MATEMATICKÉ UVAŽOVÁNÍ V ŘÍZENÍ KREDITNÍCH RIZIK

Petr Veselý

Raiffeisenbank, a.s.

Přírodovědecká fakulta MU
Ústav matematiky a statistiky
31. 10. 2012

Obsah

I. Vybrané aspekty úvěrování

II. Ilustrativní příklady matematických přístupů

- Skóring a rating
- Kreditní rizikové náklady
- Optimální hladina zamítání
- PIT a TTC odhady pravděpodobností defaultu

Vybrané aspekty úvěrování

Vybrané aspekty úvěrování 1/3

Schvalování:

- komu ano / komu ne? („risk appetite“ & profitabilita)
- skóring a rating (kvantifikace podstupované míry rizika)
- jaký produkt, v jaké výši, s jakým zajištěním, na jak dlouho, za kolik?
- restrukturalizace a konsolidace existujících kontraktů

Data:

- rozsah užívaných dat:
  - demografická data o žadatelích a klientech
  - údaje o obchodech (typ produktu, schvalovací kanál, parametry, zajištění, vazby na „předchůdce“ a následníky“)
  - časové řady vývoje zůstatků a platební morálky
  - klientská transakční data
  - externí data
- datová úložiště a jejich správa, dokumentace a rozvoj
- data quality management
- data mining

Vybrané aspekty úvěrování 2/3

Portfolio management:

- profitabilita a náklady na kapitál
- očekávané a neočekávané riziko
- makroekonomické vlivy
- predikce budoucího vývoje
- limit management (koncentrace, objemové limity)
- stress testing

Monitoring a reporting:

- sledované metriky
- systém včasného varování (early warning system)

Vymáhání problematických pohledávek:

- vymáhací strategie
- odprodej pohledávek / zajištění
- výpočet recovery

Vybrané aspekty úvěrování 3/3

Marketing a relationship management:

- získání a udržení klientů
- komunikace a distribuce
- design a vývoj produktové nabídky

Organizace:

- organizační a řídící struktury, vnitrobankovní komunikace
- motivace jednotlivých útvarů (sladění rozporných zájmů)

Regulatorní aktivity:

- Basel II (RWA, PD, LGD, CF, EL, ICAAP)
- účtování a opravné položky
- dokumentace modelů a procesů
- Compliance

Ilustrace 1:
Skóring a rating

Skóringové funkce – obecný pohled 1/3

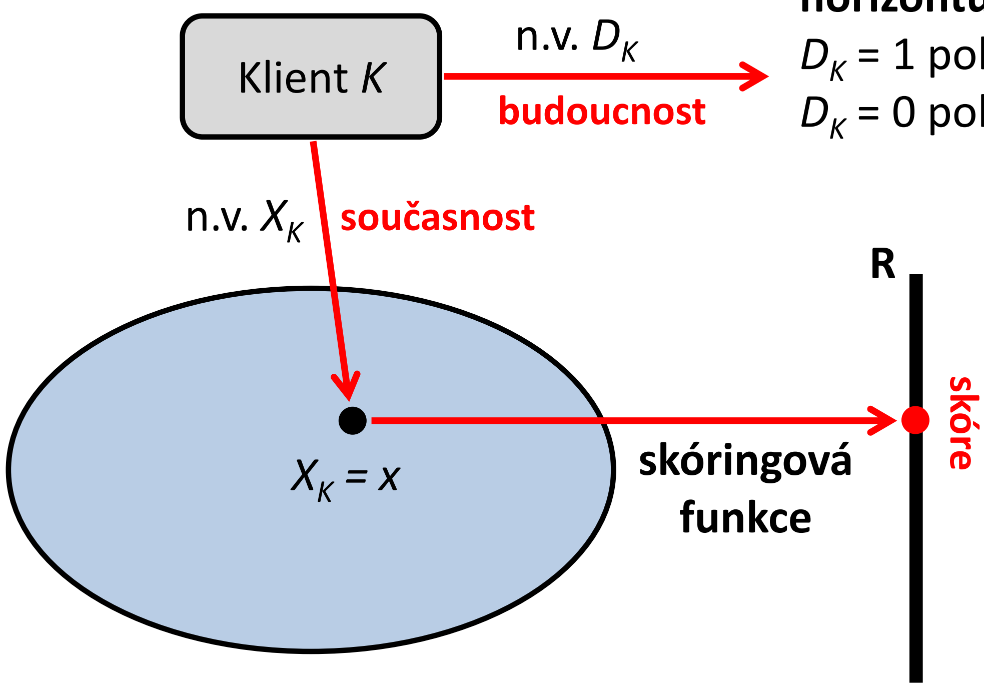
Problém:

Jak na základě známých informací o (potenciálním) dlužníkovi odhadnout, zda dostojí svým závazkům vůči bance?

budoucí chování klienta *K* v daném časovém horizontu: 0-1 náhodná veličina D_K

$D_K = 1$ pokud *K* dostojí závazkům

$D_K = 0$ pokud *K* nedostojí závazkům **(default)**



stavový informační prostor ($\mathbf{V},\mathfrak{F}$)

V – množina možných stavů

$\mathfrak{F}$ – příslušná σ-algebra

Cílem je zkonstruovat takovou skóringovou funkci, která umožní pomocí skóre co nejlépe „diskriminovat“ budoucí špatné klienty v tom smyslu, že dobrým klientům přiřazuje převážně vysoká skóre a špatným klientům převážně nízká skóre.

Skóringové funkce – obecný pohled 2/3

Základní typy vyhodnocovaných informací – příklady dimenzí prostoru ($\mathbf{V}$,$\mathfrak{F}$)**:**

1. **aplikační data:**
   - fyzické osoby: demografické údaje a údaje typu požadovaný produkt, prodejní kanál apod.)
   - podnikatelské subjekty: účetní výkazy (popř. daňové přiznání), odvětví apod.
2. **behaviorální data:** charakteristiky založené na časových řadách platební morálky, vývoje zůstatků na účtech apod.
3. **transakční data:** charakteristiky založené na transakční historii (obraty na účtech a termínovaných vkladech a jejich variabilita, transakce kreditními kartami apod.)

Informační zdroje:

- údaje poskytnuté klientem
- interní databáze
- „credit bureaus“ (Czech Credit Bureau, Solus, Centrální registr úvěrů ČNB)
- jiné veřejně dostupné zdroje (obchodní rejstřík, insolvenční rejstřík apod.)

Časový horizont pro náhodnou veličinu D_K:

Obvykle 12 měsíců.

Skóringové funkce – obecný pohled 3/3

Přirozená volba:

$$s(x) = \mathrm{P}[\, D_K = 1 \mid X_K = x\,] = \mathrm{E}[\, D_K \mid X_K = x\,]\,, \quad x \in \mathbf{V}.$$

Vlastnosti:

1. pravděpodobnost neselhání klienta je rovna jeho skóre: $\mathrm{P}[\, D_K = 1 \mid s(X_K) = a\,] = a$
2. minimalizace střední čtvercové odchylky $\mathrm{E}\big(D_K - s(X_K)\big)^2$
3. funkce minimalizující váženou střední čtvercovou odchylku $\mathrm{E}\big(W_K(D_K - s(X_K))\big)^2$, kde $W_K = w_0$ pro $D_K = 0$ a $W_K = w_1$ pro $D_K = 1$, je monotónní transformací funkce s
4. funkce minimalizující střední $\mathrm{L_p}$-normu $\mathrm{E}|D_K - s(X_K)|^{\mathrm{p}}$ je pro každé $\mathrm{p} \geq 1$ monotónní transformací funkce s
5. nejlepší možné rozlišení dobrých a špatných klientů:

   Neexistuje skóringová funkce r různá od funkce s nebo její monotónní transformace taková, že by pro nějaké $a \in \mathbf{R}$ platilo

   $$\mathrm{P}[\, r(X_K) \leq a \mid D_K = 0\,] \geq \mathrm{P}[\, s(X_K) \leq a \mid D_K = 0\,]$$

   a

   $$\mathrm{P}[\, r(X_K) > a \mid D_K = 1\,] \geq \mathrm{P}[\, s(X_K) > a \mid D_K = 1\,]$$

   a přitom některá z těchto nerovností byla ostrá.

Skóringové funkce – statistické odhady

Jak skóringovou funkci zkonstruovat na základě reálných dat?

LOGISTICKÁ REGRESE

Další metody (rozhodovací stromy, neuronové sítě, …) nejsou příliš časté.

Skóringové funkce – kvantifikace výkonnosti 1/3

Giniho koeficient:

Pro libovolnou skóringovou funkci s je definován jako

$$\mathrm{GC}(s) = 2 \int |F_0(a) - \mathrm{F}_1(a)| \, \mathrm{d}F_0(a) = 2 \int |F_0(a) - F_1(a)| \, \mathrm{d}F_1(a) ,$$

