

Π Ε Ρ Ι Ε Χ Ο Μ Ε Ν Α

Περιεχόμενα.....	1
Κατάλογος Σχημάτων	9
Κατάλογος Πινάκων.....	12
Πρόλογος.....	13
Μερικές εκδοτικές-γλωσσικές και άλλες παρατηρήσεις.....	18

ΜΕΡΟΣ Ι

Σύντομη Επισκόπηση βασικών εννοιών Μεθοδολογίας Επιστημονικής Έρευνας & Περιγραφικής Στατιστικής

ΕΝΟΤΗΤΑ Ι

Μεθοδολογία Επιστημονικής Έρευνας.

<i>Μια συνοπτική παρουσίαση</i>	21
<i>Εισαγωγή</i>	21
1. ► Ερευνητικό πλαίσιο, Ερευνητικό δείγμα, Κλίμακες μέτρησης.....	24
1.1 ► Είδη και χαρακτηριστικά επιστημονικής έρευνας.....	24
1.2 ► Επιλογή ερευνητικού θέματος.....	25
1.3 ► Το ερευνητικό δείγμα - συλλογή ερευνητικού υλικού.....	26
1.4 ► Κλίμακες μέτρησης	28
2. ► Μεταβλητές και Μετρήσεις, Εννοιολογικός και Λειτουργικός ορισμός	29
2.1 ► Μεταβλητές.....	29

2.2 ▶ Ανεξάρτητες και Εξαρτημένες μεταβλητές και μετρήσεις, Εννοιολογικός και Λειτουργικός ορισμός.....	32
3. ▶ Ζητήματα Μετρικής Θεωρίας.....	37
3.1 ▶ Αξιοπιστία και Εγκυρότητα	37
3.2 ▶ Ψυχομετρικά εργαλεία.....	38
4.. ▶ Ερευνητικές στρατηγικές – Εσωτερική εγκυρότητα.....	40
4.1. ▶ Σχεδιασμός ερευνητικής στρατηγικής.....	40
4.1.1 Πειραματική ερευνητική στρατηγική (experimental design).....	41
4.1.2 Εκ των υστέρων εξίσωση διαφορικών ομάδων (ex post facto research design).....	41
4.1.3 Συναφειακή ερευνητική στρατηγική (correlational design-studies)	42
4.1.4 Διερευνητική στρατηγική (descriptive studies).....	43
4.2. ▶ Ερευνητικές στρατηγικές και Εξωγενείς παράγοντες.....	43
4.2.1 Πειραματική μεθοδολογία και Εξωγενείς παράγοντες	43
4.2.2 Βασικός πειραματικός σχεδιασμός.....	45
4.2.3 Εξωγενείς παράγοντες	48
4.2.4 Έλεγχος εξωγενών παραγόντων	53
4.2.5 Πειραματικά σχέδια με έναν συμμετέχοντα	58
5. ▶ Εξωτερική εγκυρότητα, Δεοντολογία στην έρευνα και, Συγγραφή ερευνητικής μελέτης	61
5.1 ▶ Εξωτερική εγκυρότητα	61
5.2 ▶ Θέματα δεοντολογίας στην έρευνα	64
5.3 ▶ Κοινοποίηση αποτελεσμάτων εμπειρικής έρευνας.....	68
Επίλογος ενότητας Ι.....	76
Ανακεφαλαίωση ενότητας Ι.....	76

ΕΝΟΤΗΤΑ II

Μέθοδοι Περιγραφικής Στατιστικής. Μια συνοπτική παρουσίαση

Εισαγωγή.....	79
---------------	----

6. ▶ Ερευνητικά δεδομένα – Κατανομές συχνότητας	79
6.1 ▶ Αρχικές (πρωτογενείς) τιμές	79
6.2 ▶ Κατανομές συχνότητας αρχικών τιμών	82
7. ▶ Η Κεντρική Τάση.....	88
7.1 ▶ Μέτρα θέσης	88
7.2 ▶ Πέραν του μέσου όρου, τι άλλο;	90
8. ▶ Η Διασπορά των τιμών	93
8.1 ▶ Στατιστικοί δείκτες διακύμανσης των τιμών.....	93
8.1.1 Το ζήτημα της επιλογής χρήσης δειγματικής (εκτιμητικής) τυπικής απόκλισης	96
8.1.2 Εκ των υστέρων εκτιμητική διόρθωση και εξ αρχής εκτιμητική προσέγγιση	97
8.2 ▶ Χρησιμότητα της τυπικής απόκλισης	99
8.3 ▶ Άλλα μέτρα διασποράς.....	102
9 ▶ Μετατροπές-αντιστροφές αρχικών τιμών και θέματα γραφικής απεικόνισης στατιστικών δεικτών.....	104
10 ▶ Ομαδοποιημένες κατανομές συχνότητας	111
11. ▶ Αριθμητικοί στατιστικοί δείκτες συσχέτισης – Συνάφεια μεταβλητών	117
11.1 ▶ Συνάφεια για δύο αριθμητικές μετρήσεις: δείκτης r κατά <i>Pearson</i>	118
11.2 ▶ Δείκτης μερικής συνάφειας.....	126
11.3 ▶ Δείκτης πολλαπλής συνάφειας.....	127
11.4 ▶ Συνάφεια μεταξύ μιας δίτιμης και μιας αριθμητικής μέτρησης	128
11.5 ▶ Δύο δίτιμες μετρήσεις και η συνάφειά τους	129
11.6 ▶ Δείκτης συνάφειας C	130
11.7 ▶ Δείκτες Rho του <i>Spearman</i> και <i>Tau-b</i> του <i>Kendall</i>	130
11.8 ▶ Δείκτης W του <i>Kendall</i>	132
12 ▶ Μετατροπή αρχικών τιμών σε δευτερογενείς	134
Επίλογος ενότητας II	138
Ανακεφαλαίωση ενότητας II.....	139

