Kapitoly z pojistné matematiky

Teorie kredibility

Silvie Zlatošová

2021

Obsah

1. Úvod
2. Teorie kredibility omezených fluktuací
3. Optimální teorie kredibility

Obsah

1 Úvod

2 Teorie kredibility omezených fluktuací

3 Optimální teorie kredibility

Základní pojmy

- **Teorie kredibility** je nástroj, který pojišťovnám umožňuje upravovat budoucí pojistné klientů v závislosti na jejich historii či rizikové skupině, do níž klient patří.
- Jestliže klient dosahuje trvale **lepších výsledků** (nenárokuje pojistné plnění) než průměrný klient, který platí základní pojistné, pak by bylo spravedlivé, aby takový klient získal **redukci** svého pojistného (slevu).
- Podle stejné logiky by také klienti s **vyšší úrovní rizika** měli platit **vyšší pojistné**.

Základní pojmy

- Tabulková hodnota pojistného je navržena tak, aby odrážela očekávané zkušenosti **celé skupiny klientů** a předpokládá **homogenní riziko**.
- Ve skupinách však díky **nedokonalosti systému** stále zůstává jistá míra **heterogenity** v úrovních rizika.
- Tedy někteří klienti představují nižší riziko někteří naopak vyšší riziko, než předpokládají tabulky.
- **Pojistitel tedy musí zvážit**, do jaké míry jsou odlišnosti ve zkušenostech klientů ovlivněny náhodnou odchylkou a do jaké míry tím, že klient skutečně představuje větší či menší riziko než je průměr pro danou skupinu.
- Také musí zvážit, jak důvěryhodné jsou vlastní zkušenosti klienta.

Základní pojmy

Budeme předpokládat následující:

- Čím více informací o minulých zkušenostech klientů pojistitel má, tím důvěryhodnější jsou vlastní zkušenosti klienta. Zkušenosti větší skupiny klientů jsou důvěryhodnější než zkušenosti malé skupiny.
- Konkurenční tlak může pojistitele přimět k tomu, aby dal plnou důvěru vlastním zkušenostem klienta. Pojišťovny se snaží si klienta udržet.

Základní pojmy

- Teorie kredibility se dostává do **sporu s klasickou statistikou**.
- Jestliže máme k dispozici zkušenosti nějaké skupiny klientů, pak by na základě přístupu klasické statistiky mělo být pojistné určeno jako střední hodnota nebo nějaký jiný nestranný odhad.
- Teorie kredibility říká, že by takto vypočtené pojistné mělo mít pouze částečnou váhu a zbývající část by měla být dopočtena na základě jiných informací.

Obsah

1 Úvod

2 Teorie kredibility omezených fluktuací

3 Optimální teorie kredibility

Formulace problému

- Poskytuje mechanismus, který umožňuje stanovení plné nebo částečné kredibility pro zkušenost pojištěného.
- Pokouší se omezit vliv náhodného kolísání pozorování na odhady.
- Tento přístup byl navržen na počátku 20. století ve spojení s pojištěním zaměstnanců.
- První článek zabývající se touto problematikou napsal Mowbray v roce 1914.

Formulace problému

- Předpokládejme klienta i, který má zkušenosti z T_i po sobě jdoucích období a v každém nahlásil N_{it} pojistných nároků, kde $t \in \{1, 2, \ldots, T_i\}$.
- Budeme předpokládat, že škody u jednotlivých pojistných nároků jsou stejné (můžeme je tedy pro další výpočty zanedbat) a že pro všechna $t \in \{1, 2, \ldots, T_i\}$ platí

$$\mathrm{E}(N_{it}) = \xi,$$

  což znamená, že střední hodnota je v čase konstantní.

Formulace problému

- Dále budeme pro všechna $t \in \{1, 2, \ldots, T_i\}$ předpokládat konstantní
$$\mathrm{Var}(N_{it}) = \sigma^2.$$
- Pro tohoto klienta pak můžeme vyjádřit jeho průměrnou zkušenost za T_i sledovaných období jako
$$\bar{N}_i = \frac{1}{T_i}(N_{i1} + N_{i2} + \ldots + N_{iT_i}).$$
- Pro střední hodnotu $\bar{N}_i$ pak platí $\mathrm{E}\bar{N}_i = \xi$.
- Budeme-li předpokládat nezávislost jednotlivých N_{it} pak můžeme psát
$$\mathrm{Var}\bar{N}_i = \frac{\sigma^2}{T_i}.$$
- Cílem pojistitele je určení hodnoty parametru ξ.

Přístupy k určení pojistného

- Jednou z možností je ignorovat zkušenosti klientů, v tom případě můžeme mluvit o **nulové kredibilitě**, a účtovat všem stejné základní pojistné μ.
- Toto pojistné nazýváme také jako *tabulkové pojistné*, v angličtině bývá označováno jako *manual rate*.
- Další možností je ignorovat μ a účtovat $\bar{N}_i$. Pak hovoříme o **plné kredibilitě**.
- Třetí možností je účtovat pojistné, které bude kombinací μ a $\bar{N}_i$. Jedná se o **částečnou kredibilitu**.

