

Model konkurujících si rizik. Sledování může být ukončeno z k příčin, které nastávají v časech

$$X_1, \dots, X_k,$$

což jsou nezávislé n.v. pozorovatelné pouze hypoteticky. Známe jen dobu sledování a příčinu ukončení

$$W = \min X_j, \quad (I_1, \dots, I_k), \quad \text{kde } I_j = \chi_{[W=X_j]}.$$

Jsou-li rozdělení X_j spojitá, můžeme je charakterizovat intenzitami

$$\lambda_j(x) = \lim_{h \rightarrow 0} P[X_j \in (x, x+h) \mid X_j > x]/h$$

tj. $\lambda_j = f_j/(1 - F_j)$, $F_j(t) = 1 - \exp(-\int_0^t \lambda_j)$, kde F_j , f_j jsou distribuční funkce a hustota X_j .

Je $P[I_j = 1 \mid W] = \lambda_j(W) / \sum_i \lambda_i(W)$.

Příklady. Zaměstnanci je při ukončení zaměstnání přiznána částka závisající na době zaměstnání a příčině odchodu (výpověď, vznik invalidity, úmrtí, odchod do důchodu).

Předčasné ukončení splátek pojistného může znamenat různé pojistné plnění v závislosti na příčině ukončení (smrt, rozhodnutí pojištěného).

V pojištění invalidity je vyplácena různá částka podle způsobu vzniku invalidity (nemocí, úrazem).

Ve zdravotním pojištění se zajímáme nejen o analýzu úmrtnosti, ale také o příčiny smrti (kardiovaskulární choroby, rakovina, úraz, jiné).

Model. Mějme n . výběr k -tic nezávislých $\text{Exp}(\lambda_1), \dots, \text{Exp}(\lambda_k)$ rozdělených n.v. rozsahu n , resp. z každé k -tice je pozorováno jen její minimum, včetně jeho identifikace,

$$W^i, \quad (I_1^i, \dots, I_k^i), \quad i = 1, \dots, n,$$

což vede k věrohodnostní funkci

$$L = e^{-\sum \lambda_j W} \prod_j \lambda_j^{I_j}, \quad \text{kde } W = \sum_i W^i, \quad I_j = \sum_i I_j^i$$

pro $\lambda_j, \dots, \lambda_k$. Další potřebné parametrické funkce jsou

$$\theta_j = 1/\lambda_j, \quad R_j(t) = e^{-\lambda_j t}, \quad \lambda = \sum \lambda_j, \quad p_j = \lambda_j/\lambda,$$

případně $\gamma_j = (\lambda - \lambda_j)/\lambda_j = 1/p_j - 1$.

Odhady. Přirozený konjugovaný systém je tvořen rozděleními, kde $\lambda_1, \dots, \lambda_k$ jsou nezávislé, $\lambda_j \sim G(a, q_j)$. Pak také $\lambda \sim G(a, q)$, kde $q = \sum q_j$, a $(p_1, \dots, p_k) \sim D(q_1, \dots, q_k)$ jsou nezávislé. Při takové apriorní hustotě a při kvadratické ztrátové funkci dostáváme bayesovské odhady

$$\begin{aligned} \hat{\lambda}_j &= \frac{I_j + q_j}{W + a}, \quad \hat{\theta}_j = \frac{W + a}{I_j + q_j - 1}, \quad \widehat{R_j(t)} = \left(1 + \frac{t}{W+a}\right)^{-(I_j+q_j)}, \\ \hat{\lambda} &= \frac{n + q}{W + a}, \quad \hat{p}_j = \frac{I_j + q_j}{n + q}, \quad \hat{\gamma}_j = \frac{n + q_{xj} - I_j}{I_j + q_j - 1}. \end{aligned}$$

Vezmeme-li apriorní rozdělení s tímž rozdělením (a pak i odhady) pro $p_1, \dots, p_k$, nezávislým s $\lambda \sim G(a, q + r)$, nebudou už $\lambda_1, \dots, \lambda_k$ nezávislé. Máme $(\lambda_j | p_j) \sim G(a_j/p_j, q + r)$ a odhady

$$\hat{\lambda} = \frac{n + q + r}{W + a}, \quad \hat{\lambda}_j = \hat{\lambda} \hat{p}_j, \quad \text{a} \quad \hat{\theta}_j = \widehat{(1/\lambda)} \widehat{(1/p_j)}.$$

Hodnoty $a = 0, q_j = 1/2, r = -k/2$ odpovídají Jeffreysově apriorní hustotě.

