ΤΕΧΝΟΛΟΓΙΚΟ ΕΚΠΑΙΔΕΥΤΙΚΟ ΙΔΡΥΜΑ ΔΥΤΙΚΗΣ ΕΛΛΑΔΑΣ
ΣΧΟΛΗ ΔΙΟΙΚΗΣΗΣ & ΟΙΚΟΝΟΜΙΑΣ
ΤΜΗΜΑ ΛΟΓΙΣΤΙΚΗΣ & ΧΡΗΜΑΤΟΟΙΚΟΝΟΜΙΚΗΣ

ΠΤΥΧΙΑΚΗ ΕΡΓΑΣΙΑ

Θεωρία πιθανοτήτων ως το μαθηματικό θεμέλιο της στατιστικής – Έλεγχοι στατιστικών υποθέσεων

ΚΟΥΡΙΔΟΥ ΙΩ. ΚΩΝΣΤΑΝΤΙΝΑ
ΕΥΑΓΓΕΛΟΠΟΥΛΟΥ ΙΩ. ΑΡΤΕΜΙΣ

ΕΠΟΠΤΕΥΩΝ ΕΚΠΑΙΔΕΥΤΙΚΟΣ
ΜΕΓΑΡΙΤΗΣ ΑΘΑΝΑΣΙΟΣ
ΕΙΔΙΚΟΤΗΤΑ : ΜΑΘΗΜΑΤΙΚΑ - ΣΤΑΤΙΣΤΙΚΗ

ΜΕΣΟΛΟΓΓΙ 2017

ΠΡΟΛΟΓΟΣ

Η πεμπτουσία ενός απαιτητικού προγράμματος σπουδών, όπως αυτό του ΤΕΙ Δυτικής Ελλάδος είναι η συγγραφή της πτυχιακής εργασίας. Και αυτό διότι απαιτεί πολύ χρόνο αλλά και σύνθετες δεξιότητες ανάλυσης, επιλογής στοιχείων, κριτικής σκέψης αλλά και συνεργασίας για να μπορέσει να ολοκληρωθεί. Αυτή ακριβώς η διαδικασία την καθιστά απαιτητική, αλλά και εξαιρετικά ενδιαφέρουσα για εμάς που δουλέψαμε τον τελευταίο χρόνο πάνω στο συγκεκριμένο θέμα.

Θέλουμε να ευχαριστήσουμε από καρδιάς τον Καθηγητή μας κύριο Μεγαρίτη για την υποστήριξη, την βοήθεια, την καθοδήγηση αλλά και την υπομονή που επέδειξε. Χωρίς τη βοήθεια του δεν θα ήταν δυνατή η ολοκλήρωση της δουλειάς μας. Θα πρέπει να ληφθεί υπόψη το γεγονός ότι το μάθημα της Στατιστικής δεν μας ήταν τόσο οικείο, διότι το μεγαλύτερο κομμάτι του προγράμματος σπουδών μας είναι Λογιστική και Οικονομία. Συνεπώς η ολοκλήρωση της πτυχιακής μας έχει ακόμη μεγαλύτερη σημασία εξαιτίας της συνθετότητας του αντικειμένου στο οποίο αναφερόταν.

Για το τέλος θα θέλαμε να εκφράσουμε την αγάπη και την ευγνωμοσύνη στις οικογένειες μας για την αγάπη και την στήριξη όλα αυτά τα χρόνια. Πίστεψαν και επένδυσαν σε εμάς, σε ιδιαίτερα δύσκολες οικονομικές συνθήκες, και μας έδωσαν αγάπη, υποστήριξη και εμπιστοσύνη σε στιγμές δύσκολες, όταν αμφισβητούσαμε τους ίδιους μας τους εαυτούς.

ΠΕΡΙΛΗΨΗ

Η εργασία αυτή εστιάζεται στην ανάδειξη του ρόλου της έννοιας της θεωρίας των πιθανοτήτων ως το μαθηματικό θεμέλιο της στατιστικής. Βασικός σκοπός της εργασίας μας αποτελεί η ανάλυση των βασικών της εννοιών και ο τρόπος με τον οποίο μπορεί αυτή να εφαρμοστεί σε ένα μεγάλο μέρος σύγχρονων επιστημονικών και τεχνολογικών περιοχών.