Přístupy k určení pojistného

- Z pohledu pojistitele je výhodnější účtovat $\bar{N}_i$ v případě, kdy jsou zkušenosti klienta méně volatilní (hodnota σ^2 je malá).
- Naopak v případě, kdy jsou zkušenosti klienta více volatilní je vhodnější účtovat pojistné ve výši μ.
- Pokud má pojistitel důvod věřit, že se vybraný klient liší od ostatních, podle kterých je určeno pojistné μ, měla by být větší váha přiřazena pojistnému $\bar{N}_i$, neboť nám může dát užitečnou informaci o výši ξ.
- V případě, že se domníváme, že všichni ostatní klienti mají stejnou hodnotu ξ, není důvod se přiklánět ke zkušenostem jednoho klienta, když μ dokáže dobře vystihnout budoucí pojistné nároky.

Plná kredibilita

- Abychom zjistili, zda je vhodné účtovat pojistné $\bar{N}_i$, tedy zvolit plnou kredibilitu, musíme se zabývat stabilitou $\bar{N}_i$.
- Můžeme říci, že $\bar{N}_i$ je stabilní, pokud pravděpodobnost, že rozdíl mezi $\bar{N}_i$ a ξ je malý vzhledem ke ξ, je blízká jedné.

Definice

Řekneme, že $\bar{N}_i$ je stabilní, pokud pro reálná čísla r kladná blízká nule a $0 < p < 1$ blízká jedné platí

$$\mathrm{P}(-r\xi \leq \bar{N}_i - \xi \leq r\xi) \geq p. \tag{1}$$

Pak je možné $\bar{N}_i$ přiřadit plnou kredibilitu.

Za p a r z Definice 1 volíme obvykle $r = 0,05$ a $p = 0,9$.

Plná kredibilita

Nerovnost (1) můžeme upravit jako

$$\mathrm{P}\left(\left|\frac{\bar{N}_i-\xi}{\frac{\sigma}{\sqrt{T_i}}}\right|\leq\frac{r\xi\sqrt{T_i}}{\sigma}\right)\geq p. \tag{2}$$

Označme

$$y_p=\inf_y\left\{\mathrm{P}\left(\frac{\bar{N}_i-\xi}{\frac{\sigma}{\sqrt{T_i}}}\leq y\right)\geq p\right\}. \tag{3}$$

V případě, že se $\bar{N}_i$ v (3) řídí spojitým rozdělením, můžeme znaménko „$\geq$“ nahradit „$=$“ a y_p pak bude splňovat

$$\mathrm{P}\left(\frac{\bar{N}_i-\xi}{\frac{\sigma}{\sqrt{T_i}}}\leq y_p\right)=p.$$

Plná kredibilita

Aby byla splněna podmínka plné kredibility, musí z nerovnosti (2) platit také $y_p \leq \frac{r\xi\sqrt{T_i}}{\sigma}$. Po úpravě dostaneme **podmínku pro plnou kredibilitu**

$$\frac{\sigma}{\xi} \leq \frac{r}{y_p}\sqrt{T_i} = \sqrt{\frac{T_i}{\lambda_0}}, \tag{4}$$

kde $\lambda_0 = (\frac{y_p}{r})^2$. Z podmínky (4) vyplývá také

$$\mathrm{Var}(\bar{N}_i) = \frac{\sigma^2}{T_i} \leq \frac{\xi^2}{\lambda_0}. \tag{5}$$

Z nerovnosti (4) je možné vyjádřit **požadavek na počet období** T_i, který musí být splněn, aby bylo možné využít plnou kredibilitu. Tedy

$$T_i \geq \lambda_0 \left(\frac{\sigma}{\xi}\right)^2. \tag{6}$$

Plná kredibilita

- V mnoha případech se rozdělení $\bar{N}_i$ aproximuje normálním rozdělením se střední hodnotou ξ a rozptylem $\frac{\sigma^2}{T_i}$.
- Je-li T_i velké můžeme aplikovat centrální limitní větu a pak

$$U = \frac{\bar{N}_i - \xi}{\sigma/\sqrt{T_i}}$$

má aproximativně standardní normální rozložení.

- Odtud dostáváme

$$p = 2\Phi(y_p) - 1,$$

kde $\Phi(x)$ označuje distribuční funkci standardizovaného normálního rozdělení.

Plná kredibilita

Dostáváme tedy

$$\Phi(y_p) = \frac{1+p}{2}$$

a y_p je $\left(\frac{(1+p)}{2}\right)$ kvantilem standardizovaného normálního rozdělení.

Plná kredibilita

Příklad

Předpokládejme, že pro každého klienta i máme k dispozici údaje o minulých pojistných nárocích $N_{i1}, N_{i2}, ..., N_{iT_i}$. Ze vzorku dat byla odhadnuta střední hodnota $\xi = \mathrm{E}(N_{it})$. Bylo zaznamenáno 10 pozorování, kde prvních 6 bylo nulových a pak byly hlášeny škody ve výši 253, 398, 439 a 756. Jaké množství dat T_i je potřeba, aby bylo možné použít plnou kredibilitu, jestliže $r = 0,05$ a $p = 0,9$.