Jiné zobecnění předpokládá nezávislá $\lambda_j \sim G(a_j, q_j)$, $j = 1, \dots, k$, pak $(\lambda \mid p_1, \dots, p_k) \sim G(\sum a_j p_j, q)$ a $(p_1, \dots, p_k) \sim (\prod p_j^{q_j-1}) / (\sum a_j p_j)^q$. Máme

$$\hat{\lambda}_j = \frac{I_j + q_j}{W + a_j}, \quad \hat{\theta}_j = \frac{W + a_j}{I_j + q_j - 1}, \quad \widehat{R_j(t)} = \left(1 + \frac{t}{W+a_j}\right)^{-(I_j+q_j)}.$$

$\hat{\lambda} = \sum \hat{\lambda}_j$ a $\hat{\gamma}_j = \hat{\theta}_j \sum_{\ell \neq j} \hat{\lambda}_\ell$. Odhad p_j lze pro $k = 2$ vyjádřit pomocí hypergeometrické funkce jako

$$\hat{p}_1 = \frac{n+q}{I_1+q_1} {}_2F_1(1, I_2+q_2, n+q+1, 1 - (W+a_1)/(W+a_2)),$$

kde ${}_2F_1(a, b, c, z) = \sum_{k=0}^{\infty} \frac{(a)_k (b)_k}{(c)_k} \frac{z^k}{k!}$, $(x)_k = x(x+1) \dots (x+k-1)$.

Vezmeme-li v úvahu všechna rozdělení uvedená výše, při

$$\pi_{a_1, \dots, a_k, q_1, \dots, q_k, r} \propto \prod \lambda_j^{q_j-1} (\sum \lambda_j)^r \exp(-\sum a_j \lambda_j)$$

docházíme při odhadování k integrálům

$$I_{q_1, \dots, q_k, m} = \int \dots \int_{\substack{p_1, \dots, p_k \geq 0 \\ \sum p_j = 1}} \frac{\prod p_j^{q_j-1}}{(\sum a_j p_j)^{q+m}}.$$

Speciálně pro $m \in \mathbf{N}$ je můžeme explicitně vyjádřit, nebo pro $k = 2$ docházíme k ${}_2F_1$.

Konvergence. ZVČ a CLV pro W a I při $n \rightarrow \infty$ poskytují konvergenci skoro jistě a v distribuci, např.

$$\begin{aligned} \sqrt{n}(\hat{\lambda}_j - \lambda_j) &= \sqrt{n} \left(\frac{I_j + q_j}{W + a} - \lambda_j \right) \\ &\sim \frac{1}{W/n} \left[\sqrt{n} \left(\frac{I_j}{n} - p_j \right) - \lambda_j \sqrt{n} \left(\frac{W}{n} - \frac{1}{\lambda} \right) \right] \sim \lambda N(0, p_j). \end{aligned}$$

Dokonce pro $\widehat{R}_j(t) = (\frac{W+a_j}{W+a_j+t})^{I_j+q_j}$ uvažovaný jako náhodný proces máme

$$\sqrt{n}(\widehat{R}_j - R_j) \xrightarrow[n \rightarrow \infty]{\mathcal{D}} Z \quad \text{v } C([0, \infty]),$$

kde Z je Gaussovský s $\text{cov}(Z(s), Z(t)) = s t e^{-\lambda_j(s+t)} \lambda_j^2 / p_j$, a to porovnáním s $e^{-\lambda_j t} \sqrt{n}(\frac{I_1+q_j}{W+a_j} t - \lambda_j t)$.

