ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ
ATHENS UNIVERSITY OF ECONOMICS AND BUSINESS

ΣΧΟΛΗ ΕΠΙΣΤΗΜΩΝ & ΤΕΧΝΟΛΟΓΙΑΣ ΤΗΣ ΠΛΗΡΟΦΟΡΙΑΣ
SCHOOL OF INFORMATION SCIENCES & TECHNOLOGY

ΤΜΗΜΑ ΣΤΑΤΙΣΤΙΚΗΣ
DEPARTMENT OF STATISTICS

ΚΥΚΛΟΣ ΣΕΜΙΝΑΡΙΩΝ ΣΤΑΤΙΣΤΙΚΗΣ - ΑΠΡΙΛΙΟΣ 2015

Χαράλαμπος Χανιαλίδης

School of Mathematical Sciences, University College Dublin, Ireland.

Flexible regression models for count data

ΠΑΡΑΣΚΕΥΗ 3/4/2015
13:00 – 15:00

ΑΙΘΟΥΣΑ 607, 6ος ΟΡΟΦΟΣ,
ΚΤΙΡΙΟ ΜΕΤΑΠΤΥΧΙΑΚΩΝ ΣΠΟΥΔΩΝ
(ΕΥΕΛΠΙΔΩΝ & ΛΕΥΚΑΔΟΣ)

ΠΕΡΙΛΗΨΗ (ΣΤΑ ΑΓΓΛΙΚΑ)

Count data is a common type of data found in observational and epidemiological studies. Counts represent the number of occurrences of an event within a fixed period, e.g. number of goals in a football game, number of emergency hospital admissions, etc. A natural choice for modelling this type of data is the Poisson distribution. The default model for count data is the Poisson regression model in which one models the mean of the counts as a log-linear function of the covariates. In the presence of overdispersion (having variance larger than the mean) the Poisson regression model becomes limited due to its assumption of equidispersion (the mean has to be equal to the variance). The most commonly used model for treating overdispersed data is the negative binomial regression model which allows the variance to be larger than the mean.

In this talk i) we will focus on an even more flexible regression model for count data, the COM-Poisson regression model. This model allows modelling the mean and the variance explicitly, thus identifying covariates with different effects on the mean and the variance of the counts. It is flexible enough to handle underdispersion, something that neither of the previous models can do. The main reason why this model is not as widely used is that the normalisation constant of the COM-Poisson distribution, with respect to the observed data, does not have a closed form.

ii) we will propose two exact MCMC algorithms which address this problem; and

iii) present a much more flexible Bayesian density regression model for count data which can be thought of as an extension to the "simple" COM-Poisson regression model.

ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

ATHENS UNIVERSITY OF ECONOMICS AND BUSINESS

ΣΧΟΛΗ ΕΠΙΣΤΗΜΩΝ & ΤΕΧΝΟΛΟΓΙΑΣ ΤΗΣ ΠΛΗΡΟΦΟΡΙΑΣ
SCHOOL OF INFORMATION SCIENCES & TECHNOLOGY

ΤΜΗΜΑ ΣΤΑΤΙΣΤΙΚΗΣ
DEPARTMENT OF STATISTICS

AUEB STATISTICS SEMINAR SERIES - APRIL 2015

Charalampos Chanialidis

School of Mathematical Sciences, University College Dublin, Ireland.

Flexible regression models for count data

FRIDAY 3/4/2015
13:00 – 15:00

ROOM 607, 6th FLOOR,
POSTGRADUATE STUDIES BUILDING
(EVELPIDON & LEFKADOS)

ABSTRACT

Count data is a common type of data found in observational and epidemiological studies. Counts represent the number of occurrences of an event within a fixed period, e.g. number of goals in a football game, number of emergency hospital admissions, etc. A natural choice for modelling this type of data is the Poisson distribution. The default model for count data is the Poisson regression model in which one models the mean of the counts as a log-linear function of the covariates. In the presence of overdispersion (having variance larger than the mean) the Poisson regression model becomes limited due to its assumption of equidispersion (the mean has to be equal to the variance). The most commonly used model for treating overdispersed data is the negative binomial regression model which allows the variance to be larger than the mean.

In this talk i) we will focus on an even more flexible regression model for count data, the COM-Poisson regression model. This model allows modelling the mean and the variance explicitly, thus identifying covariates with different effects on the mean and the variance of the counts. It is flexible enough to handle underdispersion, something that neither of the previous models can do. The main reason why this model is not as widely used is that the normalisation constant of the COM-Poisson distribution, with respect to the observed data, does not have a closed form.

ii) we will propose two exact MCMC algorithms which address this problem; and

iii) present a much more flexible Bayesian density regression model for count data which can be thought of as an extension to the "simple" COM-Poisson regression model.