Matematicko-statistické metody v pojišťovnictví

Martin Kolář

Ústav matematiky a statistiky PřF MU

March 4, 2021

Úvod

Pojistná = aktuárská matematika

Pojistný matematik = aktuár

NYT Best jobs (2015) ... 1. Actuary

Aktuárské organizace: SOA, AAE, ČSpA

ČSpA "uznává" studium na MFF UK, na PřF MU a na VŠE

Aktuárské zkoušky ... Core Syllabus AAE, www.actuary.eu

Historie pojistné matematiky u nás:

1900 Matyáš Lerch (zakladatel ÚMS)

1990 Petr Mandl, obnovení ČSpA

2016 Solvency 2

Životní pojištění ... deterministické metody, Life tables

Neživotní pojištění ... stochastické metody

Teorie rizika, teorie ruinování, teorie kredibility, pricing (GLM, decision trees, neural nets)

Hlavní cíl : předpovídat pomocí pravděpodobnostního modelu budoucí výdaje pojišťovny.

Model ... zjednodušený matematický popis nějakého systému

Sestavený na základě znalostí a zkušeností aktuára a dat z minulosti.

Modelovací proces:

– výběr modelu

– kalibrace

– validace

– porovnání modelů

– Adaptace modelu – úprava parametrů, např. vliv inflace

Data slouží ke kalibraci parametrů modelu

“Alternativa” : empirický přístup, budoucnost bude stejná jako minulost (až na inflaci).

Rovnováha dvou faktorů: jednoduchost x soulad s daty (Occamova břitva)

“Pokud pro nějaký jev existuje vícero vysvětlení, je lépe upřednostňovat to nejméně komplikované”.

Základní nástroj - teorie pravděpodobnosti

Příklady náhodných veličin v pojistné matematice:

– zda nastala pojistná událost z dané smlouvy (0 nebo 1)

– čas kdy nastala pojistná událost

– velikost ztráty z pojistné události

– celkový počet pojistných nároků z jedné smlouvy (portfolia)

– velikost pojistného plnění z jedné smlouvy (portfolia)

Musíme tedy umět:

1. Modelovat celkový počet nároků
2. Modelovat velikost jednotlivých nároků
3. Dát to dohromady – Kolektivní teorie rizika

(+ závislost na čase ... stochastické procesy)

Literatura: Klugman, Panjer, Willmot: Loss models (KPW)
DP Mai Truongové (DMT)

Příklady modelů

Model 1.

Počet pojistných nároků za rok z jedné pojistné smlouvy.

$P(X = n) =$

– $0,5$ pro $n = 0$

– $0,25$ pro $n = 1$

– $0,12$ pro $n = 2$

– $0,08$ pro $n = 3$

– $0,05$ pro $n = 4$

Může být např. empirický výběr z teoretického rozdělení (např. Poissonova).

Model 2.

X ... Suma vyplacená při pojistném nároku z povinného ručení

Distribuční funkce je

$$F(x) = 1 - \left(\frac{2000}{x + 2000}\right)^3$$

pro $x \geq 0$, jinak $F(x) = 0$.

Paretovo rozdělení (silné chvosty)

Model 3.

X ... celková suma pojistného plnění vyplacená na pojistné smlouvě za 1 rok.

Diskrétní část:

$$P(X = 0) = 0,9$$

Spojitá část: Pro $x > 0$ má $P(X = x)$ distribuční funkci

$$F(x) = 1 - 0,1e^{-0,001x}$$

Tedy "hustota" je $f(x) = 0,0001e^{-0,001x}$

– Integrál z $f(x)$ není 1, ale 0,1.

Opakování z teorie pravděpodobnosti

Diskrétní náhodná proměnná (náhodná veličina) je funkce

$$X : \Omega \to \{x_1, x_2, \dots\} \subseteq \mathbb{R},$$

kde $\{x_1, x_2, \dots\}$ je diskrétní podmnožina $\mathbb{R}$.

Definice: *Pravděpodobnostní funkce náhodné veličiny X je definována jako*

$$f(x) = P(X = x).$$

Definice: *Distribuční funkce náhodné veličiny X je*

$$F(x) = P(X \leq x).$$

Definice: *Očekávání* (střední hodnota) náhodné veličiny X s pravděpodobnostní funkcí $f(x)$ je definována jako

$$E(X) = \sum_{x:\, f(x)>0} x f(x),$$

Očekávání můžeme vypočítat také pomocí vztahu

$$E(X) = \sum_{\omega \in \Omega} X(\omega) P(\omega).$$

Definice: Je-li k přirozené číslo, *k-tý moment* m_k náhodné veličiny X je definován jako

$$m_k = E(X^k).$$

Definice:*k-tý centrální moment* σ_k je definován jako

$$\sigma_k = E((X - m_1)^k).$$

Speciálně,

$$m_1 = \mu = E(X)$$

je *střední hodnota* a

$$\sigma_2 = E((X - E(X))^2)$$

je *rozptyl* (variance). Tedy $\sigma_2 = \sigma^2$, kde $\sigma = \sqrt{\sigma_2}$ je střední směrodatná odchylka.

Šikmost

$$\gamma_1 = \frac{\sigma_3}{\sigma^3}$$

Symetrické rozdělení má šikmost nulovou. Záporná - levý chvost je silnější, kladná naopak.

Špičatost

$$\gamma_2 = \frac{\sigma_4}{\sigma^4}$$

– Normální rozdělení má špičatost rovnu třem.

– Nejmenší špičatost má alternativní rozdělení, rovnu jedné.