Bayesovská rizika. V případě přirozeného konjugovaného apriorního rozdělení, kdy $W \sim a \text{B2}(q, n)$, např.

$$\begin{aligned} \text{BR } \widehat{\lambda}_j &= \text{E} \frac{I_j + q_j}{(W + a)^2} = (n \frac{q_j}{q} + q_j) \frac{q(q+1)}{a^2(n+q)(n+q+1)}, \\ \text{BR } \widehat{\theta}_j &= \text{E} \frac{(W + a)^2}{(I_j + q_j - 1)^2(I_j + q_j - 2)} \\ &= \frac{1}{n} \frac{a^2}{(q-1)(q-2)} \text{E} p_j^{-3} + O(1/n^2), \end{aligned}$$

nebo po vícenásobném odhadování

$$\begin{aligned} \text{BR } \widehat{R}_j(t) &= \text{E} \left[\left(1 + \frac{2t}{W+a}\right)^{-(I_j+q_j)} - \left(1 + \frac{t}{W+a}\right)^{-2(I_j+q_j)} \right] \\ &= \text{E} \left(t^2 e^{-2t(I_j+q_j)/(W+a)} \frac{I_j + q_j}{(W+a)^2} \right) + O(1/n^2) \\ &= \frac{1}{n} t^2 \text{E} e^{-2t\lambda_j} p_j \lambda^2 + O(1/n^2). \end{aligned}$$

Použijeme-li místo bayesovského $\widehat{\lambda}_j$ jiný odhad $\widetilde{\lambda}_j$, je

$$\text{BR } \widetilde{\lambda}_j = \text{E}(\lambda_j - \widetilde{\lambda}_j)^2 = \text{BR } \widehat{\lambda}_j + \text{E}(\widehat{\lambda}_j - \widetilde{\lambda}_j)^2.$$

Dají-li se rizika rozvinout ve tvaru

$$\text{BR } \widehat{\lambda}_j = \frac{c}{n} + \frac{a}{n^2} + o(\frac{1}{n^2}), \quad \text{BR } \widetilde{\lambda}_j = \frac{c}{n} + \frac{b}{n^2} + o(\frac{1}{n^2}),$$

definujeme asymptotickou deficienci $\tilde{\lambda}_j$ vzhledem k $\hat{\lambda}_j$ jako $(b - a)/c$. Např. pro $\tilde{\lambda} = (I_j + \tilde{q}_j)/(W + \tilde{a})$

$$\begin{aligned} \mathbb{E}(\hat{\lambda} - \tilde{\lambda})^2 &= \mathbb{E} \left(\frac{W(q_j - \tilde{q}_j) - I_j(a - \tilde{a}) + q_j\tilde{a} - \tilde{q}_ja}{(W + a)(W + \tilde{a})} \right)^2 \\ &= \frac{1}{n^2} \mathbb{E} [(q_j - \tilde{q}_j)^2 \lambda^2 - 2(q_j - \tilde{q}_j)(a - \tilde{a})p_j \lambda^3 + (a - \tilde{a})^2 p_j^2 \lambda^4] \\ &\quad + O(1/n^3), \end{aligned}$$

což vůči $\hat{\lambda}_j$ znamená deficienci

$$\begin{aligned} \frac{q}{q_j}(q_j - \tilde{q}_j)^2 - 2\frac{q+2}{a}(q_j - \tilde{q}_j)(a - \tilde{a}) \\ + \frac{1}{a^2}\frac{q_j}{q+1}(q+2)(q+3)(a - \tilde{a})^2. \end{aligned}$$

Obecněji. Pro intenzity tvaru $\lambda_j(t) = \theta_j \mu_j(t)$, kde $\mu_j(t)$ je známá, je situace podobná, např. při nezávislých $G(a_j, q_j)$ je

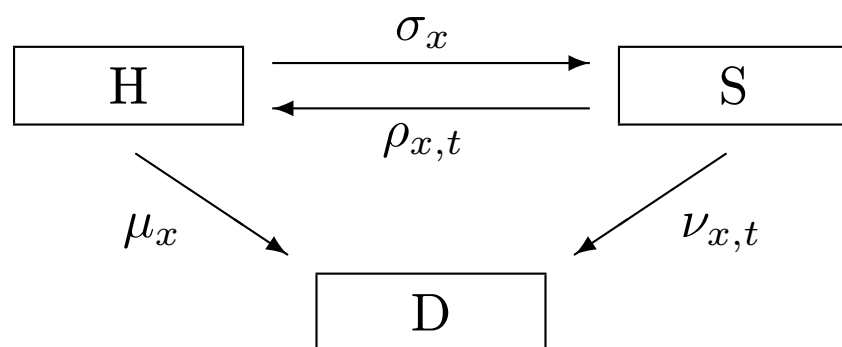
$$\hat{\theta}_j = (I_j + q_j)/(W_j + a_j), \quad \text{kde } W_j = \sum_{i=1}^n \int_0^{W^i} \mu_j(t) dt.$$

