Kapitoly z pojistné matematiky

Kolektivní model rizika - 2. část

Silvie Zlatošová

2021

Obsah

1. Rozdělení počtu pojistných událostí N
2. Rozdělení výše pojistných nároků X

Obsah

1. Rozdělení počtu pojistných událostí N
2. Rozdělení výše pojistných nároků X

Poissonovo rozdělení

Základním rozdělením počtu pojistných událostí je **Poissonovo rozdělení**. Popisuje výskyt řídkých jevů za určitou jednotku času.

Definice

Diskrétní náhodná veličina N má Poissonovo rozdělení s parametrem $\lambda > 0$, píšeme $N \sim Po(\lambda)$, jestliže je pravděpodobnostní funkce tvaru

$$\mathrm{P}(N = k) = \begin{cases} \frac{\lambda^k \mathrm{e}^{-\lambda}}{k!}, & k = 0, 1, 2, \ldots, \\ 0, & \textit{jinak}. \end{cases}$$

.

Poissonovo rozdělení

Generující funkce:

$$G_N(s) = \mathrm{e}^{\lambda(s-1)}$$

Střední hodnota a rozptyl:

$$\mathrm{E}(N) = \lambda$$

$$\mathrm{Var}(N) = \lambda.$$

Věta

Nechť $N_1, N_2, ..., N_n$ jsou nezávislé náhodné veličiny z Poissonova rozdělení s parametry $\lambda_1, \lambda_2, ..., \lambda_n$. Pak $N = N_1 + N_2 + ... + N_n$ má také Poissonovo rozdělení s parametrem $\lambda = \lambda_1 + \lambda_2 + ... + \lambda_n$.

Geometrické rozdělení

Vyjadřuje počet neúspěchů před prvním úspěchem.

Definice

Diskrétní náhodná veličina N má geometrické rozdělení s parametrem $p \in (0,1)$, píšeme $N \sim Ge(p)$, jestliže je pravděpodobnostní funkce tvaru

$$\mathrm{P}(N = k) = \begin{cases} p(1-p)^k, & k = 0, 1, 2, \ldots, \\ 0, & \text{jinak}, \end{cases}$$

resp. tvaru

$$\mathrm{P}(N = k) = \begin{cases} \frac{\beta^k}{(1+\beta)^{k+1}}, & k = 0, 1, 2, \ldots, \\ 0, & \text{jinak} \end{cases}$$

pro $p = \frac{1}{1+\beta}$ *tj.* $\beta = \frac{1-p}{p} > 0$.

Geometrické rozdělení

Generující funkce:

$$G_N(s) = \frac{p}{1 - s(1-p)} = \frac{1}{1 - \beta(s-1)}$$

Střední hodnota a rozptyl:

$$\mathrm{E}(N) = \frac{1-p}{p} = \beta$$

$$\mathrm{Var}(N) = \frac{1-p}{p^2} = \beta(1+\beta)$$

Binomické rozdělení

Definice

Diskrétní náhodná veličina N má binomické rozdělení s parametry $n \in \mathbb{N}$ a $p \in (0,1)$, píšeme $N \sim Bi(n,p)$, pokud je pravděpodobnostní funkce tvaru

$$\mathrm{P}(N=k) = \begin{cases} \binom{n}{k} p^k (1-p)^{n-k}, & k = 0,1,2,\dots,n \\ 0, & \textit{jinak.} \end{cases}$$

Generující funkce:

$$G_N(s) = (1 + p(s-1))^n$$

Střední hodnota a rozptyl:

$$\mathrm{E}(N) = np$$

$$\mathrm{Var}(N) = np(1-p).$$

Příklad

Dokažte, že platí:
Jestliže $n \to \infty$ a $p \to 0$ takovým způsobem, že $np = \lambda$, kde λ je kladná konstanta, pak k-tý člen binomického rozdělení