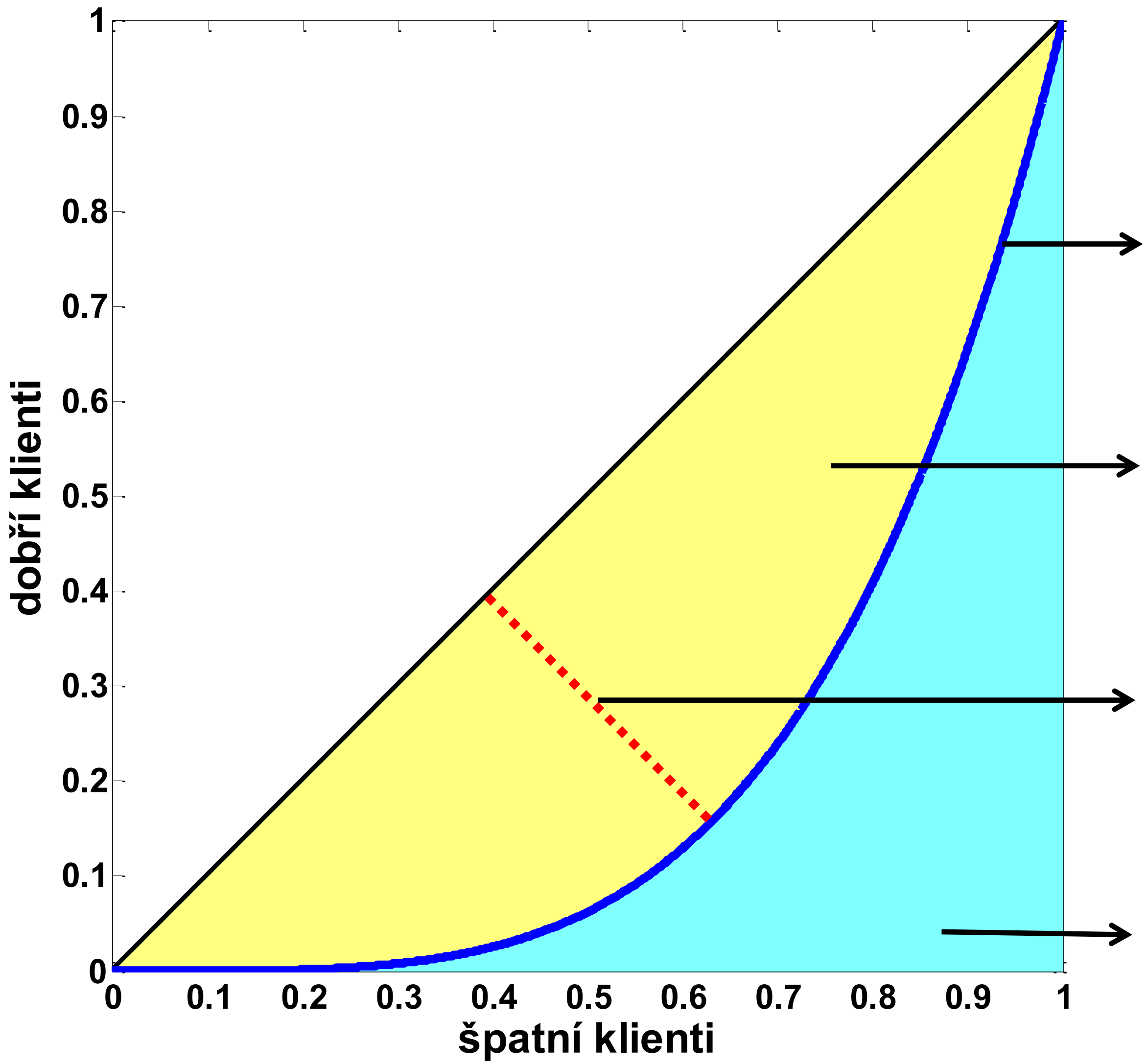
kde F_0 a F_1 jsou distribuční funkce rozdělení skóre špatných a dobrých klientů, tj.

$$F_0(a) = \mathrm{P}[\, s(X_K) \leq a \mid D_K = 0 \,] \text{ a } F_1(a) = \mathrm{P}[\, s(X_K) \leq a \mid D_K = 1 \,] , a \in \mathbf{R}.$$

Vlastnosti Giniho koeficientu:

- Nabývá hodnot z intervalu [0,1], kde mezní hodnoty 0 a 1 odpovídají náhodnému skóre (nulová diskriminační výkonnost) a dokonalému rozlišení dobrých a špatných klientů.
- Maximální hodnoty nabývá pro podmíněnou pravděpodobnost (viz slide č. 10).
- Geometrická interpretace: Je roven dvojnásobku plochy mezi diagonálou a ROC křivkou, tj. křivkou danou parametricky předpisem $\{ [F_0(a), F_1(a)], a \in \mathbf{R} \}$.
- Pravděpodobnostní interpretace: Pokud $F_0(a) \geq F_1(a)$ pro všechna $a \in \mathbf{R}$ (přirozený požadavek kladený na skóringovou funkci), pak je hodnota $(1 + \mathrm{GC}(s))/2$ rovna pravděpodobnosti toho, že náhodně vybraný dobrý klient má větší skóre než náhodně vybraný špatný klient.
- Je invariantní vůči monotónním transformacím skóre.
- Je invariantní vůči hodnotám četností dobrých a špatných klientů.

Skóringové funkce – kvantifikace výkonnosti 2/3



ROC křivka

1/2 Giniho koeficientu

√2/2 × Kolmogorovova-Smirnovova statistika

P[dobrý klient má menší skóre než špatný klient]

Skóringové funkce – kvantifikace výkonnosti 3/3

Kolmogorovova-Smirnovova statistika:

$$\mathrm{KS}(s) = \sup_a |F_0(a) - F_1(a)|$$

Nižší citlivost proti Giniho koeficientu (změna ROC křivky ovlivňující diskriminační schopnost se nemusí promítnout v hodnotě KS statistiky).

Information value:

$$\mathrm{IV}(s) = \textstyle\sum_i (b_i - g_i) \log(b_i / g_i) ,$$

kde b_i a g_i jsou relativní četnosti špatných a dobrých klientů pro jednotlivé kategorie hodnot skóre. Menší robustnost (citlivost na počty pozorování pro jednotlivé kategorie hodnot), problém při porovnávání dvou skóringových funkcí lišících se počtem shod.

Kendallův korelační koeficient τ: závisí na relativních četnostech dobrých a špatných klientů ve vzorku (riziko chybného porovnání výkonnosti skóringové funkce na dvou časových oknech)

Kruskallovo-Wallisovo γ: riziko chybného porovnání dvou skóringových funkcí lišících se počtem shod

Poznámka:
Histogram rozdělení dobrých a špatných klientů (popř. odhad hustot) pro posouzení diskriminační výkonnosti nestačí (obtížené porovnání dvou funkcí mezi sebou).

Skóringové funkce – nástrahy a problémy 1/4

Teoreticky nejlepší skóringová funkce je známá,

existují metody jejího odhadu,

existují míry její kvality,

takže problematika skóringu je z matematického hlediska jen otázkou vhodného výběru prediktorů a co nejlepších statistických odhadů?

Reálný svět je složitější…

Skóringové funkce – nástrahy a problémy 2/4

Příklad 1:

datový vzorek

⬇

vývoj scoringové funkce

⬇

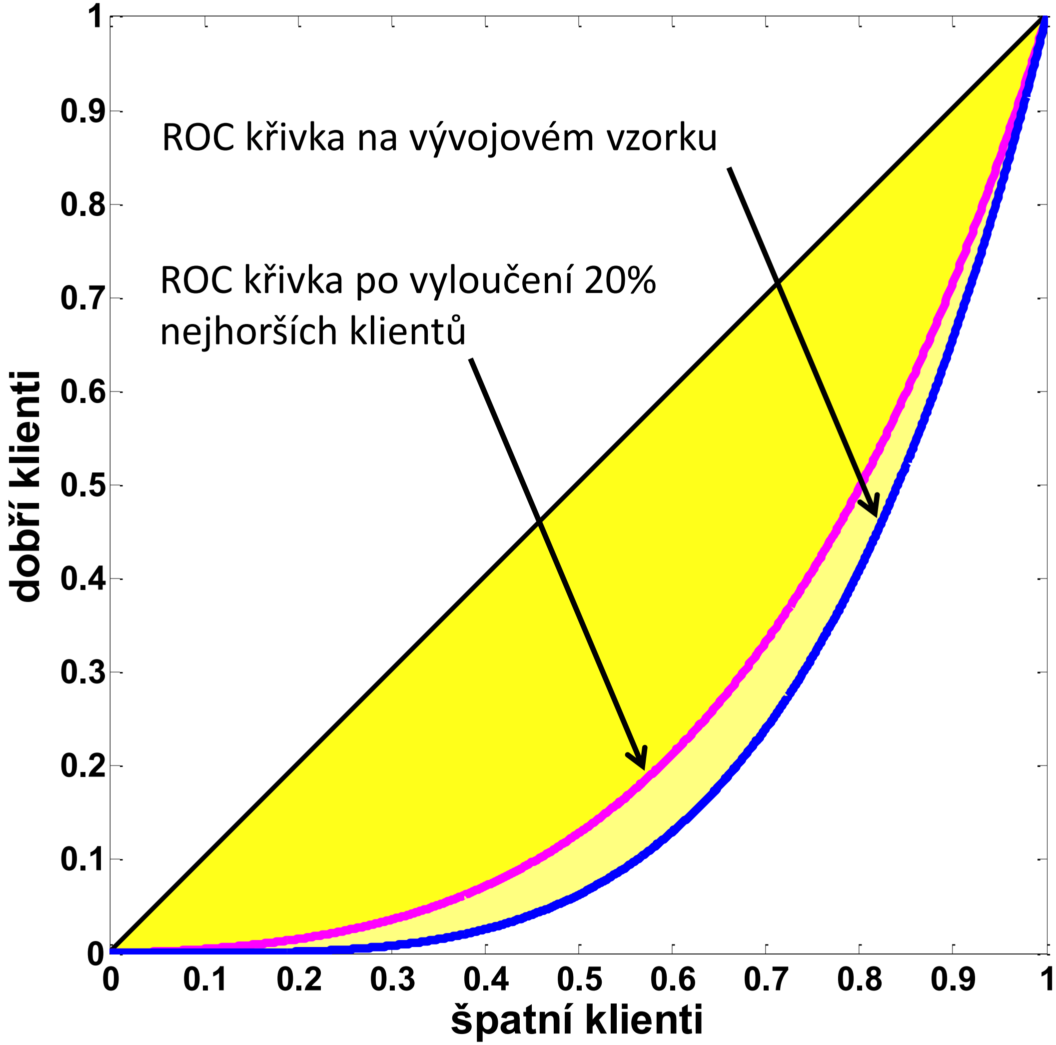
nasazení do praxe

⬇

zamítání žádostí se skóre menším než zvolená hodnota (tzv. cut off)

⬇

zdánlivý pokles výkonnosti scoringové funkce + ochuzení datového vzorku pro budoucí vývoj modelů



Skóringové funkce – nástrahy a problémy 3/4

Příklad 2:

Skóringová funkce je součástí komplexního procesu poskytování obchodů. Důsledky:

1. Monitoring výkonnosti i příprava datových vývojových vzorků musí brát v úvahu metodické změny v obchodních a schvalovacích procesech (tzv. credit policy).
2. Není jednoduché odlišit výkonnost skóringové funkce jako takové od výkonnosti celého schvalovacího procesu.
3. Schvalovací proces „cenzoruje" data pro budoucí update skóringové funkce. Čím dokonalejší schvalovací proces, tím složitější vývoj nové skóringové funkce.

Kroky prováděné ve schvalovacím procesu bývají závislé na skóringovém/ratingovém hodnocení (míra automatizace rozhodnutí o poskytnutí obchodu a jeho parametrech, stupeň verifikace údajů poskytnutých žadatelem apod.), což dále komplikuje dopady uvedené výše.