ΜΕΡΟΣ II

ΣΤΑΤΙΣΤΙΚΗ ΑΝΑΛΥΣΗ ΕΠΑΓΩΓΩΝ

ΕΝΟΤΗΤΑ III

Στατιστικές Επαγωγές: βασική θεωρία

<i>Εισαγωγή</i>	143
13.. ▶ Θεωρία Συνόλων, Συνδυαστική, Πιθανότητες	147
13.1. ▶ Βασική Θεωρία Συνόλων	147
13.1.1 Ένωση συνόλων	148
13.1.2 Τομή συνόλων	148
13.2. ▶ Συνδυαστική θεωρία	148
13.2.1 Καρτεσιανό γινόμενο	148
13.2.2 Διατάξεις	149
13.2.3 Συνδυασμοί	150
13.3. ▶ Βασική Θεωρία Πιθανοτήτων-ενδεχόμενα	150
13.3.1 Δειγματοχώρος και πειράματα τύχης	150
13.3.2 Το διώνυμο του Newton	152
13.3.3 Τυχαίοι αριθμοί και Ερευνητική Δειγματοληψία	152
14.. ▶ Θεωρητικές ασυνεχείς κατανομές	156
14.1 ▶ Διωνυμική ασυνεχής θεωρητική κατανομή κατά <i>Bernoulli</i>	156
14.2 ▶ Θεωρητική ασυνεχής κατανομή κατά <i>Poisson</i>	162
14.3 ▶ Θεωρητική ασυνεχής Υπεργεωμετρική κατανομή	165
14.4 ▶ Θεωρητική ασυνεχής Αρνητική Διωνυμική κατανομή (<i>Pascal</i>)	167
14.5. ▶ Θεωρητικές ασυνεχείς κατανομές: συγκεντρωτική θεώρηση και εισαγωγή στο λογισμικό 	169
14.5.1 Εφαρμογές για τη θεωρητική Διωνυμική κατανομή	174
14.5.2 Εφαρμογές για τη θεωρητική <i>Poisson</i> κατανομή	175

14.5.3 Εφαρμογές για τη θεωρητική Υπεργεωμετρική κατανομή.....	175
14.5.4 Εφαρμογές για τη θεωρητική Αρνητική Διωνυμική κατανομή	176
15.. ► Θεωρητικές συνεχείς κατανομές.....	176
15.1 ► Δείκτες και παράμετροι	177
15.2 ► Άπειρα τυχαία ισοπληθή δείγματα.....	178
15.3. ► Πιθανοθεωρητική τυχειότητα & συνεχείς κατανομές.....	183
15.3.1 Δειγματοληπτική κατανομή τού Μέσου Όρου	187
15.3.2 Δειγματοληπτική κατανομή τής Διαφοράς δύο μέσων όρων	192
15.3.3 Δειγματοληπτική κατανομή <i>Snedecor</i> (για λόγους διασπορών)	198
15.3.4 Δειγματοληπτική κατανομή τής Διασποράς.....	200
16.. ► Στατιστικές υποθέσεις και Εκτιμήσεις, Στατιστικός Έλεγχος και Σφάλματα, Στατιστικά κριτήρια και Ισχύς τους.....	203
16.1 ► Στατιστικές υποθέσεις	203
16.2. ► Εκτιμητικές διαδικασίες.....	207
16.2.1 Εκτίμηση σημείου	207
16.2.2 Εκτίμηση διαστήματος	209
16.3 ► Στατιστικός έλεγχος- επιλογή στατιστικού κριτηρίου και, στατιστική σημαντικότητα.....	215
16.4 ► Στατιστικά σφάλματα	223
16.5 ► Ισχύς στατιστικών κριτηρίων και παράγοντες που την επηρεάζουν.....	227
Επίλογος ενότητας III.....	234
Ανακεφαλαίωση ενότητας III	235

