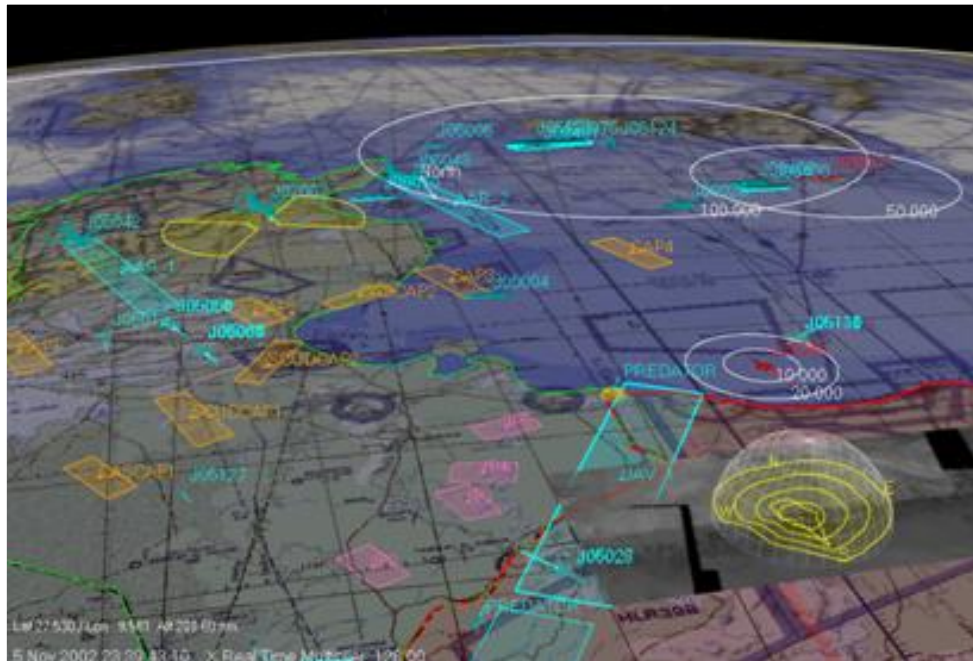




**ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΑΓΡΟΝΟΜΩΝ & ΤΟΠΟΓΡΑΦΩΝ ΜΗΧΑΝΙΚΩΝ
ΤΟΜΕΑΣ ΓΕΩΓΡΑΦΙΑΣ ΚΑΙ ΠΕΡΙΦΕΡΕΙΑΚΟΥ ΣΧΕΔΙΑΣΜΟΥ**

ΘΕΜΑ: ΑΝΑΠΤΥΞΗ ΕΚΠΑΙΔΕΥΤΙΚΟΥ ΥΛΙΚΟΥ ΔΕΙΓΜΑΤΟΛΗΨΙΑΣ ΣΕ ΠΕΡΙΒΑΛΛΟΝ ΓΕΩΓΡΑΦΙΚΩΝ ΣΥΣΤΗΜΑΤΩΝ ΠΛΗΡΟΦΟΡΙΩΝ ARC GIS 10



ΠΤΥΧΙΑΚΗ ΕΡΓΑΣΙΑ ΤΟΥ ΝΙΚΑ Π. ΣΠΥΡΙΔΩΝ

ΕΠΙΒΛΕΠΟΝΤΕΣ ΚΑΘΗΓΗΤΕΣ:

ΒΛΑΣΤΟΣ ΑΘΑΝΑΣΙΟΣ ΚΑΘΗΓΗΤΗΣ Ε.Μ.Π.

ΣΙΟΛΑΣ ΑΓΓΕΛΟΣ ΚΑΘΗΓΗΤΗΣ Ε.Μ.Π.

ΣΑΓΙΑΣ ΙΩΝ ΕΠΙΚΟΥΡΟΣ ΚΑΘΗΓΗΤΗΣ Ε.Μ.Π.

ΦΕΒΡΟΥΑΡΙΟΣ 2012

ΠΡΟΛΟΓΟΣ - ΕΥΧΑΡΙΣΤΙΕΣ

Η παρούσα διπλωματική εργασία εκπονήθηκε στο Εργαστήριο Γεωγραφίας του Τομέα Γεωγραφίας & Περιφερειακού Σχεδιασμού, της Σχολής Αγρονόμων & Τοπογράφων Μηχανικών, κατά τη χρονική περίοδο Φεβρουαρίου 2011 - Φεβρουαρίου 2012.

Στο σημείο αυτό, θα ήθελα να εκφράσω τις θερμές μου ευχαριστίες στον Διευθυντή του Εργαστηρίου Γεωγραφίας Καθηγητή Κ. Κωνσταντίνο Κουτσόπουλο για την τιμή που μου έκανε να μου αναθέσει ένα τόσο ενδιαφέρον θέμα, καθώς και στον επιβλέποντα Δρα Κ. Θωμά Χατζηχρήστο για την πολύτιμη βοήθεια που μου προσέφερε κατά την συγγραφή της μελέτης αυτής.

Επιπλέον θα ήθελα να ευχαριστήσω τον Καθηγητή Κ. Κωνσταντίνο Κασσιό καθώς και τα επιστημονικά μέλη Ε.Ε.ΔΙ.Π. Κ. Δ. Δημητρίου και Κ. Δ. Παπακωνσταντίνου για την πολύτιμη βοήθεια και τις συστάσεις τους για την εύρεση διπλωματικής εργασίας στο συγκεκριμένο τομέα της σχολής.

Ένα μεγάλο ευχαριστώ θα ήθελα να εκφράσω στον απόφοιτο της σχολής μας και μεταπτυχιακό φοιτητή Κ. Στέλιο Σκαράκη για την πολύτιμη βοήθειά του που μου έδωσε με τις γνώσεις του στο λογισμικό Arc Gis 10 καθώς και στον Κ. Ιωσήφ Μπάτσο υπάλληλο της εταιρείας Marathon Data Systems για την υπομονή του και την κατάρτιση που μου έδωσε για την πρώτη επαφή με το καινούριο λογισμικό.

Τέλος, θα ήθελα να ευχαριστήσω τους γονείς για την οικονομική βοήθεια που μου προσέφεραν για την αποπεράτωση των σπουδών μου και τους Κ. Χαράλαμπο Βασιλόπουλο & Αλέξανδρο Ευστρατίου για τη συνεχή υποστήριξη του κατά τη διάρκεια των σπουδών μου.

ΚΑΤΑΛΟΓΟΣ ΠΕΡΙΕΧΟΜΕΝΩΝ

A/A	Τίτλος	Σελίδα
	ΠΡΟΛΟΓΟΣ - ΕΥΧΑΡΙΣΤΙΕΣ	2
	ΠΕΡΙΛΗΨΗ	7
	SUMMARY	9
	ΕΙΣΑΓΩΓΗ	11
1	Γεωγραφία και Γ.Σ.Π. στην Εκπαίδευση	14
1.1	Σύντομη Ιστορία της Εκπαίδευσης της Γεωγραφίας	14
1.2	Εισαγωγή των Ηλεκτρονικών Υπολογιστών στα Μαθήματα Γεωγραφίας	16
1.3	Τα Γεωγραφικά Συστήματα Πληροφοριών στα Μαθήματα Γεωγραφίας	18
1.3.1	Ορισμοί Γεωγραφικών Συστημάτων Πληροφοριών - Ιστορική Αναδρομή	18
1.3.2	Επιλογή και Χρησιμότητα των Γ.Σ.Π. στη Διδακτική της Γεωγραφίας	19
1.3.3	Γεωγραφικά Δεδομένα – Απλές Εφαρμογές Γ.Σ.Π. στα Μαθήματα Γεωγραφίας	20
1.4	Η χρησιμότητα των Γ.Σ.Π. στην Εκπαίδευση	23
2	Βασικές Έννοιες της Δειγματοληψίας και της Στατιστικής	27
2.1	Γενικά	27
2.2	Πλεονεκτήματα της Δειγματοληψίας	30
2.3	Πηγές Σφαλμάτων στη Δειγματοληψία	31
2.4	Σχέδια Δειγματοληψίας	33
2.5	Η Αντιπροσωπευτικότητα του Δείγματος	34

2.6	Η Δειγματοληψία ως Μέρος της Γεωγραφικής Μελέτης	35
2.7	Χωρική Δειγματοληψία	36
2.7.1	Βήματα Δειγματοληπτικής Διαδικασίας	37
2.7.2	Δειγματοληπτικό υπόβαθρο	38
2.7.3	Επιλογή δειγματοληπτικής μεθόδου	39
2.8	Τυχαία Δειγματοληψία	40
2.8.1	Σημειακή	42
2.8.2	Γραμμική	43
2.8.3	Επιφανειακή	44
2.9	Συστηματική Δειγματοληψία	46
2.9.1	Σημειακή	46
2.9.2	Γραμμική	47
2.9.3	Επιφανειακή	48
2.10	Στρωματοποιημένη Δειγματοληψία	50
2.10.1	Στρωματοποιημένη Τυχαία	51
2.10.2	Στρωματοποιημένη Συστηματική	51
2.10.3	Στρωματοποιημένη Συστηματική μη Γραμμική Δειγματοληψία	52
2.11	Εφαρμογή της Κανονικής Κατανομής	53
2.11.1	Μέτρα Ακρίβειας - Διαστήματα Εμπιστοσύνης	56
2.11.2	Εκτιμήσεις Διαστήματος Εμπιστοσύνης για Μετρήσεις	58
2.11.3	Εκτιμήσεις Διαστήματος Εμπιστοσύνης για Απαριθμήσεις	59
2.12	Μέγεθος του Δείγματος και Εκτίμηση του	60
2.13	Έλεγχος Υποθέσεων – Test χ^2	62

3	Γεωγραφικά Συστήματα Πληροφοριών	65
3.1	Γενικά	65
3.2	Τα Μέρη ενός Γ.Σ.Π	67
3.3	Δομές Δεδομένων	69
3.4	Πηγές Δεδομένων για τα Γ.Σ.Π.	72
3.5	Επεξεργασία Δεδομένων – Χωρική Ανάλυση	73
4	Ανάπτυξη Εκπαιδευτικού Υλικού	75
4.1	Γενικά	75
4.2	Περιγραφή Άσκησης	77
4.3	Συνοπτικές οδηγίες για την Εκπόνηση της Άσκησης	79
4.4	Αναλυτικά Βήματα Εφαρμογής Μεθόδων Δειγματοληψίας	83
4.5	Εφαρμογή Δοκιμαστικής Δειγματοληψίας	85
4.6	Εύρεση Απαιτούμενου Μεγέθους Δείγματος	94
4.7	Εφαρμογή Απλής Τυχαίας Δειγματοληψίας	95
4.8	Εφαρμογή Στρωματοποιημένης Τυχαίας Δειγματοληψίας	104
4.9	Συμπεράσματα Άσκησης	140
5	Γενικά Συμπεράσματα	141
	ΠΑΡΑΡΤΗΜΑ	142
	Περιγραφή Άσκησης	142
	Συνοπτικές Οδηγίες για την Εκπόνηση της Άσκησης	144
	Σύντομη Τεχνική Έκθεση	148
	Το Λογισμικό Πακέτο ArcGis 10	151

	Πίνακας Κανονικής Κατανομής	154
	Πίνακας Κρίσιμων Τιμών χ^2	155
	Πίνακας Τυχαίων Αριθμών	156
	Πίνακες Αποτελεσμάτων Δειγματοληψιών	157
	Αρτικόλεξο	159
	ΒΙΒΛΙΟΓΡΑΦΙΑ	160

ΠΕΡΙΛΗΨΗ

Η παρούσα διπλωματική εργασία έχει ως αντικείμενο, την ανάπτυξη εκπαιδευτικού υλικού για την εφαρμογή των δύο κύριων μεθόδων χωρικής δειγματοληψίας (απλή τυχαία, στρωματοποιημένη) σε περιβάλλον γεωγραφικών συστημάτων πληροφοριών Arc Gis 10. Το υλικό αυτό αποτελείται από χωρικά δεδομένα και αναλυτικές οδηγίες εφαρμογής των δειγματοληπτικών μεθόδων.

Βασικό μέρος κάθε μελέτης είναι η συλλογή στοιχείων που θα επεξεργαστούν, θα αναλυθούν και θα αποδώσουν στο τελικό στάδιο τη χρήσιμη πληροφορία. Τα στοιχεία μπορεί να έχουν προέλευση από αρχειακές καταγραφές ή εργασίες πεδίου που περιλαμβάνουν απογραφές, μετρήσεις, παρατηρήσεις και δειγματοληψίες.

Συχνά η δυνατότητα να συλλέξουμε χαρακτηριστικά από το σύνολο των στοιχείων (πληθυσμού) που μας ενδιαφέρουν καθίσταται δύσκολο έως αδύνατο. Όμως με την τεχνική της δειγματοληψίας καθίσταται δυνατό συλλέγοντας χαρακτηριστικά από ένα πεπερασμένο σύνολο στοιχείων (τμήμα πληθυσμού) και εξαγοντας συμπεράσματα για τα χαρακτηριστικά του συνόλου στοιχείων με χρήση ορισμένων ορίων αβεβαιότητας.

Κύριος σκοπός της εργασίας είναι η παρουσίαση και η εφαρμογή των δύο δειγματοληψιών σε ένα χάρτη χρήσεων γης, ώστε λαμβάνοντας δείγματα με τις διάφορες μεθόδους να εκτιμηθούν τα ποσοστά των χρήσεων γης του χάρτη. Η εφαρμογή έχει αναπτυχθεί με συνοπτικό και αναλυτικό τρόπο (κατά βήματα) με στόχο να είναι εύκολα εφαρμόσιμη από τους φοιτητές. Η διαδικασία εφαρμογής δεν απαιτεί εξειδικευμένες γνώσεις στα συστήματα γεωγραφικών πληροφοριών αλλά έχει ως σκοπό να φέρει τους φοιτητές αρχικά σε μια πρώτη επαφή με κάποιες βασικές έννοιες των Γ.Σ.Π..

Στο πρώτο κεφάλαιο γίνεται αρχικά μια αναφορά στη διδακτική της Γεωγραφίας από τα παλαιότερα χρόνια μέχρι σήμερα, όπου οι ηλεκτρονικοί υπολογιστές έχουν γίνει ένα αναπόσπαστο κομμάτι στη διδασκαλία και εκμάθησή της. Ακολουθεί μια εκτενής αναφορά για τη χρήση των Γ.Σ.Π. στα μαθήματα της Γεωγραφίας και σημειώνονται όλα εκείνα που μπορεί να προσφέρει η χρήση τους στην εκπαίδευση γενικότερα.

Στο δεύτερο κεφάλαιο της εργασίας αναπτύσσονται αρχικά οι βασικές έννοιες της δειγματοληψίας και της επαγωγικής στατιστικής, δεύτερον αναλύονται τα πλεονεκτήματα και τα μειονεκτήματα που παρουσιάζονται από τη χρήση δειγμάτων έναντι της καθολικής απογραφικής μεθόδου και τελικά αναφέρονται οι κυριότερες χωρικές δειγματοληπτικές μέθοδοι, τα πλεονεκτήματα – μειονεκτήματά τους, οι τρόποι εφαρμογής τους. Επίσης, αναφέρονται οι σχέσεις που χρησιμοποιούνται για τον υπολογισμό του μεγέθους του δείγματος συναρτήσει του επιλεγόμενου επιπέδου εμπιστοσύνης και περιθωρίου σφάλματος.

Στο τρίτο κεφάλαιο δίνονται οι ορισμοί, τα χαρακτηριστικά και οι βασικές λειτουργίες των γεωγραφικών συστημάτων πληροφοριών.

Στο τέταρτο κεφάλαιο αναπτύσσεται συνοπτικά και αναλυτικά η εφαρμογή των δειγματοληπτικών μεθόδων, (απλή τυχαία, στρωματοποιημένη δειγματοληψία) δίνονται στοιχεία για τα θεματικά επίπεδα του χάρτη – υπόβαθρο και παρουσιάζονται οι λειτουργίες

των εργαλείων που χρησιμοποιούνται μέσω βασικών εντολών. Για την εκτίμηση του μεγέθους του δείγματος και την εξαγωγή χρήσιμων στατιστικών συμπερασμάτων, χρησιμοποιήθηκε λογιστικό φύλλο Excel έτσι ώστε να έχουμε άμεση δυνατότητα αλλαγής παραμέτρων υπολογισμού του μεγέθους του δείγματος βάσει των απαιτήσεων ακρίβειας.

Η εργασία ολοκληρώνεται με την παράθεση συμπερασμάτων (γενικών και ειδικών) που προέκυψαν από την εφαρμογή των δειγματοληπτικών μεθόδων.

SUBJECT: SPATIAL SAMPLING EXERCISE DESIGN IN ARC GIS 10 ENVIRONMENT

SUMMARY

The present diplomatic work has as an object, the development of an educational tool for the application of two main methods of spatial sampling (simple random point sampling, stratified sampling) using Geographic Information Systems Arc Gis 10. This material is constituted by spatial data and analytic directives of application of sampling methods.

The main part of each study is the collection of elements which will be processed, analysed and they will provide the useful information in the final stages. The elements can originate from archival recordings or work of field that includes inventories, measurements, observations and samplings. Often the possibility of collecting characteristics from the total of elements (that is of population) in which we are interested is rendered difficult if not impossible. However, with the technique of sampling it is rendered possible to collect characteristics from a fraction of population and to draw conclusions for the statistic population within limits that are nevertheless somewhat uncertain.

The main aim of work is the presentation and the application of two samplings in a land use map, using different sampling methods to estimate the rates of land uses. The application has been developed with concise and analytic way (via steps) in order to be easily applicable by the students. The process of application does not require specialized knowledge in the systems of geographic information but has the aim to bring the students initially in contact with certain basic significances of G.I.S..

In the first Chapter initially a report is presented regarding the didactics of Geography from earlier years up to present, where computers have become an integral part of teaching and learning Geography. Following this, thorough reference is made to the use of G.I.S. in all courses pertaining to Geography and all benefits G.I.S. has to offer in Geography related courses in general.

The second chapter of the work develops firstly the basic concepts of sampling and inferential statistics, secondly the advantages and the disadvantages of sampling towards board census methodology and finally it mentions the main spatial sampling methods, their advantages – disadvantages and their application methods. The relations that are used for the calculation of the size of the sample in relation to the chosen confidence level and error are also presented.

In the third chapter terms, the characteristics, and the basic operations of geographic information systems are given.

In the fourth chapter the application of sampling methods (simple random point sampling, stratified sampling) is developed concisely and analytically elements for the thematic levels of map – background are given and the operations of tools that are used via basic commands are presented. For the estimate size of sample and the export of useful statistical conclusions,

accountant leaf Excel was used thus enabling us to change parameters of calculation of sample based on the requirements of precision set.

The work is completed by the display of general and special conclusions on the sampling applications.

ΕΙΣΑΓΩΓΗ

Η παρούσα διπλωματική εργασία έχει ως αντικείμενο, την ανάπτυξη εκπαιδευτικού υλικού για την εφαρμογή των δύο κύριων μεθόδων χωρικής δειγματοληψίας (απλή τυχαία, στρωματοποιημένη) σε περιβάλλον γεωγραφικών συστημάτων πληροφοριών Arc Gis 10. Το υλικό αυτό αποτελείται από χωρικά δεδομένα και αναλυτικές οδηγίες εφαρμογής των δειγματοληπτικών μεθόδων και μπορεί να αποτελέσει μέρος μιας σειράς ασκήσεων του υποχρεωτικού προπτυχιακού μαθήματος **Γεωγραφία και Ανάλυση Χώρου** που διδάσκεται στο 6^ο εξάμηνο της Σχολής Αγρονόμων και Τοπογράφων Μηχανικών του ΕΜΠ.

Η δειγματοληψία αποτελεί το θέμα διότι είναι ένα βασικό ερευνητικό εργαλείο της γεωγραφικής επιστήμης καθώς και πολλών άλλων επιστημών (π.χ. Ιατρική, Στατιστική κ.α.). Επίσης, η δειγματοληψία καλύπτει σημαντικό μέρος της ύλης του μαθήματος της Γεωγραφίας και Ανάλυσης του Χώρου ενώ είναι απαραίτητη και για την εκπόνηση της συγκεκριμένης εφαρμογής.

Παλαιότερες εργασίες σχετικά με το θέμα εφαρμογής δειγματοληπτικών μεθόδων στα πλαίσια διδασκαλίας τους, παρουσίαζαν βασικές μεθόδους συλλογής στοιχείων, οι περισσότερες των οποίων ήταν αναλογικά σχεδιασμένες, χωρίς τη χρήση Γ.Σ.Π.. Οι μέθοδοι αυτοί εφαρμόζονταν απευθείας πάνω στον τυπωμένο χάρτη και μόνο μια από αυτές ήταν σχεδιασμένη στο λογισμικό πακέτο Arc Gis 8, και αντιμετώπιζε προβλήματα συμβατότητας στις καινούργιες εκδόσεις του λογισμικού πακέτου. Επιπλέον, άλλες εργασίες που έγιναν με θέμα εφαρμογής τη δειγματοληψία χρησιμοποιούσαν ένα στατιστικό πακέτο (ανεξάρτητο από το GIS το οποίο ενσωματωνόταν) και προσέφερε μια σειρά εργαλείων χωρικής ανάλυσης παρέχοντας έτσι τη δυνατότητα σε μη εξοικειωμένους χρήστες να πραγματοποιήσουν σύνθετες λειτουργίες δειγματοληψίας απλούστερα. Περαιτέρω σε παλαιότερες πτυχιακές εργασίες του μεταπτυχιακού προγράμματος Γεωπληροφορικής της Σχολής Αγρονόμων και Τοπογράφων Μηχανικών έχει συνταχθεί εργαλειοθήκη με βάση την γεννήτρια τυχαίων σημείων (Random Point – in – Polygon Generation). Η εργαλειοθήκη έχει ειδικό σκοπό την εκτέλεση δειγματοληπτικών εργασιών απλών ή στρωματοποιημένων. Κάθε πλήκτρο της εργαλειοθήκης εκτελεί ένα συγκεκριμένο σενάριο κώδικα (script) που αφορά κάποιο στάδιο της δειγματοληψίας (π.χ. υπολογισμός απαιτούμενου αριθμού σημείων κ.τ.λ.). Μόλις εκτελεσθεί η δοκιμαστική δειγματοληψία η εργαλειοθήκη αποδίδει αυτόματα τα απαιτούμενα σημεία για την εκτέλεση της απλής σημειακής δειγματοληψίας. Το γεγονός αυτό αποτελεί σημαντικό πλεονέκτημα στην χρήση της συγκεκριμένης εργαλειοθήκης.

Τέλος, εκμεταλλευόμενοι στο μέγιστο δυνατό όριο τις δυνατότητες του λογισμικού πακέτου υλοποιήσαμε τις μεθόδους δειγματοληψίας μόνο με την χρήση εντολών εγγενώς εγκατεστημένων στο Arc Gis 10. Σκοπός αυτής της προσπάθειας είναι η απαλλαγή των μελλοντικών χρηστών από την αναζήτηση εξωτερικών προγραμμάτων. Έτσι όχι μόνον κερδίζεται χρόνος από την αποφυγή της αναζήτησης αλλά αποφεύγονται κατά το μέγιστο δυνατόν πιθανές ασυμβατότητες σε μελλοντικές εκδόσεις του Arc Gis.

Οποιαδήποτε προσέγγιση – επέμβαση στο χώρο συνδέεται άμεσα με την έννοια της γεωγραφικής μελέτης. Η γεωγραφική μελέτη απαρτίζεται από τον καθορισμό του προβλήματος, τη διαδικασία από στοιχεία σε πληροφορία και τα συμπεράσματα. Η διαδικασία από στοιχεία σε πληροφορία δημιουργεί και συνδυάζει την απαραίτητη πληροφορία για την επίλυση του προβλήματος, όπως αυτό ορίστηκε και οριοθετήθηκε στο

πρώτο βήμα της μελέτης, τον καθορισμό του προβλήματος. Τελικά η υλοποίηση του στόχου της μελέτης, οι εναλλακτικές προσεγγίσεις και οι απαντήσεις στα ερωτήματα που τίθενται στο πρώτο βήμα αποτελούν τα συμπεράσματα της μελέτης, τα οποία με τη σειρά τους μπορούν να λειτουργήσουν σαν νέα δεδομένα για επανεξέταση, επέκταση, αναθεώρηση του υπάρχοντος προβλήματος (αναδραστικός μηχανισμός).

Βασικό μέρος κάθε μελέτης είναι η συλλογή στοιχείων που θα επεξεργαστούν, θα αναλυθούν και θα αποδώσουν στο τελικό στάδιο τη χρήσιμη πληροφορία. Τα στοιχεία μπορεί να έχουν προέλευση από αρχειακές καταγραφές ή εργασίες πεδίου που περιλαμβάνουν απογραφές, μετρήσεις, παρατηρήσεις και δειγματοληψίες.

Συχνά η δυνατότητα να συλλέξουμε χαρακτηριστικά από το σύνολο των στοιχείων (πληθυσμού) που μας ενδιαφέρουν καθίσταται δύσκολο έως αδύνατο. Όμως με την τεχνική της δειγματοληψίας καθίσταται δυνατό συλλέγοντας χαρακτηριστικά από ένα πεπερασμένο σύνολο στοιχείων (τμήμα πληθυσμού) και εξάγοντας συμπεράσματα για τα χαρακτηριστικά του συνόλου στοιχείων με χρήση ορισμένων ορίων αβεβαιότητας.

Κύριος σκοπός της εργασίας είναι η παρουσίαση και η εφαρμογή των δύο δειγματοληψιών σε ένα χάρτη χρήσεων γης, ώστε λαμβάνοντας δείγματα με τις διάφορες μεθόδους να εκτιμηθούν τα ποσοστά των χρήσεων γης του χάρτη. Η εφαρμογή έχει αναπτυχθεί με συνοπτικό και αναλυτικό τρόπο (κατά βήματα) με στόχο να είναι εύκολα εφαρμόσιμη από τους φοιτητές. Η διαδικασία εφαρμογής δεν απαιτεί εξειδικευμένες γνώσεις στα συστήματα γεωγραφικών πληροφοριών αλλά έχει ως σκοπό να φέρει τους φοιτητές αρχικά σε μια πρώτη επαφή με κάποιες βασικές έννοιες των Γ.Σ.Π..

Στο πρώτο κεφάλαιο γίνεται αρχικά μια αναφορά στη διδακτική της Γεωγραφίας από τα παλαιότερα χρόνια μέχρι σήμερα, όπου οι ηλεκτρονικοί υπολογιστές έχουν γίνει ένα αναπόσπαστο κομμάτι στη διδασκαλία και εκμάθησή της. Ακολουθεί μια εκτενής αναφορά για τη χρήση των Γ.Σ.Π. στα μαθήματα της Γεωγραφίας και σημειώνονται όλα εκείνα που μπορεί να προσφέρει η χρήση τους στην εκπαίδευση γενικότερα.

Στο δεύτερο κεφάλαιο της εργασίας αναπτύσσονται αρχικά οι βασικές έννοιες της δειγματοληψίας και της επαγωγικής στατιστικής, δεύτερον αναλύονται τα πλεονεκτήματα και τα μειονεκτήματα που παρουσιάζονται από τη χρήση δειγμάτων έναντι της καθολικής απογραφικής μεθόδου και τελικά αναφέρονται οι κυριότερες χωρικές δειγματοληπτικές μέθοδοι, τα πλεονεκτήματα – μειονεκτήματά τους, οι τρόποι εφαρμογής τους. Επίσης, αναφέρονται οι σχέσεις που χρησιμοποιούνται για τον υπολογισμό του μεγέθους του δείγματος συναρτήσει του επιλεγόμενου επιπέδου εμπιστοσύνης και περιθωρίου σφάλματος.

Στο τρίτο κεφάλαιο δίνονται οι ορισμοί, τα χαρακτηριστικά και οι βασικές λειτουργίες των γεωγραφικών συστημάτων πληροφοριών.

Στο τέταρτο κεφάλαιο αναπτύσσεται συνοπτικά και αναλυτικά η εφαρμογή των δειγματοληπτικών μεθόδων, (απλή τυχαία, στρωματοποιημένη δειγματοληψία) δίνονται στοιχεία για τα θεματικά επίπεδα του χάρτη – υπόβαθρο και παρουσιάζονται οι λειτουργίες των εργαλείων που χρησιμοποιούνται μέσω βασικών εντολών. Για την εκτίμηση του μεγέθους του δείγματος και την εξαγωγή χρήσιμων στατιστικών συμπερασμάτων, χρησιμοποιήθηκε λογιστικό φύλλο Excel έτσι ώστε να έχουμε άμεση δυνατότητα αλλαγής παραμέτρων υπολογισμού του μεγέθους του δείγματος βάσει των απαιτήσεων ακρίβειας.

Η εργασία ολοκληρώνεται με την παράθεση συμπερασμάτων (γενικών και ειδικών) που προέκυψαν από την εφαρμογή των δειγματοληπτικών μεθόδων.

ΚΕΦΑΛΑΙΟ 1: Γεωγραφία και Γ.Σ.Π. στην Εκπαίδευση

1.1 Σύντομη Ιστορία της Εκπαίδευσης της Γεωγραφίας

Πηγαίνοντας πενήντα χρόνια πίσω οι γεωγραφικές πληροφορίες στην εκπαίδευση απεικόνιζαν την εγκυκλοπαιδική φύση του μαθήματος της Γεωγραφίας. Η παγκόσμια κατάσταση είχε μια έννοια στασιμότητας και διάρκειας. Το καουτσούκ προερχόταν από τη Μαλαισία, ο καφές από τη Ν. Αμερική κ.ο.κ.. Στην περίπτωση του Ηνωμένου Βασιλείου ένα μεγάλο μέρος από τις εξαγωγές κατέληγε στα έθνη της Κοινοπολιτείας, σε πολλές χώρες το Ηνωμένο Βασίλειο ήταν ο μεγαλύτερος προμηθευτής βιομηχανικών αγαθών. Στα ενδότερα του Βασιλείου, η κεντρική χώρα ήταν η βιομηχανικά ζωτική και κεντρική περιοχή, στη βορειοδυτική και βορειανατολική χώρα όπως επίσης και στη Σκωτία κατασκευάζονταν πλοία για τον ακμάζοντα Βρετανικό στόλο ενώ στη Νότια Ουαλία βρίσκονταν τα μεγαλύτερα μεταλλεία άνθρακα.

Η χαρτογραφία απεικόνιζε τη φαινομενική μονιμότητα και όριζε τη λεπτομέρεια των συνοριακών γραμμών, των μεγάλων ποταμών, τη θέση μεγάλων πόλεων τα είδη και τη θέση των βιομηχανιών και των καλλιεργειών. Στις αίθουσες διδασκαλίας το ενδιαφέρον εστιαζόταν, από συνήθεια κυρίως στο να εξασφαλιστεί η γνώση των παραπάνω και μόνο σε εξαιρετικές περιπτώσεις υπήρχαν αρχές και διαδικασίες ανάλυσης των λόγων πίσω από γεγονότα που σχετίζονταν μεταξύ τους. Καθημερινά θέματα όπως το περιβάλλον και η μόλυνση του δεν αποτελούσαν θέματα εκείνη την εποχή. Στις περισσότερες περιπτώσεις η ιδέα της επίσκεψης στο πεδίο και του βιώματος της γεωγραφίας από τους σπουδαστές δεν έβρισκε εφαρμογή.

Η γεωγραφία αρχικά εντυπώνονταν στο νου των σπουδαστών μέσω των διαλέξεων και του πίνακα της αίθουσας. Μερικές δεξιότητες ήταν ιδιαίτερα ανεπτυγμένες. Τυπικά οι καθηγητές γεωγραφίας μπορούσαν να σχεδιάσουν έναν τέλειο κύκλο στον πίνακα οπότε υπήρχαν παγκόσμια θέματα συζήτησης. Ένας μαύρος και ένας άσπρος κόσμος ήταν τα μόνα χρώματα που χρησιμοποιούσαν στις εκπαιδευτικές αίθουσες. Οι πηγές γεωγραφικών πληροφοριών ήταν κάποια προκαθορισμένα κείμενα και χάρτες που περιορίζονταν σε έναν της χώρας και έναν παγκόσμιο.

Στις αρχές της δεκαετίας του 1970 άρχισε η ανάπτυξη πρακτικής εξάσκησης με εργασίες και έρευνα στο πεδίο ως μέρος της διδακτικής γεωγραφίας. Αυτό είχε ως αποτέλεσμα τη μετατροπή του μαθήματος της γεωγραφίας από ένα στατικό και συνηθισμένο μάθημα, σε ένα μάθημα με στοιχεία δράσης και ανάλυσης. Οι σπουδαστές που παρακολουθούσαν τα μαθήματα γεωγραφίας, συνειδητοποίησαν ότι η πραγματικότητα δεν χωράει σε ένα καλό βιβλίο. Η συζήτηση με ειδικούς στο πεδίο αποκάλυψε σύγκρουση ενδιαφερόντων και διαφορές οι οποίες δεν προσδιορίζονταν από την ανάγνωση ενός απλού χάρτη ή ενός απλού κειμένου. Εθνικές και παγκόσμιες οικονομικές, πολιτικές και κοινωνικές μελέτες αποτελούσαν πια μέρος της ανάλυσης του μαθήματος της γεωγραφίας.

Μετέπειτα, οι καθηγητές απελευθερωμένοι από τη ρουτίνα της εκπαιδευτικής αίθουσας και ενδυναμωμένοι από την άφιξη του φωτοαντιγραφικού, άρχισαν να παράγουν το δικό τους πληροφοριακό υλικό το οποίο περιελάμβανε παιχνίδια, παίξιμο ρόλων (π.χ. μελετητών) και προσομοιώσεις τα οποία ανέπτυξαν την εξεταστική και αναλυτική προσέγγιση του μαθήματος.

Οι πληροφορίες για παγκόσμιου ενδιαφέροντος θέματος ήταν εύκολες να βρεθούν και τα βιβλία που μέχρι εκείνη τη στιγμή αποτελούσαν τη μεγαλύτερη πηγή πληροφοριών άρχισαν να εξαφανίζονται.

Στις εκπαιδευτικές αίθουσες αν και ακόμα υπήρχε το προκαθορισμένο κείμενο και οι κλασικοί χάρτες – άτλαντες, καινούργιες πηγές γνώσεων άρχισαν να εισέρχονται σε αυτές όπως άρθρα εφημερίδων, πανεπιστημιακές εργασίες και έρευνες, έρευνες πεδίου, κασέτες ήχου και τηλεόραση. Το χρώμα εισήλθε στην περιγραφή του κόσμου όπως στην πραγματικότητα. Η εποχή των πληροφοριών στη γεωγραφία άρχισε.

1.2 Εισαγωγή των Ηλεκτρονικών Υπολογιστών στα Μαθήματα Γεωγραφίας

Τα τελευταία είκοσι χρόνια ένα μεγάλο εύρος υπολογιστικών πόρων αναπτύχθηκε ειδικά για εκπαιδευτικούς λόγους, ενώ άλλοι σχεδιάστηκαν για διαφορετικούς σκοπούς οι οποίοι με τη σειρά τους μπήκαν στην υπηρεσία της εκπαίδευσης. Οι πόροι αυτοί ποικίλλουν από το βαθμό γενίκευσης. Μερικοί για παράδειγμα, έχουν γεωγραφικά θέματα και ικανότητες ενσωματωμένες, ενώ άλλοι όχι, κάποιοι είναι σχεδιασμένοι ειδικά για εκπαιδευτικό σκοπό και υιοθετούν μια ειδική εκπαιδευτική νοοτροπία, ενώ άλλοι πάλι όχι, μερικοί μπορεί να χρησιμοποιούνται με ένα στενά ορισμένο τρόπο ενώ άλλοι είναι σχεδιασμένοι για μια πιο ευέλικτη χρήση. Παρακάτω παρουσιάζονται οι κυριότεροι υπολογιστικοί πόροι που χρησιμοποιούνται στα μαθήματα της Γεωγραφίας:

- **Εξάσκηση (drive and practice)**: Μηχανιστικές ασκήσεις σε υπολογιστή, σχεδιασμένες για να παρέχουν πρακτική εξάσκηση και ενίσχυση των σπουδαστών σε συγκεκριμένες γνώσεις και επιδεξιότητες.
- **Υπολογιστικά υποβοηθούμενη διδασκαλία (Computer assisted instruction)**: Ένα είδος διδασκαλίας το οποίο περιλαμβάνει από την επαναλαμβανόμενη προγραμματισμένη μάθηση (programmed learning) μέχρι τα λογικά διδακτικά συστήματα (intelligent tutoring systems) βασισμένα σε αρχές ανεπτυγμένες με τεχνητή νοημοσύνη.
- **Υπολογιστικά διαχειριζόμενη διδασκαλία (Computer managed instruction)**: Λογισμικό που διαχειρίζεται συμβατικό ή βασισμένο σε υπολογιστή διδακτικό υλικό. Διατηρεί αρχείο για την πρόοδο των σπουδαστών και την απόδοση τους βάση τακτικών τεστ, προτείνει μελλοντικές εκπαιδευτικές δραστηριότητες και παρουσιάζει στους διδάσκοντες συνοπτικές πληροφορίες.
- **Λογισμικό παρουσίασης (Presentation software)**: Λογισμικό (συνήθως γραφιστικό) σχεδιασμένο για χρήση ως εργαλείο επίδειξης ή παρουσίασης.
- **Παιχνίδια σε υπολογιστή (Computer – based games)**: Προγράμματα τα οποία παρουσιάζουν προσομοιωμένες καταστάσεις, στις οποίες οι σπουδαστές έχουν ένα ρόλο και αλληλεπιδρούν με τον υπολογιστή.
- **Εργαλεία δημιουργίας μοντέλων (Model – building tools)**: Λογισμικό (π.χ. γλώσσες προγραμματισμού, προγράμματα λογιστικού φύλλου, εργαλεία γραφικών μοντέλων κ.α.) που χρησιμοποιούνται για τη δημιουργία μοντέλων του πραγματικού κόσμου.
- **Λογισμικό βάσης δεδομένων (Database software)**: Λογισμικό που επιτρέπει τη δημιουργία βάσης δεδομένων και τη διαχείριση αυτής, ώστε τα γεωγραφικά δεδομένα να είναι σε ηλεκτρονική μορφή.
- **Λογισμικό διαχείρισης γεωγραφικών δεδομένων (Geographical information handling software)**: Προγράμματα που χρησιμοποιούνται για την εξέταση και έρευνα γεωγραφικών δεδομένων χρησιμοποιώντας καθιερωμένες μεθόδους χωρικής ανάλυσης

και απεικόνισης (π.χ. λογισμικό Θεματικής Χαρτογραφίας, λογισμικό Γεωγραφικών Συστημάτων Πληροφοριών, λογισμικό Ψηφιακής Τηλεπισκόπησης).

1.3 Τα Γεωγραφικά Συστήματα Πληροφοριών στα Μαθήματα Γεωγραφίας

1.3.1 Ορισμοί Γεωγραφικών Συστημάτων Πληροφοριών – Ιστορική Αναδρομή

Πολλοί είναι οι ορισμοί που κατά καιρούς έχουν δοθεί για τα Γεωγραφικά Συστήματα Πληροφοριών (GIS).

Σύμφωνα με τον Parker τα ΣΓΠ είναι μια << τεχνολογία πληροφοριών η οποία αποθηκεύει, αναλύει και παρουσιάζει ταυτόχρονα χωρικές και μη χωρικές πληροφορίες >> (Parker, 1988) ενώ σύμφωνα με τον Aronoff GIS αποτελεί << οποιοδήποτε, καθοδηγούμενο από τον άνθρωπο ή βασισμένο σε υπολογιστή, σύνολο διαδικασιών που χρησιμοποιείται για την αποθήκευση και διαχείριση γεωγραφικών δεδομένων >>.

Ο Χατζόπουλος Ι. ορίζει τα ΣΓΠ ως << την επιστήμη, την τέχνη και την τεχνολογία για τη διαχείριση γεωγραφικών δεδομένων και πληροφοριών >> (Πηγή: Χατζόπουλος Ι. 2006).

Ένας συνοπτικός ορισμός που θα μπορούσε να δοθεί είναι πως τα GIS είναι ο τομέας εκείνος της γεωγραφικής επιστήμης όπου χρησιμοποιούνται οι νέες τεχνολογίες με σκοπό την οπτικοποίηση και ανάλυση των γεωγραφικών δεδομένων.

Τα Γεωγραφικά Πληροφοριακά Συστήματα έχουν πλήθος εφαρμογών στη Γεωγραφία, την υγεία, τη πυροσβεστική, το κτηματολόγιο, τη πολεοδομία, το περιβάλλον, την εγκληματολογία, τα κτηματομεσιτικά, τον τουρισμό, την αρχαιολογία κ.α.

Η ανάπτυξη των GIS είναι παράλληλη με την ανάπτυξη των υπολογιστικών μηχανών και της τεχνολογίας λογισμικού. Διέπονται από αυξανόμενα πολύπλοκες αρχές και κανόνες με τάση να αποτελέσουν στο μέλλον ξεχωριστή επιστήμη.

Το πρώτο πραγματικό ΣΓΠ δημιουργήθηκε το 1962 στην Οτάβα του Καναδά από τον Δρ. Roger Tomlinson και ονομάστηκε << Canada Geographic Information System (CGIS) >> και χρησιμοποιήθηκε για ανάλυση δεδομένων χρήσεων γης σε κλίμακα 1:15000. Το CGIS σταμάτησε να χρησιμοποιείται στα μέσα της δεκαετίας του 1990.

Το 1964 ο Howard T. Fischer δημιούργησε το Laboratory for Computer Graphics and Spatial Analysis στο Harvard Graduate School of Design, όπου και αναπτύχθηκε σειρά θεωρητικών προσεγγίσεων στην ανάλυση χωρικών δεδομένων. Από τα μέσα της δεκαετίας του 70 το εργαστήριο ανέπτυξε σειρά λογισμικών εφαρμογών όπως τα “ Symap “ “ Grid “ “ Odyssey “ που χρησιμοποιήθηκαν ως βάσεις για την ανάπτυξη εμπορικών και μη εφαρμογών ΣΓΠ.

Στις αρχές της δεκαετίας του 1980 οι εταιρείες M&S Computing (γνωστή αργότερα ως Intergraph), Environmental Systems Research Institute (ESRI) και (CARIS), ανέπτυξαν εμπορικές εφαρμογές ΓΣΠ με πολλά στοιχεία δανεισμένα από το CGIS και συνδυάζοντας στοιχεία από την πρώτη γενιά ΓΣΠ με τη δεύτερη, οργανώνοντας τα δεδομένα σε βάσεις δεδομένων. Παράλληλα, το U.S. Army Corp of Engineering Research Laboratory (USA – CERL) δημιούργησε τον πρόγονο του GRASS, ένα δωρεάν πρόγραμμα ανοικτού

κώδικα. Τα περισσότερα προγράμματα στις αρχές του 80 έτρεχαν σε περιβάλλον UNIX, κάτι το οποίο τη σημερινή περίοδο και ιδιαίτερα από το 2000 και έπειτα έχει αλλάξει άρδην. Στην πλειονότητα τους, προγράμματα ΓΣΠ τρέχουν σε όλα σχεδόν τα λειτουργικά περιβάλλοντα με έμφαση στα Windows.

1.3.2 Επιλογή και Χρησιμότητα των Γ.Σ.Π. στη Διδακτική της Γεωγραφίας

Τα Γεωγραφικά Συστήματα Πληροφοριών αποτελούν ένα ιδανικό μέσο για τη διδασκαλία θεμάτων γεωγραφίας δίνοντας τη δυνατότητα στους σπουδαστές να αναπτύξουν σημαντικές δεξιότητες στην επίλυση προβλημάτων. Τα προηγούμενα χρόνια περιοριζόταν σε επαγγελματίες με πρόσβαση σε υψηλού επιπέδου λογισμικό και υπολογιστικό εξοπλισμό. Σήμερα τα Γ.Σ.Π. είναι προσιτά και χρησιμοποιούνται ευρέως από σπουδαστές ακόμη και στη δευτεροβάθμια εκπαίδευση (κυρίως στο εξωτερικό). Τα Γ.Σ.Π. όταν χρησιμοποιούνται ως διδακτικό εργαλείο, έχουν ένα μεγάλο αριθμό πλεονεκτημάτων σε σύγκριση με τα παραδοσιακά διδακτικά εργαλεία. Αυτά τα πλεονεκτήματα περιλαμβάνουν τα παρακάτω στοιχεία:

- **Οπτικοποίηση Δεδομένων (Data Visualization)**: Έρευνες βασισμένες σε Γ.Σ.Π., συμπεραίνουν ότι επιτρέπουν στους σπουδαστές να αναγνωρίσουν και να χαρακτηρίσουν σχέσεις και αλληλεπιδράσεις, διαχειρίζοντας επιδέξια πολλαπλές οπτικές απεικονίσεις των δεδομένων, περιλαμβανομένων δυναμικών και προσαρμοζόμενων χαρτών, πινάκων, γραφημάτων και κινουμένων σχεδίων.
- **Ανάλυση Δεδομένων (Data Analysis)**: Το αναλυτικό εργαλείο ενός Γ.Σ.Π. παρέχει τη δυνατότητα στους σπουδαστές να ποσοτικοποιούν σχέσεις και αλληλεπιδράσεις εντός και μεταξύ ομάδων χωρικών δεδομένων, χρησιμοποιώντας λειτουργίες βάσεων δεδομένων, στατιστική ανάλυση και εφαρμογές χωρικής επικοινωνίας.
- **Ενσωμάτωση Πολυμέσων (Multimedia Integration)**: Άλλες μορφές ψηφιακής πληροφορίας, συμπεριλαμβανομένων κινουμένων σχεδίων, βίντεο, ηχητικών αρχείων (audio) και ψηφιακής φωτογραφίας, μπορούν να συμπεριληφθούν στις δραστηριότητες των Γ.Σ.Π., εμπλουτίζοντας τα σε μεγάλο βαθμό και επεκτείνοντας τη διδακτική δυναμικότητά τους.
- **Γνώση Τεχνολογίας και Μεταβιβάσιμες Ικανότητες (Technology Literacy and Transferable Skills)**: Η χρήση των Γ.Σ.Π. προάγει τη γνώση της τεχνολογίας και εφοδιάζει τους σπουδαστές με ικανότητες που μπορούν να προσαρμοσθούν στη δική τους έρευνα, σε άλλα μαθήματα και στη γενικότερη σφαίρα ανταγωνισμού.

Οι βασισμένες σε Γ.Σ.Π. εκπαιδευτικές μονάδες, έχουν τη δυνατότητα να επιδρούν στην εκπαίδευση των σπουδαστών ενισχύοντας τις αρχές τους, μέσω της εξερεύνησης και της βελτίωσης τους στη δυνατότητα επίλυσης προβλημάτων, της οπτικοποίησης δεδομένων και της ανάπτυξης υπολογιστικών ικανοτήτων.

Τα Γ.Σ.Π. μπορούν να εξασφαλίσουν ένα πολύτιμο επίκεντρο προσοχής στη γεωγραφία. Παρέχουν τη δυνατότητα στους σπουδαστές να συνενώσουν φαινομενικώς ασύνδετα δεδομένα σε ένα κατανοητό σύνολο. Με τα Γ.Σ.Π. οι σπουδαστές αποκτούν την ικανότητα να

βλέπουν με ποιο τρόπο τα γεωγραφικά εργαλεία, οι τεχνικές και οι ικανότητες μπορούν να χρησιμοποιηθούν για τη συλλογή, διαχείριση, συνένωση, ανάλυση και εμφάνιση γεωγραφικών δεδομένων – πληροφοριών. Οι σπουδαστές θα πρέπει να διακατέχονται από την ικανότητα να ξέρουν γιατί, πώς και πότε να χρησιμοποιούν συγκεκριμένες τεχνικές και αρχές, καθώς και να αντιλαμβάνονται τι συμβαίνει στα γεωγραφικά δεδομένα όταν αυτά διαχειρίζονται και αναλύονται. Επιπλέον τα Γ.Σ.Π. έχουν τη δυνατότητα να φέρνουν στο προσκήνιο τη θεωρία της γεωγραφίας μέσω πρακτικών εφαρμογών, όπως για παράδειγμα τη χρήση συντεταγμένων, δεδομένων χωρικής ανάλυσης και χαρτών. Είναι σημαντικό για τους σπουδαστές να μπορούν να χρησιμοποιούν γεωγραφικές ικανότητες στην πράξη, παρά να μένουν στο άκουσμα για αυτές.

Τα Γ.Σ.Π. αφήνουν τους σπουδαστές να κάνουν πλήρη χρήση των γνώσεων και τεχνικών που έχουν διδαχτεί σε άλλα μαθήματα όπως για παράδειγμα τα Μαθηματικά και τα μαθήματα των ηλεκτρονικών υπολογιστών. Η διαθεσιμότητα των ηλεκτρονικών υπολογιστών και τα απλά λογισμικά Γ.Σ.Π. μπορούν να αξιοποιήσουν τις γνώσεις ηλεκτρονικών υπολογιστών για εφαρμογή στο περιβάλλον των γεωγραφικών εφαρμογών. Τα Γ.Σ.Π. είναι επίσης σημαντικά στις περιβαλλοντικές επιστήμες οι οποίες αποτελούν ένα πεδίο ενδιαφέροντος στο μέλλον της Γεωγραφίας. Ακόμη τα Γ.Σ.Π. θα μπορούσαν να βοηθήσουν στην ανάπτυξη λογικής σκέψης σε διαδικασίες ερώτησης – απάντησης και λύσης προβλημάτων.

Η παρουσίαση των Γ.Σ.Π. σε ένα πρώιμο στάδιο της εκπαίδευσης, είναι δυνατόν να θέσει τις βάσεις για τη διδασκαλία της γεωγραφίας με τη χρήση χωρικών δεδομένων. Αυτό θα επενεργήσει ευεργετικά τόσο τους σπουδαστές, όσο και τους καθηγητές στην Ανώτατη Εκπαίδευση.

Εν κατακλείδι, οι τεχνικές των Γ.Σ.Π. μπορούν να εφοδιάσουν τον σπουδαστή με μια διορατικότητα στο πως τα γεωγραφικά δεδομένα μπορούν να χρησιμοποιηθούν ως βάση για το σχεδιασμό και τη λήψη αποφάσεων.

1.3.3 Γεωγραφικά Δεδομένα – Απλές Εφαρμογές Γ.Σ.Π. στα Μαθήματα Γεωγραφίας

Τα Γεωγραφικά δεδομένα αποτελούν μια ειδική κατηγορία δεδομένων τα οποία κατανέμονται στο χώρο και μεταβάλλονται στο χρόνο. Για την περιγραφή της πραγματικότητας γίνεται η θεώρηση ότι αυτή συντίθεται από ένα σύνολο διακριτών αλλά παράλληλα αλληλοσυσχετιζόμενων μονάδων που καλούνται οντότητες. Οντότητα ορίζεται κάθε μονάδα ή αντικείμενο με φυσική ή εννοιολογική υπόσταση (Πηγή: Στεφανάκης 2003).

Τα γεωγραφικά δεδομένα αποτελούν για τα Γ.Σ.Π. ότι τα καύσιμα για τα αυτοκίνητα. Για να έχουν λόγο ύπαρξης τα Γ.Σ.Π. χρειάζονται καύσιμα στη μορφή γεωγραφικών δεδομένων ή πληροφοριών για το χώρο. Σε αντίθεση με τα καύσιμα των αυτοκινήτων στα γεωγραφικά δεδομένα των Γ.Σ.Π. δεν παρουσιάζεται έλλειψη. Υπάρχουν διάφορες μορφές δεδομένων αρκετές από τις οποίες μπορούν να μελετηθούν ακόμη και από αρχάριους χρήστες.

Κατ' αρχήν υπάρχουν τα γεωγραφικά δεδομένα που αναφέρονται στο φυσικό χώρο, δηλαδή σε σημεία, γραμμές και επιφάνειες. Χρησιμοποιώντας αυτές τις ομάδες δεδομένων μπορούμε να σχεδιάσουμε τοπογραφικούς χάρτες – υπόβαθρα. Για παράδειγμα ένας παγκόσμιος χάρτης σεισμών θα μπορούσε να έχει ένα θεματικό επίπεδο γραμμών (line layer) με τα γεωγραφικά μήκη και πλάτη, ένα θεματικό επίπεδο επιφανειών (polygon layer) με τις

ηπείρους και ένα θεματικό επίπεδο σημείων (point layer) με τις θέσεις των επικέντρων των σεισμών. Επιπλέον ο χάρτης θα μπορούσε ακόμα να περιέχει και άλλα χρήσιμα θεματικά επίπεδα, όπως ένα γραμμικό με τα οδικό δίκτυο, ένα επιφανειακό με τις χώρες και ένα σημειακό που θα απεικόνιζε τις πόλεις με πληθυσμό περισσότερο από ένα εκατομμύριο κατοίκους. Όλα αυτά τα σετ γεωγραφικών δεδομένων του παραπάνω παραδείγματος, δείχνουν που είναι το κάθε τι στον πλανήτη και ορίζουν ένα μοναδικό όνομα ή αριθμό για το κάθε χαρακτηριστικό του χάρτη.

Ύστερα υπάρχουν οι γεωγραφικές πληροφορίες, που αποδίδουν τα χαρακτηριστικά των στοιχείων που συνθέτουν το χάρτη (attribute data). Αυτά τα σετ δεδομένων είναι πίνακες (data tables) δημιουργημένοι από γραμμές ή εγγραφές (records) και στήλες ή πεδία (fields) πληροφοριών. Για παράδειγμα, σε έναν πίνακα χωρών κάθε χώρα θα αντιστοιχούσε σε μια εγγραφή και κάθε πληροφορία για τη χώρα (π.χ. πληθυσμός) θα υπήρχε στο αντίστοιχο πεδίο. Έτσι γρήγορα αυτός ο πίνακας θα μπορούσε να έχει πάρα πολλές πληροφορίες για πολλές χώρες. Ακόμη θα μπορούσαν να υπάρχουν πολλοί πίνακες με πληροφορίες για αυτές τις χώρες. Η σύνδεση κάθε μοναδικής εγγραφής με την αντίστοιχη ονομασία στα γεωγραφικά δεδομένα έχει ως αποτέλεσμα την μετατροπή αδιάφορων πινάκων άμεσα σε δυναμικές οπτικές απεικονίσεις. Οι χώρες δεν θα απεικονίζονται πλέον με το ίδιο χρώμα, αλλά θα έχουν για παράδειγμα διαφορετικές αποχρώσεις ανάλογα με το δείκτη γήρανσης του πληθυσμού. Τα ποτάμια δεν θα απεικονίζονται πλέον ως μπλε γραμμές αλλά θα έχουν μια διαβάθμιση ανάλογα με την ποσότητα του νερού που ρέει σε αυτά. Τέλος, οι πόλεις δεν θα είναι απλές κουκίδες στο χάρτη αλλά θα σχηματίζονται , θα αποκτούν μέγεθος και χρώμα ανάλογα ώστε να περιγράφουν κάποια γεωγραφική πληροφορία.

Άλλες μορφές πληροφοριών που πρέπει να λάβουμε υπόψη είναι οι πληροφορίες σε μορφή κειμένου (text document). Αυτές υπάρχουν σε αφθονία και αξίζουν ιδιαίτερης προσοχής. Ο προγραμματισμός του λογισμικού που χρησιμοποιούμε είναι απλός και γίνεται με το πάτημα ενός κουμπιού που εμφανίζει αυτόματα τα αρχεία κειμένου, που περιέχουν πληροφορίες για συγκεκριμένα χαρακτηριστικά πάνω στο χάρτη μας. Αναλόγως, ψηφιακές εικόνες, φωτογραφίες και βίντεο μπορούν να προγραμματιστούν έτσι ώστε με την επιλογή μιας θέσης στο χάρτη να δίνεται εντολή στο πρόγραμμα να εμφανίσει το επιθυμητό αρχείο. Με λίγα λόγια κάθε πληροφορία που μπορεί να εμφανιστεί σε ψηφιακή μορφή, μπορεί να χρησιμοποιηθεί με τον παραπάνω τρόπο.

Τα γεωγραφικά δεδομένα σχηματίζουν τη βάση των δεδομένων που είναι το θεμέλιο των Γ.Σ.Π. στην Ανώτατη Εκπαίδευση. Στα πλαίσια των μαθημάτων της Γεωγραφίας τα Γ.Σ.Π. μπορούν να χρησιμοποιηθούν με απλό και αποτελεσματικό τρόπο.

Η σύνδεση πινάκων με γεωγραφικές πληροφορίες με τα αντίστοιχα θεματικά επίπεδα γεωγραφικών πληροφοριών, δίνει τη δυνατότητα στους σπουδαστές – χρήστες των Γ.Σ.Π. να χαρτογραφήσουν μια ατελείωτη σειρά πληροφοριών. Για την εξερεύνηση γεωγραφικών δεδομένων, οι σπουδαστές μπορούν να χρησιμοποιούν απλά εργαλεία για να εστιάζουν (zoom), να μετακινούνται μέσα στο χάρτη (pan), να μετρούν, να επιλέγουν και να εξαγάγουν πληροφορίες για συγκεκριμένα χαρακτηριστικά του χάρτη. Όλα αυτά γίνονται με μικρό βαθμό δυσκολίας. Εύκολη θεωρείται εξίσου η χωρική ανάλυση των δεδομένων και ο υπολογισμός βασικών στατιστικών μεγεθών.

