ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

ATHENS UNIVERSITY OF ECONOMICS AND BUSINESS EXOAH ETIETHMON & TEXNOAOTIAE THE THEPOOPOPIAE SCHOOL OF INFORMATION SCIENCES & TECHNOLOGY

TMHMA ΣΤΑΤΙΣΤΙΚΗΣ DEPARTMENT OF STATISTICS

ΚΥΚΛΟΣ ΣΕΜΙΝΑΡΙΩΝ ΣΤΑΤΙΣΤΙΚΗΣ ΙΑΝΟΥΑΡΙΟΣ 2016

Luca Tardella

Dipartimento di Scienze Statistiche – Sapienza Universita di Roma (Italy)

A novel Bayesian mixture of Poisson estimation for expressed sequence tag sampling

TETAPTH 27/1/2016 13:00

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

ΠΕΡΙΛΗΨΗ

Expressed sequence tag (EST) clustering is a process that identifies and assembles ESTs which correspond to the same gene. A cDNA library can be sampled and the number of tags observed from each distinct gene are often summarized in terms of a gene cluster profile $(n_1, ..., n_j, ..., n_T)$ where n_j represents the number of genes that had j ESTs in the sample. This profile has been used for gene capture prediction and overlap estimation in EST sequencing based on the appropriate modelling of the so-called frequency of frequencies (Susko, E. and Roger, A.J., 2004; Wang et al. 2005). We focus on the estimation of the unobserved expressed genes and we propose a novel moment-based flexible Bayesian approach comparing the new estimator with those already available and we discuss computational and inferential challenges.

ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

ATHENS UNIVERSITY OF ECONOMICS AND BUSINESS EXOAH ERIETHMON & TEXNOAORIAS THE TAHPOФOPIAS SCHOOL OF INFORMATION SCIENCES & TFCHNOLOGY

TMHMA ΣΤΑΤΙΣΤΙΚΗΣ DEPARTMENT OF STATISTICS

AUEB STATISTICS SEMINAR SERIES JANUARY 2016

Luca Tardella

Dipartimento di Scienze Statistiche – Sapienza Universita di Roma (Italy)

A novel Bayesian mixture of Poisson estimation for expressed sequence tag sampling

Wednesday 27/1/2016 15:00 - 17:00

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

ABSTRACT

Expressed sequence tag (EST) clustering is a process that identifies and assembles ESTs which correspond to the same gene. A cDNA library can be sampled and the number of tags observed from each distinct gene are often summarized in terms of a gene cluster profile $(n_1, ..., n_j, ..., n_T)$ where n_j represents the number of genes that had j ESTs in the sample. This profile has been used for gene capture prediction and overlap estimation in EST sequencing based on the appropriate modelling of the so-called frequency of frequencies (Susko, E. and Roger, A.J., 2004; Wang et al. 2005). We focus on the estimation of the unobserved expressed genes and we propose a novel moment-based flexible Bayesian approach comparing the new estimator with those already available and we discuss computational and inferential challenges.