Příklad 3:

Výkonnost skóringového modelu je korelovaná s celkovou hospodářskou situací. Vývoj skóringové funkce může probíhat na datech z období jiné makroekonomické situace, než je ta, v níž je model aplikován v praxi.

Skóringové funkce – nástrahy a problémy 4/4

Příklad 4:

Nejen banka si vybírá, komu půjčí, ale i klienti si vybírají, za jakých podmínek, tj. od které banky, si půjčí.

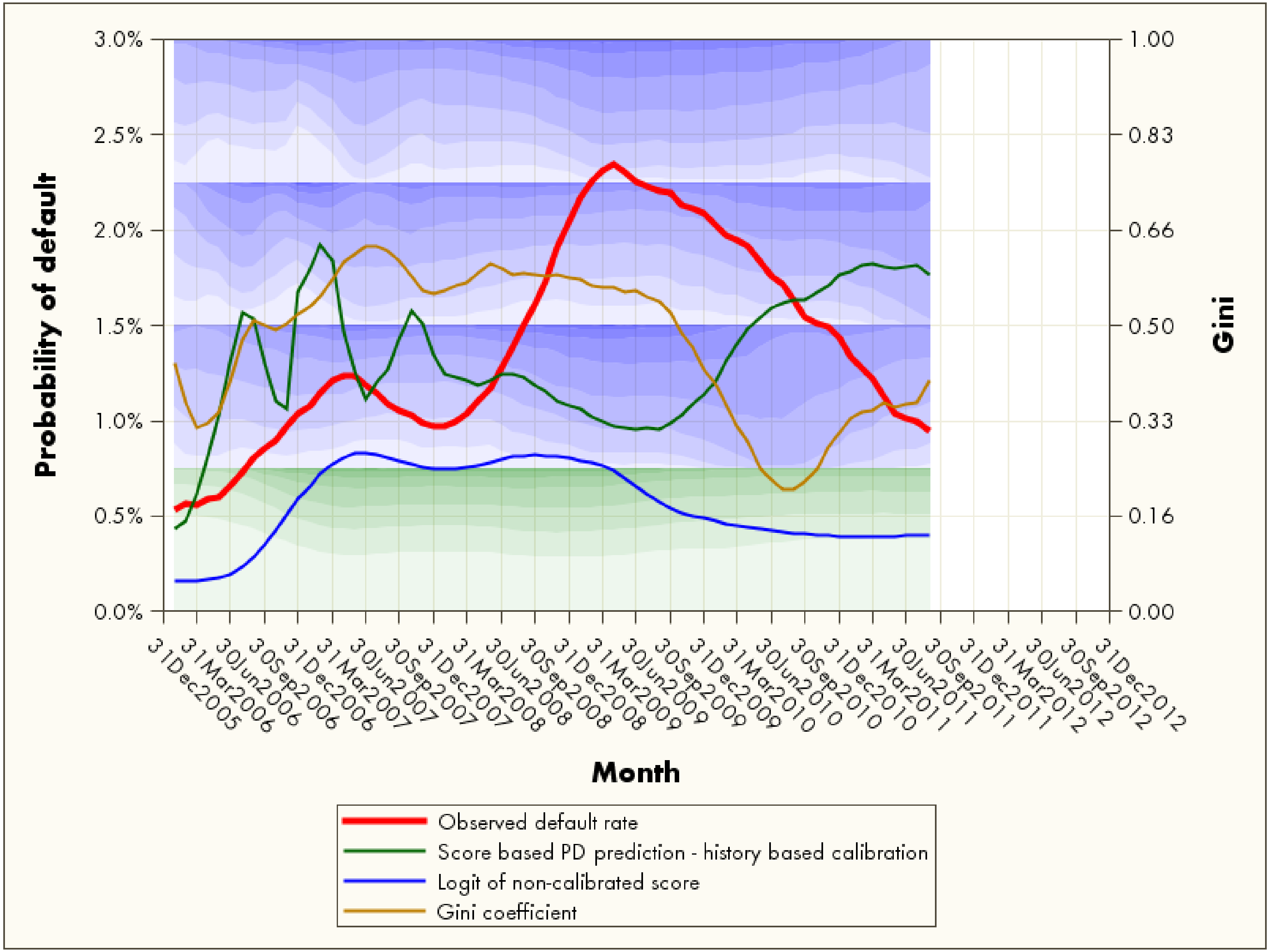
Důsledek: Pozorovaná míra nesplácení (default rate) je ovlivněna také skladbou těch klientů, kteří akceptují podmínky půjčky. Tento vliv může být někdy zásadní (fenomén tzv. „negativní selekce“).

Příklad 5:

Klientská skóringová funkce *nebo* produktové skóringové funkce ?

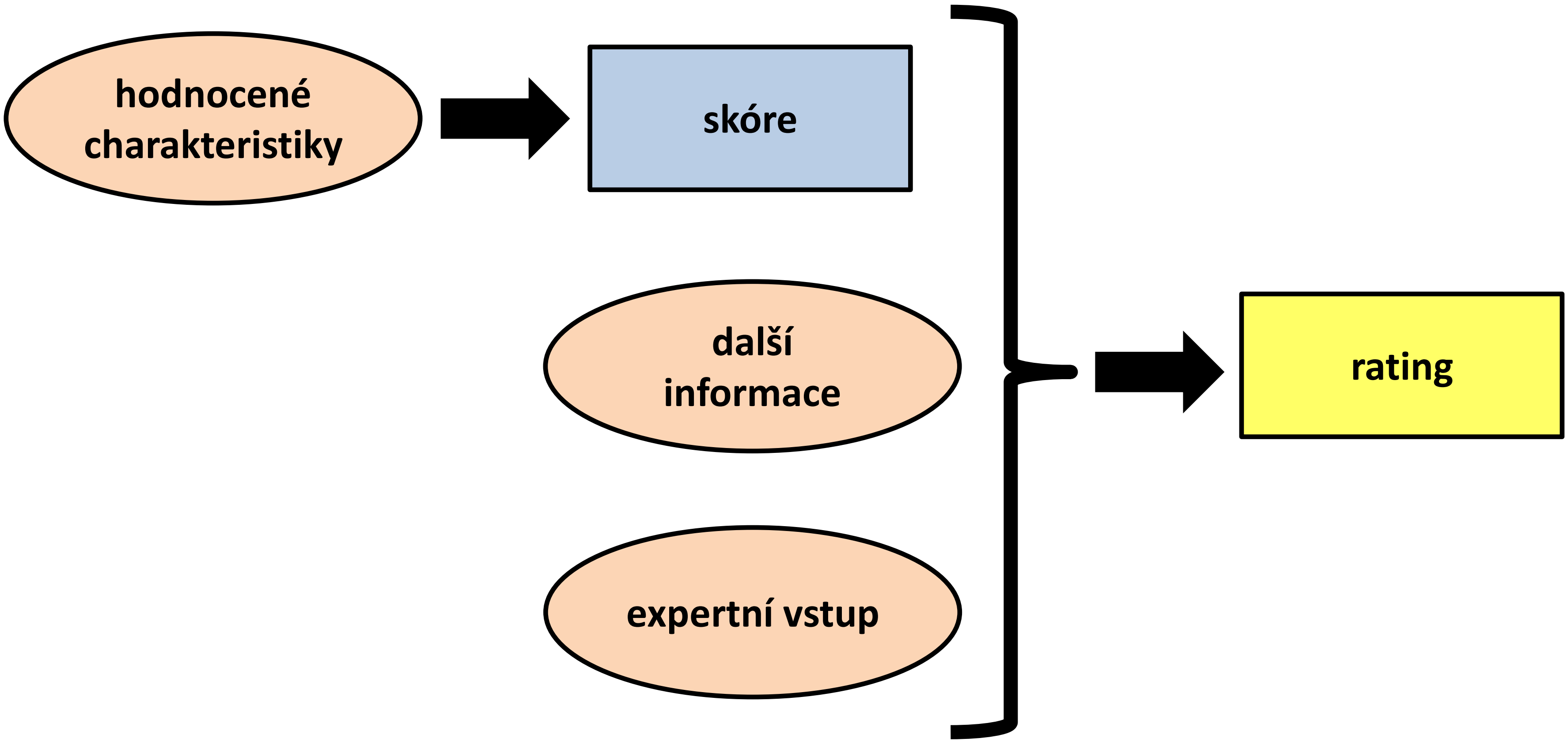
- klientský pohled: jednotný pohled na klienta, nižší pracnost tvorby a údržby modelů
- produktový pohled: zohlednění produktových specifik a vyšší výkonnost, nutnost vyvíjet a udržovat více modelů, roztříštěné vnímání klienta

Skóringové funkce – komplexita interpretace výkonnosti



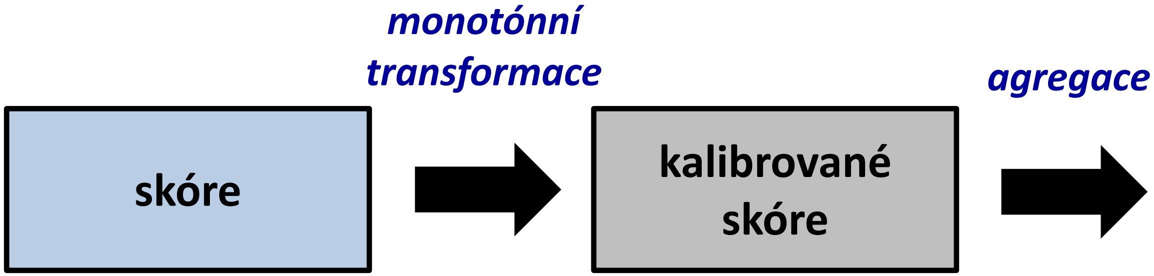
Rating

Rating = míra rizikovosti klienta vyjádřená hodnotou na zvolené diskrétní škále (tzv. **ratingová stupnice**)



Rating – kalibrace

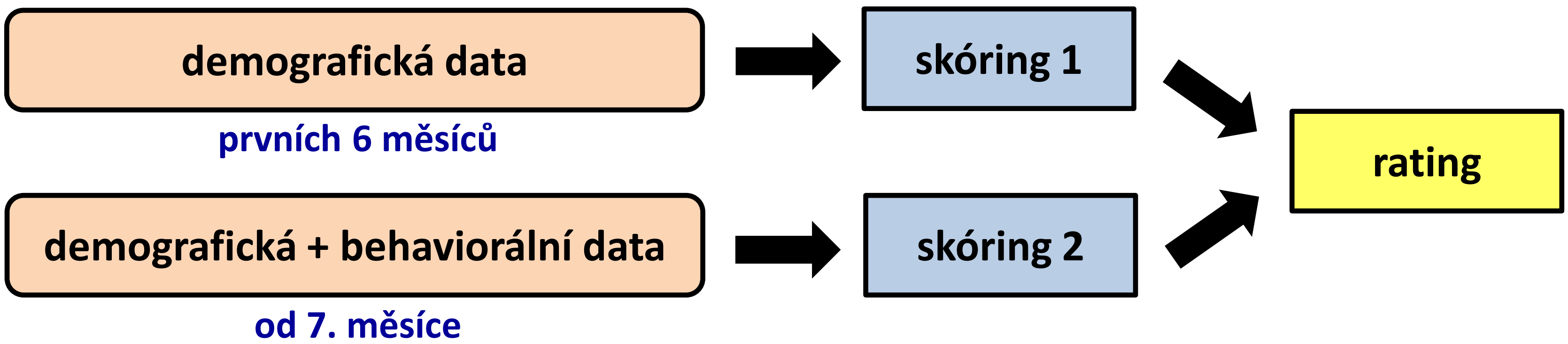
Jednotlivé ratingové stupně mohou mít předdefinovaný pravděpodobnostní význam (např. roční pravděpodobnost selhání a její možné rozmezí) – tzv. **master scale**.