ΕΝΟΤΗΤΑ IV

Στατιστική Ανάλυση Επαγωγών: μονομεταβλητές μέθοδοι & εφαρμογές

Εισαγωγή.....	237
---------------	-----

17 ▶ Βάση δεδομένων για τις στατιστικές αναλύσεις	237
18. ▶ Μονομεταβλητός στατιστικός έλεγχος χωρίς ανεξάρτητους παράγοντες (ομάδες).....	242
18.1 ▶ Σύγκριση ενός δείγματος με τον πληθυσμό	243
18.2 ▶ Στατιστικός έλεγχος κανονικότητας μιας κατανομής συχνότητας	245
18.3 ▶ Στατιστικός έλεγχος επαρκούς σύγκλισης (<i>goodness of fit</i>) για μια μη-αριθμητική μέτρηση	253
19 ▶ Στατιστικός έλεγχος ανεξαρτησίας (<i>independence</i>) μεταξύ δύο μη-αριθμητικών μετρήσεων	254
20. ▶ Στατιστικός έλεγχος για τη διαφορά μεταξύ δύο ερευνητικών ομάδων-μετρήσεων.....	255
20.1 ▶ Στατιστικός έλεγχος για τη διαφορά μεταξύ δύο ανεξαρτητών ανισοπληθών ομάδων.....	256
20.2 ▶ Στατιστικός έλεγχος για τη διαφορά μεταξύ δύο εξαρτημένων ομάδων ή δύο μετρήσεων	262
20.3 ▶ Μη-παραμετρικό στατιστικό κριτήριο <i>Mann-Whitney U</i>	265
20.4 ▶ <i>Wilcoxon signed-ranks</i> για δύο εξαρτημένα δείγματα	267
21.. ▶ Ανάλυση διακύμανσης.....	269
21.1. ▶ Μονοπαραγοντικός μονομεταβλητός στατιστικός έλεγχος για τρία ή περισσότερα δείγματα συμμετεχόντων: παραμετρική, μονοπαραγοντική ανάλυση διακύμανσης.....	270
21.1.1 <i>Εκ των προτέρων συγκρίσεις</i> (<i>planned comparisons</i>).....	279
21.1.2 <i>Εκ των υστέρων συγκρίσεις</i>	282
21.2 ▶ Μονομεταβλητή μονοπαραγοντική παραμετρική ανάλυση διακύμανσης με έναν εντός των ατόμων παράγοντα (εξαρτημένα δείγματα)	284
21.3 ▶ Μονομεταβλητή παραμετρική ανάλυση διακύμανσης με δύο εντός των ατόμων παράγοντες (εξαρτημένα δείγματα με εμφωλευμένες μετρήσεις)	292
21.4 ▶ Μη-παραμετρικά στατιστικά κριτήρια στη μονομεταβλητή μονοπαραγοντική ανάλυση διακύμανσης (για ανεξάρτητα και εξαρτημένα δείγματα): Τα κριτήρια <i>Kruskal-Wallis H</i> και <i>Friedman</i>	299
21.5 ▶ Διπαραγοντικά ερευνητικά σχέδια μονομεταβλητής παραμετρικής ανάλυσης διακύμανσης για ανεξάρτητα δείγματα	302

21.6 ▶ Τριπαραγοντικά (και πολυπαραγοντικά) ερευνητικά σχέδια μονομεταβλητής παραμετρικής ανάλυσης διακύμανσης για ανεξάρτητα δείγματα	317
21.7. ▶ Υποδείγματα μικτών επιδράσεων στη μονομεταβλητή διπαραγοντική και μονοπαραγοντική παραμετρική ανάλυση διακύμανσης με έναν εντός των ατόμων παράγοντα	320
21.7.1 Υπόδειγμα μικτών επιδράσεων με δύο μεταξύ των ομάδων παράγοντες και έναν εντός των ατόμων παράγοντα	320
21.7.2 Υπόδειγμα μικτών επιδράσεων με έναν μεταξύ των ομάδων παράγοντα και έναν εντός των ατόμων παράγοντα	329
22 ▶ Ανάλυση απλής γραμμικής παλινδρόμησης.....	332
23 ▶ Παραγοντική ανάλυση συνδιακύμανσης	337
Επίλογος ενότητας IV	344
Ανακεφαλαίωση ενότητας IV	345

ΕΝΟΤΗΤΑ V

Πολυμεταβλητά Στατιστικά Υποδείγματα και εφαρμογές

24 ▶ Πολυμεταβλητή προσέγγιση στη στατιστική ανάλυση επαγωγών.....	347
25 ▶ Διερευνητική πολυμεταβλητή ανάλυση· Ανάλυση Κυρίων Συνιστωσών: η πορεία προς την στατιστική επαγωγή	353
26 ▶ Διερευνητική πολυμεταβλητή Ανάλυση Παραγόντων: η πορεία προς την στατιστική επαγωγή συνεχίζεται.....	366
27▶ Επιβεβαιωτική Ανάλυση Παραγόντων: στατιστικές επαγωγές ως προς τη δομή και τις λανθάνουσες μεταβλητές	369
28. ▶ Πολυμεταβλητή ανάλυση διακύμανσης	377
28.1 ▶ Μονοπαραγοντική πολυμεταβλητή ανάλυση διακύμανσης για ανεξάρτητα δείγματα.....	378
28.2 ▶ Πολυμεταβλητή ανάλυση διακύμανσης με ανεξάρτητα και εξαρτημένα δείγματα (<i>multivariate</i> <i>profile analysis with repeated measures</i>)	384

