



ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

Πατησίων 76, 10434 Αθήνα - τηλ: 210-8203911, fax: 210-8226204
url: www.aueb.gr

ΤΜΗΜΑ ΣΤΑΤΙΣΤΙΚΗΣ

Αντωνιάδου 1-3, τηλ: 210-8203111-3, url: www.stat-athens.aueb.gr

Καθηγητής Ι. Πανάρετος

τηλ./fax: 210-8210509

e-mail: jpan@aub.gr

url: <http://www.panaretos.eu>

29 Μαρτίου 2013

Προς: την ΓΣΕΣ του Τμήματος Στατιστικής

Κοιν.: Γραμματεία Τμήματος Στατιστικής

Η Δημιουργία Σχολών στο Οικονομικό Πανεπιστήμιο Αθηνών

Με αφορμή τη συζήτηση και τους προβληματισμούς που έχουν αναπτυχθεί για την δημιουργία Σχολών στο Οικονομικό Πανεπιστήμιο Αθηνών θα ήθελα να παρατηρήσω τα εξής:

Το Οικονομικό Πανεπιστήμιο Αθηνών είναι ένα μικρό ίδρυμα με σαφώς προδιαγεγραμμένο αντικείμενο. Δεδομένου ότι έχει αλλάξει πλήρως ο ρόλος της Σχολής σε σχέση με το αρχικό σχέδιο νόμου, η δημιουργία Σχολών απλώς θα δημιουργήσει νέες γραφειοκρατικές δομές χωρίς να προσθέσει τίποτα το ουσιαστικό.

Στην περίπτωση όμως που επιλεγεί η δημιουργία Σχολών, θεωρώ ότι τα υφιστάμενα τμήματα Στατιστικής και Εφαρμοσμένης Πληροφορικής θα έπρεπε να ανήκουν στην ίδια Σχολή.

Ένα άλλο θέμα που πάντοτε προκαλεί προβληματισμούς και αντιθέσεις είναι η ονομασία. Αντιλαμβάνομαι ότι το Υπουργείο Παιδείας στην πρόταση του πράγματι τοποθετεί τα τμήματα Στατιστικής και Πληροφορικής σε μία Σχολή με τίτλο **Σχολή Θετικών Επιστημών**. Στο θέμα του τίτλου θα επανέλθω αργότερα.

Μια ιστορική αναδρομή

Όπως είναι γνωστό μέχρι το 1984 στην ΑΣΟΕΕ υπήρχαν δύο προγράμματα σπουδών, Οικονομικών και Οργάνωσης και Διοίκησης των Επιχειρήσεων. Οι τρεις καθηγητές Στατιστικής που κατείχαν αντίστοιχες έδρες, κάλυπταν τις ανάγκες της Στατιστικής για τα συγκεκριμένα προγράμματα. Η Στατιστική στην ΑΣΟΕΕ τότε εξελαμβάνετο απλώς ως εργαλείο στις επιστήμες της Οικονομίας και της Διοίκησης. Το 1984 δημιουργήθηκαν τρία τμήματα. Τα δύο ήταν η μετεξέλιξη των προαναφερθέντων προγραμμάτων σπουδών σε αυτοτελή τμήματα και το τρίτο ήταν το τμήμα Στατιστικής και Πληροφορικής. Αυτό ήταν συνέπεια της αναγνώρισης ότι το Ίδρυμα έπρεπε να δώσει έμφαση στα δύο αυτά επιστημονικά αντικείμενα αυτοτελώς. Στο τμήμα αυτό ενετάγησαν οι υπηρετούντες καθηγητές Στατιστικής. Επίσης ενετάγησαν σε αρκετοί καθηγητές οικονομετρίας (κάτι που δυσκόλεψε για κάποιο διάστημα το τμήμα να καθορίσει την φυσιογνωμία του).

Το 1989 το τμήμα αυτό χωρίστηκε σε δύο τμήματα, τμήμα Στατιστικής και τμήμα Εφαρμοσμένης Πληροφορικής. Το ίδιο έγινε και με άλλα τμήματα του Ιδρύματος. (Οι καθηγητές Οικονομετρίας του κοινού Τμήματος ενετάγησαν στο Τμήμα Εφαρμοσμένης Πληροφορικής, κάτι που επιβεβαίωσε την πλήρη διαφοροποίηση του Τμήματος Στατιστικής από τα Οικονομικά).