Částečná kredibilita

- Částečnou kredibilitu je možné využít v případě, kdy nejsou splněny podmínky pro využití plné kredibility.
- Model částečné kredibility pro stanovení výše netto pojistného využívá jak minulé zkušenosti $\bar{N}_i$ tak tabulkové pojistné μ.
- Pojistné P_c pak můžeme vyjádřit jako jejich vážený průměr ve tvaru

$$P_c = Z\bar{N}_i + (1 - Z)\mu,$$

kde je potřeba určit hodnotu váhy Z z intervalu $[0, 1]$.
- Z budeme dále nazývat **kredibilitním faktorem**.
- Je-li $Z = 1$, pak se jedná o model plné kredibility.

Částečná kredibilita

- Jednou z možností, jak určit Z, je omezení volatility pojistného $\bar{N}_i$ podobně jako u plné kredibility.
- Dostáváme

$$\begin{aligned} \frac{\xi^2}{\lambda_0} &= \mathrm{Var}(P_c) \\ &= \mathrm{Var}[Z\bar{N}_i + (1-Z)\mu] \\ &= Z^2\,\mathrm{Var}(\bar{N}_i) \\ &= Z^2\frac{\sigma^2}{T_i}. \end{aligned}$$

- Odtud dostáváme $Z = \frac{\xi}{\sigma}\sqrt{\frac{T_i}{\lambda_0}}$. Musí být splněno $Z \in [0, 1]$.

$$Z = \min\left\{\frac{\xi}{\sigma}\sqrt{\frac{T_i}{\lambda_0}}, 1\right\}.$$

Částečná kredibilita

- Na rozdíl od $\bar{N}_i$ není P_C nestranným odhadem ξ.
- Kredibilita však umožňuje pracovat s vychýlenými odhady, neboť co obětuje z hlediska vychýlení, to naopak získá ve smyslu redukce střední kvadratické chyby.
- Tedy u vychýlených odhadů je míra jejich kvality dána nikoliv rozptylem, ale střední kvadratickou chybou.
- Tu však u uvedeného postupu pro odhad Z neznáme.
- A to je problém přístupu souvisejícího s teorií omezených fluktuací.

Problémy teorie s omezenými fluktuacemi

- Mezi důležité problémy patří fakt, že neexistuje teoretický model rozdělení N_{it} a tedy není důvod dávat přednost pojistnému ve tvaru $P_c = Z\bar{N}_i + (1 - Z)\mu$ před tabulkovým pojistným μ.
- Volba Z v rovnici pro P_c také není určena přesně a tudíž není přesně určeno ani P_c.
- Také nemáme k dispozici žádné vodítko, jak zvolit r a p.
- Teorie omezených fluktuací nezkoumá rozdíl mezi ξ a μ. Je potřeba si uvědomit, že také μ je pouze odhadem a tudíž nemusí být přesné.
- Měli bychom se tedy zabývat také srovnáním spolehlivosti $\bar{N}_i$ vzhledem k μ a ne pouze otázkou, jak moc je přesné $\bar{N}_i$.

Částečná kredibilita

Příklad

Uvažujme tabulkové pojistné μ ve výši 225. Určete kredibilitní odhad pojistného, uvažujeme-li, že $T_i = 10, \xi = 184.6, \lambda_0 = 1082.41, \sigma = 267.89$.

Obsah

1 Úvod

2 Teorie kredibility omezených fluktuací

3 Optimální teorie kredibility

Rizikový parametr θ

- V každé tarifní skupině zůstává jistá míra heterogenity. Proto existuje možnost, že se pojištěnec bude odlišovat od toho, co očekáváme.
- Předpokládejme, že úroveň rizika každého klienta můžeme charakterizovat rizikovým faktorem θ, přičemž θ se u jednotlivých pojištěných liší.
- Θ můžeme chápat jako vyjádření nepozorovatelných rizikových faktorů, které pak způsobují odlišnou rizikovost klienta uvnitř tarifní skupiny. Musíme si ale uvědomit, že Θ je nepozorovatelné a tudíž neznáme jeho přesnou hodnotu.
- V každé skupině jsme však schopni určit rozdělení $\pi(\theta)$ které udává pravděpodobnost jednotlivých hodnot rizikového faktoru Θ uvnitř tarifní skupiny.

Rizikový parametr θ

- **Distribuční funkce**
$$F_{\Theta}(\theta) = \mathrm{P}(\Theta \leq \theta)$$
náhodné veličiny Θ reprezentuje pravděpodobnost, že náhodně vybraný pojištěnec z dané třídy bude mít hodnotu rizikového parametru menší nebo rovnu θ.
- **Zkušenost** jednotlivých pojištěnců je ovlivněna právě hodnotou θ.
- Škody X pak vychází z podmíněného rozdělení X při daném θ.
- Tuto **podmíněnou hustotu**, resp. pravděpodobnostní funkci budeme značit
$$f_{X|\Theta}(x|\theta).$$

Rizikový parametr θ

Příklad

Předpokládejme automobilové pojištění, kde máme 2 skupiny řidičů. Dobří řidiči tvoří 75% pojištěných a mají 0,1 nebo 2 nehody s pravděpodobností 0.7, 0.2, resp. 0.1. Špatní řidiči tvoří 25% pojištěných a mají 0,1 nebo 2 nehody s pravděpodobností 0.5, 0.3, resp. 0.2. Popište tento proces a modelujte jeho závislost na neznámém parametru θ.