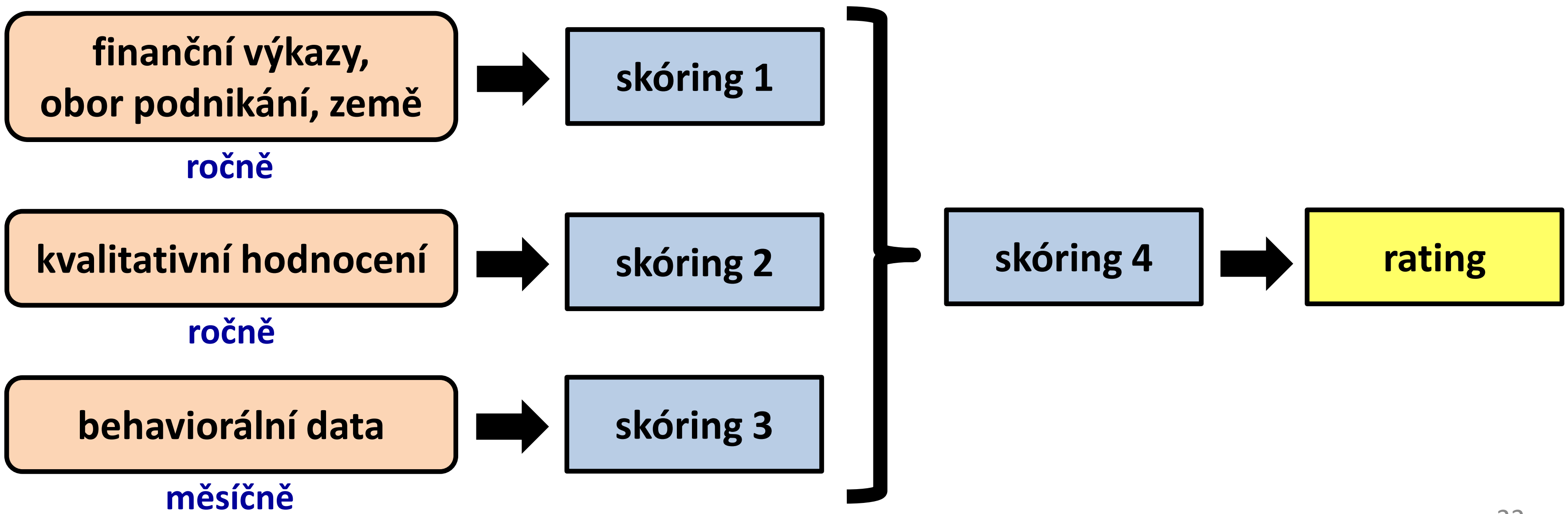
rating	pravděpodobnost od - do
1	0.00% - 0.17%
2	0.17% - 0.35%
3	0.35% - 0.69%
4	0.69% - 1.37%
5	1.37% - 2.70%
6	2.70% - 5.26%
7	5.26% - 10.00%
8	10.00% - 18.18%
9	18.18% - 100.00%

Rating a skóring – „architektura“

Fyzické osoby – příklad:



Podnikatelské subjekty – příklad:



Ilustrace 2:
Kreditní rizikové náklady

Kreditní rizikové náklady – obecný pohled 1/4

Cíl:

Stanovit takovou přirážku k úrokové sazbě, aby se celkový výnos z jejího výběru rovnal celkové očekávané ztrátě.

Příklad: Pokud je pravděpodobnost defaultu 4%, má být riziková přirážka rovna 4% ?

Formalizace problému:

$\xi(t)$ – dlužná částka v čase t podle smluvního splátkového kalendáře ($\xi(0)$ = výše úvěru, $\xi(t) = 0$ pro $t \geq$ datum splatnosti úvěru)

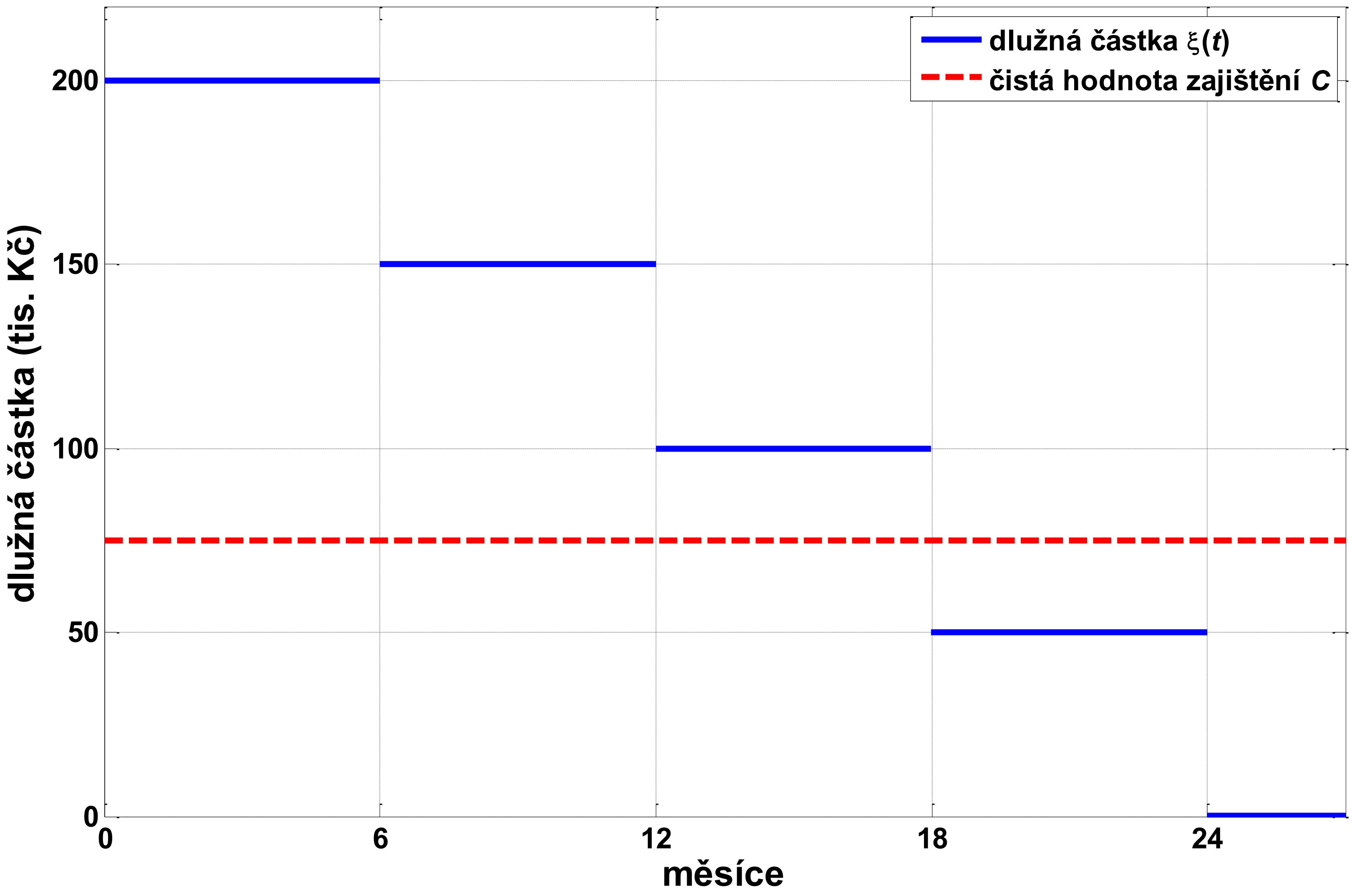
C – čistá hodnota zajištění (celková výtěžnost očištěná o náklady a očekávané riziko při realizaci zajištění)

T – náhodná veličina doby do defaultu ($T \geq 0$)

$j(t)$ – diskontní funkce, tj. výnos A v čase t má stejnou hodnotu jako výnos $e^{-j(t)}A$ v čase 0 (obvykle $j(t) = \alpha t$, kde α je diskontní úroková míra)

r – hledaná úroková sazba pokrývající očekávanou ztrátu, tzv. kreditní rizikové náklady (risk charges, risk costs apod.)