Οι προσπάθειες που έγιναν σε ξένα πανεπιστήμια για διδασκαλία περαιτέρω ικανοτήτων και δεξιότητων των Γ.Σ.Π. στα πλαίσια του μαθήματος της Γεωγραφίας όπως η δημιουργία υπομνημάτων σε χάρτες, η επέμβαση σε πίνακες δεδομένων, η δημιουργία πινάκων

δεδομένων και η ψηφιοποίηση δεν είχαν τα αναμενόμενα αποτελέσματα. Οι σπουδαστές αισθάνθηκαν ότι αποσπάρθηκαν από την ικανότητα να μάθουν το περιεχόμενο της επιστήμης της Γεωγραφίας και απογοητεύτηκαν εύκολα. Επομένως συμπεραίνουμε στα πλαίσια του μαθήματος της Γεωγραφίας πρέπει:

- Να παραμένουμε στη διδασκαλία του περιεχομένου της επιστήμης της Γεωγραφίας.
- Να προετοιμάζουμε και να εμφανίζουμε τα γεωγραφικά δεδομένα χρησιμοποιώντας τα Γ.Σ.Π., με τέτοιο τρόπο που να μην απαιτείται λεπτομερής χρήση και επιδεξιότητα στο χειρισμό του εργαλείου.

Βασιζόμενοι στα παραπάνω συμπεράσματα και με σταδιακή αύξηση της χρήσης των Γ.Σ.Π. στα μαθήματα της Γεωγραφίας είναι πολύ πιθανό στο άμεσο μέλλον, τα Γ.Σ.Π. να αναβαθμίσουν ακόμη περισσότερο το ρόλο τους στη διδακτική της Γεωγραφίας.

Στο σημείο αυτό αξίζει να παραθέσουμε μια χαρακτηριστική φράση για τη σπουδαιότητα των Γ.Σ.Π. στα μαθήματα της Γεωγραφίας η οποία διατυπώθηκε από το Geography Education Standards Projects (1994): “ Η δύναμη ενός Γ.Σ.Π. έγκειται στο γεγονός ότι μας επιτρέπει να κάνουμε ερωτήσεις στα γεωγραφικά δεδομένα / πληροφορίες.” Οι σπουδαστές, χρησιμοποιώντας αυτές τις ερωτήσεις έρευνας, προσέγγισης και αναζήτησης πληροφοριών, αναπτύσσουν μια μεθοδολογία, συσσωρεύουν και αναλύουν γεωγραφικά δεδομένα και καταλήγουν σε συμπεράσματα.

1.4 Η Χρησιμότητα των Γ.Σ.Π. στην Εκπαίδευση

Στην ενότητα αυτή θα αναφερθούν όλα εκείνα τα στοιχεία που μπορούν να προσφέρουν τα Γ.Σ.Π. στον τομέα της εκπαίδευσης γενικότερα, αφού μπορούν να εφευρεθούν εφαρμογές και σε άλλα πεδία εκτός του μαθήματος της Γεωγραφίας. Αυτό προκύπτει από το γεγονός ότι σε πολλές χώρες της Ευρώπης και κυρίως στις Η.Π.Α. τα Γ.Σ.Π. χρησιμοποιούνται στη δευτεροβάθμια εκπαίδευση όχι μόνο σε μαθήματα Γεωγραφίας.

Η αποτελεσματικότητα των Γ.Σ.Π. δεν έγκειται μόνο σε ένα απλό πάτημα πλήκτρου για τη δημιουργία ενός θεματικού χάρτη. Από τη μεριά της εκπαίδευσης αυτό σημαίνει ανάπτυξη της ενεργητικής μάθησης που απαιτεί όμως κριτική σκέψη. Τα Γ.Σ.Π. μπορούν να επηρεάσουν και να δομήσουν την εμπειρία της εκπαίδευσης για τους φοιτητές, τους καθηγητές και την κοινωνία σε ένα διαφορετικό πλαίσιο από το συνηθισμένο. Στη συνέχεια παρουσιάζονται μερικές από τις βασικές απαιτήσεις, μελλοντικές αξίες και επιπτώσεις από τη χρήση των Γ.Σ.Π. στην Ανώτατη Εκπαίδευση.

- Τα Γ.Σ.Π. μπορούν να κατέχουν σημαντικό ρόλο στην αναβάθμιση των σπουδών.
 - Η προαγωγή της αλλαγής και της ανάπτυξης στην εκπαίδευση από τα Γ.Σ.Π. με οφέλη τόσο για τους σπουδαστές, όσο και για τους καθηγητές μπορεί να συγκροτήσει ένα ισχυρό εργαλείο. Συγχρόνως παρέχεται η δυνατότητα για αξιοποίηση ικανοτήτων, εκπαιδευτική μεθοδολογία, περιεχόμενο σπουδών και κοινωνική συμμετοχή.
 - Το λογισμικό και οι γεωγραφικές πληροφορίες δεν περιέχουν, ούτε δείχνουν την απάντηση. Ο ορισμός των απαντήσεων γίνεται βάσει των ερωτήσεων που έχουν τεθεί και των παραμέτρων που θέλουμε να επαληθεύσουμε από τους χρήστες των Γ.Σ.Π.. Τα Γ.Σ.Π. ορίζουν επίσης μεθόδους με τις οποίες έχουμε τη δυνατότητα να μελετήσουμε εναλλακτικές απαντήσεις και σενάρια για συγκεκριμένα προβλήματα και καταστάσεις. Οι χρήστες πρέπει ακόμη να προσδιορίσουν τι απαρτίζει μια ικανοποιητική απάντηση στο ερώτημα που έχει τεθεί από τους ίδιους. Ο θεμελιώδης ρόλος της κριτικής σκέψης στην αποτελεσματική χρήση των Γ.Σ.Π. έχει ως αποτέλεσμα η εξερεύνηση να αποτελεί μεγάλη πρόκληση για τους σπουδαστές.
 - Οι σπουδαστές και οι καθηγητές μπορεί να είναι ταυτόχρονα ενεργητικοί μαθητές. Η καλλιέργεια καινούργιων ικανοτήτων και η εξερεύνηση νέων ερμηνειών σε μια σειρά ποικίλων θεμάτων από τους καθηγητές μπορούν να εντρυφήσουν τους σπουδαστές στη διαδικασία και στην αξία της διαρκούς μάθησης πέρα από τα στενά όρια των εκπαιδευτικών ιδρυμάτων.
 - Αφού ο ηλεκτρονικός υπολογιστής αποτελεί ένα πολύ δυνατό εργαλείο για την εξερεύνηση του ίδιου περιεχομένου μέσω εναλλακτικών διαδρομών, οι σπουδαστές που χρησιμοποιούν τα Γ.Σ.Π., μπορούν να προοδεύσουν με διάφορους τρόπους, με ένα προσωπικό τρόπο και ρυθμό που αρμόζει για τις προσωπικές τους ανάγκες, δυνάμεις και ενδιαφέροντα τους. Η ενεργητική εξερεύνηση με τη χρήση Γ.Σ.Π. μπορεί εύκολα να συνδυάσει τις πολλαπλές μεθόδους πρόσβασης δεδομένων τις οποίες οι διάφοροι σπουδαστές χρειάζονται, ενώ επιπλέον είναι δυνατόν η συμμετοχή σε ομαδικές δραστηριότητες και η εξασφάλιση μιας δυναμικής ευκαιρίας για το σχεδιασμό ξεχωριστών οραμάτων για τον κόσμο.

- Η αποτίμηση της προόδου και των επιτευγμάτων των σπουδαστών μπορεί να πραγματοποιηθεί με πολλαπλούς τρόπους, κατάλληλους για τους ίδιους τους σπουδαστές, τα ενδιαφέροντά τους και το πρόγραμμά τους.
- Η χρήση Γ.Σ.Π. μπορεί να βοηθήσει τόσο τους σπουδαστές, όσο και τους καθηγητές στην ενασχόληση με τα τοπικά, κοινωνικά και παγκόσμια δρώμενα. Η συνεργασία με άλλους σπουδαστές – χρήστες Γ.Σ.Π. θεσπίζει την αρχή, ότι ο χώρος εκπαίδευσης είναι μια μικρή κοινωνία. Οι σπουδαστές, τα εκπαιδευτικά ιδρύματα και πάνω από όλα η κοινωνία ευεργετούνται από αυτή τη συνεργασία αφού ο καθένας ενδιαφέρεται για τις ανάγκες των άλλων.
- Τα Γ.Σ.Π. αποτελούν ένα επαγγελματικό εργαλείο για τους σπουδαστές.
 - Τα Γ.Σ.Π. βοηθούν στην ανάπτυξη βασικών υπολογιστικών γνώσεων.
 - Η παροχή ολοκληρωμένης εκπαίδευσης στη διαδικασία της έρευνας, αλλά και στη συλλογή, προετοιμασία, αποθήκευση, ανάλυση και παρουσίαση των δεδομένων επιτυγχάνεται μέσω της αποτελεσματικής χρήσης των Γ.Σ.Π..
 - Οι δραστηριότητες των Γ.Σ.Π. παρέχουν πραγματική εκπαίδευση για πολλά επαγγέλματα. Τα Γ.Σ.Π. είναι το εργαλείο του 21^{ου} αιώνα.
- Τα Γ.Σ.Π. συνδυάζουν και εξασκούν πολλαπλές ικανότητες και ευφυΐες.
 - Κριτική σκέψη (ικανότητα ανάλυσης, σύνθεσης και αποτίμησης).
 - Λογική - μαθηματική ευφυΐα.
 - α. Αριθμολογία (ικανότητα στην ερμηνεία και χρήση αριθμών και αριθμητικών δεξιοτήτων).
 - β. Τεχνολογική δυναμικότητα (ικανότητα στην κατανόηση και χρήση εργαλείων, τα οποία διευκολύνουν την απόκτηση, επεξεργασία και μεταβίβαση πληροφοριών).
 - Γλωσσική ευφυΐα.
 - α. Γνώση γραπτού λόγου (ικανότητα στην ερμηνεία και παρουσίαση πληροφοριών σε μορφή κειμένου).
 - β. Γνώση γραφικών (ικανότητα στην ανάγνωση και χρήση οπτικών συμβόλων).
 - Κατανόηση της έννοιας του χώρου – ανάγνωση χάρτη (ικανότητα στη μετατροπή της πραγματικής ζωής σε νοερή ή οπτική εικόνα ή εικονική πραγματικότητα σε πολλαπλές κλίμακες).
 - Διαπροσωπική ευφυΐα – επικοινωνία (ικανότητα στην αποτελεσματική μεταφορά πληροφοριών και γνώσεων που αποκτήθηκαν από ερευνητική διαδικασία).

- Τα Γ.Σ.Π. βοηθούν στην ανάπτυξη της εξερεύνησης στη μάθηση.
 - Αποκτούν διάθεση για ενεργητική μάθηση.
 - Ψάχνουν για νέες δυνατότητες.
 - Ερευνούν πολλαπλές όψεις ενός συγκεκριμένου θέματος ή μιας ομάδας πληροφοριών.
 - Αναγνωρίζουν ότι σπάνια υπάρχει μια “σωστή απάντηση” σε μια δοσμένη ερώτηση ή πρόβλημα.
 - Αντιλαμβάνονται ότι ο καθηγητής είναι αυτός που δείχνει το δρόμο προς τη διαρκή μάθηση, παρά αυτός που παρουσιάζει πληροφορίες.
- Τα Γ.Σ.Π. προωθούν την έρευνα για την απόκτηση πληροφοριών.
 - Οι αποτελεσματικοί χρήστες είναι ικανοί για την αναγνώριση των κατάλληλων τύπων και πηγών πληροφοριών που χρειάζονται, για να δοθεί λύση σε κάποιο πρόβλημα.
 - Οι αποτελεσματικοί χρήστες ενοποιούν πληροφορίες διαφορετικού τύπου και προέλευσης.
 - Οι αποτελεσματικοί χρήστες αναγνωρίζουν για να επιτευχθεί η σωστή χρήση μιας οποιασδήποτε πληροφορίας, προϋποτίθεται η κατανόηση της φύσης και της ποιότητας των δεδομένων. Επίσης, οι αποτελεσματικοί χρήστες μπορούν να αναγνωρίσουν παράγοντες οι οποίοι επιδρούν στην ποιότητα των δεδομένων και ξέρουν πώς να συσχετίσουν και να προσαρμόσουν κατάλληλα τα δεδομένα.
- Τα Γ.Σ.Π. προωθούν τη χωρική κατανόηση.
 - Οι αποτελεσματικοί χρήστες μπορούν να διακρίνουν χωρικά πρότυπα σε πραγματική τοπογραφία και σε συμβολικές αναπαραστάσεις.
 - Οι αποτελεσματικοί χρήστες κάνουν επαναληπτικές ερωτήσεις για να περιγράψουν και να εξηγήσουν χωρικά πρότυπα.
 - Οι αποτελεσματικοί χρήστες ερευνούν χωρικά πρότυπα, αναγνωρίσιμα μέσα από μια πληθώρα πηγών δεδομένων και διαφορετικού είδους αναπαραστάσεις.
 - Οι αποτελεσματικοί χρήστες προχωρούν στην ενοποίηση δεδομένων διαφορετικής κλίμακας και αναγνωρίζουν πρότυπα και διαδικασίες μέσα σε μια μεγάλη ακτίνα δράσης, από πολύ μικρό ως πολύ μεγάλο.

- Τα Γ.Σ.Π. προωθούν τη γνώση ηλεκτρονικών υπολογιστών.
 - Επιδέξιος χειρισμός βάσεων δεδομένων (κριτική ικανότητα).
 - Χειρισμός προγραμμάτων λογιστικού φύλλου.
 - Χρήση γραφικών εργαλείων.
 - Χρήση δεδομένων τηλεπισκόπησης (αεροφωτογραφίες και δορυφορικές εικόνες).
 - Χρήση του διαδικτύου για εύρεση δεδομένων.
 - Χρήση λογισμικού παρουσιάσεων.
 - Παραγωγή εργασιών με χρήση πολυμέσων.

- Η αποτελεσματική χρήση των Γ.Σ.Π. απαιτεί γνώση προχωρημένων διαδικασιών του λογισμικού.
 - Με την πάροδο του χρόνου οι αποτελεσματικοί χρήστες μαθαίνουν τον τρόπο με τον οποίο το λογισμικό εξυπηρετεί ακριβώς το σκοπό τους.
 - Οι αποτελεσματικοί χρήστες εκμεταλλεύονται τα διάφορα επίπεδα του λογισμικού. Μερικές διαδικασίες είναι πιο περίπλοκες από άλλες, ενώ μερικά εργαλεία είναι ευρύτερα εφαρμόσιμα. Οι αποτελεσματικοί χρήστες αρχικά μαθαίνουν να χρησιμοποιούν βασικά εργαλεία και με τον καιρό στην εργαλειοθήκη τους προσθέτουν πιο προχωρημένα.

ΚΕΦΑΛΑΙΟ 2: Βασικές Έννοιες της Δειγματοληψίας και της Στατιστικής

2.1 Γενικά

Στατιστική είναι η επιστήμη η οποία ασχολείται με τη συλλογή, τη συνοπτική παρουσίαση και ανάλυση – ερμηνεία παρατηρήσεων που υπόκεινται σε τυχαίες μεταβολές με τελικό στόχο την εξαγωγή βασικών συμπερασμάτων και τη λήψη βέλτιστων αποφάσεων σε συνθήκες αβεβαιότητας. (Πηγή: Κοκολάκης Γ. , Σπηλιώτης Ι., 1999)

Η επιστήμη της στατιστικής είναι ένα σύνολο μεθόδων και κανόνων για:

- Την οργάνωση μιας συλλογής δεδομένων.
- Τη συνοπτική και αποτελεσματική παρουσίαση και ανάλυση τους.
- Την εξαγωγή χρησιμων και γενικευμένων συμπερασμάτων για τους πληθυσμούς με βάση τις πληροφορίες που συλλέγονται από τα δείγματα.

Ο κλάδος της στατιστικής που ασχολείται με τον πρώτο στόχο ονομάζεται σχεδιασμός πειραμάτων (experimental design).

Η δεύτερη περίπτωση αναφέρεται στον κλάδο της στατιστικής που ονομάζεται περιγραφική στατιστική (descriptive statistics).

Ο τρίτος τύπος προβλημάτων ονομάζεται επαγωγική στατιστική ή στατιστική συμπερασματολογία (inferential statistics).

Στη στατιστική ορολογία το σύνολο των στοιχείων που θέλουμε να περιγράψουμε και να μελετήσουμε ως προς μια συγκεκριμένη ιδιότητα ή χαρακτηριστικό ονομάζεται **πληθυσμός**. Κάθε πεπερασμένο υποσύνολο του πληθυσμού ονομάζεται **δείγμα**. Όπως εξυπακούεται από τη λέξη υποσύνολο τα μέλη του δείγματος αποτελούν και μέλη του πληθυσμού. Οποιοδήποτε χαρακτηριστικό ή κατάσταση παρουσιάζει αλλαγή, ή έχει διαφορετική τιμή για διαφορετικά υποκείμενα τα οποία ανήκουν σε ένα πληθυσμό ή ένα δείγμα ονομάζεται **μεταβλητή**.

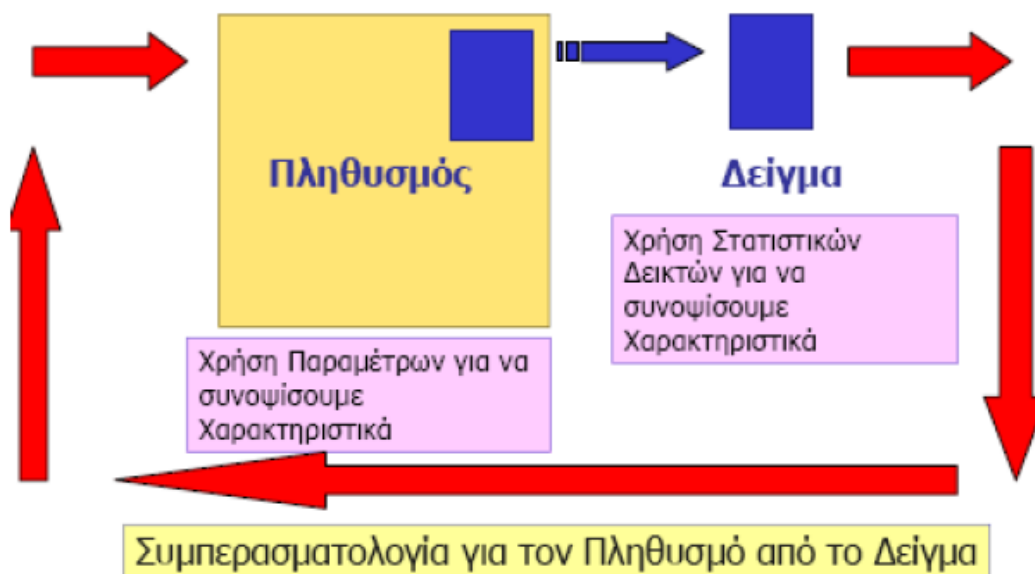


Η στατιστική σήμερα χρησιμοποιείται ευρύτατα σε όλους σχεδόν τους τομείς της ανθρώπινης δραστηριότητας. Βασικές έννοιες και παράμετροι έχουν εισχωρήσει και στη συνέχεια έχουν ενσωματωθεί σε όλες τις επιστήμες. Από τις Ανθρωπιστικές, Νομικές και Κοινωνικές επιστήμες (Αρχαιολογία, Κοινωνιολογία κ.α.), τις Φυσικές Επιστήμες, τις επιστήμες Υγείας, τις Τεχνολογικές και φυσικά τις επιστήμες Οικονομίας και Διοίκησης. Η ανάλυση στατιστικών ερευνών είναι το κυριότερο εργαλείο έρευνας σε ένα μεγάλο φάσμα εφαρμογών των παραπάνω επιστημών. Σπουδαίες πηγές βασικής γνώσης των κοινωνικών επιστημών αποτελούν οι έρευνες των ανθρώπινων πληθυσμών που είναι γνωστές με τον όρο δημοσκοπήσεις.

Ένα παράδειγμα επαγωγικής στατιστικής σε δείγμα 100 ανθρώπων μιας πόλης 20.000 κατοίκων που ερωτάται τι πρόκειται να ψηφίσουν στις επόμενες εκλογές.

Αρχικά πρέπει να αποσαφηνιστούν κάποιες βασικές έννοιες:

- Με ποιον τρόπο επιλέχτηκε το δείγμα (π.χ. απλή τυχαία δειγματοληψία).
- Γιατί επιλέχτηκε το συγκεκριμένο μέγεθος δείγματος (100 ατόμων).
- Ποιος είναι ο πληθυσμός από τον οποίο επιλέχτηκε το δείγμα και στον οποίο θα αναφέρονται τα αποτελέσματα του δείγματος (οι μόνιμοι κάτοικοι της πόλης που έχουν δικαίωμα ψήφου;).
- Η μεταβλητή που προφανώς μας ενδιαφέρει (προφανώς η πρόθεση ψήφου).
- Επιμέρους μεταβλητές που θα μπορούσαν να μας ενδιαφέρουν (πρόθεση ψήφου ανά ηλικιακή ομάδα)κ.τ.λ.
- Ακρίβεια των αποτελεσμάτων.



Ένας τρόπος για να λάβουμε τις απαραίτητες πληροφορίες που χρειαζόμαστε για κάποιο πληθυσμό είναι να εξετάσουμε όλα τα άτομα (στοιχεία) του πληθυσμού ως προς το χαρακτηριστικό που μας ενδιαφέρει. Η μέθοδος συλλογής των δεδομένων με αυτόν τον τρόπο ονομάζεται απογραφή (census). Για παράδειγμα η Εθνική Στατιστική Υπηρεσία (ΕΣΥΕ) της χώρας μας διενεργεί απογραφή πληθυσμού κάθε 10 χρόνια η οποία αποτελεί κύρια πηγή δεδομένων δημογραφικού, οικονομικού, εμπορικού και βιομηχανικού χαρακτήρα. Όμως τις περισσότερες φορές η μελέτη ολόκληρου του πληθυσμού δεν είναι πρακτικά δυνατή, γιατί το μέγεθος του είναι πολύ μεγάλο. Για αυτό το λόγο αντικείμενο της μελέτης αποτελεί ένα υποσύνολο του πληθυσμού που ονομάζεται **δείγμα** και από το οποίο ο ερευνητής συλλέγει πληροφορίες. Κάνει τις παρατηρήσεις του στο δείγμα και ύστερα τις γενικεύει για ολόκληρο τον πληθυσμό. Οι αρχές και οι μέθοδοι για τη συλλογή και ανάλυση δεδομένων από πεπερασμένους πληθυσμούς είναι αντικείμενο της **Δειγματοληψίας** (sampling), που αποτελεί τη βάση της Στατιστικής. Η οργάνωση της συλλογής και επεξεργασίας των σχετικών δεδομένων και πληροφοριών πραγματοποιείται με τέτοιο τρόπο που για δεδομένη ακρίβεια να επιτυγχάνεται το χαμηλότερο δυνατό κόστος ή αντιστρόφως να εξασφαλίζεται η μέγιστη δυνατή ακρίβεια σε σχέση με τα μέσα που έχουμε στη διάθεση μας.

2.2 Πλεονεκτήματα της Δειγματοληψίας

Ο βασικός σκοπός της επαγωγικής στατιστικής είναι η εξαγωγή χρήσιμων και γενικευμένων συμπερασμάτων για τους πληθυσμούς, με βάση τις πληροφορίες που συλλέγονται από τα δείγματα. Τόσο πρακτικοί όσο και θεωρητικοί λόγοι δίνουν προβάδισμα στη δειγματοληψία έναντι της απαρίθμησης και της απογραφής ολόκληρου του πληθυσμού. Τα πλεονεκτήματα της δειγματοληψίας παρουσιάζονται συνοπτικά παρακάτω.

Η δειγματοληψία είναι απαραίτητη σε πολλά προβλήματα της γεωγραφικής έρευνας. Εάν ο πληθυσμός που μελετάται είναι υπερβολικά μεγάλος το να επιτευχθεί μια πλήρης απαρίθμηση - απογραφή δεν είναι δυνατό. Για παράδειγμα η δειγματοληψία είναι η μόνη εναλλακτική λύση όταν μελετώνται οι λόγοι για τους οποίους τμήματα του πληθυσμού μετακινούνται από μια περιοχή σε μια άλλη περιοχή αφού δεν είναι δυνατό ο μελετητής να έρθει σε επαφή με τα τμήματα του πληθυσμού που άλλαξαν τόπο κατοικίας σε ένα ορισμένο χρονικό διάστημα. Επίσης, ένας γεωγράφος που αναλύει τις παγκόσμιες αλλαγές στη φύση και στην έκταση των περιοχών που συμβαίνουν τροπικές καταιγίδες, θα λάβει δείγματα από αυτές τις περιοχές μιας και είναι αδύνατο να έχει στοιχεία για το σύνολο αυτών των περιοχών που μπορεί να είναι μεγάλες και πρακτικά άπειρες.

Η δειγματοληψία είναι αποτελεσματική και οικονομική μέθοδος συλλογής δεδομένων. Με τη βοήθεια ενός δείγματος μια κατάλληλη ποσότητα δεδομένων που αφορά τον πληθυσμό μπορεί να αποκτηθεί και να αναλυθεί σε ελάχιστο χρονικό διάστημα. Συνεπώς, η δειγματοληψία μειώνει τα έξοδα, τον χρόνο και ελαχιστοποιεί μια σειρά από λογιστικά προβλήματα σε αντίθεση με την εξέταση του συνόλου του πληθυσμού.

Η παροχή υψηλής ακρίβειας στην πληροφορία χαρακτηρίζει τη δειγματοληψία. Σε γεωγραφικά προβλήματα στα οποία η σε βάθος ανάλυση είναι απαραίτητη μόνο ένας μικρός αριθμός ατόμων ή τοποθεσιών μπορεί να συμπεριληφθεί στη μελέτη. Αυτά τα ελάχιστα στοιχεία του δείγματος μπορεί να περάσουν από εκτεταμένο έλεγχο σε συνδυασμό με μια σειρά από πρόσθετες πληροφορίες για κάθε μέλος του δείγματος. Για παράδειγμα μια μελέτη για την συμπεριφορά του καταναλωτικού κοινού θα μπορούσε να περιέχει ερωτήσεις όπως πόσες φορές μέσα στην εβδομάδα πηγαίνετε για ψώνια στο σουπερμάρκετ κ.τ.λ..

Η επανάληψη της συλλογής της πληροφορίας γρήγορα και με μικρό κόστος είναι ένα άλλο δικαίωμα που παρέχει η δειγματοληψία σε αντίθεση με πολλές άλλες γεωγραφικές έρευνες που απαιτούν την πρόσθετη συλλογή πληροφορίας μετά από κάποιο χρονικό διάστημα ώστε να εξεταστούν οι όποιες διαφορές προέκυψαν. Για παράδειγμα ένας αστικός γεωγράφος που μελετά την ανάπτυξη μιας περιοχής θα ήθελε να επικεντρωθεί το ενδιαφέρον του στις απόψεις των κατοίκων πριν και μετά την κατασκευή ενός μεγάλου εμπορικού κέντρου στην περιοχή τους. Η μελέτη τέτοιου είδους προβλημάτων δυναμικά και όχι στατικά είναι η μόνη λύση ώστε τα έξοδα σε χρόνο και χρήματα της δειγματοληψίας να μην είναι μεγάλα. Εν κατακλείδι η δειγματοληψία συχνά παρέχει υψηλής ακρίβειας αποτελέσματα που είναι πιο ακριβή από εκείνα της απογραφής η οποία αναλώνει πολυάριθμο ανθρώπινο προσωπικό για την διεξαγωγή που δεν είναι άρτια εκπαιδευμένο και καταρτισμένο και χωρίς να πετυχαίνει τους στόχους της.

2.3 Πηγές Σφαλμάτων στη Δειγματοληψία

Όπως αναφέρθηκε, κύριος σκοπός της δειγματοληψίας είναι η εξαγωγή χρήσιμων συμπερασμάτων για τα χαρακτηριστικά του συνόλου του πληθυσμού μελετώντας ένα πεπερασμένο υποσύνολο των μελών του πληθυσμού.

Για να διασφαλίσουμε ότι το δείγμα μας θα είναι αμερόληπτο και αντιπροσωπευτικό έννοιες που θα ερμηνευθούν περαιτέρω σε επόμενες παραγράφους πρέπει στη διαδικασία να εμπεριέχεται το στοιχείο της τύχης.

Πάντως, η τύχη από μόνη της δεν εξασφαλίζει την αμεροληψία του δείγματος, υπάρχουν και άλλες πιθανές πηγές σφαλμάτων.

Δυστυχώς, σπάνια γνωρίζουμε εάν το δείγμα μας είναι απολύτως αντιπροσωπευτικό. Αυτό γίνεται επειδή δεν είναι δυνατό να είναι γνωστές όλες οι παράμετροι από όπου προέρχεται το δείγμα. Οι όποιες αβεβαιότητες υπάρχουν έχουν ως αποτέλεσμα την άμεση συσχέτιση με το δείγμα και τη δημιουργία σφαλμάτων σε αυτό.

Κάνοντας μετρήσεις ακρίβειας και ορθότητας ωφελούν στην κατηγοριοποίηση των πηγών των σφαλμάτων. Τα αποτελέσματα από ένα μικρό δείγμα εμπεριέχουν το ενδεχόμενο να μην είναι ακριβή και ορθά. Σε αυτή την περίπτωση αυξάνοντας το μέγεθος του δείγματος μπορεί να αυξηθεί και η ακρίβεια. Σε αντίθετη περίπτωση μεγαλύτερο δείγμα δεν παρουσιάζει μεγάλες διαφορές μεταξύ των στοιχείων του αλλά έχει τα μειονεκτήματα της δαπάνης και του μεγαλύτερου χρόνου για τη συλλογή του. Επομένως, θα πρέπει να επιλέγεται μια ισορροπία ανάμεσα στην ακρίβεια του δείγματος και στη προσπάθεια που απαιτείται για την απόκτησή του.

Οι ανακρίβειες που παρουσιάζει ένα δείγμα είναι ένα εξαιρετικά περίπλοκο θέμα. Συστηματική μεροληπτικότητα (Systematic bias) μπορεί να εισαχθεί στη δειγματοληπτική διαδικασία με πολλούς και διαφορετικούς τρόπους. Πολυάριθμες ανακρίβειες προέρχονται από την διαδικασία που ακολουθείται. Για παράδειγμα τα στοιχεία ή τα μέλη του πληθυσμού που θέλουμε να εξετάσουμε μπορεί να μην ταιριάζουν με τα στοιχεία ή μέλη του δείγματος που παίρνουμε. Ένας αστικός γεωγράφος που μελετά τη χωροθέτηση των νέων οικοδομών σε μια περιοχή θα μπορούσε να χρησιμοποιήσει ως πηγή για τα δεδομένα του τη λίστα οικοδομικών αδειών. Αν όμως η λίστα αυτή εμπεριέχει τις ανακαινίσεις οικοδομών ως νέες οικοδομές, κάποιες από αυτές θα χρησιμοποιηθούν στο δείγμα ως νέες οικοδομές.

Η επιλογή λανθασμένης μεθόδου αποτελεί επίσης μια σημαντική πηγή για ανακριβή αποτελέσματα. Η δυσκολία ενός γεωγραφικού προβλήματος απαιτεί συνήθως πιο εξειδικευμένο σχέδιο από την απλή τυχαία δειγματοληψία.

Σε άλλα γεωγραφικά προβλήματα μια ακατάλληλη για την περίπτωση μέθοδος συλλογής δεδομένων μπορεί να έχει ως αποτέλεσμα την δημιουργία σφαλμάτων και τη λανθασμένη απεικόνιση της πραγματικότητας (π.χ. ερωτηματολόγιο δια αλληλογραφίας αντί τηλεφωνικής συνέντευξης).

Άλλες ανακρίβειες μπορεί να έχουν προέλευση από λογιστικά ή σφάλματα προσωπικού που δε σχετίζονται καθόλου με τη διαδικασία της δειγματοληψίας. Στη συλλογή των δεδομένων ή στις συνεντεύξεις μπορεί να προκύψουν ανακολουθίες που θα επηρεάσουν σε μεγάλο βαθμό τα αποτελέσματα. Επίσης, λάθη μπορεί να προκύψουν κατά την εγγραφή και την πινακοποίηση της πληροφορίας.

Ο μελετητής λοιπόν θα πρέπει να αναπτύξει και να σχεδιάσει τη διαδικασία της δειγματοληψίας προσεκτικά ώστε να περιοριστούν τα σφάλματα στο ελάχιστο δυνατό επίπεδο ή και καθόλου.

2.4 Σχέδια Δειγματοληψίας

Τα σχέδια δειγματοληψίας είναι:

- **Τυχαία Δειγματοληψία**
- **Μη Τυχαία Δειγματοληψία**

Η τυχαία δειγματοληψία (Random Sampling) αναφέρεται και σαν δειγματοληψία με γνωστή πιθανότητα επιλογής (Probability Sampling), διότι εάν η δειγματοληπτική διαδικασία είναι τυχαία ή τυχαιοποιημένη μπορεί να εφαρμοστεί η θεωρία των πιθανοτήτων καθώς και το ότι υπάρχει δυνατότητα υπολογισμού του τυχαίου σφάλματος. Έτσι το υπόδειγμα της κατανομής δειγματοληψίας είναι διαθέσιμο.

Ο όρος τυχαίο δείγμα (Random Sample) δεν χρησιμοποιείται για την περιγραφή των δεδομένων του δείγματος αλλά για τη διαδικασία που εφαρμόστηκε για την επιλογή του δείγματος. Επομένως η τυχαιότητα είναι μια ιδιότητα που χαρακτηρίζει την δειγματοληπτική διαδικασία και όχι ένα συγκεκριμένο δείγμα.

Μη τυχαία δειγματοληψία (Non Random Sampling) αναφέρεται και σαν δειγματοληψία με άγνωστη πιθανότητα επιλογής (Non Probability Sampling) δηλαδή το δείγμα επιλέγεται χωρίς τη χρήση της τυχαιότητας. Ένα δείγμα που επιλέγεται με αυτόν τον τρόπο στηρίζεται σε άλλες θεωρήσεις όπως για παράδειγμα την εμπειρική κρίση του ερευνητή. Τα σπουδαιότερο χαρακτηριστικό της μη τυχαίας δειγματοληψίας είναι ότι υφίσταται δειγματοληπτική μεταβλητότητα (διακυμάνσεις της δειγματοληψίας) και δεν υπάρχει τρόπος να γνωρίζουμε το υπόδειγμα της μεταβλητότητας κατά την διαδικασία.

Στη συγκεκριμένη εργασία γίνεται αναφορά και μελέτη της τυχαίας δειγματοληψίας και των μεθόδων της απλής και στρωματοποιημένης σημειακής δειγματοληψίας.

2.5 Η Αντιπροσωπευτικότητα του Δείγματος

Η παράθεση της δειγματοληπτικής μεθοδολογίας στοχεύει στον προσδιορισμό ενός δείγματος με το μικρότερο δυνατό μέγεθος έτσι ώστε να μειωθεί ανάλογα και το κόστος της δειγματοληψίας. Βασική προϋπόθεση είναι το δείγμα να αναπαράγει με τον πιο ικανοποιητικό τρόπο τα χαρακτηριστικά του αρχικού πληθυσμού από τον οποίο επιλέχτηκε και παράλληλα οι όποιες διαφοροποιήσεις του πληθυσμού να αντιπροσωπεύονται όσο το δυνατόν πιο πιστά.

Δυστυχώς η επιλογή αντιπροσωπευτικού δείγματος παρ' ότι είναι επιθυμητή τις περισσότερες φορές είναι σχεδόν αδύνατη. Το γεγονός αυτό πιστοποιείται με τα ακόλουθα δεδομένα. Εάν η χαρακτηριστική ιδιότητα ενός πληθυσμού έχει τις τιμές x_1, x_2, \dots, x_r με αντίστοιχες απόλυτες συχνότητες:

$$N_1, N_2, \dots, N_r \quad \text{όπου} \quad \sum_{i=1}^r N_i = N$$

τότε για ένα αντιπροσωπευτικό δείγμα οι αντίστοιχες απόλυτες συχνότητες θα πρέπει να είναι:

$$v_1, v_2, \dots, v_r \quad \text{όπου} \quad v_i = v \frac{N_i}{N}, \quad v = \sum_{i=1}^r v_i$$

Αυτό σημαίνει ότι για τον προσδιορισμό του δείγματος θα πρέπει να είναι γνωστές οι απόλυτες συχνότητες N_i ($i = 1, 2, \dots, r$) του αρχικού πληθυσμού.

Όταν όμως οι πληροφορίες για την κατανομή του αρχικού πληθυσμού δεν είναι τόσο ακριβείς σταματάει να έχει νόημα η εφαρμογή δειγματοληπτικών μεθόδων. Μπορεί να αναφερθεί ότι μια λύση στο πρόβλημα αυτό, της λήψης ενός καλού δείγματος, προσφέρει η εκλογή που στηρίζεται στα αποτελέσματα πειραμάτων τύχης, των οποίων η μεθοδολογία καθορίζεται αυστηρά με την επαναληπτική εκτέλεση των ίδιων των πειραμάτων. Συνεπώς η αξιολόγηση των αποτελεσμάτων συνιστά αντικείμενο της επαγωγικής στατιστικής, όπου η εξαγωγή συμπερασμάτων για τις ιδιότητες του αρχικού πληθυσμού και κυρίως για τον τρόπο κατανομής του διερευνώμενου χαρακτηριστικού βασίζεται στα αποτελέσματα της τυχαίας δειγματοληψίας.

2.6 Η Δειγματοληψία ως Μέρος της Γεωγραφικής Μελέτης

Η γεωγραφική μελέτη είναι μια μελέτη που αναλύει τη δομή του χώρου, τις αλληλεξαρτήσεις των στοιχείων και τις διαδικασίες αλλαγής του. Αναπόσπαστο κομμάτι μιας γεωγραφικής μελέτης είναι η συλλογή στοιχείων όπου από την επεξεργασία τους θα προκύψει η πληροφορία. Η παραπάνω σειρά ενεργειών αναφέρεται και ως διαδικασία από στοιχεία σε πληροφορία.

Τα στοιχεία μπορούν να διαιρεθούν σε δύο βασικές κατηγορίες: (Πηγή: Κουτσόπουλος Κ., 2006)

- **Ποσοτικά στοιχεία (quantitative)**
- **Ποιοτικά ή κατηγορικά (qualitative – categorical)**

Η διάκρισή τους σε θεωρητικό επίπεδο είναι προφανής. Τα μεν ποσοτικά είναι εκείνα που μπορούν να μετρηθούν, να πάρουν μια τιμή, να αναλυθούν κατά τάξη μεγέθους (π.χ. εισόδημα, πληθυσμός κ.τ.λ.) τα δε ποιοτικά ή κατηγορικά είναι εκείνα που δεν έχουν μέγεθος αλλά εκφράζουν κάποια ιδιότητα (π.χ. η κατάσταση των φυσικών πόρων, η ποιότητά τους κ.τ.λ.) και συνήθως ταξινομούνται.

Στη πραγματικότητα η συγκεκριμένη διάκριση δημιουργεί αρκετές φορές προβλήματα. Άλλωστε αυτό που ενδιαφέρει περισσότερο κατά τη συλλογή των στοιχείων για κάποια γεωγραφική μελέτη είναι το επίπεδο μέτρησης των στοιχείων. Επιπλέον τα στοιχεία υπάρχουν σε διαφορετικά επίπεδα μέτρησης. Τα επίπεδα μέτρησης των στοιχείων είναι: (Πηγή: Κουτσόπουλος Κ., 2006)

- **Ποιοτικό (nominal)**: κατηγοριοποιεί τα αντικείμενα παρά μετρά τις ιδιότητες τους.
- **Ιεραρχικό (ordinal)**: η κλίμακα μέτρησης εκφράζει μια σύγκριση ανάμεσα σε κατηγορίες (π.χ. μικρότερου ή μεγαλύτερου, κοντότερου ή ψηλότερου).
- **Ποσοτικό (interval)**: εκφράζει ποσοτικά μια ιεραρχική σχέση.
- **Αναλογικό (ratio)**: έχει το πλεονέκτημα μιας φυσικής αφετηρίας δηλαδή ξεκινά από το σταθερό μηδέν (π.χ. το βάρος, το ύψος, η μάζα, το μήκος κ.τ.λ.)

Από τη στιγμή που έχουν προσδιοριστεί το περιεχόμενο, οι πηγές άντλησης και το επίπεδο μέτρησης των στοιχείων θα πρέπει να αποφασιστεί η μέθοδος συλλογής τους. Αυτή μπορεί να είναι η απογραφική (καθολική – πλήρης) ή δειγματοληπτική (δείγμα – κομμάτι του πληθυσμού).

2.7 Χωρική Δειγματοληψία

Στη χωρική δειγματοληψία, η χωρική αναφορά κάθε στοιχείου, αντικειμένου είναι μεγάλης σημασίας. Η αναφορά αυτή μπορεί να εκφράζεται με τη βοήθεια ενός συστήματος αναφοράς το οποίο αποδίδει μοναδικές συντεταγμένες σε κάθε σημείο στο χώρο. Για παράδειγμα χωρική δειγματοληψία χρησιμοποιείται όταν έχουμε έναν χάρτη από συνεχώς κατανεμημένες μεταβλητές (όπως το είδος της βλάστησης, το είδος της πανίδας, ο τύπος του εδάφους κ.τ.λ.) και γίνεται επιλογή ενός δείγματος από τον χάρτη.

Ο σκοπός χρησιμοποίησης δειγμάτων πληθυσμού γίνεται για την αποκόμιση μιας γενικής εντύπωσης της περιοχής που μας ενδιαφέρει και για την εκτίμηση ορισμένων χαρακτηριστικών της. Η μοντέρνα γεωγραφία ασχολείται κυρίως με τη στατιστική χρήση των δειγμάτων για την εκτίμηση των χαρακτηριστικών της περιοχής που μελετάται.

Δύο βασικές παράμετροι κάθε χωρικής δειγματοληπτικής διαδικασίας είναι:

- **Ο Πληθυσμός**
- **Το Δείγμα**

Με την έννοια πληθυσμός ορίζεται ένα σύνολο στοιχείων που πρόκειται να μελετηθεί ως προς μία ή περισσότερες ιδιότητες - χαρακτηριστικά (μεταβλητές). Αξίζει να σημειωθεί ότι τα στοιχεία του πληθυσμού μπορεί να είναι οποιοδήποτε σύνολο αντικειμένων ή τιμών που προέκυψαν από πειράματα ή παρατηρήσεις

Για να γίνει πιο κατανοητή η έννοια του πληθυσμού ας υποθεθεί ότι υπάρχει ανάγκη να εξακριβωθούν διάφορα χαρακτηριστικά μιας περιοχής (π.χ. το μέγεθος των αγροτεμαχίων και το είδος της καλλιέργειάς τους). Η καταγραφή καθενός από αυτά τα χαρακτηριστικά δημιουργεί δύο πληθυσμούς. Ο ένας αποτελείται από το σύνολο των **μετρήσεων** (measurements) για το μέγεθος όλων των αγροτεμαχίων της περιοχής και ο δεύτερος από το σύνολο των **απαριθμήσεων** (counts) για τα είδη των καλλιεργειών στην περιοχή.

Συνεπώς στη στατιστική πληθυσμός νοείται ένα σύνολο από μετρήσεις ή απαριθμήσεις που προέρχονται από όλα τα αντικείμενα που έχουν ένα ή περισσότερα κοινά χαρακτηριστικά. Στο παραπάνω παράδειγμα το κοινό χαρακτηριστικό αναφερόταν στο αντικείμενο << **αγροτεμάχια μιας ορισμένης περιοχής** >>. (Πηγή: Κουτσόπουλος Κ., 2006)

Στη στατιστική **δείγμα** λέγεται κάθε πεπερασμένο υποσύνολο του πληθυσμού. Αναμφισβήτητα ο καλύτερος τρόπος μελέτης των ιδιοτήτων ενός γεωγραφικού πληθυσμού είναι η μελέτη κάθε αντικειμένου ξεχωριστά. Τις περισσότερες φορές δεν είναι αναγκαίο αλλά ούτε και πρακτικό να αναζητούνται μετρήσεις και απαριθμήσεις για ολόκληρο τον πληθυσμό. Έτσι χρησιμοποιείται μόνο ένα τμήμα - δείγμα του πληθυσμού. Το τμήμα – δείγμα αντικαθιστά τον πληθυσμό για να εξοικονομηθεί χρόνος, χρήματα και προσπάθεια ή όταν οι μετρήσεις σε ολόκληρο τον πληθυσμό δεν είναι δυνατόν ή πρακτικό να γίνουν.

Για να αντικαθιστά αξιόπιστα ένα δείγμα τον πληθυσμό θα πρέπει να είναι **αντιπροσωπευτικό**, δηλαδή να θα πρέπει να εκφράζει τις διαφοροποιήσεις του πληθυσμού όσο γίνεται πιο πιστά. Επίσης, θα πρέπει να μην εμπεριέχει **μεροληπτικότητα** (bias). Η μεροληπτικότητα είναι το αποτέλεσμα της επιλογής μη αντιπροσωπευτικών τμημάτων του πληθυσμού για δειγματοληψία. Τα δείγματα θα πρέπει να αποφεύγονται να επιλεγούν με βάση την προσωπική κρίση του ερευνητή γιατί είναι σίγουρο ότι θα εμπεριέχουν μεροληπτικότητα. Ο πιο εύκολος τρόπος λήψης ενός καλού δείγματος που αναπαράγει τον αρχικό πληθυσμό κατά τον καλύτερο τρόπο είναι εκείνος που βασίζεται στα αποτελέσματα πειραμάτων τύχης. Άλλωστε, σκοπός της δειγματοληψίας δεν είναι η μέτρηση των χαρακτηριστικών ιδιοτήτων του δείγματος αλλά η εκτίμηση των πραγματικών μεγεθών του πληθυσμού.

2.7.1 Βήματα Δειγματοληπτικής Διαδικασίας

Στις γεωγραφικές μελέτες στις οποίες συλλέγεται πληροφορία από δείγματα, η διαδικασία περιλαμβάνει μια σειρά από βήματα που αναφέρονται παρακάτω:

Βήμα 1^ο: Καθορισμός πληθυσμού και περιοχής αναφοράς.

- Πληθυσμός αναφοράς ορίζεται το σύνολο των μελών από όπου συλλέγεται η πληροφορία.
- Περιοχή αναφοράς ορίζεται ολόκληρη η περιοχή ή το σύνολο των θέσεων από όπου συλλέγεται η πληροφορία.

Βήμα 2^ο: Προσδιορισμός του δειγματοληπτικού πληθυσμού και της δειγματοληπτικής περιοχής από το δειγματοληπτικό υπόβαθρο.

- Δειγματοληπτικός πληθυσμός ορίζεται το σύνολο των μελών από τα οποία στη πραγματικότητα επιλέγεται το δείγμα.
- Δειγματοληπτική περιοχή ορίζεται η περιοχή ή το σύνολο των θέσεων που βρίσκονται τα στοιχεία τα οποία λαμβάνουν μέρος στη διαδικασία.

Βήμα 3^ο: Επιλογή δειγματοληπτικής μεθόδου:

- Η τυχαία δειγματοληψία (probability sampling) προτιμάται έναντι της μη τυχαίας (non probability sampling).
- Τα κυριότερα είδη δειγματοληψίας είναι: η απλή τυχαία, η συστηματική και η στρωματοποιημένη.

Βήμα 4^ο: Οι μέθοδοι συλλογής των δεδομένων περιλαμβάνουν: άμεσες παρατηρήσεις, μετρήσεις πεδίου, τηλεφωνικές συνεντεύξεις, προσωπικές συνεντεύξεις, αποστολή ερωτηματολογίων.

Βήμα 5^ο: Διεξαγωγή προκαταρκτικού τεστ και έλεγχος της λειτουργίας της μεθόδου που επιλέχτηκε, διόρθωση των προβλημάτων που μπορούν να οδηγήσουν σε σφάλματα τη δειγματοληψία και τελική απόφαση για το μέγεθος του δείγματος.

Βήμα 6^ο: Συλλογή δείγματος:

- Η συνοχή στη μέθοδο συλλογής του δείγματος και στη διαδικασία (procedure) είναι μεγάλης σημασίας.
- Εξασφάλιση υψηλού επιπέδου αποτελεσμάτων.

2.7.2 Δειγματοληπτικό Υπόβαθρο

Το δειγματοληπτικό υπόβαθρο απαρτίζει τον στατιστικό πληθυσμό από όπου λαμβάνεται το δείγμα και μπορεί να είναι χωρικό ή μη χωρικό.

Όταν η χωρική διάσταση των στοιχείων, αντικειμένων, φαινομένων που μελετώνται και κατ' επέκταση η θέση τους είναι καθοριστικής σημασίας, είναι απαραίτητο το δείγμα να ληφθεί από ένα χωρικό δειγματοληπτικό υπόβαθρο. Τέτοιο υπόβαθρο μπορεί να είναι ένας χάρτης στον οποίο παρουσιάζεται η μεταβλητή που εξετάζουμε.

Σε περίπτωση που η χωρική διάσταση των στοιχείων, αντικειμένων, φαινομένων που μελετώνται είναι μεν υπαρκτή αλλά όχι καθοριστική, το υπόβαθρο μπορεί να είναι μη χωρικό, δηλαδή ένας κατάλογος από τα μέλη του οποίου επιλέγεται το δείγμα . Φυσικά, τα μέλη του καταλόγου ως σύνολο, δηλαδή ο στατιστικός πληθυσμός, θα εντοπίζονται στο χώρο του προβλήματος χωρίς όμως να είναι γνωστή η θέση τους.

Στο **χωρικό δειγματοληπτικό υπόβαθρο**, η θέση είναι ένα βασικό τμήμα της μεταβλητότητας καθεμιάς από τις παρατηρήσεις και συνεπώς πρέπει να αντιπροσωπεύει το συνολικό φάσμα της χωρικής μεταβλητότητας (spatial variability).

Αν για παράδειγμα πρέπει να υπολογιστεί το ποσοστό χρήσεων γης ανά κατηγορία σε μια περιοχή που απεικονίζεται σε ένα χάρτη, λαμβάνοντας ένα δείγμα από τον χάρτη της περιοχής θα πρέπει όλα τα τμήματα του χάρτη να αντιπροσωπεύονται στο δειγματοληπτικό υπόβαθρο. Μη χωρικά δειγματοληπτικά υπόβαθρα είναι αποδεκτά στη περίπτωση που το ενδιαφέρον εστιάζεται σε ένα χαρακτηριστικό της περιοχής ως σύνολο.

Στην περίπτωση των χωρικών δειγματοληπτικών υποβάθρων θα πρέπει να καθοριστεί το είδος των παρατηρήσεων, δηλαδή να οριστεί η χωρική δειγματοληπτική μονάδα. Βασικά υπάρχουν τρεις τέτοιες μονάδες: σημεία, γραμμές και επιφάνειες καννάβων (quadrats). Αυτό έχει ως αποτέλεσμα να υπάρχουν σημειακές, γραμμικές και επιφανειακές δειγματοληψίες καννάβων. Σε σχετική βιβλιογραφία το υπόβαθρο συναντάται και ως πλαίσιο.

2.7.3 Επιλογή δειγματοληπτικής μεθόδου

Το πρώτο στάδιο της δειγματοληπτικής διαδικασίας τελειώνει με την επιλογή του κατάλληλου δειγματοληπτικού υποβάθρου.

Ακολουθεί η επιλογή των παρατηρήσεων που θα αποτελέσουν το δείγμα. Οι επόμενες παράγραφοι αναφέρονται σε μεθόδους τυχαίας δειγματοληψίας (probability sampling) καθώς αυτές χρησιμοποιούνται περισσότερο, εξαιτίας και των βασικών πλεονεκτημάτων που παρουσιάζουν (τυχειότητα, μεγαλύτερη αντικειμενικότητα).

Η διαδικασία αυτή μπορεί να γίνει με τρεις βασικούς τρόπους : Τυχαία (απλή), Συστηματική, Στρωματοποιημένη. Υπάρχουν και διάφορες πιο σύνθετες μορφές, για πιο σύνθετα προβλήματα όπως η Ιεραρχική, η Πολυπαραγοντική, η Χωρο – Χρονική δειγματοληψία.

Στις επόμενες παραγράφους αναλύονται οι βασικές αρχές της κάθε μεθόδου και έπειτα επικεντρώνονται στην εφαρμογή της απλής τυχαίας και της στρωματοποιημένης σημειακής δειγματοληψίας.

2.8 Τυχαία Δειγματοληψία

Μια διαδικασία δειγματοληψίας ονομάζεται απλή τυχαία εάν σε ένα δείγμα μεγέθους n , όλοι οι δυνατοί συνδυασμοί των n στοιχειωδών μονάδων που μπορούν να σχηματιστούν από τον πληθυσμό N στοιχειωδών μονάδων έχουν την ίδια πιθανότητα επιλογής.

Η τυχαία δειγματοληψία αποτελεί τη βάση κάθε στατιστικής επεξεργασίας στοιχείων και αναφέρεται σε μια διαδικασία που πληροί δύο βασικές στατιστικές υποθέσεις:

- Τα μέλη ενός πληθυσμού έχουν ίσες πιθανότητες να επιλεγούν για το δείγμα.
- Η επιλογή ενός μέλους του πληθυσμού για το δείγμα με κανένα τρόπο δεν επηρεάζει την επιλογή ενός οποιουδήποτε άλλου.

Εξασφαλίζεται λοιπόν **ίση πιθανότητα επιλογής** για τις παρατηρήσεις και **ανεξαρτησία**.

Στη περίπτωση που η χωρική διάσταση ενός πληθυσμού φαινομένων, αντικειμένων είναι υπαρκτή αλλά όχι και καθοριστική (περιοχή ως σύνολο), όπως προαναφέρθηκε, το δειγματοληπτικό υπόβαθρο μπορεί να είναι ένας κατάλογος. Η επιλογή ενός τυχαίου δείγματος από το δειγματοληπτικό αυτό υπόβαθρο επιτυγχάνεται με τη χρήση ενός πίνακα τυχαίων αριθμών ο οποίος πληροί τις δύο συνθήκες τυχειότητας που αναφέρθηκαν προηγουμένως.

Στη περίπτωση που ο αρχικός πληθυσμός είναι άπειρος, η λήψη ενός στοιχείου με επανάθεση, δηλαδή επανατοποθέτηση του επιλεγέντος στοιχείου κατά την προηγούμενη επιλογή, η χωρίς επανάθεση, δεν αλλοιώνει την αρχική σύνθεση του πληθυσμού. Έτσι στη περίπτωση αυτή η λήψη του τυχαίου δείγματος συνεπάγεται ισονομία - ίση πιθανότητα επιλογής και ανεξαρτησίας μεταξύ τους. Η συνάρτηση συχνότητας του δείγματος, δηλαδή η πιθανότητα να ληφθούν οι τιμές n_1, n_2, n_v θα είναι:

$$f(n_1, n_2, \dots, n_v) = \prod_{i=1}^v f_n(n_i)$$

Εάν ο αρχικός πληθυσμός είναι πεπερασμένος, τότε στη περίπτωση που η δειγματοληψία γίνεται με επανάθεση, οι τυχαίες μεταβλητές θα είναι ισόνομες και ανεξάρτητες. Εάν όμως η δειγματοληψία γίνεται χωρίς επανάθεση, οι τυχαίες μεταβλητές δεν είναι στατιστικά ανεξάρτητες.

Χρήση Πινάκων Τυχαίων Αριθμών

Για τη λήψη τυχαίας δειγματοληψίας χρησιμοποιούνται ευρύτατα στη πράξη οι πίνακες τυχαίων αριθμών, οι οποίοι κατασκευάζονται με βάση την αρχή της τυχαίας εκλογής με επανάθεση. Για παράδειγμα για την λήψη ενός δείγματος, μεγέθους n , από ένα πεπερασμένο αριθμό N στοιχείων, με βάση την αρχή της τυχαίας εκλογής με επανάθεση, θα έπρεπε να τεθούν στη κάλπη N στοιχεία με συγκεκριμένη αρίθμηση και να γίνει εξαγωγή με επανάθεση των n στοιχείων του δείγματος. Για να αποφευχθεί η πρακτική επανάληψη του πειράματος χρησιμοποιούνται οι πίνακες των τυχαίων αριθμών, οι οποίοι περιέχουν, κατά σειρά, αριθμούς με ορισμένο πλήθος στοιχείων που επιλέχτηκαν με τυχαίο τρόπο. Στην περίπτωση μονοψήφιων αριθμών εξάγεται τυχαία ένα στοιχείο από τους αριθμούς 0, 1, 2, 3, 4, 5, 6, 7, 8, 9 με επανάθεση και σημειώνεται ο αριθμός κατά σειρά εξαγωγής. Για διψήφιους εξάγονται δύο στοιχεία και σημειώνονται οι αντίστοιχοι αριθμοί. Για τετραψήφιους τυχαίους μπορούν να χρησιμοποιηθούν ανά δύο οι διψήφιοι τυχαίοι που έχουν δημιουργηθεί κ.ο.κ. Για τη χρήση των πινάκων στη λήψη τυχαίας δειγματοληψίας ακολουθείται η εξής διαδικασία:

Αριθμούνται όλα τα στοιχεία του πληθυσμού από τον οποίο θα ληφθεί το δείγμα και στη συνέχεια επιλέγεται μία τυχαία αφετηρία στο πίνακα και λαμβάνεται ένα πλήθος τυχαίων αριθμών ίσο προς το μέγεθος του δείγματος.

