ugyanezt igazolhattuk volna közvetlen számolással is:

$$E(\lambda|\eta_t = k) = \int_0^\infty s \cdot s^{\alpha+k-1} \cdot e^{-(1+\beta)s} \cdot \frac{\beta^\alpha}{k!P(\eta_t = k)} ds =$$

$$\frac{\beta^\alpha \cdot \Gamma(\alpha+k+1)}{(1+\beta)^{\alpha+k+1} \cdot k! \cdot P(\eta_t = k)} \cdot \underbrace{\int_0^\infty \frac{s^{(\alpha+k+1)-1} \cdot e^{-(1+\beta)s} \cdot (1+\beta)^{k+\alpha+1}}{\Gamma(\alpha+k+1)} ds}_{=1} =$$

$$= \frac{\alpha+k}{\beta+1} \cdot \underbrace{\frac{\beta^\alpha \cdot \Gamma(\alpha+k)}{(1+\beta)^{\alpha+k} \cdot k! \cdot P(\eta_t = k)}}_{(*)} = \frac{\alpha+k}{\beta+1},$$

ugyanis

$$P(\eta_t = k) = \int_0^\infty \frac{s^k}{k!} e^{-s} \beta^\alpha s^{\alpha-1} e^{-\beta s} ds =$$

$$= \frac{\beta^\alpha}{k!} \cdot \frac{\Gamma(\alpha+k)}{(1+\beta)^{\alpha+k}} \cdot \underbrace{\int_0^\infty \frac{s^{\alpha+k-1} \cdot e^{-(1+\beta)s} \cdot (1+\beta)^{\alpha+k}}{\Gamma(\alpha+k)} ds}_{=1}.$$

Teljesen hasonlóan látható be általánosan, hogy az előző néhány évi kártörténet ismeretében a becslés $\frac{\alpha+\eta_1+\ldots+\eta_t}{\beta+t}$ lesz, ahol $\eta_1, \ldots, \eta_t$ az előző t év kárszámai (ez nem feltétlenül jelenti azt, hogy t éve van biztosítása a szerződőnek).

3.6. Összehasonlítás

Felmerül a kérdés, hogy vajon az eltelt évek és a legutolsó állapot - ezt jelöljük (1)-gyel -, vagy az utolsó évi kárszám - (2) - alapján lehet-e jobb becslést adni az illető λ paraméterére. Ezen kérdés megválaszolásához gondoljunk arra, hogy az előzőekben olyan $\lambda \in \mathbb{R}_0^+$ meghatározásán fáradoztunk, amely minimalizálja az $\int_0^\infty (\lambda_t - \lambda)^2 dQ(\lambda|\text{feltétel})$-t, ahol a feltétel helyére az előbbi két információ valamelyike helyettesíthető. Más szóval négyzetes veszteségfüggvénnyel dolgozva veszünk várható értéket a feltételes eloszlás szerint. Erről könnyen láthatjuk, hogy a $Q(\lambda|\text{feltétel})$ eloszlás szerint vett szórásnégyzete λ-nak, hiszen $\lambda_t = E(\lambda|\text{feltétel})$. A (2) esetben ez explicite ismert, hiszen a $\Gamma(\alpha+k, \beta+1)$ valószínűségi változók szórásnégyzete $\frac{\alpha+k}{(1+\beta)^2}$

3.6.1 Megjegyzés Ha szeretnénk kiszámolni, a következőképpen járunk el:

$$\int_0^\infty (\lambda_t - \lambda)^2 dQ(\lambda|X_t = k) = \int_0^\infty \left(\frac{\alpha+k}{\beta+1} - \lambda\right)^2 \cdot \frac{(1+\beta)^{\alpha+k} \cdot \lambda^{\alpha+k-1} \cdot e^{-(1+\beta)\lambda}}{\Gamma(\alpha+k)} d\lambda =$$

$$= \int_0^\infty \frac{(\alpha+k)^2 - 2(\alpha+k)(1+\beta)\lambda + (1+\beta)^2\lambda^2}{(1+\beta)^2} \cdot \frac{(1+\beta)^{\alpha+k} \cdot \lambda^{\alpha+k-1} e^{-(1+\beta)\lambda}}{\Gamma(\alpha+k)} d\lambda =$$

$$= \int_0^\infty \frac{1}{\Gamma(\alpha+k)} \Big[(\alpha+k)^2 \cdot (1+\beta)^{\alpha+k-2} \cdot \lambda^{\alpha+k-1} \cdot e^{-(1+\beta)\lambda} -$$

$$-2(\alpha+k)(1+\beta)^{\alpha+k-1}\lambda^{\alpha+k} \cdot e^{-(1+\beta)\lambda} + (1+\beta)^{\alpha+k}\lambda^{\alpha+k-1} e^{-(1+\beta)\lambda}\Big] d\lambda =$$

$$= \frac{(\alpha+k)^2}{(1+\beta)^2} \cdot \int f_{\Gamma(\alpha+k,1+\beta)} - 2 \cdot \frac{(\alpha+k)\Gamma(\alpha+k+1)}{(1+\beta)^2\Gamma(\alpha+k)} \int f_{\Gamma(\alpha+k+1,1+\beta)} +$$

$$+\frac{\Gamma(\alpha+k+2)}{(1+\beta)^2\Gamma(\alpha+k)}\cdot\int f_{\Gamma(\alpha+k+2,1+\beta)},$$

ahol az integrálok értéke 1, hiszen sűrűségfüggvényeket integrálunk a teljes tartón. Innen átrendezéssel és az $\alpha\Gamma(\alpha)=\Gamma(\alpha+1)$ azonosság felhasználásával megkapjuk $\frac{\alpha+k}{(1+\beta)^2}$-et.

Az (1) esetben pedig láttuk, hogy a $D_1^2 := D^2(\lambda|B_t=c)$ feltételes szórásnégyzetre nincs explicit formulánk, ezért csak közelíteni tudjuk. Annak eldöntéséhez, hogy melyik becslést tekinthetjük jobbnak a másiknál, el kell döntenünk, hogy az adott esetben a $D_1^2 \leq D_2^2$, agy a $D_2^2 \leq D_1^2$ egyenlőtlenség teljesül-e. Ahol kisebb a szórásnégyzet, azt jobbnak mondjuk. A közelítést a már látott feltételes várható értékhez hasonlóan végezzük, most

$$\frac{\sum\limits_{n=1}^{N}\xi_n^2\eta_n}{\sum\limits_{n=1}^{N}\eta_n}-\left(\frac{\sum\limits_{n=1}^{N}\xi_n\eta_n}{\sum\limits_{n=1}^{N}\eta_n}\right)^2\xrightarrow{p}D_1^2.$$

Teljes táblázatok mellékelésétől itt eltekintünk, melyek (1) és (2) összehasonlítását szolgálják. Ellenben egy példán megvizsgáljuk a kettő közötti különbséget, és ennek mintájára tetszőleges információ birtokában elvégezhető a szórásnégyzetek összevetése. Az $\alpha=1,7$ és $\beta=18$ paramétereket használjuk továbbra is.