Με τα δεδομένα της εποχής, η απόφαση αυτή ήταν σωστή. Όπως είναι γνωστό, επί δεκαετίες και μέχρι το τέλος του προηγούμενου αιώνα οι επιστήμες ήταν άκρως εξειδικευμένες. Η ανάπτυξή τους εξαρτιόταν από την αυτοτελή βελτίωση και εξειδίκευση των επιστημονικών δραστηριοτήτων σε κάθε αντικείμενο.

Το τμήμα Στατιστικής, μοναδικό μέχρι σήμερα στην χώρα με αντικείμενο αποκλειστικά την Στατιστική, υιοθέτησε την δεκαετία του '90 μια στρατηγική κάλυψης όλων των αντικειμένων που είναι σχετικά με την Στατιστική. Δεν έκανε καμιά πρόσληψη μέλους ΔΕΠ που το γνωστικό του αντικείμενο να συνδέεται -έστω και οριακά- με τα Οικονομικά ή με οποιαδήποτε άλλη επιστήμη που θεραπευόταν από άλλο Τμήμα του Ιδρύματος (σύμφωνα με απόφαση που είχε ληφθεί τότε και δυστυχώς δεν εφαρμόστηκε από τα υπόλοιπα Τμήματα). Καταξιώθηκε διεθνώς, τόσο λόγω του επιστημονικού του προσωπικού αλλά κυρίως λόγω των πτυχιούχων του, πολλοί από τους οποίους έγιναν δεκτοί για μεταπτυχιακές σπουδές στα καλύτερα πανεπιστήμια του εξωτερικού. Μερικοί από αυτούς μάλιστα έγιναν δεκτοί απ' ευθείας για διδακτορικό λόγω του εξαιρετικού προπτυχιακού προγράμματος που είχαν παρακολουθήσει. Ο αριθμός των πτυχιούχων της εποχής αυτής που στελεχώνουν σήμερα Αμερικανικά και Ευρωπαϊκά Πανεπιστήμια είναι εντυπωσιακός με οποιαδήποτε στάνταρντς.

Η ενδεχόμενη ένταξη επομένως του Τμήματος Στατιστικής σε μια Σχολή Οικονομικών Επιστημών θα αποτελούσε ένα ανεπίτρεπτο περιορισμό του ρόλου της Στατιστικής στο πανεπιστήμιό μας και επιστροφή στα δεδομένα της δεκαετίας του 70.

Η συνάφεια Στατιστικής και η Πληροφορικής

Διεθνώς, στα τέλη της δεκαετίας του 80 και η Στατιστική βρισκόταν σε ένα πεδίο αναζητήσεων. Από τα μέσα της δεκαετίας του '50 και για μεγάλο διάστημα η Στατιστική έδωσε μεγάλη έμφαση στα Μαθηματικά, ενώ αργότερα στην δεκαετία του '80 κινήθηκε περισσότερο προς την υπολογιστική κατεύθυνση και αργότερα προς τις εφαρμογές. Ένας από τους κορυφαίους Στατιστικούς του 20^{ου} αιώνα που συνέβαλε σε μεγάλο βαθμό στον καθορισμό των τάσεων της επιστήμης, ο καθηγητής του Stanford Brad Efron, είχε επισημάνει ότι η Στατιστική θα πρέπει να συνεργάζεται εξ ίσου αποδοτικά με τα Μαθηματικά και τις επιστήμες των υπολογιστών προκειμένου να αξιοποιεί στο μέγιστο δυνατόν τις δυνατότητές της και να προσφέρει το μέγιστο δυνατόν στις άλλες επιστήμες. (<http://stat-athens.aueb.gr/~jpan/volume-Panaretos/Efron.pdf>).