Kreditní rizikové náklady – obecný pohled 2/4



Kreditní rizikové náklady – obecný pohled 3/4

výběr generovaný sazbou *r* (náhodná veličina)**:**

$$\int_0^T e^{-j(t)}\, r\,\xi(t)\,\mathrm{d}t$$

ztráta (náhodná veličina)**:**

$$\max\left(\, e^{-j(T)}\xi(T) - C,\ 0\,\right)$$

požadavek:

$$\mathrm{E}\int_0^T e^{-j(t)}\, r\,\xi(t)\,\mathrm{d}t \overset{!}{=} \mathrm{E}\max\left(\, e^{-j(T)}\xi(T) - C,\ 0\,\right)$$

odtud:

$$r = \frac{\mathrm{E}\max\left(\, e^{-j(T)}\xi(T) - C,\ 0\,\right)}{\mathrm{E}\int_0^T e^{-j(t)}\,\xi(t)\,\mathrm{d}t}$$

Úvěrové produkty s náhodnou výší čerpání (kreditní karty, kontokorenty):

$\xi(t) \approx$ průměrná relativní výše čerpání daného produktu (napříč časem a klienty, kteří nejsou v selhání)

$\xi(T) \approx$ průměrná relativní výše čerpání klientů v defaultu (blízká 100%)

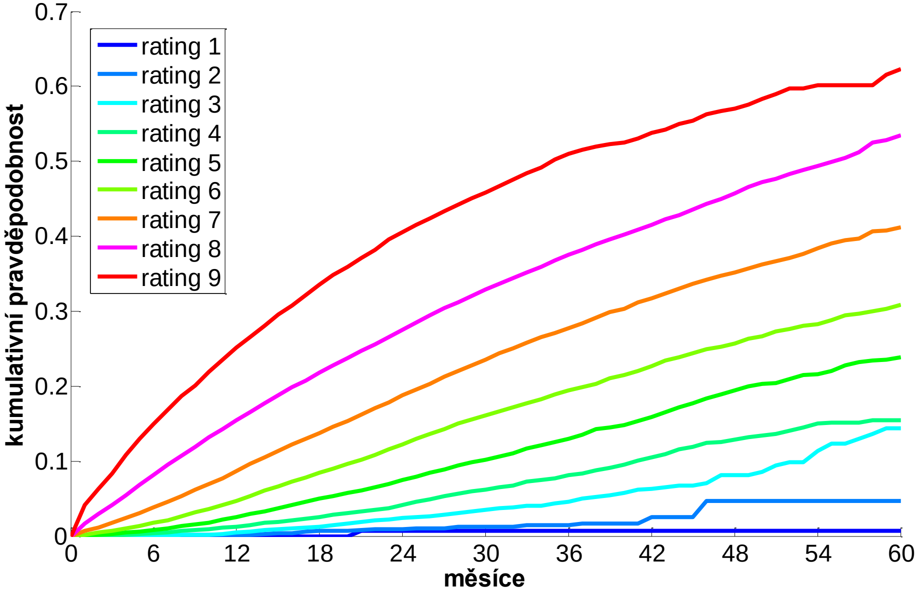
Kreditní rizikové náklady – obecný pohled 4/4

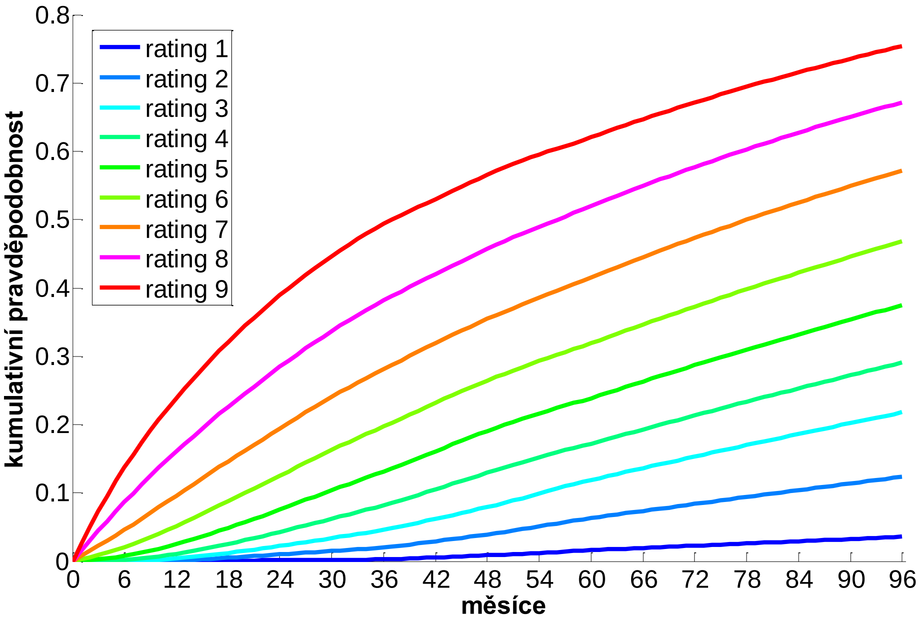
Poznámky:

- Rizikové náklady *r* je možné snadno vyjádřit explicitně s využitím distribuční funkce či hustoty a intenzity poruch náhodné veličiny doby do defaultu *T*. Alternativně je možné užít Markovské řetězce.
- Za předpokladu $P[\,T \leq t\,] = 1 - e^{-\lambda t}$ (exponenciálního rozdělení doby do defaultu) platí $r \equiv \lambda$.
- Rizikové náklady obecně nejsou monotónní vzhledem ke splatnosti úvěru. Obvykle bývají rostoucí funkcí doby do splatnosti pro málo rizikové klienty a klesající funkcí doby do splatnosti pro více rizikové klienty.

	splatnost v letech									
rating	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10+**
1	0.00%	0.03%	0.07%	0.14%	0.19%	0.22%	0.26%	0.30%	0.34%	0.38%
2	0.14%	0.28%	0.43%	0.56%	0.66%	0.71%	0.75%	0.79%	0.82%	0.84%
3	0.30%	0.61%	0.91%	1.16%	1.33%	1.42%	1.48%	1.52%	1.54%	1.56%
4	0.56%	1.17%	1.72%	2.13%	2.40%	2.54%	2.62%	2.65%	2.66%	2.67%
5	1.41%	2.36%	3.14%	3.70%	4.08%	4.26%	4.35%	4.39%	4.41%	4.41%
6	3.80%	4.87%	5.69%	6.29%	6.68%	6.89%	7.01%	7.08%	7.12%	7.14%
7	9.18%	9.83%	10.36%	10.81%	11.12%	11.29%	11.37%	11.42%	11.44%	11.45%
8	19.12%	18.70%	18.52%	18.68%	18.86%	18.95%	18.98%	18.99%	18.98%	18.97%
9	33.65%	32.28%	31.09%	30.88%	30.99%	31.11%	31.24%	31.40%	31.54%	31.66%

Kreditní rizikové náklady – příklad CDF





Empirické CDF doby do defaultu:

- nespolehlivé nebo chybějící odhady pro delší časová období
- možnost narušení uspořádání podle ratingu

Vyhlazení pomocí kvadratického programování za okrajových podmínek:

- monotonie v čase
- monotonie podle ratingu
- hladkost

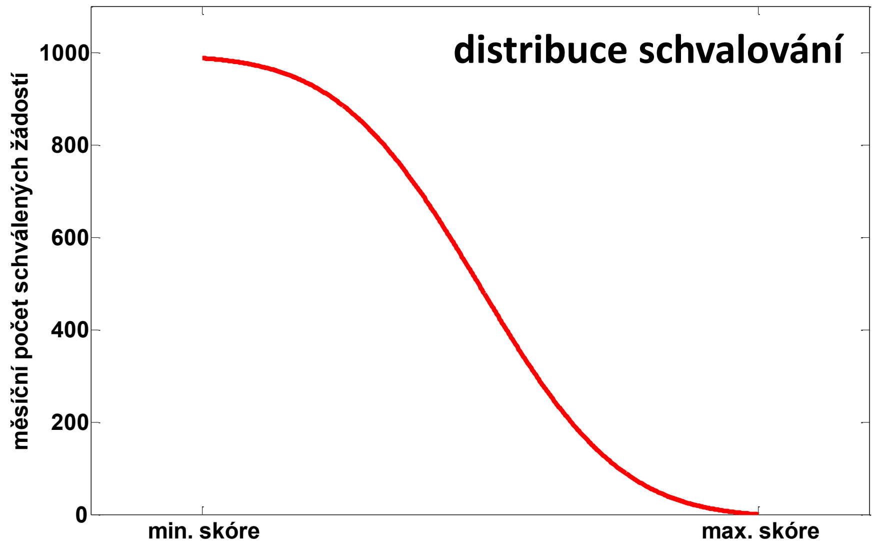
Ilustrace 3:
Optimální hladina zamítání

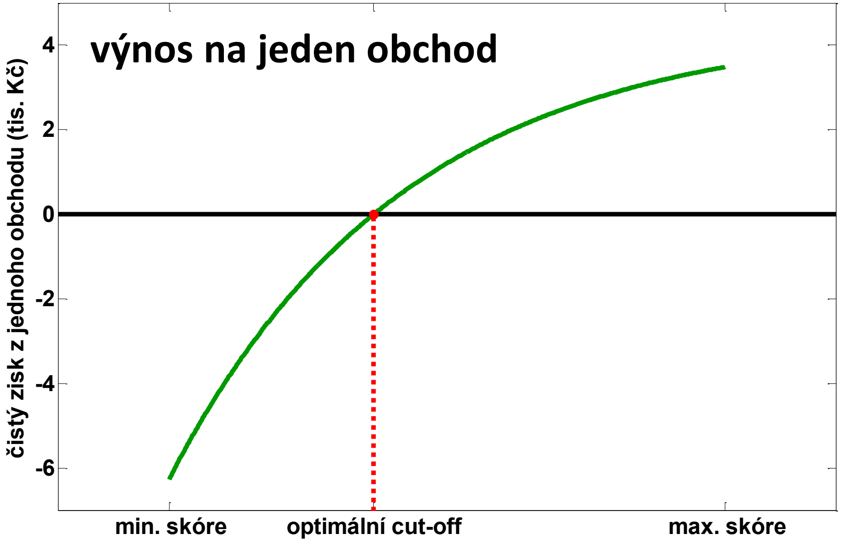
Optimální hladina zamítání – obecný pohled 1/2

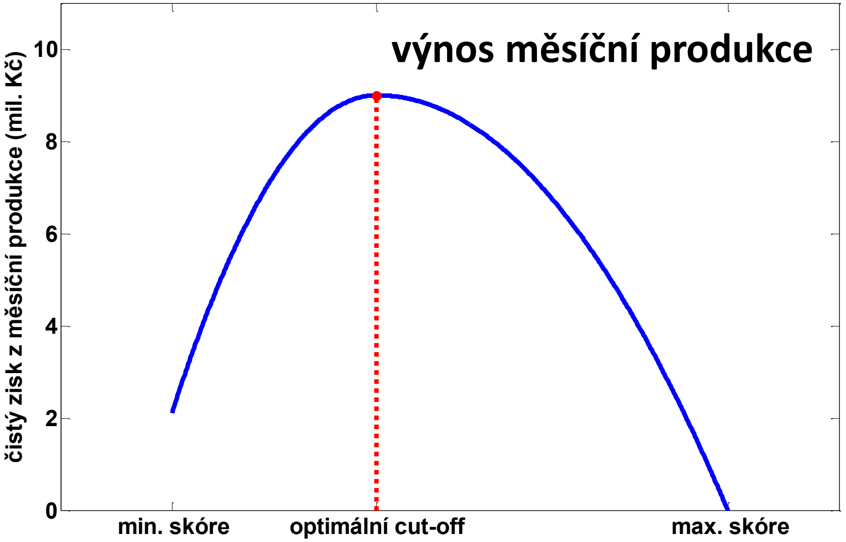
Problém: Pod jakou hladinou skóre zamítat žadatele o daný produkt?