Rizikový parametr θ

Příklad

Počet pojistných nároků se řídí exponenciálním rozdělením se střední hodnotou $1/\Theta$. Uvnitř tarifní třídy pojištěných má parametr Θ gama rozdělení s parametrem $\alpha = 4$ a parametrem šikmosti $\beta = 0.001$. Popište matematicky tento model.

Bayesovská metodologie

- Nechť pro konkrétního pojištěného máme pozorování $\boldsymbol{X} = \boldsymbol{x}$, kde $\boldsymbol{X} = (X_1, X_2, ..., X_n)$ a $\boldsymbol{x} = (x_1, x_2, ..., x_n)$.
- Snažíme se stanovit takovou sazbu, abychom pokryli pojistný nárok nadcházejícího období, X_{n+1}.
- Budeme předpokládat, že rizikový parametr pojištěného je θ, ale jeho hodnotu neznáme.
- A dále předpokládáme nezávislost $X_1, ..., X_n$ za podmínky θ.
- Nechť X_j má podmíněné rozdělení

$$f_{X_j|\Theta}(x_j|\theta)$$

pro $j = 1, 2, ..., n, n+1$.

Bayesovská metodologie

- Pokud bychom znali hodnotu θ, pro předpověď škody X_{n+1} by bylo možné použít $f_{X_{n+1}|\Theta}(x_{n+1}|\theta)$.
- Místo toho ovšem známe $\boldsymbol{x}$, které můžeme využít k výpočtu **prediktivní distribuce**, kterou udává podmíněné rozdělení X_{n+1} při daném $\boldsymbol{X} = \boldsymbol{x}$.
- Z Bayesovy věty a předpokladu nezávislosti zkušeností z jednotlivých období za podmínky $\Theta = \theta$ dostáváme

$$f_{\mathbf{X},\Theta}(\boldsymbol{x}, \theta) = f(x_1, \ldots, x_n|\theta)\pi(\theta) = \left[\prod_{j=1}^{n} f_{X_j|\Theta}(x_j|\theta)\right]\pi(\theta).$$

Bayesovská metodologie

- Sdružené rozdělení $\boldsymbol{X}$ potom získáme integrací přes všechny hodnoty parametru θ, tedy

$$f_X(x) = \int \left[\prod_{j=1}^{n} f_{X_j|\Theta}(x_j|\theta) \right] \pi(\theta)\,\mathrm{d}\theta.$$

- Pokud budeme chtít získat sdružené rozdělení náhodného vektoru $\boldsymbol{X}' = (X_1, X_2, \ldots, X_n, X_{n+1})$, nahradíme ve vzorci v součinu na pravé straně n za $n+1$.
- Z těchto vztahů a s použitím Bayesovy věty získáme prediktivní hustotu

$$f_{X_{n+1}|\mathbf{X}}(x_{n+1}|\mathbf{x}) = \frac{1}{f_{\mathbf{X}}(\mathbf{x})} \int \left[\prod_{j=1}^{n+1} f_{X_j|\Theta}(x_j|\theta) \right] \pi(\theta)\,\mathrm{d}\theta.$$

Bayesovská metodologie

- Pro posteriorní funkci náhodné veličiny Θ platí

$$\pi_{\Theta|\mathbf{X}}(\theta|\mathbf{x}) = \frac{f_{\mathbf{X},\Theta}(\mathbf{x},\theta)}{f_{\mathbf{X}}(\mathbf{x})} = \frac{1}{f_{\mathbf{X}}(\mathbf{x})}\left[\prod_{j=1}^{n} f_{X_j|\Theta}(x_j|\theta)\right]\pi(\theta).$$

- Dosazením tohoto vztahu do prediktivní hustoty $f_{X_{n+1}|\mathbf{X}}(x_{n+1}|\mathbf{x})$ získáme

$$f_{X_{n+1}|\mathbf{X}}(x_{n+1}|\mathbf{x}) = \int f_{X_{n+1}|\Theta}(x_{n+1}|\theta)\pi_{\Theta|\mathbf{X}}(\theta|\mathbf{x})\,\mathrm{d}\theta.$$

Bayesovská metodologie

Příklad

Předpokládejme automobilové pojištění, kde máme 2 skupiny řidičů. Dobří řidiči tvoří 75% pojištěných a mají 0,1 nebo 2 nehody s pravděpodobností 0.7, 0.2, resp. 0.1. Špatní řidiči tvoří 25% pojištěných a mají 0,1 nebo 2 nehody s pravděpodobností 0.5, 0.3, resp. 0.2. Pro konkrétního pojištěného známe hodnoty $x_1 = 0$ a $x_2 = 1$. Určete prediktivní rozdělení $(X_3 | X_1 = 0, X_2 = 1)$ a pro $(\Theta | X_1 = 0, X_2 = 1)$.

Bayesovská metodologie

Příklad

Počet pojistných nároků se řídí exponenciálním rozdělením se střední hodnotou $1/\Theta$. Uvnitř tarifní třídy pojištěných má parametr Θ gama rozdělení s parametrem $\alpha = 4$ a parametrem šikmosti $\beta = 0.001$. Předpokládejme osobu se škodami 100, 950, 450. Určete prediktivní rozdělení čtvrté škody.