Η Θεωρία των Πιθανοτήτων μας επιτρέπει να υπολογίσουμε ένα μέτρο της βεβαιότητας για τα συμπεράσματα στα οποία καταλήγουμε από την εφαρμογή στατιστικών μεθόδων ανάλυσης και πρόβλεψης.

Η Επαγωγική Στατιστική παρέχει τις μεθόδους εκείνες που μας επιτρέπουν να γενικεύουμε δειγματοληπτικά συμπεράσματα, δηλ. από τη μελέτη των στατιστικών παραμέτρων ενός αντιπροσωπευτικού δείγματος να εξάγουμε συμπεράσματα για τις αντίστοιχες παραμέτρους του πληθυσμού από τον οποίο έχει ληφθεί το δείγμα.

Η Συνδυαστική ασχολείται με τη μέτρηση του πλήθους των σχηματισμών που προκύπτουν από ένα σύνολο στοιχείων και έχουν καθορισμένη δομή και ιδιότητες. Στόχος της Συνδυαστικής είναι η ανάπτυξη μεθόδων, αναλυτικών και αλγοριθμικών τεχνικών ώστε η μέτρηση του πλήθους των σχηματισμών να γίνεται όσο το δυνατόν αποτελεσματικότερα. Με απλά λόγια η Συνδυαστική απαντά σε προβλήματα του τύπου: (α) Με πόσους τρόπους μπορώ να κάνω κάτι (β) Πόσα αντικείμενα υπάρχουν με μια δοσμένη ιδιότητα

Καλείται διάταξη των n στοιχείων ενός συνόλου Ε, λαμβανόμενων ανά m, κάθε διατεταγμένη ακολουθία m στοιχείων που λαμβάνεται από τα n στοιχεία του συνόλου Ε. Αν ορισμένα από τα m στοιχεία συμπίπτουν, τότε μιλούμε για διατάξεις με επανάληψη.

Έστω Ε ένα σύνολο με n στοιχεία. Καλούμε συνδυασμό των n στοιχείων ανά m κάθε υποσύνολο του Ε που έχει m στοιχεία. Δυο συνδυασμοί που έχουν τον ίδιο αριθμό στοιχείων μπορεί να διαφέρουν μόνο ως προς τα στοιχεία αυτά καθ' αυτά και όχι λόγω της θέσεώς τους.

Έστω Ε ένα σύνολο με n στοιχεία. Μια διάταξη των n στοιχείων του Ε ανά n λέγεται μετάθεση. Μετάθεση είναι μια ακολουθία των n στοιχείων του Ε, η απλούστερα μια διατεταγμένη n-ιάδα.

Τυχαία μεταβλητή είναι η συνάρτηση που απεικονίζει κάθε απλό ενδεχόμενο ενός πειράματος σε έναν μοναδικό πραγματικό αριθμό, απεικονίζει δηλαδή κάθε στοιχείο ενός δειγματικού χώρου σε έναν πραγματικό αριθμό. Η τυχαία μεταβλητή Χ λοιπόν αντιστοιχίζει το στοιχείο s του δειγματοχώρου στον αριθμό x με X(s) =x.

Οι μεταβλητές διακρίνονται στις διακριτές και στις συνεχείς. Διακριτές είναι οι μεταβλητές που οι τιμές τους εκφράζονται μόνο με ακέραιους αριθμούς. Τέτοιες μεταβλητές είναι ο αριθμός των ανέργων, ο αριθμός των ποδηλάτων που κυκλοφορούν σε μία πόλη κλπ. Συνεχείς είναι οι μεταβλητές που παίρνουν οποιαδήποτε τιμή και στην πράξη εκφράζονται είτε με ακέραιους είτε με δεκαδικούς αριθμούς.

Ενώ για τις τυχαίες διακριτές μεταβλητές εργαζόμαστε με αθροίσματα πιθανοτήτων και μπορούμε φυσικά να υπολογίσουμε πιθανότητες για συγκεκριμένες τιμές της τυχαίας μεταβλητής, για τις συνεχείς τυχαίες μεταβλητές χρησιμοποιούμε ολοκληρώματα.

