

Testy dobré shody pro model zrychleného času v analýze přežití

Petr Novák

3. února 2010

Analýza přežití - o co jde

- Studujeme data reprezentující doby do nějaké většinou negativní předem definované události (úmrť pacienta, porucha součástky)
- Cenzorování zprava - někteří jedinci vyňati ze sledování před dosažením události
- T_i^* : skutečné časy událostí, C_i : časy cenzorování
- $T_i = \min(T_i^*, C_i)$: pozorované časy konce sledování, $\Delta_i = I(T_i \leq C_i)$: indikátor necenzorované události
- \mathbf{X}_i , hodnoty regresorů

- Data ve tvaru $(T_i, \Delta_i, \mathbf{X}_i)_{i=1}^n$
- Chceme rozumný popis dat - pomocí regresního modelu

Coxův model proporcionálního rizika (Cox,1972)



$$\alpha_i(t) = \alpha_0(t) \exp(\mathbf{X}_i^T \beta)$$

- $\alpha_i(t) = \lim_{h \rightarrow 0^+} P(t \leq T_i^* < t + h | T_i^* \geq t) / h$ riziková funkce
- $\alpha_0(t)$ základní riziko

Model zrychleného času (Accelerated Failure Time - AFT, Buckley&James, 1979)



$$\log T_i^* = -\mathbf{X}_i^T \beta + \epsilon_i, \quad \epsilon_i(\text{iid})$$

- Virtuální čas každé jednotky plyne rychleji/pomaleji dle hodnot regresorů

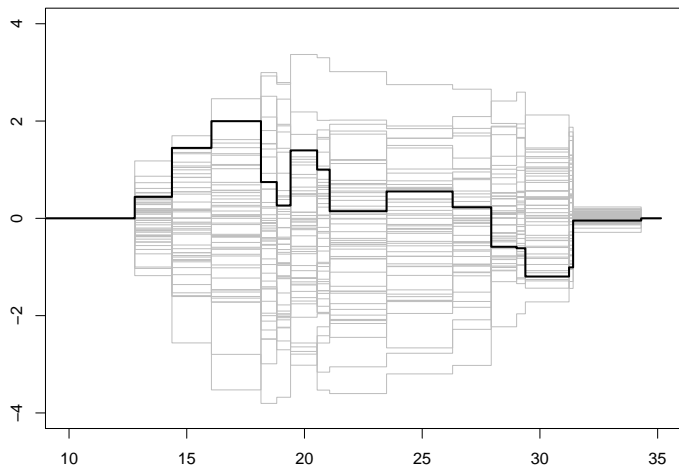
Testy dobré shody pro Coxův model

- Reprezentace pomocí čítacích procesů
 $N_i(t) = I(t \geq T_i, \Delta_i = 1), Y_i(t) = I(T_i \geq t)$
- $M_i(t) = N_i(t) - \Lambda_i(t) = N_i(t) - \int_0^t Y_i(s)\alpha_i(s)ds$ martingaly
vzhledem k $\mathcal{F}_{t-} = \sigma \{N_i(s), Y_i(s), \mathbf{X}_i(t), 0 \leq s < t, i = 1, \dots, n\}$
- Užitím martingalové CLV lze najít proces který jde snadno simulovat a je asymptoticky ekvivalentní se skórovým procesem

$$\tilde{U}(\hat{\beta}, t) = \sum_{i=1}^n \mathbf{X}_i \hat{M}_i(t)$$

(Bagdonavičius&Nikulin,2002)

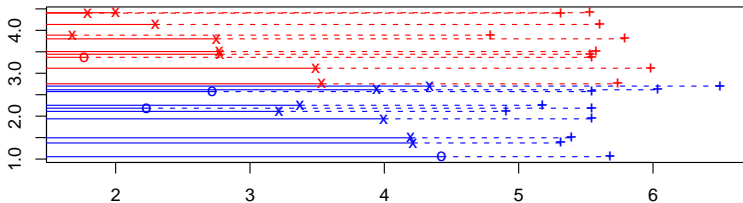
- Test dobré shody pomocí porovnání pozorované hodnoty z dat a hodnot replikovaných za hypotézy



skórový proces a jeho 50 replikací za hypotézy Coxova modelu

Testy dobré shody pro AFT model

- Spočteme reziduály $r_i = \log T_i + \mathbf{X}_i^T \hat{\beta}$
- Provedeme korekci u cenzorovaných pozorování
- Za hypotézy reziduály asymptoticky nestranné odhady ϵ_i se stejnou střední hodnotou.
- Rozdělíme reziduály do dvou skupin podle hodnot regresorů a testujeme shodu středních hodnot
- Wilcoxonův test a asymptotický t-test



rezidua AFT modelu (+) a $\log(T)$ (x – pozorovaná událost, o – cenzorování)

- Simulovali jsme podle Coxova modelu a testovali AFT a naopak
- Replikováno vždy mnoho výběrů, zkoumána empirická síla (podíl zamítnutých výběrů)
- Různé základní rozdělení, typy regresorů, koeficienty, míra cenzorování...
- Výsledky viz poster

Děkuji za pozornost

Reference

- [1]]Cox D.R.: *Regression models and life tables*, J. Roy. Statist. Soc. Ser. B 34, 187–220, 1972.
- [2]]Buckley J., James I.R.: *Linear regression with censored data*, Biometrika 66, 429–436, 1979.
- [3]]Nikulin M., Bagdonavičius V.: *Accelerated Life Models*, Chapman&Hall, 2002.