Η έρευνα

Σήμερα η Στατιστική και η επιστήμη των υπολογιστών (αυτός είναι ο όρος που έχει επικρατήσει και όχι ο αδόκιμος όρος Πληροφορική), έχουν πολύ περισσότερα κοινά ενδιαφέροντα. Παραδοσιακά οι δύο αυτές επιστήμες είχαν μία σημαντική επικάλυψη, αλλά την τελευταία δεκαετία η προσέγγιση αυτή έχει γίνει πλησιέστερη από ποτέ. Αρκεί να αναφέρει κανείς την περιοχή του Machine Learning (Artificial Intelligence είναι ένας εναλλακτικός όρος που χρησιμοποιείται συχνά). Πρόκειται ουσιαστικά για έναν κλάδο της επιστήμης των υπολογιστών που μελετά στατιστικά προβλήματα. Η διαφορά βρίσκεται στην προσέγγιση. Η προσέγγιση του Machine Learning από τους επιστήμονες της Πληροφορικής είναι περισσότερο αλγοριθμική με στόχο την βελτιστοποίηση ενός κριτηρίου όπως για παράδειγμα ένα mean prediction error. Με αυτή την έννοια μοιάζει περισσότερο με μια προσέγγιση μαύρου κουτιού (black box approach). Για παράδειγμα δεν ενδιαφέρει να κατανοήσουμε πώς λειτουργούν τα χρηματιστήρια, αλλά ενδιαφέρει να προβλέψουμε τις διακυμάνσεις τους προκειμένου να βελτιώσουμε την οικονομική απόδοση. Η Στατιστική από το άλλο μέρος, δίνει έμφαση στα μοντέλα με στόχο να προσδιοριστούν και να καθοριστούν συγκεκριμένες κρίσιμες παράμετροι του υπό μελέτη προβλήματος, για τις οποίες υπάρχει η πεποίθηση ότι διαφωτίζουν το φαινόμενο που γίνεται προσπάθεια να μοντελοποιηθεί. Παράδειγμα αποτελεί το πρόβλημα της εκτίμησης μίας παραμέτρου β σε μια παλινδρόμηση, σε αντίθεση με την βέλτιστη πρόβλεψη της εξαρτημένης μεταβλητής Y . Υπάρχουν και πολλά άλλα παραδείγματα όπου οι επιστήμονες της Πληροφορικής χρησιμοποιούν διαφορετική ορολογία από ότι οι Στατιστικοί. Π.χ. support vector machines, neural networks κτλ.

Τα δύο αντικείμενα της Στατιστικής και της επιστήμης των υπολογιστών πρωταγωνιστούν στην εποχή μας σε μια από τις πλέον ενδιαφέρουσες ερευνητικές δραστηριότητες, αυτή που ονομάζεται «Επανάσταση των Μεγάλων Δεδομένων» (Big Data Revolution). Πρόκειται για τεράστια σετ δεδομένων όπου ο αριθμός των

παραμέτρων είναι της ίδιας (ή και μεγαλύτερης) τάξης μεγέθους από τον αριθμό των παρατηρήσεων. Στην επιστήμη των υπολογιστών οι επιστήμονες ενδιαφέρονται να αναπτύξουν αποτελεσματικούς αλγορίθμους που διαμορφώνονται κατάλληλα με το διευρυμένο μέγεθος των δεδομένων, ενώ οι Στατιστικοί ενδιαφέρονται για την κατανόηση του πώς οι παραδοσιακές στατιστικές μέθοδοι θα πρέπει να διαφοροποιηθούν για να δώσουν ικανοποιητικές απαντήσεις στο συγκεκριμένο πλαίσιο. Είναι προφανές ότι οι Στατιστικοί μπορούν να προσφέρουν πολλά και χρήσιμα πράγματα στους αλγορίθμους των επιστημόνων της Πληροφορικής.

Την ίδια ώρα και οι δύο επιστήμες ενδιαφέρονται για προβλήματα παρόμοιας φύσης τα οποία όμως έχουν διαφορετικό όνομα στις δύο επιστημονικές περιοχές: π.χ. επιλογή μεταβλητών (variable selection) στην Στατιστική ή compressed sensing (Πληροφορική). Οι Στατιστικοί ενδιαφέρονται να συμπεράνουν ποιες μεταβλητές παίζουν ρόλο μεταξύ ενός μεγάλου αριθμού πιθανών υποψηφίων μεταβλητών σε μία παλινδρόμηση ενώ οι επιστήμονες της Πληροφορικής προσπαθούν να συμπιέσουν πολύ μεγάλα αντικείμενα (very large objects) με μερικές μόνο γραμμικές μετρήσεις. Μέθοδοι όπως η LASSO (L1 regularization) βρίσκονται στην πρώτη γραμμή των προσπαθειών αυτών.