Vícestavový model. Rozlišujeme stavy, ve kterých se osoba může nacházet a sledujeme změny stavů v čase. Chování je popsáno intenzitami přechodu

$$\lambda_{ij}(x, t) = \lim_{h \rightarrow 0} P[\text{stav } j \text{ v čase } x + h \mid \text{stav } i \text{ v čase } x \text{ trvá } t] / h$$

Pro $\lambda_{ij}(x, t) = \lambda_{ij}(x)$ se jedná o markovský proces.

Příklady. V nemocenském pojištění může pobývat osoba ve 3 stavech: zdravý, nemocný, mrtvý.



Přitom intenzity výstupu ze stavu nemocných budou záviset nejen na věku, ale také na době od začátku nemoci.

V pojištění invalidity je rozlišena invalidita v důsledku nemoci a úrazu, podle charakteru invalidity může být předpovězena trvalá či dočasná invalidita, či rozlišeno více stupňů invalidity.

Model s konstantními intenzitami. Pro konstantní intenzity přechodu λ_{ij} mezi m přechodnými a k absorpčními stavy vede sledování n jedinců k věrohodnostní funkci

$$\left(\prod_{i=1}^m p_{0i}^{I_{0i}} \right) \prod_{i=1}^m \left(e^{-\lambda_i W_i} \prod_{j \neq i} \lambda_{ij}^{I_{ij}} \right), \quad \lambda_i = \sum_{j \neq i} \lambda_{ij}.$$

Výstupy ze stavu probíhají jako v modelu konkurujících rizik, ovšem počet výstupů ze stavu i je náhodný

$$N_i = \sum_{j \neq i} I_{ij}, \quad N_i / n \rightarrow p_{0i}^* / (1 - p_{ii}^*) \text{ s. j.}$$

Předpokládáme-li obdobná apriorní rozdělení, např. nezávislá $\lambda_{ij} \sim G(a_i, q_{ij})$, ihned

$$\hat{\lambda}_{ij} = \frac{I_{ij} + q_{ij}}{W_i + a_i}, \quad \hat{\theta}_{ij} = \frac{W_i + a_i}{I_{ij} + q_{ij} - 1}, \quad \hat{\gamma}_{ij} = \frac{N_i + q_{i \times j} - I_{ij}}{I_{ij} + q_{ij} - 1},$$

$$\hat{\lambda}_i = \frac{N_i + q_i}{W_i + a_i}, \quad \hat{p}_{ij} = \frac{I_{ij} + q_{ij}}{N_i + q_i}, \quad \widehat{R_{ij}(t)} = \left(1 + \frac{t}{W_i + a_i}\right)^{-(I_{ij} + q_{ij})},$$

přičemž konvergence s. j. i v distribuci zůstává zachována. Pro $m = 2$ dostaneme bayesovská rizika

$$\text{BR } \hat{\lambda}_{1j} = \frac{1}{n}(ab + cd) + O\left(\frac{1}{n^2}\right), \quad \text{BR } \hat{p}_{1j} = \frac{1}{n}(ae + cf) + O\left(\frac{1}{n^2}\right),$$

kde b, d, e, f jsou zlomky, např. $b = \frac{q_{1j}q_{1 \times j}}{q_1(q_1+1)}$, a pro a a c máme $1 \leq a \leq \min(E \frac{1}{p_{01}}, \frac{q_2-1}{q_{21}-1})$ a $\frac{q_{21}}{q_2} \leq c \leq \min(1, \frac{q_{21}}{q_2} E \frac{1}{p_{01}})$. Pro $p_{01} = 1$ je $a = 1, c = q_{21}/q_2$.