$$\binom{n}{k} p^k (1-p)^{n-k},$$

konverguje ke k-tému členu Poissonova rozdělení

$$\mathrm{e}^{-\lambda} \frac{\lambda^k}{k!}.$$

Negativně binomické rozdělení

Definice

Diskrétní náhodná veličina N má negativně binomické rozdělení s parametry $m > 0$ a $p \in (0,1)$, píšeme $N \sim NeBi(m,p)$, jestliže je pravděpodobnostní funkce tvaru

$$\mathrm{P}(N=k) = \begin{cases} \binom{k+m-1}{k} p^m (1-p)^k, & k = 0,1,2,\dots, \\ 0, & \textit{jinak,} \end{cases}$$

resp. tvaru

$$\mathrm{P}(N=k) = \begin{cases} \binom{k+m-1}{k} \left(\frac{1}{1+\beta}\right)^m \left(\frac{\beta}{1+\beta}\right)^k, & k = 0,1,2,\dots, \\ 0, & \textit{jinak} \end{cases}$$

pro $p = \frac{1}{1+\beta}$ *tj.* $\beta = \frac{1-p}{p} > 0$.

Negativně binomické rozdělení

Vyjadřuje počet neúspěchů před m-tým úspěchem. Jedná se o zobecnění geometrického rozdělení.

Generující funkce:

$$G_N(s) = \left(\frac{p}{1-s(1-p)}\right)^m = \left(\frac{1}{1-\beta(s-1)}\right)^m.$$

Střední hodnota a rozptyl:

$$\mathrm{E}(N) = m\frac{1-p}{p} = m\beta$$

$$\mathrm{Var}(N) = m\frac{1-p}{p^2} = m\beta(1+\beta).$$

Modely počtu pojistných událostí

Rozhodujeme-li se při modelování počtu pojistných událostí jaké pravděpodobnostní rozdělení použít, pak k tomu můžeme využít vztah mezi číselnými charakteristikami náhodné veličiny N.

- Poissonovo rozdělení: $\mathrm{E}(N) = \mathrm{Var}(N)$ ... **equidispersion**,
- Negativně binomické rozdělení: $\mathrm{E}(N) < \mathrm{Var}(N)$ ... **overdispersion**,
- Binomické rozdělení: $\mathrm{E}(N) > Var(N)$ ... **underdispersion**.

Třída rozdělení $(a, b, 0)$

Definice

Nechť $p_N(k) = P(N = k)$ je pravděpodobnostní funkce diskrétní náhodné veličiny N. Řekneme, že je členem třídy rozdělení $(a, b, 0)$ jestliže existují reálné konstanty a, b takové, že platí

$$\frac{p_N(k)}{p_N(k-1)} = a + \frac{b}{k}, \quad pro\ k = 1, 2, 3, \ldots \tag{1}$$

- Pravděpodobnost $p_N(0)$ dopočítáme z $\sum_{k=0}^{\infty} p_N(k) = 1$
- Do třídy obecných rozdělení $(a, b, 0)$ patří právě Poissonovo rozdělení, geometrické rozdělení, binomické rozdělení a negativně binomické rozdělení.

Třída rozdělení $(a, b, 0)$

Příklad

Určete reálné koeficienty a, b

- Poissonova rozdělení,
- negativně binomického rozdělení,
- binomického rozdělení.

Třída rozdělení $(a, b, 0)$

- Pro konkrétní datový soubor s velkým množstvím pozorování lze určit vhodný model pomocí vzorce (1).
- Vzorec přepíšeme do tvaru

$$\frac{p_N(k)}{p_N(k-1)} \cdot k = ak + b, \quad k = 1, 2, 3, \ldots$$

- podíl $\frac{p_N(k)}{p_N(k-1)}$ odhadneme na základě pozorovaných četností n_k a n_{k-1} hodnot k a $k-1$

$$\frac{\widehat{p_N(k)}}{\widehat{p_N(k-1)}} \cdot k = \frac{n_k}{n_{k-1}} \cdot k.$$

Třída rozdělení $(a, b, 0)$

- Graf procházející body $\left[k, k \cdot \frac{n_k}{n_{k-1}}\right]$ by měl přibližně vykazovat lineární průběh
- podle směrnice a dané přímky zvolíme vhodný model:
  - nulová směrnice - Poissonovo rozdělení,
  - záporná směrnice - binomické rozdělení,
  - kladná směrnice - negativně binomické rozdělení.