Střední hodnota škod

- Kromě prediktivní distribuce může pojišťovna požadovat také určení střední hodnoty škod nebo ztrát v příštím zkušenostním období.
- Pokud o klientovi nemáme žádné informace, pro střední hodnotu bude platit

$$\mu_{n+1} = \mathrm{E}(X_{n+1}) = \mathrm{E}[\,\mathrm{E}(X_{n+1}|\Theta)] = \mathrm{E}[\mu_{n+1}(\Theta)],$$

kde $\mu_{n+1}(\Theta)$ pro $\Theta = \theta$ je dáno vztahem

$$\mu_{n+1}(\theta) = \mathrm{E}(X_{n+1}|\Theta = \theta) = \int_0^\infty x_{n+1} f_{X_{n+1}|\Theta}(x_{n+1}|\theta)\,\mathrm{d}x_{n+1},$$

kde meze integrálu odpovídají faktu, že škody mohou nabývat jen nezáporných hodnot.

Střední hodnota škod

Vyjadřuje-li náhodná veličina X_i celkovou ztrátu v i-tém zkušenostním období pro $i = 1, 2, ..., n$, pak vztah

$$\mu_{n+1} = \mathrm{E}(X_{n+1}) = \mathrm{E}[\,\mathrm{E}(X_{n+1}|\Theta)] = \mathrm{E}[\mu_{n+1}(\Theta)],$$

udává **kolektivní pojistné** a vztah

$$\mu_{n+1}(\theta) = \mathrm{E}(X_{n+1}|\Theta = \theta) = \int_0^\infty x_{n+1} f_{X_{n+1}|\Theta}(x_{n+1}|\theta)\,\mathrm{d}x_{n+1}\,,$$

individuální pojistné.

Individuální pojistné

Definice

Individuální pojistné *$\mu_{n+1}(\theta)$ je pojistné, které by bylo účtováno pojištěnému s rizikovým parametrem θ v případě, že by hodnota tohoto parametru byla známá. Jedná se o očekávanou hodnotu agregovaných ztrát pojištěného v následujícím zkušenostním období při jeho dané úrovni rizika.*

- V tomto případě je rizikový parametr spojený přímo s konkrétním pojištěným. Toto pojistné je tak nejreálnějším odrazem budoucích událostí jednotlivých pojištěných.
- Střední hodnota $\mu_{n+1}(\theta)$ bývá označována jako **hypotetická**.

Individuální pojistné

- Problém s individuálním pojistným spočívá v hodnotě rizikového parametru θ který nejsme schopni v praxi vypozorovat.
- Individuální pojistné tedy nedokážeme přesně stanovit a jedinou možností je odhadnout jej z dat.

Kolektivní pojistné

Definice

Kolektivní pojistné μ_{n+1} *je pojistné, které bude účtováno pojištěnému v případě, že nevíme nic o jeho úrovni rizika. Jedná se o očekávanou hodnotu náhodné veličiny vyjadřující výši individuálního pojistného.*

- Využívá se v situacích, kdy o pojištěném nemáme žádné informace, tedy například u nového pojištěného při stanovení pojistného na první zkušenostní období.

Bayesovské pojistné

Definice

Nechť $X_1, X_2, ..., X_n$ označují zkušenost pojištěného za n zkušenostních období. ***Bayesovské pojistné*** *$B(X_1, X_2, ..., X_n)$ je potom dáno jako*

$$B(X_1, X_2, ..., X_n) = \arg\min_{g(.)} \mathrm{E}\left[(\mu_{n+1}(\Theta) - g(X_1, X_2, ..., X_n))^2\right],$$

kde $g(.)$ je nějakou funkcí dat $X_1, X_2, ..., X_n$.

- Dá se ukázat, že funkce minimalizující střední kvadratickou chybu v předchozí definici je tvaru

$$B(X_1, X_2, ..., X_n) = \mathrm{E}[\mu_{n+1}(\Theta)|X_1, X_2, ..., X_n].$$

Bayesovské pojistné

- Předpokládáme-li nezávislost X_i pro $i = 1, 2, ..., n$ při podmíněnosti náhodnou veličinou Θ, dostáváme

$$\mu_{n+1}(\Theta) = \mathrm{E}(X_{n+1}|\Theta) = \mathrm{E}(X_{n+1}|\Theta, X_1, X_2, ..., X_n).$$

- Odtud pro Bayesovské pojistné

$$\begin{aligned} B(X_1, X_2, ..., X_n) &= \mathrm{E}[\mathrm{E}(X_{n+1}|\Theta, X_1, X_2, ..., X_n)|X_1, X_2, ..., X_n] \\ &= \mathrm{E}(X_{n+1}|X_1, X_2, ..., X_n), \end{aligned}$$

tedy podmíněná střední hodnota $\mathrm{E}[\mu_{n+1}(\Theta)|\boldsymbol{X}] = \mathrm{E}(X_{n+1}|\boldsymbol{X})$ bude minimalizovat také střední kvadratickou chybu

$$\mathrm{E}\left[(X_{n+1} - g(X_1, X_2, ..., X_n))^2\right].$$

Bayesovské pojistné

Příklad

Předpokládejme automobilové pojištění, kde máme 2 skupiny řidičů. Dobří řidiči tvoří 75% pojištěných a mají 0,1 nebo 2 nehody s pravděpodobností 0.7, 0.2, resp. 0.1. Špatní řidiči tvoří 25% pojištěných a mají 0,1 nebo 2 nehody s pravděpodobností 0.5, 0.3, resp. 0.2. Pro konkrétního pojištěného známe hodnoty $x_1 = 0$ a $x_2 = 1$. Určete Bayesovské pojistné.