Η μέση τιμή μιας διακριτής τυχαίας μεταβλητής υπολογίζεται ως το άθροισμα των γινομένων των τιμών της τυχαίας μεταβλητής επί την πιθανότητα της κάθε τιμής. Εκφράζει ποια αναμένεται να είναι η τιμή της Χ μετά από μεγάλο αριθμό επαναλήψεων ενός πειράματος

Για τις συνεχείς τυχαίες μεταβλητές ορίζουμε μια Συνάρτηση Πυκνότητας Πιθανότητας, δηλαδή μια συνάρτηση που αναπαριστά την πυκνότητα – συγκέντρωση τιμών μιας συνεχούς τυχαίας μεταβλητής Χ. Ανάλογα με την πυκνότητα των τιμών σε συγκεκριμένο διάστημα η συνάρτηση πυκνότητας πιθανότητας αποδίδει μια πιθανότητα. Η πιθανότητα αυτή ισούται με το ορισμένο ολοκλήρωμα της συνάρτησης πυκνότητας πιθανότητας, από το ένα άκρο του διαστήματος ως το άλλο. Και όπως γνωρίζουμε για τα ολοκληρώματα, το ορισμένο ολοκλήρωμα από το ένα άκρο έως το άλλο, ισούται με το εμβαδό του χωρίου που περικλείεται ανάμεσα στην καμπύλη της γραφικής παράστασης της συνάρτησης, τον άξονα των χ και τις δύο κατακόρυφες γραμμές που φέρουμε στα δύο άκρα του διαστήματος για το οποίο υπολογίζουμε την πιθανότητα

Οι τιμές μιας διακριτής μεταβλητής για τα μέλη ενός πληθυσμού ή ενός δείγματος μπορούν να παρουσιαστούν συνοπτικά σε έναν πίνακα κατανομής συχνοτήτων. Οι σχετικές συχνότητες της μεταβλητής, όπως έχει ήδη αναφερθεί είναι οι απόλυτες συχνότητες f διαιρεμένες με το μέγεθος ν του πληθυσμού ή του δείγματος , f/ν, πολλαπλασιασμένες επί 100, δηλαδή 100 f/ν, εφόσον θέλουμε να τις εκφράζουμε ως ποσοστά επί τοις εκατό.

Όταν η πιθανότητα για μια τυχαία μεταβλητή Χ να πάρει μια συγκεκριμένη τιμή x ακολουθεί τον παραπάνω τύπο με x=0,1,2,3,…ν, τότε λέμε ότι η τυχαία μεταβλητή ακολουθεί τη διωνυμική κατανομή με παράμετρο p και χρησιμοποιούμε το συμβολισμό Χ~Β(ν,p) , (όπου το Β προέρχεται από την Αγγλική λέξη binomial = διωνυμική).

Η γεωμετρική κατανομή είναι μια διακριτή συνάρτηση κατανομής τυχαίας μεταβλητής. Περιγράφει ένα τυχαίο πείραμα με δυο πιθανά αποτελέσματα (επιτυχία - αποτυχία) και πιθανότητα επιτυχίας p που επαναλαμβάνεται μέχρι να έχουμε μια επιτυχία.

Η κατανομή Poisson χρησιμοποιείται πολύ συχνά σε περιπτώσεις όπου η επιτυχία είναι ένα σχετικά σπάνιο ενδεχόμενο, σε μια σειρά πολλών ίσως άπειρων δοκιμών. Στην περίπτωση δειγματοληψίας χωρίς επανάθεση από ένα «μικρό» πληθυσμό χρησιμοποιούμε την υπεργεωμετρική κατανομή. Αν Χ μια τυχαία μεταβλητή η οποία παίρνει τις τιμές x = 1,2…k με σταθερή πιθανότητα 1/k τότε λέμε ότι η Χ ακολουθεί την διακριτή ομοιόμορφη κατανομή.

Η κανονική κατανομή είναι η σημαντικότερη κατανομή από τις κατανομές των συνεχών αλλά και των διακριτών τυχαίων μεταβλητών. Αποτελεί σύμφωνα με αρκετούς συγγραφείς και ερευνητές τη σημαντικότερη κατανομή πιθανότητας τόσο από θεωρητική άποψη όσο και από άποψη εφαρμογών. Αναφέρεται σε συνεχείς τυχαίες μεταβλητές που περιγράφουν φαινόμενα στα οποία επιδρά ένας μεγάλος αριθμός παραγόντων.