Ο τρόπος κατασκευής των πινάκων τυχαίων αριθμών ποικίλει και δεν υπάρχει δυνατότητα σύγκρισης μεταξύ τους. Ως βασικό κριτήριο καταλληλότητας των πινάκων είναι η κατά προσέγγιση ίση συχνότητα εμφάνισης των 10 βασικών αριθμών. Η κατασκευή τους είναι ιδιαίτερα δύσκολη με τη χρήση Η/Υ.

Οι πίνακες τυχαίων αριθμών πληρούν τις δυο συνθήκες ίση πιθανότητα – ανεξαρτησία. Βέβαια, η μέθοδος αυτή είναι επίπονη και για το λόγο αυτό συναντά πρακτικές δυσκολίες σε μεγάλες εφαρμογές (π.χ. μεγάλα δείγματα). Ο πίνακας τυχαίων αριθμών υπάρχει στο παράρτημα.

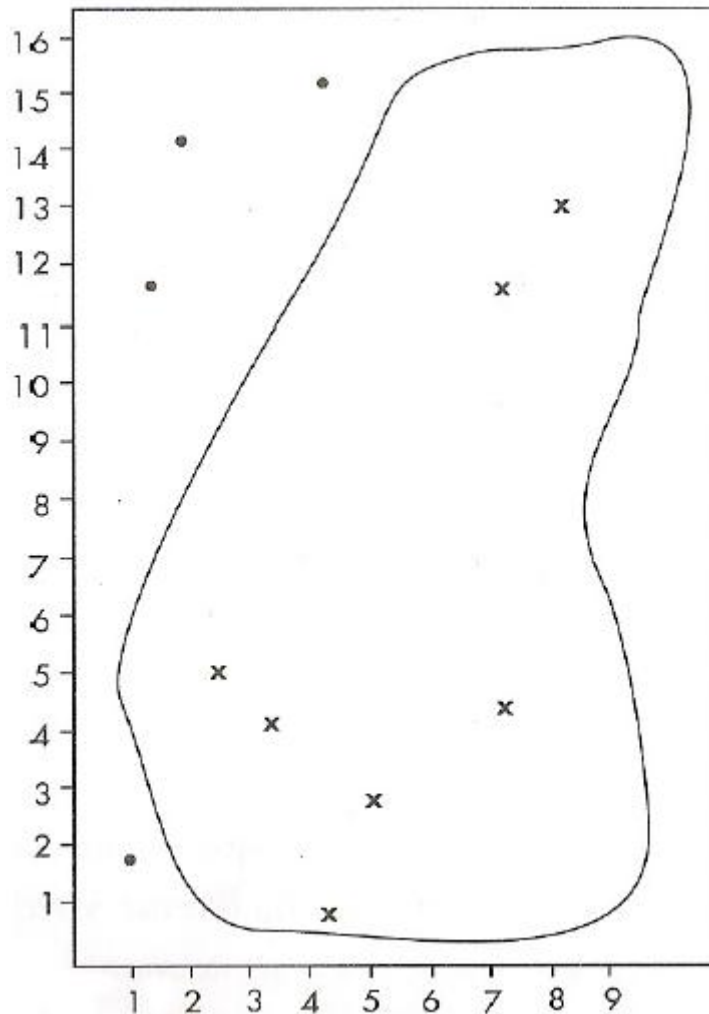
90057	12880	89522	81453	41294	14603
97431	01310	04978	25833	61253	48959
03346	61125	91434	84200	73664	45914
04812	76028	20044	78244	58015	05867
74532	60769	26928	29666	84318	21283
43016	56171	27506	72769	60939	50508
90227	88560	86908	32283	83958	92590
04809	23138	34852	45986	73823	08534
59332	70137	24535	17957	21012	51279
20195	14420	27684	35554	90414	80832
43144	85543	85639	91044	58951	85251
14952	50079	08851	56337	76769	83059
12700	60802	93248	12585	88835	97516
88126	63220	09217	89809	41827	63143
83873	30245	69153	36222	71188	50184

Σχήμα 2.1: Απόσπασμα Πίνακα Τυχαίων Αριθμών

2.8.1 Σημειακή Δειγματοληψία

Στη περίπτωση της σημειακής δειγματοληψίας κάθε μέλος του πληθυσμού αντιπροσωπεύεται από ένα ζεύγος συντεταγμένων (x , y). Χρησιμοποιώντας ένα σύστημα συντεταγμένων πάνω στο δειγματοληπτικό υπόβαθρο, δημιουργείται ένα σύνολο (πληθυσμός) από σημεία που ορίζονται από τις x και y συντεταγμένες τους. Στη συνέχεια με τη χρήση πινάκων τυχαίων αριθμών γίνεται η επιλογή του τυχαίου δείγματος.

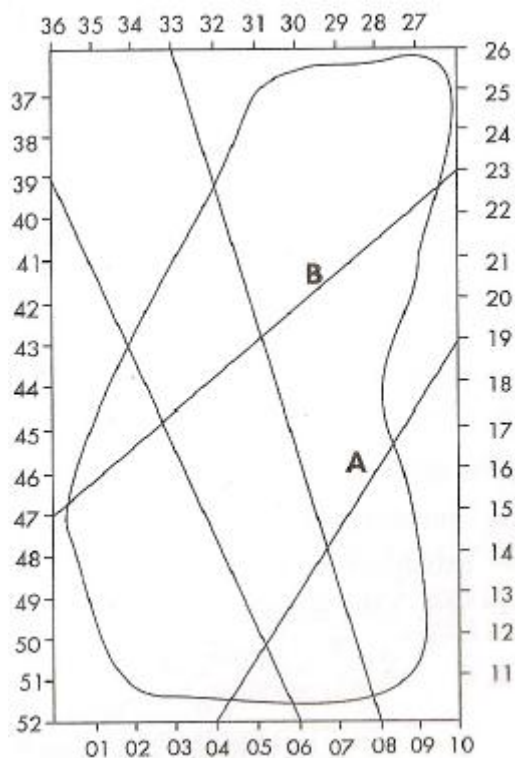
Όπως προαναφέρθηκε, ως αφετηρία για τη χρήση του πίνακα τυχαίων αριθμών επιλέγεται τυχαία ένας αριθμός και έπειτα με ένα προκαθορισμένο τρόπο (π.χ. κάθε νιοστός αριθμός διαγωνίως) επιλέγονται οι υπόλοιποι έως τη συμπλήρωση του δείγματος. Σε περίπτωση που ο αριθμός που λαμβάνεται με τον τρόπο που αναφέρθηκε δεν βρίσκεται εντός του υποβάθρου ή επιλέγεται για δεύτερη φορά το σημείο αυτό δεν συμπεριλαμβάνεται στο δείγμα. Η σημειακή δειγματοληψία χρησιμοποιείται στις περισσότερες περιπτώσεις χαρτών που απεικονίζουν γεωλογικούς σχηματισμούς ή χρήσεις γης.



Σχήμα 2.2: Σημειακή Τυχαία Χωρική Δειγματοληψία

2.8.2 Γραμμική Δειγματοληψία

Στη περίπτωση της γραμμικής δειγματοληψίας, η πληροφορία που αντλούμε λαμβάνεται μέσω ενός αριθμού γραμμών τυχαία κατανεμημένων στο χάρτη. Έτσι σε αντίθεση με τη σημειακή όπου ένα σημείο αντιστοιχεί σε μια χρήση γης, όταν ένας χάρτης από όπου αντλούμε τη πληροφορία είναι χάρτης χρήσεων γης, στη γραμμική θα μιλάμε για ποσοστό της γραμμής που περνά από τις διάφορες χρήσεις. Οι γραμμές προέρχονται από τυχαία επιλογή. Ένας τρόπος για τη δημιουργία τυχαίας γραμμικής δειγματοληψίας είναι να βαθμονομηθούν τα περιθώρια του υποβάθρου και στη συνέχεια με τον πίνακα τυχαίων αριθμών καθορίζεται η αρχή και το τέλος των γραμμών.



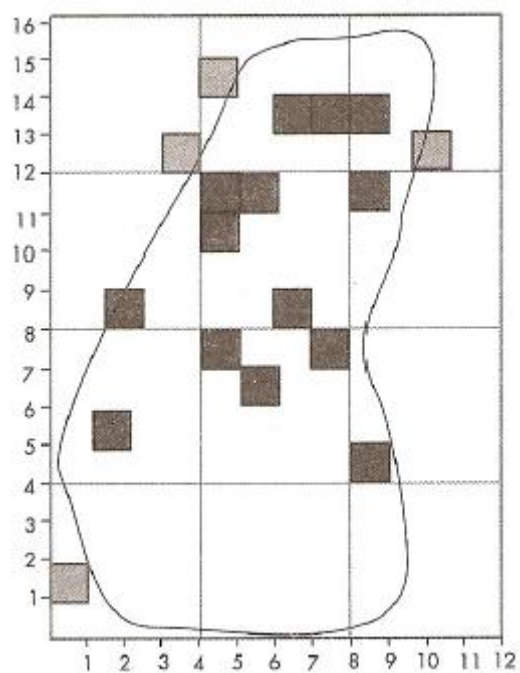
Σχήμα 2.3: Γραμμική Τυχαία Χωρική Δειγματοληψία

2.8.3 Επιφανειακή Δειγματοληψία

Στη περίπτωση που το ενδιαφέρον εστιάζεται σε χωρικά φαινόμενα όπου η επιφάνεια αποτελεί συστατικό στοιχείο μπορεί να χρησιμοποιηθεί ή επιφανειακή - κανναβική δειγματοληψία.

Στην επιφανειακή δειγματοληψία, το υπόβαθρο θεωρείται ότι αποτελείται από φατνία ενός καννάβου από όπου επιλέγεται τυχαία ένας μικρότερος αριθμός για να αποτελέσει το δείγμα. Αρχικά δημιουργείται ένα σύστημα συντεταγμένων που οριοθετεί το δειγματοληπτικό υπόβαθρο. Με το γνωστό τρόπο επιλέγονται από τον πίνακα τυχαίων αριθμών ζεύγη τιμών που αντιστοιχούν στη κάτω αριστερή γωνία του κάθε φατνίου.

Καθοριστικής σημασίας για το αποτέλεσμα της δειγματοληψίας είναι το σχήμα και το μέγεθος του κάθε φατνίου. Συνήθως χρησιμοποιείται τετράγωνο μικρού μεγέθους.



Σχήμα 2.4: Επιφανειακή Τυχαία Χωρική Δειγματοληψία

2.9 Συστηματική Δειγματοληψία

Δύο σπουδαία πλεονεκτήματα της τυχαίας δειγματοληψίας είναι η ανεξαρτησία και η ίση πιθανότητα επιλογής. Πρακτικά όμως η τυχαία δειγματοληψία παρουσιάζει το μειονέκτημα της ελλιπούς κάλυψης του δειγματοληπτικού υποβάθρου, με αποτέλεσμα κάποιες περιοχές του υποβάθρου να μην αντιπροσωπεύονται από σημεία δειγματοληψίας (στη σημειακή) και έτσι να υποβαθμίζεται η ύπαρξη τους.

Η συστηματική δειγματοληψία περιορίζει αυτό το μειονέκτημα της τυχαίας, αφού η επιλογή των επιμέρους παρατηρήσεων που αποτελούν το δείγμα γίνεται με ένα σταθερό και προκαθορισμένο (συστηματικό) τρόπο.

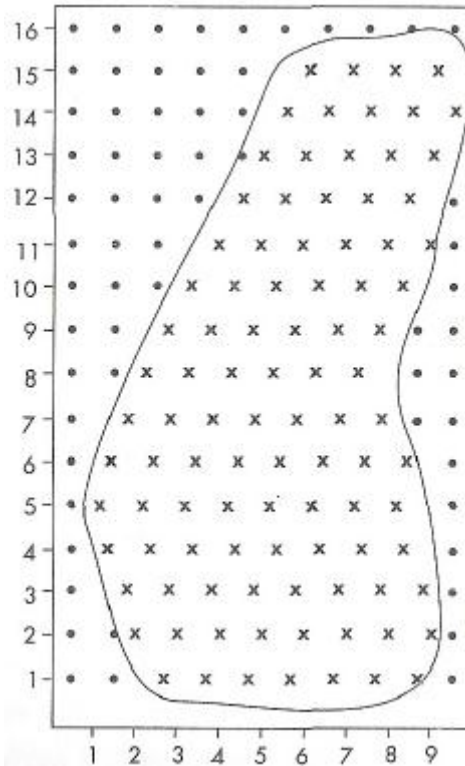
Στη συστηματική δειγματοληψία επιλέγεται ένα δείγμα μεγέθους n , από ένα πληθυσμό με N στοιχεία που είναι αριθμημένα από 1 έως N . Συνήθως, τίθεται η υπόθεση $N = v * k$, όπου k φυσικός αριθμός. Με τη χρήση τυχαίων αριθμών επιλέγεται ένας αριθμός $1 \leq j \leq k$, και διαχωρίζονται ως δείγμα εκείνα τα στοιχεία του πληθυσμού που έχουν αύξοντες αριθμούς $j + h * k$, $h = 0, 1, 2, \dots, v-1$.

Η συστηματική δειγματοληψία βρίσκει εφαρμογή σε μη χωρικά αλλά και σε χωρικά υπόβαθρα. Όταν το υπόβαθρο είναι ένας κατάλογος (μη χωρικό), επιλέγεται ένα σταθερό βήμα k , που ορίζεται ως το αντίστροφο της αναλογίας $\frac{\text{δείγμα } (n)}{\text{Πληθυσμός } (N)}$. Το δείγμα σχηματίζεται επιλέγοντας μια τυχαία αφετηρία, δηλαδή έναν αριθμό μεταξύ 1 και k και στη συνέχεια επιλέγεται κάθε φορά ο k επόμενος αριθμός.

Για την περίπτωση των χωρικών υποβάθρων που συνήθως είναι κάποιοι χάρτης ακολουθούν τα επόμενα:

2.9.1 Σημειακή Δειγματοληψία

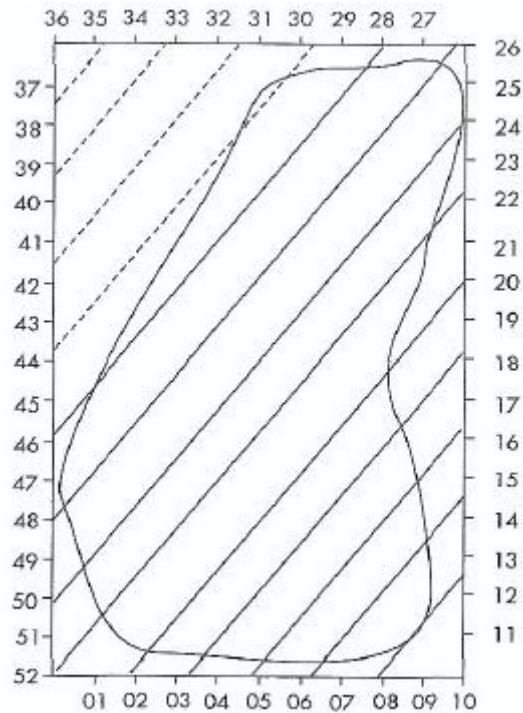
Η δειγματοληψία ξεκινά από ένα τυχαίο επιλεγμένο σημείο (δύο αριθμοί από πίνακα τυχαίων αριθμών) ώστε να προσδιοριστούν οι συντεταγμένες του. Τα υπόλοιπα σημεία επιλέγονται με συστηματικό τρόπο, για παράδειγμα κάθε σημείο σε ίση απόσταση, κατά τους άξονες x και y , από το άλλο.



Σχήμα 2.5: Συστηματική Σημειακή Χωρική Δειγματοληψία

2.9.2 Γραμμική Δειγματοληψία

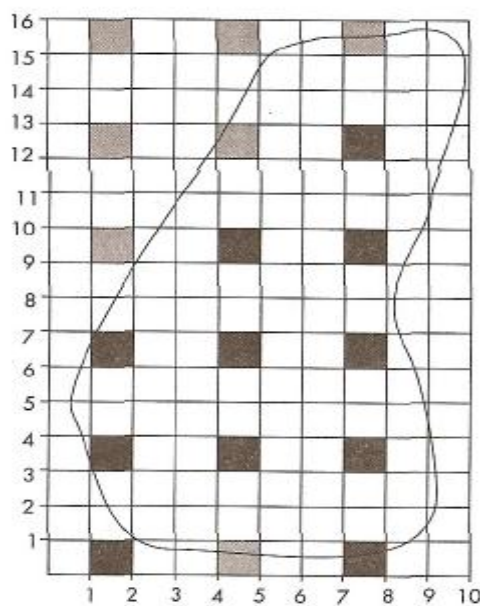
Ίδια μέθοδος ακολουθείται και στη γραμμική. Εδώ επιλέγεται τυχαία η πρώτη γραμμή και όλες οι υπόλοιπες γραμμές τοποθετούνται παράλληλες προς τη αρχική σε μια συγκεκριμένη απόσταση (προκαθορισμένος τρόπος) μετρούμενη στον άξονα της αρχής της πρώτης γραμμής.



Σχήμα 2.6: Συστηματική Γραμμική Χωρική Δειγματοληψία

2.9.3 Επιφανειακή (Κανναβική) Δειγματοληψία

Κατά αντιστοιχία με τα προηγούμενα στη περίπτωση της επιφανειακής συστηματικής, αρχικά επιλέγεται η πρώτη επιφάνεια του δείγματος τυχαία (όπως στην επιφανειακή τυχαία δειγματοληψία) και έπειτα τα επόμενα μέλη του δείγματος επιλέγονται με τέτοιο τρόπο ώστε η κάτω αριστερή γωνία κάθε φατνίου να βρίσκεται σε μια σταθερή και προκαθορισμένη απόσταση, κατά τους άξονες x και y , από το προηγούμενο φατνίο.



Σχήμα 2.7: Συστηματική Επιφανειακή Χωρική Δειγματοληψία

Η συστηματική δειγματοληψία παρέχει το σημαντικό πλεονέκτημα της ομοιόμορφης κάλυψης του δειγματοληπτικού υποβάθρου. Υπάρχουν βέβαια και δύο σημαντικά μειονεκτήματα.

Αν ο πληθυσμός περικλείει έναν περιοδικό τύπο διακύμανσης και εάν το διάστημα μεταξύ διαδοχικών μονάδων δειγματοληψίας συμπίπτει με το μήκος κύματος της περιοδικότητας μπορεί, να πάρουμε ένα δείγμα πάρα πολύ μεροληπτικό. Για παράδειγμα ένα συστηματικό δείγμα των κατοίκων μιας πόλης θα μπορούσε να περιέχει πάρα πολλές ή πολύ λίγες γωνιακές κατοικίες.

Το άλλο βασικό μειονέκτημα είναι ότι το αρχικό σημείο καθορίζει και την επιλογή των υπολοίπων σημείων. Έτσι μετά την τυχαία επιλογή του πρώτου σημείου, όλα τα σημεία του υποβάθρου δεν έχουν την ίδια πιθανότητα επιλογής, αφού τα υπόλοιπα επιλέγονται με συστηματικό τρόπο.

Ωστόσο η συστηματική δειγματοληψία συνήθως είναι μέρος ενός περισσότερο σύνθετου σχεδίου στο οποίο είναι δυνατό να λάβουμε αμερόληπτες εκτιμήσεις. Στα σχέδια αυτά η τυχαία διαδικασία λαμβάνει μέρος όχι μόνο για τον προσδιορισμό ενός αρχικού σημείου (στα χωρικά υπόβαθρα) ή ενός μέλους του καταλόγου (στα μη χωρικά) αλλά και καθ' όλη την εφαρμογή της.

2.10 Στρωματοποιημένη Δειγματοληψία (Απλή Τυχαία)

Στη στρωματοποιημένη δειγματοληψία ο πληθυσμός N διαιρείται σε k υποπληθυσμούς (στρώματα): $N_1, N_2, N_3, \dots, N_k$ έτσι ώστε να μην υπάρχει επικάλυψη μεταξύ των υποπληθυσμών (στρωμάτων) και το σύνολο των επιμέρους στρωμάτων να συνιστούν τον αρχικό πληθυσμό N . [$N_1 + N_2 + N_3 + \dots + N_k = N$]

Η διαδικασία κατά την οποία λαμβάνεται απλό τυχαίο δείγμα από κάθε στρώμα ξεχωριστά είναι γνωστή σαν στρωματοποιημένη δειγματοληψία (ή κατά στρώματα τυχαία).

Η στρωματοποιημένη δειγματοληψία είναι πολύ σημαντική όταν ο πληθυσμός αποτελείται από διακεκριμένες ομάδες που χαρακτηρίζονται από διαφορετικό μέγεθος ή χαρακτηριστικά .

Άλλες καταστάσεις που προκρίνουν τη στρωματοποιημένη δειγματοληψία για παράδειγμα για τον καθορισμό της έκτασης που καταλαμβάνουν τα καταστήματα λιανικής πώλησης, είναι πολύ λογικό να περιφερειοποιηθεί η πόλη με βάση την απόσταση από το κέντρο και μετά να ληφθούν δείγματα σε κάθε ζώνη. Και αυτό επειδή το ποσοστό γης που καταλαμβάνεται από τα καταστήματα ελαττώνεται, όσο κάποιος απομακρύνεται από το κέντρο και επομένως θα πρέπει και το μέγεθος του δείγματος να ελαττώνεται ανάλογα σε κάθε ζώνη ώστε να απεικονίζει την πραγματικότητα. Επιπλέον, η σύγκριση κάποιου **χωρικού** χαρακτηριστικού δύο ή περισσότερων περιοχών καθώς και η σύγκριση δύο μεταβλητών οι οποίες δεν διαφοροποιούνται ομοιόμορφα στο χώρο, αποτελούν παραδείγματα στα οποία η στρωματοποίηση του δείγματος είναι μια σωστή προσέγγιση.

Τα χαρακτηριστικά πλεονεκτήματα της στρωματοποιημένης δειγματοληψίας είναι δύο. Πρώτον, η συμμετοχή των διαφόρων υποομάδων του πληθυσμού στο δείγμα ώστε να είναι αντιπροσωπευτικό και δεύτερον η αύξηση της ποιότητας των εκτιμήσεων για τον πληθυσμό που είναι και σε τελική ανάλυση ο αποκλειστικός σκοπός κάθε δειγματοληψίας.

Στη χωρική δειγματοληψία που επιλέγεται η στρωματοποίηση υποβάθρου υπάρχουν δύο προσεγγίσεις:

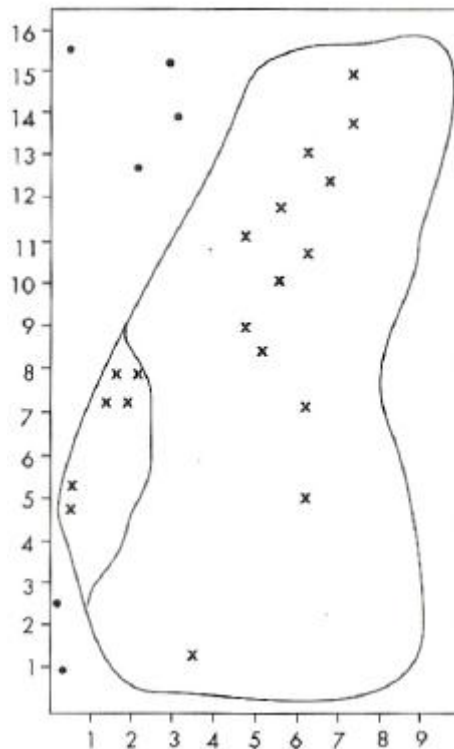
- **Φυσική στρωματοποίηση**
- **Κανναβική στρωματοποίηση**

Στη φυσική στρωματοποίηση μπορούν να διακριθούν υποπεριοχές γνωστής έκτασης στο δειγματοληπτικό υπόβαθρο. Οι υποπεριοχές αυτές αποτελούν τα στρώματα στα οποία θα γίνει η ανεξάρτητη λήψη δειγμάτων. Το μέγεθος του δείγματος σε κάθε υποπεριοχή είναι ανάλογο της επιφάνειάς της, η οποία λόγω της φυσικής ύπαρξής της είναι γνωστή, ενώ το είδος (βασική μορφή) της δειγματοληψίας εξαρτάται από το φαινόμενο που μελετάται.

Στη κανναβική στρωματοποίηση η περιοχή μελέτης επικαλύπτεται από ένα κάνναβο με φατνία ίσου μεγέθους και για κάθε υποπεριοχή – στρώμα υπολογίζεται το μέγεθος του δείγματος το οποίο καθορίζει τον αριθμό των σημείων σε κάθε φατνίο. Η επιλογή των σημείων σε κάθε φατνίο γίνεται με το γνωστό τρόπο τυχαίας δειγματοληψίας.

2.10.1 Στρωματοποιημένη Τυχαία Δειγματοληψία

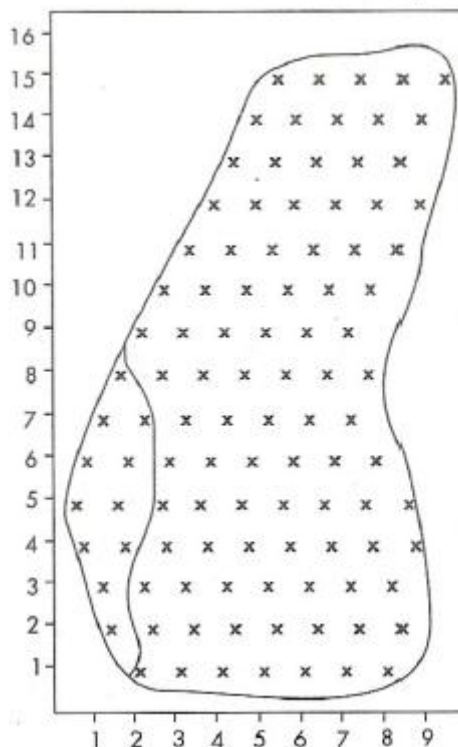
Στην περίπτωση αυτή το υπόβαθρο διαιρείται σε υποπεριοχές που θα αποτελέσουν και τα στρώματα της δειγματοληψίας. Στη συνέχεια σε καθεμιά από αυτές τις υποπεριοχές γίνεται τυχαία δειγματοληψία, με αριθμό σημείων – εάν πρόκειται για σημειακή που έχει ήδη καθοριστεί (π.χ. ανάλογα με την επιφάνεια κάθε στρώματος).



Σχήμα 2.8: Στρωματοποιημένη Τυχαία Δειγματοληψία

2.10.2 Στρωματοποιημένη Συστηματική Δειγματοληψία

Η δειγματοληψία αυτή διαφέρει από την προηγούμενη μόνο ως προς τον τρόπο με τον οποίο γίνεται η επιλογή του δείγματος στις υποπεριοχές - στρώματα που είναι συστηματική και όχι τυχαία καθώς και το ότι προσφέρει τα πλεονεκτήματα της συστηματικής δειγματοληψίας που έχουν αναφερθεί (π.χ. η πλήρης κάλυψη του δειγματικού υποβάθρου) με ότι αυτό συνεπάγεται.



Σχήμα 2.9: Στρωματοποιημένη Συστηματική Δειγματοληψία

2.10.3 Στρωματοποιημένη Συστηματική μη Γραμμική Δειγματοληψία

Είναι το πιο σύνθετο είδος δειγματοληψίας που συνδυάζει τα πλεονεκτήματα της τυχαίας επιλογής, της στρωματοποίησης και του συστηματικού δείγματος. Η σύνθετη αυτή διαδικασία εφαρμογής περιλαμβάνει αρχικά τη διαίρεση του υποβάθρου με βάση ένα σύστημα συντεταγμένων σε κάρναβο με φατνία ίσου μεγέθους. Το πρώτο σημείο επιλέγεται τυχαία στο πρώτο γωνιακό φατνίο με τυχαίο τρόπο. Τα υπόλοιπα σημεία επιλέγονται ένα σημείο σε κάθε φατνίο από τις συντεταγμένες του πρώτου σημείου που έχουν καθορισθεί (συστηματικότητα). Συγκεκριμένα η τετμημένη του πρώτου σημείου κρατιέται σταθερή για όλα τα σημεία που επιλέγονται στα φατνία της πρώτης σειράς ενώ η τεταγμένη τους επιλέγεται τυχαία από τους πίνακες των τυχαίων αριθμών. Παράλληλα, όλα τα σημεία της πρώτης στήλης φατνίων θα έχουν την ίδια τεταγμένη με αυτή του πρώτου σημείου και τετμημένες που επιλέγονται με τυχαίο τρόπο στο καθένα. Η επιλογή των επόμενων σημείων έχει πλέον καθορισθεί.

2.11 Εφαρμογή της Κανονικής Κατανομής

Στις περισσότερες περιπτώσεις δειγματοληψιών, τόσο ο αρχικός πληθυσμός όσο και τα επιλεγόμενα δείγματα, είναι αρκετά μεγάλα. Το γεγονός αυτό έχει ως αποτέλεσμα την θεμελίωση του Κεντρικού Οριακού Θεωρήματος που είναι ένα από τα πιο ουσιαστικά στοιχεία της Θεωρίας των Πιθανοτήτων και οδηγεί στην εφαρμογή της κανονικής κατανομής για ένα άθροισμα ανεξάρτητων τυχαίων μεταβλητών.

Η κανονική κατανομή είναι μία από τις πλέον χρήσιμες συνεχείς κατανομές στη θεωρία Πιθανοτήτων αλλά και στη Στατιστική. Η πρώτη ιστορικά εφαρμογή της κατανομής αυτής οφείλεται στον de Moivre (1733) ο οποίος διαπίστωσε ότι οι Διωνυμικές πιθανότητες προσεγγίζονται ικανοποιητικά από την Κανονική σ.π.π. Μια δεύτερη ιστορικά σημαντικά εφαρμογή της Κανονικής κατανομής οφείλεται στον Gauss (1777 – 1855), ο οποίος διαπίστωσε ότι τα τυχαία σφάλματα στις μετρήσεις μιας ποσότητας ακολουθούν Κανονική κατανομή. Για το λόγο αυτό η Κανονική κατανομή είναι επίσης γνωστή ως Γκαουσιανή.

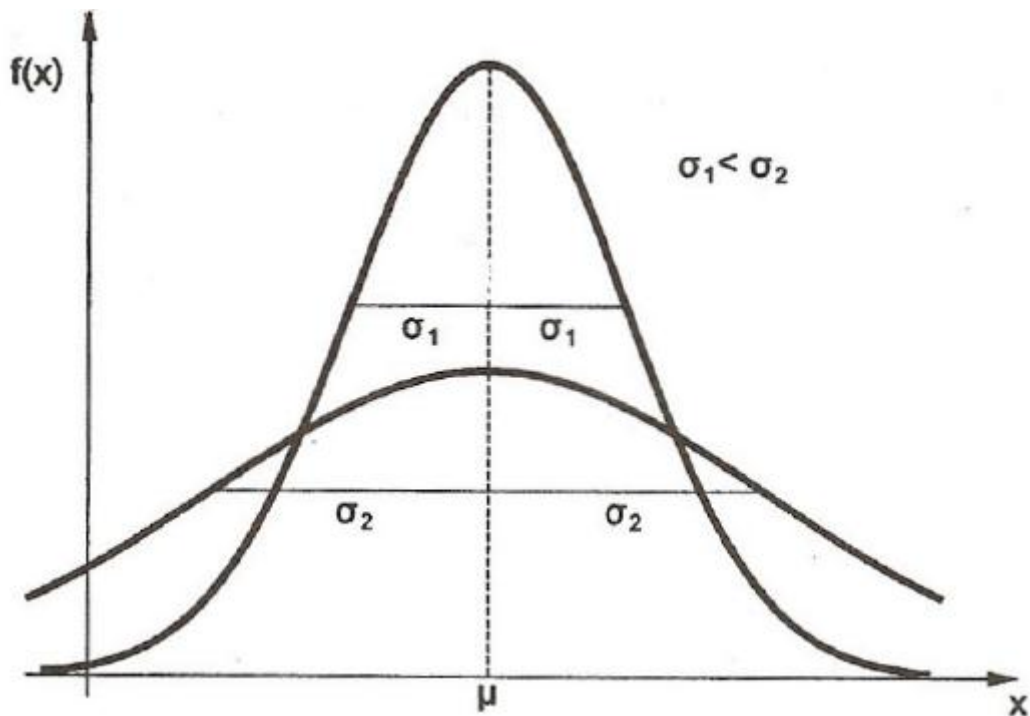
Η συνάρτηση πυκνοτήτων της κανονικής κατανομής είναι:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad \text{για } -\infty < x < +\infty$$

Οι παράμετροι μ , σ είναι η ακριβής τιμή της τυχαίας μεταβλητής και το τυπικό σφάλμα των μετρήσεων.

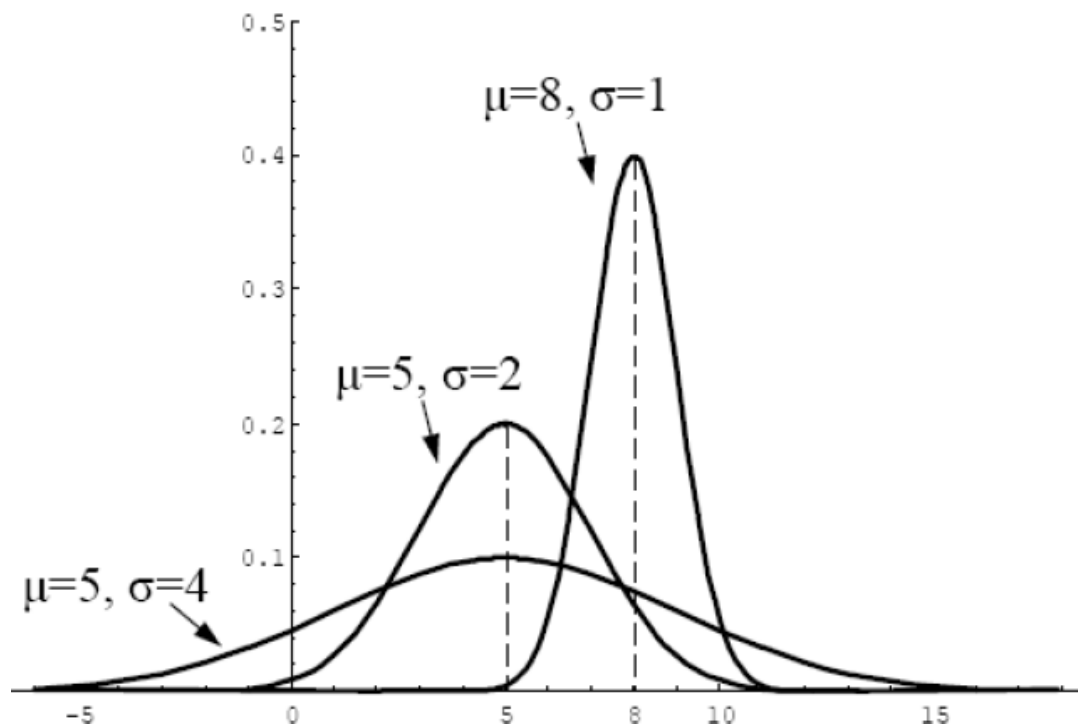
Η συνάρτηση κατανομής της κανονικής κατανομής είναι η:

$$F(x) = \int_{-\infty}^x f(x)dx = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx$$



Σχήμα 2.10: Γραφικές Παραστάσεις Συναρτήσεων Πυκνοτήτων Κανονικής Κατανομής ($\sigma_1 < \sigma_2$)

Η συνάρτηση $f(x)$ είναι συμμετρική ως προς την ακριβή τιμή μ , ενώ τα σημεία καμπής της βρίσκονται στις θέσεις $x = \mu - \sigma$ και $x = \mu + \sigma$.



Σχήμα 2.11: Γράφημα της Συνάρτησης Πυκνότητας Πιθανότητας για Διάφορα (μ, σ)

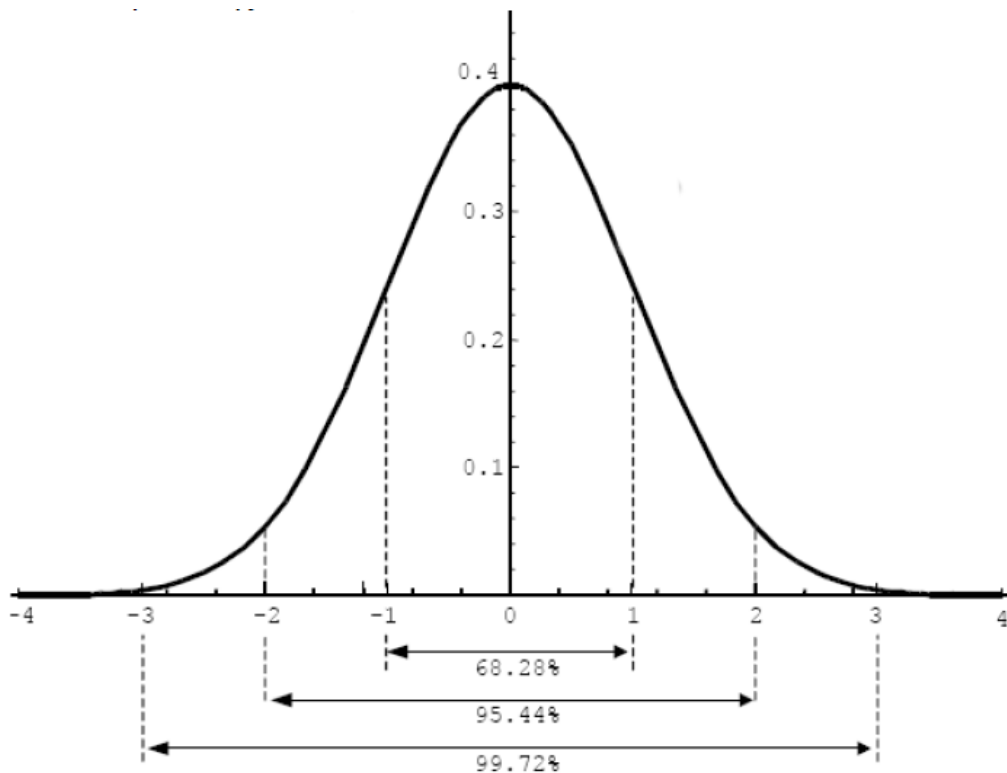
Τυποποιημένη Κανονική Κατανομή

Αν αντί της τυχαίας μεταβλητής x θεωρήσουμε την $Z = \frac{x - \mu}{\sigma}$ καταλήγουμε σε ακριβή τιμή $\mu = 0$ και μεταβλητότητα $\sigma = 1$ έτσι:

$$f(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}$$

$$F(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^z e^{-\frac{z^2}{2}} dz = F(x)$$

Η τυχαία μεταβλητή z ονομάζεται τυποποιημένη μεταβλητή και η συνάρτηση πυκνοτήτων της:



Σχήμα 2.12: Συνάρτηση Πυκνότητας Πιθανότητας της Τυπικής Κανονικής Κατανομής

Όπως φαίνεται η $f(z)$ προκύπτει από την $f(x)$ αν τεθεί $\mu = 0$ και $\sigma = 1$.

Έτσι οι $f(z)$ και $F(z)$ πινακοποιούνται για $z \geq 0$ και από αυτές υπολογίζονται οι πιθανότητες που ενδιαφέρουν.

Είναι προφανές ότι: $f(-z) = f(z)$ και $F(-z) = 1 - F(z)$ ενώ $P(-z_i \leq z \leq z_i) = 2F(z_i) - 1$

2.11.1 Μέτρα Ακρίβειας – Διαστήματα Εμπιστοσύνης

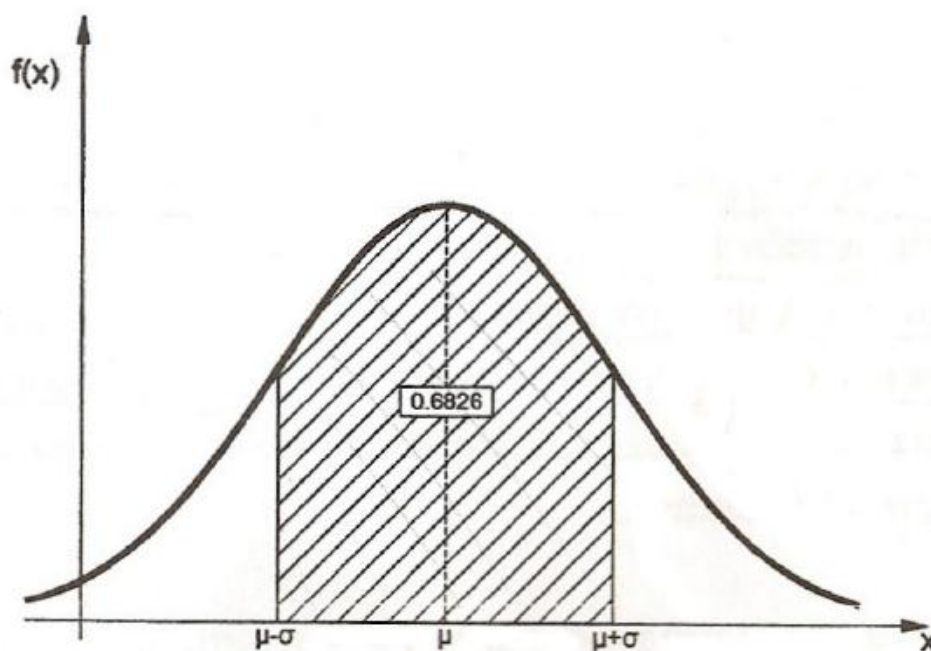
Ένα μέτρο λοιπόν της ακρίβειας των μετρήσεων είναι το τυπικό σφάλμα σ , γιατί όσο μεγαλώνει το σ τόσο πιο πεπλατυσμένη γίνεται η συνάρτηση πυκνοτήτων $f(x)$, που σημαίνει ότι η ακρίβεια είναι αντίστροφη του τυπικού σφάλματος.

Έτσι δικαιολογείται απόλυτα και το όνομα τυπικό σφάλμα σαν ένα μέγεθος που χαρακτηρίζει (τυποποιεί) μια σειρά μετρήσεων.

Θεωρώντας την τυποποιημένη μεταβλητή z προκύπτει:

$$P(\mu - \sigma < \chi < \mu + \sigma) = P(-1 < z < 1) = 0.6826 = 68\%$$

Από τις παραπάνω σχέσεις συμπεραίνεται πως η διαφορά μιας εκτίμησης χ της μεταβλητής από την ακριβή της τιμή θα κυμαίνεται μεταξύ $-\sigma$ και $+\sigma$ (αν χ είναι η μέση τιμή τότε σ θα είναι το τυπικό σφάλμα της μέσης τιμής) με πιθανότητα 68% που σημαίνει πως αν η εκτίμηση που έχουμε κάνει είναι χ η ακριβής τιμή της μεταβλητής θα βρίσκεται στο διάστημα $\chi - \sigma$ έως $\chi + \sigma$ με πιθανότητα 68%.



Σχήμα 2.13: Εύρεση Διαστήματος Τυπικού Σφάλματος Διαφοράς $\chi - \mu$ με πιθανότητα P

Σαν μέτρα ακρίβειας των μετρήσεων χρησιμοποιούνται και διαστήματα πολλαπλάσια του τυπικού σφάλματος, όπως το 2σ ή 3σ , για τα οποία βρίσκονται οι αντίστοιχες πιθανότητες.

$$P(\mu - 2\sigma < \chi < \mu + 2\sigma) = P(-2 < z < 2) = 95.4\%$$

$$P(\mu - 3\sigma < \chi < \mu + 3\sigma) = P(-3 < z < 3) = 99.7\%$$

Συχνά όμως παρουσιάζεται το πρόβλημα της εύρεσης του διαστήματος μέσα στο οποίο βρίσκεται η διαφορά $\chi - \mu$ με ορισμένη πιθανότητα P .

Η πιθανότητα ονομάζεται **επίπεδο εμπιστοσύνης** και το διάστημα που συνδέεται με αυτό **διάστημα εμπιστοσύνης**. (Πηγή: Αγατζά – Μπαλοδήμου Α.Μ., 2007)

z_p	$P = P(x - z\sigma < \mu < x + z\sigma) = 2F(z) - 1$
0.6745	0.500
1.000	0.683
1.645	0.900
1.960	0.950
2.000	0.954
2.576	0.990
3.000	0.997

Σχήμα 2.14: Τιμές z για Διάφορα Διαστήματα Εμπιστοσύνης

Π.χ. για επίπεδο εμπιστοσύνης $p = 99\%$ έχουμε $z_p = 2.576$ που σημαίνει διάστημα εμπιστοσύνης για τη διαφορά $\chi - \mu$ από -2.576σ έως 2.576σ ή

$$P(x - 2.576\sigma < \mu < x + 2.576\sigma) = 99\%$$

Το επίπεδο εμπιστοσύνης γενικά συμβολίζεται και με $p = 1 - \alpha$ όπου α η πιθανότητα να βρίσκεται η $\chi - \mu$ έξω από το διάστημα εμπιστοσύνης.

Ο όρος παράμετρος πληθυσμού (population parameter) χρησιμοποιείται για να δηλωθεί το χαρακτηριστικό ενός πληθυσμού, ενώ ο όρος εκτιμητής δείγματος (sample statistic) για να δηλωθεί το χαρακτηριστικό ενός δείγματος. Η δειγματοληψία στηρίζεται στον προσδιορισμό εκτιμητών του δείγματος που στη περίπτωση που το δείγμα αυτό είναι αντιπροσωπευτικό και αμερόληπτο θα αποτελεί εκτίμηση για την αντίστοιχη παράμετρο του πληθυσμού.

Ακριβής εκτίμηση των παραμέτρων του πληθυσμού (μ , σ) (ακριβής τιμή και τυπική απόκλιση του πληθυσμού) από τα M και s (μέση τιμή και τυπική απόκλιση του δείγματος), δηλαδή τους εκτιμητές του δείγματος είναι αδύνατη, όπως αντίστοιχα δύο δείγματα να έχουν ίδιες τιμές για τα M και s .

Η κατανομή των εκτιμητών των δειγμάτων (π.χ. η μέση τιμή των δειγμάτων που λαμβάνονται από έναν πληθυσμό) είναι η δειγματοληπτική κατανομή τους. Αν τα δείγματα είναι αντιπροσωπευτικά και αμερόληπτα αναμένεται ο κάθε εκτιμητής τους να κατανέμεται συμμετρικά γύρω από την αντίστοιχη παράμετρο του πληθυσμού.

Παράδειγμα: Έστω ότι από τον έλεγχο ενός τυχαίου δείγματος εκτιμάται ότι το 10% των τετραδίων μιας συγκεκριμένης έκδοσης είναι ελαττωματικά. Απέναντι σε αυτό το νούμερο, απόλυτα λογικές και αναμενόμενες είναι ερωτήσεις όπως: μπορούμε να ισχυριστούμε με βεβαιότητα ότι το πραγματικό ποσοστό ελαττωματικών τετραδίων κυμαίνεται μεταξύ 5 – 15 %; ή πόσο πιθανόν είναι το ποσοστό των ελαττωματικών τετραδίων να είναι μεταξύ 9 – 11%;

2.11.2 Εκτιμήσεις Διαστήματος Εμπιστοσύνης για Μετρήσεις

Η διαδικασία εκτίμησης τουλάχιστον για τις περιπτώσεις που αντιμετωπίζουν οι περισσότεροι ερευνητές, βασίζεται σε ένα θεμελιώδες θεώρημα που ονομάζεται το **θεώρημα του κεντρικού ορίου** (central limit theorem). Το θεώρημα αυτό λέει ότι αν θεωρηθεί ότι εάν ένα τυχαίο δείγμα n παρατηρήσεων ληφθεί από έναν πληθυσμό με ακριβή τιμή μ και μεταβλητότητα σ^2 , τότε εφόσον οι παρατηρήσεις n είναι πολλές, η δειγματοληπτική κατανομή των μέσων τιμών θα προσεγγίζει την κανονική κατανομή.

Έτσι το τυπικό σφάλμα όπως ονομάζεται αλλιώς η τυπική απόκλιση των δειγματοληπτικών κατανομών θα δίνεται από τη σχέση:

$$\sigma_M = \frac{\sigma}{\sqrt{n}}$$

όπου: σ_M : το τυπικό σφάλμα της δειγματοληπτικής κατανομής των M

σ : η τυπική απόκλιση του πληθυσμού

n : το μέγεθος του δείγματος

Στην τυποποιημένη της μορφή:

$$Z_M = \frac{M - \mu}{\sigma_M} = \frac{M - \mu}{\sigma / \sqrt{n}}$$

Η Z_M για επίπεδο εμπιστοσύνης $(100 - \alpha)\%$:

$$P(-Z_{\alpha/2} \leq Z_M \leq Z_{\alpha/2}) = 100 - \alpha$$

$$P(M - Z_{\alpha/2} \sigma_M \leq \mu \leq M + Z_{\alpha/2} \sigma_M) = 100 - \alpha$$

Έτσι το διάστημα μέσα στο οποίο βρίσκεται η πραγματική τιμή του πληθυσμού, είμαστε $(100 - \alpha)\%$ σίγουροι ότι είναι:

$$M \pm Z_{\alpha/2} \sigma_M = M \pm Z_{\alpha/2} (\sigma / \sqrt{n})$$

όπου $Z_{\alpha/2}$: η τιμή από τους πίνακες κανονικής κατανομής που αντιστοιχεί στην αθροιστική συχνότητα $100 - (\alpha/2)$.

Καθότι η τυπική απόκλιση του πληθυσμού δεν είναι γνωστή (σ) για μεγάλα δείγματα (μεγαλύτερο των 30 παρατηρήσεων) χρησιμοποιείται ένας εκτιμητής της τυπικής απόκλισης του πληθυσμού που είναι η τυπική απόκλιση του δείγματος (s) και έτσι έχουμε το διάστημα εμπιστοσύνης το οποίο δίνεται από τη σχέση:

$$M \pm Z_{\alpha/2} (s / \sqrt{n})$$

2.11.3 Εκτιμήσεις Διαστήματος Εμπιστοσύνης από Απαριθμήσεις

Οι παρατηρήσεις δεν αφορούν μετρήσεις (π.χ. μέσο μέγεθος επιφάνειας κατοικιών) αλλά απαριθμήσεις (π.χ. αριθμός κτιρίων σε χρήση “ αμιγής κατοικία “), τότε αυτό που ενδιαφέρει είναι η αναλογία των διαφόρων κατηγοριών στο δείγμα, έτσι ώστε να βγάλουμε ασφαλή συμπεράσματα για τις αναλογίες χρήσεων των κτιρίων ολόκληρου του πληθυσμού.

Η δειγματοληπτική κατανομή που εφαρμόζεται σε τέτοιες περιπτώσεις είναι η διωνυμική η οποία σχετίζεται με την επανάληψη τυχαίων γεγονότων που μπορούν να λάβουν δύο μόνο μορφές (π.χ. κτίρια σε βιομηχανική χρήση και κτίρια σε μη βιομηχανική χρήση).

Το διάστημα εμπιστοσύνης δίνεται από τη σχέση: (Πηγή: Κουτσόπουλος Κ., 2006)

$$np \pm Z_{\alpha/2} \sqrt{npq}$$

Επειδή το ενδιαφέρον εστιάζεται περισσότερο στις αναλογίες παρά στον ακριβή αριθμό το διάστημα εμπιστοσύνης δίνεται σε επί τοις εκατό αναλογίες ως εξής:

$$np \pm Z_{\alpha/2} \sqrt{100p \cdot 100q/n}$$

2.12 Μέγεθος του Δείγματος και Εκτίμηση του

Η απαρίθμηση και η απογραφή του πληθυσμού προσφέρει προφανώς τη μέγιστη ακρίβεια για όποιες παραμέτρους και χαρακτηριστικά που μελετώνται (φυσικά με την προϋπόθεση ότι είναι δυνατό να διενεργηθεί απογραφή από εξειδικευμένο και έμπειρο προσωπικό). Όπως έχει αναφερθεί στις περισσότερες περιπτώσεις είναι αδύνατο να πραγματοποιηθεί απαρίθμηση - απογραφή ολόκληρου του πληθυσμού διότι είναι αντισυμβατικό σε χρόνο και χρήμα. Αυτό έχει ως αποτέλεσμα την επιλογή ενός μικρού αριθμού των μελών του πληθυσμού για τη μελέτη των διαφόρων χαρακτηριστικών του (επιλογή δείγματος). Το δείγμα αυτό θα πρέπει να αντιπροσωπεύει τον πληθυσμό έτσι ώστε οι εκτιμήσεις που γίνονται για το σύνολο του πληθυσμού να είναι ασφαλείς. Καθώς, όμως, το δείγμα ελαττώνεται το περιθώριο λάθους στις εκτιμήσεις για τον πληθυσμό που γίνονται με βάση αυτό μεγαλώνει. Αυτό που πρέπει να καθορισθεί είναι το ελάχιστο μέγεθος του δείγματος που ικανοποιεί ορισμένες απαιτήσεις ακρίβειας.

Η μέθοδος αυτή απαιτεί τη συλλογή ενός δοκιμαστικού δείγματος που καλό είναι να περιλαμβάνει 30 παρατηρήσεις. Το δοκιμαστικό δείγμα μπορεί, και συνήθως αποτελεί, τμήμα του οριστικού δείγματος με την προϋπόθεση ότι η επιλογή του δεν έχει επιπτώσεις στην επιλογή των υπόλοιπων παρατηρήσεων για το οριστικό δείγμα. Η δοκιμαστική δειγματοληψία είναι μια δειγματοληψία μειωμένης ακρίβειας που δίνει μια εικόνα για τα χαρακτηριστικά του πληθυσμού που μελετάται και τα αποτελέσματά της χρησιμοποιούνται για τον υπολογισμό του οριστικού δείγματος.

Ο τρόπος εύρεσης του ελάχιστου μεγέθους του δείγματος διαχωρίζεται όπως προαναφέρθηκε στα διαστήματα εμπιστοσύνης ανάλογα με το αν η δειγματοληψία αναφέρεται σε μετρήσεις ή απαριθμήσεις.

Στη περίπτωση που η δειγματοληψία αναφέρεται σε **μετρήσεις** (π.χ. εκτίμηση μέσου μεγέθους ιδιοκτησιών), το απαιτούμενο μέγεθος δείγματος (ή ελάχιστο μέγεθος δείγματος) εξαρτάται από το **περιθώριο λάθους θ** που θα θέσουμε καθώς και το **διάστημα εμπιστοσύνης** που θα επιλέξουμε. Η δοκιμαστική δειγματοληψία που έχει προηγηθεί έχει δώσει τιμή για την τυπική απόκλιση. Σύμφωνα με τη βιβλιογραφία το απαιτούμενο μέγεθος δείγματος για μετρήσεις δίνεται από τον τύπο: (Πηγή: Κουτσόπουλος Κ., 2006)

$$n = (Z_{\alpha/2} \cdot s / \theta)^2$$

όπου:

n: το απαιτούμενο μέγεθος του δείγματος

$Z_{\alpha/2}$: η τιμή του Z για διάστημα εμπιστοσύνης (100 – α)%

S: η τυπική απόκλιση του δείγματος

θ : το περιθώριο λάθους

Στη περίπτωση που η δειγματοληψία αφορά απαριθμήσεις – αναλογίες (εκτίμηση ποσοστών χρήσεων γης μιας περιοχής), το απαιτούμενο μέγεθος δείγματος δίνεται από τον τύπο: (Πηγή: Κουτσόπουλος Κ., 2006)

$$n = 100 p \cdot 100 q \cdot (Z_{\alpha/2} / \theta)^2$$

όπου:

p και q: οι αναλογίες του δείγματος που ανήκουν και δεν ανήκουν στη συγκεκριμένη κατηγορία.

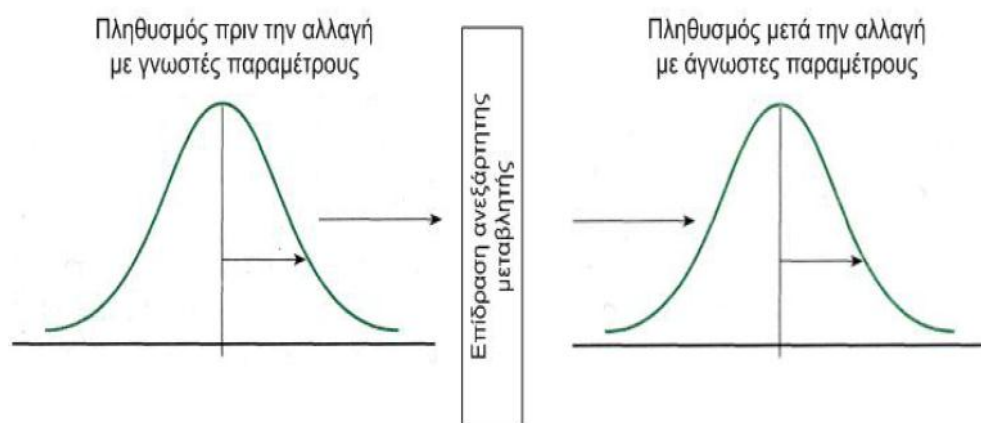
2.13 Έλεγχος Υποθέσεων – Test χ^2 (καλής προσαρμογής)

Από τη μελέτη ενός δείγματος μπορούμε να εξαγάγουμε κάποια συμπεράσματα σχετικά με τα χαρακτηριστικά του που τα ανάγουμε στο σύνολο του πληθυσμού είτε με τον υπολογισμό των σημειακών εκτιμητριών είτε με τη χρήση διαστημάτων εμπιστοσύνης.