Bayesovské pojistné

Příklad

Počet pojistných nároků se řídí exponenciálním rozdělením se střední hodnotou $1/\Theta$. Uvnitř tarifní třídy pojištěných má parametr Θ gama rozdělení s parametrem $\alpha = 4$ a parametrem šikmosti $\beta = 0.001$. Předpokládejme osobu se škodami 100, 950, 450. Určete Bayesovské pojistné, víme-li, že

$$\pi(\theta|100, 950, 450) = \frac{\theta^6 \, \mathrm{e}^{-2500\theta} 2500^7}{\Gamma(7)}$$

Kredibilitní pojistné

- Jedná se o alternativu k Bayesovskému a individuálnímu pojistnému.
- Dochází k aproximaci $\mu_{n+1}(\theta)$ pomocí lineární funkce dat z minulosti pojištěného, tedy

$$\alpha_0 + \sum_{i=1}^{n} \alpha_i X_i,$$

  kde $\alpha_0, \alpha_1, ..., \alpha_n$ musíme určit.
- Vycházíme z definice bayesovského pojistného, tedy neznámé parametry α_i pro $i = 0, 1, ..., n$ zvolíme tak, abychom minimalizovali střední kvadratickou chybu.

Kredibilitní pojistné

Definice

Nechť $X_1, X_2, ..., X_n$ označují zkušenost pojištěnce z n zkušenostních období. ***Kredibilitní pojistné*** *$C(X_1, X_2, ..., X_n)$ je potom lineární funkcí $X_1, X_2, ..., X_n$ tedy*

$$C(X_1, X_2, ..., X_n) = \tilde{\alpha}_0 + \sum_{i=1}^{n} \tilde{\alpha}_i X_i,$$

kde $\tilde{\alpha}_i$ pro $i = 1, 2, ..., n$ minimalizují střední kvadratickou chybu

$$Q = \mathrm{E}\left\{\left[\mu_{n+1}(\Theta) - \alpha_0 - \sum_{i=1}^{n} \alpha_i X_i\right]^2\right\}.$$

Kredibilitní pojistné

- Naší snahou je minimalizovat Q z rovnice

$$Q = \mathrm{E}\left\{\left[\mu_{n+1}(\Theta) - \alpha_0 - \sum_{i=1}^{n} \alpha_i X_i\right]^2\right\}.$$

- Použijeme parciální derivace, které položíme rovny nule

$$\begin{aligned} \frac{\partial Q}{\partial \alpha_0} &= 0 \\ \mathrm{E}\left\{2\left[\mu_{n+1}(\Theta) - \alpha_0 - \sum_{i=1}^{n} \alpha_i X_i\right](-1)\right\} &= 0 \\ \mathrm{E}[\mu_{n+1}(\Theta)] &= \tilde{\alpha}_0 + \sum_{i=1}^{n} \tilde{\alpha}_i \,\mathrm{E}(X_i), \end{aligned}$$

kde $\tilde{\alpha}_0, \tilde{\alpha}_1, \ldots, \tilde{\alpha}_n$ jsou hodnoty $\alpha_0, \alpha_1, \ldots, \alpha_n$ minimalizující Q.

Kredibilitní pojistné

Vzhledem k platnosti

$$\mathrm{E}(X_{n+1}) = \mathrm{E}[\,\mathrm{E}(X_{n+1}|\Theta)] = \mathrm{E}[\mu_{n+1}(\Theta)],$$

můžeme psát

$$\mathrm{E}(X_{n+1}) = \tilde{\alpha}_0 + \sum_{i=1}^{n} \tilde{\alpha}_i\,\mathrm{E}(X_i).$$

Kredibilitní pojistné

Analogicky můžeme postupovat při parciálních derivacích podle ostatních α_j, kde $j = 1, 2, \ldots n$, tedy

$$
\begin{aligned}
\frac{\partial Q}{\partial \alpha_j} &= 0 \\
\mathrm{E}\left\{2\left[\mu_{n+1}(\Theta) - \alpha_0 - \sum_{i=1}^{n} \alpha_i X_i\right](-X_j)\right\} &= 0 \\
\mathrm{E}[\mu_{n+1}(\Theta) X_j] = \tilde{\alpha}_0\, \mathrm{E}(X_j) &+ \sum_{i=1}^{n} \tilde{\alpha}_i\, \mathrm{E}(X_i X_j),
\end{aligned}
$$

