

Kalibrace kredit skóringových modelů

Eva Horáčková

A decorative graphic element consisting of a thick teal horizontal bar that spans the width of the slide. Below this bar, on the right side, are three thin, parallel white horizontal lines that extend to the right edge of the slide.

Obsah

- Definice kalibrace
- Využití kalibrace
- Způsoby použití kalibrace
- Metody kalibrace
- Testování kalibrace
 - Parametrické testování

Kalibrace

Prvotním cílem CS karet byla pomoc při seřazení klientů a rozhodování o přijetí/nepřijetí

Dnes je důraz kladen i na přesnost odhadu pravděpodobnosti selhání klienta.

„Použití bandingu, scalingu a dalších metod tak, aby bylo zajištěno, že výsledné skóre mají stejný význam napříč různými skóringovými kartami a znázorňují pravděpodobnost selhání.“

Využití kalibrace

Konzistence mezi modely - porovnání skóre vytvořené různými CS modely

Změna plošných podmínek ve společnosti - změna, která se projeví stejně u všech klientů

Jinak definovaná pravděpodobnost selhání – PS známe, ale potřebujeme znát i pro jinou definici selhání

Odhad pravděpodobnosti selhání klienta (Basel II)

Odhad pravděpodobnosti selhání

Klíčový faktor pro oceňování úvěru a stanovení kapitálových rezerv. Směrnice Basel II doporučuje, aby byl kapitálový požadavek stanoven na základě ratingu a skóringu

Basel II – možnosti výpočtu kapitálového požadavku

Standardizovaný přístup - plošný výpočet, používaný v Basel I
Externí rating – využití ratingu ratingových agentur (stanovují rating pouze pro velké společnosti)

Interní rating (IRB – Internal Ratings-based Approach) – využití vlastních odhadů pravděpodobnosti selhání

Použití kalibrace

- **Je vypočítávána na základě skutečné průměrné pravděpodobnosti selhání vzorku**
- Během vývoje modelu, v závěru
- Rozsáhlost je závislá na technice tvorby CS modelu, čase, metrických rozdílech
- Modely tvořené logistickou regresí – kalibrace není nutná (v praxi vždy neplatí)

Způsoby kalibrace

- Změna koeficientů modelu (pouze regresní modely)
- Rozdělení do skupin (banding)
- Lineární posun a škálování
- Transformace pomocí vhodné funkce (beta funkce)
- Lineární programování (složitě, nepoužívané v praxi)

Banding - libovolné skupiny

Rozdělení do 5-25 (30) skupin, 5-10 skupin pro homogenní nebo omezená data, skupiny nejsou nijak definovány

Stanovení počtu skupin:

- Calinski-Harabaszova statistika

$$CH = \frac{\sum_{k=1}^g n_k (p_k - p)^2}{g-1} \bigg/ \frac{\sum_{k=1}^g \sum_{i=1}^{n^k} n^k (P_{i,k} - P_k)^2}{n-g}$$

- Nesrovnatelné s výsledky jiného modelu

Banding - Benchmarking

- Pevná průměrná PS skupiny
- Minimalizace součtu čtverců
- mezi danou PS a PS klientů
- Třídy ratingových agentur
- Postup
 - Počet skupin
 - Hranice shruba stejných skupin
 - Součet čtverců
 - Posun hranic o bod, součet čtverců

Investment Grade	'AAA'	Extremely strong capacity to meet financial commitments. Highest rating
	'AA'	Very strong capacity to meet financial commitments
	'A'	Strong capacity to meet financial commitments, but somewhat susceptible to adverse economic conditions and changes in circumstances
	'BBB'	Adequate capacity to meet financial commitments, but more subject to adverse economic conditions
Speculative Grade	'BBB-'	<i>Considered lowest investment grade by market participants</i>
	'BB+'	<i>Considered highest speculative grade by market participants</i>
	'BB'	Less vulnerable in the near-term but faces major ongoing uncertainties to adverse business, financial and economic conditions
	'B'	More vulnerable to adverse business, financial and economic conditions but currently has the capacity to meet financial commitments
	'CCC'	Currently vulnerable and dependent on favorable business, financial and economic conditions to meet financial commitments
	'CC'	Currently highly vulnerable
	'C'	A bankruptcy petition has been filed or similar action taken, but payments of financial commitments are continued
	'D'	Payment default on financial commitments

Ratings from 'AA' to 'CCC' may be modified by the addition of a plus (+) or minus (-) sign to show relative standing within the major rating categories.

