

Ι. ΠΑΝΑΡΕΤΟΥ & Ε. ΞΕΚΑΛΑΚΗ
Καθηγητών του Τμήματος Στατιστικής
του Οικονομικού Πανεπιστημίου Αθηνών

ΕΙΣΑΓΩΓΗ
στη
ΣΤΑΤΙΣΤΙΚΗ ΣΚΕΨΗ

Τόμος 1
(Περιγραφική Στατιστική)

ΑΘΗΝΑ

Στους φοιτητές μας

ΠΕΡΙΕΧΟΜΕΝΑ

ΠΡΟΛΟΓΟΣ

ΚΕΦΑΛΑΙΟ 1

ΕΙΣΑΓΩΓΗ ΣΤΑ ΒΑΣΙΚΑ ΠΡΟΒΛΗΜΑΤΑ ΤΗΣ ΣΤΑΤΙΣΤΙΚΗΣ

1.1	Αντικείμενο της Στατιστικής και Στατιστική Σκέψη	1
1.2	Ενδεικτικές περιπτώσεις εφαρμογής της Στατιστικής	4
1.3	Μορφές και Είδη Δεδομένων	7
1.4	Οι βασικές έννοιες της εφαρμοσμένης Στατιστικής	16
	Τα βασικά στοιχεία της Στατιστικής	28
1.5	Αξιολόγηση των Στατιστικών Μελετών	30
1.6	Ο ρόλος των Στατιστικών για την εξαγωγή Στατιστικών Συμπερασμάτων	35
1.7	Τρόποι συλλογής στοιχείων	37
	Τυχαία δείγματα	37
	Τυχαιοποιημένα πειράματα	38
	Διαθέσιμα δεδομένα	39
	Λογικές υποομάδες	40
1.8	Στατιστικός τρόπος σκέψης για Διαδικασίες Λήψης Αποφάσεων	41
1.9	Χάρτες ροής	54
1.10	Εισαγωγή στο Σχεδιασμό και Ανάλυση Πειραμάτων	55
	Κύρια σημεία για το σχεδιασμό ενός πειράματος	65
1.11	Χρήση των υπολογιστών στη Στατιστική Ανάλυση	65
	Το πακέτο MINITAB	67
	Η εντολή READ	68
	Η εντολή SET	70

Η εντολή PRINT	71
Διορθώσεις: οι εντολές LET, DELETE, INSERT	71
Αριθμητικές πράξεις: οι εντολή LET	73
Η εντολή NAME	73
Οι εντολές SAVE και RETRIEVE	74
Αποθήκευση και ανάσυρση δεδομένων σε δισκέττα	74
Υποεντολές	75
Οι εντολές HELP και STOP	75
Το στατιστικό πακέτο SAS	76
Εισαγωγή στοιχείων	76
Εκτύπωση δεδομένων	79
Αριθμητικές πράξεις στο SAS	79
ΑΣΚΗΣΕΙΣ	81

ΚΕΦΑΛΑΙΟ 2

ΠΕΡΙΓΡΑΦΗ ΔΕΔΟΜΕΝΩΝ ΜΕ ΓΡΑΦΙΚΕΣ ΜΕΘΟΔΟΥΣ

2.1 Κατανομή Δεδομένων	87
2.2 Διαγράμματα Σημείων	88
2.3 Διαγράμματα Μίσχου-Φύλλου	90
2.4 Χρήση του MINITAB και του SAS για την κατασκευή διαγραμμάτων μίσχου - φύλλου	97
Χρήση του Minitab	97
Χρήση του SAS	99
2.5 Κατανομές Συχνότητας και Ιστογράμματα	100
2.6 Χρήση του MINITAB και του SAS για την κατασκευή ιστογραμμάτων	106
Χρήση του Minitab	106

Χρήση του SAS	109
2.7 Πολύγωνα Συχνότητας	111
2.8 Κυκλικά Διαγράμματα, Ραβδογράμματα, Εικονογράμματα και Γραμμογραφήματα	115
Κυκλικά διαγράμματα	116
Ραβδογράμματα	118
Εικονογράμματα	124
Γραμμογραφήματα	126
2.9 Παραποιήσεις με Γραφικές Παραστάσεις	127
ΑΣΚΗΣΕΙΣ	136