Kredibilitní pojistné

Zároveň však využitím vlastností střední hodnoty a nezávislosti X_j a X_{n+1} při podmíněnosti Θ můžeme psát

$$
\begin{aligned}
\mathrm{E}[\mu_{n+1}(\Theta)X_j] &= \mathrm{E}\{\,\mathrm{E}[X_j\mu_{n+1}(\Theta)|\Theta]\} \\
&= \mathrm{E}\{\mu_{n+1}(\Theta)\,\mathrm{E}[X_j|\Theta]\} \\
&= \mathrm{E}[\,\mathrm{E}(X_{n+1}|\Theta)\,\mathrm{E}(X_j|\Theta)] \\
&= \mathrm{E}[\,\mathrm{E}(X_{n+1}X_j|\Theta)] \\
&= \mathrm{E}(X_jX_{n+1}),
\end{aligned}
$$

Z toho získáme

$$
\mathrm{E}(X_jX_{n+1}) = \tilde{\alpha}_0\,\mathrm{E}(X_j) + \sum_{i=1}^{n} \tilde{\alpha}_i\,\mathrm{E}(X_iX_j).
$$

Kredibilitní pojistné

Vynásobením

$$\mathrm{E}(X_{n+1}) = \tilde{\alpha}_0 + \sum_{i=1}^{n} \tilde{\alpha}_i \, \mathrm{E}(X_i).$$

střední hodnotou $\mathrm{E}(X_j)$ a odečtením od

$$\mathrm{E}(X_j X_{n+1}) = \tilde{\alpha}_0 \, \mathrm{E}(X_j) + \sum_{i=1}^{n} \tilde{\alpha}_i \, \mathrm{E}(X_i X_j).$$

získáme

$$\mathrm{Cov}(X_j, X_{n+1}) = \sum_{i=1}^{n} \tilde{\alpha}_i \, \mathrm{Cov}(X_i, X_j), \quad j = 1, \ldots, n.$$

Kredibilitní pojistné

Pokud dáme dohromady rovnici

$$\mathrm{E}(X_{n+1}) = \tilde{\alpha}_0 + \sum_{i=1}^{n} \tilde{\alpha}_i \, \mathrm{E}(X_i)$$

a n rovnic

$$\mathrm{Cov}(X_j, X_{n+1}) = \sum_{i=1}^{n} \tilde{\alpha}_i \, \mathrm{Cov}(X_i, X_j), \quad j = 1, \ldots, n,$$

získáme soustavu $n+1$ tzv. **normálních rovnic.** Jejím řešením obdržíme $\tilde{\alpha}_0, \tilde{\alpha}_1, \ldots, \tilde{\alpha}_n$ a dostaneme tak kredibilitní pojistné

$$C(X_1, X_2, \ldots, X_n) = \tilde{\alpha}_0 + \sum_{i=1}^{n} \tilde{\alpha}_i X_i.$$

Kredibilitní pojistné

- Parciálním derivováním podle $\alpha_0, \alpha_1, \ldots, \alpha_n$ vztahů
$$Q_1 = \mathrm{E}\left\{\left[\mathrm{E}(X_{n+1}|\mathbf{X}) - \alpha_0 - \sum_{i=1}^{n} \alpha_i X_i\right]^2\right\}$$
a
$$Q_2 = \mathrm{E}\left[\left(X_{n+1} - \alpha_0 - \sum_{i=1}^{n} \alpha_i X_i\right)^2\right]$$
získáme tutéž soustavu $n+1$ normálních rovnic.
- Z toho vyplývá, že stejné hodnoty $\tilde{\alpha}_0, \tilde{\alpha}_1, \ldots, \tilde{\alpha}_n$, které minimalizují Q minimalizují i Q_1 a Q_2.
- Tedy kredibilitní pojistné je nejlepším lineárním odhadem nejen pro $\mu_{n+1}(\Theta)$, ale i pro bayesovské pojistné $\mathrm{E}(X_{n+1}|\boldsymbol{X})$ a náhodnou veličinu X_{n+1}.

Kredibilitní pojistné

Příklad

Nechť pro všechna X_i, kde $i = 1, 2, ..., n + 1$, platí, že $\mathrm{E}(X_i) = \mu$, $\mathrm{Var}(X_i) = \sigma^2$ a pro $i \neq j$ platí $\mathrm{Cov}(X_i, X_j) = \rho\sigma^2$, tedy $-1 < \rho < 1$ odpovídá korelačnímu koeficientu. Určete kredibilitní pojistné $\tilde{\alpha}_0 + \sum_{i=1}^{n} \tilde{\alpha}_i X_i$.

Bühlmannův model

- **Bühlmannův model** je nejjednodušším kredibilitiním modelem.
- Předpokládá, že u všech pojištěnců je střední hodnota a rozptyl minulých škod $X_1, X_2, ..., X_n$ stejný při podmíněnosti náhodnou veličinou Θ.
- Zároveň se ještě požaduje, aby veličiny $X_1, X_2, ..., X_n$ měly identické rozdělení a byly nezávislé za podmíněnosti Θ.
- Tedy můžeme definovat **hypotetickou** střední hodnotu

$$\mu(\theta) = \mathrm{E}(X_i|\Theta = \theta)$$

a **procesní** rozptyl

$$\nu(\theta) = \mathrm{Var}(X_i|\Theta = \theta).$$

Bühlmannův model

Dále definujme očekávanou hodnotu hypotetických středních hodnot jako

$$\mu = \mathrm{E}[\mu(\Theta)],$$

očekávanou hodnotu procesního rozptylu jako

$$\nu = \mathrm{E}[\nu(\Theta)]$$

a rozptyl hypotetických středních hodnot

$$a = \mathrm{Var}[\mu(\Theta)].$$

Přitom μ volíme jako pojistné tehdy, když nemáme žádné informace o klientovi a tedy žádné informace o $\mu(\theta)$.