Μια κανονική κατανομή προσδιορίζεται πλήρως, όταν είναι γνωστές οι τιμές του μέσου όρου και της τυπική απόκλισης. Τότε με ολοκλήρωση της συνάρτησης πυκνότητας πιθανότητας της κανονικής κατανομής μπορούμε να υπολογίσουμε οποιαδήποτε πιθανότητα της μορφής $P(x_1 \leq X \leq x_2)$.

Η τυπική κανονική κατανομή είναι χρήσιμη γιατί ενώ γενικά η συνάρτηση πυκνότητας πιθανότητας είναι αρκετά σύνθετη και δύσκολα ολοκληρώσιμη ώστε να υπολογιστούν οι ζητούμενες κατά περίπτωση πιθανότητες, υπάρχουν ωστόσο πίνακες αθροιστικής συνάρτησης πιθανότητας διαθέσιμοι για την τυπική κανονική κατανομή.

Αν και η κανονική κατανομή χρησιμοποιείται σε πάρα πολλές εφαρμογές, δεν καλύπτει όλες τις ανάγκες της επαγωγικής στατιστικής. Υπάρχουν και ορισμένες άλλες βασικές κατανομές, οι οποίες έχουν σημαντικές εφαρμογές στους στασιαστικούς ελέγχους που παρουσιάζονται. Η πλέον σημαντική είναι η κατανομή t ή t του Student.

Η κατανομή t είναι στην πραγματικότητα μια οικογένεια κατανομών που διαφοροποιούνται επειδή έχουν διαφορετικούς βαθμούς ελευθερίας. Μοιάζει με την τυπική κατανομή και όταν έχει άπειρους βαθμούς ελευθερίας δηλαδή πρακτικά πολλούς βαθμούς ελευθερίας, προσεγγίζεται από την τυπική κανονική κατανομή. Άρα ως πρακτικό κανόνα μπορούμε να θυμόμαστε ότι όταν σε ένα πρόβλημα εκτίμησης της μέσης τιμής, εξαιτίας μικρού δείγματος, έχουμε λίγους βαθμούς ελευθερίας, μπορούμε να χρησιμοποιήσουμε την t κατανομή, ενώ όταν οι βαθμοί ελευθερίας είναι περισσότεροι από 30 χρησιμοποιούμε την τυπική κανονική κατανομή.

Με δύο πραγματικούς αριθμούς α,β με α<β και X μια συνεχή τυχαία μεταβλητή η οποία μπορεί να πάρει τιμές μόνο μεταξύ των α και β. ας ορίσουμε μια κατανομή τέτοια ώστε σε υποδιαστήματα του [α,β] με ίσο πλάτος να αντιστοιχεί ίδια πιθανότητα, ή με άλλα λόγια για κάθε χ Ε (α,β) η πιθανότητα η τυχαία μεταβλητή να βρίσκεται κοντά στο χ να είναι η ίδια. Αυτή είναι η ομοιόμορφη κατανομή.

Ο Στατιστικός έλεγχος υποθέσεων συνίσταται στον έλεγχο κάποιας υπόθεσης όσον αφορά την τιμή κάποιας παραμέτρου ενός πληθυσμού. Ένα παράδειγμα αποτελεί αυτό του μέσου όρου με βάση την τιμή της παραμέτρου η οποία προέκυψε από κάποιο δείγμα του συγκεκριμένου πληθυσμού. Συγκεκριμένα τίθεται μια υπόθεση για την τιμή της συγκεκριμένης παραμέτρου στον πληθυσμό. Στη συνέχεια λαμβάνουμε ένα δείγμα από τον παραπάνω πληθυσμό και υπολογίζουμε την αντίστοιχη τιμή της παραμέτρου στο δείγμα. Κατόπιν συγκρίνοντας τη διαφορά μεταξύ της τιμής της παραμέτρου που προέρχεται από το δείγμα με αυτή της υποθέσεως που έχουμε θέσει για την τιμή της παραμέτρου του πληθυσμού, αποφασίζουμε με κάποια πιθανότητα σφάλματος, κατά πόσο η διαφορά η οποία παρατηρείται οφείλεται σε δειγματοληπτικά σφάλματα (είναι δηλαδή φαινομενική) ή όχι.

