

# BAYESOVSKÉ ODHADY V EXPONENCIÁLNÍM MODELU KONKURUJÍCÍCH SI RIZIK

MICHAL FRIESL

*Katedra matematiky, Fakulta aplikovaných věd,  
Západočeská univerzita  
<http://home.zcu.cz/~friesl>*

**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

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

Př.: spolehlivost součástek v zařízení, doba života a příčiny smrti, při  $k = 2$  náhodné cenzorování.

Jsou-li rozdělení  $X_j$  spojitá, můžeme je charakterizovat intenzitami

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

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

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

**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  $\lambda_j, \dots, \lambda_k$ . Může nás zajímat také

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

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

Při společném zobecnění s hustotou

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

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  $\widehat{R}_j(t) = (\frac{W+a_j}{W+a_j+t})^{I_j+q_j}$  jako 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

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

Mají-li rizika rozvoje

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

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

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

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

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