29. ► Πολυμεταβλητή ανάλυση παλινδρόμησης.....	389
29.1 ► Πλήρες υπόδειγμα παλινδρόμησης, η μέθοδος "enter"	392
29.2 ► Ελέγχοντας την καταλληλότητα των μετρήσεων πρόβλεψης: η μέθοδος "stepwise"	394
29.3 ► Πολυμεταβλητή ιεραρχική ανάλυση παλινδρόμησης: η χρήση των δεσμευμένων πιθανοτήτων	398
30 ► Λογιστική ανάλυση παλινδρόμησης: προβλέποντας το γεγονός	405
31. ► Δομικές Αιτιακές Εξισώσεις: μια εισαγωγή και δύο εφαρμογές.....	408
31.1 ► Εισαγωγή.....	408
31.2 ► Δομικές Αιτιακές Εξισώσεις παλινδρομικού τύπου.....	409
31.3 ► Δομικές Αιτιακές Εξισώσεις: η ανάλυση διαμεσολαβητικής επίδρασης (<i>mediation analysis</i>)	413
32 ► Πολυδιάστατη Γεωμετρική Βαθμονόμηση Ομοιοτήτων (<i>Multidimensional Scaling</i>)	417
33 ► Άλλες πολυμεταβλητές τεχνικές: μια εισαγωγή.....	429
Επίλογος ενότητας V.....	435
Ανακεφαλαίωση ενότητας V	436
Επίλογος.....	439
ΒΙΒΛΙΟΓΡΑΦΙΑ	441
ΠΑΡΑΡΤΗΜΑ Ι : ΒΑΣΗ ΔΕΔΟΜΕΝΩΝ ΣΑΕπ.....	449
ΠΑΡΑΡΤΗΜΑ ΙΙ : ΣΤΑΤΙΣΤΙΚΟΙ ΠΙΝΑΚΕΣ.....	457
Ευρετήριο Όρων & Διεθνής Ορολογία.....	475

Πρόλογος

Ο Jacob Cohen έγραφε το 1994 "*The earth is round ($p < .05$)*", δηλαδή "η γη είναι στρογγυλή (αν και δεχόμαστε πιθανότητα στατιστικού σφάλματος, έστω, μικρότερη από 5%)". Ίσως να ακούγεται αστείο ή ακατανόητο στην αρχή, αλλά στην πραγματικότητα ο Cohen προσπάθησε να ασκήσει κριτική στην υπέρ το δέον πολλές φορές χρήση των στατιστικών μεθόδων. Η κριτική του αφορούσε στην κατάχρηση των πιθανοθεωρητικών ορίων ως προς το στατιστικό σφάλμα και στο γεγονός ότι η για πολλές δεκαετίες οριοθέτηση των στατιστικών σφαλμάτων –όπως αυτή προκύπτει μέσω του ελέγχου τής στατιστικής σημαντικότητας για τη μηδενική υπόθεση– θα μπορούσε να θεωρηθεί αυθαίρετη ή τουλάχιστον παραπλανητική. Σε κάποιο ποσοστό, αρκετά χρόνια αργότερα, οι επισημάνσεις του Cohen αλλά και άλλων θεωρητικών τής μεθοδολογίας και στατιστικής για την αναγκαιότητα ικανής θεωρητικής στήριξης πριν τη στατιστική ανάλυση έχουν αρχίσει να αποδίδουν καρπούς καθώς όλο και περισσότεροι ερευνητές στην Ψυχολογία ενδιαφέρονται πλέον για την συστηματική τους ενημέρωση και εκπαίδευση ως προς τη θεωρία και χρήση των στατιστικών τεχνικών προκειμένου να καταλήγουν σε όσο το δυνατόν ασφαλέστερα συμπεράσματα. Αν μάλιστα αυτό είναι επιθυμητό στην περιγραφική στατιστική (*descriptive statistics*) στα πλαίσια τής προσπάθειας κάθε ερευνητή να απεικονίσει ή/και να συνδυάσει δεδομένα εντός των ερευνητικών του δειγμάτων, στην επαγωγική στατιστική (*inferential statistics*) η αναγκαιότητα των παραινήσεων του Cohen είναι αδιαμφισβήτητη και ιδιαιτέρως μεγάλης σημασίας. Τις παραινήσεις αυτές ακολούθησαν και άλλοι θεωρητικοί και ερευνητές στο γνωστικό αντικείμενο της στατιστικής ανάλυσης όπως ο Johnson (1999) ο οποίος συνιστά *κατανόηση* των στατιστικών τεχνικών και *ορθή χρήση τους* σε συνάρτηση με το είδος των ερευνητικών δεδομένων που κάποιος ερευνητής διαθέτει. Άλλοι βεβαίως συγγραφείς (π.χ. Berkson, 2003) ασκούν ηπιότερη κριτική και αποδίδουν στην πιθανότητα στατιστικού σφάλματος πλήρη αποδεικτική ισχύ.