Εναλλακτικά πληροφορίες που προέρχονται από ένα δείγμα μπορούν να χρησιμοποιηθούν ώστε να ελεγχθεί ως προς την αξιοπιστία της μια υπόθεση που ο ερευνητής έχει κάνει για τον πληθυσμό που εξετάζει.

Έτσι γίνεται μια υπόθεση αναφορικά με έναν πληθυσμό και τα συμπεράσματα για την ορθότητα αυτής της υπόθεσης βασίζονται στην επεξεργασία πληροφοριών που προέρχονται από το δείγμα.

Αν και υπάρχουν πολλές παραλλαγές της μεθόδου ελέγχου υποθέσεων για κάθε ερευνητική περίπτωση, η βασική λογική που ακολουθείται είναι πάντα η ίδια. Αξίζει να σημειωθεί ότι συνήθως ένας ερευνητής μελετά πληθυσμούς οι οποίοι προσεγγίζουν μία κανονική κατανομή (normal distribution), δηλαδή οι περισσότερες παρατηρήσεις συγκεντρώνονται γύρω από τη μέση τιμή, και ο αριθμός τους μειώνεται συμμετρικά και προς τις δύο μεριές.



Σχήμα 2.15: Η Λογική του Ελέγχου Υποθέσεων

Η λογική του ελέγχου υποθέσεων μπορεί να συνοψιστεί στα ακόλουθα τέσσερα βήματα: διατύπωση υποθέσεων, διαμόρφωση των κριτηρίων για τη λήψη μιας απόφασης, συλλογή δεδομένων από δείγματα και αξιολόγηση της μηδενικής υπόθεσης.

Η μηδενική υπόθεση είναι μία πρόταση που προβλέπει ότι η αλλαγή της ανεξάρτητης μεταβλητής (μεταβλητή που ελέγχω) δεν έχει καμία επίδραση στην εξαρτημένη μεταβλητή (μεταβλητή που παρατηρώ). Με αυτόν τον τρόπο μπορούμε να θεωρήσουμε γνωστές τις παραμέτρους του νέου πληθυσμού που δημιουργείται από την αλλαγή της τιμής της ανεξάρτητης μεταβλητής, καθώς υποθέτουμε ότι αυτή η αλλαγή δεν έχει καμία επίδραση. Επομένως οι τιμές των παραμέτρων του πληθυσμού μετά την αλλαγή της ανεξάρτητης μεταβλητής είναι ίδιες με αυτές πριν την αλλαγή.

Η δεύτερη υπόθεση που διατυπώνει ο ερευνητής ονομάζεται εναλλακτική υπόθεση και είναι η ακριβώς αντίθετη από τη μηδενική υπόθεση, δηλαδή δηλώνει ότι η αλλαγή της ανεξάρτητης μεταβλητής επιδρά στην εξαρτημένη μεταβλητή. Ο στόχος της μεθόδου ελέγχου υποθέσεων είναι να αποδείξει ότι η μηδενική υπόθεση μπορεί να απορριφθεί με κάποιο βαθμό βεβαιότητας.

Εκτός των παραμετρικών υποθέσεων, ιδιαίτερα χρησιμοποιούνται οι έλεγχοι καλής προσαρμογής (goodness of fit test) μια πρώτη προσέγγιση έγινε από τον Karl Pearson στις αρχές του 20^{ου} αιώνα.

Το κριτήριο που αφορά τον έλεγχο των υποθέσεων:

$H_0 : p_i = p_{oi} \ (i = 1, \dots, k)$ μηδενική υπόθεση

$H_1 : p_i \neq p_{oi}$ για ένα τουλάχιστον l εναλλακτική υπόθεση – απόρριψη μηδενικής
Όπου $p_i \ (i = 1, \dots, k)$ οι πιθανότητες k αμοιβαίως αποκλεισμένων ενδεχομένων A_i .

Σε n ανεξάρτητες επαναλήψεις εάν $N_i \ (i = 1, \dots, k)$ είναι οι αριθμοί εμφανίσεων των ενδεχομένων A_i $\left(\sum_{i=1}^k N_i = n \right)$, τότε η κατανομή της πολυδιάστατης τυχάιας μεταβλητής $N = (N_1, \dots, N_k)$ είναι η $(k - 1)$ διάστατη Πολυωνυμική κατανομή.

$$X^2 = \sum_{i=1}^k \frac{(O_i - E_i)^2}{E_i} \text{ (test statistic)}$$

όπου: O_i : η παρατηρούμενη συχνότητα της i κατηγορίας
 E_i : η αναμενόμενη / πραγματική συχνότητα της i κατηγορίας
 k : ο αριθμός των κατηγοριών

Περιοχή απόρριψης: $\chi^2 > \chi_a^2$ όπου χ_a^2 έχει $(k - 1)$ βαθμούς ελευθερίας

Σημαντική προϋπόθεση για να ισχύει η προσέγγιση του test: η αναμενόμενη τιμή για κάθε κατηγορία να είναι μεγαλύτερη από 5. $[E (n_i) \geq 5]$

Σε περίπτωση που η αναμενόμενη συχνότητα μιας κατηγορίας είναι μικρότερη από 5 συγχωνεύεται με άλλη κατηγορία.

Παράδειγμα

Συγκρίνουμε το μερίδιο αγοράς πριν και μετά από μια διαφημιστική καμπάνια για να δούμε αν υπάρχει διαφορά.

Μερίδιο αγοράς πριν τη διαφημιστική καμπάνια

ΕΤΑΙΡΕΙΕΣ	ΜΕΡΙΔΙΟ ΑΓΟΡΑΣ ΠΡΙΝ %
ΕΤΑΙΡΕΙΑ Α	45
ΕΤΑΙΡΕΙΑ Β	40
ΛΟΙΠΕΣ	15

200 πελάτες καταμετρήθηκαν μετά την εκστρατεία, τα αποτελέσματα είναι:

ΕΤΑΙΡΕΙΕΣ	Αριθμός Πελατών
ΕΤΑΙΡΕΙΑ Α	102
ΕΤΑΙΡΕΙΑ Β	82
ΛΟΙΠΕΣ	16

Εξετάζουμε αν η εκστρατεία ήταν αποτελεσματική σε συγκεκριμένο επίπεδο σημαντικότητας α .

- Η μηδενική υπόθεση (H_0): $P(\text{εταιρεία Α}) = 0.45$, $P(\text{εταιρεία Β}) = 0.40$, $P(\text{λοιπές}) = 0.15$.
- Η εναλλακτική υπόθεση (H_a): τουλάχιστον μια P_i διαφορετική.

Το στατιστικό τεστ που χρησιμοποιούμε είναι το test χ^2

Η περιοχή απόρριψης $\chi^2 > \chi_{\alpha}^2$

Τελικά βρίσκουμε $\chi^2 = 8.18$ και για επίπεδο σημαντικότητας $\alpha = 0.05$ και βαθμούς ελευθερίας $d = 3 - 1 = 2$, από πίνακα $\chi_{\alpha}^2 = 5.99$

Έτσι $\chi^2 > \chi_{\alpha}^2$ επομένως απορρίπτεται η μηδενική υπόθεση.

Υπάρχουν λοιπόν σαφείς ενδείξεις ότι η διαφήμιση άλλαξε το μερίδιο αγοράς.

ΚΕΦΑΛΑΙΟ 3: Γεωγραφικά Συστήματα Πληροφοριών

3.1 Γενικά

Τα Γεωγραφικά Συστήματα Πληροφοριών (Γ.Σ.Π.) είναι ευρέως γνωστά με την ονομασία GIS (Geographical Information Systems) που είναι συστήματα που παρέχουν ένα σύνολο εργαλείων κατάλληλων για τη διαχείριση χωρικών δεδομένων (spatial data) και συσχετισμένων ιδιοτήτων. Ένα ψηφιακό σύστημα έχει την ικανότητα να ενσωματώσει, αποθηκεύσει, προσαρμόσει, αναλύσει και παρουσιάσει γεωγραφικά συσχετισμένες πληροφορίες. Σε μια γενική όψη ένα Γ.Σ.Π. είναι ένα εργαλείο “ **έξυπνου χάρτη** ” το οποίο επιτρέπει στους χρήστες του να αποτυπώσουν μια περίληψη του πραγματικού κόσμου, να δημιουργήσουν ερωτήσεις χωρικού ή περιγραφικού χαρακτήρα, να αναλύσουν τα χωρικά δεδομένα (spatial data), να τα προσαρμόσουν και να τα αποδώσουν σε αναλογικά μέσα (εκτυπώσεις χαρτών και διαγραμμάτων) ή σε ψηφιακά μέσα (αρχεία χωρικών δεδομένων, διαδραστικοί χάρτες στο Διαδίκτυο).

Τα GIS αποτυπώνουν χωρικά δεδομένα σε γεωγραφικό ή χαρτογραφικό ή καρτεσιανό σύστημα συντεταγμένων. Κύριο χαρακτηριστικό των συστημάτων αυτών είναι η σύνδεση των χωρικών με των περιγραφικών δεδομένων. Για παράδειγμα μια ομάδα σημείων που αναπαριστά θέσεις πόλεων συνδέεται με ένα πίνακα όπου κάθε εγγραφή εκτός από τη θέση περιέχει πληροφορίες όπως ονομασία, πληθυσμός κ.τ.λ..

Τα Γεωγραφικά Συστήματα Πληροφοριών είναι πληροφοριακά συστήματα που παρέχουν τη δυνατότητα συλλογής, διαχείρισης, αποθήκευσης, επεξεργασίας, ανάλυσης και οπτικοποίησης σε ψηφιακό περιβάλλον των δεδομένων που σχετίζονται με το χώρο. Τα δεδομένα αυτά ονομάζονται γεωγραφικά ή χαρτογραφικά ή χωρικά (spatial) και μπορεί να συσχετίζονται με μια σειρά από περιγραφικά δεδομένα τα οποία και τα χαρακτηρίζουν μοναδικά.

Χαρακτηριστική ιδιότητα που προσφέρουν τα GIS είναι αυτή της σύνδεσης της χωρικής με την περιγραφική πληροφορία η οποία από μόνη της δεν έχει χωρική υπόσταση. Η τεχνολογία που χρησιμοποιείται για αυτή τη λειτουργία βασίζεται:

- Είτε στο σχεσιακό μοντέλο δεδομένων (relational) όπου τα περιγραφικά δεδομένα πινακοποιούνται χωριστά και αργότερα συσχετίζονται με τα χωρικά δεδομένα μέσω κάποιων μοναδικών τιμών που είναι κοινές και στα δύο είδη δεδομένων.
- Είτε στο αντικειμενοστραφές (object – oriented) μοντέλο δεδομένων, όπου τόσο τα χωρικά όσο και τα περιγραφικά δεδομένα συγχωνεύονται σε αντικείμενα τα οποία μπορεί να μοντελοποιούν κάποια αντικείμενα με φυσική υπόσταση (π.χ. κατηγορία = “ δρόμος “, όνομα = “ Πανεπιστημίου “, γεωμετρία = “ [X1, Y1], [X2, Y2]... “, πλάτος = “ 20 μέτρα “).

Τα τελευταία χρόνια το αντικειμενοστραφές μοντέλο τείνει να χρησιμοποιείται περισσότερο σε εφαρμογές GIS εξαιτίας των αυξημένων δυνατοτήτων του σε σχέση με το σχεσιακό μοντέλο καθώς και για τη δυνατότητα παροχής για εύκολη και απλοποιημένη μοντελοποίηση σύνθετων φυσικών φαινομένων και αντικειμένων με χωρική διάσταση.

Πολλές φορές η ολοκληρωμένη έννοια των GIS (integrated GIS concept) επεκτείνεται για να συμπεριλάβει τόσο τα δεδομένα (που αποτελούν ουσιαστικά τον πυρήνα τους), το λογισμικό και τον μηχανικό εξοπλισμό, όσο και τις διαδικασίες και το ανθρώπινο δυναμικό, που αποτελούν αναπόσπαστα τμήματα ενός οργανισμού, ο οποίος έχει σαν πρωταρχική του δραστηριότητα την διαχείριση πληροφορίας με την βοήθεια GIS.

3.2 Τα Μέρη ενός Γ.Σ.Π.

Τα Γ.Σ.Π. έχουν τρία βασικά χαρακτηριστικά τα οποία βρίσκονται σε συνεχή ισορροπία και αλληλεξάρτηση. Αυτά είναι τα μηχανήματα (hardware) οι αλγόριθμοι (software) και τα διαθέσιμα (resourceware).

Μηχανήματα (hardware)

Τα μηχανικά μέρη ενός Γ.Σ.Π. είναι τρία: η κεντρική μονάδα (CPU), τα περιφερειακά και το τερματικό (V.D.U.). Η κεντρική μονάδα επεξεργασίας είναι η καρδιά του συστήματος και εξυπηρετεί όλες τις υπολογιστικές διαδικασίες. Τα περιφερειακά διακρίνονται σε αυτά της εισόδου που επιτρέπουν την είσοδο των στοιχείων (π.χ. ψηφιοποιητές και σαρωτές), σε περιφερειακά εξόδου που συμμετέχουν στην παρουσίαση των στοιχείων (π.χ. σχεδιαστές) και σε περιφερειακά διαχείρισης που βοηθούν στην αποθήκευση και διαχείριση των στοιχείων (disk and tape drivers).

Αλγόριθμοι (software)

Οι αλγόριθμοι μπορεί να κατηγοριοποιηθούν σε έξι βασικές ομάδες:

- Εισαγωγή δεδομένων και έλεγχο – επαλήθευσή τους που καλύπτει τις ανάγκες μετασχηματισμού των στοιχείων από την αρχική τους μορφή (χάρτες, τηλεσκοπικά προϊόντα κ.λ.π.) σε αναγνωρίσιμη ψηφιακή μορφή.
- Διαχείριση της γεωγραφικής βάσης δεδομένων που αναφέρεται στον τρόπο δόμησης και οργάνωσης των χωρικών και μη χωρικών δεδομένων.
- Εξαγωγή πληροφοριών και παρουσίαση που εστιάζεται στην παρουσίαση στοιχείων και αποτελεσμάτων των αναλυτικών διαδικασιών.
- Μετασχηματισμό των δεδομένων που στοχεύουν στο συντονισμό και την ανάλυση των στοιχείων.
- Αναζήτηση που βοηθούν το χρήστη να επικοινωνεί με τον Η/Υ αναζητώντας λύση μέσα από μια σειρά ερωτήσεων.
- Εργαλεία χωρικής ανάλυσης απαραίτητα για την κάλυψη αναγκών σε εφαρμογές που αναφέρονται στην ανάλυση χώρου.

Σε γενικές γραμμές τα Γ.Σ.Π. μπορούν να παραλληλισθούν με μια εργαλειοθήκη. Μερικά από τα εργαλεία ή συνδυασμοί τους μπορούν να χρησιμοποιηθούν για τη επίλυση αρκετών χωρικών προβλημάτων. Σύμφωνα με τον P. Burrough τα κυριότερα χωρικά προβλήματα που επιλύει ένα Γ.Σ.Π. είναι τα παρακάτω:

- Τι είναι στη θέση A;
- Που είναι τα X1 και X2;
- Που είναι το A σε σχέση με το B;
- Πόσο μεγάλο είναι το A (εμβαδόν, περίμετρος);
- Ποιο είναι το αποτέλεσμα της αλληλεπίθεσης μεταξύ δύο ή περισσότερων επιπέδων (overlay);
- Πόσες εμφανίσεις τύπου A βρίσκονται σε απόσταση ρ από το B;
- Προσδιορισμός της ζώνης σε απόσταση ρ από το A (buffering);

Τα κυριότερα εργαλεία που είναι ενσωματωμένα στα περισσότερα λογισμικά πακέτα Γ.Σ.Π. είναι η αλληλεπίθεση και η δημιουργία ζώνης. Καθώς όμως εξελίσσονται αυτά τα πακέτα συναντάμε περισσότερα και πιο ευέλικτα εργαλεία.

Διαθέσιμα (resourceware)

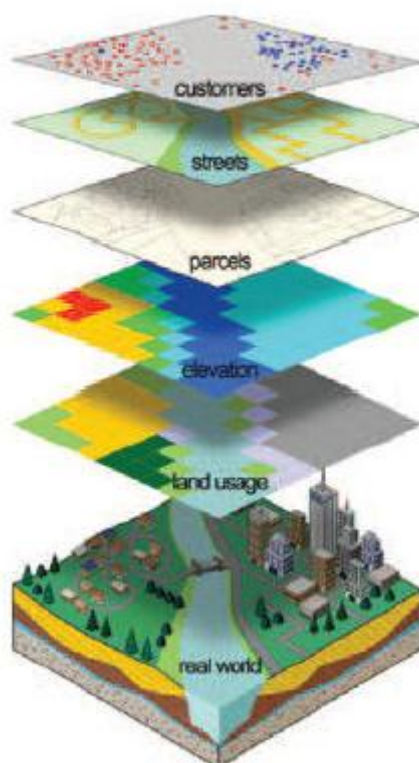
Η αγορά ενός υπολογιστικού συστήματος με το αναγκαίο λογισμικό δεν εξασφαλίζει καμία επιτυχία σε οποιαδήποτε προσπάθεια αν δεν υπάρχουν τα κατάλληλα στοιχεία, οι εξειδικευμένοι χειριστές και αναλυτές χώρου και βέβαια ένας οργανισμός που να υποστηρίζει το σύνολο των διαδικασιών που απαιτεί η χρήση ενός Γ.Σ.Π..

Τέλος, για την ομαλή λειτουργία ενός Γ.Σ.Π. χρειαζόμαστε πολλά περισσότερα από το υλικό μέρος, το λογισμικό και τα δεδομένα. Χρειαζόμαστε ανθρώπινο δυναμικό που είναι ο σημαντικότερος παράγοντας για την ομαλή λειτουργία του συστήματος.

3.3 Δομές Δεδομένων

Σε ένα Σ.Γ.Π. τα χωρικά δεδομένα αναπαρίστανται με δύο βασικές δομές: την διανυσματική και ψηφιδωτή δομή. Σε όλα τα Σ.Γ.Π. οι δύο δομές αποδίδονται ταυτόχρονα σε κοινές απεικονίσεις ενώ πολλά λογισμικά πακέτα GIS προσφέρουν τη δυνατότητα μετάβασης από τη μία δομή στην άλλη.

Ένας από τους πιο συνηθέστερους τρόπους δόμησης γεωγραφικών και χαρτογραφικών δεδομένων είναι σε θεματικά επίπεδα (layers) δηλαδή κάθε επίπεδο αντιπροσωπεύει μια θεματική προσέγγιση για ένα συγκεκριμένο σκοπό ή ομάδα αναγκών μπορεί να περιλαμβάνει ένα ή περισσότερα είδη πληροφοριών.



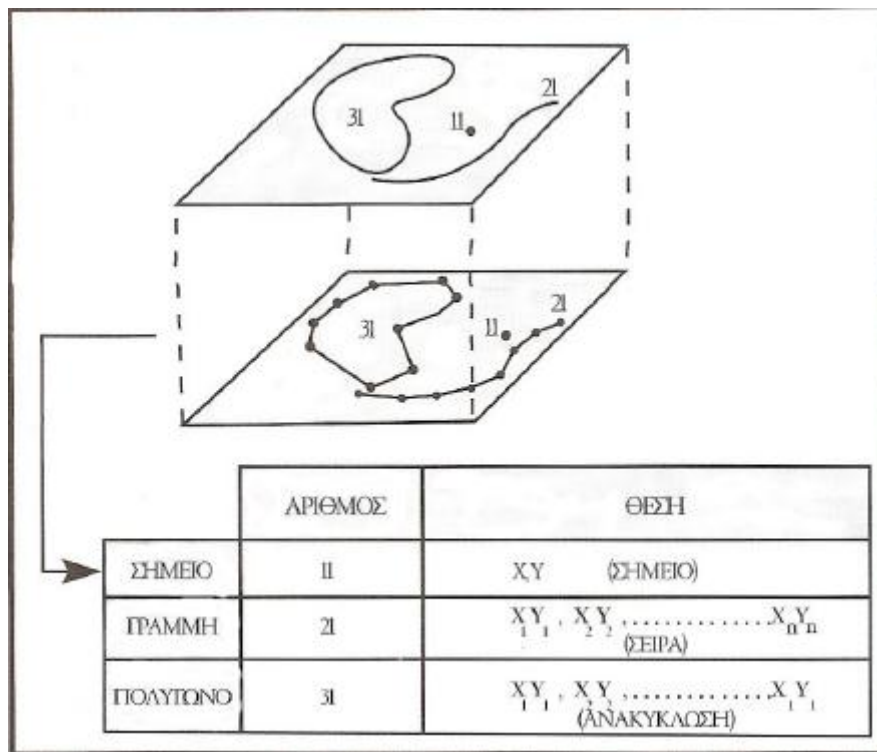
Σχήμα 3.1: Οργάνωση σε layers (θεματικά επίπεδα)

Διανυσματική Δομή Δεδομένων (vector data structure): Τα χωρικά δεδομένα αποδίδονται με τρεις βασικούς τύπους γεωμετριών **σημεία**, **γραμμές** και **πολύγωνα**. Ένα διανυσματικό σύστημα απεικονίζει τα δεδομένα ως σύνολα χωρικών αρχετύπων. Στα δισδιάστατα μοντέλα τα χωρικά αρχέτυπα είναι σημεία, γραμμές και πολύγωνα ενώ στα τρισδιάστατα χρησιμοποιούνται επιπλέον οι επιφάνειες και οι όγκοι. Η χρήση του κάθε χωρικού αρχέτυπου εξαρτάται από τη κλίμακα παρατήρησης ή το επίπεδο γενίκευσης. Σε μικρή κλίμακα οι πόλεις αναπαρίστανται από σημεία, ενώ οι ποταμοί και οι δρόμοι αναπαρίστανται από γραμμές. Καθώς, όμως αυξάνεται η κλίμακα παρατήρησης, πρέπει να λαμβάνεται υπόψη και η επιφανειακή διάσταση των φαινομένων. Συνεπώς, σε μεγαλύτερες κλίμακες, η πόλη αναπαρίστανται από ένα πολύπλοκο σύνολο χωρικών αρχετύπων που ορίζουν τα όρια των κτιρίων, τους δρόμους, τα πάρκα και τα άλλα φυσικά και τεχνητά φαινόμενα από τα οποία αποτελείται.

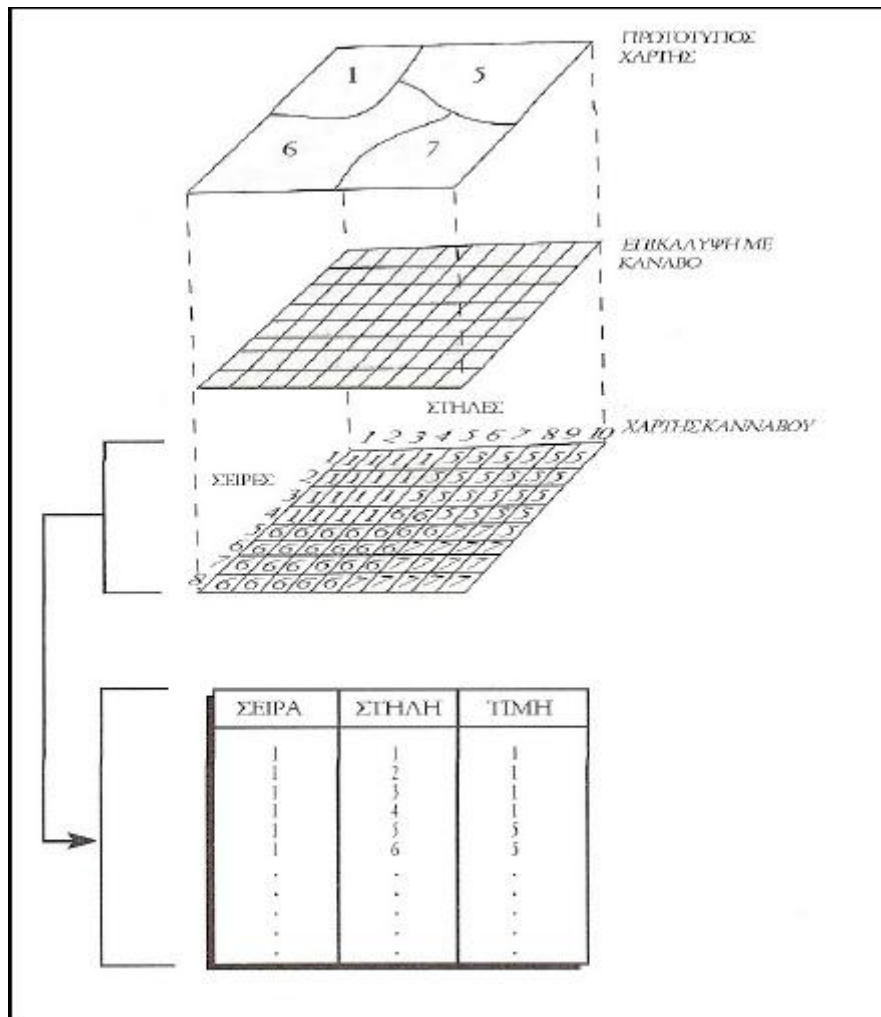
Οι γραμμές ή τα τόξα είναι σειρές διατεταγμένων σημείων των οποίων καταγράφονται οι συντεταγμένες. Οι επιφάνειες ή τα πολύγωνα αποθηκεύονται επίσης ως διατεταγμένες σειρές σημείων, όμως το αρχικό και τελικό σημείο ταυτίζονται έτσι ώστε να ορίζεται ένα κλειστό σχήμα. Τα σημεία σε μια διανυσματική δομή δεδομένων περιγράφονται ως γραμμές μηδενικού μήκους.

Κανονικοποιημένη Δομή Δεδομένων (raster data structure): Αντιλαμβάνεται το χώρο ως μωσαϊκή διαίρεση φατνίων – ψηφίδων (tessellation of cells), δηλαδή τοποθετεί και αποθηκεύει τα δεδομένα χρησιμοποιώντας ένα πίνακα ή ένα κάρναβο φατνίων. Συνήθως αναφερόμαστε στο φατνίο με τον όρο εικονοστοιχείο (pixel = picture element).

Πρέπει επίσης να σημειωθεί ότι η κατηγορία των κανονικοποιημένων χωρικών δομών δεν περιλαμβάνει μόνο τις δομές που βασίζονται σε ένα ορθογωνικό ή τετραγωνικό πλέγμα. Στην πραγματικότητα αυτή η κατηγορία περιλαμβάνει κάθε συνεχώς επαναλαμβανόμενο πρότυπο ενός κανονικού πολυγώνου ή πολυέδρου. Ο όρος που χρησιμοποιείται στη γεωμετρία γι' αυτό το σκοπό είναι η κανονική μωσαϊκή διαίρεση (regular tessellation). Πιο συγκεκριμένα, οι τρεις τύποι κανονικών διαιρέσεων που χρησιμοποιούνται στα χωρικά μοντέλα δεδομένων είναι οι τετραγωνικές, οι τριγωνικές και οι εξαγωνικές. Καθεμία από αυτές έχει διαφορετικά λειτουργικά χαρακτηριστικά που βασίζονται στη διαφορετική γεωμετρία των στοιχειωδών πολυγώνων.



Σχήμα 3.2: Διανυσματική Δομή Δεδομένων



Σχήμα 3.3: Κανονικοποιημένη Δομή Δεδομένων

Τοπολογία

Η τοπολογία είναι μια από τις σημαντικότερες σχέσεις που υποστηρίζονται στις χωρικές βάσεις δεδομένων. Μπορεί να οριστεί ως η μαθηματική διεργασία της συνδεσιμότητας (connectivity) και της γειτνίασης (adjacency) των σημείων και γραμμών που καθορίζει τις χωρικές σχέσεις σε ένα Σ.Γ.Π.. Η τοπολογική δομή στο διανυσματικό σύστημα προσδιορίζει με ακρίβεια πώς και πού τα σημεία και οι γραμμές ενός χάρτη συνδέονται μέσω των **κόμβων** (τοπολογικά σημεία σύνδεσης). Η σειρά της συνδεσιμότητας ορίζει το σχήμα ενός τόξου ή ενός πολυγώνου.

3.4 Πηγές Δεδομένων για τα ΣΓΠ

Τα γεωγραφικά δεδομένα μπορούν να ταξινομηθούν σε δύο κατηγορίες:

- **Δεδομένα για το περιβάλλον και φυσικά διαθέσιμα:**
 - Θεματικά (έδαφος, χρήση / κάλυψη γης, βλάστηση, γεωλογία).
 - Τοπογραφικά (χώροι καθορισμένης χρήσης, κτίρια, εγκαταστάσεις, τοπογραφικά δεδομένα, ακτές, υδρολογία, διοικητικά όρια, δίκτυα μεταφορών / υποδομών, ανάγλυφο).
- **Κοινωνικο – οικονομικά δεδομένα:**
 - Πληθυσμιακές / οικονομικές / γεωργικές απογραφές, στατιστικές (εργατικό δυναμικό, εγγραφές κτηματολογίου), αρχεία μεταφορών και παροχών, διοικητικά αρχεία).

Τα δεδομένα αυτά συλλέγονται από διάφορες πρωτογενείς και δευτερογενείς πηγές δεδομένων από τις οποίες οι πιο σημαντικές επιγραμματικά είναι:

- Τοπογραφικές μετρήσεις
- Κτηματολόγιο
- Δορυφορικές εικόνες
- Υπάρχοντες χάρτες
- Γραφικά χαρακτηριστικά
- Διαθέσιμα ψηφιακά δεδομένα
- Αεροφωτογραφίες

3.5 Επεξεργασία Δεδομένων – Χωρική Ανάλυση

Η επεξεργασία και η χωρική ανάλυση των δεδομένων αναφέρεται σ' ένα ευρύ φάσμα αναλυτικών εργασιών στις οποίες συχνά περιλαμβάνονται στατιστικές ή μαθηματικές τεχνικές. Ένας ορισμός που θα μπορούσε να δοθεί για τη χωρική ανάλυση είναι ο εξής: << **μια γενική ικανότητα να χειριστεί κανείς τη χωρική πληροφορία σε διαφορετικές μορφές με αποτέλεσμα να εξάγει επιπρόσθετη σημασία** >>. (Πηγή: Κάβουρας Μ., 2007) Κατά τη διαδικασία της χωρικής ανάλυσης ο χρήστης έχει ένα συγκεκριμένο πρόβλημα ή ερώτηση. Η βάση δεδομένων περιέχει πληροφορία (συχνά με τη μορφή χαρτών) που μπορεί να χρησιμοποιηθεί για την επίλυση του προβλήματος. Έτσι απαιτείται μια διαδικασία για τη μετατροπή των αρχικών δεδομένων της βάσης σε παράγωγα δεδομένα ή παράγωγο χάρτη, πίνακα ή σχήμα που θα παρέχει την απάντηση στην ερώτηση του χρήστη.

Οι στόχοι της χωρικής ανάλυσης είναι οι παρακάτω: (Πηγή: Κάβουρας Μ., 2007)

- Ακριβής περιγραφή των γεγονότων στο γεωγραφικό χώρο (συμπεριλαμβάνεται και η περιγραφή των χωρικών προτύπων).
- Συστηματική διερεύνηση των χωρικών προτύπων των γεγονότων και συσχέτιση των γεγονότων στο χώρο, έτσι ώστε να γίνονται πιο κατανοητές οι διαδικασίες που είναι υπεύθυνες για τη κατανομή αυτών των γεγονότων.
- Βελτίωση της ικανότητας πρόβλεψης και ελέγχου των γεγονότων που προκύπτουν στο γεωγραφικό χώρο.

Επίσης τα Γ.Σ.Π. παρέχουν δυνατότητες για μετασχηματισμό των αρχικών δεδομένων και απάντηση συγκεκριμένων ερωτήσεων που διατυπώνουν οι χρήστες. Μερικές από τις δυνατότητες μετασχηματισμού, όπως η διαγραφή και η ενημέρωση των δεδομένων, καθώς και η αλλαγή κλίμακας και προβολικού συστήματος είναι κοινές, τόσο στα Γ.Σ.Π. όσο και στα σχεδιαστικά συστήματα CAD. Περισσότερο τα Γ.Σ.Π. παρέχουν ένα ευρύ φάσμα αναλυτικών δυνατοτήτων που αφορούν την τοπολογία και τα χωρικά χαρακτηριστικά των γεωγραφικών δεδομένων, τα μη χωρικά χαρακτηριστικά ή συνδυασμούς χωρικών και μη χωρικών χαρακτηριστικών.

Παρακάτω παρουσιάζουμε συνοπτικά τις τεχνικές διαχείρισης και χωρικής ανάλυσης των δεδομένων:

- **Ανάκτηση Δεδομένων** (Data Retrieval) μέσω διατύπωσης ερωτήσεων και χρήση λογικών τιμών 0 ή 1 (Boolean manipulation) στην αποθηκευμένη πληροφορία.
- **Αφαιρετική Διαδικασία** (Map Abstraction) είναι παρόμοια με αυτή της γενίκευσης αλλά διαφέρει στον τρόπο υλοποίησης. Οι αφαιρετικές λειτουργίες είναι ο υπολογισμός κεντροειδών, οι ισαριθμικές γραμμές, τα πολύγωνα Thiessen, η αναταξινόμηση και η κανονικοποίηση.

- **Διαχείριση Φύλλων Χάρτη** (Map sheet manipulation) περιγράφει μια σειρά τεχνικών οι οποίες διαχειρίζονται τις καρτεσιανές συντεταγμένες ενός φύλλου χάρτη. Οι τεχνικές είναι η μεταβολή κλίμακας, η απαλοιφή παραμορφώσεων, η μεταβολή προβολής και ο μετασχηματισμός συντεταγμένων.
- **Δημιουργία Ζωνών Επιρροής** (Buffer Generation) έγκειται στην παραγωγή νέων πολυγώνων γύρω από σημειακές, γραμμικές και επιφανειακές οντότητες με κάποιο εύρος.
- **Επίθεση και Διαχωρισμός Πολυγώνων** (Polygon Overlay and Dissolve) περιλαμβάνει την σύνθεση ή τη διάσπαση πολλαπλών χαρτών (δύο ή περισσότερων) ώστε να παραχθεί μια νέα ομάδα δεδομένων. Οι λειτουργίες είναι η συνθετική επίθεση, η απαλοιφή ορίων και ο υπολογισμός επιφανειών.
- **Μετρήσεις** (Measurements) αφορούν σημεία, γραμμές, επιφάνειες και όγκους.
- **Ανάλυση Δικτύου** (Network Analysis) με την έννοια δίκτυο είναι ένα σύνολο αλληλοσυνδεδεμένων γραμμών που σχηματίζουν ένα σύνολο στοιχείων μέσα από τα οποία ρέουν πόροι (πρόβλημα εύρεσης βέλτιστης διαδρομής, πρόβλημα περιοδεύοντα πωλητή, χωροθετήσεις, καταγραφή πορείας).
- **Ανάλυση Πινακοποιημένων Δεδομένων** (Grid Cell Analysis) περιλαμβάνει παρόμοιες τεχνικές με εκείνες της ανάλυσης του χάρτη που πραγματοποιούνται σε συντεταγμένες x, y, αλλά χρησιμοποιούν μεγαλύτερη γενίκευση. Οι βασικές αναλυτικές λειτουργίες είναι η επίθεση πινακοποιημένων δεδομένων, η επίθεση περιοχών, η συνένωση με ακτινική αναζήτηση, ο υπολογισμός απόστασης και η επιλογή της βέλτιστης διαδρομής.
- **Ανάλυση Ψηφιακού Αναγλύφου** (Digital Terrain Analysis) περιλαμβάνει τον υπολογισμό ποικίλων προϊόντων από ένα ψηφιακό μοντέλο υψομέτρων (D.E.M.) Οι μορφές ανάλυσης των ψηφιακών μοντέλων υψομέτρων είναι η ορατότητα, η ένταση του φωτός, η παρεμβολή, η διατομή, η δημιουργία ισούψων, η κλίση, ο υπολογισμός της λεκάνης απορροής, ο προσανατολισμός και οι 3 – D όψεις.

Τα τέσσερα βασικά προϊόντα που παίρνουμε από ένα Γ.Σ.Π. είναι οι αναλογικοί χάρτες, οι πίνακες με στατιστικά δεδομένα, η εμφάνιση στην οθόνη γραφικών και μη γραφικών προϊόντων και τέλος τα αρχεία δεδομένων που προκύπτουν ύστερα από την επεξεργασία των γεωγραφικών δεδομένων.

ΚΕΦΑΛΑΙΟ 4: Ανάπτυξη Εκπαιδευτικού Υλικού

4.1 Γενικά

Στο κεφάλαιο αυτό παρουσιάζεται αναλυτικά και συνοπτικά το εκπαιδευτικό υλικό που αναπτύχθηκε για το μάθημα “ Γεωγραφία και Ανάλυση του Χώρου “ του 6^{ου} εξαμήνου που διδάσκεται στη ΣΑΤΜ.

Συγκεκριμένα αναπτύχθηκε μια άσκηση εφαρμογής δύο βασικών δειγματοληπτικών μεθόδων σε περιβάλλον γεωγραφικού συστήματος πληροφοριών, σε χωρικό υπόβαθρο που αποτελεί έναν διανυσματικό χάρτη χρήσεων γης μιας περιοχής. Συγκεκριμένα εφαρμόστηκαν οι παρακάτω μέθοδοι:

- Της απλής τυχαίας σημειακής χωρικής δειγματοληψίας,
- Της στρωματοποιημένης τυχαίας σημειακής χωρικής δειγματοληψίας.

Αντικείμενο της δειγματοληψίας είναι η εκτίμηση των ποσοστών των επιμέρους χρήσεων γης που παρουσιάζονται στους χάρτες.

Οι δύο σημειακές μέθοδοι πραγματοποιούνται σε δύο ψηφιακούς χάρτες με τη βοήθεια μιας γεννήτριας τυχαίων σημείων. Ουσιαστικά η γεννήτρια παράγει τυχαίες συντεταγμένες στα όρια της έκτασης που υποδεικνύει ο χρήστης και δημιουργεί νέο θεματικό επίπεδο (σημειακό), με στοιχεία τον αριθμό των τυχαίων σημείων που επιλέχθηκε να δημιουργηθούν.

Η αυξανόμενη ανάγκη για διαχείριση, ανάλυση, επεξεργασία και παρουσίαση των διάφορων γεωγραφικών δεδομένων, έχει κάνει τη χρήση των GIS (Geographic Information Systems) απαραίτητη στις περισσότερες γεωγραφικές μελέτες.

Τα γεωγραφικά συστήματα πληροφοριών αποτελούν ένα ιδανικό μέσο και για τη διδασκαλία θεμάτων γεωγραφίας δίνοντας τη δυνατότητα στους σπουδαστές να αναπτύξουν σημαντικές δεξιότητες στην επίλυση των προβλημάτων. Τα προηγούμενα χρόνια περιορίζονταν σε επαγγελματίες με πρόσβαση σε υψηλού επιπέδου λογισμικό και υπολογιστικό εξοπλισμό. Τα Γ.Σ.Π. σήμερα είναι πλέον προσιτά και χρησιμοποιούνται ευρέως. Η εφαρμογή επεξεργάστηκε στο λογισμικό ArcGIS 10.

Η χρήση των Γ.Σ.Π. στην εφαρμογή, εκτός του ότι συμβάλλει στην καλύτερη και πιο φιλική παρουσίαση με στόχο την κατανόηση των βασικών εννοιών της δειγματοληψίας μπορεί να θεωρηθεί ότι βοηθά:

- Στην εξοικείωση των σπουδαστών με ένα σύγχρονο λογισμικό πακέτο Γ.Σ.Π.,
- Στην ανάπτυξη ικανότητας ερμηνείας χαρτών και την επέμβαση σε αυτούς με τα διάφορα εργαλεία που προσφέρει το λογισμικό,

- Στη χρήση διάφορων χαρτογραφικών συμβολισμών και στον πειραματισμό με αυτούς, ώστε να προκύψει ένα κατά το δυνατό βέλτιστο οπτικά προϊόν και
- Τέλος στην εισαγωγή των σπουδαστών στην ανάλυση του χώρου μέσα από ρεαλιστικά παραδείγματα και στην αντιμετώπιση των διαφόρων προβλημάτων που προκύπτουν.

Τέλος, με τη χρήση των Γ.Σ.Π. για τη διενέργεια των δειγματοληψιών μειώνεται στο ελάχιστο ο χρόνος εφαρμογής (π.χ. παραγωγή τυχαίων σημείων) επιτρέποντας στους σπουδαστές να εμβαθύνουν περισσότερο στο θεωρητικό κομμάτι αυτών των διαδικασιών.

Ένας αναλογικός σχεδιασμός της εφαρμογής, δηλαδή η πραγματοποίηση δειγματοληψιών πάνω σε αναλογικό χάρτη θα έθετε σημαντικούς περιορισμούς οι σημαντικότεροι εκ των οποίων είναι οι εξής:

- Αρχικά το μέγεθος του δείγματος δε θα μπορούσε να ξεπεράσει τις 20 με 30 παρατηρήσεις (σημεία), έτσι θα ήταν αδύνατο να υπάρξει σε εύλογο χρονικό διάστημα η ακρίβεια των δειγματοληψιών που επιτυγχάνεται με τις παρακάτω διαδικασίες,
- Τα τυχαία σημεία πάνω σε έναν αναλογικό χάρτη θα τοποθετούνταν ένα - ένα με τη βοήθεια πινάκων τυχαίων αριθμών και προφανώς ο υπολογισμός των σημείων που αντιστοιχούν σε κάθε κατηγορία του πειράματος θα γίνονταν πάλι χειροκίνητα, με πιθανότητα πάντα χονδροειδών σφαλμάτων (σφαλμάτων καταγραφής κ.α.) και
- Τέλος θα έπρεπε να υπάρχουν υπολογισμένα ήδη τα πραγματικά – αναμενόμενα ποσοστά της κάθε κατηγορίας, ώστε να γίνει έλεγχος καλής προσαρμογής.

4.2 Περιγραφή Άσκησης

Η άσκηση αναφέρεται σε δυο κύριες μεθόδους χωρικής δειγματοληψίας, την Απλή Τυχαία Δειγματοληψία και τη Στρωματοποιημένη Τυχαία Δειγματοληψία. Οι σπουδαστές καλούνται να εφαρμόσουν και να εξετάσουν τις δυο παραπάνω μεθόδους, σε δυο κατάλληλα σχεδιασμένους ψηφιακούς χάρτες, που δημιουργήθηκαν για το σκοπό αυτό.

Οι δυο ψηφιακοί χάρτες αναφέρονται σε μια παραθαλάσσια περιοχή και απεικονίζουν τις παρακάτω χρήσεις γης:

- Αρδευόμενη Γεωργική Γη (Irrigating Agricultural Soil)
- Θαμνώδεις Εκτάσεις (Bushy Area)
- Βοσκότοποι (Grazing Ground)
- Δασικές Εκτάσεις (Forest)
- Δένδρο – Καλλιέργειες

Επιπλέον ο δεύτερος χάρτης της περιοχής, απεικονίζει και τρεις υψομετρικές ζώνες της. Αυτές είναι:

- η υψομετρική ζώνη των 0 – 100 μέτρων.
- η υψομετρική ζώνη των 100 – 200 μέτρων.
- η υψομετρική ζώνη των 200 – 300 μέτρων.

Ζητούμενα:

1. Να εκτιμήσετε τα ποσοστά χρήσεων γης της περιοχής, εφαρμόζοντας μια δοκιμαστική Απλή Τυχαία Δειγματοληψία, με μέγεθος δείγματος 30 σημεία. Να χρησιμοποιηθεί ο πρώτος χάρτης.
2. Να υπολογιστεί το απαιτούμενο μέγεθος δείγματος (απαιτούμενος αριθμός σημείων), ώστε τα ποσοστά χρήσεων γης να προκύψουν με ακρίβεια 2%, για επίπεδο εμπιστοσύνης 95%.
3. Με το νέο μέγεθος δείγματος που υπολογίσατε, εφαρμόστε μια νέα Απλή Τυχαία Δειγματοληψία, ώστε να προκύψουν τα νέα ποσοστά χρήσεων γης με τη ζητούμενη ακρίβεια. Να χρησιμοποιηθεί πάλι ο πρώτος χάρτης.

4. Με το ίδιο μέγεθος δείγματος, εφαρμόστε τη Στρωματοποιημένη Τυχαία Δειγματοληψία στο δεύτερο χάρτη, ώστε να προκύψουν τα ποσοστά χρήσεων γης της κάθε υψομετρικής ζώνης, αλλά και τα συνολικά ποσοστά χρήσεων γης της περιοχής. Υπενθυμίζεται ότι η Στρωματοποιημένη Τυχαία Δειγματοληψία εφαρμόζεται με την εκτέλεση μιας απλής τυχαίας δειγματοληψίας σε κάθε στρώμα, δηλαδή σε κάθε υψομετρική ζώνη. Ο αριθμός των σημείων της κάθε μίας από τρεις απλές τυχαίες δειγματοληψίες, προκύπτει αναλογικά από το εμβαδόν της ζώνης σε σχέση με το συνολικό του δειγματοληπτικού υποβάθρου.

5. Στη συνέχεια, δίνονται τα πραγματικά ποσοστά χρήσεων γης της περιοχής και ζητείται:

α. Η σύγκριση των τριών δειγματοληψιών που εφαρμόστηκαν με τη χρήση του Test χ^2 . Υπενθυμίζεται ότι το Test χ^2 συγκρίνει την κατανομή του πληθυσμού με την κατανομή του δείγματος, οπότε όσο κοντύτερα η εξαγόμενη τιμή στο 0, τόσο καλύτερα η κατανομή του δείγματος προσεγγίζει την κατανομή του πληθυσμού.

β. Ποια από τις δυο μεθόδους δειγματοληψίας που εξετάστηκαν δίνει καλύτερα αποτελέσματα; Που υπερτερεί η Στρωματοποιημένη Τυχαία Δειγματοληψία και που υστερεί σε σχέση με την Απλή Τυχαία Δειγματοληψία;

Παρακάτω παρατίθεται και ο πίνακας με τα πραγματικά ποσοστά χρήσεων γης της περιοχής που αναφέραμε παραπάνω

Αρδευόμενη Γεωργική Γη	3.6%
Θαμνώδεις Εκτάσεις	48.9%
Βοσκότοποι	5.0%
Δασικές Εκτάσεις	41.6%
Δένδρο - Καλλιέργειες	0.9%

4.3 Συνοπτικές οδηγίες για την Εκπόνηση της Άσκησης

Στον αποθηκευτικό χώρο κάθε σπουδαστή θα πρέπει να υπάρχουν οι φάκελοι της άσκησης, δηλαδή οι Map 1 και Map 2. Για την εκκίνηση του προγράμματος Arc Map: Start > Programs > ArcGis > ArcMap.

Ζητούμενο 1

Στο menu file > open > (πηγαίνετε στη θέση του φακέλου που βρίσκονται τα δεδομένα της άσκησης) > map1 (σε μορφή ArcMap document)]. Έτσι στην οθόνη εμφανίζεται πλέον ο χάρτης. Στη συνέχεια στο menu Selection > Options > επιλέξτε “ Select features that the box or graphic are completely within “ > ok.

Στον πίνακα περιεχομένων των θεματικών επιπέδων (Layers), κάντε δεξί κλικ στο Layer randompoly > Open attribute table και click στο μοναδικό πολύγωνο με FID = 0. Σκοπός η επιλογή του εξωτερικού πολυγώνου του εδάφους του χάρτη (ακτογραμμή και όρια χάρτη). Τώρα στο πολύγωνο αυτό θα έχει τονιστεί το περίγραμμα με κάποιο προεπιλεγμένο χρώμα. Θα πρέπει επίσης στο message bar (κάτω μέρος οθόνης) να γράφεται “ Number of Features selected: 1 “

Για τη διενέργεια της δοκιμαστικής τυχαίας δειγματοληψίας με μέγεθος δείγματος 30 σημεία: Αυτό γίνεται με την εντολή “ Create Random Points “ από την εργαλειοθήκη Data Management Tools.

[ArcToolBox > Data Management Tools > Feature Class > Create Random Points].

Εμφανίζεται το παράθυρο αποθήκευσης του νέου θεματικού επιπέδου που θα σχηματιστεί. Επιλέξτε την ονομασία “ simeia_1 “ και αποθηκεύστε το αρχείο στο φάκελο E:\efarmogi_3\map1. Μετά την ολοκλήρωση των παραπάνω στον πίνακα των θεματικών επιπέδων θα έχει εμφανιστεί το νέο επίπεδο “ simeia_1 “ και στο χάρτη τα 30 τυχαία σημεία της δοκιμαστικής δειγματοληψίας.

Στη συνέχεια κάντε τομή του θεματικού επιπέδου των σημείων με αυτών των χρήσεων γης (simeia_1 & theme_1) με την εντολή “ Intersect “ από την εργαλειοθήκη Analysis Tools.

[ArcToolBox > Analysis Tools > Overlay > Intersect]. Αυτό που προκύπτει από την τομή είναι ένας πίνακας τιμών (simeia_1_Intersect1) όπου κάθε σημείο αντιστοιχίζεται σε μια συγκεκριμένη κατηγορία χρήσης γης πάνω στην οποία πέφτει.

Παρακάτω πατήστε δεξί κλικ στο πεδίο LU CODE στον πίνακα simeia_1_Intersect1 επιλέξτε την εντολή Summarize όπου εμφανίζεται ένα παράθυρο διαλόγου το οποίο το συμπληρώνετε και στο τέλος του δίνετε μια ονομασία (Sum_Output_1). Ο πίνακας Sum_Output_1 δίνει το πλήθος των σημείων που πέφτουν σε κάθε χρήση γης.

Τέλος, εξάγετε τα αποτελέσματα σε λογιστικό φύλλο Excel όπου κάνετε τους αντίστοιχους υπολογισμούς για τα ποσοστά χρήσεων γης πρώτον και δεύτερον για το test χ^2 (υπολογισμένο με 2 τρόπους) χρησιμοποιώντας τη στατιστική συνάρτηση CHISQ.TEST καθώς και πίνακα όπου εισάγετε το επίπεδο σημαντικότητας – βαθμοί ελευθερίας (χρήσεις γης) όπως δίνεται από την εκφώνηση και προκύπτει η τιμή του test χ^2 αυτόματα.

Ζητούμενο 2

Ο υπολογισμός του απαιτούμενου μεγέθους δείγματος για την απλή τυχαία δειγματοληψία προκύπτει αυτόματα από το λογιστικό φύλλο Excel της δοκιμαστικής τυχαίας δειγματοληψίας με παραμέτρους το επίπεδο εμπιστοσύνης (p), το περιθώριο λάθους (e) τα οποία δίνονται από την εκφώνηση της άσκησης, το διάστημα εμπιστοσύνης z για το συγκεκριμένο επίπεδο εμπιστοσύνης και τα $p\%$, $q\%$ (όπου $p\%$ είναι η καλύτερη εκτίμηση του ποσοστού κατηγορίας χρήσεων γης, για την οποία, είναι το πλησιέστερο στο 50%, σε σχέση με τις καλύτερες εκτιμήσεις των ποσοστών των άλλων χρήσεων και το $q\% = 100 - p\%$).

Η σχέση που χρησιμοποιείται είναι η εξής:

$$n_{\text{new}} = \frac{z^2 * p\% * q\%}{e^2}$$

Ζητούμενο 3

Για τη διενέργεια της απλής τυχαίας δειγματοληψίας με μέγεθος δείγματος 2401 σημεία εφαρμόστε την ίδια διαδικασία. Αυτό γίνεται με την εντολή “ Create Random Points “ από την εργαλειοθήκη Data Management Tools.

[ArcToolBox > Data Management Tools > Feature Class > Create Random Points].

Εμφανίζεται το παράθυρο αποθήκευσης του νέου θεματικού επιπέδου που θα σχηματιστεί. Επιλέξτε την ονομασία “ simeia_2 “ και αποθηκεύστε το αρχείο στο φάκελο E:\efarmogi_3\map1. Μετά την ολοκλήρωση των παραπάνω στον πίνακα των θεματικών επιπέδων θα έχει εμφανιστεί το νέο επίπεδο “ simeia_2 “ και στο χάρτη τα 2401 τυχαία σημεία της απλής τυχαίας δειγματοληψίας.

Στη συνέχεια κάντε τομή του θεματικού επιπέδου των σημείων με αυτών των χρήσεων γης (simeia_2 & theme_1) με την εντολή “ Intersect “ από την εργαλειοθήκη Analysis Tools.

[ArcToolBox > Analysis Tools > Overlay > Intersect]. Αυτό που προκύπτει από την τομή είναι ένας πίνακας τιμών (simeia_2_Intersect2) όπου κάθε σημείο αντιστοιχίζεται σε μια συγκεκριμένη κατηγορία χρήσης γης πάνω στην οποία πέφτει.

Παρακάτω πατήστε δεξί κλικ στο πεδίο LU CODE στον πίνακα simeia_2_Intersect2 επιλέξτε την εντολή Summarize όπου εμφανίζεται ένα παράθυρο διαλόγου το οποίο το συμπληρώνετε και στο τέλος του δίνετε μια ονομασία (Sum_Output_2). Ο πίνακας Sum_Output_2 δίνει το πλήθος των σημείων που πέφτουν σε κάθε χρήση γης.

Τέλος, εξάγετε τα αποτελέσματα σε λογιστικό φύλλο Excel όπου κάνετε τους αντίστοιχους υπολογισμούς για τα ποσοστά χρήσεων γης πρώτον και δεύτερον για το test χ^2 (υπολογισμένο με 2 τρόπους) χρησιμοποιώντας τη στατιστική συνάρτηση CHISQ.TEST καθώς και πίνακα όπου εισάγετε το επίπεδο σημαντικότητας – βαθμοί ελευθερίας (χρήσεις γης) όπως δίνεται από την εκφώνηση και προκύπτει η τιμή του test χ^2 αυτόματα.

Ζητούμενο 4

Η στρωματοποιημένη τυχαία δειγματοληψία πραγματοποιείται στο δεύτερο χάρτη. Ο χάρτης αποτελείται από τρεις υψομετρικές ζώνες.

Στο menu file > open > (πηγαίνετε στη θέση του φακέλου που βρίσκονται τα δεδομένα της άσκησης) > map2 (σε μορφή ArcMap document)]. Έτσι στην οθόνη εμφανίζεται πλέον ο χάρτης. Στη συνέχεια στο menu Selection > Options > επιλέξτε “ Select features that the box or graphic are completely within “ > ok.

Στον πίνακα περιεχομένων των θεματικών επιπέδων (Layers), κάντε δεξί κλικ στο Layer contourpoly > Open attribute table και click στις υψομετρικές ζώνες με FID = 0 - 2.

Αρχικά έχοντας ως γνωστό τον αριθμό των σημείων της απλής τυχαίας δειγματοληψίας η διαδικασία που ακολουθείται για την παραγωγή του αριθμού των σημείων σε κάθε υψομετρική ζώνη είναι η εξής:

Για τον υπολογισμό του εμβαδού της κάθε ζώνης ξεχωριστά χρησιμοποιήστε την εντολή “ Calculate Areas “ από την εργαλειοθήκη Spatial Statistics Tools.

[ArcToolBox > Spatial Statistics Tools > Utilities > Calculate Areas]. Στη συνέχεια εξάγετε τα αποτελέσματα του πίνακα contourpoly_CalculateAreas9 (πεδίο F Area) σε λογιστικό φύλλο Excel όπου με την συνάρτηση Sum παίρνετε το συνολικό εμβαδόν. Έτσι από τον τύπο:

$n = n_{\text{new}} * E / \Sigma E$ προκύπτει στο φύλλο Excel αυτόματα ο αριθμός των σημείων για κάθε υψομετρική ζώνη.

Παρακάτω δώστε υψόμετρο στις χρήσεις γης. Για να γίνει αυτό χρησιμοποιήστε την εντολή “ Union “ από την εργαλειοθήκη Analysis Tools.

[ArcToolbox > Analysis Tools > Overlay > Union]. Τα επίπεδα που ενώνονται είναι τα υψομετρικά πολύγωνα (contourpoly_CalculateAreas9) και οι χρήσεις γης (theme 2). Το αποτέλεσμα που προκύπτει από την ένωση είναι το επίπεδο (union_stroma) και ο πίνακας με την ίδια ονομασία. Με προσεκτική παρατήρηση διαπιστώνεται ότι κάποια πολύγωνα έχουν αρνητικό υψόμετρο. Έτσι χρησιμοποιήστε την εντολή “ Delete Rows “ από την εργαλειοθήκη Data Management Tools.

[ArcToolbox > Data Management Tools > Table > Delete Rows] με την οποία διαγράψτε τις γραμμές με τα πολύγωνα που έχουν αρνητικό υψόμετρο.