Třída rozdělení $(a, b, 0)$

Příklad

počet nehod k	počet smluv n_k
0	7840
1	1317
2	239
3	42
4	14
5	4
6	4
7	1
8 a více	0
Celkem	9461

Pro pojistný kmen o 9461 klientech máme v tabulce zaznamenány pozorované četnosti škod na jednu smlouvu.

Určete rozdělení z třídy $(a, b, 0)$ vhodné k modelování počtu nehod.

Třída rozdělení $(a, b, 1)$

- Rozdělení třídy $(a, b, 0)$ často nepopisuje adekvátně data, s nimiž se v praxi setkáváme.
- Hlavní příčinou je neschopnost rozdělení třídy $(a, b, 0)$ vystihnout tvar dat v jistých částech rozdělení, zejména **hodnotu v nule**.
- Budeme se věnovat rozložení pravděpodobnosti v nule, tedy pravděpodobnosti, že nenastane **žádná pojistná událost** během stanoveného časového období.
- U pojištění s malou pravděpodobností výskytu škod (pojištění odpovědnosti, pojištění nemovitosti, aj.) je pravděpodobnost v nule největší.
- Úpravou pravděpodobnosti v nule lze třídu rozdělení $(a, b, 0)$ rozšířit na třídu $(a, b, 1)$.

Třída rozdělení $(a, b, 1)$

Definice

Nechť $p_N(k)$ je pravděpodobnostní funkce diskrétní náhodné veličiny N. Řekneme, že je členem třídy rozdělení $(a, b, 1)$ za předpokladu, že existují konstanty $a, b \in \mathbb{R}$ takové, že

$$\frac{p_N(k)}{p_N(k-1)} = a + \frac{b}{k}, \quad pro\ k = 2, 3, \ldots \tag{2}$$

Suma $\sum_{k=1}^{\infty} p_N(k)$ může nabývat libovolných hodnot na intervalu $(0, 1\rangle$, zbývající pravděpodobnost je v $k = 0$, jelikož

$$p_N(0) + \sum_{k=1}^{\infty} p_N(k) = 1$$

Třída rozdělení $(a, b, 1)$

U třídy $(a, b, 1)$ rozlišujeme dvě podtřídy:

- Je-li $p_N(0) = 0$ jedná se o **rozdělení useknuté v nule** s pravděpodobnostní funkcí $p_N^T(k)$.
- Je-li $p_N(0) > 0$ jedná se o **rozdělení modifikované v nule** s pravděpodobnostní funkcí $p_N^M(k)$.

Rozdělení modifikovaná v nule

Na tento typ rozdělení lze pohlížet jako na směs rozdělení třídy $(a, b, 0)$ a degenerovaného rozdělení se všemi pravděpodobnostmi soustředěnými v nule.

Definice

Diskrétní náhodná veličina N má degenerované rozdělení s parametrem $\mu \in \mathbb{R}$, píšeme $N \sim Dg(\mu)$, jestliže je pravděpodobnostní funkce tvaru

$$\mathrm{P}(N = k) = \begin{cases} 1 & k = \mu, \mu \in \mathbb{R} \\ 0, & \text{jinak.} \end{cases}$$

.

Rozdělení modifikovaná v nule

Označme

- $G_N(s) = \sum_{k=0}^{\infty} p_N(k) \cdot s^k$ generující funkci rozdělení třídy $(a, b, 0)$,
- $G_N^M(s) = \sum_{k=0}^{\infty} p_N^M(k) \cdot s^k$ generující funkci příslušného v nule modifikovaného rozdělení třídy $(a, b, 1)$.

Platí, že

$$p_N^M(k) = c \cdot p_N(k), \quad \text{pro } k = 1, 2, \ldots; \text{ a } c \in \mathbb{R}^+$$

a $p_N^M(0)$ je libovolně zvolené z intervalu $(0, 1)$. Musíme vypočítat hodnotu c.