Θα επιστρέφω στον Cohen, ο οποίος, τότε το 1994, πρότεινε στο άρθρο του μεταξύ άλλων τη *λελογισμένη* (*sensible*) και προσεκτική χρήση και *ερμηνεία* των στατιστικών κριτηρίων καθώς και τη *συσσώρευση* των αποδεικτικών στοιχείων γύρω από αυτά μέσω της *περιγραφής* των *ορίων εμπιστοσύνης* (*confidence intervals*), του *μεγέθους επίδρασης* (*effect size*), του *αυστηρού ελέγχου τής μηδενικής υπόθεσης* (*hypothesis testing*), υπό το φως τής *αρχής τής διαψευσσιμότητας* του K. Popper και της *δυνατότητας επανελέγχου* και *επανεύρεσης-επανάληψης* των αρχικών αποτελεσμάτων (*replication*).

"Κινέζικα"; Ακατανόητη ορολογία; **Όχι**, αρκεί να προχωρήσουμε προσεκτικά και συστηματικά στην επεξήγηση των βασικών τεχνικών και μεθόδων τής επαγωγικής στατιστικής ανάλυσης, αλλιώς επονομαζομένης **Στατιστικής Ανάλυσης Επαγωγών (Inferential Statistical Analysis)** και τις εφαρμογές της στις επιστήμες τής συμπεριφοράς, αφού πρώτα κατανοήσουμε επαρκώς τη θεωρία που διέπει, οριοθετεί και καθορίζει τις μεθόδους αυτές. Χρειάζεται, βεβαίως, πρώτα να δούμε προσεκτικά τους επιμέρους θεωρητικούς άξονες και

να τους συνδέσουμε μεταξύ τους, ώστε να είμαστε σε θέση να κατανοήσουμε το γενικότερο σύστημα που θα εφαρμόζουμε. Χρειάζεται υπομονή εκ μέρους των αναγνωστών στην αρχή, ιδιαιτέρως των πρωτόπειρων, όταν τα διάφορα τμήματα του κειμένου δεν θα "δένουν" ακόμη μεταξύ τους. Αργότερα, όλα τα "κομμάτια του παζλ" θα πέσουν σιγά-σιγά στη θέση τους. Οι αναγνώστες χρειάζεται να γνωρίζουν και να θυμούνται ότι στις σελίδες τού παρόντος συγγράμματος, **ο κύριος στόχος θα είναι ένας**: πώς να εντοπίζουμε και να χρησιμοποιούμε-ερμηνεύουμε προσεκτικά και με σωστό τρόπο τις τεχνικές που διαθέτουμε στη στατιστική ανάλυση επαγωγών, πάντα στα πλαίσια τής επαγωγικής στατιστικής θεωρίας καθώς και στα πλαίσια των θεμελιωδών αρχών τής Ψυχομετρίας και της Μεθοδολογίας Επιστημονικής Έρευνας.

Στο σημείο αυτό θα παραθέσω μερικά ιστορικά στοιχεία περί στατιστικής σημαντικότητας, στατιστικών υποθέσεων και *probability values* (p).

Το 1914, δόθηκε στη δημοσιότητα από τον Karl Pearson ένας πίνακας με πιθανοθεωρητικές τιμές. *Ναι*, δεν προέκυψε από το πουθενά ο γνωστός δείκτης συνάφειας και πολλές άλλες στατιστικές μέθοδοι και ο ερευνητής θα πρέπει να αναφέρεται κατά την κοινοποίηση της έρευνάς του στους δημιουργούς των στατιστικών μεθόδων που χρησιμοποιεί ή να παραπέμπει σε συγγράμματα ή άρθρα που αναφέρονται στις μεθόδους αυτές. Στον πίνακα του K. Pearson και σε έναν ακόμη που ακολούθησε από τον Ronald Fisher (F , κατά Fisher, στατιστικό κριτήριο –θα μας απασχολήσει ιδιαιτέρως), η βασική συλλογιστική αφορούσε στην προσπάθεια να ελεγχθεί η μηδενική υπόθεση για τη στατιστική σημαντικότητά της. Αυτή ήταν μόνον η αρχή μιας συζήτησης που κράτησε ως το τέλος τού 20^{ου} αιώνα (και μάλλον δεν έχει λήξει ακόμη). Την πρώτη αυτή προσπάθεια ακολούθησε η συλλογιστική των Neyman-Pearson (ο Egon Pearson, γιος τού Karl) που εισήγαγε και την εναλλακτική υπόθεση. Προσπαθώντας να αντικρούσουν τη συλλογιστική τού Fisher ως προς τις τιμές p που αφορούν στη στατιστική σημαντικότητα της μηδενικής υπόθεσης, στην πραγματικότητα άθελά τους οι Jerzy Neyman & Egon Pearson την παγίωσαν (Kern, 2014). Η συζήτηση που αναπτύχθηκε κατόπιν, αφορούσε στη χρησιμότητα των δεικτών p (*probability values*) που οδηγούν στην απόφαση για τον αν θα πρέπει ή όχι να απορρίψουμε τη μηδενική υπόθεση με κάποια μικρή πιθανότητα στατιστικού σφάλματος. Επίσης, αντιπαράτίθενται άλλες προσεγγίσεις που οι αντιμαχόμενοι την θεώρηση Fisher θεωρούν καταλληλότερες, αρχίζοντας από την απλή περιγραφή μέσω γραφικών παραστάσεων και φτάνοντας στην Μπεϋζιανή πιθανοθεωρία (Thomas Bayes), όπου τα αποτελέσματα μιας έρευνας διαμορφώνουν εκ των υστέρων την πιθανότητα για μία υπόθεση. Συγχρόνως, πολλά σφάλματα γύρω από τη χρήση των p values διαιωνίστηκαν στα αμφιθέατρα (Gigerenzer, 2004), στις δημοσιευμένες εργασίες, στα συγγράμματα, κ.ο.κ., λόγω απλής αναπαραγωγής τους, λόγω του φόβου τού διαφορετικού, λόγω αδιαφορίας ή/και λόγω άγνοιας· ακόμη και λόγω μιας τελετουργικής –με μορφή *ritual* και σχεδόν εμμονικής– προσέγγισης στον υπολογισμό και την ερμηνεία των p values.

