

Ryzyko kredytowe towarzyszy nieodłącznie działalności bankowej. Tradycyjne podejście do oceny ryzyka kredytowego polega na stosowaniu analizy dyskryminacyjnej (dla przedsiębiorstw) i regresji logistycznej (dla klientów indywidualnych) w modelach kredytowych. Współczesne oczekiwania, dotyczące analizy ryzyka, są coraz większe. Obecnie stosowane metody okazują się niewystarczające i banki poszukują innych, które można byłoby użyć w modelach ryzyka kredytowego. Jedną z nich jest analiza przetrwania, której popularność w ostatnich latach wzrosła. Niniejsza książka w całości poświęcona jest możliwości zastosowania analizy przetrwania w ocenie ryzyka kredytowego klientów indywidualnych oraz pokazuje korzyści wynikające z używania tej metody.

W pracy znalazło się wiele zagadnień związanych z tą metodą, w szczególności:

- dane cenzurowane,
- tabele trwania życia,
- estymacja krzywej przetrwania metodą Kaplana-Meiera,
- modele parametryczne,
- model proporcjonalnego hazardu Coxa,
- ryzyka konkurujące.

W badaniach, przeprowadzonych na użytek niniejszej publikacji, skupiono się na modelach przewidujących niewypłacalność klienta. Jednakże zaprezentowane w książce przykłady użycia analizy przetrwania nie wykorzystują wszystkich możliwości jej zastosowania w działalności bankowej.

Publikacja skierowana jest do osób, które zajmują się budową modeli kredytowych w instytucjach finansowych, jak również innych, chcących poszerzyć wiedzę w zakresie analizy przetrwania, a następnie wykorzystać ją w praktyce.

## Wstęp 9

### Część I

#### Analiza przetrwania

#### Rozdział 1

##### Ogólna charakterystyka analizy przetrwania 17

##### 1.1. Definicja 17

##### 1.2. Dane cenzurowane i dane obcięte 19

##### 1.2.1. Dane cenzurowane 19

##### 1.2.2. Dane obcięte 22

##### 1.3. Możliwości zastosowania analizy przetrwania 24

##### 1.4. Podstawowe definicje związane z analizą przetrwania 25

#### Rozdział 2

##### Przegląd zastosowania analizy przetrwania w literaturze fachowej 31

#### Rozdział 3

##### Modele w analizie przetrwania 57

##### 3.1. Tabele trwania życia 57

##### 3.2. Przykład estymacji krzywej przetrwania za pomocą metody K-M 63

- 3.3. Inne podejście nieparametryczne 64
  - 3.3.1. Porównanie krzywych przetrwania 65
  - 3.3.2. Współczynnik hazardu (hazard ratio - HR) 66
  - 3.3.3. Test logarytmiczny rang (Log-rank test) 67
- 3.4. Modele parametryczne 73
  - 3.4.1. Modele AFT 74
  - 3.4.2. Rozkład wykładniczy 76
  - 3.4.3. Rozkład Weibulla 77
  - 3.4.4. Rozkład log-normalny 78
  - 3.4.5. Pozostałe modele parametryczne 80
- 3.5. Model proporcjonalnego hazardu Coxa 82
- 3.6. Porównywanie modeli 84
  - 3.6.1. Reszty 84
  - 3.6.2. Wskaźnik wiarygodności (likelihood ratio - LR) 85
- 3.7. Model proporcjonalnych szans (proportional odds) 86
- 3.8. Podsumowanie 88

## Część II

Zastosowanie analizy przetrwania w modelowaniu ryzyka kredytowego

### Rozdział 4

Ryzyko kredytowe związane z kredytowaniem klienta indywidualnego 93

- 4.1. Pojęcie ryzyka kredytowego 93
- 4.2. Kredyty oferowane klientom indywidualnym 94
- 4.3. Zdolność kredytowa 97
- 4.4. Ocena ryzyka kredytowego związanego z kredytowaniem klienta indywidualnego 98
- 4.5. Credit scoring 99
- 4.6. Regresja logistyczna 103
- 4.7. Zastosowanie regresji logistycznej w modelu ryzyka kredytowego 106