Banding - marginal risk boundaries

Třídy stanovené nejen průměrnou pravděpodobností selhání, ale také min a max hranicemi

Využívá se metodologie založená na odhadu mezního rizika
– složitý úkol, vztah mezi skóre a rizikem není zpravidla přes celou škálu totožný

Metoda nejbližšího souseda

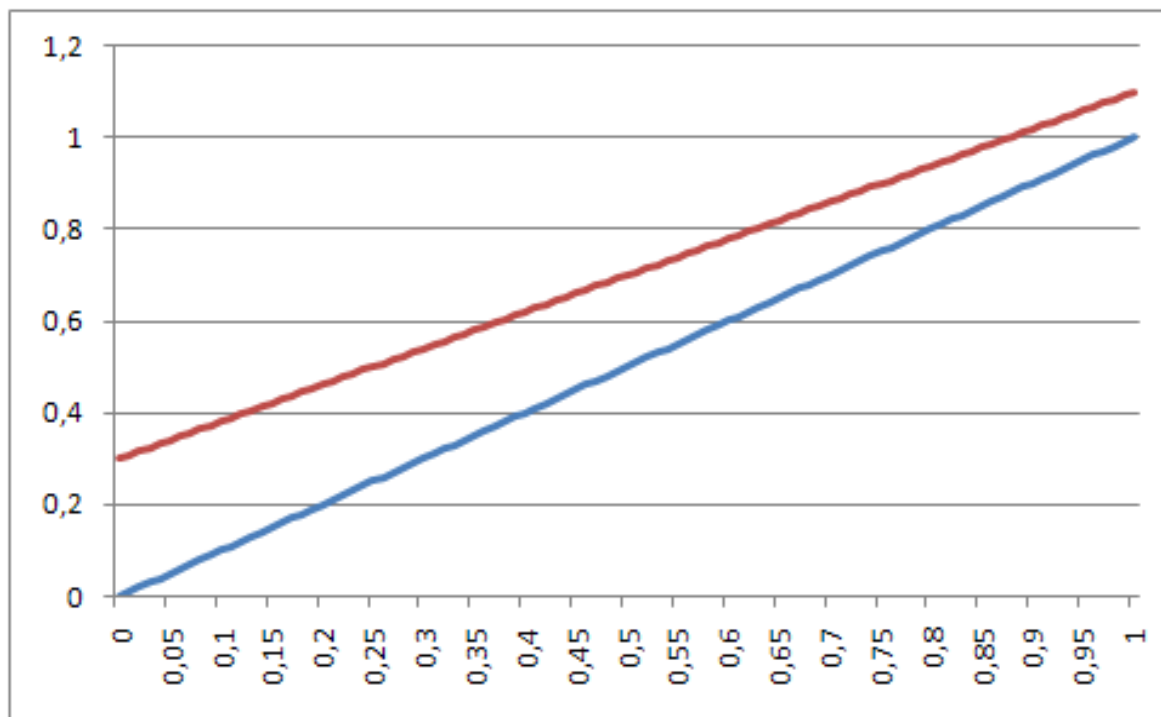
Rozdělení do skupin a interpolace

Lineární transformace

Shift, scaling

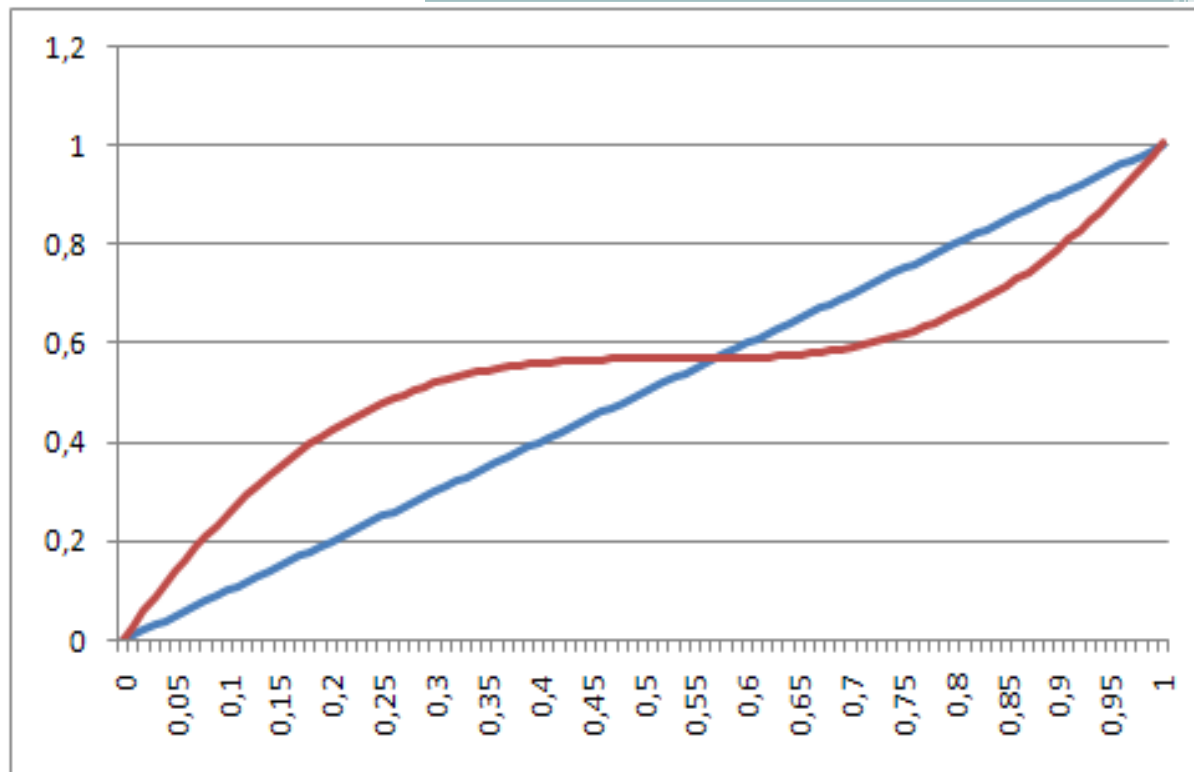
$$PD = a + b * S$$

Lineární vztah mezi PD a skóre



Beta funkce

Nelineární vztah mezi
PD a skóre
Nutno nalézt vhodnou
transformační funkci



Často využívaná je beta funkce

$$B(\alpha, \beta, x) = \frac{1}{\int_0^1 t^{\alpha-1} (1-t)^{\beta-1} dt} x^{\alpha-1} (1-x)^{\beta-1}$$

Testování kalibrace

Testování podmíněných odhadů-

Závislost odhadů na daném prostředí (kovariance s HDP, nezaměstnaností), přijímáme předpoklad nezávislosti selhání jednotl. klientů

Odhad PD ozn. PIT (Point-in-time)

Testování nepodmíněných odhadů-

Nezávislost na prostředí, odhad PD ozn. TTC (Through-the-Cycle). Nelze přijmout předpoklad nezávislosti

Testování kalibrace

Podmíněné testování

- Binomický test
 - Normální aproximace
- Chí kvadrát testy (Hosmer-Lemeshovův test)
- Spiegelhalterův test

Nepodmíněné testování

Binomický test

- Testování kalibrace jedné skupiny
- Selhání zákazníků jsou nezávislá

H_0 : Odhadnutá PS skupiny = pozorované PS

