

ΠΑΝΕΠΙΣΤΗΜΙΟ ΚΡΗΤΗΣ  
ΤΜΗΜΑ ΜΑΘΗΜΑΤΙΚΩΝ ΚΑΙ ΕΦΑΡΜΟΣΜΕΝΩΝ ΜΑΘΗΜΑΤΙΚΩΝ

ΟΜΙΛΙΕΣ ΣΤΑΤΙΣΤΙΚΗΣ ΚΑΙ ΠΙΘΑΝΟΤΗΤΩΝ

15:15, Παρασκευή, 28 Απριλίου 2017  
Αίθουσα Α-303

**Χριστίνα Παρπούλα**, Τμήμα Στατιστικής και Ασφαλιστικής Επιστήμης,  
Πανεπιστήμιο Πειραιώς

*Αλγοριθμική Προσέγγιση στην Ανάλυση Υπερκορεσμένων Σχεδιασμών με  
Μέτρα της Θεωρίας Πληροφορίας*

Στην εργασία αυτή παρουσιάζεται η γενίκευση του προβλήματος ανάλυσης υπερκορεσμένων σχεδιασμών κάτω από την παραδοχή των γενικευμένων γραμμικών μοντέλων. Συγκεκριμένα αναπτύσσεται ένας συνεκτικός αλγόριθμος της Θεωρίας Πληροφορίας (Information Theoretical Algorithm-ITA) για την αναζήτηση και εύρεση των σημαντικών παραγόντων, αναλύοντας δεδομένα από υπερκορεσμένους σχεδιασμούς και θεωρώντας γενικευμένα γραμμικά μοντέλα των οποίων τα δεδομένα απόκρισης προέρχονται από την Bernoulli, Poisson, γεωμετρική και αρνητική διωνυμική κατανομή. Η συγκριτική απόδοση του προτεινόμενου αλγορίθμου, με βάση την εκτενή μελέτη προσομοίωσης που διεξήχθη, υποδεικνύει ότι η προτεινόμενη μέθοδος αποδίδει πολύ καλά παρουσιάζοντας χαμηλά ποσοστά για τα δύο υπό εκτίμηση σφάλματα (Τύπου I και Τύπου II) και μπορεί να θεωρηθεί εύρωστη από τη στιγμή που είτε με τη χρήση διαφορετικών μέτρων εντροπίας (Shannon, Rényi, Tsallis και Havrda-Charvát εντροπία) είτε με τη χρήση διαφορετικών τιμών κατωφλιού για το κριτήριο της συμμετρικής αβεβαιότητας προέκυψαν παρόμοια αποτελέσματα. Αξίζει να σημειωθεί ότι η προτεινόμενη μέθοδος υπερτερεί δύο γνωστών μεθόδων επιλογής μεταβλητών, των αλγορίθμων Conditional Mutual Information Maximization (CMIM) και minimal-redundancy-maximal-relevance (mRMR) κάτω από την υπόθεση ενός λογιστικού μοντέλου παλινδρόμησης.

Στην ανάλυση υπερκορεσμένων σχεδιασμών δεν είναι εφικτό να αναγνωρίζονται οι σημαντικοί παράγοντες και να εκτιμώνται με αμελητέο σφάλμα, ένα πρόβλημα το οποίο απορρέει εξαιτίας της μη ορθογωνιότητας του πίνακα σχεδιασμού η οποία καθιστά αναπόφευκτο το να γίνεται ανεκτή η ύπαρξη μιας μικρής μεροληψίας ανάμεσα σε όλες τις εκτιμημένες επιδράσεις. Βάσει των παραπάνω, είναι αξιοσημείωτο το γεγονός ότι η προτεινόμενη μέθοδος επιτυγχάνει να ανιχνεύσει όλες τις πραγματικά ενεργές επιδράσεις με σχεδόν αμελητέο σφάλμα, και πόσο μάλλον κάτω από την υπόθεση ενός γενικευμένου γραμμικού μοντέλου, μέθοδος η οποία είναι μη τετριμμένη και μη αναπτυγμένη στην έως σήμερα διεθνή βιβλιογραφία όπου συναντάται το πρόβλημα ανάλυσης των υπερκορεσμένων σχεδιασμών μόνο κάτω από την υπόθεση του κλασικού γραμμικού μοντέλου.