### Rozdział 5

Wykorzystanie analizy przetrwania przy grupowaniu atrybutów (coarse-classifying) 109

- 5.1. Statystyka D Somersa 110
- 5.2. Grupowanie zmiennych: ciągłej i skokowej 110
  - 5.2.1. Grupowanie zmiennej ciągłej: okres współpracy z bankiem 111
  - 5.2.2. Grupowanie zmiennej skokowej: zawód 112

### Rozdział 6

Zastosowanie analizy przetrwania w modelowaniu niewypłacalności klienta 115

- 6.1. Dane stosowane do budowy modeli wykorzystujących analizę przetrwania 115
- 6.2. Analiza wybranych przypadków w bazie danych 117
- 6.3. Estymacja krzywej przetrwania metodą Kaplana-Meiera (K-M) 119
- 6.4. Sprawdzenie hipotez o jednakowym rozkładzie krzywych przetrwania dla dwóch grup 120
- 6.5. Sprawdzenie hipotez o jednakowym rozkładzie krzywych przetrwania dla trzech grup 122
- 6.6. Tabela trwania życia 124
  - 6.6.1. Wybór zmiennych 128
  - 6.6.2. Wybór rozkładu funkcji hazardu 129
  - 6.6.3. Wyniki modeli 137

## Rozdział 7

Sprawdzanie założenia proporcjonalności hazardu zmiennych wykorzystywanych w modelu Coxa 141

7.1. Metoda graficzna 142

7.1.1. Porównanie estymowanych funkcji  $[-\ln(-\ln(S(t)))]$  dla rozważanych kategorii zmiennych 142

7.1.2. Porównanie obserwowanej i przewidywanej krzywej przetrwania 143

7.2. Test dopasowania 144

7.3. Zastosowanie zmiennej zależnej od czasu 145

7.4. Testowanie proporcjonalności zmiennych w analizowanym przykładzie 146

7.4.1. Metoda graficzna 147

7.4.2. Testowanie proporcjonalności zmiennych za pomocą testu zgodności 149

7.4.3. Testowanie proporcjonalności zmiennych za pomocą procedury ASSESS, wykorzystującej reszty martyngałowe 151

7.5. Zestawienie wyników 155

7.6. Budowa modelu Coxa 156

7.7. Uwzględnianie charakterystyk zmiennych w czasie 158

7.8. Budowa modelu przy wykorzystaniu regresji logistycznej 164

7.8.1. Model warstwowy proporcjonalnego hazardu (stratified proportional hazard model) 166

7.8.2. Porównanie metod: regresji logistycznej i analizy przetrwania 171

## Rozdział 8

Model analizy przetrwania z sieciami neuronowymi 173

8.1. Sieci neuronowe 174

8.2. Model analizy przetrwania z sieciami neuronowymi 176

8.3. Opis danych wykorzystanych w badaniu 177

8.4. Wyniki zbudowanych modeli 180

8.4.1. Wyniki modeli dla pierwszego okresu 180

8.4.2. Wyniki modeli dla drugiego okresu 185

8.4.3. Wyniki modeli dla trzeciego okresu 189

8.5. Wnioski 192

## Rozdział 9

Wykorzystanie analizy przetrwania w modelu dla portfela kredytów hipotecznych 193

9.1. Opis modelu i zmiennych 194

9.2. Budowa modelu 196

9.3. Struktura modelu 198

9.4. Wynik modelu 200

## Rozdział 10

Ryzyka konkurujące (Competing risks) 203

10.1. Wprowadzenie 203

10.2. Określone rodzaje hazardów (type-specific hazards) 204

10.3. Podsumowanie 217

## Zakończenie

Aneks 231

Spis rysunków 237  
Spis tabel 241