Legyen az (1) esetben $c=B8$ rögzített, és az évek száma $t=10,\ldots,28$ közötti. Hamar kiderül, hogy ekkor minden esetben az (1) szolgáltatja a jobb becslést, hiszen (2)-ben tetszőleges előző évi kárszám esetén a szórásnégyzet legalább $\frac{1{,}7+0}{(1+18)^2}\approx 0{,}0075$, ami bármely (1)-ben található szórásnégyzetnél nagyobb. Ugyanez történik akkor, ha a (2) esetben az egyén előző $2,3,\ldots,6$ évi kárszámát ismerjük. Javulást (2) részéről akkor tapasztalunk, amikor 7 évre visszamenőleg ismerjük a kártörténetet. A példa teljességéért egy kivonat a szórásnégyzetekről:

	(t = 11)				(t=15)				
D_1^2	0.0032	0.0030	0.0034	0.0036	0.0033	0.0037	0.0039	0.0035	0.0040
	(t = 20)					(t=25)			
D_1^2	0.0042	0.0038	0.0042	0.0044	0.0040	0.0043	0.0045	0.0041	0.0044

és a (2) esetben k előző t évi összkár jelöléssel

D_2^2:	$k=0$	$k=1$	$k=2$	$k=3$	$k=4$	$k=5$	$k=6$	$k=7$	$k=8$	$k=9$
$t=1$	0.0075	0.0102	0.0130	0.0158	0.0186	0.0213	0.0241	0.0269	0.0296	0.0324
$t=2$	0.0068	0.0093	0.0118	0.0142	0.0168	0.0192	0.0217	0.0242	0.0268	0.0292
$t=3$	0.0061	0.0084	0.0107	0.0129	0.0152	0.0175	0.0197	0.0220	0.0243	0.0265
$t=4$	0.0056	0.0076	0.0097	0.0118	0.0138	0.0159	0.0180	0.0200	0.0221	0.0242
$t=5$	0.0051	0.0070	0.0089	0.0108	0.0127	0.0146	0.0164	0.0183	0.0202	0.0221
$t=6$	0.0047	0.0064	0.0082	0.0099	0.0116	0.0134	0.0151	0.0168	0.0186	0.0203
$t=7$	0.0043	0.0059	0.0075	0.0091	0.0107	0.0123	0.0139	0.0155	0.0171	0.0187
$t=8$	0.0040	0.0055	0.0070	0.0084	0.0099	0.0114	0.0129	0.0143	0.0158	0.0173
$t=9$	0.0037	0.0051	0.0064	0.0078	0.0092	0.0106	0.0119	0.0133	0.0147	0.0160
$t=10$	0.0034	0.0047	0.0060	0.0073	0.0085	0.0098	0.0111	0.0124	0.0136	0.0149

3.6.2 Megjegyzés Amennyiben az eltelt évek száma és még néhány előző évi kárszám rendelkezésünkre áll, módosíthatjuk a becslésünket a következő módon. A Monte-Carlo típusú módszerünk során az η_i értékét pontosan akkor választjuk 1-nek, ha t év alatt a c osztályba jutunk, és az ismert $j_1, \dots, j_l$-edik évek kárszámai megegyeznek az ismert $k_{j_1}, \dots, k_{j_l}$-ekkel. Tehát amikor egy Markov-láncot generálunk véletlenszerűen, akkor a megfelelő j_i-edik lépésekben kapott károk megegyeznek az illető ismert kártörténetében szereplő k_{j_i} értékekkel. Ezzel jobb becslést kapunk a kárgyakoriságra, aminek az az ára, hogy jóval kevesebb η_i lesz 1, azaz több véletlen mintavételezést kell szimulálnunk.

3.7. Jó sofőr, rossz sofőr - egy diszkrét eset

Tegyük fel, hogy Λ eloszlása, azaz a keverőeloszlás diszkrét. Kezdjük azzal az egyszerű példával, hogy pontosan két értéket vehet fel: $P(\Lambda = \lambda_i) = p_i$ $(i = 1, 2)$, azaz λ_i-t p_i valószínűséggel, ahol $p_1 + p_2 = 1$. Ezt felfoghatjuk úgy is, hogy a szerződők p_1 hányada jó sofőr, viszonylag kicsi λ_i kárgyakorisággal, másik p_2 részük pedig rossz sofőr, nagyobb λ_2-vel. Tegyük fel, hogy a vizsgált szerződőre teljesül $B_t = c$, azaz t év alatt A0-ból a c osztályba jutott el. Ezzel a Bayes-tételből adódik

$$P(\Lambda = \lambda_i | B_t = c) = \frac{P(B_t = c|\Lambda = \lambda_i) \cdot P(\Lambda = \lambda_i)}{\sum\limits_{j=1,2} P(B_t = c|\Lambda = \lambda_j) \cdot P(\Lambda = \lambda_j)} = \frac{M^t(\lambda_i)_{(5,|c|)} \cdot p_i}{\sum\limits_{j=1,2} M^t(\lambda_j)_{(5,|c|)} \cdot p_j}$$

a jó, illetve rossz sofőr voltának valószínűsége az ismert feltétel mellett. Hiszen a $\{\Lambda = \lambda_1\}$ esemény megegyezik azzal az eseménnyel, hogy az illető jó sofőr. Az így kapott két pontra koncentrált feltételes eloszlás szerint várható értéket véve egy lehetséges becslést kapunk az illető következő évi kárszámára, más szóval λ-ra:

$$E(\Lambda | B_t = c) = \frac{\sum\limits_{i=1,2} \lambda_i \cdot M^t(\lambda_i)_{(5,|c|)} \cdot p_i}{\sum\limits_{i=1,2} M^t(\lambda_i)_{(5,|c|)} \cdot p_i}.$$

Példaként legyen a biztosítottak 75%-a jó sofőr $\lambda_1 = 0,04$, a maradék 25%-uk pedig rossz sofőr $\lambda_2 = 0,2$ kárgyakorisággal. Legyen $t = 12$ év - azaz vegyünk 12 lépést a rendszerben, és erről a szerződőről próbáljuk megmondani, hogy milyen valószínűséggel jó sofőr. Azért szemléletesebb 12 évtől néznünk, mert ez az első olyan t, amelyre egyszerre az összes osztály valószínűsége pozitív. Az összes osztályra az alábbi táblázat szolgáltatja a valószínűségeket (P(a 12 év alatt c-be kerülő sofőr jó)).