Pro intenzity tvaru $\lambda_{ij}(x, t) = \theta_{ij}\mu_{ij}(x, t)$, kde $\mu_{ij}(x, t)$ jsou známé, s apriorními nezávislými $\theta_{ij} \sim G(a_{ij}, q_{ij})$ je

$$\hat{\theta}_{ij} = (I_{ij} + q_{ij})/(W_{ij} + a_{ij}) \quad \text{s } W_{ij} = \sum_{\ell=1}^{N_i} \int_0^{W_i^\ell} \mu_{ij}(T_i^\ell, t) dt.$$

Příklad. V případě ukončení pojištění před časem t z příčiny 1 dojde k výplatě jednotkové částky. Pojistné dle principu střední hodnoty je

$$P = (1 + \alpha)(1 - e^{-\lambda t})p_1.$$

Předpokládáme-li nezávislá $G(a, q_j)$ rozdělení pro intenzity λ_j , pak bayesovský odhad

$$\begin{aligned}\hat{P} &= (1 + \alpha)(1 - E(e^{-\lambda t} | W)) E(p_1 | I_1) \\ &= (1 + \alpha) \left[1 - \left(\frac{W + a}{W + a + t} \right)^{q+n} \right] \frac{I_1 + q_1}{n + q}\end{aligned}$$

má bayesovské riziko

$$\begin{aligned}\frac{\text{BR } \hat{P}}{(1 + \alpha)^2} &= E \frac{(I_1 + q_1)(I_1 + q_1 + 1)}{(n + q)(n + q + 1)} \left[1 - 2 \left(1 + \frac{t}{W + a} \right)^{n+q} \right. \\ &\quad \left. + \left(1 + \frac{2t}{W + a} \right)^{n+q} \right] - \frac{(I_1 + q_1)^2}{(n + q)^2} \left[1 - \left(1 + \frac{t}{W + a} \right)^{n+q} \right]^2 \\ &= E e^{-2t(n+q)/(W+a)} \frac{2t^2(I_1 + q_1)^2}{(W + a)^2(n + q + 1)} + O(1/n^2) \\ &= \frac{t^2}{n} E e^{-2\lambda t} \lambda^2 p_1^2 + O(1/n^2) = \frac{t^2}{n} \frac{a^q q_1 (q_1 + 1)}{(a + 2t)^{q+2}} + O(1/n^2).\end{aligned}$$

Příklad. $W^1, W^2, \dots \sim \text{Exp}(\lambda)$ doby mezi škodami nezávislé s velikostmi škod $Y^1, Y^2, \dots$, přičemž $P[Y^i = c_j] = p_j$, $j = 1, \dots, k$.

Pozorováním za dobu t zjistíme počty I_j škod ve výši c_j , věrohodnostní funkce pro $\lambda, p_1, \dots, p_k$ je

$$L = e^{-\lambda t} \prod \lambda_j^{I_j} \quad (\lambda_j = \lambda p_j).$$

Celkový počet škod za dobu t je $N = \sum_j I_j \sim \text{Po}(\lambda t)$ a celková škoda $S = \sum_{i=1}^N Y^i$ má

$$E S = \lambda t \sum c_j p_j = t \sum c_j \lambda_j, \quad \text{var } S = \lambda t \sum c_j^2 p_j = t \sum c_j^2 \lambda_j.$$

Veličiny $X_1^i \sim \text{Exp}(\lambda_1), \dots, X_k^i \sim \text{Exp}(\lambda_k)$ z modelu konkurujících si rizik představují i -tou dobu čekání (s možnou škodou $c_1, \dots, c_k$), celková doba pozorování je $W = t$.

Za předpokladu nezávislých $G(a_j, q_j)$ apriorních rozdělení pro λ_j pak např. pro pojistné na jednotkové období dle principu rozptylu,

$$P = \sum c_j \lambda_j + \alpha \sum c_j^2 \lambda_j = \sum (c_j + \alpha c_j^2) \lambda_j,$$

máme bayesovský odhad

$$\hat{P} = \sum (c_j + \alpha c_j^2) \hat{\lambda}_j = \sum (c_j + \alpha c_j^2) \frac{I_j + q_j}{t + a_j}$$

s bayesovským rizikem ($d_j = c_j + \alpha c_j^2$)

$$\begin{aligned} \text{BR } \hat{P} &= E \sum d_j^2 \text{var}(\lambda_j \mid I_1, \dots, I_k) = E \sum d_j^2 \frac{I_j + q_j}{(t + a_j)^2} \\ &= \sum d_j^2 \frac{E N p_j + q_j}{(t + a_j)^2} = \sum d_j^2 \frac{t q_j / a_j + q_j}{(t + a_j)^2} = \sum d_j^2 \frac{q_j}{a_j (t + a_j)}. \end{aligned}$$