Για να κλείσω προς το παρόν τη συζήτηση περί πιθανότητας στατιστικού σφάλματος, το αδιαμφισβήτητο γεγονός είναι ότι ο χειρισμός των p values από

τους ερευνητές δέχθηκε –όχι εντελώς άδικα κατά την άποψή μου– μια ισχυρή επίθεση κατά το μεγαλύτερο μέρος του 20^{ου} αιώνα. Η αναγκαιότητα *ορθής* χρήσης τους μοιάζει είναι το επίμαχο σημείο, καθώς αποτελούν ακρογωνιαίο λίθο και χρησιμοποιούνται παντού στην επαγωγική στατιστική. Μια γενική προσωπική μου θέση είναι ότι τα *p values* δεν μπορούμε να τα αποφύγουμε, αλλά αυτό που μπορούμε να κάνουμε είναι να προσέχουμε τα περιγραφικά στατιστικά στοιχεία ενός δείγματος περισσότερο, πάντα να τα παριστούμε γραφικώς με διάφορους τρόπους, να *επιλέγουμε* εάν θα προχωρήσουμε σε στατιστική ανάλυση επαγωγών καθώς αυτό δεν είναι πάντοτε αυτό αναγκαίο ούτε ορθό, ενώ εάν προχωρήσουμε τελικώς σε στατιστική ανάλυση επαγωγών, να διερευνήσουμε πρώτα το κατά πόσον διαθέτουμε ένα κατάλληλο δείγμα στατιστικού πληθυσμού (αντιπροσωπευτικό αυτού), ώστε να αναγάγουμε στο πληθυσμό τα όσα προκύπτουν σε δειγματικό επίπεδο¹. Πολύ σημαντικά επιστημονικά περιοδικά στην Ψυχολογία όπως το *Memory & Cognition* και αργότερα το *Journal of Personality and Social Psychology* έδωσαν έμφαση ή και απαίτησαν από τους συγγραφείς-ερευνητές παρουσίαση ή έκθεση των περιγραφικών στατιστικών τους στοιχείων *πριν* την εφαρμογή επαγωγικών στατιστικών μεθόδων, όμως αυτές οι προσπάθειες ακόμη δεν έχουν δυστυχώς ούτε παγιωθεί ούτε γενικευθεί σε μεγάλο βαθμό.

Βασικοί άξονες του συγγράμματος

Θα πρέπει να επισημάνω στο σημείο αυτό ότι στο παρόν σύγγραμμα περιέχονται μερικές μόνον από τις πολυπληθείς μεθόδους και τεχνικές τής στατιστικής ανάλυσης επαγωγών, με στόχο αυτές οι μέθοδοι να χρησιμεύσουν τόσο στους πρωτόπειρους ερευνητές όσο και στους περισσότερο ως πολύ έμπειρους, καθώς περιλαμβάνονται πολλά τμήματα με ευρείες προεκτάσεις. Έχει βεβαίως δοθεί ιδιαίτερη έμφαση στην εφαρμογή των επαγωγικών στατιστικών μεθόδων στις Κοινωνικές Επιστήμες και ειδικώς στην Ψυχολογία. Αυτό δεν σημαίνει ότι στην Κοινωνιολογία, Ανθρωπολογία, στη Γεωπονική επιστήμη, στην Ιατρική και σε άλλες επιστήμες, οι μέθοδοι είναι διαφορετικές· απλώς, ο εκάστοτε ερευνητής προσαρμόζει τις τεχνικές αναλόγως των ερωτημάτων που θέτει προς απάντηση μέσω της επιστήμης του. Σημαντικό είναι επίσης να τονίσω την ισχυρή διασύνδεση της Στατιστικής Ανάλυσης Επαγωγών (αλλά και της Περιγραφικής) με τη Μεθοδολογία Επιστημονικής Έρευνας και την Ψυχομετρία στα πλαίσια τής πιθανοθεωρητικής και μαθηματικής τους υπόστασης. Ο ενδιαφερόμενος ερευνητής μπορεί βεβαίως να προστρέξει και σε πολλές άλλες πηγές συμπεριλαμβανομένου και του συγγράμματος *Στατιστική Θεωρία και Εφαρμογές με τον MS-Excel* (Μυλωνάς, 2012 & 2018 –αναθεωρημένη/διορθωμένη έκδοση) με τους στατιστικούς αλγορίθμους *StatsExcel^{GR}*, συνδυάζοντας γνώσεις που θα οδηγήσουν σε μια γενικότερη και πληρέστερη εικόνα των ερευνητικών μεθόδων στις κοινωνικές επιστήμες και ειδικότερα στην Ψυχολογία.