Myšlenka modelu:

- vyjádřit čistý měsíční výnos portfolia jako funkci skóringového **cut-off**
- tuto funkci modelovat pomocí distribuce žádostí podle skóre a čistého příjmu z jednoho produktu podle rizikovosti žadatele







Optimální hladina zamítání – obecný pohled 2/2

Formalizace problému:

Čistý výnos z měsíční produkce (od poskytnutí do splatnosti) jako funkce cut-off *c*:

$$Z(c) = -\int_{c}^{\mathit{max.\ sk\acute{o}re}} v(s)\, \mathrm{d}F(s) \ ,$$

kde $F(s)$ je obráceně vzatá distribuční funkce schválených obchodů a $v(s)$ je současná hodnota čistého výnosu z jednoho obchodu (funkce výše úvěrového limitu, rizikových nákladů a klientské úrokové sazby, kde všechny tyto veličiny jsou funkcí skóre, očištěná o cenu zdrojů, náklady na poskytnutí a měsíční náklady na správu).

Doplňkové charakteristiky – očekávaný počet a objem defaultních obchodů:

$$-\int_{c}^{\mathit{max.\ sk\acute{o}re}} p(s)\, \mathrm{d}F(s) \quad \text{a} \quad -\int_{c}^{\mathit{max.\ sk\acute{o}re}} p(s)\, L(s)\, \mathrm{d}F(s) \ ,$$

kde $p(s)$ je pravděpodobnost defaultu a $L(s)$ je průměrný schválený limit.

Doplňkové charakteristiky – očekávaný počet a objem zamítnutých obchodů:

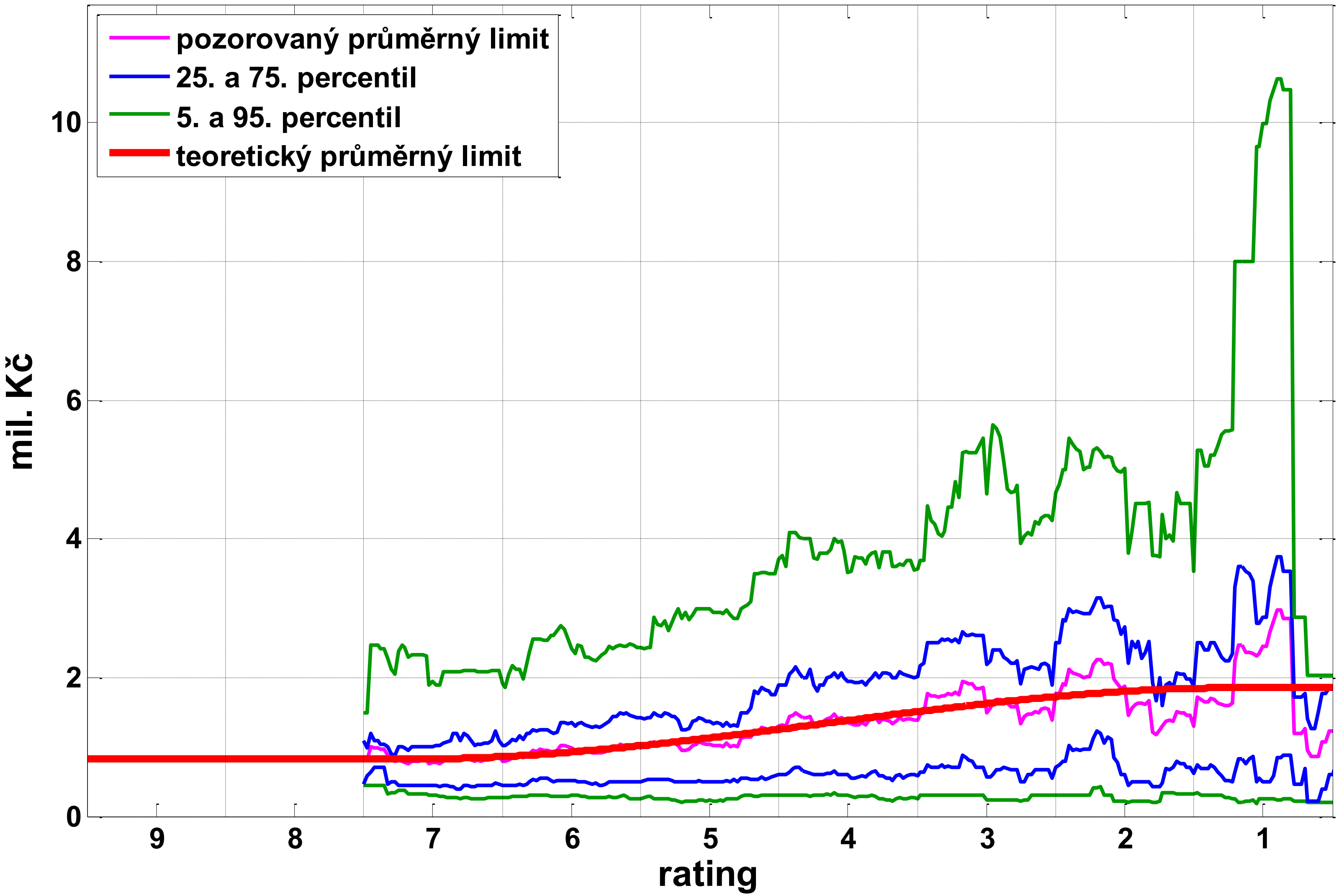
$$F(\mathit{min.\ sk\acute{o}re}) - F(c) \quad \text{a} \quad -\int_{c}^{\mathit{max.\ sk\acute{o}re}} L(s)\, \mathrm{d}F(s) \ .$$

Statistická úloha:

Odhady a vyhlazení uvedených funkcí včetně případné extrapolace do oblasti skóre, kde se dosud obchody neposkytovaly.

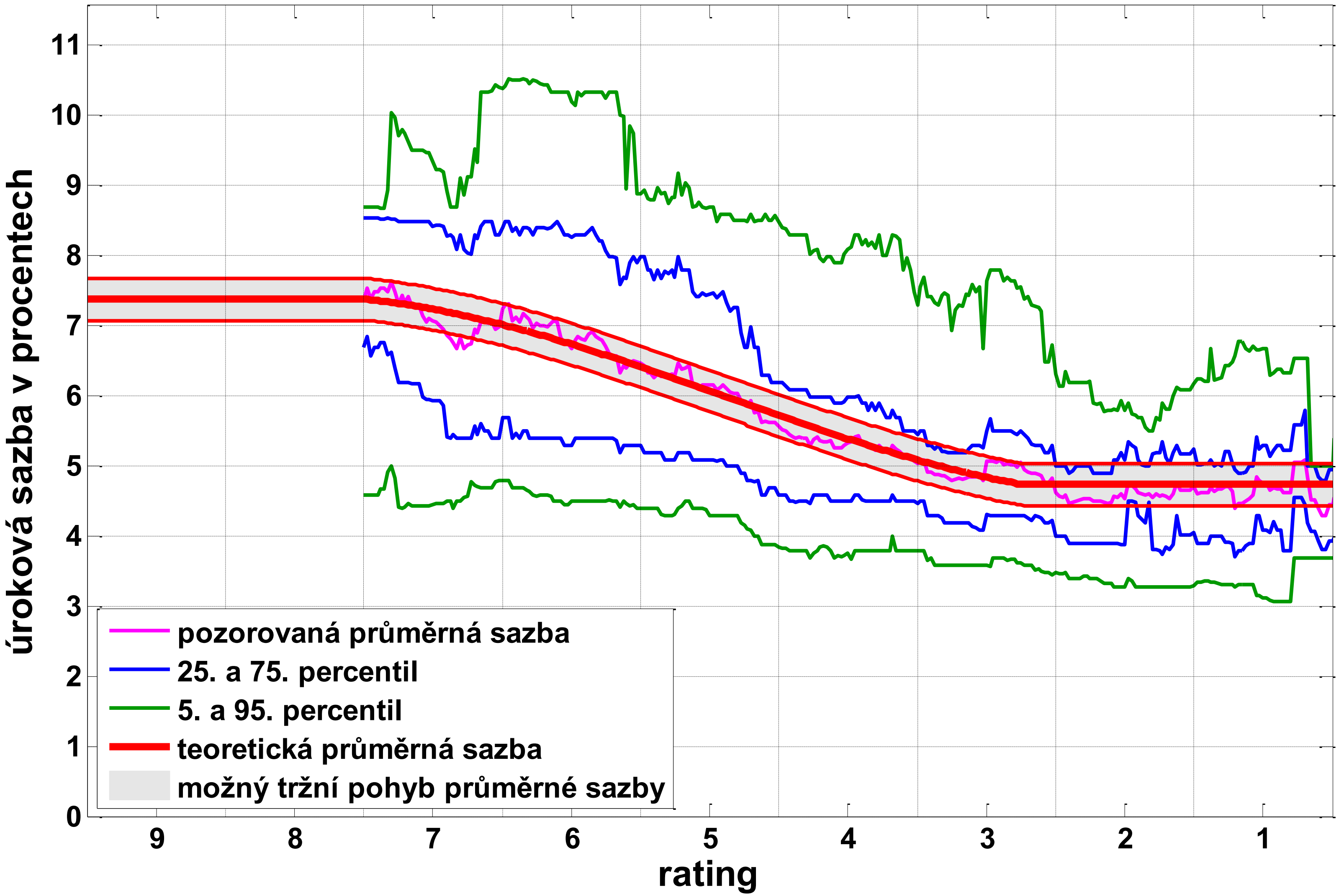
Optimální hladina zamítání – příklad

Výše úvěrového limitu jako funkce skóre:



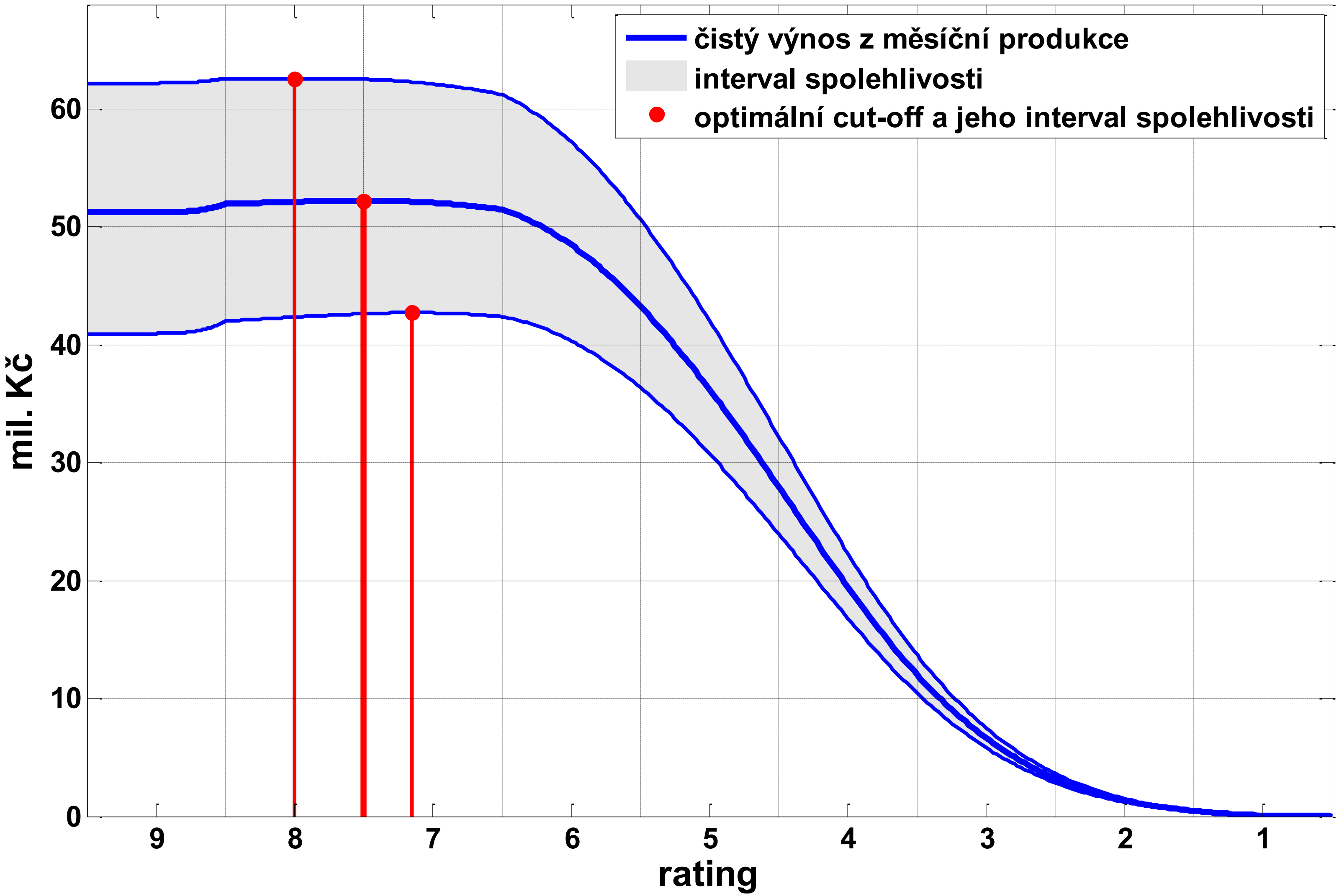
Optimální hladina zamítání – příklad 2/4

Výše úrokové sazby jako funkce skóre:



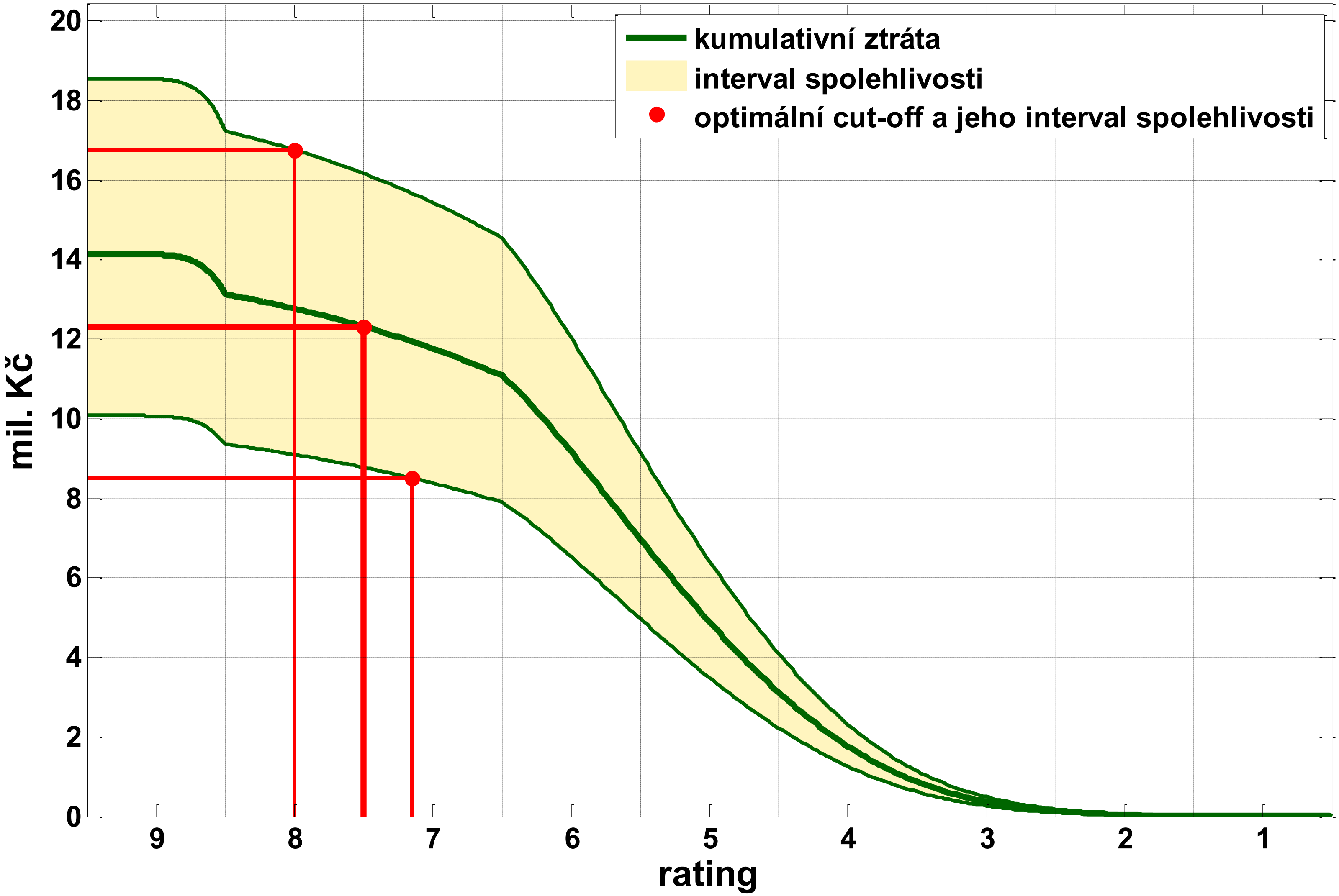
Optimální hladina zamítání – příklad 3/4

Ukázka výstupu modelu – optimální cut-off:



Optimální hladina zamítání – příklad 4/4

Ukázka výstupu modelu – roční kumulativní ztráta z měsíční produkce:



Ilustrace 4:
PIT a TTC odhady pravděpodobností defaultu

PIT a TTC odhady

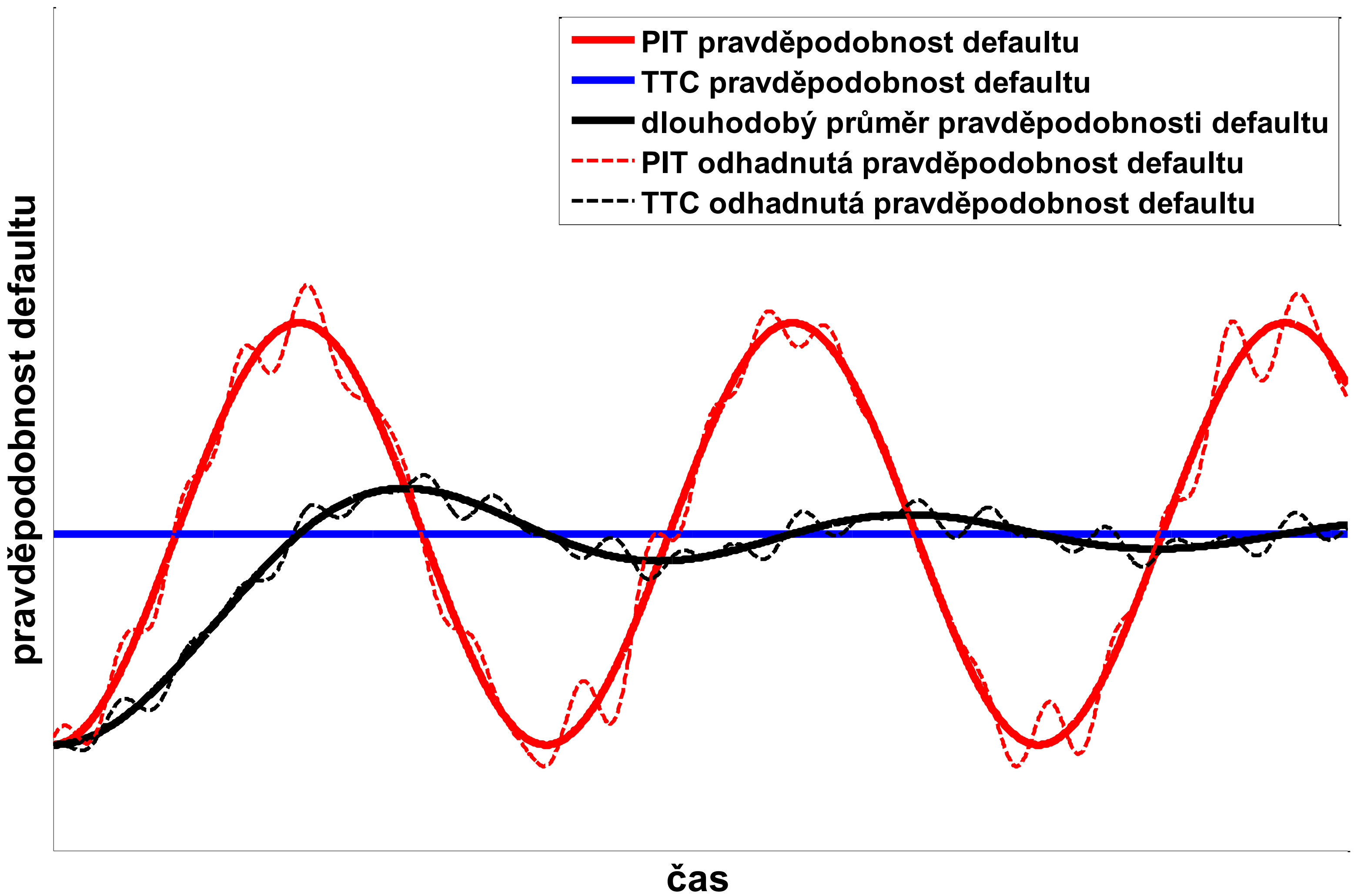
Point in Time (PIT) přístup = odhad pravděpodobnosti defaultu na základě aktuální historie, popř. korigovaný o ekonomická očekávaní