ΠΙΝΑΚΑΣ ΠΕΡΙΕΧΟΜΕΝΩΝ

Εισαγωγή

Η στατιστική είναι η επιστήμη της συλλογής, ταξινόμησης, παρουσίασης, και ανάλυσης αριθμητικών δεδομένων που αναφέρονται σε ιδιότητες πολυπληθών ομάδων. Ως συλλογή στοιχείων νοείται η διαδικασία της μετρήσεως ή απαριθμήσεως χαρακτηριστικών ιδιοτήτων των μονάδων ενός συνόλου ή τμήματός του και η καταγραφή των αριθμητικών δεδομένων που προκύπτουν.

Τα δεδομένα αυτά αναφέρονται σε μετρήσεις (ή «παρατηρήσεις») που προέρχονται από ένα πείραμα ή μία δειγματοληπτική έρευνα μιας μεταβλητής. Μεταβλητές είναι, για παράδειγμα, το ύψος ή το βάρος μιας ομάδας ανθρώπων, το εθνικό εισόδημα της Ελλάδος σε μια σειρά ετών, κ.λπ.

Ως μαθηματικό θεμέλιο της στατιστικής, η θεωρία πιθανοτήτων είναι απαραίτητη σε πολλές δραστηριότητες που περιλαμβάνουν ανάλυση μεγάλων συνόλων δεδομένων. Μέθοδοι της θεωρίας πιθανοτήτων εφαρμόζονται και στην περιγραφή πολύπλοκων συστημάτων, όπως στη στατιστική μηχανική.

Η συγκεκριμένη εργασία αποτελείται από επτά κεφάλαια που πραγματεύονται τις διαφορετικές όψεις και εφαρμογές της θεωρίας των πιθανοτήτων.

Στο πρώτο κεφάλαιο παρουσιάζονται μια ιστορική αναδρομή της θεωρίας των Πιθανοτήτων, τα στοιχεία της θεωρίας των συνόλων με έμφαση στο πείραμα τύχης και στον δειγματικό χώρο ενώ παρουσιάζονται διάφοροι ορισμοί της πιθανότητας καθώς και οι ιδιότητες της.

Στο δεύτερο κεφάλαιο παρουσιάζονται τα βασικά στοιχεία της συνδυαστικής ανάλυσης. Πιο συγκεκριμένα αναλύονται οι διατάξεις, οι μεταθέσεις και οι συνδυασμοί.

Στο τρίτο κεφάλαιο παρουσιάζεται η έννοια της μεταβλητής οι ορισμοί των τυχαίων διακριτών και τυχαίων συνεχών μεταβλητών, ενώ παρουσιάζονται η μέση τιμή και η διασπορά τυχαίας μεταβλητής, η μέση τιμή διασπορά και τυπική απόκλιση συνεχών τυχαίων μεταβλητών, τα μέτρα κεντρικής τάσης και διασποράς.

Το τέταρτο κεφάλαιο αφιερώνεται στις βασικές διακριτές κατανομές. Παρουσιάζονται η Διωνυμική κατανομή, η κατανομή Poisson, η Γεωμετρική κατανομή, η Υπεργεωμετρική κατανομή, και η Ομοιόμορφη διακριτή Κατανομή.

Το πέμπτο κεφάλαιο επικεντρώνεται στις βασικές συνεχείς κατανομές. Πιο συγκεκριμένα παρουσιάζονται η κανονική κατανομή, η κατανομή t–Student, η Ομοιόμορφη συνεχής κατανομή και η Εκθετική κατανομή,

Το έκτο κεφάλαιο στον έλεγχο στατιστικών υποθέσεων και τη σύνδεση με τη θεωρία των πιθανοτήτων. Το έβδομο κεφάλαιο επικεντρώνεται στην ανάλυση συσχέτισης μεταβλητών ενώ παρουσιάζεται και η απλή γραμμική παλινδρόμηση.