ΚΕΦΑΛΑΙΟ 3

ΑΡΙΘΜΗΤΙΚΗ ΠΕΡΙΓΡΑΦΗ ΔΕΔΟΜΕΝΩΝ

3.1 Μέτρα Θέσης ή Κεντρικής Τάσης	146
Αριθμητικός μέσος	147
Παραλλαγές του αριθμητικού μέσου	148
Διάμεσος	148
Επικρατούσα τιμή	149
Γεωμετρικός μέσος	150
Αρμονικός μέσος	152
3.2 Σύγκριση των ιδιοτήτων Μέσου, Διαμέσου και Επικρατούσας τιμής	154
3.3 Μέτρα Μεταβλητότητας ή Διασποράς	158
Εκταση ή εύρος	159
Μέση απόλυτη απόκλιση	160
Διακύμανση ή διασπορά	162
Τυπική απόκλιση	164

Υπολογισμός της δειγματικής διακύμανσης και τυπικής απόκλισης	165
Βαθμοί ελευθερίας	168
Σχέση μέσης απόλυτης απόκλισης και τυπικής απόκλισης	169
Ερμηνεία και χρήση της τυπικής απόκλισης	169
Θεώρημα του Chebyshev	170
Ενδοτεταρτημοριακό εύρος	174
3.4 Χρήση του MINITAB και του SAS για τον υπολογισμό των μέτρων θέσης και μεταβλητότητας	175
MINITAB	175
SAS	176
3.5 Προσέγγιση περιγραφικών μέσων για ομαδοποιημένα δεδομένα	178
3.6 Μέτρα σχετικής θέσης	181
Ποσοστιαία ή εκατοστιαία σημεία	181
Τυποποιημένες τιμές	189
3.7 Μέτρα σχετικής μεταβλητότητας	197
Συντελεστής μεταβλητότητας	197
Μέση διαφορά του Gini	199
3.8 Διάγραμμα Πλαισίου και Απολήξεων	200
Διάγραμμα Πλαισίου και Απολήξεων με το MINITAB και το SAS	210
MINITAB	210
SAS	211
3.9 Μέτρα ασυμμετρίας και κύρτωσης	211
ΑΣΚΗΣΕΙΣ	218
ΒΙΒΛΙΟΓΡΑΦΙΑ	251
ΕΥΡΕΤΗΡΙΟ	254
ΕΥΡΕΤΗΡΙΟ ΑΓΓΛΙΚΩΝ ΟΡΩΝ	257

ΠΡΟΛΟΓΟΣ

Ο τίτλος του βιβλίου αυτού εκφράζει με χαρακτηριστικό τρόπο την προσπάθεια των συγγραφέων να βοηθήσουν κάποιον χωρίς προηγούμενες γνώσεις στατιστικής να αντιληφθεί τις βασικές έννοιες της επιστήμης αυτής αλλά και, κυρίως, να είναι σε θέση να διαμορφώσει ένα τρόπο σκέψης ο οποίος θα τον βοηθήσει να αντιμετωπίζει στατιστικά προβλήματα.

Τα περισσότερα βιβλία με παρόμοιο περιεχόμενο χαρακτηρίζονται από την εισαγωγή που παρέχουν στις στατιστικές έννοιες. Οι συγγραφείς όμως του παρόντος συγγράμματος πιστεύουν ότι σημαντικότερο ίσως είναι να αποκτήσει κανείς τη δυνατότητα να σκέπτεται στατιστικά.

Όπως είναι γνωστό, η Στατιστική είναι μια επιστήμη που έχει πραγματοποιήσει ραγδαία ανάπτυξη τα τελευταία τριάντα χρόνια. Δεν υπάρχει σήμερα επιστήμη η οποία να μην χρησιμοποιεί με τον ένα ή με τον άλλο τρόπο στατιστικές μεθόδους προκειμένου να μελετήσει τα διαθέσιμα στοιχεία και να προχωρήσει σε συμπερασματολογία βασισμένη στα στοιχεία αυτά. Η Ιατρική, τα Οικονομικά, η Διοίκηση των Επιχειρήσεων, η Πολιτική Επιστήμη, η Ανθρωπολογία, η Ιστορία, η Βιολογία τα Νομικά είναι μερικές μόνο από τις επιστήμες που χρησιμοποιούν στατιστικές μεθόδους.