M4	M3	M2	M1	A0	B1	B2	B3
0.001	0.004	0.005	0.005	0.029	0.029	0.028	0.138

B4	B5	B6	B7	B8	B9	B10
0.138	0.128	0.449	0.449	0.804	0.804	0.953

A kárszám a posteriori várható értéke pedig ebben az esetben:

M4	M3	M2	M1	A0	B1	B2	B3
0.200	0.199	0.199	0.199	0.195	0.195	0.196	0.178

B4	B5	B6	B7	B8	B9	B10
0.178	0.179	0.128	0.128	0.071	0.071	0.047

3.7.1 Megjegyzés Ha maximum likelihood módszerrel dolgoznánk, akkor $M^t(\lambda)_{(5,|c|)}$-t kellene maximalizálni λ-ban, azaz megnézni, hogy melyik λ_i-re vesz fel nagyobb értéket. Más szóval melyik a valószínűbb, hogy t év után a c osztályba rossz, vagy jó sofőr került? Itt csak az számít, hogy milyen λ_i értékek lehetségesek, hozzájuk tartozó elsődleges p_i elképzelésünk nincs. Alkalmazva ezt a fenti $t = 12$ példa folytatásaként azt kapjuk, hogy a B7 osztálytól lefelé rossz, B8-tól felfelé jó minősítést kapnak a sofőrök.

Végül megjegyezzük, hogy a megoldás kiterjeszthető teljesen hasonló módon arra az esetre, amikor 2 helyett q-féle sofőr létezését tesszük fel. Ez interpretálható úgy, mint a biztosítottak különböző megbízhatósági szintekkel való felruházása. Gondolhatunk arra, hogy $q = 5$ esetben az 5 csoport a rossz, átlagnál rosszabb, átlagos, átlagnál jobb és jó sofőr címkékkel látható el, és ez tovább finomítható.

4. fejezet

Gyorsító módszerek

4.1. Importance Sampling

Az ismertetett Φ_N becsléssorozattal kapcsolatban felmerül a kérdés, hogy vajon nem tehetnénk-e hatékonyabbá bizonyos módosítással. Hiszen a vizsgált $\alpha = 1.7, \beta = 18$ paraméterek mellett az alacsonyabb osztályokat tekintve nagyon alacsony a $P(\eta_n = 1)$ valószínűség, azaz nagyon kicsi eséllyel választjuk be a hozzá tartozó ξ_n-et az átlag kiszámításába. Például a $B_t = M4$ feltétel mellett nagyon kis arányban találunk bele a $\Gamma(\alpha, \beta)$-ból generált paraméterrel t év alatt ebbe az osztályba. Ugyanis az $\frac{\alpha}{\beta}$ várható érték és kis szórásnégyzet miatt az esetek túlnyomó többségében nem kapunk olyan nagy ξ_n értéket, amivel t év alatt a legrosszabb osztályba kerülhetünk. Látható ugyanakkor, hogy az életben előforduló α, β értékek esetén is így van, nem csak a mi példánk mutatja ezt a rossz tulajdonságot. Ezzel kevés ξ_n-ből átlagot véve könnyen előfordulhat, hogy a szórás olyan nagy marad, amit szeretnénk elkerülni, ill. lejjebb szorítani. A következőkben egy erre konstruált módszert, az importance sampling-et ismertetjük, melyet fontossági mintavételezésnek fordíthatunk, de itt az angol megfelelőjét fogjuk használni.

Az alapötlet átgondolását kezdjük azzal a feltevéssel, hogy egy valószínűségi változó bizonyos g függvényének a várható értékét szeretnénk kiszámítani, ahol g az X értékkészletén értelmezett. Ez a várható érték felírható integrál alakban abszolút folytonos eloszlású - és az általunk vizsgált esetben nemnegatív - X esetén: $E(g(X)) = \int_0^\infty g(x) f_X(x) dx$. [1] Amennyiben ez nem számolható közvetlenül, szokásos módon adható rá $\tilde{g}_n(X) = \frac{\sum_{i=1}^{n} g(X_i)}{n}$ közelítő átlag, ahol $X_1, X_2, \ldots, X_n$ az X eloszlásából generált véletlen minta, amely sztochasztikusan tart $E(g(X))$-hez, amennyiben

[1] Teljesen hasonlóan meggondolható mindez diszkrét esetben is, amit most nem részletezünk.

az létezik. A becslés szórásnégyzete

$$D^2\left(\frac{\sum_{i=1}^{n} g(X_i)}{n}\right) = \frac{1}{n}D^2(g(X)) = \frac{1}{n}\left(E(g(X)^2) - E^2(g(X))\right).$$

Ebből is látszik, hogy a szórásnégyzet a második momentum csökkentésével lenne lehetséges, hiszen a várható értéken nem változtathatunk, végtére is arra szeretnénk közelítést adni. Képzeljük el, hogy - az $f(x) := f_X(x)$ jelöléssel - az $\int_0^\infty g(x)f(x)dx$-ben az integrandust beszorozzuk 1-gyel, méghozzá $\frac{h(x)}{h(x)}$ alakban, ahol $h(x)$ integrálja $(0,\infty)$-en 1, azaz sűrűségfüggvény. Ezzel $\int_0^\infty g(x)f(x)\frac{h(x)}{h(x)}dx = \int_0^\infty \frac{g(x)f(x)}{h(x)}h(x)dx = E_H\left(\frac{g(X)f(X)}{h(X)}\right)$-et kapunk. Mostantól lényeges, hogy milyen eloszlás szerint veszünk várható értéket, ezért indexben fogjuk jelölni az aktuális eloszlást. Legyen így H a $h(x)$, ill. F az $f(x)$ sűrűségfüggvények által meghatározott eloszlások. Ezzel pontosan azt kaptuk, hogy

$$E_F(g(X)) = E_H\left(\frac{g(X)f(X)}{h(X)}\right),$$

más szóval a jobb oldal is egy jó, torzítatlan becslés a keresett várható értékre. Ha emellett a jobb oldali várható értékben szereplő valószínűségi változó második momentuma kisebb a bal oldaliénál, akkor abból következően a szórásnégyzete is kisebb. Ez pontosan akkor teljesül, ha az

$$\int_0^\infty g(x)^2 f(x)dx > \int_0^\infty \frac{g^2(x)f^2(x)}{h^2(x)}h(x)dx$$

egyenlőtlenség is teljesül. Vegyük észre, hogy a legoptimálisabb megoldást a jobb oldal minimalizálása jelentené $h(x)$-ben, a megfelelő integrálfeltétel mellett, azaz

$$\min_{h(x)} \int_0^\infty \frac{g^2(x)f^2(x)}{h(x)}dx$$

$\int_0^\infty h(x) = 1,\ supp(f(x)) \subseteq supp(h(x))$ a feladat.