ΚΕΦΑΛΑΙΟ 1 : ΠΙΘΑΝΟΤΗΤΕΣ

1.1 Εισαγωγή

Η **θεωρία των πιθανοτήτων** αποτελεί ανεξάρτητο κλάδο των Μαθηματικών με εφαρμογές στη Φυσική, στη Βιολογία, στη Γενετική, στην Κοινωνιολογία, στη Ψυχολογία, στην Οικονομία, στις Τηλεπικοινωνίες, στους Ηλ. Υπολογιστές κ.ά.
Πρόκειται δηλαδή για τον κλάδο των μαθηματικών ο οποίος ασχολείται με την ανάλυση τυχαίων φαινομένων. Κεντρικό ρόλο στη θεωρία πιθανοτήτων παίζει η έννοια της **πιθανότητας**, ενώ σημαντικές είναι οι τυχαίες μεταβλητές, οι συναρτήσεις κατανομής, οι στοχαστικές διαδικασίες και τα γεγονότα.

Η έννοια της πιθανότητας **διαφέρει** ποιοτικά από τις άλλες παραδοσιακές **μαθηματικές έννοιες** και η διαφοροποίηση αυτή οφείλεται στην ίδια τη φύση της έννοιας σε σχέση με το επίπεδο της **διαισθητικής** της **αντίληψης**, στον τρόπο **υπολογισμού** της και στους **βασικούς νόμους** που αυτή υπακούει.
Η **πιθανότητα** εκφράζει την αβεβαιότητα του ανθρώπου για την εξέλιξη των φαινομένων. Το μέτρο δε της πιθανότητας ορίστηκε να είναι ένας θετικός αριθμός, μικρότερος ή ίσος του ένα. Αν και τα γεγονότα που μελετώνται από τη θεωρία πιθανοτήτων, όπως π.χ. η ρίψη ενός ζαριού ή το στρίψιμο ενός κέρματος, είναι τυχαία, όταν επαναλαμβάνονται πολλές φορές η αλληλουχία των τυχαίων γεγονότων παρουσιάζει ορισμένα στατιστικά μοτίβα τα οποία μπορούν να μελετηθούν και να προβλεφθούν.

Η διατύπωση της θεωρίας πιθανοτήτων έχει ως αντικείμενο την μελέτη των **μαθηματικών μοντέλων-υποδειγμάτων** με τη βοήθεια των οποίων μπορούμε να περιγράψουμε ένα **τυχαίο φαινόμενο**.
Ως μαθηματικό θεμέλιο της στατιστικής, η θεωρία πιθανοτήτων είναι απαραίτητη σε πολλές δραστηριότητες που περιλαμβάνουν ανάλυση μεγάλων συνόλων δεδομένων. Μέθοδοι της θεωρίας πιθανοτήτων εφαρμόζονται και στην περιγραφή πολύπλοκων συστημάτων, όπως στη στατιστική μηχανική. Μία μεγάλη ανακάλυψη του εικοστού αιώνα ήταν η πιθανοκρατική φύση των φυσικών νόμων σε υποατομικό επίπεδο, σύμφωνα με τα ευρήματα της κβαντομηχανικής
Η Θεωρία Πιθανοτήτων διατυπώνει **λογικούς νόμους** που μπορούν να εφαρμοστούν σε καταστάσεις **τελείως τυχαίες** οι οποίες αποτυπώνονται σε **μαθηματικά μοντέλα.**

Τα **μαθηματικά μοντέλα** διακρίνονται σε δύο κατηγορίες: τα προσδιοριστικά (deterministic) και τα στοχαστικά (stochastic). Εάν οι συνέπειες οποιασδήποτε μεταβολής σε ένα σύστημα μπορούν να προβλεφθούν με βεβαιότητα και ακρίβεια τότε το μοντέλο ονομάζεται προσδιοριστικό. Εάν όμως οι μεταβολές στο σύστημα μπορούν να εκφραστούν μαθηματικά μόνο με τη βοήθεια τυχαίων μεταβλητών τότε το μοντέλο ονομάζεται στοχαστικό.

Τις τελευταίες δεκαετίες παρατηρείται μια στροφή των ερευνητών σε όλα σχεδόν τα επιστημονικά πεδία, προς τα **στοχαστικά μοντέλα**, αφού αυτά προσφέρουν καλύτερες δυνατότητες αναπαράστασης των φαινομένων, που στην πλειονότητά τους ενσωματώνουν τυχαία στοιχεία.

1.2 Ιστορική αναδρομή της Θεωρ

ERROR: stackunderflow
OFFENDING COMMAND: exch

STACK:

/_ct_na