Η ραγδαία εξέλιξη των ηλεκτρονικών υπολογιστών και τα στατιστικά πακέτα, τα οποία έχουν αναπτυχθεί προκειμένου να αναλύουν μεγάλες ποσότητες δεδομένων, έχουν δώσει τα τελευταία δέκα χρόνια ακόμα μεγαλύτερη ώθηση στην ανάπτυξη της Στατιστικής και την χρήση της σε άλλες επιστημονικές περιοχές. Τέτοια πακέτα είναι το STATGRAPHICS, το MINITAB, το SAS, το SPSS, το BMDP, το SYSTAT κ.λ.π. Κάθε ένα από τα πακέτα αυτά έχει τα ιδιαίτερα χαρακτηριστικά του και τα πλεονεκτήματά του. Είναι λοιπόν απαραίτητο ένα τέτοιο εισαγωγικό σύγγραμμα να δίνει τη δυνατότητα στον αναγνώστη του να εξοικειωθεί με ένα από τα πακέτα αυτά. Το πακέτο που επελέγη από τους συγγραφείς για το σκοπό αυτό είναι το MINITAB. Οι λόγοι είναι πολλοί: το πακέτο αυτό είναι από τα ευκολότερα που υπάρχουν στην αγορά. Επίσης, το πακέτο αυτό χρησιμοποιείται στο μάθημα "Περιγραφική Στατιστική", που διδάσκεται στους πρωτοετείς φοιτητές του Τμήματος Στατιστικής του Οικονομικού Πανεπιστημίου Αθηνών. Είναι επίσης το πακέτο εκείνο που χρησιμοποιείται στις περισσότερες χώρες του κόσμου που έχουν αναπτύξει την επιστήμη της Στατιστικής και στα περισσότερα προγράμματα σπουδών για να συμπληρώνει ένα εισαγωγικό μάθημα Στατιστικής.

Ταυτόχρονα όμως με την χρήση του MINITAB δίνονται και οι βασικές εντολές χρήσης του SAS. Αυτό γιατί το SAS είναι το ισχυρότερο, ίσως, στατιστικό πακέτο που κυκλοφορεί στη διεθνή αγορά, ικανό να αναλύσει τεράστιους όγκους στοιχείων με τη χρήση ακόμα και των πιο πολύπλοκων στατιστικών τεχνικών.

Για πολλά χρόνια τώρα υπάρχει μια φιλοσοφική αντίθεση για αυτή καθαυτή την επιστήμη της Στατιστικής και για το κατά πόσον θα πρέπει να θεωρείται ένας κλάδος των Μαθηματικών ή αν θα πρέπει

να αποσυνδεθεί από τα Μαθηματικά. Κατά τη γνώμη των συγγραφέων η Στατιστική αποτελεί μια επιστήμη που έχει τη δική της δομή και φιλοσοφία. Τα Μαθηματικά είναι βέβαια ένα απαραίτητο και χρήσιμο εργαλείο για την ανάπτυξή της. Χωρίς τη γνώση Μαθηματικών ο χρήστης των στατιστικών μεθόδων είναι ένας απλός χρήστης εργαλείων με τις βασικές τους προδιαγραφές. Η ανάπτυξη όμως και η παρουσίαση των θεμάτων της στατιστικής χρειάζεται τον δικό της τρόπο. Για το λόγο αυτό άλλωστε και οι συγγραφείς επέλεξαν ως τίτλο του βιβλίου τον τίτλο "Εισαγωγή στην Στατιστική Σκέψη".

Όπως θα διαπιστώσει ο αναγνώστης, η ανάπτυξη των εννοιών και η έμφαση που δίνεται σ'αυτές είναι κάπως διαφορετική στο βιβλίο αυτό από άλλα αντίστοιχα που κυκλοφορούν στην αγορά. Ένας λόγος είναι ότι, κατά τη γνώμη των συγγραφέων, θα πρέπει να δοθεί έμφαση στο γεγονός ότι τα στοιχεία τα οποία η Στατιστική αναλύει είναι στοιχεία που προκύπτουν, συνήθως, από μια διαδικασία σε εξέλιξη και όχι στοιχεία από ένα τυχαίο δείγμα το οποίο λαμβάνεται από κάποιο πληθυσμό. Δίνεται επίσης έμφαση στην κατανόηση των εννοιών και για το λόγο αυτό δίνονται συνήθως φραστικές εξηγήσεις εννοιών και μαθηματικών τύπων.

Ο τρόπος ανάπτυξης των θεμάτων έχει επηρεαστεί σημαντικά από την διδασκαλία από τους συγγραφείς αντιστοιχών μαθημάτων σε μια σειρά από τμήματα και πανεπιστήμια της Ελλάδας και του εξωτερικού. Η διδασκαλία αντιστοιχών μαθημάτων σε φοιτητές διαφορετικών κρατών από τους συγγραφείς τα τελευταία δεκαπέντε χρόνια τους έδωσε μια πολύτιμη εμπειρία για τον τρόπο ανάπτυξης των θεμάτων. Τους έδωσε επίσης την ευκαιρία να διαπιστώσουν τις εξελίξεις και τον τρόπο που η διδασκαλία του μαθήματος αυτού επιβάλλεται να γίνεται με την πάροδο των ετών.