¹ Ασχέτως της ανωτέρω συζήτησης και ακόμη και αν τελικώς έχουμε δικαίωμα να απορρίψουμε μέσω των *probability values* την μηδενική υπόθεση, **δεν έχουμε το δικαίωμα να αναφερθούμε στη μηδενική υπόθεση ως αποδεκτή**, μόνο και μόνο επειδή τα *p values* δεν συνιστούν απόρριψη.

• Το βιβλίο που κρατάτε στα χέρια σας δεν είναι βιβλίο μαθηματικών ούτε περιέχει μαθηματικές αποδείξεις αλλά *εμπεριέχει και βασίζεται σε μαθηματικά θεωρήματα* τα οποία αναφέρονται και εξηγούνται και με θεωρητικό και με εφαρμοσμένο τρόπο, ώστε να γίνει κατανοητή η χρησιμότητά τους. Συνεπώς, το βιβλίο δεν αφορά στα *μαθηματικά* τής πιθανοθεωρητικής προσέγγισης, όμως στηρίζεται στην προσέγγιση αυτή, όπως αυτή ισχύει τόσο για τους τυχαίους αριθμούς, τις τυχαίες μετρήσεις και τις δειγματοληπτικές κατανομές που τις συνοδεύουν, αλλά και για σύνολα τυχαίων αριθμών, δηλαδή για πιθανοθεωρητικώς τυχαία δείγματα μονάδων τού στατιστικού πληθυσμού. Συνεπώς, στις κύριες Ενότητες και κεφάλαια περιλαμβάνονται στοιχεία Πιθανοθεωρίας που αφορούν στη Θεωρία των Τυχαίων Αριθμών και τις Δειγματοληπτικές Κατανομές και, η Θεωρία Δειγματοληπτικών Κατανομών με τα είδη αυτών. • Στο βιβλίο αυτό έχω αφιερώσει –υπό μορφή επιτομής– μία ολόκληρη Ενότητα στην περιγραφική στατιστική ανάλυση, καθώς είναι σημαντικό για τον αναγνώστη να γνωρίζει τουλάχιστον τις βασικές αρχές τής ανάλυσης αυτής μαζί με τους στόχους που υπηρετούν μερικές βασικές μέθοδοι περιγραφικής στατιστικής επεξεργασίας πληροφοριών, π.χ. μέτρα θέσης και διασποράς, βασικοί δείκτες συνάφειας όπως ο r κατά *Pearson*, ο Tau κατά *Kendall*, ο Rho κατά *Spearman*, κ.ο.κ. • Περιλαμβάνεται επίσης μια σύντομη κατά το δυνατόν Ενότητα με τις βασικές έννοιες Μεθοδολογίας Επιστημονικής Έρευνας αμέσως πριν από τις αναγκαίες γνώσεις Περιγραφικής Στατιστικής ανάλυσης. • Το βιβλίο αυτό δεν περιέχει όλες τις μεθόδους και τεχνικές τής στατιστικής ανάλυσης επαγωγών. Κάτι τέτοιο θα απαιτούσε πολλούς τόμους βεβαίως, όχι μόνον έναν, ενώ για κάθε μέθοδο απαιτούνται επίσης πολλοί έτεροι τόμοι όπου οι ειδικές περιπτώσεις κάθε τεχνικής μπορούν να παρουσιασθούν στην πλήρη τους έκταση. • Σε άλλα κεφάλαια ή/και Ενότητες περιλαμβάνονται η πορεία στατιστικής ανάλυσης επαγωγικού χαρακτήρα και ελέγχου στατιστικών υποθέσεων, η στατιστική σημαντικότητα, οι επαγωγικές στατιστικές εκτιμήσεις, τα στατιστικά σφάλματα και, άλλα τμήματα της θεωρίας τής επαγωγικής στατιστικής, πολλά εκ των οποίων έχουν συζητηθεί –αν και λιγότερο αναλυτικά– στο προηγούμενο σύγγραμμά μου του 2012. • Ακολούθως, περιλαμβάνεται η θεωρία και εφαρμογή, πρώτα σε μονο-μεταβλητό επίπεδο ανάλυσης, πλήθους συγκεκριμένων στατιστικών τεχνικών και μεθόδων όπως αυτές χρησιμοποιούνται στην συνήθη ψυχολογική έρευνα. • Στην τελευταία Ενότητα επιχειρώ μια σύντομη κατά το δυνατόν παρουσίαση επιλεγμένων συνθετικών μεθόδων ανάλυσης πολυ-μεταβλητών συστημάτων. • **Τέλος, ως σημειώσω εγκαίρως ότι καλό θα είναι πριν διαβάσετε μια Ενότητα, να διαβάζετε πρώτα τον Επίλογο και την Ανακεφαλαίωση της.** Είναι επίσης σημαντικό ο πρωτόπειρος επιστήμονας να μελετήσει το βιβλίο με τη σειρά των Ενότητων και των κεφαλαίων τους, όχι αποσπασματικά.

