



APPLIED STATISTICS WORKSHOP

Céline BUGLI
SMCS, UCL

Les statistiques : un atout supplémentaire pour les demandeurs d'emploi ? Histoire du développement d'une formation

Abstract:

Pour la 3ème année consécutive, le SMCS (UCL) et le FOREM s'unissent pour lutter contre le chômage. L'histoire commence en 2010, lorsque le FOREM biotechnologie à Liège dépose un projet 'Bioinformatique' dans le cadre du 6e appel à projets Biowin labellisé par le Gouvernement Wallon. Le projet initial visait à développer un parcours de formation professionnelle et qualifiante en bioinformatique à destination des demandeurs d'emploi porteurs d'un diplôme universitaire (biologie, chimie, biochimie, bioingénieur, médecin, pharmacien,...) dans le but de favoriser leur engagement dans ce nouveau créneau qui se développe et prend de plus en plus d'ampleur au sein des entreprises Biotech et des centres de recherche. Projet accepté à une condition : un comité d'accompagnement du projet devait être mis sur pied et être composé en grande partie d'entreprises du secteur afin de garantir l'adéquation entre le contenu des modules de formation et les besoins des entreprises. Très vite, l'UCL et le SMCS en particulier se sont portés volontaires pour réfléchir au parcours de formation dans un premier temps et pour participer en tant que formateur par la suite.

Depuis lors, la formation a été donnée 3 fois à des groupes d'une dizaine de participants et a beaucoup évolué : d'une formation en bioinformatique avec 20 jours de statistique la formation est devenue une formation exclusivement statistique de 34 jours de formation orientée pratique avec utilisation de logiciels tels que R, SAS et JMP.

Lors de cet atelier, nous expliquerons les enjeux et les challenges du développement d'une telle formation. Nous discuterons également des outils à notre disposition et de la philosophie d'apprentissage que nous avons développé. Enfin, nous évoquerons les possibilités d'évolution et de pérennisation d'une telle formation.

La présentation sera accompagnée d'une séance de discussion et de réflexion

March 06, 2015 - 14:30 - Room c 115



APPLIED STATISTICS WORKSHOP

Christian RITTER
SMCS, UCL

Séance de discussion et réflexion liée à l'exposé de 14:30 sur les thèmes :

Dans quelle mesure, le contenu actuel est-il adéquat ?

Quelles parties pourrait-on utiliser dans d'autres contextes ? :

- a) Ecoles doctorales (à l'image de ce que le consortium FLAMES fait en Flandre)
- b) Services universitaires (médecins, infirmiers, chercheurs)
- c) Entreprises actives dans le domaine bio-pharmaceutique

Y a-t-il des initiatives similaires aux nôtres ailleurs ? Des personnes avec qui collaborer ? (Comme par exemple FLAMES)

Comment devrions-nous gérer la propriété intellectuelle de la formation ? :

Devrait-elle être libre d'accès dans une formule 'creative commons' ?

Quelle institution/public devrait payer pour le développement et l'entretien ?

Ceux qui souhaitent poursuivre la discussion après la fin officielle (17h), se retrouveront à partir de 17:30 à la Crêperie Bretonne

March 06, 2015 - 16:00 - Room c 115



INSTITUT DE STATISTIQUE, BIOSTATISTIQUE ET SCIENCES
ACTUARIELLES

UNIVERSITE CATHOLIQUE DE LOUVAIN (UCL)

UNIVERSITE LIBRE DE BRUXELLES (ULB)

JOINT SEMINAR ISBA/ULB

Statistics Seminar in the framework of ARC "Survival Analysis"

Justin CHOWN
ISBA, UCL

A new approach in testing for heteroskedastic errors of a nonparametric regression

Abstract:

A common problem faced by statistical practitioners is for variation in the errors of a regression model to depend on variables of interest to the study, heteroskedastic errors. We model the regression function by a local polynomial smoother, a nonparametric approach. First, we study a suitable residual-based empirical distribution function that is found to be efficient in the sense of Hajek and Le Cam. Then, we introduce a test for heteroskedasticity based on this estimator. The approach uses a "detection function" that makes it possible to verify the existence of heteroskedasticity in data by exploiting the mere independence-dependence structure between the detection function and model errors; i.e. we do not require a specific model of the scale function, if it exists. Requirements on the detection function are suitable for practitioners to adapt our test to their specific needs. This procedure is asymptotically distribution free, meaning inferences made from it do not depend on unknown parameters, and it is consistent at the so-called parametric rate of convergence. Our results are extended to the case of missing data.

Friday, March 13, 2015 - 14:30

**ULB, Campus de la Plaine – Bâtiment NO – 9e étage,
Salle des Professeurs**



INSTITUT DE STATISTIQUE, BIOSTATISTIQUE ET SCIENCES
ACTUARIELLES

UNIVERSITE CATHOLIQUE DE LOUVAIN (UCL)

UNIVERSITE LIBRE DE BRUXELLES (ULB)

JOINT SEMINAR ISBA/ULB

Christine DE MOL
ULB

Sparsity for Big Data

Abstract:

The aim of this talk is to provide a general overview of the mathematical, statistical and computational challenges arising from the necessity to cope with the data deluge. We show how the concept of sparsity can be useful in addressing some of these challenges.

Friday, March 13, 2015 - 16:00

**ULB, Campus de la Plaine – Bâtiment NO – 9e étage,
Salle des Professeurs**



Statistics Seminar in the framework of ARC "Survival Analysis"

Samuel Maistre

Université Claude Bernard Lyon 1, France

A general approach to goodness-of-fit tests for regression models: application to single-index hypotheses and functional data models and bootstrap considerations

Abstract

In this talk we present a general approach to test a wide variety of regression modeling situations. This approach, based on kernel smoothing in low dimension, leads to asymptotical standard gaussian critical values. We apply it to two quite different situations : test for the goodness-of-fit of models involving functional variables for the response and/or for the covariates and to test a single-index assumptions. We also propose bootstrap corrections for small samples.

Wednesday, March 18, 2015 - 13:00 - Room c 115



Statistics Seminar in the framework of ARC "Survival Analysis"

Ursula Müller

Department of Statistics

Texas A&M University, College Station, TX, USA

Efficient estimation in regression with missing responses

Abstract

I will first review some of my results on efficient estimation in semiparametric regression with responses Y "missing at random" based on imputation. Then I will demonstrate that characteristics of the conditional distribution of Y given a covariate vector X can be estimated efficiently using complete case analysis. This means that one can simply omit incomplete cases and work with an appropriate efficient estimator without losing this desirable property. The "efficiency transfer" is a general result and holds true for all regression models for which the distribution of Y given X and the marginal distribution of X do not share common parameters. The derivation uses the "transfer principle" for obtaining limiting distributions of complete case statistics (for general missing data models) from corresponding results in the complete data model. For an illustration we consider estimation of regression parameters and of functionals of the error distribution.

Friday, March 20, 2015 - 11:30 - Room c 115