

Modèles de durée

Nom : **DE TRUCHIS**

Prénom : **Gilles**

Année : **M2**

Semestre : **9**

Nature : **CM**

Volume horaire : **24**

ECTS / Coef : **4**

Prérequis

- Cours de statistiques (propriétés des estimateurs, théorie des tests)
 - Cours d'économétrie (estimation par le maximum de vraisemblance)
-

Résumé

Ce cours a pour objet la présentation des techniques de modélisation des temps d'événement les plus utilisées. On expose successivement les estimations non paramétriques, et tout particulièrement l'estimateur de Kaplan-Meier de la survie, puis les modélisations semi-paramétriques associées à des choix de distribution de ces durées avec les tests de spécification qui peuvent être réalisés. Enfin, on traite du modèle de COX qui constitue la technique de modélisation semi-paramétrique la plus populaire du risque de survenue d'un événement.

Objectifs

- Mener et interpréter une estimation des fonctions de survie par Kaplan-Meier. Etre capable de repérer d'éventuelles hétérogénéités d'individus au sein d'une population, et de réaliser une première sélection de variables explicatives dans un cadre non paramétrique.
 - Savoir, choisir une hypothèse de distribution adaptée aux données et réaliser une estimation paramétrique de la survie, savoir mener des tests de validation et interpréter les résultats de cette estimation.
 - Etre capable d'estimer et d'interpréter les résultats d'une estimation du modèle de Cox, notamment en termes de ratios de risque, de réaliser des tests de validation du modèle, d'introduire des variables explicatives à valeurs dépendantes des durées et de réaliser si besoin des études stratifiées dans ce cadre semi-paramétrique.
-

Bibliographie

- Manuels et ouvrages d'application :
 - John D. Kalbfleisch & Ross L. Prentice The Statistical Analysis of Failure Time Data (2nd edition), 2002, Wiley.
 - Paul D. Allison, Survival Analysis using SAS: A practical guide, 2010, SAS Institute (2nd edition).
 - Alan Cantor, SAS® Survival analysis techniques for medical research (2nd edition), 2003, SAS Institute.
-

PLAN

Chapitre 1 Introduction

- La nature des données de survie
- La description de la distribution des temps de survie

Chapitre 2 L'approche non paramétrique

Section 2.1 L'estimateur de Kaplan-Meier de la fonction de survie : une présentation heuristique

Section 2.2 Kaplan-Meier comme estimateur du maximum de vraisemblance non paramétrique

Section 2.3 Les principales hypothèses

- L'hypothèse de censure non informative
- L'hypothèse d'homogénéité de la population

Section 2.4 La variance de l'estimateur de Kaplan-Meier

Section 2.5 La construction d'IC sur la survie intervalles ponctuels

- Les intervalles de confiance intervalles ponctuels
- Les bandes de confiance

Section 2.6 L'estimation de la fonction de risque cumulée

Section 2.7 L'estimation kernel du risque instantané

Section 2.8 Comparaison des courbes de survie

- Statistique du LogRank
- Test de Wilcoxon
- Les tests stratifiés
- Tests d'association entre une variable continue et la survie

Section 2.9 Les tables de survie – La méthode actuarielle

Section 2.10 La Procédure LIFETEST

Chapitre 3 L'approche paramétrique

Section 3.1 Les modèles AFT et les modèles PH

- Les modèles à temps de vie accélérée
- Les modèles à risque proportionnels

Section 3.2 Les principales modélisations AFT

- Les distributions extrêmes

Section 3.3 Estimation sous divers types de censure

Section 3.4 Choix d'une distribution et tests de spécification

- Sélection au moyen d'un test LRT
- Les aides graphiques

Section 3.5 Estimation de fractiles sur les durées

Section 3.6 Données censurées à gauche, à droite et par intervalle

- La structuration des données
- Exemple : estimation d'un modèle Tobit via LIFEREG

Section 3.7 La procédure LIFEREG

Chapitre 4 L'approche semi-paramétrique

Section 4.1 Le modèle de Cox et son estimation

- La fonction de vraisemblance partielle
- La correction de Firth en cas de monotonie de PL
- La prise en compte d'événements simultanés
- La spécification de l'équation à estimer : commandes Model et Class

Section 4.2 Les ratios de risque

- Interprétation des coefficients et ratios de risques
- Commandes Hazardratio et Contrast
- Des exemples de sorties

Section 4.3 L'estimation de la survie de base

Section 4.4 L'analyse stratifiée avec le modèle de Cox

Section 4.5 Variables explicatives non constantes dans le temps

- Données entrées selon un processus de comptage
- Explicatives non constantes créées par programme

Section 4.6 Les tests de validation

- La qualité de l'ajustement
- Étude de spécification : résidus de martingale, régression locale et sommes partielles cumulées
- Repérage des outliers : les résidus de déviance, les statistiques DFBETA et LD
- Test de l'hypothèse PH : introduction d'interactions avec le temps, résidus de Schoenfeld

Section 4.7 La sélection automatique des variables explicatives

Chapitre 5 Quelques compléments

Section 5.1 Statistiques complémentaires ou alternatives aux ratios de risque

Section 5.2 L'ajustement du Residual Mean Survival Time en présence de variables explicatives

Section 5.3 Estimation non paramétrique avec censure par intervalle