$$B = \min (b | P(X \leq b) > \alpha) = \min \left(b | \sum_{i=b}^I \binom{I}{i} P_S^i (1 - P_S)^{I-i} > \alpha \right)$$

- X má binomické rozdělení, b – počet špatných uchazečů

Normální aproximace binomického testu

Lze využít pokud je v dané skupině 10 a více klientů (přijímáme předpoklad normálního rozložení klientů)

$$z = \frac{b - IP_S}{\sqrt{IP_S(1 - P_S)}}$$

I – počet prvků ve skupině

Z-test o střední hodnotě při známém rozptylu

Chí kvadrát testy

- Hosmer-Lemeshowův test
- Testování všech skupin současně
- H_0 : Očekávané a pozorované PS nejsou stejné
- Chí-kvadrát rozdělení s $I - 2$ stupni volnosti

$$HL = \sum_{i=1}^I z_i^2 = \sum_{i=1}^I \left(n_i \frac{(P_i - O_i)^2}{O_i(1 - O_i)} \right)$$

Spiegelhalterův test

- Testuje do skupin nerozřazená data
- H_0 : střední hodnota skóre = střední hodnota kalibrovaných výsledků

$$Z = \frac{MSE - E[MSE]}{\sqrt{\text{var}[MSE]}}$$

$$MSE = \frac{1}{n} \sum_{i=1}^n (B_i - p_i)^2$$

$$E[MSE] = \frac{1}{n} \sum_{i=1}^n p_i(1 - p_i), \text{var}[MSE] = \frac{1}{n^2} \sum_{i=1}^n p_i(1 - p_i)(1 - 2p_i)^2$$

Nepodmíněné testování

Není potřeba předpokladu nezávislosti mezi skupinami dat,
předpokládáme nezávislost v čase

Nulovou hypotézu, že pravdivá PD je menší nebo rovna q
zamítáme, pokud

$$\frac{1}{T} \sum_{\tau=1}^T \frac{d_{\tau}}{n_{\tau}} > q + \frac{\hat{\sigma}}{\sqrt{T}} \Phi^{-1}(1 - \alpha)$$

T – počet let, n_t – počet klientů, d_t – počet špatných, q –
odhad PD

Děkuji za pozornost