Σημαντικά εξάλλου επηρέασε την προσέγγιση της παρουσίασης των θεμάτων, η παρακολούθηση ενός συνεδρίου με τον τίτλο "Making Statistics more effective in schools of business" που έγινε στο Knoxville των Ηνωμένων Πολιτειών, τον Ιούνιο του 1992, στο οποίο και συμμετείχαν οι συγγραφείς. (Το συνέδριο αυτό γίνεται μια φορά το χρόνο και καταγράφει όλες τις σύγχρονες αντιλήψεις για τον τρόπο και τις μεθόδους που πρέπει να χρησιμοποιούνται για τη διδασκαλία ενός εισαγωγικού μαθήματος της Στατιστικής κυρίως σε ιδρύματα οικονομικής κατεύθυνσης).

Το βιβλίο αυτό απευθύνεται κυρίως στους πρωτοετείς φοιτητές του Τμήματος Στατιστικής του Οικονομικού Πανεπιστημίου Αθηνών. Επίσης απευθύνεται στους φοιτητές των υπολοίπων τμημάτων του ίδιου Πανεπιστημίου, αλλά και των άλλων Πανεπιστημίων της χώρας, οι οποίοι παρακολουθούν ένα εισαγωγικό μάθημα της Στατιστικής. Ο τρόπος όμως ανάπτυξης των εννοιών και η διάρθρωσή τους είναι τέτοια που οι συγγραφείς πιστεύουν ότι το καθιστά κατάλληλο βοήθημα για φοιτητές οποιασδήποτε επιστήμης οι οποίοι έχουν την πρώτη γνωριμία τους με την επιστήμη της Στατιστικής μέσω κάποιου εισαγωγικού μαθήματος Στατιστικής του προγράμματος σπουδών τους. Γενικότερα, οι συγγραφείς πιστεύουν ότι το βιβλίο αυτό είναι κατάλληλο για οποιονδήποτε επιθυμεί να εξοικειωθεί με τη Στατιστική επιστήμη.

Στον πρώτο αυτό τόμο παρουσιάζονται οι βασικές έννοιες της Περιγραφικής Στατιστικής.

Στο πρώτο κεφάλαιο γίνεται μια εισαγωγή στα βασικά προβλήματα της Στατιστικής, στον τρόπο σκέψης που απαιτείται για την αντιμετώπισή τους και στη συνεισφορά των στατιστικών πακέτων για την επίλυσή τους.

Στο δεύτερο κεφάλαιο παρουσιάζονται και αναλύονται οι κυριότερες μέθοδοι περιγραφής δεδομένων με γραφικές μεθόδους και επισημαίνεται η προσοχή που πρέπει να δίνεται στην εξαγωγή συμπερασμάτων από αυτά. Οι μέθοδοι αυτές εξηγούνται με τα στατιστικά πακέτα Minitab και SAS.

Στο τρίτο κεφάλαιο ορίζονται και μελετώνται τα κυριότερα αριθμητικά μέτρα περιγραφής δεδομένων. Γίνεται επίσης συγκριτική αξιολόγηση των μέτρων αυτών και αναπτύσσεται η διαδικασία υπολογισμού τους μέσω των στατιστικών πακέτων Minitab και SAS.

Οι συγγραφείς θα ήθελαν να ευχαριστήσουν τους φοιτητές, στους οποίους, σε διαφορετικές χρονικές περιόδους και σε διαφορετικά τμήματα, δίδαξαν κατά τη διάρκεια της καριέρας τους εισαγωγικά μαθήματα Στατιστικής. Οι φοιτητές των Πανεπιστημίων του Sheffield και Bradford Αγγλίας, του Trinity College του Δουβλίνου, του Πανεπιστημίου του Missouri και της Iowa των Ηνωμένων Πολιτειών, των Πανεπιστημίων Κρήτης, Πάτρας και του Οικονομικού Πανεπιστημίου Αθηνών με τα σχόλια και τις υποδείξεις τους βοήθησαν ώστε να βελτιωθούν οι παραδόσεις και να οδηγήσουν στο σύγγραμμα αυτό. Ευχαριστούν επίσης τους συναδέλφους τους στα Πανεπιστήμια αυτά, που με συζητήσεις και ανταλλαγή απόψεων συνεισέφεραν στην καλύτερη παρουσίαση των εννοιών.

I. Πανάρετος

Αθήνα 1993

E. Ξεκαλάκη