Το επίπεδο όπου εφαρμόζεται η εντολή είναι το union_stroma. Ύστερα από όλη αυτή την διαδικασία προκύπτει το νέο επίπεδο union_stroma και ο πίνακας με τα πολύγωνα χωρίς αρνητικό υψόμετρο.

Περισσότερο έχοντας υπολογίσει τον αριθμό των σημείων που αντιστοιχούν σε κάθε υψομετρική ζώνη εφαρμόστε τρεις απλές τυχαίες δειγματοληψίες.

Για τη διενέργεια της απλής τυχαίας δειγματοληψίας στην πρώτη υψομετρική ζώνη με μέγεθος δείγματος 614 σημεία: Αυτό γίνεται με την εντολή “ Create Random Points “ από την εργαλειοθήκη Data Management Tools.

[ArcToolBox > Data Management Tools > Feature Class > Create Random Points].

Εμφανίζεται το παράθυρο αποθήκευσης του νέου θεματικού επιπέδου που θα σχηματιστεί. Επιλέξτε την ονομασία “ stroma_1 “ και αποθηκεύστε το αρχείο στο φάκελο E:\efarmogi_3\map2. Μετά την ολοκλήρωση των παραπάνω στον πίνακα των θεματικών επιπέδων θα έχει εμφανιστεί το νέο επίπεδο “ stroma_1 “ και στο χάρτη τα 614 τυχαία σημεία της απλής τυχαίας δειγματοληψίας.

Στη συνέχεια κάντε τομή του θεματικού επιπέδου των σημείων με αυτών των χρήσεων γης (stroma_1 & theme_2) με την εντολή “ Intersect “ από την εργαλειοθήκη Analysis Tools. [ArcToolBox > Analysis Tools > Overlay > Intersect]. Αυτό που προκύπτει από την τομή είναι ένας πίνακας τιμών (Intersect_stroma_1) όπου κάθε σημείο αντιστοιχίζεται σε μια συγκεκριμένη κατηγορία χρήσης γης πάνω στην οποία πέφτει.

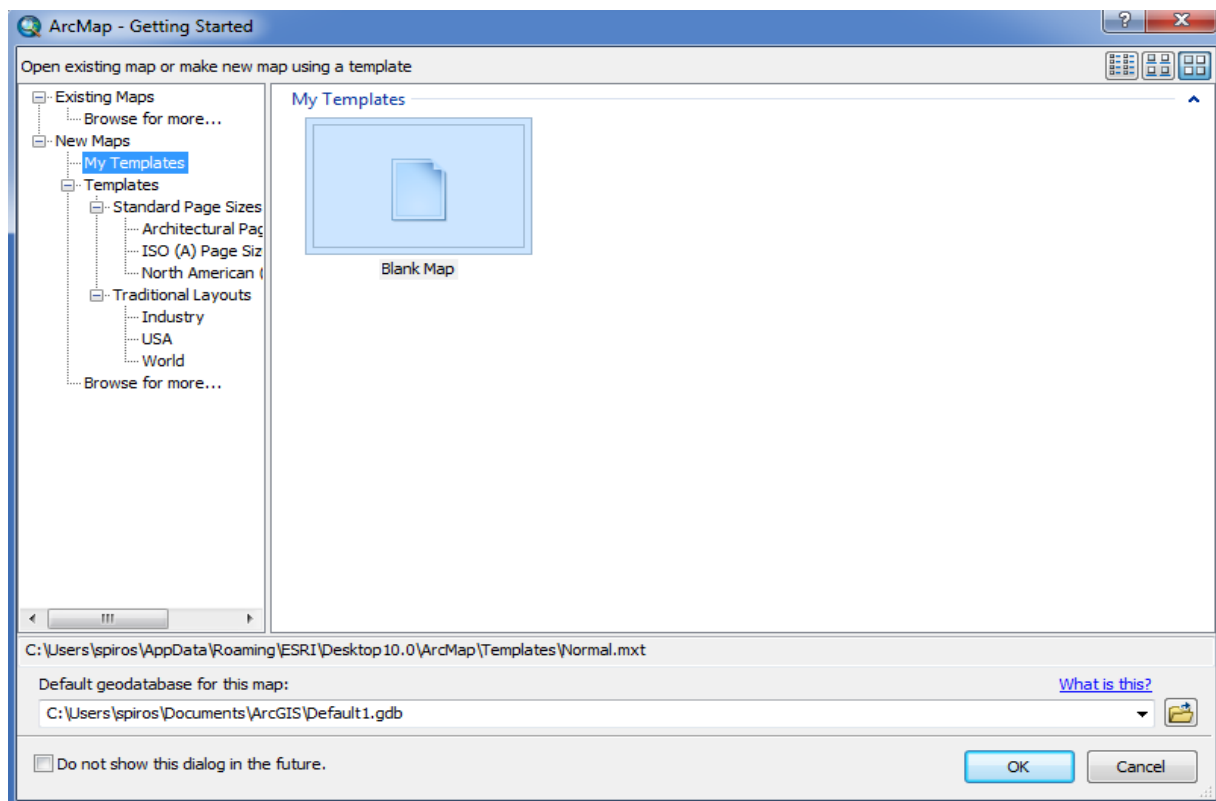
Παρακάτω πατήστε δεξί κλικ στο πεδίο LU CODE στον πίνακα Intersect_stroma_1 επιλέξτε την εντολή Summarize όπου εμφανίζεται ένα παράθυρο διαλόγου το οποίο το συμπληρώνετε και στο τέλος του δίνετε μια ονομασία (Sum_Output_1). Ο πίνακας Sum_Output_1 μας δίνει το πλήθος των σημείων που πέφτουν σε κάθε χρήση γης.

Τέλος, εξάγετε τα αποτελέσματα σε λογιστικό φύλλο Excel όπου κάνετε τους αντίστοιχους υπολογισμούς για τα ποσοστά χρήσεων γης πρώτον και δεύτερον για το test χ^2 (υπολογισμένο με 2 τρόπους) χρησιμοποιώντας τη στατιστική συνάρτηση CHISQ.TEST καθώς και πίνακα όπου εισάγετε το επίπεδο σημαντικότητας – βαθμοί ελευθερίας (χρήσεις γης) όπως δίνετε από την εκφώνηση και προκύπτει η τιμή του test χ^2 αυτόματα.

Εν κατακλείδι παρόμοια διαδικασία ακολουθήστε και στις άλλες δυο υψομετρικές ζώνες.

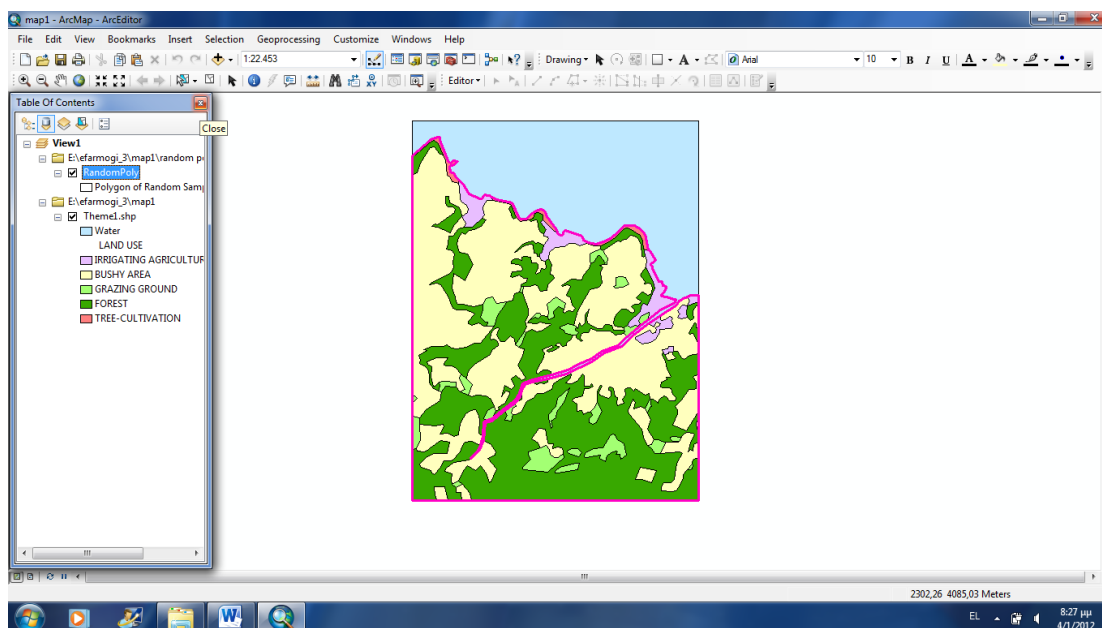
4.4 Αναλυτικά Βήματα Εφαρμογής Μεθόδων Δειγματοληψίας

Αρχικά γίνεται εκκίνηση του ArcMap [έναρξη > προγράμματα > ArcGis > ArcMap]. Στο παράθυρο που εμφανίζεται επιλέγεται εκκίνηση του ArcMap με ένα νέο χάρτη στο My Templates και στη συνέχεια κάνουμε κλικ στην επιλογή OK.



Σχήμα 4.1: Αρχικό Παράθυρο Arc Map

Για να ανοίξουν οι χάρτες της εφαρμογής [file > open > (πηγαίνοντας στη θέση του φακέλου που βρίσκονται τα δεδομένα της άσκησης) > map (σε μορφή ArcMap document)]. Έτσι στην οθόνη εμφανίζονται πλέον οι χάρτες.



Σχήμα 4.2: Περιβάλλον Arcmap – Αρχικά Δεδομένα

Ο πρώτος χάρτης αποτελείται από μια σειρά θεματικών επιπέδων (layers) που βρίσκονται στον πίνακα περιεχομένων, το επίπεδο [theme 1] που περιλαμβάνει τα πολύγωνα των 5 χρήσεων γης που θα πάρουν μέρος στη δειγματοληψία και τις υδάτινες περιοχές (1 πολύγωνο) που δεν συμμετέχουν. Το επίπεδο [RandomPoly] που είναι το υπόβαθρο για την δοκιμαστική και την απλή τυχαία δειγματοληψία (1 πολύγωνο). Αντίστοιχα ο δεύτερος χάρτης αποτελείται από τα επίπεδα [theme 2] και [contourpoly] που περιλαμβάνουν τα αντίστοιχα πλαίσια που θα χρησιμοποιηθούν για την διενέργεια της στρωματοποιημένης δειγματοληψίας.

Στον ψηφιακό χάρτη 1 [theme_1] απεικονίζονται οι παρακάτω χρήσεις γης:

- Αρδευόμενη Γεωργική Γη (Irrigating Agricultural Soil)
- Θαμνώδεις Εκτάσεις (Bushy Area)
- Βοσκότοποι (Grazing Ground)
- Δασικές Εκτάσεις (Forest)
- Δένδρο – καλλιέργειες (Tree - Cultivation)

Επιπλέον ο δεύτερος χάρτης της περιοχής απεικονίζει και τις τρεις υψομετρικές ζώνες της. Αυτές είναι:

- η υψομετρική ζώνη των 0 – 100 μέτρων.
- η υψομετρική ζώνη των 100 – 200 μέτρων.
- η υψομετρική ζώνη των 200 – 300 μέτρων.

4.5 Εφαρμογή Δοκιμαστικής Δειγματοληψίας

Αρχικό στάδιο της εφαρμογής των δειγματοληψιών είναι ο καθορισμός του μεγέθους του δείγματος. Για τον καθορισμό αυτόν συνήθως επιλέγεται ένα δοκιμαστικό δείγμα που καλό είναι να περιλαμβάνει το λιγότερο 30 παρατηρήσεις.

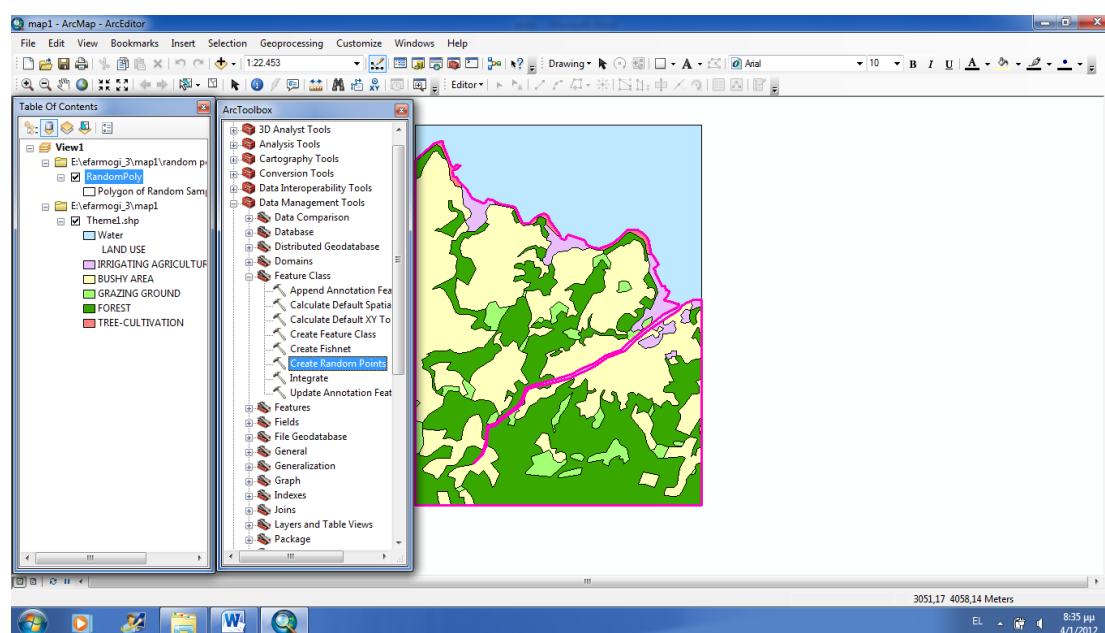
Το δειγματοληπτικό υπόβαθρο για την δοκιμαστική τυχαία δειγματοληψία είναι όλος ο χάρτης πλην των υδάτινων επιφανειών. Στον πίνακα περιεχομένων εμφανίζεται ως RandomPoly.

Το υπόβαθρο δημιουργήθηκε με διαδικασία επιλογής από το επίπεδο [theme 1] όλων των πολυγώνων πλην αυτού που απεικονίζει τις υδάτινες περιοχές. Επίσης έγινε απλοποίηση των γραμμών ώστε να παρουσιάζεται ως ένα ενιαίο πολύγωνο.

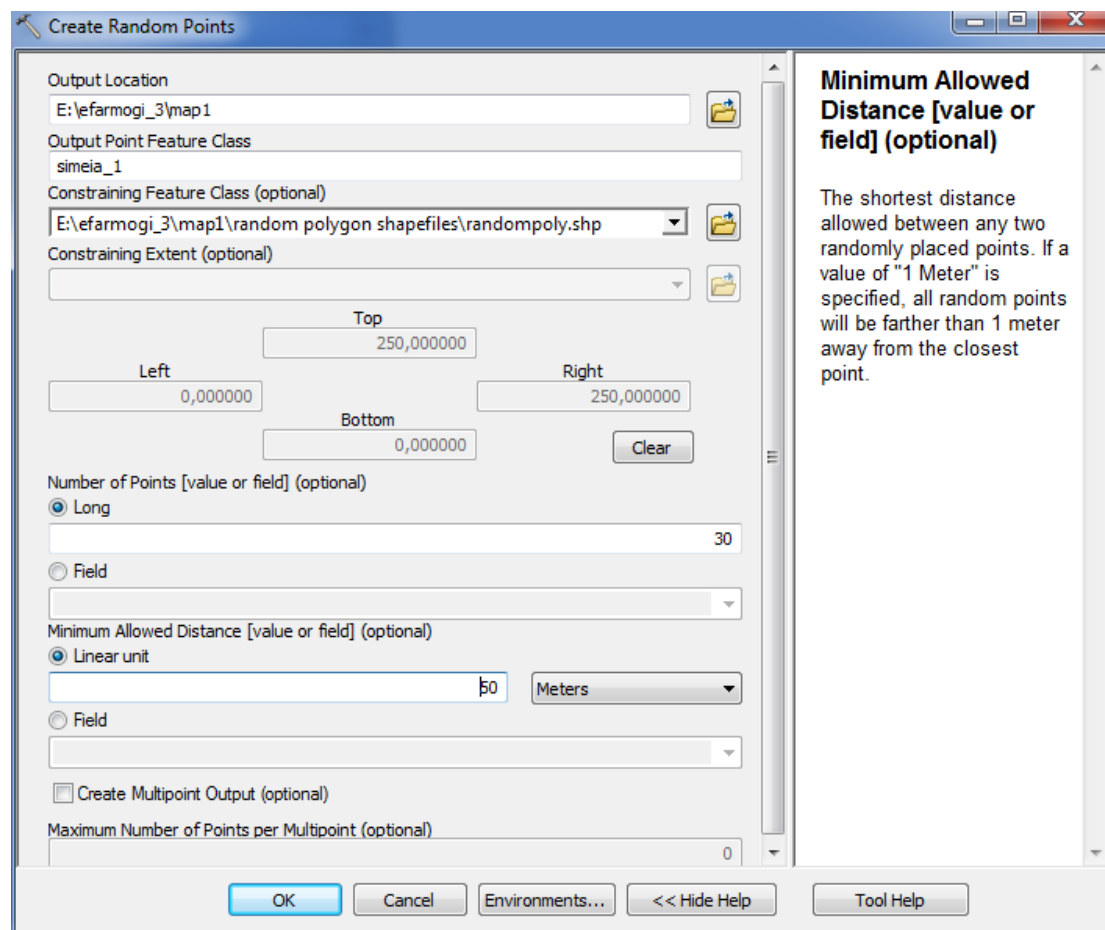
Η εφαρμογή ξεκινάει με τη λήψη ενός δοκιμαστικού δείγματος 30 παρατηρήσεων. Αυτό γίνεται με την εντολή “ Create Random Points “ από την εργαλειοθήκη Data Management Tools.

[ArcToolBox > Data Management Tools > Feature Class > Create Random Points].

Το παράθυρο διαλόγου εμφανίζεται και συμπληρώνεται όπως παρακάτω:



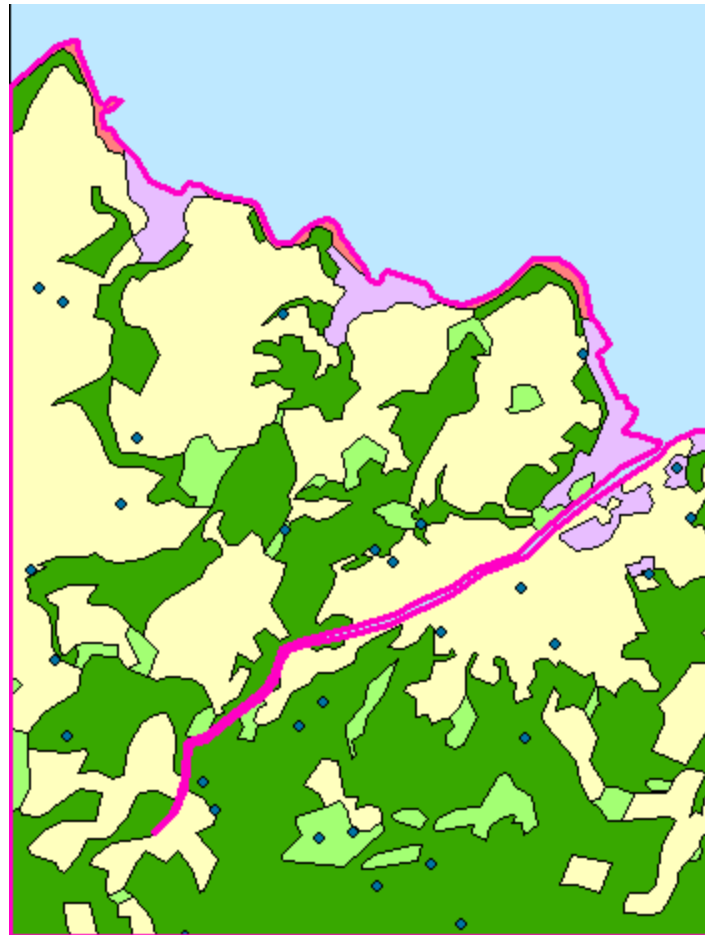
Σχήμα 4.3: Παράθυρο ArcToolbox Create Random Points Δοκιμαστικής Τυχαίας Δειγματοληψίας



Σχήμα 4.4: Παράμετροι Παραθύρου Create Random Points Δοκιμαστικής Τυχαίας Δειγματοληψίας

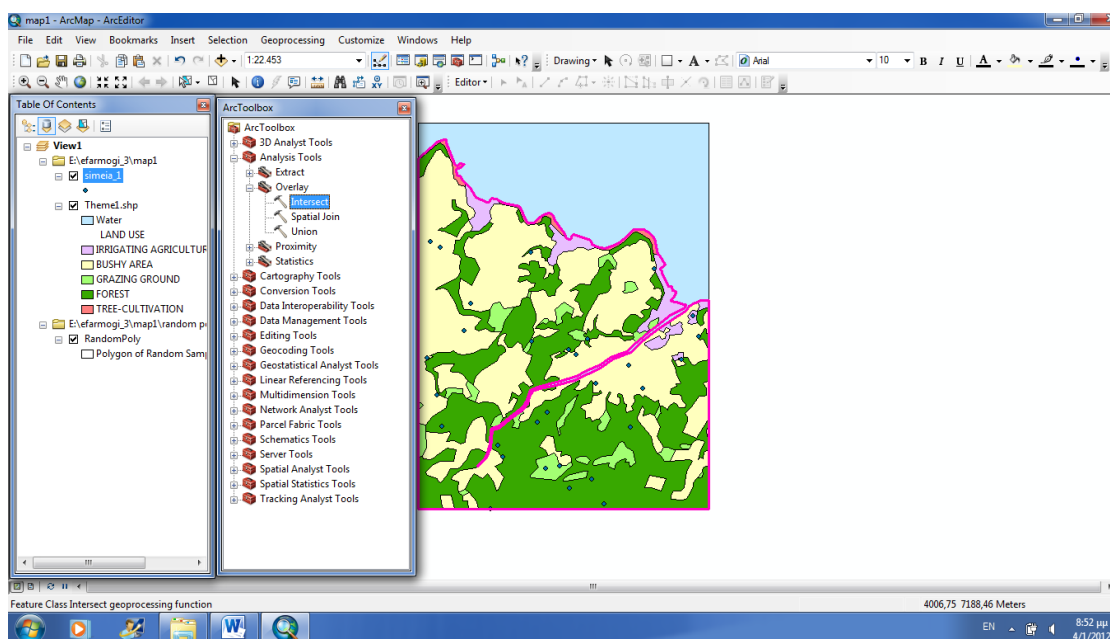
Στο πρώτο πεδίο επιλέγουμε την τοποθεσία στην οποία τα σημεία θα δημιουργηθούν. Στο δεύτερο πεδίο επιλέγουμε το όνομα που θα προκύψει από την γέννηση των σημείων (simeia_1) και στο τρίτο πεδίο επιλέγουμε περιορισμό κατηγορίας χαρακτηριστικών γνωρισμάτων στο συγκεκριμένο πολύγωνο.

Προχωρώντας στο πεδίο number of points εισάγουμε τον αριθμό των σημείων που θέλουμε να δημιουργήσουμε (30) και στο πεδίο minimum allowed distance ορίζουμε την ελάχιστη απόσταση μεταξύ δύο σημείων (50 meters). Για παράδειγμα αν βάλουμε την τιμή 1 όλα τα σημεία θα πρέπει να απέχουν περισσότερο από 1 meter από το πλησιέστερο σημείο.

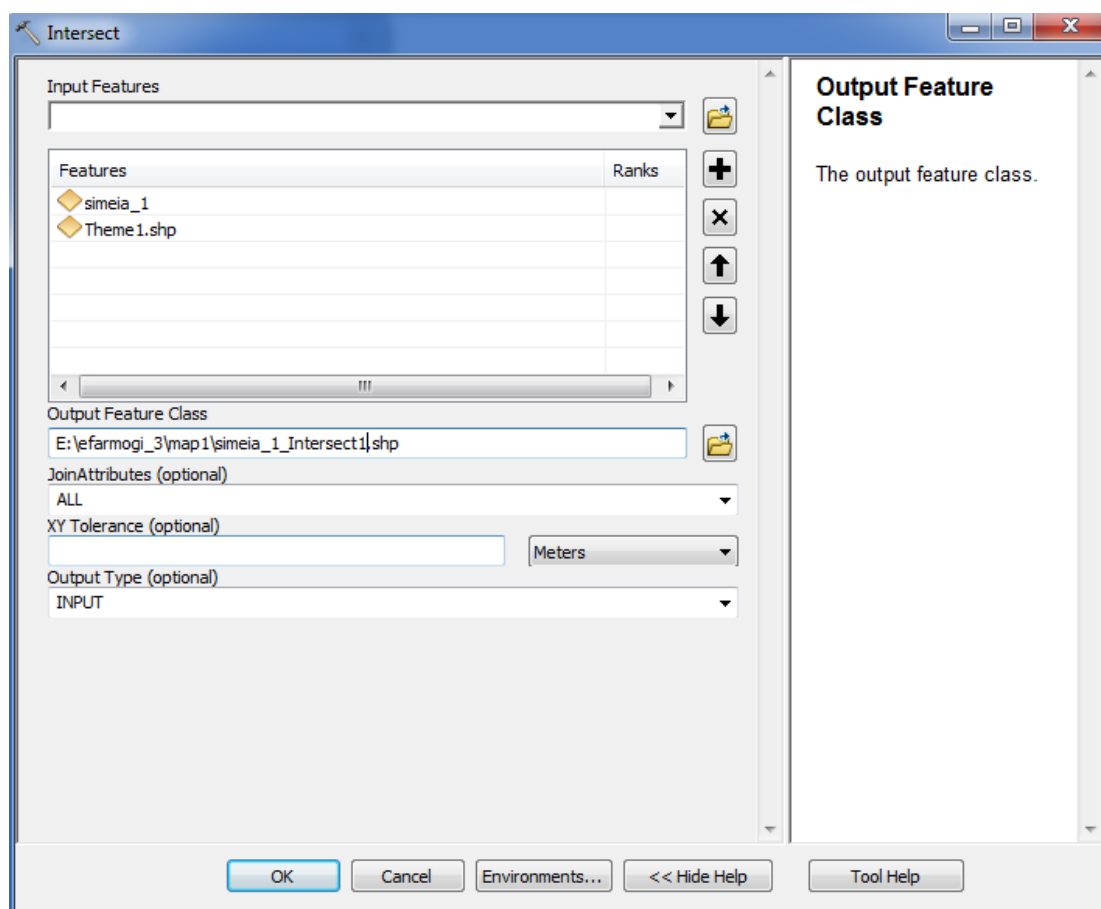


Σχήμα 4.5: Δοκιμαστικό δείγμα 30 παρατηρήσεων simeia_1

Έπειτα με την εντολή Intersect κατατάσσουμε τα σημεία στην χρήση γης στην οποία ανήκουν. Η διαδικασία που ακολουθούμε είναι [ArcToolBox > Analysis Tools > Overlay > Intersect]. Το παράθυρο διαλόγου εμφανίζεται και συμπληρώνεται όπως παρακάτω:



Σχήμα 4.6: Παράθυρο ArcToolbox Intersect Δοκιμαστικής Τυχαίας Δειγματοληψίας



Σχήμα 4.7: Παράμετροι Παραθύρου Intersect Δοκιμαστικής Τυχαίας Δειγματοληψίας

Στο πεδίο Input Features επιλέγουμε τα επίπεδα στα οποία θέλουμε να εφαρμόσουμε την τομή (simeia_1 & theme 1) και στο πεδίο Output Feature Class επιλέγουμε το όνομα και την τοποθεσία που θα προκύψει από την τομή (simeia_1_Intersect1.shp).

Ο χάρτης απεικονίζεται ως εξής:



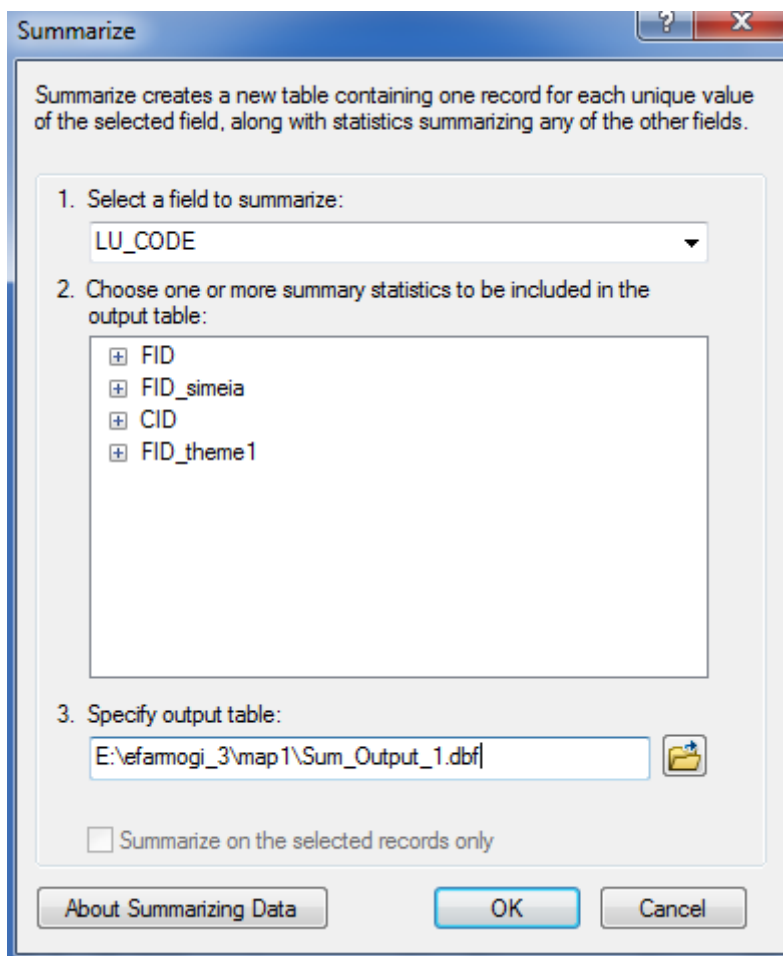
Σχήμα 4.8: Δοκιμαστικό Δείγμα 30 παρατηρήσεων simeia _1_Intersect1

Περισσότερο αυτό που προκύπτει από την τομή είναι ένας πίνακας τιμών όπου κάθε σημείο αντιστοιχίζεται σε μια συγκεκριμένη κατηγορία χρήσης γης πάνω στην οποία πέφτει.

FID	Shape *	FID_simeia	CID	FID_theme1	LU CODE
0	Point	4	0	45	400
1	Point	29	0	45	400
2	Point	3	0	45	400
3	Point	1	0	45	400
4	Point	13	0	68	300
5	Point	11	0	68	300
6	Point	20	0	45	400
7	Point	12	0	45	400
8	Point	22	0	50	400
9	Point	19	0	45	400
10	Point	6	0	45	400
11	Point	21	0	45	400
12	Point	17	0	3	200
13	Point	8	0	26	200
14	Point	0	0	26	200
15	Point	16	0	26	200
16	Point	9	0	3	200
17	Point	18	0	42	200
18	Point	14	0	37	200
19	Point	5	0	37	200
20	Point	23	0	39	300
21	Point	26	0	31	400
22	Point	25	0	26	200
23	Point	2	0	3	200
24	Point	27	0	30	100
25	Point	28	0	7	200
26	Point	24	0	14	200
27	Point	10	0	7	200
28	Point	15	0	3	200
29	Point	7	0	3	200

Σχήμα 4.9: Πίνακας Κατηγοριοποίησης Σημείων ανά Χρήση Γης simeia_1_Intersect1

Στη συνέχεια κάνοντας δεξί κλικ στο πεδίο LU CODE στον πίνακα simeia_1_Intersect1 επιλέγουμε την εντολή Summarize όπου μας εμφανίζεται ένα παράθυρο διαλόγου το οποίο το συμπληρώνουμε ως εξής:



Σχήμα 4.10: Παράθυρο Διαλόγου Summarize Δοκιμαστικής Τυχαίας Δειγματοληψίας

Στο πρώτο πεδίο επιλέγουμε το πεδίο ως προς το οποίο θέλουμε να γίνει η κατηγοριοποίηση (LU CODE), στο δεύτερο πεδίο δεν επιλέγουμε αθροιστικά στατιστικά διότι το sum αθροίζει αύξοντες αριθμούς (simeia_1) στον πίνακα και στο τρίτο δίνουμε στον πίνακα ένα όνομα και μια κατάληξη (dbf αρχείο).

Αξίζει να σημειωθεί ότι στον πίνακα (Sum_Output_1) μας δίνεται το πλήθος των σημείων που πέφτουν σε κάθε χρήση γης που είναι εκείνο που μας ενδιαφέρει περισσότερο και θα χρησιμοποιήσουμε αργότερα στο φύλλο Excel για τον υπολογισμό των ποσοστών των επιμέρους χρήσεων γης ανά κατηγορία.

OID	LU CODE	Count	LU CODE
0	100	1	
1	200	15	
2	300	3	
3	400	11	

Σχήμα 4.11: Πίνακας Σημείων ανά Χρήση Γης Δοκιμαστικής Τυχαίας Δειγματοληψίας Sum_Output_1

Τέλος, εξάγουμε τα αποτελέσματα του πίνακα Sum_Output_1 σε λογιστικό φύλλο Excel όπου κάνουμε τους αντίστοιχους υπολογισμούς για τα ποσοστά χρήσεων γης και το test χ^2 αυτόματα.

Δοκιμαστική Τυχαία Σημειακή Δειγματοληψία			
Χρήσεις Γης	Αριθμός Σημείων	Ποσοστά Χρήσεων Γης (%)	
100	1	3,33333	
200	15	50,00000	
300	3	10,00000	
400	11	36,66667	
700	0	0,00000	
Σύνολο	30	100,00000	
e=	2,00		
z	1,96		
p%=	50,00		
q%=	50,00		
nnew=	2401		
	Αναμενόμενη Τιμή (%)	Πραγματική Τιμή (%)	Τέστ χ^2
p100	3,3	3,6	0,019753
p200	50,0	48,9	0,024744
p300	10,0	5	5
p400	36,7	41,6	0,585043
p700	0,0	0,9	0,9
		χ^2 Rtest	6,52954

Σχήμα 4.12: Φύλλο Excel για τον αυτόματο υπολογισμό ποσοστών χρήσεων γης ανά κατηγορία στη δοκιμαστική δειγματοληψία, νέου απαιτούμενου μεγέθους δείγματος στην απλή δειγματοληψία και του test χ^2

Παράλληλα η τιμή του τεστ χ^2_d μπορεί να προκύψει αυτόματα για την παραπάνω δειγματοληψία. Συγκρίνοντας την τιμή του τεστ χ^2 για δεδομένο επίπεδο σημαντικότητας ($\alpha = 0.05$) και βαθμούς ελευθερίας $\nu = k - 1$ στη περίπτωση της δοκιμαστικής τυχαίας δειγματοληψίας έχουμε $\nu = 4$. Επομένως το $\chi^2_d = 9.48773$.

Όπου: k = κατηγορία χρήσης γης (5)

Συμπεραίνουμε ότι στα δεδομένα του δείγματος προσαρμόζεται ικανοποιητικά η κατανομή αφού $\chi^2 < \chi^2_d$.

4.6 Εύρεση Απαιτούμενου Μεγέθους Δείγματος

Για την εύρεση του απαιτούμενου μεγέθους δείγματος ή πιο σωστά των επιπλέον σημείων που απαιτούνται διατίθεται λογιστικό φύλλο Excel στο οποίο μεταφέρονται τα αποτελέσματα της δοκιμαστικής τυχαίας δειγματοληψίας (στήλη Count LU CODE – σχήμα 4.11).

Αντιγράφοντας τα σημεία ανά χρήση γης όπως βρέθηκαν στην δοκιμαστική τυχαία δειγματοληψία και βάζοντας το επίπεδο εμπιστοσύνης και περιθώριο λάθους που μας δίνεται από την εκφώνηση της άσκησης υπολογίζεται αυτόματα το απαιτούμενο μέγεθος δείγματος της απλής τυχαίας σημειακής δειγματοληψίας (n_{new}) το οποίο θα το χρησιμοποιήσουμε. Ο υπολογισμός του απαιτούμενου μεγέθους δείγματος γίνεται μετά την πραγματοποίηση της δοκιμαστικής δειγματοληψίας. Ο τύπος υπολογισμού είναι ο εξής:

$$n_{new} = \frac{z^2 * p\% * q\%}{e^2}$$

Όπου: n_{new} : το απαιτούμενο μέγεθος δείγματος για επίπεδο εμπιστοσύνης α , Z και περιθώριο λάθους e (π.χ. $\alpha = 95\%$, $Z = 1.96$ και $e = 2\%$)

$p\%$: η καλύτερη εκτίμηση του ποσοστού της κατηγορίας χρήσεων γης για την οποία είναι το πλησιέστερο 50% σε σχέση με τις καλύτερες εκτιμήσεις των ποσοστών των άλλων χρήσεων

$$q\% = 100 - p\%$$

Στη συγκεκριμένη εφαρμογή για επίπεδο εμπιστοσύνης $\alpha = 95\%$ έχουμε $Z = 1.96$ και περιθώριο σφάλματος ($e = 2\%$).

Με αυτές τις επιλογές λαμβάνουμε το απαιτούμενο μέγεθος δείγματος για την απλή τυχαία δειγματοληψία.

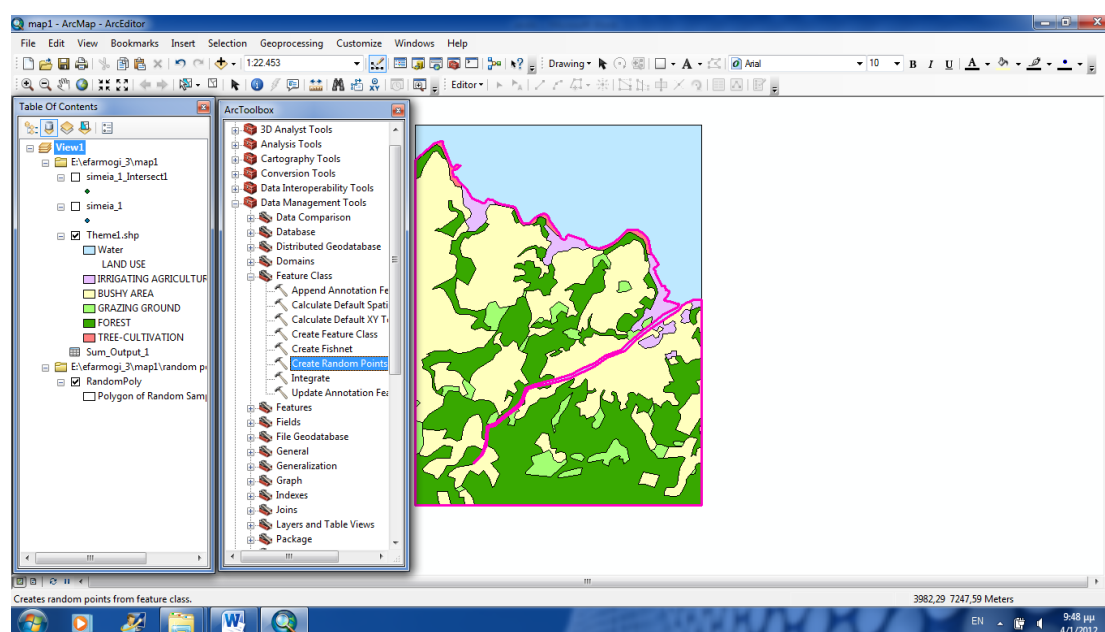
4.7 Εφαρμογή Απλής Τυχαίας Δειγματοληψίας

Ύστερα από την εφαρμογή της δοκιμαστικής τυχαίας δειγματοληψίας και την εύρεση του απαιτούμενου μεγέθους δείγματος που ικανοποιεί τις απαιτήσεις μας, μπορεί να πραγματοποιηθεί η απλή τυχαία δειγματοληψία. Η διαδικασία που ακολουθείται είναι παρόμοια με την προηγούμενη.

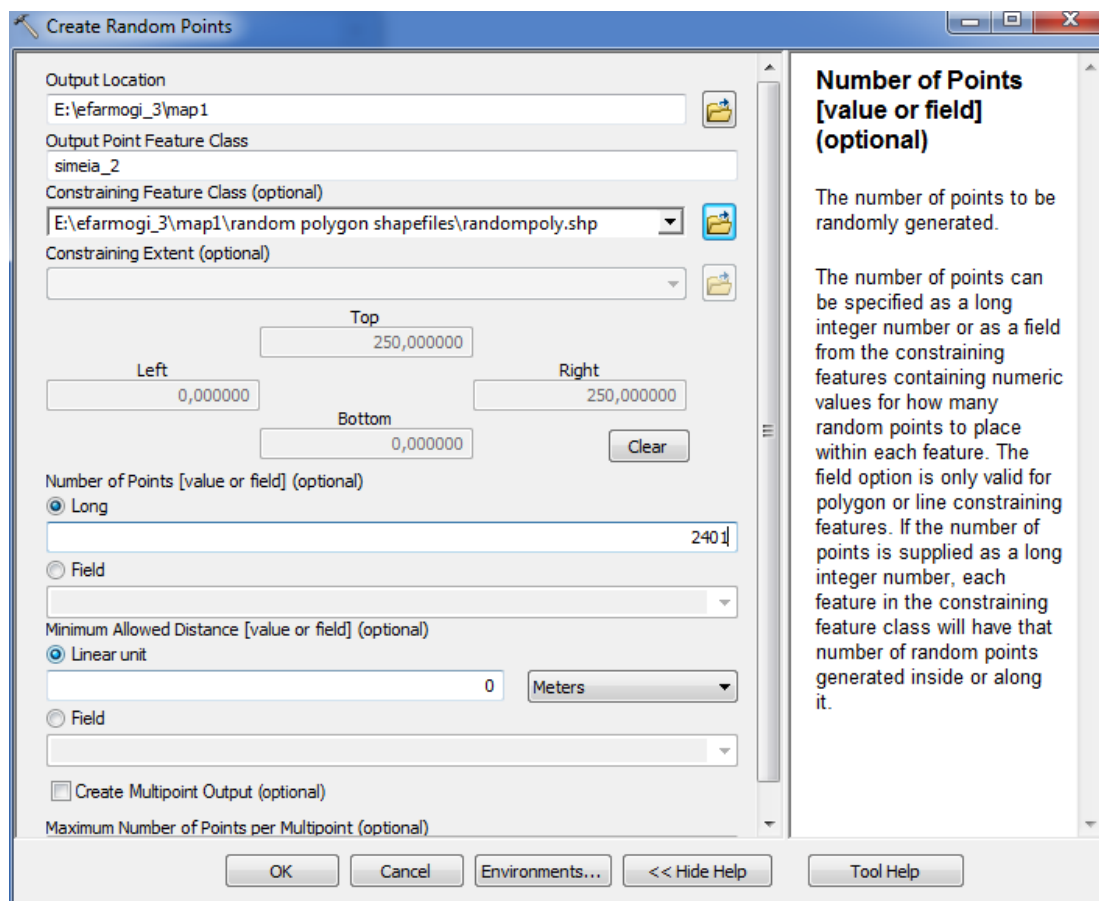
Η εφαρμογή ξεκινάει με τη λήψη ενός δοκιμαστικού δείγματος 2401 παρατηρήσεων. Αυτό γίνεται με την εντολή “ Create Random Points “ από την εργαλειοθήκη Data Management Tools.

[ArcToolBox > Data Management Tools > Feature Class > Create Random Points].

Το παράθυρο διαλόγου εμφανίζεται και συμπληρώνεται όπως παρακάτω:



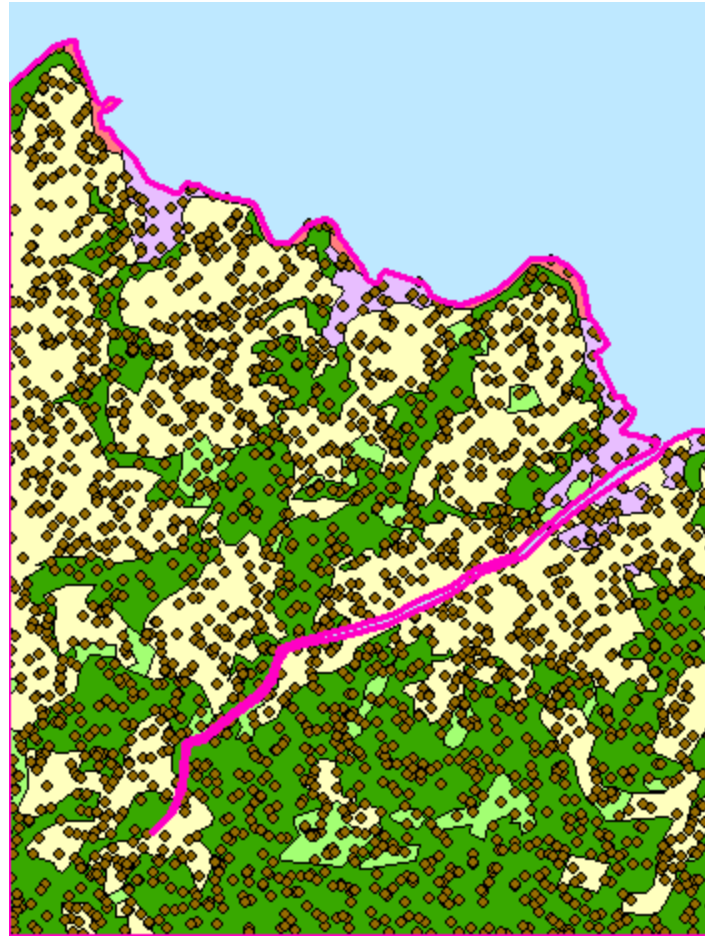
Σχήμα 4.13: Παράθυρο ArcToolbox Create Random Points Απλής Τυχαίας Δειγματοληψίας



Σχήμα 4.14: Παράμετροι Παραθύρου Create Random Points Απλής Τυχαίας Δειγματοληψίας

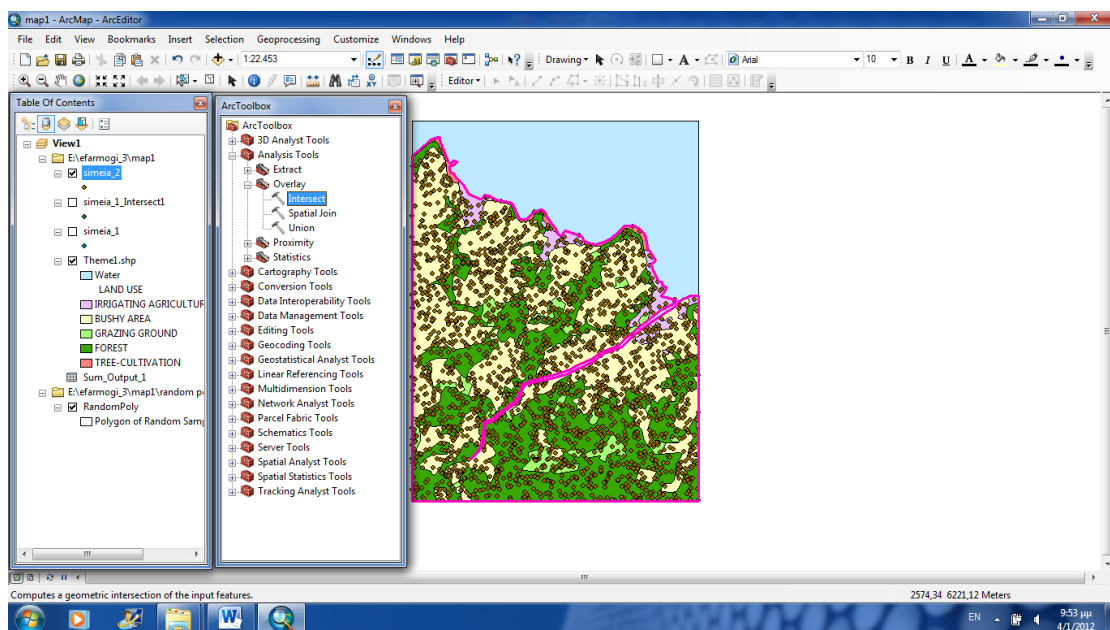
Στο πρώτο πεδίο επιλέγουμε την τοποθεσία στην οποία τα σημεία θα δημιουργηθούν. Στο δεύτερο πεδίο επιλέγουμε το όνομα που θα προκύψει από την γέννηση των σημείων (simeia_2) και στο τρίτο πεδίο επιλέγουμε περιορισμό κατηγορίας χαρακτηριστικών γνωρισμάτων στο συγκεκριμένο πολύγωνο.

Προχωρώντας στο πεδίο number of points εισάγουμε τον αριθμό των σημείων που θέλουμε να δημιουργήσουμε (2401) και στο πεδίο minimum allowed distance ορίζουμε την ελάχιστη απόσταση μεταξύ δύο σημείων (0 meters).

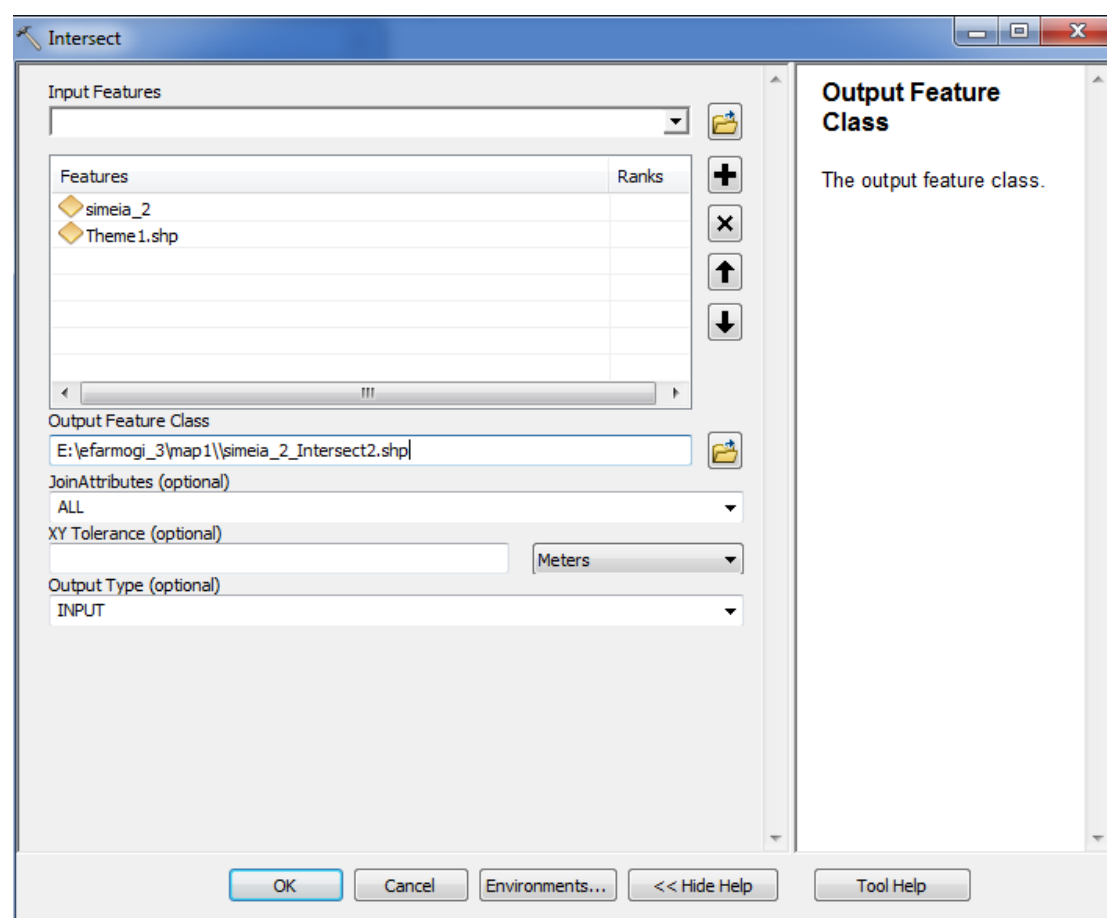


Σχήμα 4.15: Δείγμα 2401 παρατηρήσεων simeia_2

Έπειτα με την εντολή Intersect κατατάσσουμε τα σημεία στην χρήση γης στην οποία ανήκουν. Η διαδικασία που ακολουθούμε είναι [ArcToolBox > Analysis Tools > Overlay > Intersect]. Το παράθυρο διαλόγου εμφανίζεται και συμπληρώνεται όπως παρακάτω:



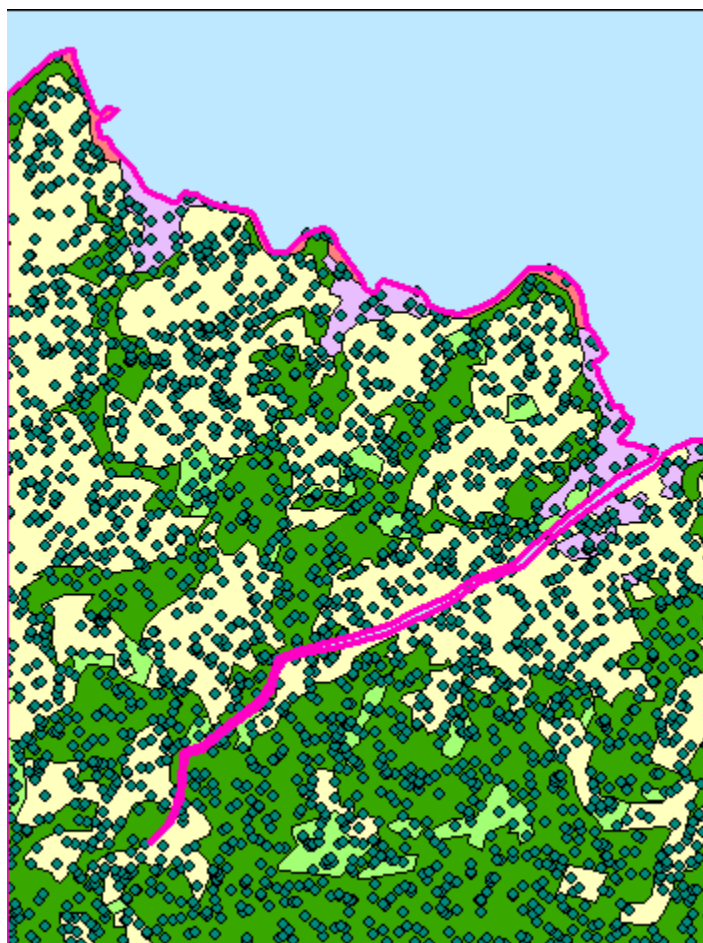
Σχήμα 4.16: Παράθυρο ArcToolbox Intersect Απλής Τυχαίας Δειγματοληψίας



Σχήμα 4.17: Παράμετροι Παραθύρου Intersect Απλής Τυχαίας Δειγματοληψίας

Στο πεδίο Input Features επιλέγουμε τα επίπεδα στα οποία θέλουμε να εφαρμόσουμε την τομή (simeia_2 & theme 1) και στο πεδίο Output Feature Class επιλέγουμε το όνομα και την τοποθεσία που θα προκύψει από την τομή (simeia _2_ Intersect2.shp).

Ο χάρτης απεικονίζεται ως εξής:



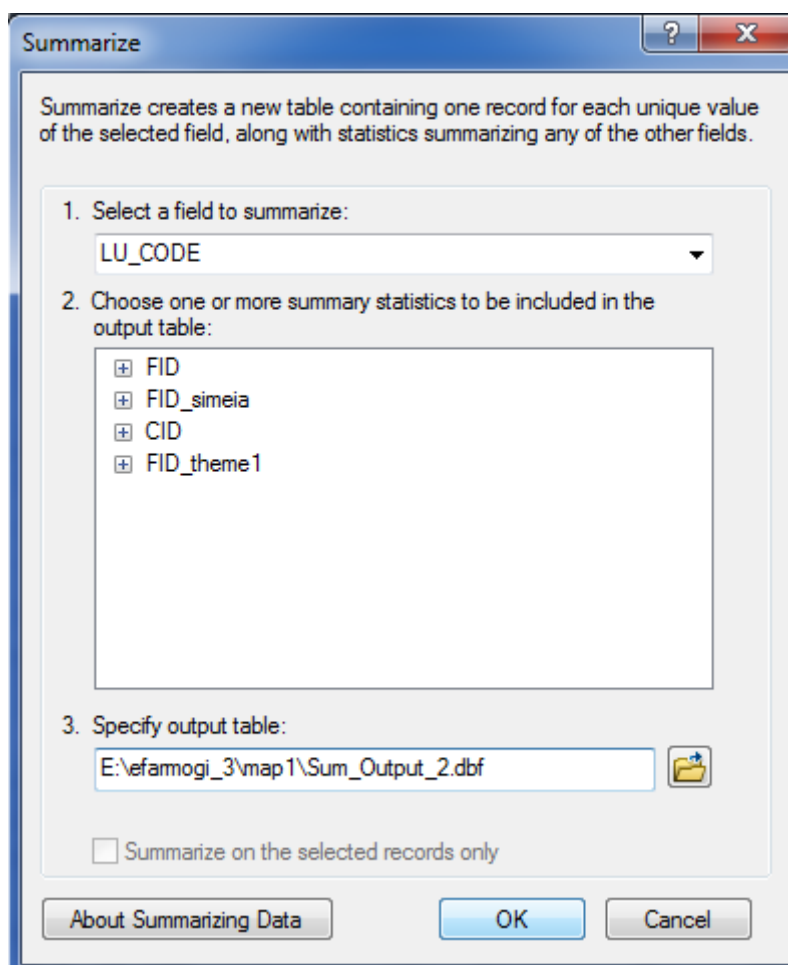
Σχήμα 4.18: Δείγμα 2401 παρατηρήσεων simeia _2_Intersect2

Περισσότερο αυτό που προκύπτει από την τομή είναι ένας πίνακας τιμών όπου κάθε σημείο αντιστοιχίζεται σε μια συγκεκριμένη κατηγορία χρήσης γης πάνω στην οποία πέφτει.

