

Στατιστική Μεθοδολογία και Αναλυτική Δεδομένων (Statistical Methods for Data Analytics)

ΔΙΔΑΣΚΟΝΤΕΣ: Ι.ΒΡΟΝΤΟΣ – ΣΤ.ΨΑΡΑΚΗΣ

Γενικά Στοιχεία Μαθήματος

Κωδικός: 62102

Τύπος: Υποχρεωτικό Κορμού

Επίπεδο: Μεταπτυχιακό

Έτος σπουδών: Α'

Εξάμηνο σπουδών: 1^ο

ECTS: 5

Γλώσσα διδασκαλίας: Ελληνική

Περιεχόμενο Μαθήματος

Το μάθημα αυτό αποτελείται από δύο μέρη. Στο πρώτο μέρος παρουσιάζεται η βασική θεωρία των διαστημάτων εμπιστοσύνης και των ελέγχων υποθέσεων ενώ στο δεύτερο μέρος γίνεται παρουσίαση της βασικής θεωρίας των στατιστικών υποδειγμάτων.

Ειδικότερα στο πρώτο μέρος γίνεται ανάλυση της βασικής θεωρίας των διαστημάτων εμπιστοσύνης για μέσες τιμές, αναλογίες και διακυμάνσεις στην περίπτωση ενός και δύο πληθυσμών στην περίπτωση κανονικών αλλά και μη-κανονικών δεδομένων. Στη συνέχεια αναλύεται η βασική θεωρία των στατιστικών ελέγχων υποθέσεων για παραμέτρους πληθυσμών όπως μέσες τιμές, αναλογίες, διακυμάνσεις, σύγκριση παραμέτρων σε δύο πληθυσμούς, επίπεδο στατιστικής σημαντικότητας, το παρατηρούμενο επίπεδο στατιστικής σημαντικότητας (β -τιμή), η ισχύς ενός στατιστικού ελέγχου, και ο καθορισμός μεγέθους δείγματος.

Στο δεύτερο μέρος το μάθημα εισάγει και παρουσιάζει τη θεμελιώδη θεωρία στατιστικών υποδειγμάτων, μεθόδων και τεχνικών, οι οποίες είναι απαραίτητες στην εμπειρική ανάλυση δεδομένων. Παρουσιάζεται η θεωρία των υποδειγμάτων παλινδρόμησης, απλής και πολλαπλής γραμμικής παλινδρόμησης. Εξετάζονται θέματα όπως η επιλογή μεταβλητών/μοντέλων, η χρήση ψευδομεταβλητών, και η πολυσυγγραμμικότητα. Δίνεται έμφαση στην εφαρμογή της θεωρίας, εξετάζοντας τις υποθέσεις των καταλοίπων με τη χρήση διαγνωστικών ελέγχων, και παρουσιάζεται η ερμηνεία των αποτελεσμάτων. Παρουσιάζεται και αναπτύσσεται η ανάλυση διακύμανσης καθώς και εμπειρικές εφαρμογές με το πακέτο R.

Προαπαιτούμενα

Εφαρμοσμένες Πιθανότητες

Επιδιωκόμενα Μαθησιακά Αποτελέσματα

Στόχος του μαθήματος είναι να παρέχει στους φοιτητές την εκμάθηση της χρήσης κατάλληλων στατιστικών μεθόδων, μοντέλων και τεχνικών που απαιτούνται για την ανάλυση δεδομένων. Μετά την επιτυχή ολοκλήρωση του μαθήματος, οι φοιτητές θα είναι σε θέση:

- Να κατασκευάζουν διαστήματα εμπιστοσύνης για τις παραμέτρους του πληθυσμού
- Να διεξάγουν στατιστικούς ελέγχους υποθέσεων για συγκεκριμένα προβλήματα
- Να χρησιμοποιούν τις βασικές θεωρίες της ανάλυσης παλινδρόμησης
- Να μάθουν τις αρχές της στατιστικής συμπερασματολογίας, έτσι ώστε να είναι σε θέση να κατανοήσουν την ανάλυση που είναι απαραίτητη για ένα συγκεκριμένο σετ δεδομένων, και πώς μπορεί να εφαρμοστεί σωστά
- Να εκτιμούν τις παραμέτρους των στατιστικών υποδειγμάτων
- Να διεξάγουν ελέγχους υποθέσεων και να κατασκευάζουν διαστήματα εμπιστοσύνης για τις παραμέτρους του πληθυσμού

- Να εκτιμούν υποδείγματα παλινδρόμησης, να κατασκευάζουν προβλέψεις και να ερμηνεύουν κατάλληλα τα αποτελέσματα της ανάλυσης που διεξάγουν
- Να μπορούν να εφαρμόσουν, με τη χρήση του πακέτου R, τα στατιστικά υποδείγματα σε εμπειρικά προβλήματα

Συνιστώμενη Βιβλιογραφία

- Draper N, Smith H (1998) Applied Regression Analysis 3rd Edition Wiley
- Montgomery D (2012) Introduction to Linear regression Analysis, 5th Edition Wiley.
- Montgomery D. and Runger GC (2018) Applied Statistics and Probability for Engineers 7th Edition Wiley
- Weisberg, S. (2005). Applied Linear Regression, 3rd edition, Wiley
- Fox, J., and Weisberg, S. (2011). An R Companion to Applied Regression, 2nd edition, SAGE Publications Inc.

Διδακτικές και Μαθησιακές Μέθοδοι

Μια διάλεξη τριών ωρών εβδομαδιαίως και ασκήσεις μελέτης στο σπίτι (ορισμένες προς παράδοση).

Μέθοδοι Αξιολόγησης και Βαθμολόγησης

90% Γραπτή εξέταση.

10% εργασία βασισμένη σε προσομοιωμένα δεδομένα όπου εφαρμόζονται οι μεθοδολογίες οι αναπτύχθηκαν κατά τη διάρκεια του μαθήματος η οποία συνοδεύεται από συγγραφή επιστημονικής έκθεσης.