Bühlmannův model

Pro náhodnou veličinu X_i vyjadřující škody klientů platí

$$\mathrm{E}(X_i) = \mathrm{E}[\,\mathrm{E}(X_i|\Theta)] = \mathrm{E}[\mu(\Theta)] = \mu$$

a dále

$$\begin{aligned} \mathrm{Var}(X_i) &= \mathrm{E}[\,\mathrm{Var}(X_i|\Theta)] + \mathrm{Var}[\,\mathrm{E}(X_i|\Theta)] \\ &= \mathrm{E}[\nu(\Theta)] + \mathrm{Var}[\mu(\Theta)] \\ &= \nu + a \end{aligned}$$

a pro $i \neq j$

$$\mathrm{Cov}(X_i, X_j) = a$$

Bühlmannův model

Dosadíme-li tyto výsledky do rovnic

$$\mathrm{E}(X_{n+1}) = \tilde{\alpha}_0 + \sum_{i=1}^{n} \tilde{\alpha}_i \, \mathrm{E}(X_i)$$

a

$$\mathrm{Cov}(X_j, X_{n+1}) = \sum_{i=1}^{n} \tilde{\alpha}_i \, \mathrm{Cov}(X_i, X_j), \quad j = 1, ..., n,$$

získáme hodnoty parametrů $\tilde{\alpha}_0, \tilde{\alpha}_1, ... \tilde{\alpha}_n$, které minimalizují Q v rovnici

$$Q = \mathrm{E}\left\{ \left[\mu_{n+1}(\Theta) - \alpha_0 - \sum_{i=1}^{n} \tilde{\alpha}_i X_i \right]^2 \right\}.$$

Bühlmannův model

Bude platit

$$\tilde{\alpha}_0 = \frac{\nu}{\nu + na}\mu$$

a pro $i = 1, 2, ..., n$

$$\tilde{\alpha}_i = \frac{a}{\nu + na}.$$

Kredibilitní pojistné pak bude tvaru

$$\begin{aligned}\tilde{\alpha}_0 + \sum_{i=1}^{n} \tilde{\alpha}_i X_i &= \frac{\nu}{\nu + na}\mu + \sum_{i=1}^{n} \frac{a}{\nu + na} X_i \\ &= \frac{\nu}{\nu + na}\mu + \frac{na}{\nu + na}\bar{X}.\end{aligned}$$

Bühlmannův model

Pokud položíme

$$k = \frac{\nu}{a} \quad \text{a} \quad Z = \frac{n}{n+k}$$

a dosadíme do předchozí rovnice, dostaneme kredibilitní pojistné tvaru

$$C_B(X_1, X_2, ..., X_n) = Z\bar{X} + (1 - Z)\mu$$

což odpovídá tvaru pojistného z parciální kredibility. Kredibilitní faktor $Z = \frac{n}{n+k}$ s volbou $k = \frac{\nu}{a}$ se označuje jako **Bühlmannův kredibilitní faktor.**

Bühlmannův model

Příklad

Předpokládejme automobilové pojištění, kde máme 2 skupiny řidičů. Dobří řidiči tvoří 75% pojištěných a mají 0,1 nebo 2 nehody s pravděpodobností 0.7, 0.2, resp. 0.1. Špatní řidiči tvoří 25% pojištěných a mají 0,1 nebo 2 nehody s pravděpodobností 0.5, 0.3, resp. 0.2. Pro konkrétního pojištěného známe hodnoty $x_1 = 0$ a $x_2 = 1$. Určete Bühlmannův odhad $\mathrm{E}(X_3|0,1)$.

Bühlmannův model

Příklad

Počet škod náhodně vybraného pojištěného má Poissonovo rozdělení s parametrem θ, pro jehož priorní rozdělení platí , že $\pi(\theta) = 3\theta^{-4}$, kde $\theta > 1$. Předpokládejme, že hodnota rizikového parametru se u jednotlivých klientů v čase nemění. V předešlých dvou letech se daný klient dopustil 20 škod. Odvoďte Bühlmannův odhad pro očekávanou četnost škod daného pojištěného v následujícím období.

Bühlmannův model

Příklad

Výše škod klientů s rizikovým parametrem θ_1 je $100, 1000$ a 20000 s pravděpodobnostmi $\frac{1}{2}, \frac{3}{10}$ a $\frac{1}{5}$. U klientů s rizikovým parametrem θ_2 se pro stejné výše škod jedná postupně o pravděpodobnosti $\frac{7}{10}, \frac{1}{5}$ a $\frac{1}{10}$. Výskyt klientů s rizikovým parametrem θ_1 je dvakrát větší než výskyt klientů s parametrem θ_2. U náhodného pojištěného s neznámým rizikovým parametrem nastala v minulém zkušenostním období škoda ve výši 100. Odvoďte Bühlmannův odhad pro očekávanou výši škody daného pojištěného v následujícím období. Předpokládáme, že hodnota rizikového parametru se u pojištěného s časem nemění.