FID	Shape *	FID_simeia	CID	FID_theme1	LU CODE
0	Point	326	0	75	200
1	Point	8	0	75	200
2	Point	315	0	67	400
3	Point	1741	0	67	400
4	Point	1476	0	45	400
5	Point	1644	0	45	400
6	Point	1392	0	45	400
7	Point	2099	0	45	400
8	Point	302	0	67	400
9	Point	240	0	45	400
10	Point	1240	0	45	400
11	Point	762	0	45	400
12	Point	324	0	45	400
13	Point	235	0	45	400
14	Point	116	0	45	400
15	Point	635	0	45	400
16	Point	344	0	67	400
17	Point	75	0	45	400
18	Point	386	0	45	400
19	Point	2195	0	45	400
20	Point	431	0	45	400
21	Point	323	0	45	400
22	Point	100	0	67	400
23	Point	1900	0	45	400
24	Point	1368	0	45	400
25	Point	2172	0	45	400
26	Point	1352	0	45	400
27	Point	411	0	45	400
28	Point	1287	0	45	400
29	Point	772	0	45	400
30	Point	736	0	74	200

Σχήμα 4.19: Πίνακας Κατηγοριοποίησης Σημείων ανά Χρήση Γης simeia_2_Intersect2

Στη συνέχεια κάνοντας δεξί κλικ στο πεδίο LU CODE στον πίνακα simeia_2_Intersect2 επιλέγουμε την εντολή Summarize όπου μας εμφανίζεται ένα παράθυρο διαλόγου το οποίο το συμπληρώνουμε ως εξής:



Σχήμα 4.20: Παράθυρο Διαλόγου Summarize Απλής Τυχαίας Δειγματοληψίας

Στο πρώτο πεδίο επιλέγουμε το πεδίο ως προς το οποίο θέλουμε να γίνει η κατηγοριοποίηση (LU CODE), στο δεύτερο πεδίο δεν επιλέγουμε αθροιστικά στατιστικά διότι το sum αθροίζει αύξοντες αριθμούς (simeia_2) στον πίνακα και στο τρίτο δίνουμε στον πίνακα ένα όνομα και μια κατάληξη (dbf αρχείο).

Αξίζει να σημειωθεί ότι στον πίνακα (Sum_Output_2) μας δίνεται το πλήθος των σημείων που πέφτουν σε κάθε χρήση γης που είναι εκείνο που μας ενδιαφέρει περισσότερο και θα χρησιμοποιήσουμε αργότερα στο φύλλο Excel για τον υπολογισμό των ποσοστών των επιμέρους χρήσεων γης ανά κατηγορία.

OID	LU CODE	Count	LU CODE
0	100	78	
1	200	1181	
2	300	132	
3	400	997	
4	700	13	

Σχήμα 4.21: Πίνακας Σημείων ανά Χρήση Γης Απλής Τυχαίας Δειγματοληψίας Sum_Output_2

Τέλος, εξάγουμε τα αποτελέσματα του πίνακα Sum_Output_2 σε λογιστικό φύλλο Excel όπου κάνουμε τους αντίστοιχους υπολογισμούς για τα ποσοστά χρήσεων γης και το test χ^2 αυτόματα.

Απλή Τυχαία Σημειακή Δειγματοληψία			
Χρήσεις Γης	Αριθμός Σημείων	Ποσοστά Χρήσεων Γης (%)	
100	78	3,24865	
200	1181	49,18784	
300	132	5,49771	
400	997	41,52436	
700	13	0,54144	
Σύνολο	2401	100,00000	
e=	2,00		
z	1,96		
p%=	49,19		
q%=	50,81		
nnew=	2400		
	Αναμενόμενη Τιμή (%)	Πραγματική Τιμή (%)	Τέστ χ^2
p100	3,24865	3,6	0,034291
p200	49,18784	48,9	0,001694
p300	5,49771	5	0,049543
p400	41,52436	41,6	0,000138
p700	0,54144	0,9	0,142849
		χ^2 Rtest	0,228516

Σχήμα 4.22: Φύλλο Excel για τον αυτόματο υπολογισμό ποσοστών χρήσεων γης ανά κατηγορία στην απλή τυχαία δειγματοληψία και του test χ^2

Παράλληλα η τιμή του τεστ χ^2_d μπορεί να προκύψει αυτόματα για την παραπάνω δειγματοληψία. Συγκρίνοντας την τιμή του τεστ χ^2 για δεδομένο επίπεδο σημαντικότητας ($\alpha = 0.05$) και βαθμούς ελευθερίας $\nu = k - 1$ στη περίπτωση της απλής τυχαίας δειγματοληψίας έχουμε $\nu = 4$. Επομένως το $\chi^2_d = 9.48773$.

Όπου: k = κατηγορία χρήσης γης (5)

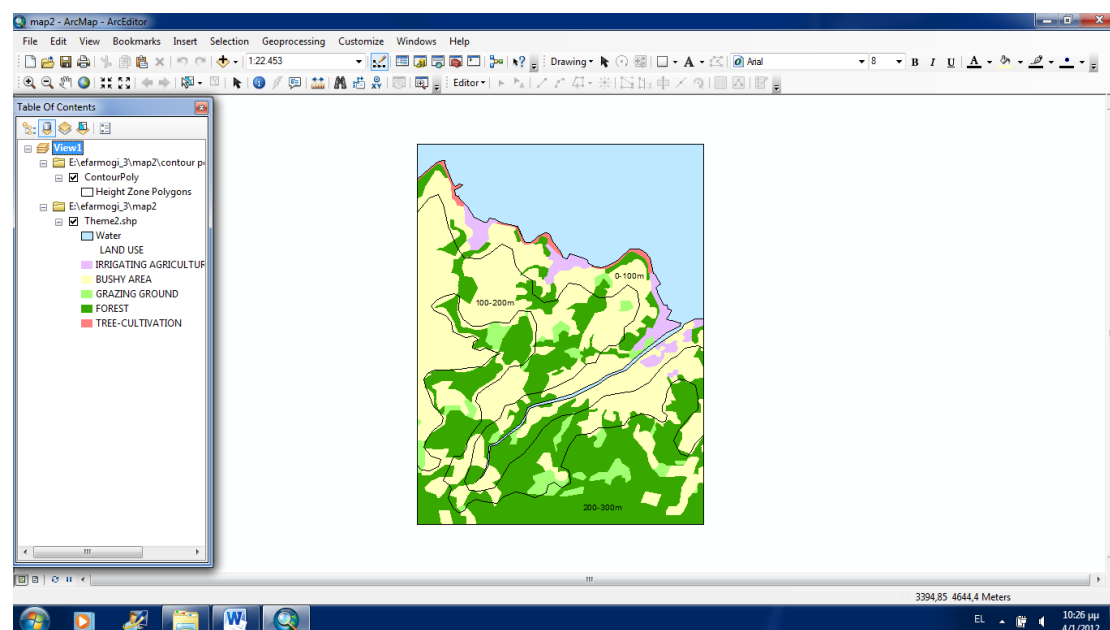
Συμπεραίνουμε ότι στα δεδομένα του δείγματος προσαρμόζεται ικανοποιητικά η κατανομή αφού $\chi^2 < \chi^2_d$.

4.8 Εφαρμογή Στρωματοποιημένης Τυχαίας Δειγματοληψίας

Στη μέθοδο αυτή η δειγματοληψία επιτυγχάνεται με τη στρωματοποίηση του υποβάθρου σε υποπεριοχές και στη συνέχεια σε κάθε μία από αυτές γίνεται απλή τυχαία δειγματοληψία με το γνωστό τρόπο.

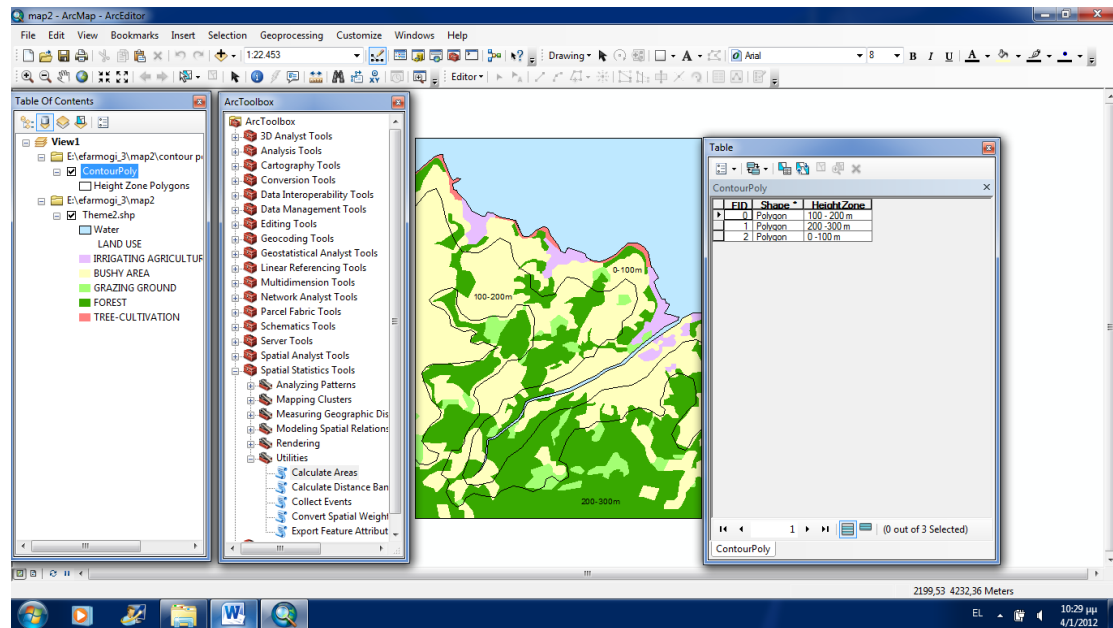
Το υπόβαθρο στρωματοποιήθηκε σε 3 υποπεριοχές που αντιστοιχούν στις τρεις υψομετρικές ζώνες του χάρτη.

- Υψομετρική ζώνη 1 (0 – 100 m)
- Υψομετρική ζώνη 2 (100 – 200 m)
- Υψομετρική ζώνη 3 (200 – 300 m)

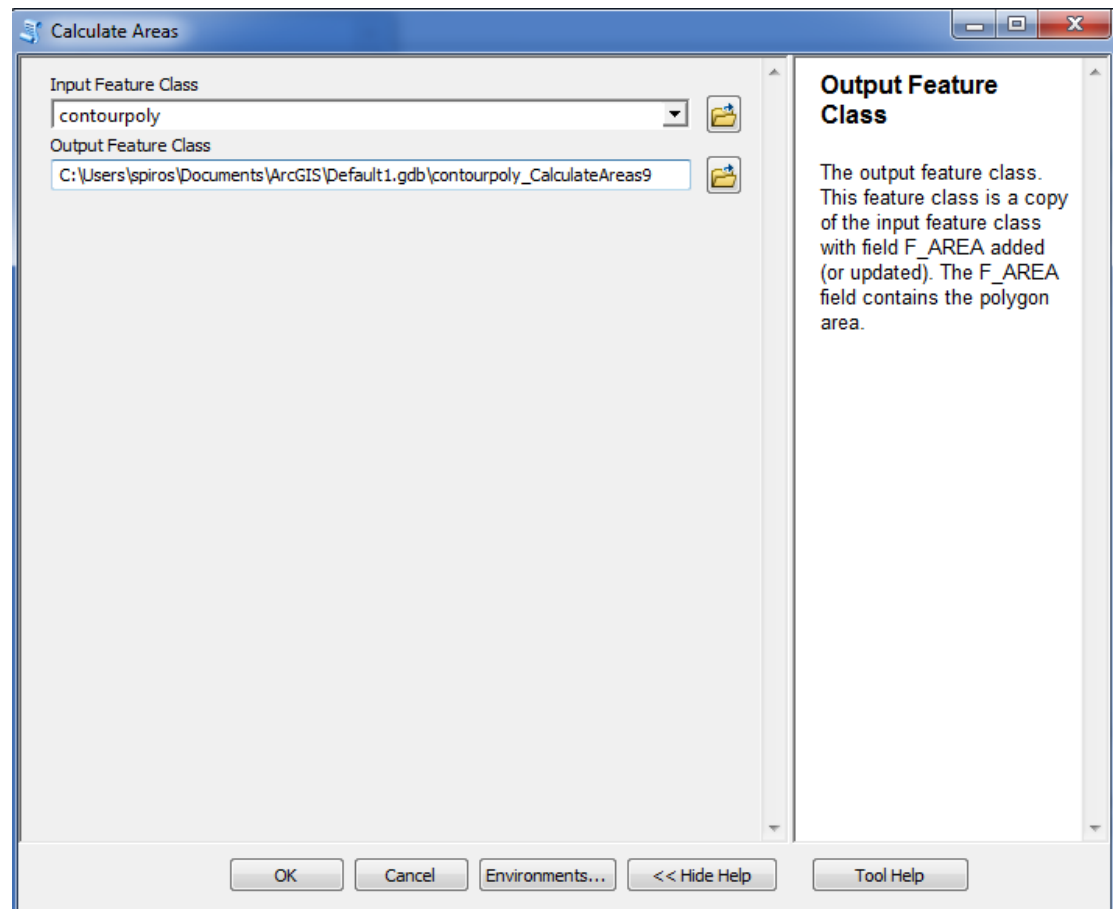


Σχήμα 4.23: Δειγματοληπτικό Υπόβαθρο Στρωματοποιημένης Τυχαίας Δειγματοληψίας

Ο υπολογισμός του εμβαδού της κάθε ζώνης ξεχωριστά αλλά και του συνολικού εμβαδού προκύπτει με την εντολή “ Calculate Areas “ από την εργαλειοθήκη Spatial Statistics Tools. [ArcToolBox > Spatial Statistics Tools > Utilities > Calculate Areas].



Σχήμα 4.24: Παράθυρο ArcToolbox Calculate Areas Στρωματοποιημένης Τυχαίας Δειγματοληψίας



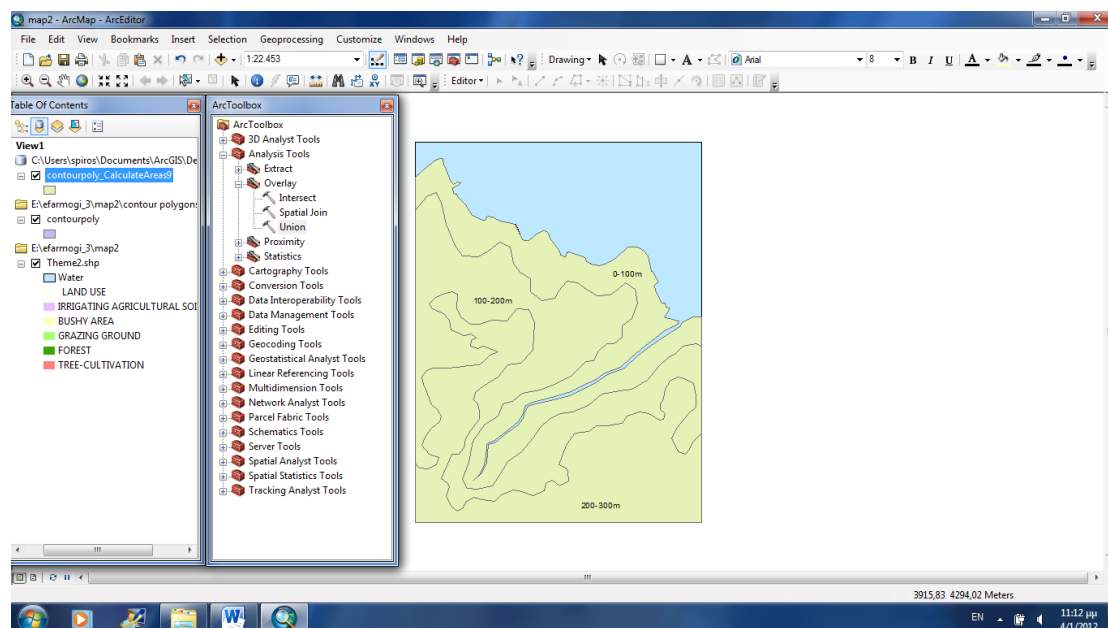
Σχήμα 4.25: Παράμετροι Παραθύρου Calculate Areas Στρωματοποιημένης Τυχαίας Δειγματοληψίας

Στο πρώτο πεδίο επιλέγουμε το όνομα του επιπέδου του οποίου θέλουμε να υπολογίσουμε τα εμβαδά (contourpoly) και στο δεύτερο πεδίο επιλέγουμε το όνομα και την τοποθεσία του επιπέδου που θα προκύψει από τον υπολογισμό των εμβαδών και θα είναι μια αντιγραφή του αρχικού επιπέδου στο πρώτο πεδίο με τη μόνη διαφορά ότι θα έχει υπολογίσει τα εμβαδά και στις 3 ζώνες (contourpoly_CalculateAreas9). Έτσι προκύπτει ο πίνακας με τον υπολογισμό των εμβαδών και στις τρεις υψομετρικές ζώνες.

OBJECTID *	Shape *	HeightZone	Shape Length	Shape Area	F AREA
1	Polygon	100 - 200 m	17905,819084	1740204,583225	1740204,583225
2	Polygon	200 -300 m	11881,017198	1432989,692646	1432989,692646
3	Polygon	0 -100 m	14535,925332	1090813,400669	1090813,400669

Σχήμα 4.26: Πίνακας Υπολογισμού Εμβαδών στις 3 υψομετρικές ζώνες contourpoly_CalculateAreas9

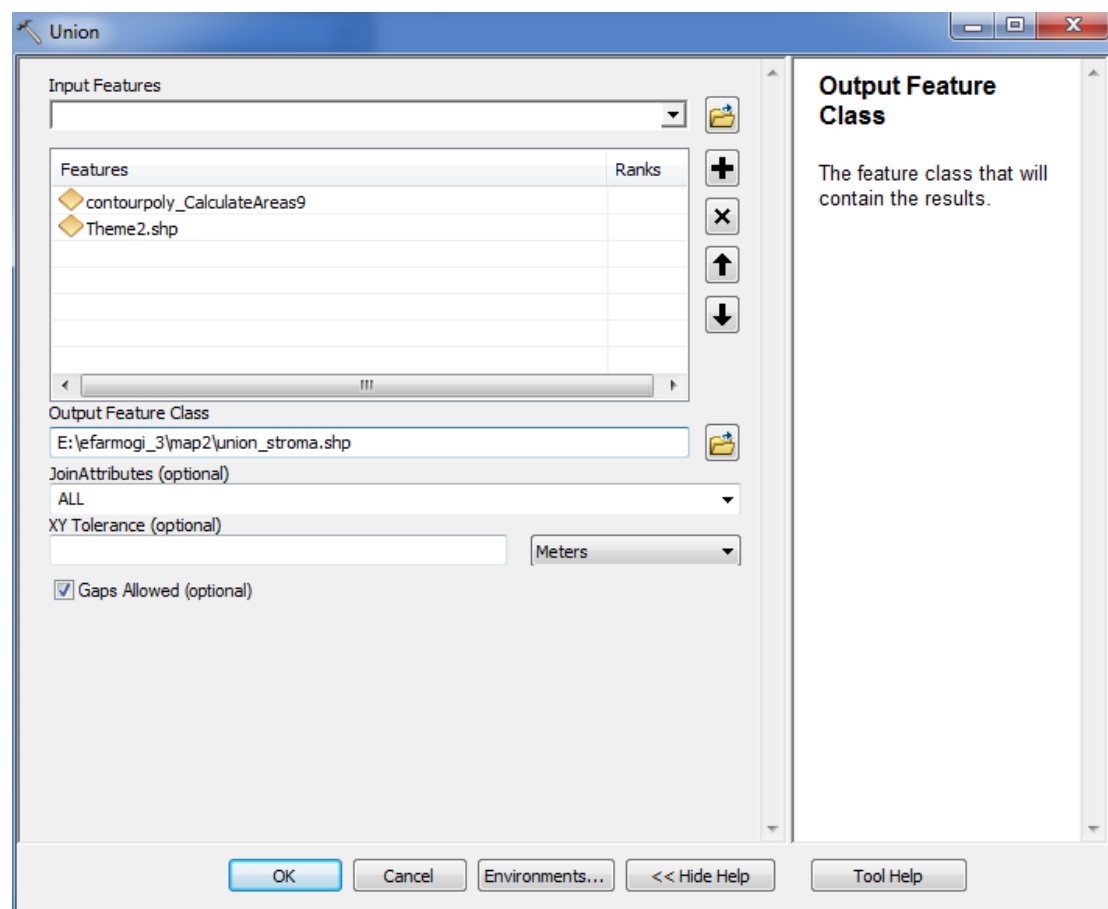
Στη στρωματοποιημένη τυχαία δειγματοληψία θέλουμε να βάλουμε υψόμετρο στις χρήσεις γης. Για γίνει αυτό κάνουμε ένωση (Union) μεταξύ του επιπέδου με τα υψομετρικά πολύγωνα (contourpoly_CalculateAreas9) και του επιπέδου με τις χρήσεις γης (theme 2). Αυτό γίνεται με την εντολή “ Union “ από την εργαλειοθήκη Analysis Tools. [ArcToolBox > Analysis Tools > Overlay > Union].



Σχήμα 4.27: Παράθυρο ArcToolbox Union Στρωματοποιημένης Τυχαίας Δειγματοληψίας

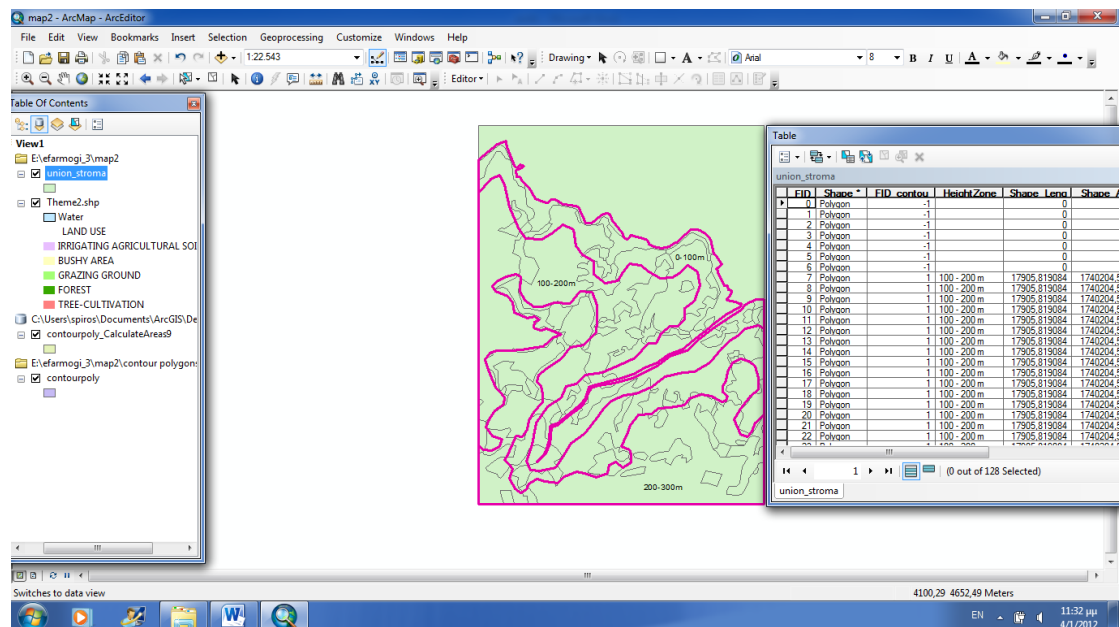
Το παράθυρο διαλόγου εμφανίζεται και συμπληρώνεται όπως παρακάτω:

Στο πεδίο Input Features εισάγουμε τα στοιχεία στα οποία θέλουμε να εφαρμόσουμε την ένωση (contourpoly_CalculateAreas9 & theme 2) και στο πεδίο Output Feature Class εισάγουμε αυτό που θα προκύψει από την ένωση (union_stroma).



Σχήμα 4.28: Παράμετροι Παραθύρου Union Στρωματοποιημένης Τυχαίας Δειγματοληψίας

Έτσι προκύπτει το επίπεδο union_stroma το οποίο περιέχει σε πίνακα πολύγωνα χρήσεων γης τα οποία βρίσκονται σε συγκεκριμένο υψόμετρο.



Σχήμα 4.29: Κατηγοριοποίηση Πολυγώνων Χρήσεων Γης σε Συγκεκριμένο Υψόμετρο (union_stroma)

Επειδή υπάρχουν λάθη στη συγκεκριμένη ψηφιοποίηση συγκεκριμένα σε ελάχιστα σημεία τα υψομετρικά πολύγωνα δεν εφάπτονται μεταξύ τους στο αρχείο union_stroma. Προκύπτουν πολύγωνα με έλλειψη υψομέτρου τα οποία δεν ανήκουν σε κάποια υψομετρική περιοχή.

Table

union_stroma

FID	Shape *	FID contou	HeightZone	Shape Leng	Shape Area	F AREA	FID theme2	LU CODE
0	Polygon	-1		0	0	0	0	500
1	Polygon	-1		0	0	0	3	200
2	Polygon	-1		0	0	0	6	700
3	Polygon	-1		0	0	0	11	100
4	Polygon	-1		0	0	0	17	100
5	Polygon	-1		0	0	0	24	400
6	Polygon	-1		0	0	0	38	200
7	Polygon	1	100 - 200 m	17905.819084	1740204.58323	1740204.58323	3	200
8	Polygon	1	100 - 200 m	17905.819084	1740204.58323	1740204.58323	5	400
9	Polygon	1	100 - 200 m	17905.819084	1740204.58323	1740204.58323	7	200
10	Polygon	1	100 - 200 m	17905.819084	1740204.58323	1740204.58323	10	400
11	Polygon	1	100 - 200 m	17905.819084	1740204.58323	1740204.58323	14	200
12	Polygon	1	100 - 200 m	17905.819084	1740204.58323	1740204.58323	15	400
13	Polygon	1	100 - 200 m	17905.819084	1740204.58323	1740204.58323	16	200
14	Polygon	1	100 - 200 m	17905.819084	1740204.58323	1740204.58323	19	400
15	Polygon	1	100 - 200 m	17905.819084	1740204.58323	1740204.58323	20	400
16	Polygon	1	100 - 200 m	17905.819084	1740204.58323	1740204.58323	21	400
17	Polygon	1	100 - 200 m	17905.819084	1740204.58323	1740204.58323	22	300
18	Polygon	1	100 - 200 m	17905.819084	1740204.58323	1740204.58323	23	300
19	Polygon	1	100 - 200 m	17905.819084	1740204.58323	1740204.58323	24	400
20	Polygon	1	100 - 200 m	17905.819084	1740204.58323	1740204.58323	26	200
21	Polygon	1	100 - 200 m	17905.819084	1740204.58323	1740204.58323	27	300
22	Polygon	1	100 - 200 m	17905.819084	1740204.58323	1740204.58323	28	300
23	Polygon	1	100 - 200 m	17905.819084	1740204.58323	1740204.58323	31	400
24	Polygon	1	100 - 200 m	17905.819084	1740204.58323	1740204.58323	33	300
25	Polygon	1	100 - 200 m	17905.819084	1740204.58323	1740204.58323	35	400
26	Polygon	1	100 - 200 m	17905.819084	1740204.58323	1740204.58323	36	300
27	Polygon	1	100 - 200 m	17905.819084	1740204.58323	1740204.58323	37	200
28	Polygon	1	100 - 200 m	17905.819084	1740204.58323	1740204.58323	38	200
29	Polygon	1	100 - 200 m	17905.819084	1740204.58323	1740204.58323	39	300
30	Polygon	1	100 - 200 m	17905.819084	1740204.58323	1740204.58323	40	100
31	Polygon	1	100 - 200 m	17905.819084	1740204.58323	1740204.58323	41	400
32	Polygon	1	100 - 200 m	17905.819084	1740204.58323	1740204.58323	42	200

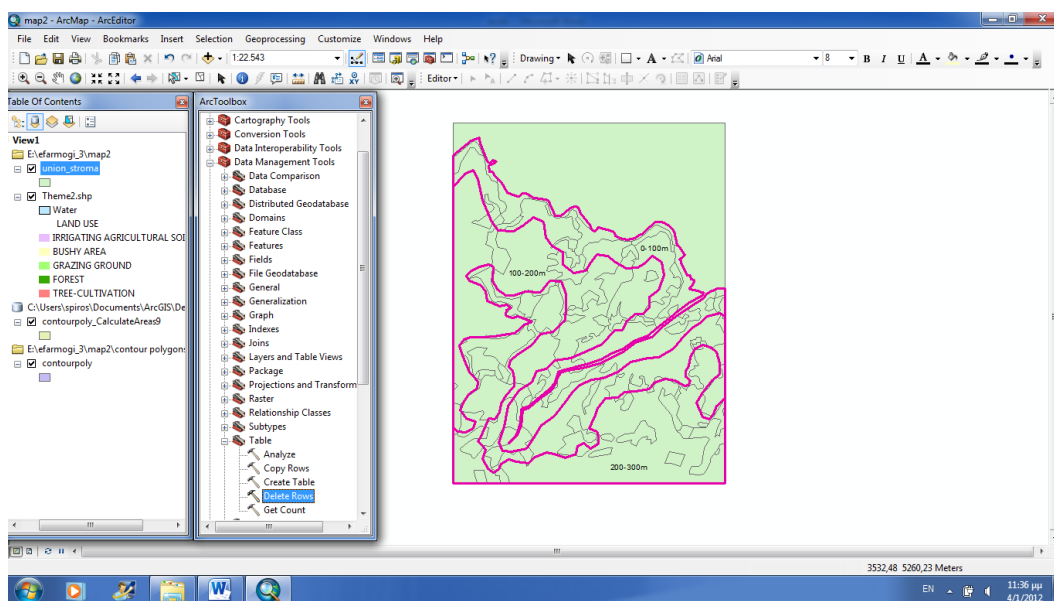
(0 out of 128 Selected)

union_stroma

Σχήμα 4.30: Αρχικός Πίνακας Κατηγοριοποίησης Πολυγώνων Χρήσεων Γης σε Συγκεκριμένο Υψόμετρο (union_stroma)

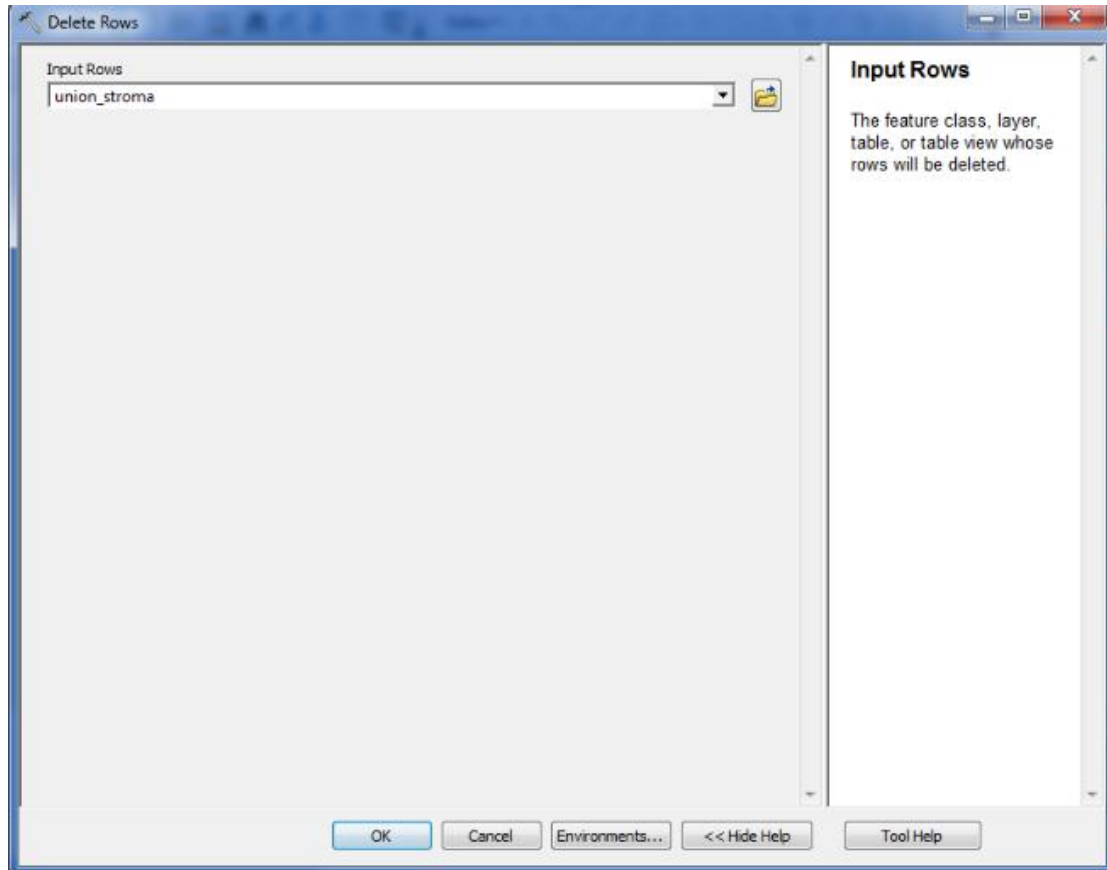
Τελικά, επειδή οι περιοχές είναι μικρές και δεν θέλουμε να τις αφαιρέσουμε κατατάσσοντας τις σε αυθαίρετη υψομετρική τιμή προτιμάμε την απαλοιφή τους. Αυτό γίνεται με την εντολή “ Delete Rows “ από την εργαλειοθήκη Data Management Tools.

[ArcToolbox> Data Management Tools> Table > Delete Rows].

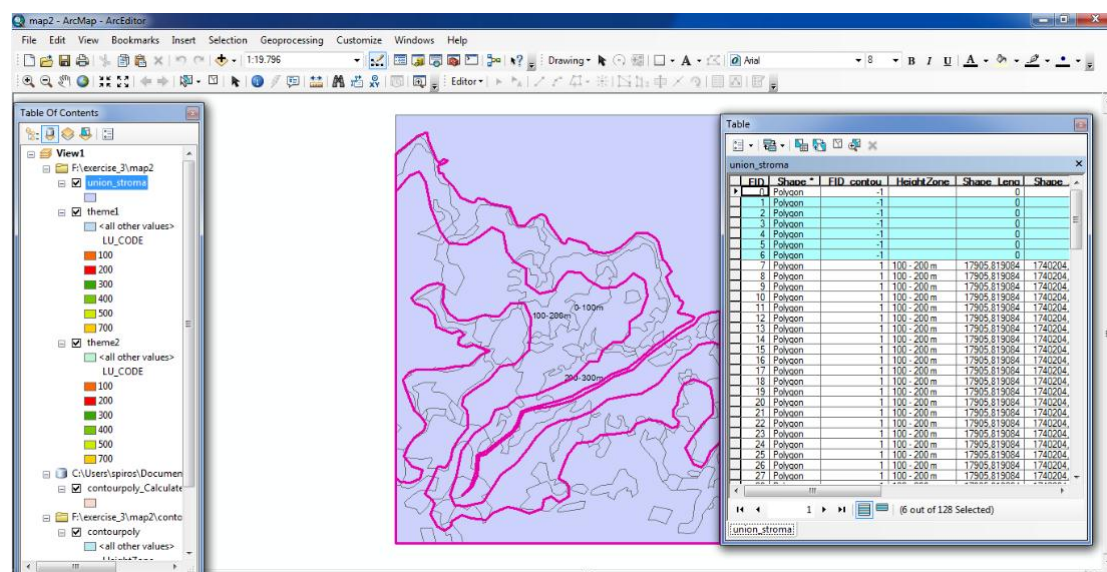


Σχήμα 4.31: Παράθυρο ArcToolbox Delete Rows Στρωματοποιημένης Τυχαίας Δειγματοληψίας

Το παράθυρο διαλόγου εμφανίζεται και συμπληρώνεται όπως παρακάτω:
 Στο πεδίο Input Features επιλέγουμε το επίπεδο στο οποία θέλουμε να εφαρμόσουμε την διαγραφή γραμμών (union_stroma).



Σχήμα 4.32: Παράμετροι Παραθύρου Delete Rows



Σχήμα 4.33: Απαλοιφή Πολυγώνων Χρήσεων Γης σε Συγκεκριμένο Υψόμετρο (union_stroma)

FID	Shape	FID contour	Height Zone	Shape Leng	Shape Area	F AREA	FID theme2	LU CODE
0	Polygon	-1		0	0	0	0	500
1	Polygon	1	100 - 200 m	17905.819084	1740204.58323	1740204.58323	3	200
2	Polygon	1	100 - 200 m	17905.819084	1740204.58323	1740204.58323	5	400
3	Polygon	1	100 - 200 m	17905.819084	1740204.58323	1740204.58323	7	200
4	Polygon	1	100 - 200 m	17905.819084	1740204.58323	1740204.58323	10	400
5	Polygon	1	100 - 200 m	17905.819084	1740204.58323	1740204.58323	14	200
6	Polygon	1	100 - 200 m	17905.819084	1740204.58323	1740204.58323	15	400
7	Polygon	1	100 - 200 m	17905.819084	1740204.58323	1740204.58323	16	200
8	Polygon	1	100 - 200 m	17905.819084	1740204.58323	1740204.58323	19	400
9	Polygon	1	100 - 200 m	17905.819084	1740204.58323	1740204.58323	20	400
10	Polygon	1	100 - 200 m	17905.819084	1740204.58323	1740204.58323	21	400
11	Polygon	1	100 - 200 m	17905.819084	1740204.58323	1740204.58323	22	300
12	Polygon	1	100 - 200 m	17905.819084	1740204.58323	1740204.58323	23	300
13	Polygon	1	100 - 200 m	17905.819084	1740204.58323	1740204.58323	24	400
14	Polygon	1	100 - 200 m	17905.819084	1740204.58323	1740204.58323	26	200
15	Polygon	1	100 - 200 m	17905.819084	1740204.58323	1740204.58323	27	300
16	Polygon	1	100 - 200 m	17905.819084	1740204.58323	1740204.58323	28	300
17	Polygon	1	100 - 200 m	17905.819084	1740204.58323	1740204.58323	31	400
18	Polygon	1	100 - 200 m	17905.819084	1740204.58323	1740204.58323	33	300
19	Polygon	1	100 - 200 m	17905.819084	1740204.58323	1740204.58323	35	400
20	Polygon	1	100 - 200 m	17905.819084	1740204.58323	1740204.58323	36	300
21	Polygon	1	100 - 200 m	17905.819084	1740204.58323	1740204.58323	37	200
22	Polygon	1	100 - 200 m	17905.819084	1740204.58323	1740204.58323	38	200
23	Polygon	1	100 - 200 m	17905.819084	1740204.58323	1740204.58323	39	300
24	Polygon	1	100 - 200 m	17905.819084	1740204.58323	1740204.58323	40	100
25	Polygon	1	100 - 200 m	17905.819084	1740204.58323	1740204.58323	41	400
26	Polygon	1	100 - 200 m	17905.819084	1740204.58323	1740204.58323	42	200
27	Polygon	1	100 - 200 m	17905.819084	1740204.58323	1740204.58323	44	200
28	Polygon	1	100 - 200 m	17905.819084	1740204.58323	1740204.58323	45	400
29	Polygon	1	100 - 200 m	17905.819084	1740204.58323	1740204.58323	49	300
30	Polygon	1	100 - 200 m	17905.819084	1740204.58323	1740204.58323	50	400
31	Polygon	1	100 - 200 m	17905.819084	1740204.58323	1740204.58323	52	300
32	Polygon	1	100 - 200 m	17905.819084	1740204.58323	1740204.58323	55	300

Σχήμα 4.34: Τελικός Πίνακας Κατηγοριοποίησης Πολυγώνων Χρήσεων Γης σε Συγκεκριμένο Υψόμετρο (union_stroma)

Το οριστικό μέγεθος δείγματος της στρωματοποιημένης τυχαίας δειγματοληψίας είναι ίδιο με αυτό που χρησιμοποιήθηκε και στην απλή τυχαία δειγματοληψία (2401 σημεία). Το επίπεδο εμπιστοσύνης και το περιθώριο λάθους δεν μεταβάλλονται. Συνεπώς, ο αριθμός των τυχαίων σημείων που θα ληφθούν σε κάθε υψομετρική ζώνη προκύπτει από τη σχέση:

$$n_1 = n_{\text{new}} * E1/ \Sigma E$$

$$n_2 = n_{\text{new}} * E2/ \Sigma E$$

$$n_3 = n_{\text{new}} * E3/ \Sigma E$$

Από τα παραπάνω συμπεραίνουμε ότι ο αριθμός των σημείων είναι ανάλογος της επιφάνειας του κάθε στρώματος – υψομετρικής ζώνης (φυσική στρωματοποίηση).

Χρήσεις γης	Δοκιμαστική	
	Αριθμός σημείων	Ποσοστά Χρήσεων γης (%)
100	1	3,33333
200	15	50,00000
300	3	10,00000
400	11	36,66667
700	0	0,00000
Σύνολο	30	100,00000
e=	2,00	
z	1,96	
p%=	50,00	
q%=	50,00	
nnew=	2401	
E1=	1090813	0 - 100m
E2=	1740205	100 - 200 m
E3=	1432990	200 - 300m
ΣΕ=	4264008	
n1=	614	
n2=	980	
n3=	807	

Σχήμα 4.35: Φύλλο Excel για τον καταμερισμό σημείων ανά υψομετρική ζώνη

Για την υψομετρική ζώνη 0 – 100 m έχουμε n_1
 Για την υψομετρική ζώνη 100 – 200 m έχουμε n_2
 Για την υψομετρική ζώνη 200 – 300 m έχουμε n_3

Άρα πλέον μένει να πραγματοποιηθούν 3 απλές τυχαίες δειγματοληψίες σε κάθε υψομετρική ζώνη.

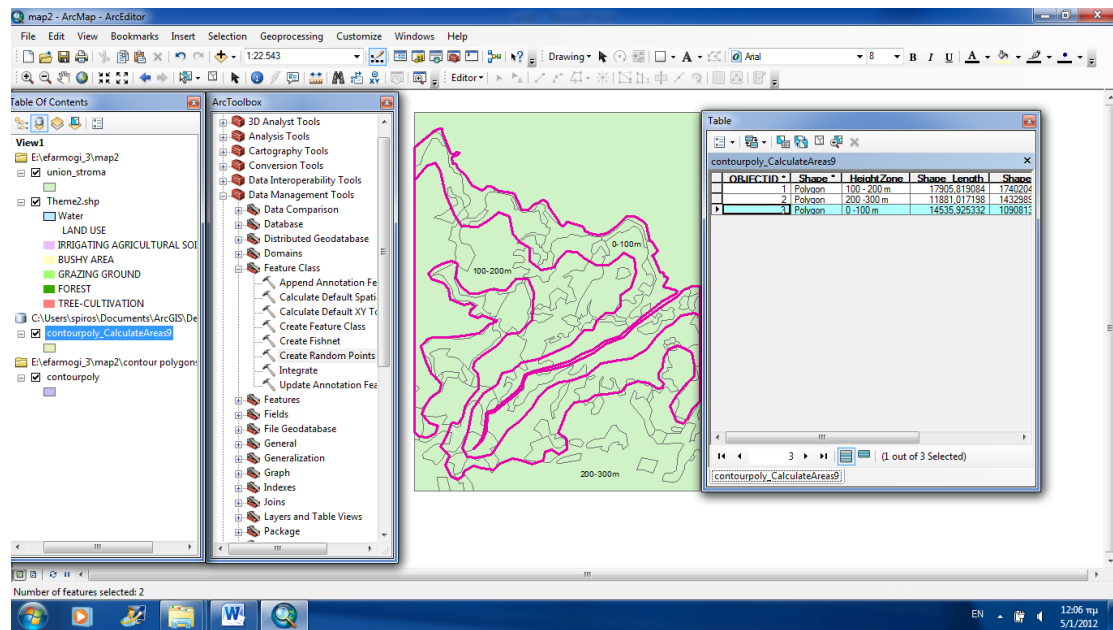
Υψομετρική Ζώνη 1 (1^ο στρώμα)

Ο αριθμός των τυχαίων σημείων που λαμβάνονται είναι $n_1 = 614$ σημεία.

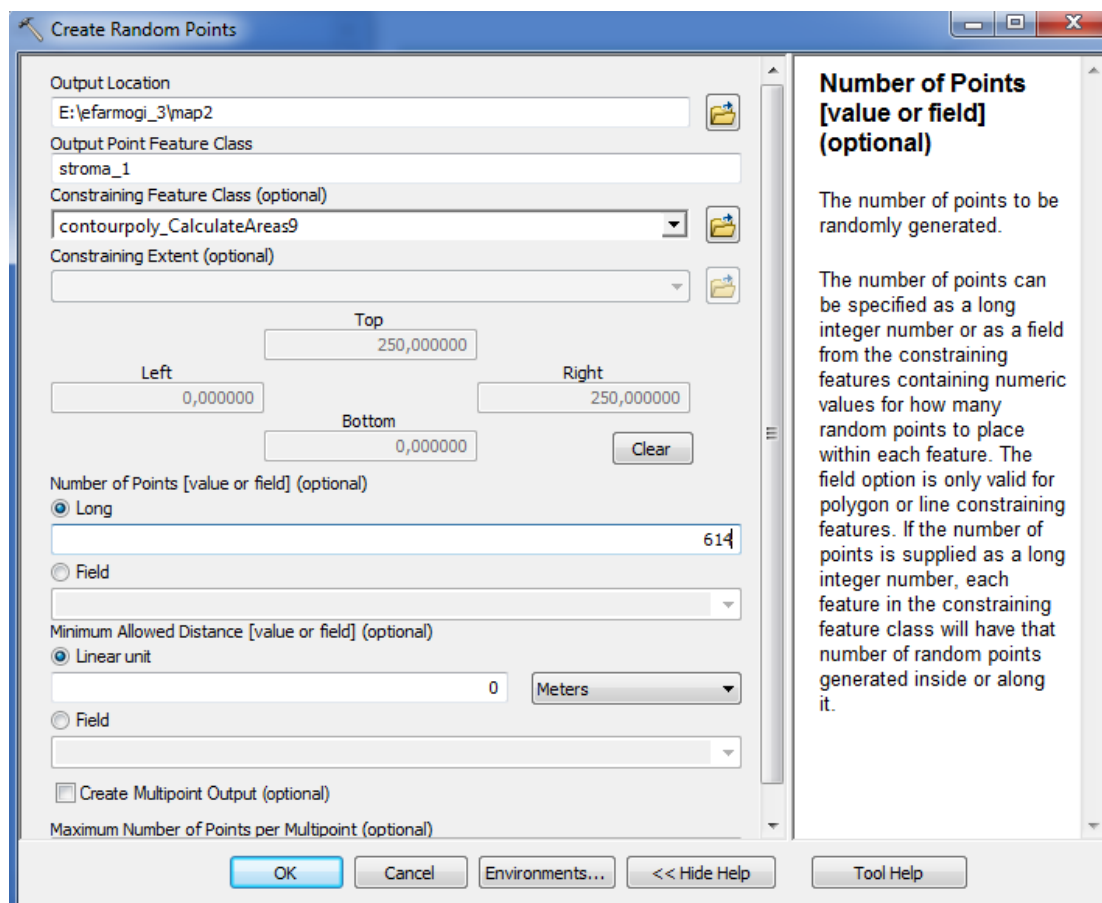
Η δειγματοληψία ξεκινάει με τη λήψη ενός δοκιμαστικού δείγματος 614 παρατηρήσεων. Αυτό γίνεται με την εντολή “ Create Random Points “ από την εργαλειοθήκη Data Management Tools.

[ArcToolBox > Data Management Tools > Feature Class > Create Random Points].

Το παράθυρο διαλόγου εμφανίζεται και συμπληρώνεται όπως παρακάτω:



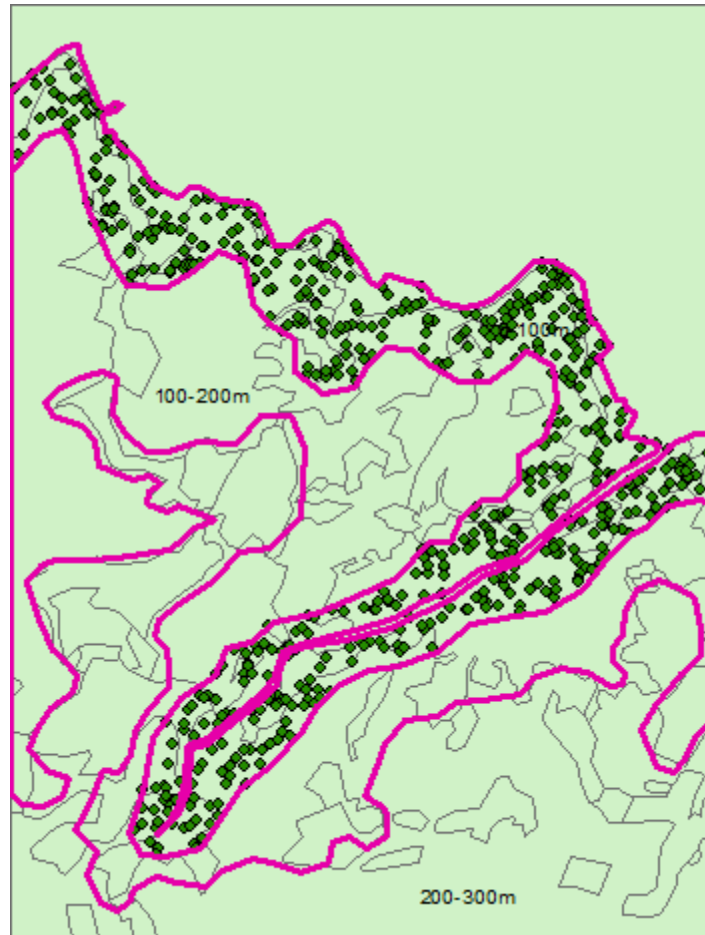
Σχήμα 4.36: Παράθυρο ArcToolbox Create Random Points Στρωματοποιημένης Τυχαίας Δειγματοληψίας στην 1^η υψομετρική ζώνη



Σχήμα 4.37: Παράμετροι Παραθύρου Create Random Points Στρωματοποιημένης Τυχαίας Δειγματοληψίας στην 1^η υψομετρική ζώνη

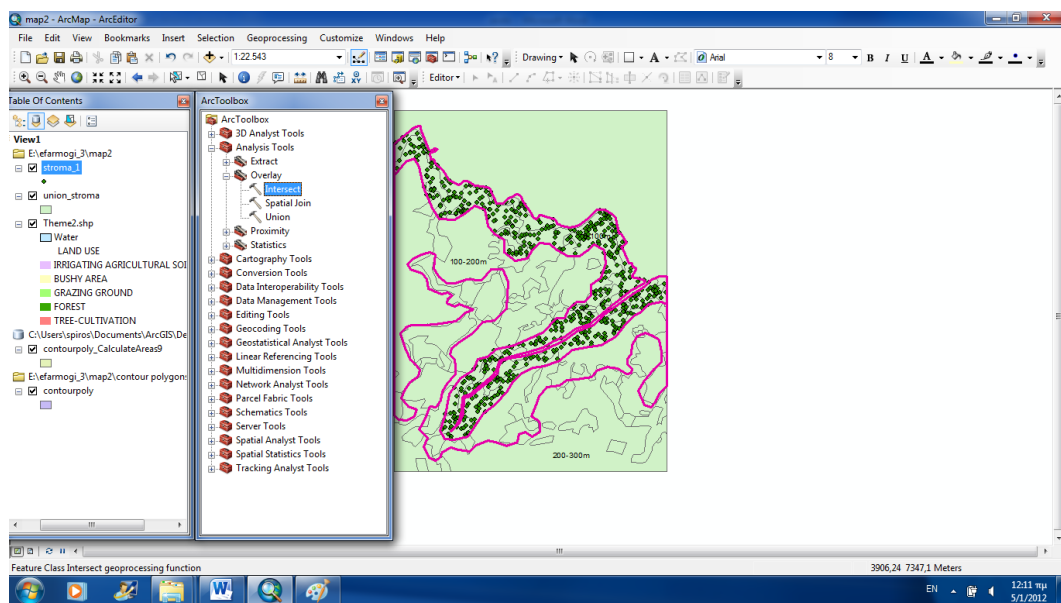
Στο πρώτο πεδίο επιλέγουμε την τοποθεσία στην οποία τα σημεία θα δημιουργηθούν. Στο δεύτερο πεδίο επιλέγουμε το όνομα που θα προκύψει από την γέννηση των σημείων (stroma_1) και στο τρίτο πεδίο επιλέγουμε περιορισμό κατηγορίας χαρακτηριστικών γνωρισμάτων στο συγκεκριμένο πολύγωνο με τις υψομετρικές ζώνες.

Προχωρώντας στο πεδίο number of points εισάγουμε τον αριθμό των σημείων που θέλουμε να δημιουργήσουμε (614) και στο πεδίο minimum allowed distance ορίζουμε την ελάχιστη απόσταση μεταξύ δύο σημείων (0 meters).

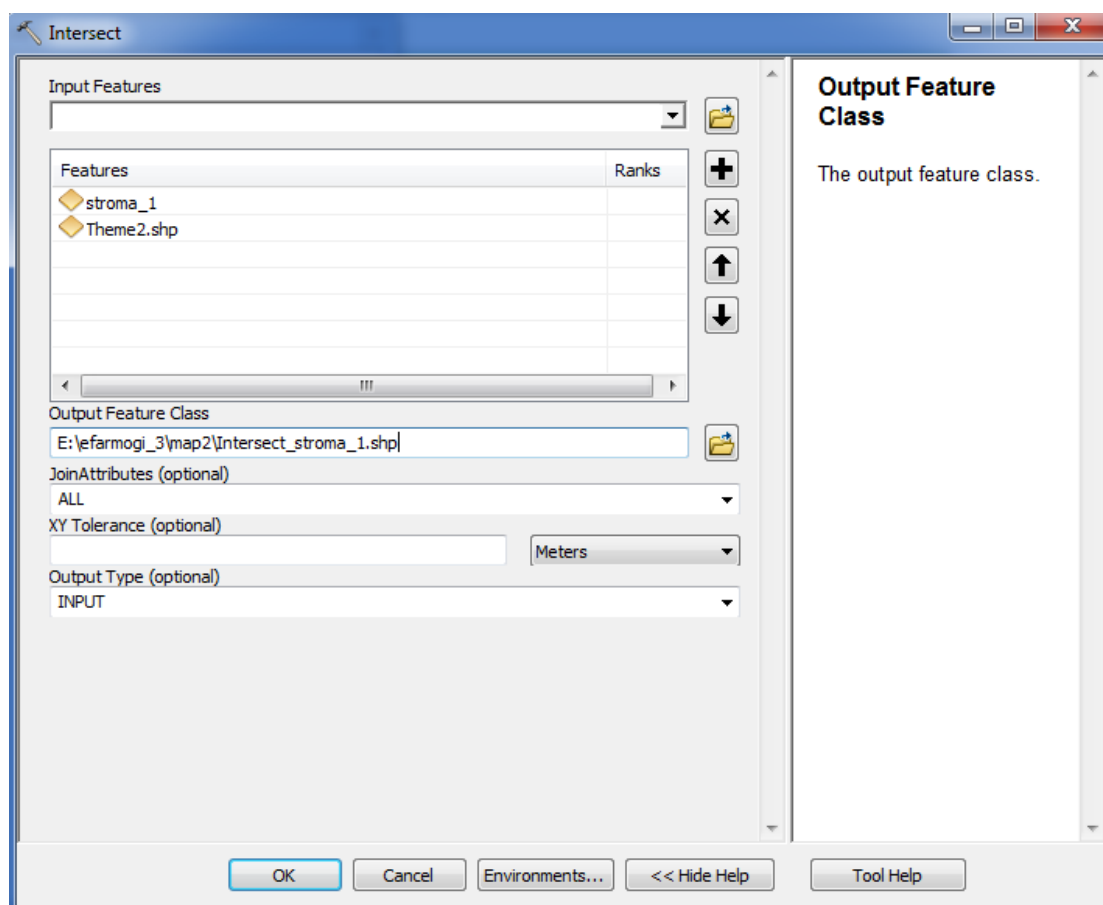


Σχήμα 4.38: Δείγμα 614 παρατηρήσεων stroma_1

Έπειτα με την εντολή Intersect κατατάσσουμε τα σημεία στην χρήση γης στην οποία ανήκουν. Η διαδικασία που ακολουθούμε είναι [ArcToolBox > Analysis Tools > Overlay > Intersect]. Το παράθυρο διαλόγου εμφανίζεται και συμπληρώνεται όπως παρακάτω:



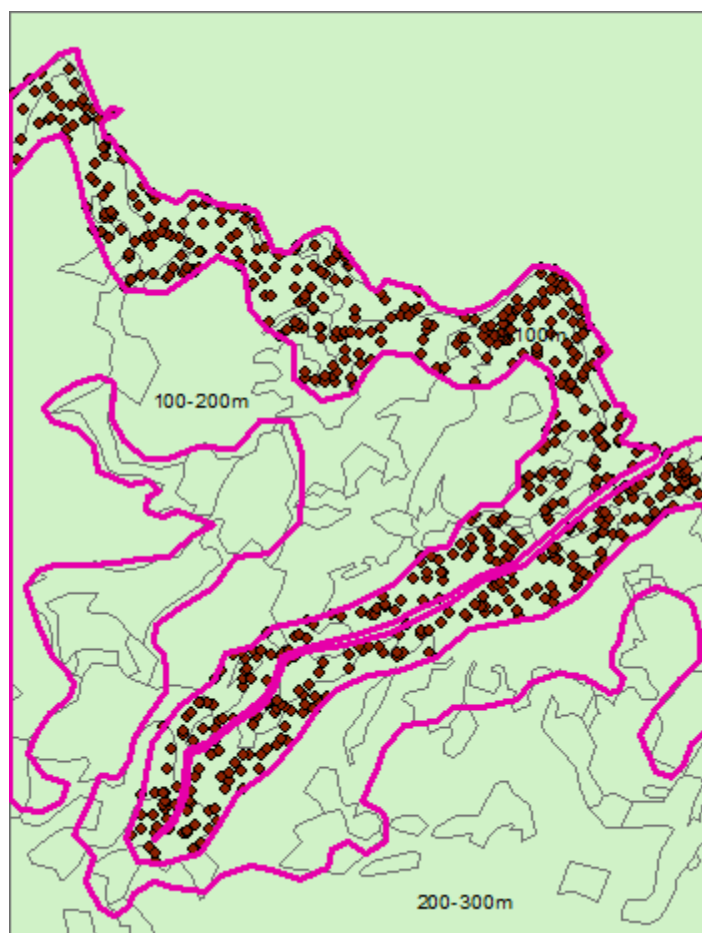
Σχήμα 4.39: Παράθυρο Arc Toolbox Intersect Στρωματοποιημένης Τυχαίας Δειγματοληψίας στην 1^η υψομετρική ζώνη



Σχήμα 4.40: Παράμετροι Παραθύρου Intersect Στρωματοποιημένης Τυχαίας Δειγματοληψίας

Στο πεδίο Input Features επιλέγουμε τα επίπεδα στα οποία θέλουμε να εφαρμόσουμε την τομή (theme 2 & stroma_1) και στο πεδίο Output Feature Class επιλέγουμε το όνομα και την τοποθεσία του επιπέδου που θα προκύψει από την τομή (Intersect_stroma_1.shp).