A legutolsó természetes feltételről még nem tettünk említést. Ez azokat az eseteket zárja ki, amikor az integrandus értéke végtelen lehet, és $h(x)$ pozitivitását fejezi ki a számunkra lényeges $(0,\infty)$ intervallumon. (Ebben már benne foglaltatik, hogy $g^2(x)$ tartóját is tartalmazza, mert most $supp(f) = (0,\infty)$.) Belátható, hogy ezt a minimumot $h(x) = \frac{f(x)g(x)}{\int_0^\infty f(x)g(x)dx}$ esetén veszi fel, többek között abból, hogy ekkor a második momentum a várható érték négyzetével lesz egyenlő, melynek következtében a szórásnégyzet eltűnik. Gondoljunk csak arra, hogy ekkor az $\frac{fg}{h}$ értéke konstans. A probléma abban áll, hogy a nevezőben megjelenik a várható érték, aminek a kiszámítására vállalkoztunk, így ha ezt tudnánk, nem lenne értelme az előbbi okoskodásnak. Meg kell elégednünk egy olyan importance sampling függvénnyel, amely - bizonyos értelemben - jól közelíti $f \cdot g$-t, így az eredeti becsléshez képest javítást tudunk előidézni. Összefoglalva az alapesetünket, olyan $h(x)$ sűrűségfüggvényt keresünk, mely pozitív a pozitív félegyenesen, közel van $f \cdot g$-hez, továbbá lényeges még az általa meghatározott eloszlásból történő könnyű szimuláció. Nem sokra mennénk ugyanis egy olyan függvénnyel, amelyből olyan nehéz mintát generálni, mint a tárgyalásunk fontos momentumát képező $f_{\Lambda|B_t=c}$ sűrűségfüggvényből.

4.2. Gyorsítás

A fenti megfontolással élve térjünk rá a problémánk megoldására, ahol kissé bonyolultabb a helyzet, ugyanis a $\frac{\sum_{n=1}^{N} \xi_n \eta_n}{\sum_{n=1}^{N} \eta_n}$ közelítésen szeretnénk javítani szórásnégyzetének csökkentésével. A $P(B_t = c|\Lambda = \lambda)$ feltételes valószínűséget rögzített $\{B_t = c\}$ mellett az átláthatóság kedvéért $r(\lambda)$-val fogjuk jelölni. Ezt megtehetjük, mert így λ-nak $\mathbb{R}^+ \to (0,1)$ függvénye. Emlékeztetünk, hogy $(r_{t,c}(\lambda) =)r(\lambda) := M^t(\lambda)_{(5,|c|)}$ egyváltozós függvény rögzített t és c értékek esetén. Megjegyezzük ugyanakkor, hogy még folytonos is, ami a $\frac{\lambda^k}{k!}e^{-\lambda}$ alakú függvények szorzatösszegéből következik. Jó tulajdonsága még, hogy nem oszcillál, hanem $M3$-tól $B9$-ig van pozitív maximuma, és ezektől balra monoton nő, jobbra pedig monoton csökken. Kivétel ez alól $M4$, melynek nincs maximuma, és λ növekedtével monoton 1-hez tart, valamint $B10$ a 0-ban 1-hez tart, onnan pedig monoton csökken. Ezt szemléltetjük a 4.1 ábrán az összes osztályra kirajzolva $t = 15$ év mellett.

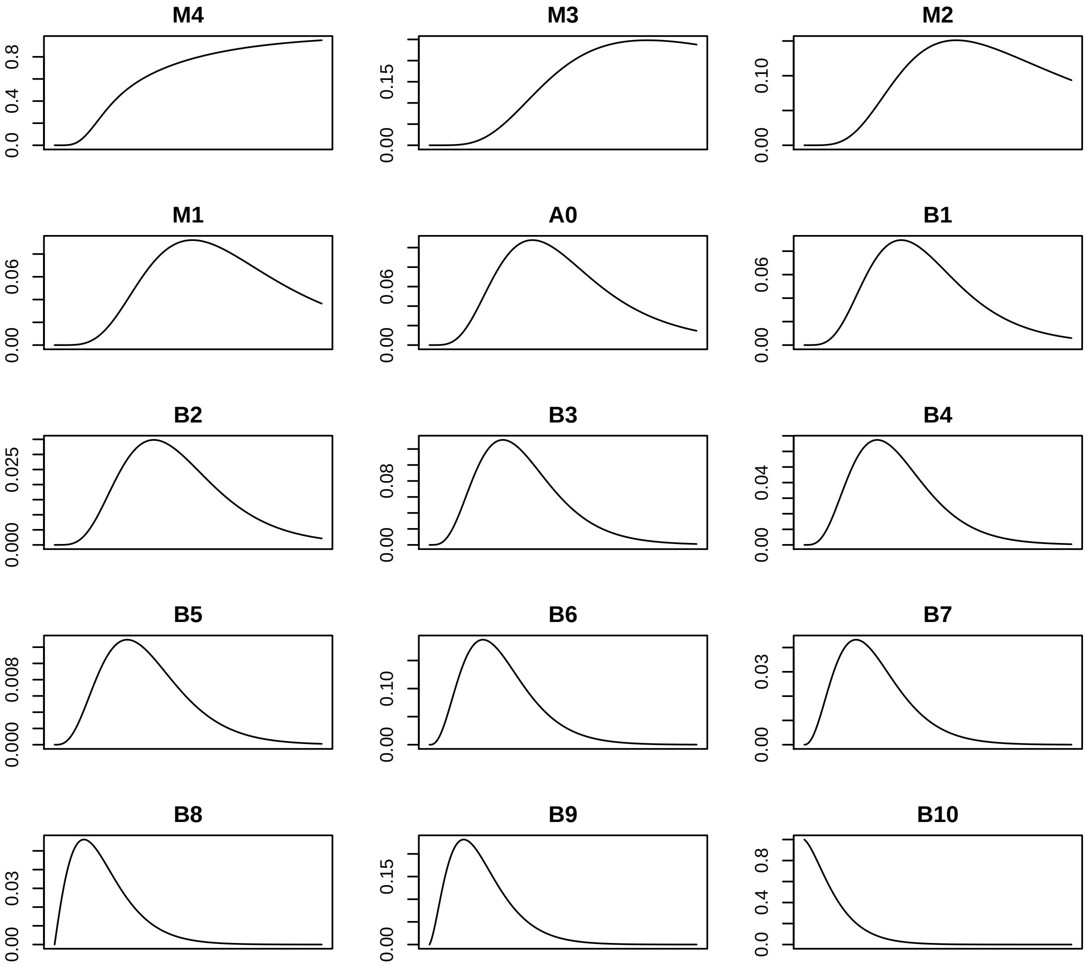
Vajon mi lesz egyáltalán az a sorozat, amellyel Λ a posteriori várható értékét fogjuk becsülni? Ehhez látnunk kell a