Ως προς τις υπολογιστικές διαδικασίες, για πολλές από τις μεθόδους που περιγράφονται στο παρόν βιβλίο, ο ενδιαφερόμενος μπορεί να αποταθεί και στο ανωτέρω σύγγραμμα *Στατιστική Θεωρία και Εφαρμογές με τον MS-Excel*. Εκεί, με βάση τη συλλογιστική τής *εκ των υστέρων εκτίμησης* (θα επεξηγηθεί εκτενώς σε άλλα σημεία τού παρόντος συγγράμματος), ο ενδιαφερόμενος μπορεί να χρησιμοποιήσει τους υπολογιστικούς αλγορίθμους *StatsExcel^{GR}* που

περιλαμβάνονται εκεί. **Στο παρόν όμως σύγγραμμα, ο υπολογισμός των μεγεθών θα στηριχθεί στην εκ των προτέρων στατιστική εκτίμηση** (θα επεξηγηθεί επίσης εκτενώς σε άλλα σημεία), με την παραδοχή ότι η Στατιστική Ανάλυση Επαγωγών αφορά σε *δειγματική έρευνα*, μια παραδοχή που δεν ισχύει υπό όλες τις ερευνητικές συνθήκες, ενώ τα υπολογιζόμενα μεγέθη θα συγκρίνονται και με τους εναλλακτικούς τρόπους προσέγγισης, όπου αυτό είναι αναγκαίο. Σπεύδω να ενημερώσω ότι τα ερευνητικά αποτελέσματα δεν επηρεάζονται παρά απειροελάχιστα από τη διαφορετικού αυτού τύπου προσέγγιση, εκτός από συγκεκριμένες περιπτώσεις που θα συζητηθούν εκτενώς. Σπεύδω επίσης να διευκρινίσω ότι και οι δύο προσεγγίσεις (εκ των προτέρων και εκ των υστέρων εκτίμηση) έχουν τη χρησιμότητά τους, *αναλόγως της δειγματικής ή πληθυσμιακής ιδιότητας του συνόλου συμμετεχόντων σε μια έρευνα*. Θα συζητήσω το σημαντικό αυτό κατά την άποψή μου ζήτημα σε μεγαλύτερο βάθος εν καιρώ.

Βεβαίως, οι μαθηματικής σημειογραφίας υπολογιστικές διαδικασίες δεν είναι ο κύριος στόχος και δεν θα δοθεί μεγάλη έμφαση σε αυτές. Εξάλλου οι κυριότερες έχουν ήδη παρουσιαστεί εξαντλητικά σε πολλά άλλα συγγράμματα ενώ, μέσω των αλγορίθμων του *StatsExcel^{GR}* καθώς των αλγορίθμων άλλων λογισμικών όπως  *SPSS, STATA, BMDP, Minitab, SYSTAT* και άλλων, μπορούμε να υπολογίσουμε τα στατιστικά μεγέθη και να προχωρήσουμε στην ερμηνεία τους χωρίς να εφαρμόζουμε –ορθώς μεν, κοπιωδώς δε– υπολογιστικούς τύπους "με το χέρι" στην αριθμομηχανή, ή ακόμη και "στο χαρτί", ή ακόμη και "στην άμμο".

Καταλήγοντας και αρχίζοντας, μας ενδιαφέρει κυρίως η μελέτη των στατιστικών τεχνικών καθώς και της θεωρίας και μεθοδολογίας τους, ώστε να μπορούν αυτές να εφαρμοσθούν μέσω οποιουδήποτε προγράμματος στατιστικής επεξεργασίας, ενώ κυρίως ενδιαφέρει και η ερμηνεία των αποτελεσμάτων τής στατιστικής ανάλυσης επαγωγών υπό τις τεχνικές που επέλεξα να παρουσιάσω στο παρόν σύγγραμμα –τις κυριότερες μονομεταβλητές και μερικές πολυμεταβλητές επαγωγικές μεθόδους.

Τα βιβλία που αφορούν στο γνωστικό αυτό αντικείμενο θα μπορούσαν -σχεδόν όλα- να έχουν τίτλο "Follow me". Το ίδιο θα μπορούσε να ισχύει και για το παρόν σύγγραμμα, καθώς είναι πράγματι ένα ταξίδι στην επαγωγική στατιστική και στις εφαρμογές της.

*Δεν θα ήθελα όμως να με ακολουθήσετε·
θα ήθελα να με συνοδεύσετε.*

Απρίλιος 2018

Κ. Μυλωνάς