Ένα άλλο κλασικό σημείο επαφής είναι αυτό της βελτιστοποίησης (optimization). Ειδικά υπό το φως της Επανάστασης των Μεγάλων Δεδομένων (Big Data Revolution) τεχνικές convex optimization είναι κεντρικής σημασίας και στις δύο επιστημονικές περιοχές. Αυτή είναι μία περιοχή όπου οι επιστήμονες της Πληροφορικής δραστηριοποιούνται για πολύ μεγαλύτερο χρόνο από ότι οι επιστήμονες της Στατιστικής και οι Στατιστικοί έχουν πολλά να ωφεληθούν από μια τέτοια συνεργασία.

Μια άλλη δραστηριότητα στις εφαρμογές είναι αυτή στα graphics & visualization, μια περιοχή όπου και οι δύο επιστήμες έχουν πολλά να προσφέρουν.

Ακόμα και σε ένα περισσότερο θεωρητικό πλαίσιο η Μαθηματική Στατιστική (π.χ. Θεωρία Αποφάσεων) έχει πολλές και βαθιές σχέσεις με την Θεωρία της Πληροφορίας (Information Theory) – και οι δύο χρησιμοποιούν την έννοια της εντροπίας με ένα θεμελιώδη τρόπο. Άλλα θέματα που εμπλέκονται σε αυτό είναι complexity κτλ. Η ανάπτυξη ορισμένων κεντρικών εννοιών στις εμπειρικές ανελίξεις (empirical processes), περιοχή κρίσιμη για την ασυμπτωτική θεωρητική μελέτη στατικών μεθόδων, έγινε μάλιστα από θεωρητικούς που δραστηριοποιούνταν σε machine learning και artificial intelligence (για παράδειγμα οι κλάσεις Vapnik-Chervonenkis).

Συμπερασματικά θα πρέπει να λεχθεί ότι η επιστήμη των υπολογιστών κινείται όλο και περισσότερο στην κατεύθυνση των τυχαιοποιημένων αλγορίθμων και της κωδικοποίησης (randomized algorithms encoding) και για τον λόγο αυτό οι επιστήμονες της Πληροφορικής μπορούν να ωφεληθούν πολλά όταν είναι σε συνεργασία με επιστήμονες που ασχολούνται με την Θεωρία Πιθανοτήτων.

Τέλος, θα πρέπει να αναφερθεί ότι υπάρχουν συνέργειες της Πληροφορικής τόσο με την κλασική όσο και με την Μπεϋζιανή Στατιστική. Το compressed sensing και το variable selection ξεκίνησαν περισσότερο από την κλασική Στατιστική θεώρηση, όμως δεν θα ήταν υπερβολή να πούμε ότι οι επιστήμονες της πληροφορικής χρησιμοποιούν σε τεράστιο βαθμό την θεωρία του Bayes (σε προβλήματα όπως επεξεργασία φυσικής γλώσσας (natural language processing), αναγνώριση αντικειμένων (object recognition) ή γενικότερα στο πρόβλημα της «όρασης υπολογιστών» (computer vision)).

Η Εκπαίδευση

Υπάρχει όμως και κάτι το πολύ σημαντικό που παρατηρείται κυρίως στα αμερικανικά πανεπιστήμια και αφορά την εκπαίδευση. Οι περισσότεροι επιστήμονες που σπουδάζουν για την απόκτηση διδακτορικού στις επιστήμες των υπολογιστών στα μεγαλύτερα αμερικανικά πανεπιστήμια σήμερα, επιλέγουν παράλληλα να εκπονήσουν ένα Master στην Στατιστική. Την ίδια στιγμή πολλοί φοιτητές που κάνουν διδακτορικό στην Στατιστική παίρνουν πολλά μαθήματα σε Computer Science. Αυτό συμβαίνει κυρίως γιατί, όσον αφορά την αγορά εργασίας, είναι εξαιρετικά χρήσιμο να έχει κανείς γνώσεις Στατιστικής και υπολογιστών. **Με δεδομένη την σημασία της Σχολής σε επίπεδο μεταπτυχιακών σπουδών με τον νέο νόμο**, η συνύπαρξη των δύο αυτών επιστημών σε μια Σχολή θα είναι ακόμα περισσότερο ουσιαστική.