$$E(\Lambda|B_t = c) = \frac{\int_0^\infty \frac{x \cdot r(x) \cdot f_\Lambda(x)}{h(x)} h(x)dx}{\int_0^\infty \frac{r(x) \cdot f_\Lambda(x)}{h(x)} h(x)dx}$$

módosított felírást. Ennek megfelelően az eredeti ξ_n változók helyett, melyeket $\Gamma(\alpha,\beta)$-ból generáltunk, most X_n-eket veszünk a H eloszlásból. Továbbá segédváltozóként definiáljuk ζ_n-eket: értékük legyen $P(B_t = c|\Lambda = X_n)$ valószínűséggel $\frac{f_\Lambda(X_n)}{h(X_n)}$, különben pedig 0. Ezek teljesen hasonlók, mint η_n-ek, csak a $h(x)$ függvény bevonásával módosítottunk rajtuk. [2] Tehát a Monte-Carlo típusú közelítésünk most a

$$\frac{\sum_{n=1}^{N} X_n \zeta_n}{\sum_{n=1}^{N} \zeta_n}$$

[2] A szakirodalomban olvashatunk erről weighted importance sampling címszó alatt, például a [2] könyvben.



4.1. ábra. $P(B_t = c|\Lambda = \lambda)$ ábrázolása λ függvényében $t = 15$ év esetén a különböző osztályokra. (Az x tengely az ábrákon a $[0,1]$ intervallum $M4$-et kivéve, ahol $[0,4]$.)

alakot ölti. Ez persze még mindig nem túl jó, hiszen a már látott jó tulajdonságokkal rendelkező h-t másnak érdemes definiálni a nevező, és másnak a számláló esetén. Megtehetjük, hogy a számlálót és a nevezőt külön kezeljük, és mindkettőhöz olyan h sűrűségfüggvényeket rendelünk, melyek a két esetben gyors konvergenciát eredményeznek. Tehát annyiban módosítunk az előző felíráson, hogy ζ_n-ek helyett $\zeta_{1,n}, \zeta_{2,n}$-eket generálunk teljesen hasonlóan, csak előbbit a h_1, utóbbit a h_2 sűrűségfüggvény alapján. Így a közelítésünk $\frac{\sum_{n=1}^{N} X_n \zeta_{1,n}}{\sum_{n=1}^{N} \zeta 2,n}$ lesz, melyekre teljesül, hogy $\frac{\sum_{n=1}^{N} X_n \zeta_{1,n}}{N} \longrightarrow \int_0^\infty \lambda P(B_t = c|\Lambda = \lambda) f_\Lambda(\lambda) d\lambda$ és $\frac{\sum_{n=1}^{N} \zeta_{2,n}}{N} \longrightarrow P(B_t = c)$ L_1-ben és 1 valószínűséggel is.

Most pedig arra fogunk törekedni, hogy megfelelő h sűrűségfüggvényt találjunk. Ehhez egy részben numerikus megoldást adunk a következőképpen. Először a nevezőre gondolva tekintsük az $r(x)f_\Lambda(x)$ függvényt rögzített t és c esetén. Vegyük a $[0,b]$ intervallum egy - az egyszerűség kedvéért egyenletes - $0 = x_1 < x_2 < \ldots < x_K = b$ felosztását. Ezekben a pontokban számítsuk ki az $r(x)f_\Lambda(x)$ függvény értékeit, melyeket jelöljön rendre $y_1, y_2, \ldots, y_K$. Technikai megjegyzésként, ezen számítás műveletigénye nem túl nagy, számítógéppel több ezerszeres felbontásra is gyorsan számolható. Ezek lesznek h-nak is az $x_1, \ldots, x_K$ pontokban felvett értékei, továbbá két szomszédos pont között lineárisan interpoláljunk. Formálisan legyen

$$h(x) = a_i x + b_i = \frac{y_{i+1} - y_i}{x_{i+1} - x_i} x + y_i - \frac{y_{i+1} - y_i}{x_{i+1} - x_i} x_i \ , \quad x_i \le x \le x_{i+1} \quad (1 \le i \le K-1)$$

Ezzel a $[0,b]$ intervallumon kaptunk egy töröttvonalat, ami finomodó felosztás mellett közelíti az integrandust. Az (x_K, ∞) intervallumon egyszerűen legyen $h(x) = e^{-x}$. A farokeloszlásnak ez a klasszikus leegyszerűsítése ugyanúgy exponenciális lecsengést eredményez, mint $P(B_t = c|\Lambda = \lambda)f_\Lambda(\lambda)$ esetén.