Ο χάρτης απεικονίζεται ως εξής:



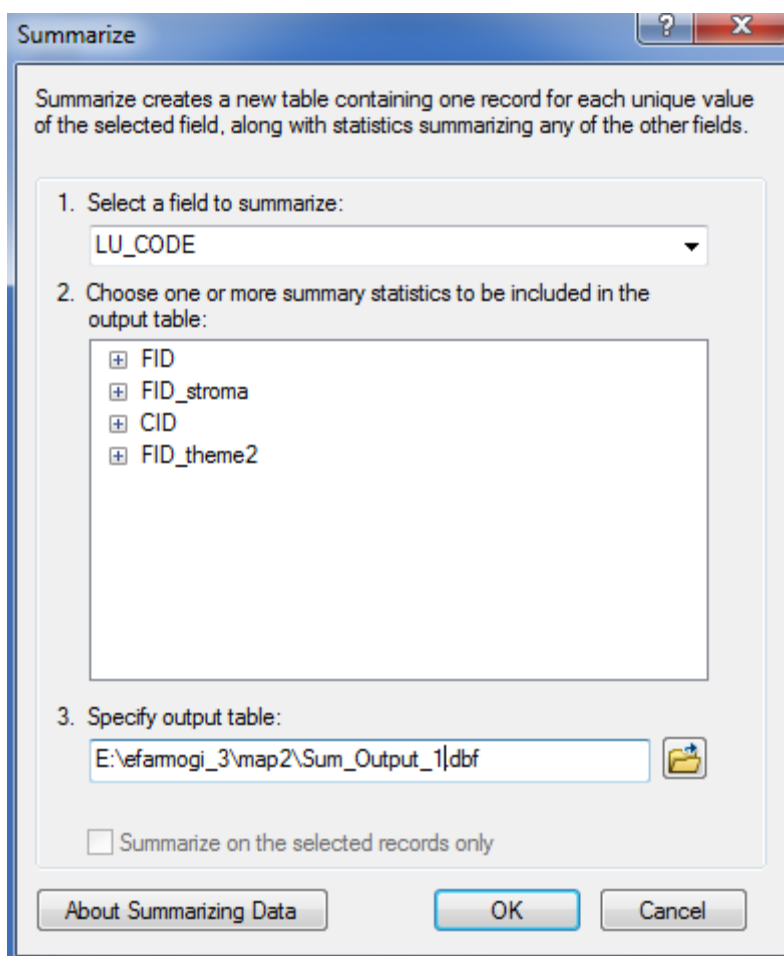
Σχήμα 4.41: Δείγμα 614 παρατηρήσεων Intersect_stroma_1

Περισσότερο αυτό που προκύπτει από την τομή είναι ένας πίνακας τιμών όπου κάθε σημείο αντιστοιχίζεται σε μια συγκεκριμένη κατηγορία χρήσης γης πάνω στην οποία πέφτει.

FID	Shape	FID_stroma	CID	FID_theme2	LU CODE
0	Point	161	3	66	200
1	Point	523	3	66	200
2	Point	115	3	66	200
3	Point	435	3	66	200
4	Point	206	3	66	200
5	Point	453	3	45	400
6	Point	407	3	67	400
7	Point	102	3	67	400
8	Point	201	3	66	200
9	Point	261	3	66	200
10	Point	404	3	66	200
11	Point	76	3	66	200
12	Point	313	3	66	200
13	Point	167	3	67	400
14	Point	207	3	67	400
15	Point	256	3	66	200
16	Point	468	3	45	400
17	Point	257	3	67	400
18	Point	296	3	67	400
19	Point	134	3	67	400
20	Point	100	3	67	400
21	Point	37	3	45	400
22	Point	442	3	66	200
23	Point	245	3	67	400
24	Point	488	3	45	400
25	Point	51	3	67	400
26	Point	216	3	58	200
27	Point	269	3	67	400
28	Point	434	3	66	200
29	Point	492	3	58	200
30	Point	279	3	66	200
31	Point	152	3	58	200
32	Point	267	3	58	200
33	Point	461	3	66	200

Σχήμα 4.42: Πίνακας Κατηγοριοποίησης Σημείων ανά Χρήση Γης Intersect_stroma_1

Στη συνέχεια κάνοντας δεξί κλικ στο πεδίο LU CODE στον πίνακα Intersect_stroma_1 επιλέγουμε την εντολή Summarize όπου μας εμφανίζεται ένα παράθυρο διαλόγου το οποίο το συμπληρώνουμε ως εξής:



Σχήμα 4.43: Παράθυρο Διαλόγου Summarize Στρωματοποιημένης Τυχαίας Δειγματοληψίας

Στο πρώτο πεδίο επιλέγουμε το πεδίο ως προς το οποίο θέλουμε να γίνει η κατηγοριοποίηση (LU CODE), στο δεύτερο πεδίο δεν επιλέγουμε αθροιστικά στατιστικά διότι το sum αθροίζει αύξοντες αριθμούς (stroma_1) στον πίνακα και στο τρίτο δίνουμε στον πίνακα ένα όνομα και μια κατάληξη (dbf αρχείο).

Αξίζει να σημειωθεί ότι στον πίνακα (Sum_Output_1) μας δίνεται το πλήθος των σημείων που πέφτουν σε κάθε χρήση γης που είναι εκείνο που μας ενδιαφέρει περισσότερο και θα χρησιμοποιήσουμε αργότερα στο φύλλο Excel για τον υπολογισμό των ποσοστών των επιμέρους χρήσεων γης ανά κατηγορία.

OID	LU CODE	Count	LU CODE
0	100	83	
1	200	340	
2	300	10	
3	400	146	
4	700	35	

Σχήμα 4.44: Πίνακας Σημείων ανά Χρήση Γης Στρωματοποιημένης Τυχαίας Δειγματοληψίας rinakas_stroma_1

Τέλος, εξάγουμε τα αποτελέσματα του πίνακα Sum_Output_1 σε λογιστικό φύλλο Excel όπου κάνουμε τους αντίστοιχους υπολογισμούς για τα ποσοστά χρήσεων γης και το test χ^2 αυτόματα.

Στρωματοποιημένη Τυχαία Σημειακή Δειγματοληψία _1			
Χρήσεις Γης	Αριθμός Σημείων	Ποσοστά Χρήσεων Γης (%)	
100	83	13,51792	
200	340	55,37459	
300	10	1,62866	
400	146	23,77850	
700	35	5,70033	
Σύνολο	614	100,00000	
	Αναμενόμενη Τιμή (%)	Πραγματική Τιμή (%)	Τέστ χ^2
p100	13,51792	3,6	27,32362
p200	55,37459	48,9	0,857267
p300	1,62866	5	2,273181
p400	23,77850	41,6	7,634755
p700	5,70033	0,9	25,60347
		$\chi^2_{Rtest} =$	63,6923

Σχήμα 4.45: Φύλλο Excel για τον αυτόματο υπολογισμό ποσοστών χρήσεων γης ανά κατηγορία στην στρωματοποιημένη τυχαία δειγματοληψία στην 1^η υψομετρική ζώνη και του test χ^2

Παράλληλα η τιμή του τεστ χ^2_d μπορεί να προκύψει αυτόματα για την παραπάνω δειγματοληψία. Συγκρίνοντας την τιμή του τεστ χ^2 για δεδομένο επίπεδο σημαντικότητας ($\alpha = 0.05$) και βαθμούς ελευθερίας $v = k - 1$ στη περίπτωση της στρωματοποιημένης τυχαίας δειγματοληψίας έχουμε $v = 4$. Επομένως το $\chi^2_d = 9.48773$.

Όπου: k = κατηγορία χρήσης γης (5)

Συμπεραίνουμε ότι στα δεδομένα του δείγματος δεν προσαρμόζεται ικανοποιητικά η κατανομή αφού $\chi^2 > \chi^2_d$.

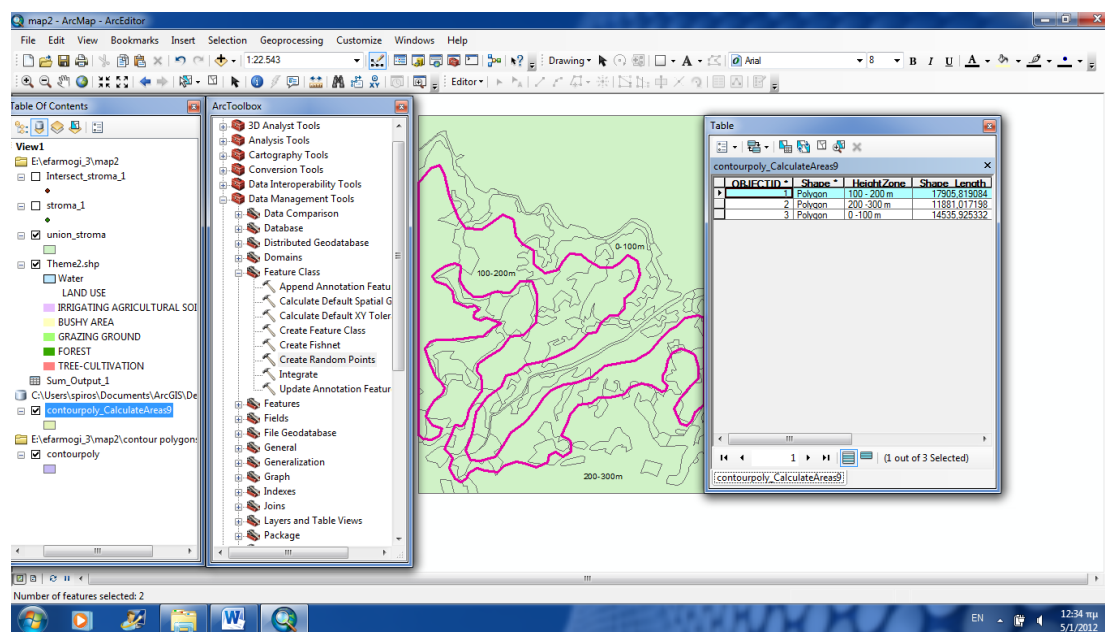
Υψομετρική Ζώνη 2 (2^ο στρώμα)

Ο αριθμός των τυχαίων σημείων που λαμβάνονται είναι $n_2 = 980$ σημεία.

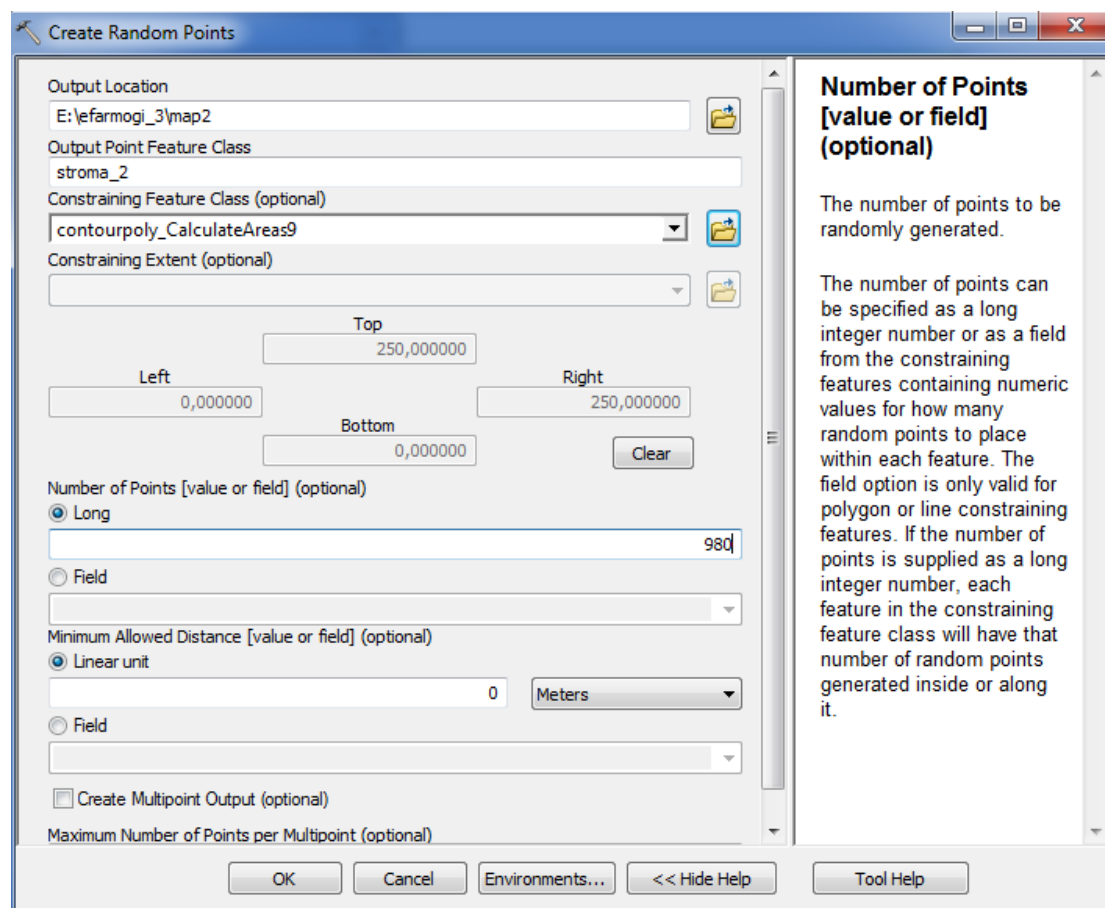
Η δειγματοληψία ξεκινάει με τη λήψη ενός δοκιμαστικού δείγματος 980 παρατηρήσεων. Αυτό γίνεται με την εντολή “ Create Random Points “ από την εργαλειοθήκη Data Management Tools.

[ArcToolBox > Data Management Tools > Feature Class > Create Random Points].

Το παράθυρο διαλόγου εμφανίζεται και συμπληρώνεται όπως παρακάτω:



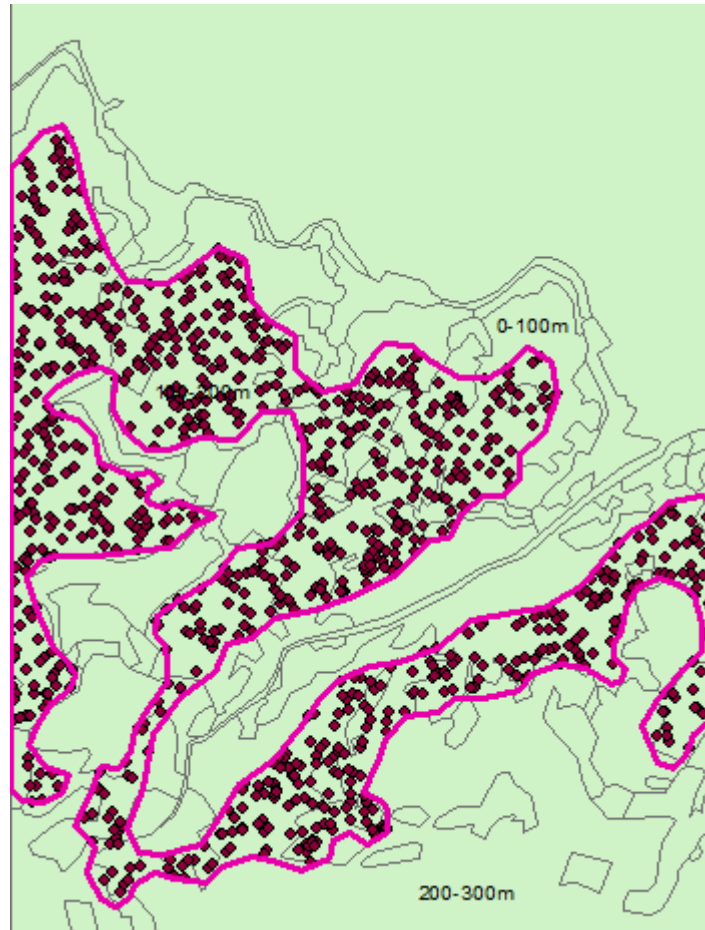
Σχήμα 4.46: Παράθυρο Arc Toolbox Create Random Points Στρωματοποιημένης Τυχαίας Δειγματοληψίας στην 2^η υψομετρική ζώνη



Σχήμα 4.47: Παράμετροι Παραθύρου Create Random Points Στρωματοποιημένης Τυχαίας Δειγματοληψίας στην 2^η υπομετρική ζώνη

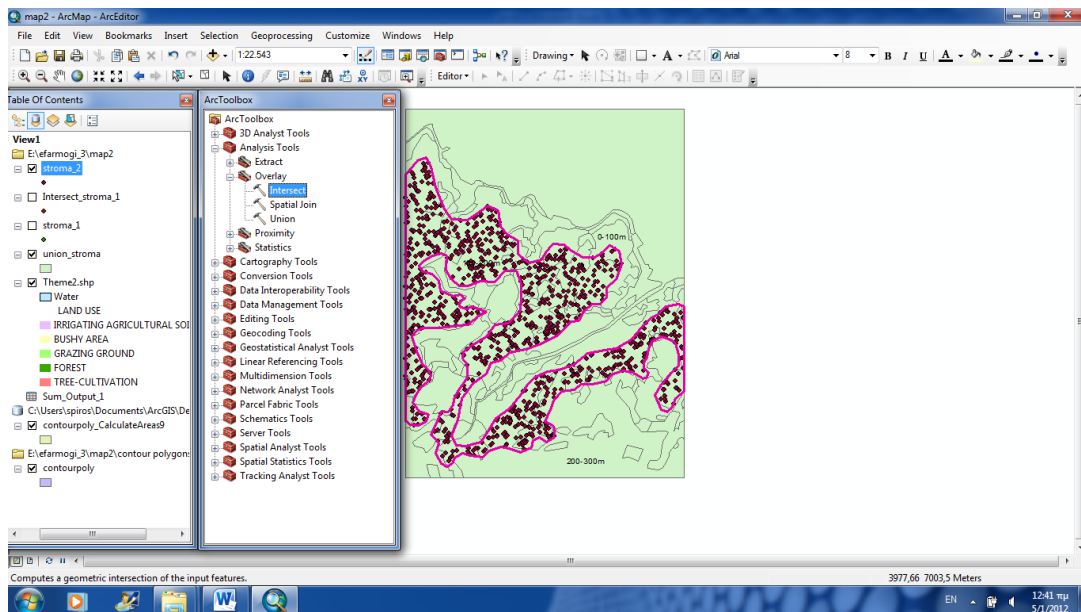
Στο πρώτο πεδίο επιλέγουμε την τοποθεσία στην οποία τα σημεία θα δημιουργηθούν. Στο δεύτερο πεδίο επιλέγουμε το όνομα που θα προκύψει από την γέννηση των σημείων (stroma_2) και στο τρίτο πεδίο επιλέγουμε περιορισμό κατηγορίας χαρακτηριστικών γνωρισμάτων στο συγκεκριμένο πολύγωνο με τις υπομετρικές ζώνες.

Προχωρώντας στο πεδίο number of points εισάγουμε τον αριθμό των σημείων που θέλουμε να δημιουργήσουμε (980) και στο πεδίο minimum allowed distance ορίζουμε την ελάχιστη απόσταση μεταξύ δύο σημείων (0 meters).

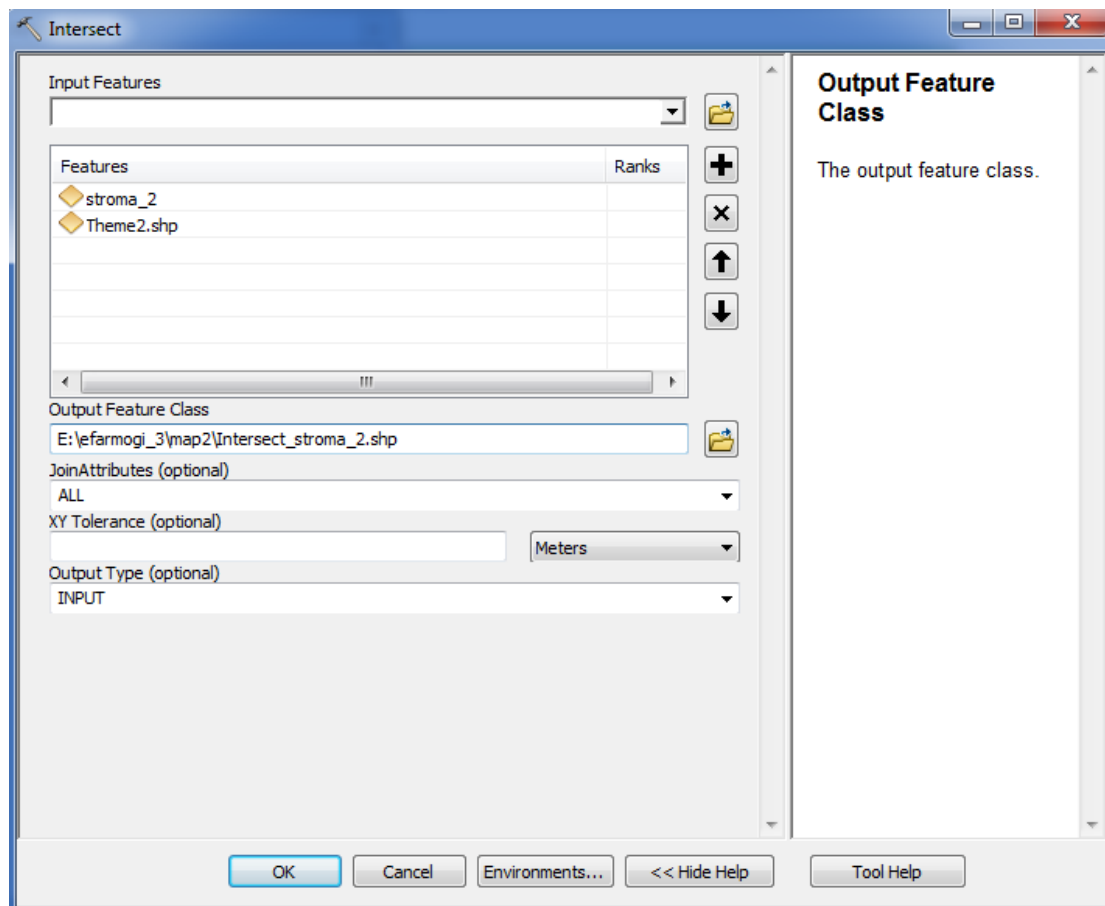


Σχήμα 4.48: Δείγμα 980 παρατηρήσεων stroma_2

Έπειτα με την εντολή Intersect κατατάσσουμε τα σημεία στην χρήση γης στην οποία ανήκουν. Η διαδικασία που ακολουθούμε είναι [ArcToolBox > Analysis Tools > Overlay > Intersect]. Το παράθυρο διαλόγου εμφανίζεται και συμπληρώνεται όπως παρακάτω:



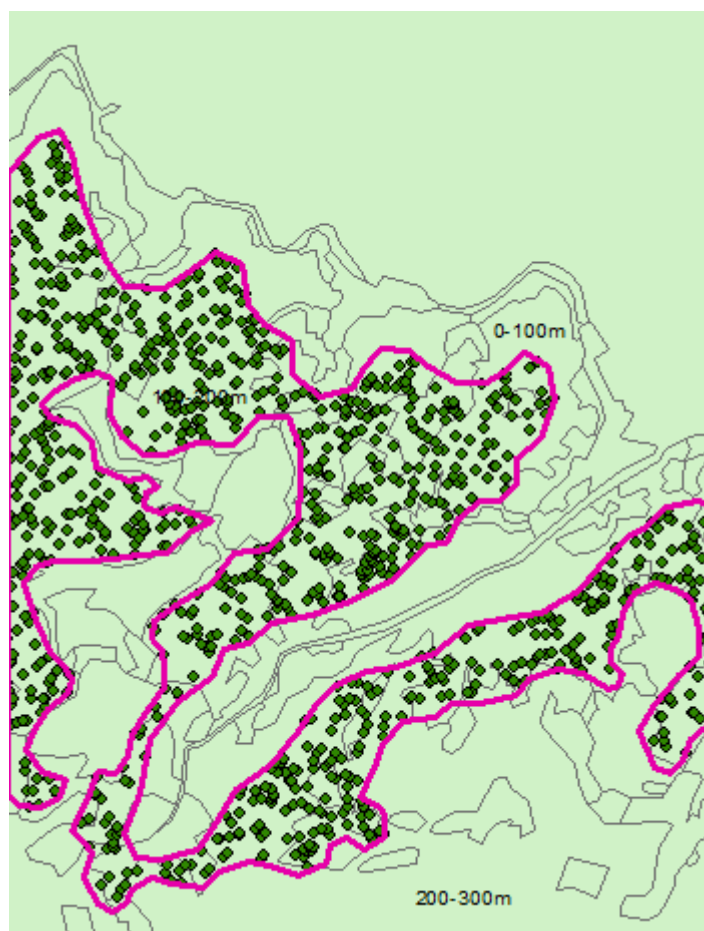
Σχήμα 4.49: Παράθυρο ArcToolbox Intersect Στρωματοποιημένης Τυχαίας Δειγματοληψίας στην 2^η υψομετρική ζώνη



Σχήμα 4.50: Παράμετροι Παραθύρου Intersect Στρωματοποιημένης Τυχαίας Δειγματοληψίας

Στο πεδίο Input Features επιλέγουμε τα στοιχεία στα οποία θέλουμε να εφαρμόσουμε την τομή (theme 2 & stroma_2) και στο πεδίο Output Feature Class επιλέγουμε το όνομα και την τοποθεσία του επιπέδου που θα προκύψει από την τομή (Intersect_stroma_2.shp).

Ο χάρτης απεικονίζεται ως εξής:



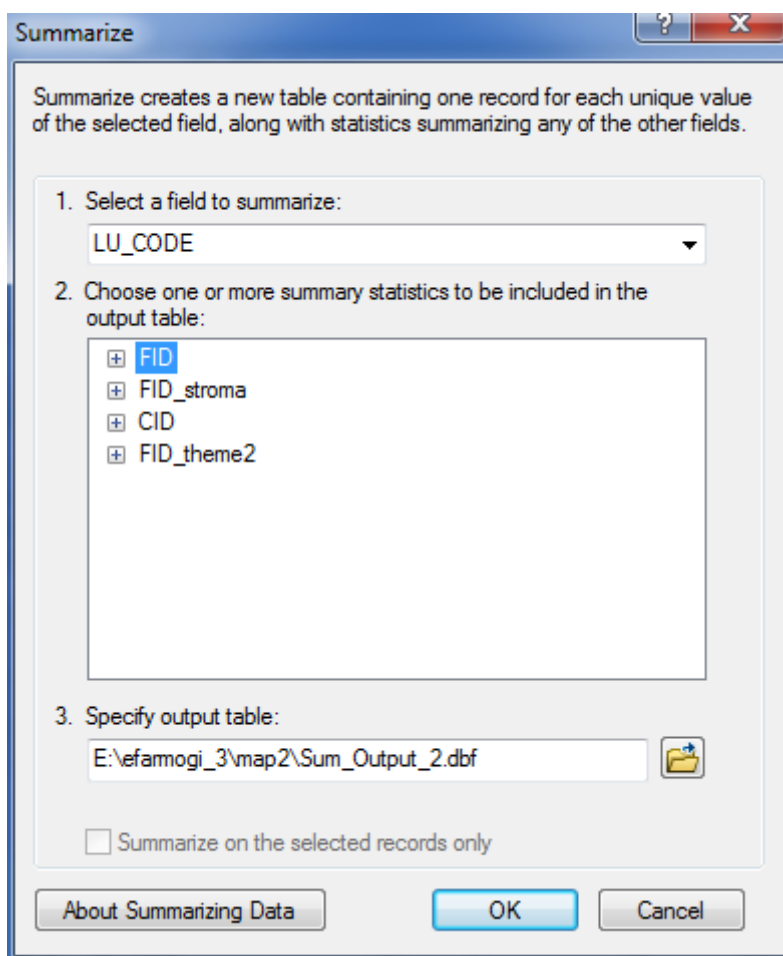
Σχήμα 4.51: Δείγμα 980 παρατηρήσεων Intersect_stroma_2

Περισσότερο αυτό που προκύπτει από την τομή είναι ένας πίνακας τιμών όπου κάθε σημείο αντιστοιχίζεται σε μια συγκεκριμένη κατηγορία χρήσης γης πάνω στην οποία πέφτει.

FID	Shape *	FID stroma	CID	FID theme2	LU CODE
0	Point	516	1	73	300
1	Point	721	1	66	200
2	Point	824	1	67	400
3	Point	861	1	45	400
4	Point	253	1	45	400
5	Point	542	1	45	400
6	Point	328	1	66	200
7	Point	14	1	66	200
8	Point	97	1	66	200
9	Point	517	1	45	400
10	Point	40	1	45	400
11	Point	44	1	45	400
12	Point	185	1	66	200
13	Point	647	1	45	400
14	Point	339	1	66	200
15	Point	836	1	66	200
16	Point	487	1	68	300
17	Point	378	1	45	400
18	Point	160	1	45	400
19	Point	71	1	68	300
20	Point	538	1	68	300
21	Point	358	1	45	400
22	Point	812	1	45	400
23	Point	154	1	66	200
24	Point	916	1	68	300
25	Point	522	1	66	200
26	Point	834	1	66	200
27	Point	153	1	68	300
28	Point	668	1	67	400
29	Point	301	1	66	200
30	Point	104	1	68	300
31	Point	860	1	68	300
32	Point	550	1	68	300
33	Point	383	1	67	400

Σχήμα 4.52: Πίνακας Κατηγοριοποίησης Σημείων ανά Χρήση Γης Intersect_stroma_2

Στη συνέχεια κάνοντας δεξί κλικ στο πεδίο LU CODE στον πίνακα Intersect_stroma_2 επιλέγουμε την εντολή Summarize όπου μας εμφανίζεται ένα παράθυρο διαλόγου το οποίο το συμπληρώνουμε ως εξής:



Σχήμα 4.53: Παράθυρο Διαλόγου Summarize Στρωματοποιημένης Τυχαίας Δειγματοληψίας

Στο πρώτο πεδίο επιλέγουμε το πεδίο ως προς το οποίο θέλουμε να γίνει η κατηγοριοποίηση (LU CODE), στο δεύτερο πεδίο δεν επιλέγουμε αθροιστικά στατιστικά διότι το sum αθροίζει αύξοντες αριθμούς (stroma_2) στον πίνακα και στο τρίτο δίνουμε στον πίνακα ένα όνομα και μια κατάληξη (dbf αρχείο).

Αξίζει να σημειωθεί ότι στον πίνακα (Sum_Output_2) μας δίνεται το πλήθος των σημείων που πέφτουν σε κάθε χρήση γης που είναι εκείνο που μας ενδιαφέρει περισσότερο και θα χρησιμοποιήσουμε αργότερα στο φύλλο Excel για τον υπολογισμό των ποσοστών των επιμέρους χρήσεων γης ανά κατηγορία.

oid	LU_CODE	Count	LU_CODE
0	100	2	2
1	200	687	687
2	300	42	42
3	400	249	249

Σχήμα 4.54: Πίνακας Σημείων ανά Χρήση Γης Στρωματοποιημένης Τυχαίας Δειγματοληψίας Sum_Output_2

Τέλος, εξάγουμε τα αποτελέσματα του πίνακα Sum_Output_2 σε λογιστικό φύλλο Excel όπου κάνουμε τους αντίστοιχους υπολογισμούς για τα ποσοστά χρήσεων γης και το test χ^2 αυτόματα.

Στρωματοποιημένη Τυχαία Σημειακή Δειγματοληψία _2			
Χρήσεις Γης	Αριθμός Σημείων	Ποσοστά Χρήσεων Γης (%)	
100	2	0,20408	
200	687	70,10204	
300	42	4,28571	
400	249	25,40816	
700	0	0,00000	
Σύνολο	980	100,00000	
	Αναμενόμενη Τιμή (%)	Πραγματική Τιμή (%)	Τέστ χ^2
p100	0,20408	3,6	3,203406
p200	70,10204	48,9	9,192772
p300	4,28571	5	0,102041
p400	25,40816	41,6	6,302298
p700	0,00000	0,9	0,9
		χ^2 Rtest=	19,70052

Σχήμα 4.55: Φύλλο Excel για τον αυτόματο υπολογισμό ποσοστών χρήσεων γης ανά κατηγορία στην στρωματοποιημένη τυχαία δειγματοληψία στην 2^η υψομετρική ζώνη και του test χ^2

Παράλληλα η τιμή του τεστ χ^2_d μπορεί να προκύψει αυτόματα για την παραπάνω δειγματοληψία. Συγκρίνοντας την τιμή του τεστ χ^2 για δεδομένο επίπεδο σημαντικότητας ($\alpha = 0.05$) και βαθμούς ελευθερίας $\nu = k - 1$ στη περίπτωση της στρωματοποιημένης τυχαίας δειγματοληψίας έχουμε $\nu = 4$. Επομένως το $\chi^2_d = 9.48773$.

Όπου: k = κατηγορία χρήσης γης (5)

Συμπεραίνουμε ότι στα δεδομένα του δείγματος δεν προσαρμόζεται ικανοποιητικά η κατανομή αφού $\chi^2 > \chi^2_d$.

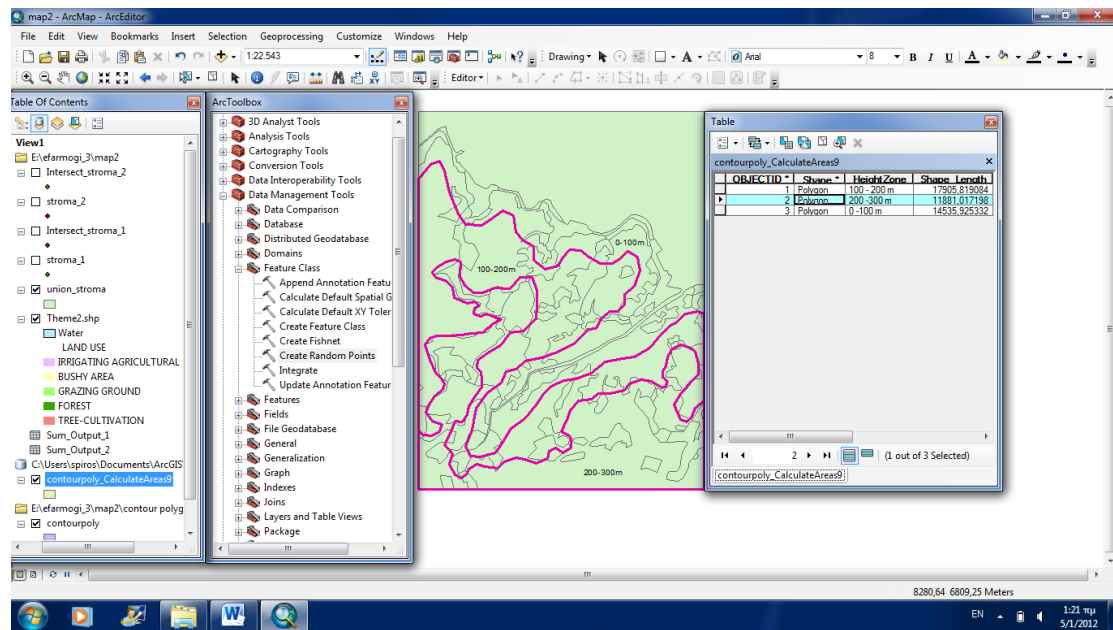
Υψομετρική Ζώνη 3 (3^ο στρώμα)

Ο αριθμός των τυχαίων σημείων που λαμβάνονται είναι $n_3 = 807$ σημεία.

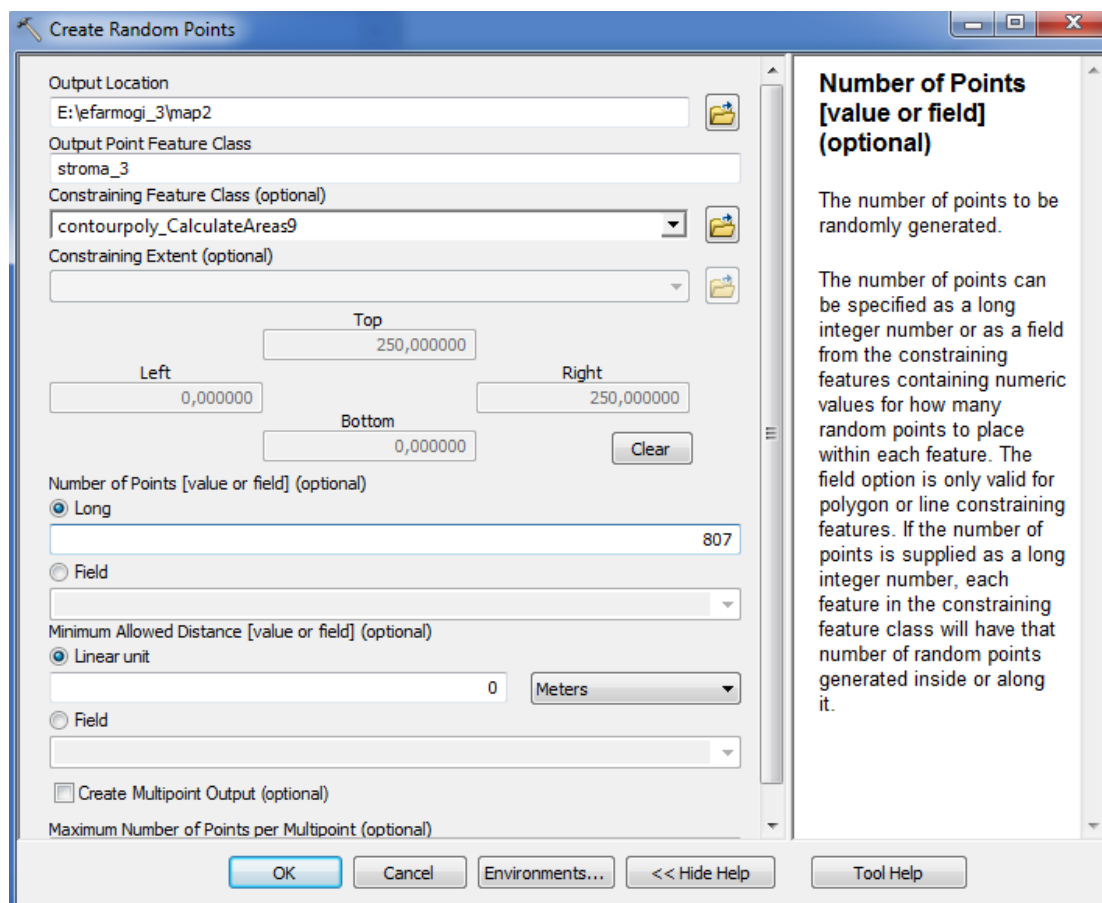
Η δειγματοληψία ξεκινάει με τη λήψη ενός δοκιμαστικού δείγματος 807 παρατηρήσεων. Αυτό γίνεται με την εντολή “ Create Random Points “ από την εργαλειοθήκη Data Management Tools.

[ArcToolBox > Data Management Tools > Feature Class > Create Random Points].

Το παράθυρο διαλόγου εμφανίζεται και συμπληρώνεται όπως παρακάτω:



Σχήμα 4.56: Παράθυρο ArcToolbox Create Random Points Στρωματοποιημένης Τυχαίας Δειγματοληψίας στην 3^η υψομετρική ζώνη



Σχήμα 4.57: Παράμετροι Παραθύρου Create Random Points Στρωματοποιημένης Τυχαίας Δειγματοληψίας στην 3^η υψομετρική ζώνη

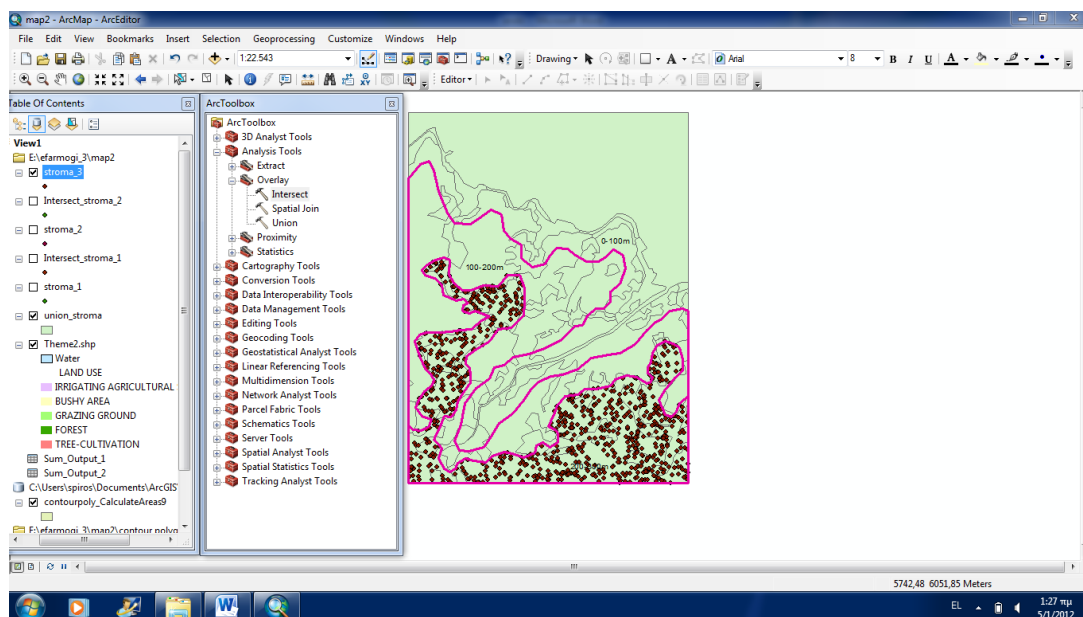
Στο πρώτο πεδίο επιλέγουμε την τοποθεσία στην οποία τα σημεία θα δημιουργηθούν. Στο δεύτερο πεδίο επιλέγουμε το όνομα που θα προκύψει από την γέννηση των σημείων (stroma_3) και στο τρίτο πεδίο περιορισμό κατηγορίας χαρακτηριστικών γνωρισμάτων στο συγκεκριμένο πολύγωνο με τις υψομετρικές ζώνες.

Προχωρώντας στο πεδίο number of points εισάγουμε τον αριθμό των σημείων που θέλουμε να δημιουργήσουμε (807) και στο πεδίο minimum allowed distance ορίζουμε την ελάχιστη απόσταση μεταξύ δύο σημείων (0 meters).

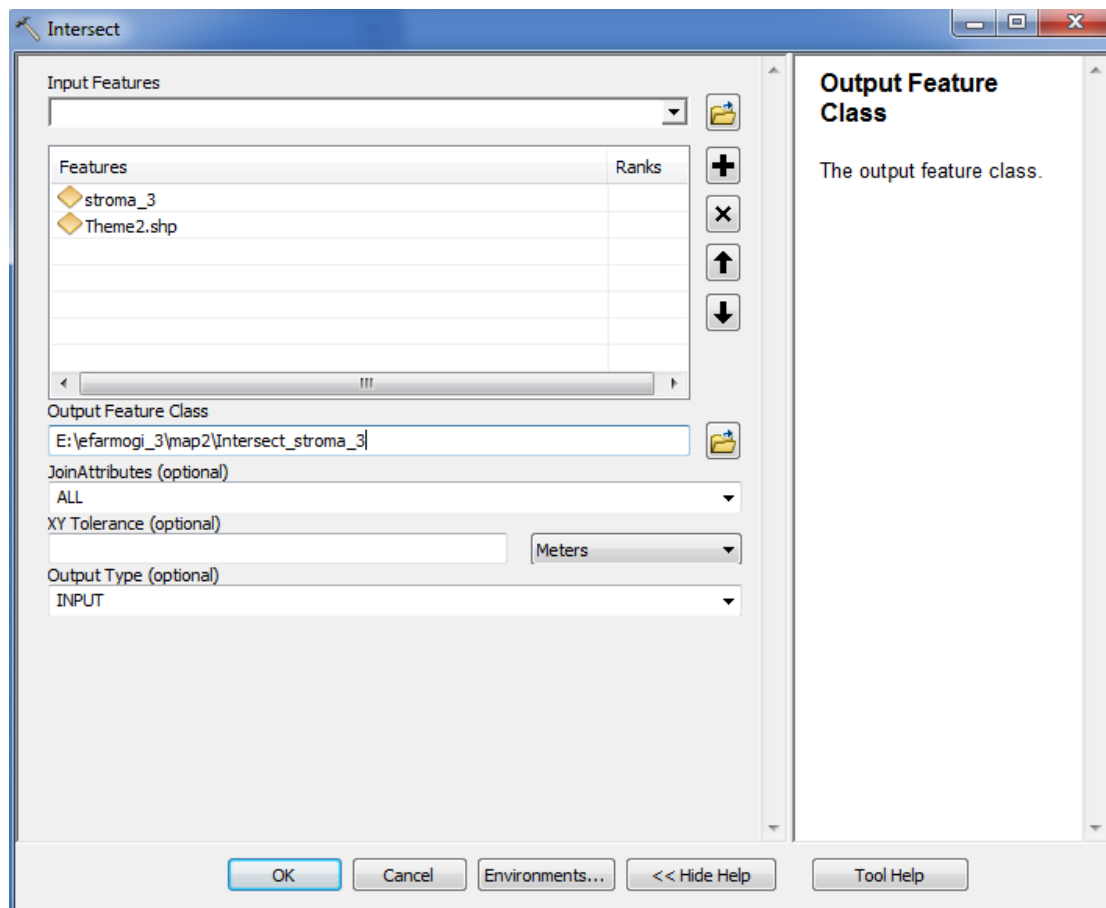


Σχήμα 4.58: Δείγμα 807 παρατηρήσεων stoma_3

Έπειτα με την εντολή Intersect κατατάσσουμε τα σημεία στην χρήση γης στην οποία ανήκουν. Η διαδικασία που ακολουθούμε είναι [ArcToolBox > Analysis Tools > Overlay > Intersect]. Το παράθυρο διαλόγου εμφανίζεται και συμπληρώνεται όπως παρακάτω:



Σχήμα 4.59: Παράθυρο ArcToolbox Intersect Στρωματοποιημένης Τυχαίας Δειγματοληψίας στην 3^η υψομετρική ζώνη



Σχήμα 4.60: Παράμετροι Παραθύρου Intersect Στρωματοποιημένης Τυχαίας Δειγματοληψίας

Στο πεδίο Input Features επιλέγουμε τα στοιχεία στα οποία θέλουμε να εφαρμόσουμε την τομή (theme 2 & stroma_3) και στο πεδίο Output Feature Class επιλέγουμε το όνομα και την τοποθεσία του επιπέδου που θα προκύψει από την τομή (Intersect_stroma_3.shp).

Ο χάρτης απεικονίζεται ως εξής:



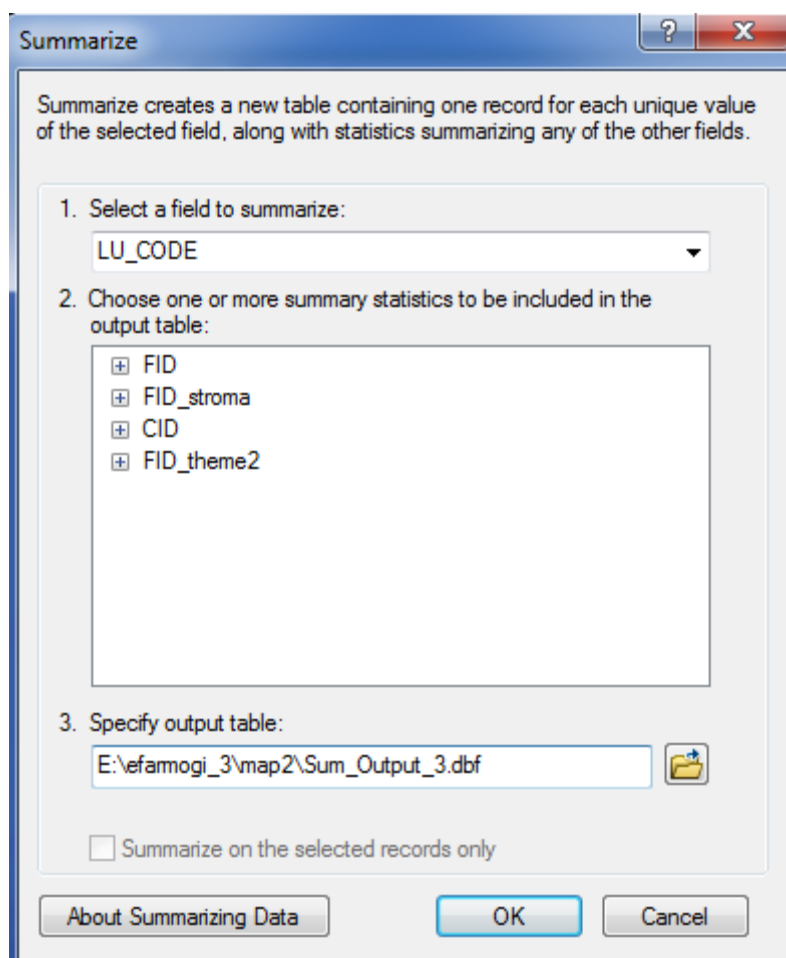
Σχήμα 4.61: Δείγμα 807 παρατηρήσεων Intersect_stroma_3

Περισσότερο αυτό που προκύπτει από την τομή είναι ένας πίνακας τιμών όπου κάθε σημείο αντιστοιχίζεται σε μια συγκεκριμένη κατηγορία χρήσης γης πάνω στην οποία πέφτει.

FID	Shape	FID_stroma	CID	FID_theme2	LU CODE
0	Point	792	2	67	400
1	Point	326	2	67	400
2	Point	126	2	45	400
3	Point	688	2	45	400
4	Point	772	2	45	400
5	Point	332	2	67	400
6	Point	46	2	45	400
7	Point	194	2	45	400
8	Point	103	2	45	400
9	Point	643	2	45	400
10	Point	586	2	45	400
11	Point	27	2	45	400
12	Point	136	2	67	400
13	Point	592	2	45	400
14	Point	462	2	45	400
15	Point	628	2	67	400
16	Point	506	2	75	200
17	Point	7	2	45	400
18	Point	274	2	45	400
19	Point	478	2	45	400
20	Point	450	2	67	400
21	Point	190	2	45	400
22	Point	795	2	45	400
23	Point	541	2	45	400
24	Point	163	2	45	400
25	Point	58	2	45	400
26	Point	102	2	45	400
27	Point	763	2	45	400
28	Point	336	2	45	400
29	Point	554	2	45	400
30	Point	639	2	45	400
31	Point	706	2	45	400
32	Point	679	2	45	400
33	Point	252	2	45	400

Σχήμα 4.62: Πίνακας Κατηγοριοποίησης Σημείων ανά Χρήση Γης Intersect_stroma_3

Στη συνέχεια κάνοντας δεξί κλικ στο πεδίο LU CODE στον πίνακα Intersect_stroma_3 επιλέγουμε την εντολή Summarize όπου μας εμφανίζεται ένα παράθυρο διαλόγου το οποίο το συμπληρώνουμε ως εξής:



Σχήμα 4.63: Παράθυρο Διαλόγου Summarize Στρωματοποιημένης Τυχαίας Δειγματοληψίας

Στο πρώτο πεδίο επιλέγουμε το πεδίο ως προς το οποίο θέλουμε να γίνει η κατηγοριοποίηση (LU CODE), στο δεύτερο πεδίο δεν επιλέγουμε αθροιστικά στατιστικά διότι το sum αθροίζει αύξοντες αριθμούς (stroma_3) στον πίνακα και στο τρίτο δίνουμε στον πίνακα ένα όνομα και μια κατάληξη (dbf αρχείο).

Αξίζει να σημειωθεί ότι στον πίνακα (Sum_Output_3) μας δίνεται το πλήθος των σημείων που πέφτουν σε κάθε χρήση γης που είναι εκείνο που μας ενδιαφέρει περισσότερο και θα χρησιμοποιήσουμε αργότερα στο φύλλο Excel για τον υπολογισμό των ποσοστών των επιμέρους χρήσεων γης ανά κατηγορία.

OID	LU CODE	Count	LU CODE
0	200	167	167
1	300	57	57
2	400	583	583

Σχήμα 4.64: Πίνακας Σημείων ανά Χρήση Γης Στρωματοποιημένης Τυχαίας Δειγματοληψίας Sum_Output_3

Τέλος, εξάγουμε τα αποτελέσματα του πίνακα Sum_Output_3 σε λογιστικό φύλλο Excel όπου κάνουμε τους αντίστοιχους υπολογισμούς για τα ποσοστά χρήσεων γης και το test χ^2 αυτόματα.

Στρωματοποιημένη Τυχαία Σημειακή Δειγματοληψία _3			
Χρήσεις Γης	Αριθμός Σημείων	Ποσοστά Χρήσεων Γης (%)	
100	0	0,00000	
200	167	20,69393	
300	57	7,06320	
400	583	72,24287	
700	0	0,00000	
Σύνολο	807	100,00000	
	Αναμενόμενη Τιμή (%)	Πραγματική Τιμή (%)	Τέστ χ^2
p100	0,00000	3,6	3,6
p200	20,69393	48,9	16,26958
p300	7,06320	5	0,851356
p400	72,24287	41,6	22,57177
p700	0,00000	0,9	0,9
		χ^2 Rtest=	44,19271

Σχήμα 4.65: Φύλλο Excel για τον αυτόματο υπολογισμό ποσοστών χρήσεων γης ανά κατηγορία στην στρωματοποιημένη τυχαία δειγματοληψία στην 3^η υψομετρική ζώνη και του test χ^2

Παράλληλα η τιμή του τεστ χ^2_d μπορεί να προκύψει αυτόματα για την παραπάνω δειγματοληψία. Συγκρίνοντας την τιμή του τεστ χ^2 για δεδομένο επίπεδο σημαντικότητας ($\alpha = 0.05$) και βαθμούς ελευθερίας $\nu = k - 1$ στη περίπτωση της στρωματοποιημένης τυχαιάς δειγματοληψίας έχουμε $\nu = 4$. Επομένως το $\chi^2_d = 9.48773$.

Όπου: k = κατηγορία χρήσης γης (5)

Συμπεραίνουμε ότι στα δεδομένα του δείγματος δεν προσαρμόζεται ικανοποιητικά η κατανομή αφού $\chi^2 > \chi^2_d$.