- Popisuje aktuální situaci obchodu/portfolia v aktuální makroekonomické situaci.
- Hodnota PD pro obchod/portfolio s konstantním rizikovým profilem fluktuuje v rámci ekonomického cyklu.
- Používá se pro kalibraci skóringových modelů.

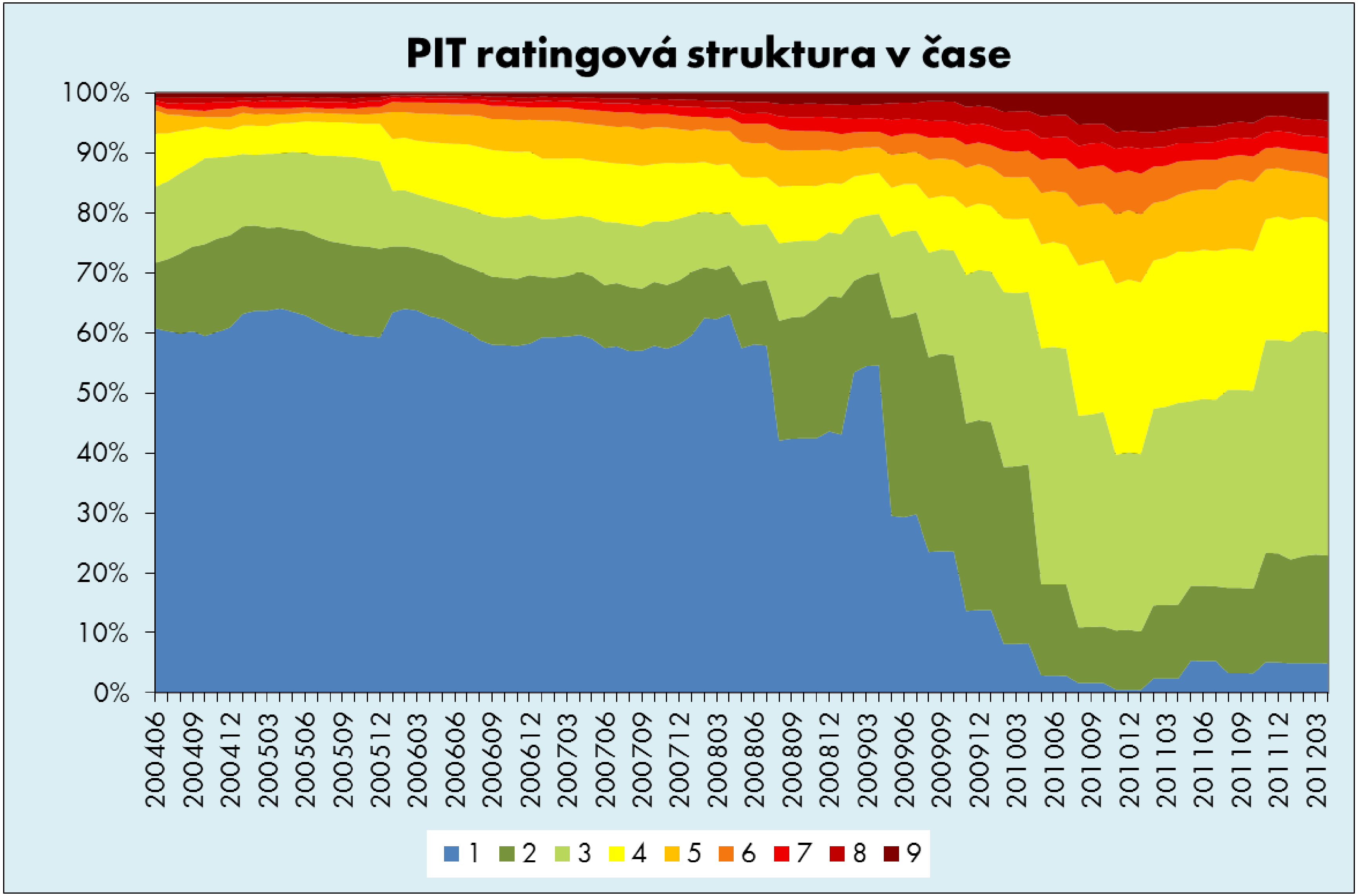
Through the Cycle (TTC) přístup = pravděpodobnost defaultu „průměrující" vliv makroekonomického cyklu

- Vyjadřuje teoretickou dlouhodobou průměrnou hladinu pravděpodobnosti defaultu.
- Pro obchodu/portfolio s konstantním rizikovým profilem je v průběhu ekonomického cyklu stabilní, a tedy v době krize nižší a naopak v době konjunktury vyšší než PIT pravděpodobnost defaultu.
- Používá se pro výpočet tzv. kapitálového požadavku (rizikově vážená aktiva).

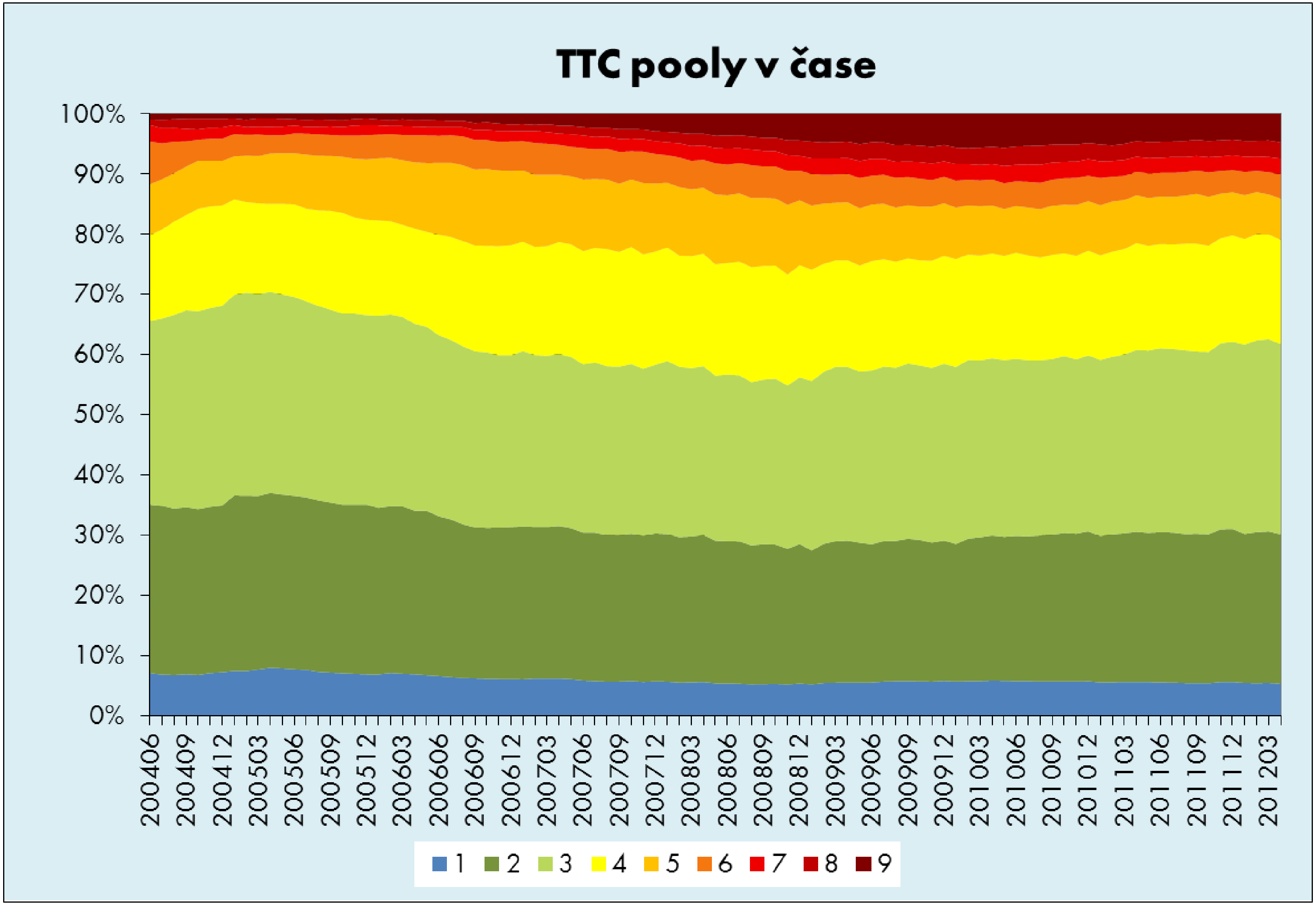
PIT a TTC odhady



PIT a TTC odhady



PIT a TTC odhady 4/5



PIT a TTC odhady 5/5