Legyen T_i az $(x_i, 0), (x_{i+1}), (x_{i+1}, y_{i+1}), (x_i, y_i)$ pontok által meghatározott trapéz területe, ahol $i = 1, \ldots, K-1$, továbbá T_K az e^{-x} grafikonja alatti terület (x_K, ∞)-en. Ezzel a jelöléssel $\int_0^\infty h(x)dx = \sum_{i=1}^{K} T_i$, melyet jelöljük ϑ-val. Ezen normáló konstans reciprokával beszorozva a $h(x)$ függvényt máris sűrűségfüggvényt kaptunk, legyen $h_\vartheta(x) = \frac{1}{\vartheta} h(x)$. Számítsuk ki a $h_\vartheta(x)$ sűrűségfüggvény által karakterizált eloszlású valószínűségi változó első és második momentumát:

$$E_{h_\vartheta}(Z) = \int_0^{x_K} x \cdot h_\vartheta(x)dx + \int_{x_K}^{\infty} x \cdot e^{-x}dx = \sum_{i=1}^{K-1} \int_{x_i}^{x_{i+1}} x h_\vartheta(x)dx + (x_K + 1)e^{-x_K} =$$

$$= \sum_{i=1}^{K-1} \int_{x_i}^{x_{i+1}} a_i x^2 + b_i x dx + (x_K + 1)e^{-x_K} = \sum_{i=1}^{K-1} \left[\frac{a_i x^2}{3} + \frac{b_i x^2}{2} \right]_{x_i}^{x_{i+1}} + (x_K + 1)e^{-x_K},$$

$$E_{h_\vartheta}(Z^2) = \int_0^{x_K} x^2 \cdot h_\vartheta(x)dx + \int_{x_K}^{\infty} x^2 \cdot e^{-x}dx = \sum_{i=1}^{K-1} \left[\frac{a_i x^4}{4} + \frac{b_i x^3}{3} \right]_{x_i}^{x_{i+1}} + (x_K^2 + 2x_K + 2)e^{-x_K}.$$

Azzal a feltételezéssel élünk, hogy az $r(x)f_\Lambda(x)$ függvény egy bizonyos paraméterű (α', β') Gamma-eloszlás sűrűségfüggvényének konstansszorosával jól közelíthető, és $h_\vartheta(x)$-et ilyennek választjuk. A paraméterek az ismert összefüggések alapján jól számolhatóak, hiszen a várható érték $E_{h_\vartheta}(Z) = \frac{\alpha'}{\beta'}$, a szórásnégyzet pedig $D^2_{h_\vartheta}(Z) = E_{h_\vartheta}(Z^2) - E^2_{h_\vartheta}(Z) = \frac{\alpha'}{\beta'^2}$. Ezek alapján a paramétereket válasszuk meg a következőképpen:

$$\alpha' = \frac{E^2_{h_\vartheta}(Z)}{E_{h_\vartheta}(Z^2) - E^2_{h_\vartheta}(Z)}, \quad \beta' = \frac{E_{h_\vartheta}(Z)}{E_{h_\vartheta}(Z^2) - E^2_{h_\vartheta}(Z)}.$$

Gondoljuk meg, hogy $P(B_t = c|\Lambda = \lambda)$-t úgy kaptuk meg, hogy összeadtunk $const_1 \cdot \lambda^x e^{-const_2 x}$ alakú tagokat. Ezekben persze előfordulhat, hogy az exponenciális tag kitevője 0, de ezt még beszoroztuk f_Λ-val, ami már biztosítja, hogy egyik ilyen kitevő sem 0. Tehát lényegében Gamma-eloszlások sűrűségfüggvényeit adtuk össze, melyek általában mindkét paraméterükben különbözők. Erre az esetre nem ismerünk általánosítást, és azt is tudjuk, hogy ily módon egy nem nevezetes eloszlás sűrűségfüggvényéhez jutunk.

Sejtésünk, hogy ezzel tulajdonképpen megfelelő felosztás mellett (gyorsan) közelíthetőek az α', β' paraméterek, melyek hányadosa a keresett a posteriori várható értéket adja. Hiszen ekkor $r(x)f_\Lambda(x) \approx \frac{x^{\alpha'-1} \cdot \beta'^{\alpha'} \cdot e^{-\beta' x}}{\Gamma(\alpha')} \cdot c$ és $xr(x)f_\Lambda(x) \approx \frac{x^{\alpha'} \cdot \beta'^{\alpha'+1} \cdot e^{-\beta' x}}{\Gamma(\alpha'+1)} \cdot \frac{\alpha'}{\beta'} \cdot c$, azaz

$$\frac{\int_0^\infty xr(x)f_\Lambda(x)dx}{\int_0^\infty r(x)f_\Lambda(x)dx} \approx \frac{\frac{\alpha'}{\beta'} \cdot c}{c} = \frac{\alpha'}{\beta'}.$$

Ha azonban egy példán keresztül megvizsgáljuk a módszer konvergenciáját, arra a megállapításra jutunk, hogy rosszabb eredményt kapunk az importance sampling nélküli közelítéshez képest. Azaz a várható érték jobban ingadozik ebben az esetben, ami mindenképpen gyanakvásra ad okot. Ezért tovább megyünk, és szemügyre vesszük a harmadik és negyedik momentumokat. Ezzel arra a következtetésre jutunk, hogy az eredeti eloszlásunk első három közelített momentuma nagyon közel van az α', β' paraméterű Gamma-eloszlás első három momentumához, ellenben a negyedik momentum jóval kisebb nála. Tehát a h_ϑ sűrűségfüggvénynél a közelíteni kívánt függvény sokkal lapultabb, mint vártuk, ugyanis a lapultságot az $\frac{E(X^4)}{D^4(X)} - 3$ formula szolgáltatja. A különbség nagyságrendjét érzékelteti, hogy a $\Gamma(\alpha', \beta')$ lapultsága $\frac{6}{\alpha'}$, ami a vizsgált esetekben 1 és 3 közötti érték, míg a közelített függvényre 40 és 100 közötti értékeket kaptunk.

A kérdés így sajnos nyitva maradt, hogy vajon lehet-e az előbbi gondolatmenet alkalmas módosításával gyorsítást elérni.

4.3. Metropolis-Hastings típusú algoritmus

A dolgozat utolsó részében szükségét érezzük egy olyan módszer ismertetésének, mely talán az összes eddiginél hatékonyabban - azaz gyorsabban - képes előállítani a feltételes eloszlásból a várható értéket. Ebben fel fogjuk használni a $P(B_t = c|\Lambda = \lambda) \cdot f_\Lambda(\lambda)$-nak, mint λ függvényének általunk feltételezett karakterisztikáját. Ebből látni fogjuk, hogy ez a $\Gamma(\alpha', \beta')$ feltételezett eloszlás jó javaslat az a posteriori eloszlásra, melynek jelentését alább tárgyaljuk.

A módszer alapja a következő. Az $f_{apost} \overset{jel}{:=} f_{\Lambda|B_t=c}(\lambda|c)$ sűrűségfüggvény által definiált eloszlásból (ez legyen F_{apost}) szeretnénk mintát generálni, melynek átlaga a nagy számok törvényének értelmében tart a várható értékhez, midőn a mintaelemszám tart végtelenbe. Ezzel a már látott problémánk adódik, hogy az eloszlás komplexitása miatt nem tudunk közvetlen mintavételezést alkalmazni. A jelenlegi keretek között nem célunk a Metropolis-Hasting-algoritmus általános formában történő ismertetése, ezzel kapcsolatban többek között a [6] könyvet ajánljuk. Mi ennek egy speciális fajtáját használjuk, melynek alapötletét Chib és Greenberg vezették be 1995-ben.