Rozdělení modifikovaná v nule

Potom

$$\begin{aligned} G_N^M(s) &= p_N^M(0) + \sum_{k=1}^{\infty} p_N^M(k) \cdot s^k \\ &= p_N^M(0) + c \cdot \sum_{k=1}^{\infty} p_N(k) \cdot s^k \\ &= p_N^M(0) + c(G_N(s) - p_N(0)). \end{aligned}$$

Z platnosti $G_N^M(1) = G_N(1) = 1$ plyne

$$1 = p_N^M(0) + c(1 - p_N(0)).$$

Odtud

$$c = \frac{1 - p_N^M(0)}{1 - p_N(0)}$$

Rozdělení modifikovaná v nule

Tudíž

$$p_N^M(k) = \frac{1 - p_N^M(0)}{1 - p_N(0)} \cdot p_N(k)$$

Generující funkce modifikovaného rozdělení je pak tvaru

$$\begin{aligned} G_N^M(s) &= p_N^M(0) + \frac{1 - p_N^M(0)}{1 - p_N(0)}(G_N(s) - p_N(0)) \\ &= \frac{p_N^M(0) - p_N(0)}{1 - p_N(0)} + \frac{1 - p_N^M(0)}{1 - p_N(0)} G_N(s) \\ &= \frac{p_N^M(0) - 1 + 1 - p_N(0)}{1 - p_N(0)} + \frac{1 - p_N^M(0)}{1 - p_N(0)} G_N(s) \\ &= \left(1 - \frac{1 - p_N^M(0)}{1 - p_N(0)}\right) + \frac{1 - p_N^M(0)}{1 - p_N(0)} G_N(s). \end{aligned}$$

Rozdělení useknutá v nule

- Tento typ rozdělení můžeme chápat jako speciální typ v nule modifikovaného rozdělení se stanovenou hodnotou $p_N^M(0) = 0$.
- Generující funkci v nule useknutého rozdělení označíme $G_N^T(s)$.
- Pak z tvaru $G_N^M(s)$, $p_N^M(k)$, a $p_N^M(0) = 0$ získáme

$$p_N^T(k) = \frac{p_N(k)}{1 - p_N(0)} \quad \text{pro } k = 1, 2, \ldots,$$

$$G_N^T = \frac{G_N(s) - p_N(0)}{1 - p_N(0)}.$$

Rozšířené useknuté negativně binomické (ETNB) rozdělení

- Množina možných hodnot parametru m je rozšířena z $m > 0$ na $m > -1$, přičemž $m \neq 0$.
- Pravděpodobnostní funkce je dána jako

$$p_N^T(k) = \begin{cases} \binom{k+m-1}{k}(1-p)^k, & k = 1, 2, ...;\ p \in (0,1) \\ 0, & \text{jinak,} \end{cases}$$

resp. tvaru

$$p_N^T(k) = \begin{cases} \frac{\binom{k+m-1}{k}\left(\frac{\beta}{1+\beta}\right)^k}{(1+\beta)^m - 1}, & k = 1, 2, ...; \beta = \frac{1-p}{p} > 0 \\ 0, & \text{jinak.} \end{cases}$$

Logaritmické rozdělení

- Toto rozdělení je limitním případem ETNB rozdělení pro $m \to 0$
- Neexistuje k němu odpovídající rozdělení ve třídě $(a, b, 0)$.
- Pravděpodobnostní funkce má tvar

$$p_N^T(k) = \begin{cases} -\frac{(1-p)^k}{k \ln(p)}, & k = 1, 2, ...; \ p \in (0, 1) \\ 0, & \text{jinak,} \end{cases}$$

resp. tvaru

$$p_N^T(k) = \begin{cases} \frac{\beta^k}{k(1+\beta)^k \ln(1+\beta)}, & k = 1, 2, ...; \beta = \frac{1-p}{p} > 0 \\ 0, & \text{jinak.} \end{cases}$$

Příklad

Uvažujme náhodnou veličinu N z negativně binomického rozdělení s parametry $m = 2,5$ a $\beta = 0,5$. Určete koeficienty $a, b \in \mathbb{R}$ a pravděpodobnostní funkci v $k = 0, 1, 2, 3$. Dále určete její v nule useknutou a v nule modifikovanou verzi jestliže $p_N^M(0) = 0,6$.