Στρωματοποιημένη Τυχαιά Σημειακή Δειγματοληψία και στις 3 Ζώνες			
Χρήσεις Γης	Αριθμός Σημείων	Ποσοστά Χρήσεων Γης (%)	
100	85	3,54019	
200	1194	49,72928	
300	109	4,53978	
400	978	40,73303	
700	35	1,45773	
Σύνολο	2401	100,00000	
	Αναμενόμενη Τιμή (%)	Πραγματική Τιμή (%)	Τέστ χ^2
p100	3,54019	3,6	0,000994
p200	49,72928	48,9	0,014063
p300	4,53978	5	0,042361
p400	40,73303	41,6	0,018068
p700	1,45773	0,9	0,34562
		$\chi^2_{Rtest} =$	0,421107

Σχήμα 4.66: Φύλλο Excel για τον αυτόματο υπολογισμό ποσοστών χρήσεων γης ανά κατηγορία στην στρωματοποιημένη τυχαιά δειγματοληψία στις 3 υψομετρικές ζώνες και του test χ^2

Παράλληλα η τιμή του τεστ χ^2_d μπορεί να προκύψει αυτόματα για την παραπάνω δειγματοληψία. Συγκρίνοντας την τιμή του τεστ χ^2 για δεδομένο επίπεδο σημαντικότητας ($\alpha = 0.05$) και βαθμούς ελευθερίας $\nu = k - 1$ στη περίπτωση της στρωματοποιημένης τυχαιάς δειγματοληψίας έχουμε $\nu = 4$. Επομένως το $\chi^2_d = 9.48773$.

Όπου: k = κατηγορία χρήσης γης (5)

Συμπεραίνουμε ότι στα δεδομένα του δείγματος προσαρμόζεται ικανοποιητικά η κατανομή αφού $\chi^2 < \chi^2_d$.

4.9 Συμπεράσματα Άσκησης

Το μέγεθος του δείγματος για τις δύο από τις τρεις μεθόδους (απλή και στρωματοποιημένη) ήταν αρκετά μεγάλο εξαιτίας των υψηλών απαιτήσεων ακρίβειας (μικρό περιθώριο σφάλματος). Η ανάγκη για τη λήψη μεγάλου δείγματος επιτάσσεται από το γεγονός ότι στον χάρτη - υπόβαθρο υπάρχει κατηγορία χρήσης γης που αντιπροσωπεύει λιγότερο από το 1% του συνόλου της έκτασης.

Η δοκιμαστική δειγματοληψία αποδεικνύεται ότι είναι αναξιόπιστη ως μέθοδος. Αυτό οφείλεται στο γεγονός του μικρού μεγέθους δείγματος (30 σημεία) και των υψηλών απαιτήσεων ακρίβειας που έχουμε θέσει (μικρό περιθώριο σφάλματος). Επίσης, διαπιστώνεται εύκολα και από την τιμή του τεστ χ^2 η οποία είναι αρκετά υψηλή (6.53) και απέχει πολύ από το μηδέν.

Η απλή τυχαία δειγματοληψία αποδεικνύεται ότι είναι πιο αξιόπιστη από την δοκιμαστική δειγματοληψία. Αυτό οφείλεται στο γεγονός ότι το μέγεθος του δείγματος είναι αρκετά μεγάλο (2401 σημεία) καθώς και στο ότι οι απαιτήσεις ακρίβειας ικανοποιούνται με αποτέλεσμα η τιμή του τεστ χ^2 να είναι (0.22).

Η φυσική στρωματοποίηση του υποβάθρου με βάση τις υψομετρικές ζώνες του χάρτη, δίνει αποτελέσματα αναφορικά με τα ποσοστά χρήσεων γης σε κάθε στρώμα. Αυτό έχει ως συνέπεια να παράγεται μια επιπλέον πληροφορία και να γίνεται σύγκριση των ποσοστών των κατηγοριών χρήσεων γης σε κάθε υψομετρική ζώνη. Με μια πρώτη ματιά διαπιστώνουμε ότι στη στρωματοποιημένη δειγματοληψία (σε κάθε ζώνη ξεχωριστά) δεν επιτυγχάνεται μεγαλύτερη ακρίβεια σε σχέση με την απλή δειγματοληψία με μικρότερο ή ίδιο μέγεθος δείγματος πρακτικά γιατί υπάρχει μεγάλη διαφοροποίηση μεταξύ των στρωμάτων ενώ θεωρητικά θα έπρεπε. Αυτό διαπιστώνεται εύκολα από τους πίνακες όπου π.χ. στην πρώτη υψομετρική ζώνη το ποσοστό των δενδροκαλλιεργειών είναι 5.70% ενώ στις υπόλοιπες δύο είναι μηδέν καθώς κανένα σημείο σε αυτές τις δύο ζώνες δεν αναπαριστά τη συγκεκριμένη χρήση γης. Επίσης, το ότι υπάρχει μεγάλη διαφοροποίηση μεταξύ των στρωμάτων φαίνεται και στις τιμές του τεστ χ^2 όπου οι τιμές που προκύπτουν είναι αρκετά μεγάλες (63.69 για την πρώτη, 19.70 για τη δεύτερη και 44.19 για την τρίτη). Αν θέλαμε να πετύχουμε καλύτερη στρωματοποίηση θα μπορούσαμε να ορίσουμε κάποιο άλλο κριτήριο (π.χ. κλίσεις εδάφους).

Επιπλέον στη συνολική στρωματοποιημένη δειγματοληψία η τιμή του τεστ χ^2 η οποία προκύπτει είναι (0.42).

Εν κατακλείδι μπορούμε να πούμε ότι η στρωματοποιημένη δειγματοληψία είναι η καλύτερη μέθοδος διότι αυξάνεται η ακρίβεια λόγω του διαχωρισμού σε υποομάδες της περιοχής μελέτης μας. Επίσης αυτός ο διαχωρισμός μας δίνει τη δυνατότητα για περαιτέρω μελέτη και ανάλυση των υποομάδων, ώστε να μπορούμε να διεξάγουμε περισσότερα συμπεράσματα για την περιοχή μας. Τα μόνα μειονεκτήματα που εμφανίζει η τελευταία μέθοδος είναι η πολυπλοκότητα και η απαίτηση περισσότερου χρόνου σε σχέση με την απλή τυχαία δειγματοληψία.

ΚΕΦΑΛΑΙΟ 5: Γενικά Συμπεράσματα

Αντικείμενο της διπλωματικής εργασίας είναι η ανάπτυξη εκπαιδευτικού υλικού για την εφαρμογή των δύο κύριων μεθόδων χωρικής σημειακής δειγματοληψίας (απλή τυχαία, στρωματοποιημένη) σε περιβάλλον γεωγραφικού συστήματος πληροφοριών Arc Gis 10. Η εφαρμογή πραγματοποιήθηκε σε έναν διανυσματικό χάρτη χρήσεων γης με σκοπό την εκτίμηση των επιμέρους ποσοστών χρήσεων γης της περιοχής.

Για την εφαρμογή των δειγματοληψιών χρησιμοποιήθηκε περιβάλλον γεωγραφικών συστημάτων πληροφοριών, έτσι ώστε να έχουμε στη διάθεση μας τα πραγματικά ποσοστά των χρήσεων γης του χάρτη υπολογισμένα από τα εμβαδά και στη συνέχεια να γίνεται αντιπαράθεση με αυτά που προέκυψαν από τις εφαρμογές των δειγματοληπτικών μεθόδων. Ταυτόχρονα η εφαρμογή τους σε Γ.Σ.Π. δίνει τη δυνατότητα στους σπουδαστές να έρθουν σε πρώτη επαφή χρησιμοποιώντας τις λειτουργίες αυτών των συστημάτων που αποτελούν πλέον αναπόσπαστο κομμάτι κάθε γεωγραφικής μελέτης.

Από τη θεωρία γνωρίζουμε ότι βασική παράμετρος κάθε δειγματοληψίας είναι ο υπολογισμός του απαιτούμενου μεγέθους του δείγματος, το οποίο στη περίπτωση εκτίμησης αναλογιών κάποιου χαρακτηριστικού ενός πληθυσμού εξαρτάται από το επίπεδο εμπιστοσύνης που θα θέσουμε και από το περιθώριο λάθους. Συγκεκριμένα στην άσκηση οι τιμές αυτές των δύο παραμέτρων μας έχουν δοθεί και σχετίζονται με την ακρίβεια που θα έχουν τα αποτελέσματά μας. Συνεπώς, για το λόγο αυτό αποφασίστηκε οι δύο αυτοί παράμετροι να καθορίζονται στην αρχή κάθε δειγματοληψίας είτε έχουμε λάβει ένα δοκιμαστικό δείγμα είτε όχι.

Περισσότερο, θα μπορούσαμε να επιλύσουμε την εφαρμογή και με τη χρήση του modelbuilder, ώστε να προκύψει μια ενιαία εφαρμογή εντός του Arc Gis η οποία θα υλοποιούσε αυτόματα τις δειγματοληπτικές μεθόδους, εκτελώντας διαδοχικά τις εντολές που παρέχει εγγενώς το Arc Gis.

Εν κατακλείδι, θα μπορούσαμε τα δύο σημειακά επίπεδα που υπάρχουν στον πίνακα περιεχομένων να τα συγχωνεύσουμε σε ένα το οποίο με τη σειρά να αποτελεί το δείγμα της οριστικής απλής τυχαίας δειγματοληψίας. Η συγχώνευση των δύο σημειακών επιπέδων θα γίνονταν με την εντολή Merge από το ArcToolbox [ArcToolbox> Data Management Tools> General> Merge].

ΠΑΡΑΡΤΗΜΑ

Περιγραφή Άσκησης

Η άσκηση αναφέρεται σε δυο κύριες μεθόδους χωρικής δειγματοληψίας, την Απλή Τυχαία Δειγματοληψία και τη Στρωματοποιημένη Τυχαία Δειγματοληψία. Οι σπουδαστές καλούνται να εφαρμόσουν και να εξετάσουν τις δυο παραπάνω μεθόδους, σε δυο κατάλληλα σχεδιασμένους ψηφιακούς χάρτες, που δημιουργήθηκαν για το σκοπό αυτό.

Οι δυο ψηφιακοί χάρτες αναφέρονται σε μια παραθαλάσσια περιοχή και απεικονίζουν τις παρακάτω χρήσεις γης:

- Αρδευόμενη Γεωργική Γη (Irrigating Agricultural Soil)
- Θαμνώδεις Εκτάσεις (Bushy Area)
- Βοσκότοποι (Grazing Ground)
- Δασικές Εκτάσεις (Forest)
- Δένδρο – Καλλιέργειες

Επιπλέον ο δεύτερος χάρτης της περιοχής, απεικονίζει και τρεις υψομετρικές ζώνες της. Αυτές είναι:

- η υψομετρική ζώνη των 0 – 100 μέτρων.
- η υψομετρική ζώνη των 100 – 200 μέτρων.
- η υψομετρική ζώνη των 200 – 300 μέτρων.

Ζητούμενα:

1. Να εκτιμήσετε τα ποσοστά χρήσεων γης της περιοχής, εφαρμόζοντας μια δοκιμαστική Απλή Τυχαία Δειγματοληψία, με μέγεθος δείγματος 30 σημεία. Να χρησιμοποιηθεί ο πρώτος χάρτης.
2. Να υπολογιστεί το απαιτούμενο μέγεθος δείγματος (απαιτούμενος αριθμός σημείων), ώστε τα ποσοστά χρήσεων γης να προκύψουν με ακρίβεια 2%, για επίπεδο εμπιστοσύνης 95%.

3. Με το νέο μέγεθος δείγματος που υπολογίσατε, εφαρμόστε μια νέα Απλή Τυχαία Δειγματοληψία, ώστε να προκύψουν τα νέα ποσοστά χρήσεων γης με τη ζητούμενη ακρίβεια. Να χρησιμοποιηθεί πάλι ο πρώτος χάρτης.

4. Με το ίδιο μέγεθος δείγματος, εφαρμόστε τη Στρωματοποιημένη Τυχαία Δειγματοληψία στο δεύτερο χάρτη, ώστε να προκύψουν τα ποσοστά χρήσεων γης της κάθε υψομετρικής ζώνης, αλλά και τα συνολικά ποσοστά χρήσεων γης της περιοχής. Υπενθυμίζεται ότι η Στρωματοποιημένη Τυχαία Δειγματοληψία εφαρμόζεται με την εκτέλεση μιας απλής τυχαίας δειγματοληψίας σε κάθε στρώμα, δηλαδή σε κάθε υψομετρική ζώνη. Ο αριθμός των σημείων της κάθε μίας από τρεις απλές τυχαίες δειγματοληψίες, προκύπτει αναλογικά από το εμβαδόν της ζώνης σε σχέση με το συνολικό του δειγματοληπτικού υποβάθρου.

5. Στη συνέχεια, δίνονται τα πραγματικά ποσοστά χρήσεων γης της περιοχής και ζητείται:

α. Η σύγκριση των τριών δειγματοληψιών που εφαρμόστηκαν με τη χρήση του Test χ^2 . Υπενθυμίζεται ότι το Test χ^2 συγκρίνει την κατανομή του πληθυσμού με την κατανομή του δείγματος, οπότε όσο κοννότερα η εξαγόμενη τιμή στο 0, τόσο καλύτερα η κατανομή του δείγματος προσεγγίζει την κατανομή του πληθυσμού.

β. Ποια από τις δυο μεθόδους δειγματοληψίας που εξετάστηκαν δίνει καλύτερα αποτελέσματα; Που υπερερεεί η Στρωματοποιημένη Τυχαία Δειγματοληψία και που υστερεί σε σχέση με την Απλή Τυχαία Δειγματοληψία;

Παρακάτω παρατίθεται και ο πίνακας με τα πραγματικά ποσοστά χρήσεων γης της περιοχής που αναφέραμε παραπάνω

Αρδευόμενη Γεωργική Γη	3.6%
Θαμνώδεις Εκτάσεις	48.9%
Βοσκότοποι	5.0%
Δασικές Εκτάσεις	41.6%
Δένδρο - Καλλιέργειες	0.9%

Συνοπτικές οδηγίες για την Εκπόνηση της Άσκησης

Στον αποθηκευτικό χώρο κάθε σπουδαστή θα πρέπει να υπάρχουν οι φάκελοι της άσκησης, δηλαδή οι Map 1 και Map 2. Για την εκκίνηση του προγράμματος Arc Map: Start > Programs > ArcGis > ArcMap.

Ζητούμενο 1

Στο menu file > open > (πηγαίνετε στη θέση του φακέλου που βρίσκονται τα δεδομένα της άσκησης) > map1 (σε μορφή ArcMap document)]. Έτσι στην οθόνη εμφανίζεται πλέον ο χάρτης. Στη συνέχεια στο menu Selection > Options > επιλέξτε “ Select features that the box or graphic are completely within “ > ok.

Στον πίνακα περιεχομένων των θεματικών επιπέδων (Layers), κάντε δεξί κλικ στο Layer randompoly > Open attribute table και click στο μοναδικό πολύγωνο με FID = 0. Σκοπός η επιλογή του εξωτερικού πολυγώνου του εδάφους του χάρτη (ακτογραμμή και όρια χάρτη). Τώρα στο πολύγωνο αυτό θα έχει τονιστεί το περίγραμμα με κάποιο προεπιλεγμένο χρώμα. Θα πρέπει επίσης στο message bar (κάτω μέρος οθόνης) να γράφεται “ Number of Features Selected: 1 “.

Για τη διενέργεια της δοκιμαστικής τυχαίας δειγματοληψίας με μέγεθος δείγματος 30 σημεία: Αυτό γίνεται με την εντολή “ Create Random Points “ από την εργαλειοθήκη Data Management Tools.

[ArcToolBox > Data Management Tools > Feature Class > Create Random Points].

Εμφανίζεται το παράθυρο αποθήκευσης του νέου θεματικού επιπέδου που θα σχηματιστεί. Επιλέξτε την ονομασία “ simeia_1 “ και αποθηκεύστε το αρχείο στο φάκελο E:\efarmogi_3\map1. Μετά την ολοκλήρωση των παραπάνω στον πίνακα των θεματικών επιπέδων θα έχει εμφανιστεί το νέο επίπεδο “ simeia_1 “ και στο χάρτη τα 30 τυχαία σημεία της δοκιμαστικής δειγματοληψίας.

Στη συνέχεια κάντε τομή του θεματικού επιπέδου των σημείων με αυτών των χρήσεων γης (simeia_1 & theme_1) με την εντολή “ Intersect “ από την εργαλειοθήκη Analysis Tools.

[ArcToolBox > Analysis Tools > Overlay > Intersect]. Αυτό που προκύπτει από την τομή είναι ένας πίνακας τιμών (simeia_1_Intersect1) όπου κάθε σημείο αντιστοιχίζεται σε μια συγκεκριμένη κατηγορία χρήσης γης πάνω στην οποία πέφτει.

Παρακάτω πατήστε δεξί κλικ στο πεδίο LU CODE στον πίνακα simeia_1_Intersect1 επιλέξτε την εντολή Summarize όπου εμφανίζεται ένα παράθυρο διαλόγου το οποίο το συμπληρώνετε και στο τέλος του δίνετε μια ονομασία (Sum_Output_1). Ο πίνακας Sum_Output_1 δίνει το πλήθος και το άθροισμα των σημείων που πέφτουν σε κάθε χρήση γης.

Τέλος, εξάγετε τα αποτελέσματα σε λογιστικό φύλλο Excel όπου κάνετε τους αντίστοιχους υπολογισμούς για τα ποσοστά χρήσεων γης πρώτον και δεύτερον για το test χ^2 (υπολογισμένο με 2 τρόπους) χρησιμοποιώντας τη στατιστική συνάρτηση CHISQ.TEST καθώς και πίνακα όπου εισάγετε το επίπεδο σημαντικότητας – βαθμοί ελευθερίας (χρήσεις γης) όπως δίνεται από την εκφώνηση και προκύπτει η τιμή του test χ^2 αυτόματα.

Ζητούμενο 2

Ο υπολογισμός του απαιτούμενου μεγέθους δείγματος για την απλή τυχαία δειγματοληψία προκύπτει αυτόματα από το λογιστικό φύλλο Excel της δοκιμαστικής τυχαίας δειγματοληψίας με παραμέτρους το επίπεδο εμπιστοσύνης (p), το περιθώριο λάθους (e) τα οποία δίνονται από την εκφώνηση της άσκησης, το διάστημα εμπιστοσύνης z για το συγκεκριμένο επίπεδο εμπιστοσύνης και τα $p\%$, $q\%$ (όπου $p\%$ είναι η καλύτερη εκτίμηση του ποσοστού κατηγορίας χρήσεων γης, για την οποία, είναι το πλησιέστερο στο 50%, σε σχέση με τις καλύτερες εκτιμήσεις των ποσοστών των άλλων χρήσεων και το $q\% = 100 - p\%$).

Η σχέση που χρησιμοποιείται είναι η εξής:

$$n_{\text{new}} = \frac{z^2 * p\% * q\%}{e^2}$$

Ζητούμενο 3

Για τη διενέργεια της απλής τυχαίας δειγματοληψίας με μέγεθος δείγματος 2401 σημεία εφαρμόστε την ίδια διαδικασία. Αυτό γίνεται με την εντολή “ Create Random Points “ από την εργαλειοθήκη Data Management Tools.

[ArcToolBox > Data Management Tools > Feature Class > Create Random Points].

Εμφανίζεται το παράθυρο αποθήκευσης του νέου θεματικού επιπέδου που θα σχηματιστεί. Επιλέξτε την ονομασία “ simeia_2 “ και αποθηκεύστε το αρχείο στο φάκελο E:\efarmogi_3\map1. Μετά την ολοκλήρωση των παραπάνω στον πίνακα των θεματικών επιπέδων θα έχει εμφανιστεί το νέο επίπεδο “ simeia_2 “ και στο χάρτη τα 2401 τυχαία σημεία της απλής τυχαίας δειγματοληψίας.

Στη συνέχεια κάντε τομή του θεματικού επιπέδου των σημείων με αυτών των χρήσεων γης (simeia_2 & theme_1) με την εντολή “ Intersect “ από την εργαλειοθήκη Analysis Tools. [ArcToolBox > Analysis Tools > Overlay > Intersect]. Αυτό που προκύπτει από την τομή είναι ένας πίνακας τιμών (simeia_2_Intersect2) όπου κάθε σημείο αντιστοιχίζεται σε μια συγκεκριμένη κατηγορία χρήσης γης πάνω στην οποία πέφτει.

Παρακάτω πατήστε δεξί κλικ στο πεδίο LU CODE στον πίνακα simeia_2_Intersect2 επιλέξτε την εντολή Summarize όπου εμφανίζεται ένα παράθυρο διαλόγου το οποίο το συμπληρώνετε και στο τέλος του δίνετε μια ονομασία (Sum_Output_2). Ο πίνακας Sum_Output_2 δίνει το πλήθος και το άθροισμα των σημείων που πέφτουν σε κάθε χρήση γης.

Τέλος, εξάγετε τα αποτελέσματα σε λογιστικό φύλλο Excel όπου κάνετε τους αντίστοιχους υπολογισμούς για τα ποσοστά χρήσεων γης πρώτον και δεύτερον για το test χ^2 (υπολογισμένο με 2 τρόπους) χρησιμοποιώντας τη στατιστική συνάρτηση CHISQ.TEST καθώς και πίνακα όπου εισάγετε το επίπεδο σημαντικότητας – βαθμοί ελευθερίας (χρήσεις γης) όπως δίνεται από την εκφώνηση και προκύπτει η τιμή του test χ^2 αυτόματα.

Ζητούμενο 4

Η στρωματοποιημένη τυχαία δειγματοληψία πραγματοποιείται στο δεύτερο χάρτη. Ο χάρτης αποτελείται από τρεις υψομετρικές ζώνες.

Στο menu file > open > (πηγαίνετε στη θέση του φακέλου που βρίσκονται τα δεδομένα της άσκησης) > map2 (σε μορφή ArcMap document)]. Έτσι στην οθόνη εμφανίζεται πλέον ο χάρτης. Στη συνέχεια στο menu Selection > Options > επιλέξτε “ Select features that the box or graphic are completely within “ > ok.

Στον πίνακα περιεχομένων των θεματικών επιπέδων (Layers), κάντε δεξί κλικ στο Layer contourpoly > Open attribute table και click στις υψομετρικές ζώνες με FID = 0 - 2.

Αρχικά έχοντας ως γνωστό τον αριθμό των σημείων της απλής τυχαίας δειγματοληψίας η διαδικασία που ακολουθείται για την παραγωγή του αριθμού των σημείων σε κάθε υψομετρική ζώνη είναι εξής:

Για τον υπολογισμό του εμβαδού της κάθε ζώνης ξεχωριστά χρησιμοποιήστε την εντολή “ Calculate Areas “ από την εργαλειοθήκη Spatial Statistics Tools.

[ArcToolBox > Spatial Statistics Tools > Utilities > Calculate Areas]. Στη συνέχεια εξάγετε τα αποτελέσματα του πίνακα contourpoly_CalculateAreas9 (πεδίο F Area) σε λογιστικό φύλλο Excel όπου με την συνάρτηση Sum παίρνετε το συνολικό εμβαδόν. Έτσι από τον τύπο:

$n = n_{new} * E / \Sigma E$ προκύπτει στο φύλλο Excel αυτόματα ο αριθμός των σημείων για κάθε υψομετρική ζώνη.

Παρακάτω δώστε υψόμετρο στις χρήσεις γης. Για να γίνει αυτό χρησιμοποιήστε την εντολή “ Union “ από την εργαλειοθήκη Analysis Tools.

[ArcToolbox > Analysis Tools > Overlay > Union]. Τα επίπεδα που ενώνονται είναι τα υψομετρικά πολύγωνα (contourpoly_CalculateAreas9) και οι χρήσεις γης (theme 2). Το αποτέλεσμα που προκύπτει από την ένωση είναι το επίπεδο (union_stroma) και ο πίνακας με την ίδια ονομασία. Με προσεκτική παρατήρηση διαπιστώνεται ότι κάποια πολύγωνα έχουν αρνητικό υψόμετρο. Έτσι χρησιμοποιήστε την εντολή “ Delete Rows “ από την εργαλειοθήκη Data Management Tools.

[ArcToolbox > Data Management Tools > Table > Delete Rows] με την οποία διαγράψτε τις γραμμές με τα πολύγωνα που έχουν αρνητικό υψόμετρο.

Το επίπεδο όπου εφαρμόζεται η εντολή είναι το union_stroma. Ύστερα από όλη αυτή την διαδικασία προκύπτει το νέο επίπεδο union_stroma και ο πίνακας με τα πολύγωνα χωρίς αρνητικό υψόμετρο.

Περισσότερο έχοντας υπολογίσει τον αριθμό των σημείων που αντιστοιχούν σε κάθε υψομετρική ζώνη εφαρμόστε τρεις απλές τυχαίες δειγματοληψίες.

Για τη διενέργεια της απλής τυχαίας δειγματοληψίας στην πρώτη υψομετρική ζώνη με μέγεθος δείγματος 614 σημεία: Αυτό γίνεται με την εντολή “ Create Random Points “ από την εργαλειοθήκη Data Management Tools.

[ArcToolBox > Data Management Tools > Feature Class > Create Random Points].

Εμφανίζεται το παράθυρο αποθήκευσης του νέου θεματικού επιπέδου που θα σχηματιστεί. Επιλέξτε την ονομασία “ stroma_1 “ και αποθηκεύστε το αρχείο στο φάκελο E:\efarmogi_3\map2. Μετά την ολοκλήρωση των παραπάνω στον πίνακα των θεματικών επιπέδων θα έχει εμφανιστεί το νέο επίπεδο “ stroma_1 “ και στο χάρτη τα 614 τυχαία σημεία της απλής τυχαίας δειγματοληψίας.

Στη συνέχεια κάντε τομή του θεματικού επιπέδου των σημείων με αυτών των χρήσεων γης (stroma_1 & theme_2) με την εντολή “ Intersect “ από την εργαλειοθήκη Analysis Tools. [ArcToolBox > Analysis Tools > Overlay > Intersect]. Αυτό που προκύπτει από την τομή είναι ένας πίνακας τιμών (Intersect_stroma_1) όπου κάθε σημείο αντιστοιχίζεται σε μια συγκεκριμένη κατηγορία χρήσης γης πάνω στην οποία πέφτει.

Παρακάτω πατήστε δεξί κλικ στο πεδίο LU CODE στον πίνακα Intersect_stroma_1 επιλέξτε την εντολή Summarize όπου εμφανίζεται ένα παράθυρο διαλόγου το οποίο το συμπληρώνετε και στο τέλος του δίνετε μια ονομασία (Sum_Output_1). Ο πίνακας Sum_Output_1 μας δίνει το πλήθος των σημείων που πέφτουν σε κάθε χρήση γης.

Τέλος, εξάγετε τα αποτελέσματα σε λογιστικό φύλλο Excel όπου κάνετε τους αντίστοιχους υπολογισμούς για τα ποσοστά χρήσεων γης πρώτον και δεύτερον για το test χ^2 (υπολογισμένο με 2 τρόπους) χρησιμοποιώντας τη στατιστική συνάρτηση CHISQ.TEST καθώς και πίνακα όπου εισάγετε το επίπεδο σημαντικότητας – βαθμοί ελευθερίας (χρήσεις γης) όπως δίνετε από την εκφώνηση και προκύπτει η τιμή του test χ^2 αυτόματα.

Εν κατακλείδι παρόμοια διαδικασία ακολουθήστε και στις άλλες δυο υψομετρικές ζώνες.

Σύντομη Τεχνική Έκθεση

Ζητούμενο 1

Στον πρώτο χάρτη δημιουργήσαμε ένα τυχαίο παραλληλόγραμμο επιλογής το οποίο πρέπει να τέμνει τουλάχιστον δύο πολύγωνα χρήσεων γης (το νερό δεν συμπεριλαμβάνεται στις χρήσεις γης). Έπειτα λειτουργώντας στο πρόγραμμα Arc Map επιλέξαμε το εξωτερικό πολύγωνο του εδάφους του MAP1. Έτσι λάβαμε με τη βοήθεια του προγράμματος 30 σημεία τα οποία απεικονίζονται στον παρακάτω πίνακα:

ΧΡΗΣΕΙΣ ΓΗΣ	ΠΛΗΘΟΣ ΣΗΜΕΙΩΝ	ΠΟΣΟΣΤΟ (%)
Αρδευόμενη Γεωργική Γη	1	3,33
Θαμνώδεις Εκτάσεις	15	50,00
Βοσκότοποι	3	10,00
Δασικές Εκτάσεις	11	36,66
Δενδροκαλλιέργειες	0	0,00

Όπου $\chi^2 = 6.53$

Η τιμή χ^2 πρέπει να τείνει στο μηδέν, πράγμα που δεν ισχύει στη συγκεκριμένη περίπτωση. Άρα συμπεραίνουμε πως το δείγμα μας δεν είναι αξιόπιστο, γεγονός που διακρίνεται αν δούμε τα πραγματικά στοιχεία που μας δίνονται.

Ζητούμενο 2

Πληκτρολογώντας σε λογιστικό φύλλο Excel τις παραμέτρους που μας δίνονται (ακρίβεια 2%, επίπεδο εμπιστοσύνης 95%) το πλήθος των σημείων που πρέπει να πάρουμε για να θεωρηθεί η δειγματοληψία αξιόπιστη είναι: **n = 2401**

Ζητούμενο 3

Εφαρμόζοντας σε αυτό το ερώτημα την ίδια μέθοδο με το 1^ο ερώτημα και με νέο δείγμα 2401 σημείων λάβαμε τα παρακάτω αποτελέσματα:

ΧΡΗΣΕΙΣ ΓΗΣ	ΠΛΗΘΟΣ ΣΗΜΕΙΩΝ	ΠΟΣΟΣΤΟ (%)
Αρδευόμενη Γεωργική Γη	78	3,25
Θαμνώδεις Εκτάσεις	1181	49,19
Βοσκότοποι	132	5,50
Δασικές Εκτάσεις	997	41,52
Δενδροκαλλιέργειες	13	0,54

Όπου $\chi^2 = 0.23$

Τώρα το χ^2 είναι πολύ κοντά στο μηδέν γεγονός που μαρτυρά πως αυτή η διαδικασία είναι πού πιο αξιόπιστη, κάτι που είναι αναμενόμενο εφόσον το δείγμα είναι κατά πολύ μεγαλύτερο από ότι ήταν στην πρώτη διαδικασία και πλησιάζει περισσότερο στα πραγματικά ποσοστά κάλυψης.

Ζητούμενο 4

Στο ερώτημα αυτό εργαζόμαστε στον MAP2 και εφαρμόζουμε τη Στρωματοποιημένη Τυχαία Δειγματοληψία. Με αυτή τη μέθοδο θα πάρουμε ποσοστά χρήσεων γης για όλη την περιοχή, αλλά και για κάθε υψομετρική ζώνη ξεχωριστά. Ήδη από το δεύτερο ερώτημα είχαμε την πληροφορία από το λογιστικό φύλλο Excel για το σύνολο των σημείων το οποίο σε συνδυασμό με τον υπολογισμό του εμβαδού κάθε ζώνης ξεχωριστά και του συνολικού εμβαδού από το πρόγραμμα Arc Map και στη συνέχεια περνώντας αυτά τα αποτελέσματα σε ένα λογιστικό φύλλο Excel και εφαρμόζοντας ένα συγκεκριμένο τύπο προέκυψε ο αριθμός των σημείων σε κάθε υψομετρική ζώνη ξεχωριστά. Έτσι έχουμε αντίστοιχα $n_1 = 614$, $n_2 = 980$ και $n_3 = 807$. Συνεπώς εργαζόμαστε σύμφωνα με τα γνωστά για κάθε ζώνη και παίρνουμε τα αποτελέσματα τα οποία αναγράφονται στους παρακάτω πίνακες:

$n_1 = 614$

ΧΡΗΣΕΙΣ ΓΗΣ	ΠΛΗΘΟΣ ΣΗΜΕΙΩΝ	ΠΟΣΟΣΤΟ (%)
Αρδευόμενη Γεωργική Γη	83	13,52
Θαμνώδεις Εκτάσεις	340	55,37
Βοσκότοποι	10	1,63
Δασικές Εκτάσεις	146	23,78
Δενδροκαλλιέργειες	35	5,70

$n_2 = 980$

ΧΡΗΣΕΙΣ ΓΗΣ	ΠΛΗΘΟΣ ΣΗΜΕΙΩΝ	ΠΟΣΟΣΤΟ (%)
Αρδευόμενη Γεωργική Γη	2	0,20
Θαμνώδεις Εκτάσεις	687	70,10
Βοσκότοποι	42	4,29
Δασικές Εκτάσεις	249	25,41
Δενδροκαλλιέργειες	0	0,00

$n_3 = 807$

ΧΡΗΣΕΙΣ ΓΗΣ	ΠΛΗΘΟΣ ΣΗΜΕΙΩΝ	ΠΟΣΟΣΤΟ (%)
Αρδευόμενη Γεωργική Γη	0	0,00
Θαμνώδεις Εκτάσεις	167	20,70
Βοσκότοποι	57	7,07
Δασικές Εκτάσεις	583	72,25
Δενδροκαλλιέργειες	0	0,00

Για όλες τις υψομετρικές ζώνες προέκυψαν τα παρακάτω αποτελέσματα:

ΧΡΗΣΕΙΣ ΓΗΣ	ΠΛΗΘΟΣ ΣΗΜΕΙΩΝ	ΠΟΣΟΣΤΟ (%)
Αρδευόμενη Γεωργική Γη	85	3,54
Θαμνώδεις Εκτάσεις	1194	49,73
Βοσκότοποι	109	4,54
Δασικές Εκτάσεις	978	40,73
Δενδροκαλλιέργειες	35	1,46

Όπου $\chi^2 = 0.42$

A) Χρησιμοποιώντας ως μέτρο σύγκρισης των μεθόδων το χ^2 προκύπτει ότι πιο αξιόπιστη δειγματοληψία είναι η τυχαία δειγματοληψία με πολύ μεγάλο μέγεθος δείγματος.

B) Ωστόσο, η τυχαία δειγματοληψία δεν είναι πάντα τόσο αξιόπιστη και έτσι προτιμάται η Στρωματοποιημένη Τυχαία Δειγματοληψία. Αυτό είναι προφανές, διότι αυξάνουμε την ακρίβεια, λόγω του διαχωρισμού σε υποομάδες της περιοχής μελέτης μας. Ακόμα αυτός ο διαχωρισμός μας δίνει τη δυνατότητα για περαιτέρω μελέτη και ανάλυση των υποομάδων, ώστε να μπορούμε να διεξάγουμε περισσότερα συμπεράσματα για την περιοχή μας. Το μειονέκτημα αυτής της μεθόδου είναι η πολυπλοκότητά της και η απαίτηση περισσότερου χρόνου από ότι η Απλή Τυχαία Δειγματοληψία. Η τελευταία είναι ευκολότερη και πιο γρήγορη, όμως υστερεί σε αξιοπιστία για τους λόγους που προαναφέραμε.

Το Λογισμικό Πακέτο ArcGis 10

Για την ανάπτυξη και την καλύτερη παρουσίαση της άσκησης χρησιμοποιήθηκε το λογισμικό πακέτο ArcGis 10 της εταιρείας ESRI. Το πακέτο αυτό είναι το πιο σύγχρονο λογισμικό στο χώρο των Γεωγραφικών Συστημάτων Πληροφοριών και αποτελεί ένα κλιμακωτό λογισμικό σύστημα για δημιουργία, διαχείριση, ενοποίηση, ανάλυση και διασπορά γεωγραφικών δεδομένων για κάθε οργανισμό, από ένα άτομο έως ένα παγκόσμιο κατανεμημένο δίκτυο ανθρώπων. Η έκδοση αυτή συνθέτει ένα κατανοητό και ολοκληρωμένο σύστημα, σχεδιασμένο για να ικανοποιεί τις ανάγκες των περισσότερων χρηστών των Γεωγραφικών Συστημάτων Πληροφοριών.

Το καινούριο λογισμικό παρουσιάζει κάποια βασικά χαρακτηριστικά:

- Είναι εύκολο στη χρήση και προσφέρει δυνατότητες γενικής χαρτογραφίας, ανάλυσης και διαχείρισης δεδομένων που είναι χρήσιμα και έξω από το ίδιο το σύστημα, ενώ επιπλέον τα δεδομένα είναι εύκολα προσαρμόσιμα.
- Είναι εξαιρετικά λειτουργικό αφού έχει ενσωματωμένες δυνατότητες επέμβασης στα γεωγραφικά δεδομένα, προχωρημένης χαρτογραφίας, βελτιωμένης διαχείρισης δεδομένων και σύνθετης χωρικής ανάλυσης.
- Είναι κλιμακωτό αφού όλα τα επιμέρους επίπεδα του λογισμικού χρησιμοποιούν τον ίδιο πυρήνα εφαρμογών, τις ίδιες λειτουργικές αρχές και το ίδιο user interface.
- Έχει δυνατότητες Web αφού οι χρήστες έχουν τη δυνατότητα να χρησιμοποιούν το διαδίκτυο για να αποκτούν άμεση πρόσβαση σε δεδομένα και υπηρεσίες.
- Είναι φιλικό προς το χρήστη αφού ο χρήστης έχει τη δυνατότητα παραγωγής δικών του επεκτάσεων, εφαρμογών, εργαλείων (scripts), σε διάφορες γλώσσες προγραμματισμού (π.χ. Visual Basic, Python κ.τ.λ.) ώστε να πετυχαίνει τη βέλτιστη απόδοση στην εργασία του.

Ο πυρήνας του λογισμικού είναι τα προγράμματα ArcMap, ArcCatalog και ArcToolbox. Το ArcMap όπως και το όνομα υποδηλώνει, αφορά εφαρμογές που σχετίζονται με χάρτες (Map). Πιο συγκεκριμένα το ArcMap δίνει τη δυνατότητα να δημιουργηθούν χάρτες από επίπεδα χωρικής πληροφορίας, να αναλυθούν χωρικές σχέσεις και να επιλεγούν μέσα από αναζητήσεις χωρικά και μη χωρικά στοιχεία. Επίσης, μπορούν να σχεδιαστούν και να δημιουργηθούν διαφορετικές απεικονίσεις ενός χάρτη, αλλάζοντας χρώματα και συμβολισμούς.

Το ArcCatalog χρησιμοποιείται για τον εντοπισμό, την οργάνωση και τη διαχείριση αρχείων χωρικών δεδομένων. Ουσιαστικά εκτελεί παρόμοιες λειτουργίες με έναν τυπικό explorer, προσανατολισμένο βέβαια στα γεωγραφικά δεδομένα.

Το ArcToolbox διαθέτει εργαλεία μετατροπής αρχείων γεωγραφικών δεδομένων και για Geoprocessing. Geoprocessing (Γεωεπεξεργασία) είναι ο μηχανισμός μέσω του οποίου επιτυγχάνεται η μετάβαση από στοιχεία σε πληροφορία (θεμελιώδης έννοια σε οποιαδήποτε

γεωγραφική μελέτη) με τη χρήση του ArcGis. Αποτελεί μια σειρά από εργαλεία για τη διαχείριση και ανάλυση των χωρικών δεδομένων. Τα δύο βασικά εργαλεία του Geoprocessing είναι τα εργαλεία του ArcToolbox και οι αναλυτικές διαδικασίες του ModelBuilder.

Το ArcToolbox στη πλήρη έκδοσή του (έκδοση ArcInfo) περιλαμβάνει περίπου 200 εργαλεία για γεωεπεξεργασία, μετατροπή δεδομένων, διαχείριση χαρτογραφικών φύλλων, αλληλεπίθεση επιπέδων, χαρτογραφικών προβολών κ.α..

Οι λειτουργίες και οι δυνατότητες του κάθε προγράμματος αυξάνονται καθώς ανεβαίνουμε από το χαμηλότερο επίπεδο στο υψηλότερο επίπεδο του λογισμικού πακέτου, δηλαδή από το ArcView, στο ArcEditor και τέλος στο ArcInfo.

Το ArcMap συνιστά τη κεντρική εφαρμογή του λογισμικού πακέτου ArcGis 10. Χρησιμοποιείται για όλες τις διεργασίες που βασίζονται σε χάρτες συμπεριλαμβανομένων της χαρτογραφίας, της χαρτογραφικής ανάλυσης και της επέμβασης στα χαρτογραφικά στοιχεία (editing). Οι χάρτες έχουν μια εικόνα εξόδου (page layout) που εμφανίζει το χάρτη με τα επιθυμητά θεματικά επίπεδα, το υπόμνημα, την κλίμακα, το σύμβολο του Βορρά και άλλα δευτερεύοντα στοιχεία. Το ArcMap προσφέρει και τη γεωγραφική όψη του χάρτη (geographic data view), όπου εκεί μπορεί να εκτελεστεί μια πληθώρα εφαρμογών των Γεωγραφικών Συστημάτων Πληροφοριών.

Αναλυτικά στο ArcMap ο χρήστης έχει τις εξής δυνατότητες:

- Να δημιουργήσει και να αλληλεπιδράσει με χάρτες.
- Να δει, να συνθέσει και να αναλύσει γεωγραφικά δεδομένα.
- Να θέσει ερωτήματα στα χωρικά δεδομένα του χάρτη, να βρει και να καταλάβει τις μεταξύ τους χωρικές σχέσεις.
- Να συμβολίσει τις γεωγραφικές πληροφορίες με μια μεγάλη ποικιλία τρόπων και συμβολισμών.
- Να δημιουργήσει γραφήματα, εκθέσεις και αναφορές για να επικοινωνήσει με άλλους χρήστες και να κάνει ακόμη πιο κατανοητά τα γεωγραφικά δεδομένα.
- Τέλος, να δημιουργήσει τελικές εικόνες εξόδου χαρτών, όπου στην οθόνη του υπολογιστή να παρουσιάζεται αυτό ακριβώς που θα εκτυπωθεί.

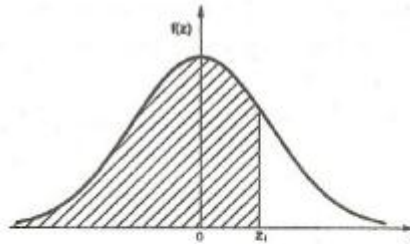
Το Arc Catalog όπως αναφέρθηκε πιο πριν είναι εφαρμογή διαχείρισης των γεωγραφικών δεδομένων. Μοιάζει σε εμφάνιση και λογική με τον Explorer των Windows. Μέσω του ArcCatalog ο χρήστης έχει τη δυνατότητα:

- Να διαχειριστεί αρχεία γεωγραφικών δεδομένων (αντιγραφή, μετακίνηση, μετονομασία, διαγραφή) και φακέλους.
- Να κάνει επισκόπηση (preview) των αρχείων γεωγραφικών δεδομένων.

- Να κάνει επισκόπηση ή τροποποίηση (edit) των μεταδεδομένων (μεταδεδομένα = δεδομένα για τα δεδομένα) που συνοδεύουν τα αρχεία γεωγραφικών δεδομένων ή και να τα δημιουργήσει εξ αρχής.
- Να εισάγει απευθείας δεδομένα στο ArcMap με drag και drop.
- Να κάνει αναζήτηση δεδομένων (search) τόσο στον δίσκο του υπολογιστή ή ακόμα και στο διαδίκτυο.
- Τέλος, να συνδεθεί με προηγμένες υπηρεσίες παροχής γεωγραφικών δεδομένων μέσω διαδικτύου.

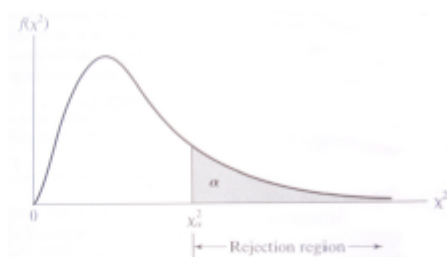
Πίνακας Κανονικής Κατανομής

$$P(-\infty < z < z_i) = \int_{-\infty}^{z_i} f(z) dz = F(z_i)$$



z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9990	0.9993	0.9995	0.9997	0.9998	0.9998	0.9999	0.9999	1.0000

Πίνακας Κρίσιμων Τιμών χ^2



Βαθμοί ελευθερίας	$\chi^2_{0.995}$	$\chi^2_{0.950}$	$\chi^2_{0.100}$	$\chi^2_{0.050}$	$\chi^2_{0.025}$
1	0.0000393	0.0039321	2.70554	3.84146	5.02389
2	0.0100251	0.102587	4.60517	5.99147	7.37776
3	0.0717212	0.351846	6.25139	7.81473	9.34840
4	0.206990	0.710721	7.77944	9.48773	11.1433
5	0.411740	1.145476	9.23635	11.0705	12.8325

Πίνακας Τυχαίων Αριθμών

08210	32973	08003	54512	64863	78634	36344	73293	77660	90199
56822	81546	04735	15228	37475	79057	48749	89898	95554	39734
89182	97110	37211	11688	69132	97318	63419	82573	72506	55631
72674	35938	58483	08612	23129	95675	47142	74290	97894	17531
90431	35286	24567	05594	75288	52097	07687	25548	36274	59709
26899	07988	28526	97601	03588	46466	55143	65558	25847	59878
18345	22317	08035	94062	39380	80045	80231	29109	50908	15304
59231	50655	21753	70621	91045	11318	73098	42457	45905	58889
25446	92080	34212	70395	64493	86192	96633	11311	62422	32794
64825	37691	90465	51579	83918	22217	55262	47065	36168	65350
56902	23614	90057	12880	89522	81453	41294	14603	39279	54730
94781	64586	97431	01310	04978	25633	61253	48959	53841	12864
55140	32776	03346	61125	91434	84200	73664	45914	90551	83073
71284	19521	04812	76028	20044	78244	58015	05867	53847	04736
12938	37575	74532	60769	26928	29666	84318	21283	17537	31834
74888	92697	43016	56171	27506	72769	60939	50508	37332	77234
92592	82165	90227	88560	86908	32283	83958	92590	97910	40983
13741	37073	04809	23138	34852	45986	73823	08534	45867	18800
84984	54147	59332	70137	24535	17957	21012	51279	67206	63087
60528	16871	20195	14420	27684	35554	90414	80832	86089	61328
74340	27455	43144	85543	85639	91044	58951	85251	69761	57118
22071	81837	14952	50079	08651	56337	76769	83059	23254	73901
46002	37283	12700	60802	93248	12585	88835	97516	69585	78330
08389	69818	88126	63220	09217	89809	41827	63143	94194	69001
80520	76260	83873	30245	69153	36222	71188	50184	69731	00382
86956	16358	66811	24846	57071	40649	92444	63235	89322	99933
53136	21244	18794	69981	67197	20656	01157	43212	71717	36254
36225	08162	77121	36757	48280	01627	42182	90286	33607	39507
94504	39905	37479	59621	90903	25258	02601	40884	18925	19143
62903	81216	82151	05623	29557	17355	75959	26697	92117	50471
49684	94165	26723	59644	97156	95754	14613	62639	19567	24652
74368	66450	66013	32765	29673	05637	18535	59269	40732	90216
65947	65189	25064	99907	41696	83361	05711	44600	31021	34714
48562	64985	78945	43273	93601	41398	47898	87513	71023	77141
56879	44941	45205	84341	92222	50193	34172	63957	65014	93068
56352	13144	30800	60403	84654	28738	48297	30614	71784	23440
02272	83016	25229	43692	36151	96505	73151	82126	30327	15817
33209	67546	92428	91522	75736	59065	00668	43097	75282	81904
04042	27494	05204	86656	04755	03880	91082	33651	63806	19334
06255	03873	64275	81962	00650	39269	73457	07644	21562	67255
02637	88632	65996	41115	33863	95116	66522	55347	30163	50594
66788	56790	14606	63123	05831	14117	04288	63610	48924	01827
75185	54242	23276	86228	68663	31051	10832	72330	70330	91769
97678	41209	34692	03790	01489	45480	18809	16099	52339	41146
98264	76970	35868	18540	25973	28110	52129	77706	18505	89583
85483	81260	62113	64994	60821	40518	94273	64432	48779	05310
11364	56369	72114	94681	74135	49670	21072	78489	39630	89270
52663	97730	76199	52766	34479	03055	92954	43490	39790	46437
86276	71122	87556	29504	32565	48447	08719	83014	23615	14317
92295	34604	61004	19707	97387	90740	48389	57337	76910	82215

Πηγή: «A Million Random Digits with 100,000 Normal Deviates» by the RAND Corp.

Πίνακες Αποτελεσμάτων Δειγματοληψιών

Παρακάτω παραβάλλουμε τους πίνακες με τα αποτελέσματα εκτιμήσεων ποσοστών χρήσεων γης του χάρτη για τις τρεις μεθόδους (δοκιμαστική, απλή τυχαία και στρωματοποιημένη σημειακή δειγματοληψία). Τα πραγματικά ποσοστά χρήσεων γης του χάρτη είναι:

Χρήσεις Γης	Ποσοστό Χρήσεων Γης
Αρδευόμενη Γεωργική Γη	3.6 %
Θαμνώδεις Εκτάσεις	48.9 %
Βοσκότοποι	5.0 %
Δασικές Εκτάσεις	41.6 %
Δένδρο - Καλλιέργειες	0.9 %

Δοκιμαστική Τυχαία Σημειακή Δειγματοληψία

Μέγεθος Δείγματος: 30 σημεία

Χρήσεις Γης	Αριθμός Σημείων	Ποσοστό Χρήσεων Γης
Αρδευόμενη Γεωργική Γη	1	3.3 %
Θαμνώδεις Εκτάσεις	15	50.0 %
Βοσκότοποι	3	10.0 %
Δασικές Εκτάσεις	11	36.7 %
Δένδρο - Καλλιέργειες	0	0.0 %

Απλή Τυχαία Σημειακή Δειγματοληψία

Μέγεθος Δείγματος: 2401 σημεία

Χρήσεις Γης	Αριθμός Σημείων	Ποσοστό Χρήσεων Γης
Αρδευόμενη Γεωργική Γη	78	3.2 %
Θαμνώδεις Εκτάσεις	1181	49.2 %
Βοσκότοποι	132	5.5 %
Δασικές Εκτάσεις	997	41.5%
Δένδρο - Καλλιέργειες	13	0.5 %

Στρωματοποιημένη Τυχαία Σημειακή Δειγματοληψία

Μέγεθος Δείγματος: 2401 σημεία

1^ο στρώμα (614 σημεία)

Χρήσεις Γης	Αριθμός Σημείων	Ποσοστό Χρήσεων Γης
Αρδευόμενη Γεωργική Γη	83	13.5 %
Θαμνώδεις Εκτάσεις	340	55.4 %
Βοσκότοποι	10	1.6 %
Δασικές Εκτάσεις	146	23.8 %
Δένδρο - Καλλιέργειες	35	5.7 %

2^ο στρώμα (980 σημεία)

Χρήσεις Γης	Αριθμός Σημείων	Ποσοστό Χρήσεων Γης
Αρδευόμενη Γεωργική Γη	2	0.2 %
Θαμνώδεις Εκτάσεις	687	70.1 %
Βοσκότοποι	42	4.3%
Δασικές Εκτάσεις	249	25.4%
Δένδρο - Καλλιέργειες	0	0.0 %

3^ο στρώμα (807 σημεία)

Χρήσεις Γης	Αριθμός Σημείων	Ποσοστό Χρήσεων Γης
Αρδευόμενη Γεωργική Γη	0	0.0 %
Θαμνώδεις Εκτάσεις	167	20.7 %
Βοσκότοποι	57	7.0 %
Δασικές Εκτάσεις	583	72.3 %
Δένδρο - Καλλιέργειες	0	0.0 %

Αρτικόλεξο

Ελληνική ορολογία

1. Γ.Σ.Π.: Γεωγραφικό Σύστημα Πληροφοριών
2. Σ.Α.Τ.Μ.: Σχολή Αγρονόμων και Τοπογράφων Μηχανικών
3. Σ.Γ.Π.: Σύστημα Γεωγραφικής Πληροφορίας
4. Σ.Π.Π.: Συνάρτηση Πυκνότητας Πιθανότητας
5. Η.Π.Α.: Ηνωμένες Πολιτείες Αμερικής

Ξένη Ορολογία

1. D.E.M.: Digital Elevation Model
2. G.I.S.: Geographic Information Systems
3. E.S.R.I.: Environmental Systems Research Institute

ΒΙΒΛΙΟΓΡΑΦΙΑ

Ελληνική

- Αγατζά – Μπαλοδήμου Α.Μ., Θεωρία Σφαλμάτων και Συνορθώσεις Ι (σημειώσεις από τις παραδόσεις του μαθήματος), Εκδόσεις Ε.Μ.Π., Αθήνα 2007
- Αραποστάθης Ε., Διπλωματική Εργασία, Διαδικτυακές Εφαρμογές Οπτικοποίησης Γεωγραφικών Δεδομένων: Η περίπτωση του Αρχαιολογικού Μουσείου Αθηνών, Τμήμα Γεωγραφίας Χαροκόπειου Πανεπιστημίου, Καλλιθέα 2010
- Κάβουρας Μ., Αρχές Γεωπληροφορικής και Συστήματα Γεωγραφικών Πληροφοριών, Εκδόσεις Ε.Μ.Π., Αθήνα 2007
- Κοκολάκης Γ., Σπηλιώτης Ι., Εισαγωγή στη Θεωρία Πιθανοτήτων Και Στατιστική, Εκδόσεις Συμεών (συγκεκριμένα 4η έκδοση), Αθήνα 1999
- Κουτσόπουλος Κ., Γεωγραφικά Συστήματα Πληροφοριών και Ανάλυση Χώρου, Εκδόσεις Παπασωτηρίου, Αθήνα 2005
- Κουτσόπουλος Κ., Ανδρουλάκης Ν., Εφαρμογές του Λογισμικού Arc Gis 9x Με Απλά Λόγια, Εκδόσεις Παπασωτηρίου, Αθήνα 2005
- Κουτσόπουλος Κ., Ανάλυση Χώρου: Θεωρία - Μεθοδολογία Και Τεχνικές, Εκδόσεις Διηλεκές, Αθήνα 2006
- Μαργαράς Γ., Διπλωματική Εργασία, Ανάπτυξη Εκπαιδευτικού Υλικού Για Δειγματοληψία Σε Περιβάλλον Γ.Σ.Π., Σχολή Αγρονόμων και Τοπογράφων Ε.Μ.Π., Φεβρουάριος 2010
- Μπένος Β., Μέθοδοι και Τεχνικές Δειγματοληψίας, Εκδόσεις Σταμούλη, Πειραιάς 1991
- Στεφανάκης Ε. Βάσεις Γεωγραφικών Δεδομένων και Συστήματα Γεωγραφικών Πληροφοριών, Εκδόσεις Παπασωτηρίου, Αθήνα 2003
- Τζιαφέτας Γ., Θεωρία και Τεχνική Δειγματοληψίας, Τομέας Εκδόσεων Ελληνικού Κέντρου Παραγωγικότητας, 1991

- Τζίγκουρας Γ., Διπλωματική Εργασία, Ανάπτυξη Εκπαιδευτικού Υλικού Για Δειγματοληπτικές Μεθόδους Σε Περιβάλλον Γ.Σ.Π., Σχολή Αγρονόμων και Τοπογράφων Ε.Μ.Π., Δ.Π.Μ. Γεωπληροφορική, Φεβρουάριος 2003
- Χατζόπουλος Ι. “ Τοπογραφία “ , (Β. Γκιούρδας Εκδοτική, Μυτιλήνη 2006)

Ξένη

- Bonham G., Geographic Information Systems for Geoscientists (modeling with Gis), Pergamon, 1994
- Cochran W., Sampling Techniques (3rd edition), John Wiley & sons, 1977
- Dixon C., Leach B., Sampling Methods for Geographical Research, Catmog, 1977
- Mendenhall W., Sincich T., Statistics for Engineering and the Sciences (fifth edition), Prentice Hall, 2007
- Monroe Ch., An introduction to statistical problem solving in geography, Wm.C.Brown Publishers, 1993
- Shaw G. and Wheeler D., Statistical Techniques in Geographical Analysis, John Wiley & sons, 1985

Διαδικτυακές Πηγές

- [http:// www.esri.com](http://www.esri.com) η επίσημη ιστοσελίδα της εταιρείας που εκδίδει το λογισμικό Arc Gis 10
- [helios\ server\ lab\ geography](http://helios\server\lab\geography) η επίσημη ιστοσελίδα του μαθήματος της γεωγραφίας και ανάλυσης χώρου από όπου προμηθευτήκαμε τα δεδομένα της άσκησης
- [http:// www.geologos.gr](http://www.geologos.gr) Γεωεπιστημονικό portal το οποίο περιέχει ορισμούς και την ιστορική αναδρομή των Γ.Σ.Π.

- <http://karagian.users.uth.gr/cscl/22-Katsanos-Avouris.pdf> ιστοσελίδα η οποία περιέχει σημειώσεις για στατιστικές μεθόδους ανάλυσης πειραματικών δεδομένων συνεργασίας
- http://www.nsph.gr/files/011_Ygeias_Paidiou/Epidimiologiki_epitirisi_mathimata/Digmatolipsia.pdf ιστοσελίδα η οποία περιέχει το πρόγραμμα εκπαίδευσης στην επιδημιολογική επιτήρηση και διερεύνηση επιδημιών ΕΣΔΥ – ΚΕΕΛΠΝΟ, 2008