Tetszőleges $x_0 \in \mathbb{R}_+$-ból kiindulva konstruálunk egy Markov-láncot, melynek stacionárius eloszlása maga az F_{apost} eloszlás, azaz megfelelő lépésszám után x_l ebből az eloszlásból származó egyelemű minta lesz. Mi most egy egyszerűbb kiinduló algoritmust készítünk, melyben X_i nemhogy $X_0, \ldots, X_{i-2}$-től nem függ, hanem még X_{i-1}-től sem, mint azt a Markov-láncokról tudjuk. Azaz véletlen bolyongásra építjük egy minta generálását, és látni fogjuk, hogy ez megfelelő lesz.

Legyen $h(x)$ az előző fejezetben kiszámított α', β' paraméterű Gamma-eloszlás sűrűségfüggvénye. Ne feledjük, hogy ez feltételes sűrűségfüggvény, ahol a feltétel a $\{B_t = c\}$, csak most ezt az egyszerűség kedvéért rögzítettük. Tegyük fel, hogy eddig megkaptuk az $x_0, \ldots, x_{i-1}$ értékeket, és most szeretnénk a következőt realizálni. Az x' értéket származtassuk ebből a $\Gamma(\alpha', \beta')$ eloszlásból, és legyen

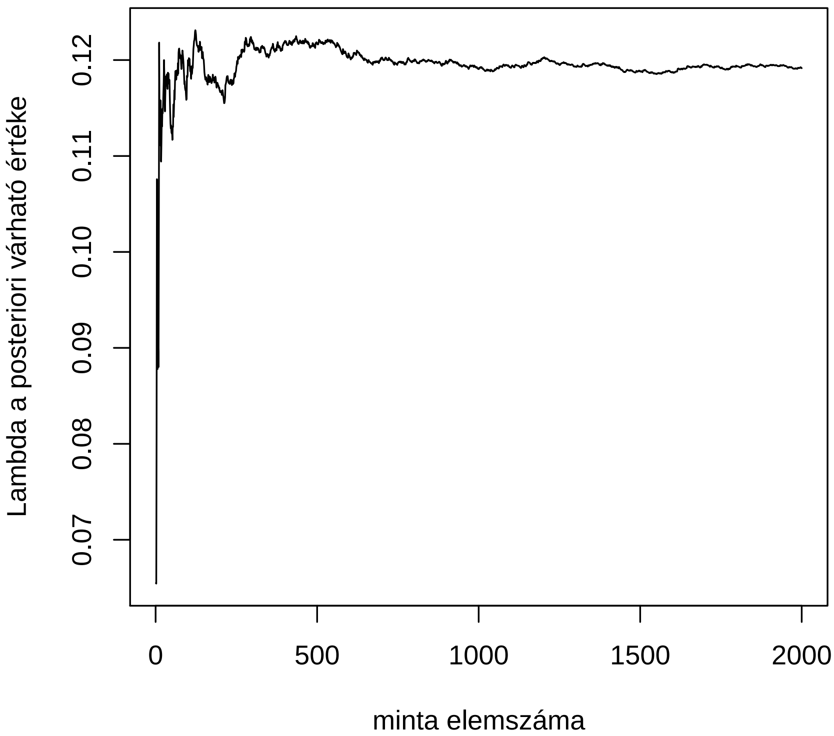
$$\alpha_i = \min\left\{1, \frac{w(x')}{w(x_{i-1})}\right\},$$

ahol $w(x) = \frac{P(B_t=c|\Lambda=x)\cdot f_\Lambda(x)}{h(x)}$ korlátos függvény. Ezt használva α valószínűséggel legyen $x_i = x'$, különben pedig maradjon az előző érték, azaz $x_i = x_{i-1}$. Ezzel az egyszerű konstrukcióval elég nagy i-re x_i az F_{apost} eloszlásból származó érték lesz.

4.3.1 Megjegyzés 1. Amikor a $P(B_t = c|\Lambda = x) \cdot f_\Lambda(x)$ szerinti mintáról beszélünk, akkor ennek normáltjára gondolunk. Azaz ebből egy konstans szorzóval, nevezetesen $\int_0^\infty P(B_t = c|\Lambda = x) \cdot f_\Lambda(x)dx$-szel történő leosztással kapott sűrűségfüggvényről. A fenti α formulájában ennek elhagyása nem okoz zavart, hiszen mind a számlálóban, mind a nevezőben le kell osztanunk vele. Más szóval a minta ezen függvény grafikonja alatti (x tengely által határolt) területrészen egyenletesen vett valószínűségi változó realizációjának x koordinátája, mely pontosan a megfelelő a posteriori eloszlásból származik.

2. Tetszőleges javasolt h sűrűségfüggvény esetén teljesül a konvergencia az a posteriori eloszláshoz. Mindamellett az algoritmus annál jobban működik, minél jobban közelíti $h(x)$ a $P(B_t = c|\Lambda = x) \cdot f_\Lambda(x)$ fenti konstansszorosát. Más szóval annál kevesebb iterációs lépésre van szükség egy Markov-lánc konstrukciójában.

Végül tekintsünk egy példát ennek működésére. Legyen $t = 13$ és az osztály $c = B7$, továbbá minden egyes mintaelem generálásakor 100 iterációt végezzünk. Ugyan ez több időt vesz igénybe, mintha közvetlenül tudnánk mintát venni egy eloszlásból, de a 4.2 ábrán szemléltetjük, hogy a jó közelítéshez már 1000 elemszám esetén megfelelő eredményre jutunk.