Příklad

Určete generující funkci, střední hodnotu a rozptyl useknutého geometrického rozdělení.

Obsah

1. Rozdělení počtu pojistných událostí N
2. Rozdělení výše pojistných nároků X

- Zaměříme se na modely výše pojistných nároků, které vycházejí z rozdělení náhodné veličiny X_i, $i = 1, 2, ..., N$.
- Výše škod je vždy nezáporná a většinou kladně zešikmená (tzn. větší četnost škod o malém rozsahu).
- Z těchto důvodů není příliš vhodné využívat normální rozdělení.

Log-normální rozdělení

- Jedná se o rozdělení náhodné veličiny $X = \mathrm{e}^{Y}$, kde $Y \sim N(\mu, \sigma^2)$.
- Nabývá pouze kladných hodnot, nejpravděpodobnější je výskyt středně velkých hodnot.

Definice

Spojitá náhodná veličina X má log-normální rozdělení s parametry $\mu \in \mathbb{R}$ a $\sigma^2 > 0$, píšeme $X \sim LN(\mu, \sigma^2)$, jestliže je její pravděpodobnostní hustota

$$f_X(x) = \begin{cases} \frac{1}{x\sigma\sqrt{2\pi}} \mathrm{e}^{-\frac{(\ln(x)-\mu)^2}{2\sigma^2}}, & x > 0 \\ 0, & \textit{jinak.} \end{cases}$$

Střední hodnota $\mathrm{E}(X) = \mathrm{e}^{\mu + \frac{\sigma^2}{2}}$

Rozptyl: $\mathrm{Var}(X) = \mathrm{e}^{2\mu+\sigma^2}(\mathrm{e}^{\sigma^2} - 1)$.

Exponenciální rozdělení

- Jedná se o rozdělení s jedním parametrem a proto nám umožňuje stanovit výši pojistných nároků nejjednodušším způsobem.
- Není vhodné ho využívat v případech, kdy mohou nastat značně vysoké škody s relativně velkou pravděpodobností.

Definice

Spojitá náhodná veličina X má exponenciální rozdělení s parametrem $\lambda > 0$, píšeme $X \sim Exp(\lambda)$, jestliže je její pravděpodobnostní hustota

$$f_X(x) = \begin{cases} \lambda\,\mathrm{e}^{-\lambda x}, & x > 0 \\ 0, & \textit{jinak.} \end{cases}$$

Střední hodnota $\mathrm{E}(X) = \frac{1}{\lambda}$ **Rozptyl:** $\mathrm{Var}(X) = \frac{1}{\lambda^2}$.

Gamma rozdělení

- Je flexibilnější než exponenciální rozdělení díky dvěma parametrům. Pro $\alpha = 1$ získáme exponenciální rozdělení.
- Nelze využít pro modelování extrémních hodnot.

Definice

Spojitá náhodná veličina X má gamma rozdělení s parametry $\alpha > 0, \beta > 0$, píšeme $X \sim Gam(\alpha, \beta)$, jestliže je její pravděpodobnostní hustota

$$f_X(x) = \begin{cases} \frac{x^{\alpha-1} e^{-\frac{x}{\beta}}}{\Gamma(\alpha)\beta^\alpha}, & x > 0 \\ 0, & \textit{jinak}, \end{cases}$$

kde $\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} e^{-x} \, \mathrm{d}x$ je gamma funkce.

Střední hodnota $\mathrm{E}(X) = \alpha\beta$ **Rozptyl:** $\mathrm{Var}(X) = \alpha\beta^2$.

Příklad

Na základě dat byla odhadnuta střední výše škod na 1 537 Kč a rozptyl na 381 764 ze 100 pojistných nároků. Odhadněte rozdělení výše škod log-normálním rozdělením. Pomocí tohoto odhadu vypočtěte pravděpodobnost, že výše škody nepřesáhne 4 800 Kč.