Δεν είναι εξ άλλου τυχαίο το τι συμβαίνει στα δύο μεγαλύτερα και γνωστότερα στον κόσμο τμήματα Στατιστικής, αυτό του Berkeley και αυτό του Stanford. Δύο από τους σημαντικότερους καθηγητές στο πανεπιστήμιο του Berkeley είναι σήμερα ο Michael Jordan και ο Martin Wainwright. Και οι δύο έχουν βασικές σπουδές στην Πληροφορική με μεταπτυχιακά στην Πληροφορική και την Στατιστική. Το κυριότερο όμως είναι ότι και οι δύο έχουν joint appointments στα τμήματα Στατιστικής και Computer Science. Το ίδιο συμβαίνει με το τμήμα Στατιστικής του Stanford, όπου ο Emmanuel Candes (πού πρόσφατα μετακινήθηκε εκεί από το CALTECH) και ο David Donoho έχουν παράλληλες δραστηριότητες σε Στατιστική και Πληροφορική.

Το όνομα της Σχολής

Η ονομασία **Σχολή Τεχνολογίας και Θετικών Επιστημών** δεν είναι κατάλληλη. Είναι αδόκιμο να υπάρχει στον τίτλο μιας Σχολής όρος (Τεχνολογία) που υπάρχει στον τίτλο Τμήματος (ΔΕΤ) που θα ανήκει μάλιστα σε άλλη Σχολή και να δημιουργεί ψευδείς εντυπώσεις για το εύρος που καλύπτει.

Η πρόταση του Υπουργείου Παιδείας (**Σχολή Θετικών Επιστημών**) εκφράζει με καλύτερο τρόπο τις συνέργειες των δύο αυτών επιστημών. Έχει όμως και αυτή το μειονέκτημα ότι δημιουργεί εντυπώσεις για το εύρος των γνωστικών αντικειμένων που καλύπτει. Ο αντίστοιχος αγγλικός όρος είναι School of Basic Science. Μια

τέτοια Σχολή όμως στο εξωτερικό καλύπτει ένα πολύ ευρύτερο φάσμα επιστημών (μαθηματικά, φυσική, χημεία, βιολογία) που προφανώς δεν θα καλύπτει αυτή η Σχολή.

Με δεδομένο ότι καθεμιά από τις επιστήμες που θα αποτελούν την νέα Σχολή έχουν πλέον διαμορφώσει πλήρως την επιστημονική φυσιογνωμία τους διεθνώς και προσδιορίζονται πλήρως από τους τίτλους των τμημάτων που θα την αποτελούν, η καλύτερη ονομασία που θα είναι συνεπής και με τις διεθνείς εξελίξεις είναι **Σχολή Στατιστικής και Επιστήμης των Υπολογιστών**. (School of Statistics and Computer Science). (Ένα τέτοιο όνομα καλύπτει και όσους ασχολούνται με την τεχνολογία των υπολογιστών). Ένα ενδιαφέρον παράδειγμα, την εξέλιξη του οποίου είχα την δυνατότητα να παρακολουθήσω, είναι το School of Computer Science and Statistics του Trinity College, Dublin (<https://www.scss.tcd.ie/welcome/index.php>).

Ιωάννης Πανάρετος

Σημείωση: Με την παράκληση να διαβιβαστεί στην Σύγκλητο του Πανεπιστημίου, το Τμήμα Εφαρμοσμένης Πληροφορικής και το Υπουργείο Παιδείας, μαζί με οποιεσδήποτε άλλες έγγραφες απόψεις έχουν διατυπωθεί για το θέμα.