4.2. ábra. Konvergencia Markov-lánc-Monte-Carlo módszerrel ($t = 13$, $c = B7$)

5. fejezet

Zárszó

Összefoglalásként a szakdolgozat teljes mértékben önálló eredményeit ismertetjük rövid áttekintéssel. Fő eredményként a kárgyakoriság ismert eltelt évek és elért osztály ismerete alapján történő Bayes-i becslését emelhetjük ki. Ezen belül a 3.4 alfejezetben ismertetett ritkításon alapuló módszert, mely az adott feltétel mellett valószínűtlenebb mintaelemeket nagy valószínűséggel kihagyja az átlagolásból. Erről a 3.4.1 és 3.4.2 állítások 3.4.3 következményeként megkaptuk, hogy aszimptotikusan torzítatlan módon becsüli az a posteriori várható értéket, és ezzel együtt kijött a torzítás mértéke is. Ezek után megvizsgáltuk, hogy a $\frac{\Phi_N}{1-q^N}$ sorozatban N értékét, azaz a generált független minta elemszámát mekkorára érdemes venni egy kívánt ε pontosságú közelítés eléréséhez. Erre a Csebisev-egyenlőtlenséget használtuk, melyhez a ritkított átlagunk szórásnégyzetének meghatározására volt szükség, erről szól a 3.4.4 lemma. A 3.4.5 következmény még nem adott kielégítő választ a becslés és az igazi várható érték eltérésére vonatkozó valószínűség felső korlátját illetően. Ennek kapcsán alsó és felső becslést adtunk az itt megjelenő $J_{N,p}$ összegre a 3.4.6 lemmán keresztül. Ezt követően igazoltuk a 3.4.8 lemmát, melyet felhasználva az a priori eloszlás α és β paraméterének, valamint a $p = P(\eta = 1)$ valószínűség és ε, δ pontossági korlátok függvényében megadtuk az N-re vonatkozó becslést.

Ezt követően konkrét példát mutattunk $\alpha = 1,7$ és $\beta = 18$ értékek mellett egy kárgyakoriságot becslő táblázatra, melyben a sorok az eltelt évek számára, az oszlopok pedig az elért osztályokra vonatkoznak. Példát adtunk továbbá a közelítés szórására grafikusan is, és megjegyeztük, hogy jobb osztályok esetén stabilabb konvergenciát kapunk, azaz itt kisebb N számú generálás is elegendő. A 3.4.11 és 3.4.12 példákban az elkészített táblázat használatát szimuláltuk különböző állományokkal, és láttuk, hogy a valóságot az előre sejthetőnél jobban közelíti, azaz a csekélynek tűnő információ többet mond a vártnál. A 3.5 alfejezetben ismertetett számítások nem új eredmények, viszont az ezt követő konkrét, szórásnégyzetre vonatkozó összehasonlításhoz szükségesek, amely önálló számításon alapul.

A fontossági mintavételezésnél a [2] és [5] irodalmak kerültek felhasználásra, majd a 4.2 Gyorsítás alfejezet önálló munka eredménye. A Metropolis-Hastings típusú algoritmushoz a [6] könyvet hívtuk segítségül, majd mutattunk rá példát.

Utólagos megjegyzésként megemlítjük, hogy az életben előforduló káreloszlások között csekély arányban ugyan, de előfordulhat, hogy a Poisson-eloszlás keverőjeként nem feltétlenül tudjuk elfogadni a Gamma-eloszlást. Azaz a hipotézisvizsgálat során elutasítjuk a minta negatív binomiális eloszlásból származóságának feltevését. Számos szakirodalom áll rendelkezésre (például [2]), melyek ilyen esetekben adnak alternatívát, többek között abban az esetben, ha a tapasztalati szórásnégyzet nagyobb, mint a feltételezett eloszlás szórásnégyzete. Ekkor a mi esetünkben érdemes megvizsgálni Gamma helyett Inverz Gauss vagy Lognormális keverővel a rendelkezésre álló kárszámokat. Ezek a Gamma-kevert Poissonnál, azaz a Negatív Binomiálisnál már jóval bonyolultabb eloszlásokat adnak, így megoldásuknál csak numerikus módszerekre hagyatkozhatunk.

Mi négyzetes veszteségfüggvénnyel dolgoztunk, amelynek minimalizálását a sokat tárgyalt Bayes-i a posteriori várható érték adta. A választás nem önkényes, hiszen esetünkben ez a legtöbbet alkalmazott függvény. Azonban meg kell említenünk, hogy még sokat használt az exponenciális veszteségfüggvény is.

A várható kárszámok mellett nem érdektelenebb kérdés az, hogy a károk nagysága hogyan fog alakulni. Hiszen a biztosítónak egyáltalán nem mindegy, hogy 200 ezer, vagy 2 millió forintos károkat kell rendeznie. Ennek vizsgálatával és a dolgozatban leírt kárszámbecslésekkel tudunk pontosabb képet adni a bevételekről és kiadásokról. Érdemes itt megjegyezni, hogy a kárt okozónak bizonyos érték alatt nem érdeke jelenteni a biztosítási eseményt, mert adott esetben kevesebbet fizet rá a másik féllel való megegyezéssel, mint a rosszabb bónuszosztállyal járó többletdíjjal. Ezt az angol nyelvű szakirodalom hunger for bonus-nak nevezi találóan. Szintén egy újabb dolgozatot megtöltene ennek tárgyalása.

Érdekes kérdés még, hogy az utolsó részben ismertetett egyszerűbb Metropolis-Hastings típusú algoritmust milyen módosítással lehetne még ennél is hatékonyabbá tenni. Azaz olyan javasolt eloszlásra (proposal distribution) cserélni a véletlen bolyongást, melyben X_i értéke függ az őt megelőző X_{i-1} értékétől.

Köszönetnyilvánítás

Szeretnék köszönetet mondani elsősorban konzulensemnek és tanáromnak, Arató Miklósnak segítő munkájáért, javaslataiért, melyekkel nagyban segítette szakdolgozatom elkészültét. Külön köszönettel tartozom Prőhle Tamás tanár úrnak, aki mindig készen állt segíteni, tanácsot adni, bármilyen kérdés merült is fel a témával kapcsolatban. Köszönöm még Karátson János, Márkus László, Móri Tamás és Prokaj Vilmos tanár uraknak a segítséget!

Irodalomjegyzék

[1] Jean Lemaire, *Automobile Insurance: Actuarial Models.* Kluwer-Nijhoff Publishing, 1985

[2] George S. Fishman, *Monte Carlo: Concepts, Algorithms, and Applications.* Springer-Verlag New York, 1996

[3] Jean-François Walhin, *Recursion for Actuaries and Applications in the Field of Reinsurance and Bonus-Malus Systems.* Doktori disszertáció, 2000

[4] Michel Denuit, Xavier Maréchal, Sandra Pitrebois, Jean-François Walhin, *Actuarial Modelling of Claim Counts.* Wiley, 2007

[5] Eric C. Anderson, *Monte Carlo Methods and Importance Sampling.* 1999

[6] Siddhartha Chib, Merlise Clyde, George Woodworth, Alan Zaslavsky, *Subjective and Objective Bayesian Statistics.* Wiley, 2003