– Velká špičatost ... silné chvosty

Koeficient variace charakterizuje variabilitu rozdělení

$$K.V. = \frac{\sigma}{\mu}$$

Spojité náhodné veličiny

Distribuční funkce jako u diskrétních náhodných veličin

3 další důležité funkce:

Pro hustotu $f(x)$ platí

$$P(X \in [a, b]) = \int_a^b f(x)dx$$

– $f(x)$ nemá význam pravděpodobnosti (může být větší než 1)

– platí ale intuitivně

$$f(x)dx \simeq P(X \in [x, x + dx]).$$

Definice: Funkce přežití $S_X(x)$ je definovaná jako

$$S_X(x) = P(X > x) = 1 - F_X(x).$$

Míra rizika

– V životním pojištění se používá termín *intenzita úmrtnosti*

Uvažujme spojitou náhodnou veličinu X a distribuční funkcí F a hustotou f.

Definice: Funkce míry rizika (hazard rate function) je definována jako

$$\lambda(t) = \frac{f(t)}{S(t)} \tag{1}$$

Představme si, že skoumáme životnost nějakého stroje (součástky), a předpokládejme, že stroj již funguje t hodin.

Chceme spočítat pravděpodobnost, že nevydrží další časový úsek dt, tedy

$$P\{X \in (t, t + dt) | X > t\}, \tag{2}$$

kde n.v. X modeluje čas poruchy stroje. Máme

$$P\{X \in (t, t + dt) | X > t\} = \frac{P(X \in (t, t + dt) \wedge X > t)}{P(X > t)} \tag{3}$$

což je rovno

$$\frac{X \in (t, t+dt)}{P(X > t)} \simeq \frac{f(t)dt}{S(t)} \tag{4}$$

$$= \lambda(t)dt. \tag{5}$$

Tedy $\lambda(t)$ reprezentuje *intenzitu pravděpodobnosti*, že t-letý stroj přestane fungovat v čase t.

Generující funkce

Uvažujeme posloupnost reálných čísel

$$a = \{a_n;\ n = 0,\ 1,\ 2,\ ..\}\,.$$

Generující funkce posloupnosti a je funkce daná součtem mocninné řady

$$G_a(s) = \sum_{n=0}^{\infty} a_n s^n$$

pro $s \in \mathbb{R}$, pro která řada konverguje.

Posloupnost a dostaneme z generující funkce G_a zpět vztahem

$$a_n = \frac{G_a^{(n)}(0)}{n!},$$

kde $G_a^{(n)}(0)$ je n-tá derivace G_a v bodě 0.

Dále budeme definovat generující funkci diskrétní náhodné veličiny.

Definice: Nechť N je diskrétní náhodná veličina s hodnotami na množině $\mathbb{N} \cup \{0\}$ a nechť $p_N(k)$ je její pravděpodobnostní funkce. Potom generující funkce náhodné veličiny N je definována vztahem

$$G_N(s) = \sum_{k=0}^{\infty} p_N(k) \cdot s^k = \mathrm{E}(s^N), \quad s \in \mathbb{R}. \tag{6}$$

Příklady generujících funkcí náhodných veličin:

1. Konstantní náhodná veličina. $P(N = k) = 1$, kde $k \in \mathbb{N} \cup \{0\}$. Máme

$$G_N(s) = 1s^k = s^k.$$

2. Bernoulliho náhodná veličina. $P(N = 1) = p$ a $P(N = 0) = 1 - p$. Tedy

$$G_N(s) = ps^1 + (1 - p)s^0 = 1 - p + ps.$$

3. Geometrické rozdělení. $P(N = k) = p(1-p)^k$ pro $k \in \mathbb{N}_{\nvdash}$.

Počet neúspěchů před prvním úspěchem

Dostaneme

$$G_N(s) = \sum_{i=0}^{\infty} p_N(n)s^n = \sum_{n=0}^{\infty} p(1-p)^n s^n = \sum_{n=0}^{\infty} p\left[(1-p)s\right]^n =$$

$$= \frac{p}{1-(1-p)s} = \frac{p}{1-s+sp}.$$

Charakteristiky náhodných veličin a generující funkce

Základní charakteristiky n.v., $E(X)$ a $Var(X)$, lze snadno spočítat pomocí $G_X(s)$.

Věta: Nechť X je náhodná veličina s generující funkcí $G_X(s)$. Pak platí:

$E(X) = G_X'(1)$.

Obecně,

$$E(X(X-1)...(X-k+1)) = G_X^{(k)}(1)$$

(tzv. k-tý faktoriální moment).

Součty náhodných veličin

Věta Nechť X a Y jsou nezávislé náhodné veličiny. Pak

$$G_{X+Y}(s) = G_X(s)G_Y(s).$$

Důkaz:

$$E(s^{X+Y}) = E(s^X s^Y) = E(s^X)E(s^Y)$$

první rovnost plyne z vlastností exponenciály, druhá z nezávislosti X a Y.

Moment generující funkce

Definice: *Moment generující funkce* náhodné veličiny X je definována vztahem

$$M(t) = E\left(e^{tX}\right)$$

pro $t \geq 0$.

Je-li X spojitá náhodná veličina s hustotou f, pak

$$M(t) = \int_{-\infty}^{\infty} e^{tx} f(x)\, dx.$$

Pro každé $k \in \mathbb{N}$ je

$$E\left(X^k\right) = M^{(k)}(0).$$

Opravdu, derivujme integrál podle parametru,

$$M'(t) = \int_{-\infty}^{\infty} x e^{tx} f(x)\, dx,$$

tedy

$$M'(0) = \int_{-\infty}^{\infty} x f(x)\, dx = E(X).$$

Analogicky k-násobným derivováním dostaneme

$$M^{(k)}(t) = \int_{-\infty}^{\infty} x^k e^{tx} f(x)\, dx.$$

Tedy

$$M^{(k)}(0) = E\left(X^k\right)$$